

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Biological Systems Engineering--Dissertations,
Theses, and Student Research

Biological Systems Engineering

Summer 7-22-2019

Design, Development, and Field Testing a VisNIR Integrated Multi-Sensing Soil Penetrometer

Nuwan K. Wijewardane

University of Nebraska - Lincoln, nkw.pdn@huskers.unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/biosysengdiss>



Part of the [Biological Engineering Commons](#), and the [Bioresource and Agricultural Engineering Commons](#)

Wijewardane, Nuwan K., "Design, Development, and Field Testing a VisNIR Integrated Multi-Sensing Soil Penetrometer" (2019). *Biological Systems Engineering--Dissertations, Theses, and Student Research*. 90.
<https://digitalcommons.unl.edu/biosysengdiss/90>

This Article is brought to you for free and open access by the Biological Systems Engineering at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Biological Systems Engineering--Dissertations, Theses, and Student Research by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

DESIGN, DEVELOPMENT, AND FIELD TESTING A VISNIR INTEGRATED
MULTI-SENSING SOIL PENETROMETER

by

Nuwan K. Wijewardane

A DISSERTATION

Presented to the Faculty of
The Graduate College at the University of Nebraska
In Partial Fulfilment of Requirements
For the Degree of Doctor of Philosophy

Major: Biological Engineering

(Agricultural and Biological Systems Engineering)

Under the Supervision of Professor Yufeng Ge

Lincoln, Nebraska

July, 2019

DESIGN, DEVELOPMENT, AND FIELD TESTING A VISNIR INTEGRATED
MULTI-SENSING SOIL PENETROMETER

Nuwan K. Wijewardane, Ph.D.

University of Nebraska, 2019

Advisor Yufeng Ge

The research community in soil science and agriculture lacks a cost-effective and rapid technology for *in situ*, high resolution vertical soil sensing. Visible and near infra-red (VisNIR) technology has the potential to be used for such sensor development due to its ability to derive multiple soil properties rapidly using a single spectrum. Such efforts must, however, overcome a few challenges: (i) a dry ground soil spectral library that can be used to predict the target soil properties accurately, (ii) a robust design which can acquire high quality VisNIR spectra of soil, (iii) an effective method that can link field intact soil spectra to the dry ground spectra in the library.

The overall goal of the work presented in this dissertation was to design, develop, and test a VisNIR integrated multi-sensing penetrometer to estimate soil properties in vertical profile. To achieve this goal, three specific objectives were developed. The first was to investigate and compare the usefulness of five approaches: External Parameter Orthogonalization (EPO), Direct Standardization (DS), Global Moisture Modeling (GMM), Slope Bias Correction (SB) and Selective Wavelength Modeling (SWM), in enabling VisNIR dry ground models to be applied directly to moist soil spectra to predict soil organic carbon and inorganic carbon. The second was to design new VisNIR probes and test them in terms of spectral quality and predictive power using an external spectral library under laboratory conditions. Third was to develop the fully integrated, multi-sensing penetrometer system for high resolution vertical soil sensing and field test the penetrometer to evaluate its performance.

The results showed that EPO, DS and GMM account satisfactorily for the effect of moisture in soil spectra. The VisNIR probe developed showed high spectral quality,

however with a systematic difference compared to standard MugLite[®] spectra which was successfully rectified by DS or spiking. The final designed fully integrated, multi-sensing penetrometer system, could estimate soil properties: total carbon, total nitrogen and bulk density, in vertical soil profile with EPO to correct for field intactness. This can lead to a rapid, robust and cost-effective penetrometer system for *in situ* high resolution vertical soil sensing in the future.

*Affectionately dedicated to all my teachers for their guidance and support
to be who I am today and my wife Niwarthana, son Punsara for their
unconditional love and inspirations to walk so far.*

ACKNOWLEDGMENTS

Many people deserve thanks for making this study and dissertation a reality. First and foremost, I must be thankful to my advisor, Dr. Yufeng Ge for his guidance and continued encouragement throughout my study at the University of Nebraska–Lincoln. His inspiration, direction and advice from the initial to the final stage of my PhD program enabled me to develop an understanding of the subject and to complete my dissertation work successfully. I should also acknowledge him for giving me this research opportunity and the generous financial support to achieve my objectives. I would also like to extend my appreciation to the committee members; Dr. John Gamon, Dr. Andy Suyker and Dr. Yeyin Shi for their excellent guidance, assistance and advising during my PhD program.

I am highly indebted to other collaborators of this project: Dr. Cristine L.S. Morgan from Soil Health Institute, and Dr. Jason Ackerson from Department of Agronomy, Purdue University for their support, constructive comments and suggestions to improve the penetrometer design and field testing. Also, I would like to thank Tyler Smith and Alan Bolt from the Department of Biological Systems Engineering, University of Nebraska-Lincoln for their assistance in the field testing of the VisNIR integrated multi-sensing penetrometer. I would like to thank USDA–NIFA for their funding and access to their resources including the spectral database to make this project a success. Also, I should acknowledge the staff of Kellogg Soil Survey Lab of the USDA-NRCS for maintaining and querying the spectral database for this study. I furthermore convey my gratitude all the fellow graduate students in the Department of Biological Systems Engineering, University of Nebraska-Lincoln for their inspirational thoughts. I should also be thankful to the entire faculty and staff of the Department of the Biological Systems Engineering for their support during my study and research.

Above all, I am indebted to my parents: Ranjani Wijesinghe and Somasiri Wijewardane, for their inspirations, teachings and hard work to make me the person I am today. I would like to express heartfelt thanks my beloved wife Niwarthana Wijewardane

and son Punsara Wijewardane for their unconditional love and continuous motivations to make my studies a success. This accomplishment would not have been possible without the support from all the aforementioned.

Finally, I should acknowledge the funding sources: USDA-NRCS (Award #: 68-7482-14-517) and USDA-NIFA (Award #: 2018-67007-28529, 2017-67021-26248), which helped this project and my research assistantship.

TABLE OF CONTENTS

LIST OF FIGURES	xi
LIST OF TABLES	xiv
CHAPTER 1: GENERAL INTRODUCTION	1
1.1 VISNIR TECHNOLOGY FOR SOIL SENSING	1
1.1.1 What is VisNIR spectroscopy	1
1.1.2 VisNIR spectral signatures of soil constituents	2
1.1.3 Model calibrations	3
1.2 VISNIR BASED SOIL SENSORS	6
1.2.1 Horizontal VisNIR based soil sensors	7
1.2.2 Vertical VisNIR based soil sensors	12
1.3 USE OF VISNIR SPECTRAL LIBRARIES IN SOIL SENSING	16
1.3.1 VisNIR spectral libraries	16
1.3.2 Challenges of using VisNIR libraries for soil sensors	17
1.4 REQUIREMENTS FOR A ROBUST VISNIR BASED SOIL SENSORS	18
CHAPTER 2: MOVING FROM LAB TO FIELD: CHALLENGE OF USING DRY GROUND VISNIR LIBRARY TO PREDICT MOIST SOILS	20
2.1 INTRODUCTION	20
2.1.1 VisNIR for soil sensing	20
2.1.2 Use of VisNIR libraries	21
2.1.3 Approaches to link dry ground VisNIR models to moist spectra	22
2.1.3.1 External parameter orthogonalization (EPO)	22
2.1.3.2 Direct standardization (DS)	23
2.1.3.3 Global moisture modelling (GMM)	23

2.1.3.4	Slope bias correction (SB)	24
2.1.3.5	Selective wavelength modelling (SWM)	24
2.1.4	Objective	25
2.2	MATERIALS AND METHODS	25
2.2.1	Soil samples, VisNIR spectral measurement and sample rewetting	25
2.2.2	Data subsets and analysis	26
2.2.2.1	Dry ground data set (S^0)	26
2.2.2.2	Transformation data set (S^1)	27
2.2.2.3	Test data set (S^2)	27
2.2.3	Analyses to link dry ground and moist spectra	28
2.3	RESULTS AND DISCUSSION	32
2.3.1	Comparison of the five approaches to account for moisture effect	32
2.3.2	Comparison between EPO, moisture-explicit DS and DS	36
2.3.3	Comparison of the results of moisture correction at different moisture contents	38
2.3.4	Practical considerations: field implementation of the moisture correction approaches	42
2.4	CONCLUSIONS	43
CHAPTER 3: DESIGN, DEVELOPMENT, AND LAB TESTING OF		
A VISNIR OPTICAL PROBE FOR SOIL SENSING		
3.1	INTRODUCTION	45
3.1.1	Soil sensing technologies	45
3.1.2	VisNIR technology for vertical soil sensing	46
3.1.3	Objective	47
3.2	MATERIALS AND METHODS	48
3.2.1	VisNIR optical probe	48
3.2.2	Soil samples and datasets	48

3.2.3	VNIR probe testing and modeling	50
3.3	RESULTS AND DISCUSSION	54
3.3.1	Comparison of D1 and D2	54
3.3.2	Qualitative Assessment of soil spectra acquired by the VNIR probe .	55
3.3.3	Prediction performance of the spectra acquired by the VisNIR probe	57
3.3.4	Development and field implementation of the VisNIR based pen- etrometer: Challenges	62
3.4	CONCLUSIONS	63
CHAPTER 4: DEVELOPMENT, AND FIELD TESTING A VISNIR INTEGRATED MULTI-SENSING SOIL PENETROMETER		64
4.1	INTRODUCTION	64
4.1.1	Soil sensing	64
4.1.2	VisNIR for <i>in situ</i> soil sensing	64
4.1.3	Objectives	66
4.2	MATERIALS AND METHODS	66
4.2.1	VisNIR integrated multi-sensing soil penetrometer system	66
4.2.2	Field experiment to test the integrated VisNIR multi-sensing pen- etrometer	68
4.2.2.1	Field testing	68
4.2.2.2	Modeling of TC and TN	72
4.2.2.3	Modeling and prediction of soil bulk density	73
4.3	RESULTS AND DISCUSSION	74
4.3.1	Performance of the VisNIR penetrometer during the field test	74
4.3.2	Using <i>in situ</i> soil spectra from the VisNIR penetrometer to predict soil TC and TN	75
4.3.3	Modeling soil bulk density	77
4.3.4	VisNIR integrated multi-sensing penetrometer	79

4.4 CONCLUSIONS	82
CHAPTER 5: GENERAL CONCLUSIONS AND FUTURE WORK .	84
REFERENCES	87

LIST OF FIGURES

Figure 1.1	Electromagnetic spectrum (Source: Viscarra Rossel, Walvoort, McBratney, Janik, and Skjemstad (2006)).	1
Figure 1.2	Soil VisNIR spectra showing approximate occurrence of the combination, first, second, and third overtone (OT) vibrations (Source: Stenberg, Viscarra Rossel, Mouazen, and Wetterlind (2010)).	3
Figure 1.3	(A) Soil VisNIR spectra and (B) the region 1100–2500 nm showing the spectra of three soils: organic agricultural soil with 40% SOC, 1% SOC with 87% sand and 4% clay, 1% SOC with 12% sand and 44% clay (Source: Stenberg, Viscarra Rossel, Mouazen, and Wetterlind (2010)). . . .	5
Figure 1.4	Cross-sectional view of the organic matter sensor and probe developed and tested by Shonk, Gaultney, Schulze, and Van Scoyoc (1991). . . .	7
Figure 1.5	Prototype portable spectrometer developed by Sudduth and Hummel (1993) to estimate soil organic matter.	8
Figure 1.6	Subsoiler-optical unit set up developed by Mouazen, De Baerdemaeker, and Ramon (2005).	9
Figure 1.7	Field testing of variable rate phosphorus applicator (VR) based on the soil sensor developed by Mouazen, De Baerdemaeker, and Ramon (2005).	10
Figure 1.8	The shank-based spectrometer demonstrated by Christy (2008) to obtain soil NIR reflectance spectra.	11
Figure 1.9	The real-time soil surface sensor developed by Kodaira and Shibusawa (2013). Left: sensor mounted on the tractor.	11
Figure 1.10	Tractor-drive chamber developed by Rodionov et al. (2015).	13
Figure 1.11	Housing assembly developed by Ben-Dor, Heller, and Chudnovsky (2008).	14

Figure 1.12 Optical probe designs tested by Poggio, Brown, and Bricklemyer (2015).	14
Figure 1.13 Soil Condition Analyses System developed by Viscarra Rossel, Lobsey, Sharman, Flick, and McLachlan (2017).	15
Figure 1.14 Veris [®] P4000 used by Veum, Parker, Sudduth, and Holan (2018). . .	16
Figure 2.1 Flow charts of the procedures of the five approaches to account for moisture in soil VisNIR reflectance.	29
Figure 2.2 Implementation of external parameter orthogonalization (EPO) transformation with the model-coupled-Cross Validation (model-coupled-CV) and Wilk's Λ method.	30
Figure 2.3 Scatterplots of soil organic carbon (OC) prediction at eight different moisture contents: (a) no correction and (b) global moisture modeling (GMM).	35
Figure 2.4 Scatterplots of soil inorganic carbon (IC) prediction at eight different moisture contents: (a) no correction and (b) moisture-explicit direct standardization (moisture-explicit DS).	36
Figure 2.5 Spectral transformation at different moisture contents by: (a) external parameter orthogonalization (EPO), (b) moisture-explicit direct standardization (moisture-explicit DS) and (c) direct standardization (DS). . . .	37
Figure 3.1 Designed VisNIR optical probes.	49
Figure 3.2 Laboratory experimental setup for the testing of the VisNIR probes. .	52
Figure 3.3 Spectra obtained from VisNIR probe designs D1, D2, dark reference corrected D1 (DRC-D1), dark reference corrected D2 and MugLite [®] (ML).	54
Figure 3.4 Qualitative assessment of the soil spectra of 15 samples from Group B scanned by the MugLite [®] and the VisNIR probe (VNIRP), and the VNIR probe scans transformed by direct standardization (VNIRP.DS).	56

Figure 3.5	The convex hull of the three spectral sets in the first two principal component space.	57
Figure 3.6	Prediction plot of Organic Carbon (OC) in Group B: (a) Library only model on VisNIR probe scans; (b) Library only model on DS transformed scans; (c) spiked model on VisNIR probe scans; and (d) spiked and extra weighted model on VisNIR probe scans.	61
Figure 4.1	Cross-sectional view of the multi-sensing penetrometer to show the key system components (left) and the solid rendering of the penetrometer with the extension shaft (right).	68
Figure 4.2	The graphical user interface of the LabVIEW program showing data collection from the VisNIR multi-sensing penetrometer system.	69
Figure 4.3	Difference in soil VisNIR spectra acquired in the lab (dry ground) and by the VisNIR multi-sensing penetrometer (<i>in situ</i>).	77
Figure 4.4	Prediction plots for dry-ground (DG) spectra, uncorrected <i>in situ</i> (VNIRP) spectra, and external parameter orthogonalization corrected (EPO) <i>in situ</i> spectra.	78
Figure 4.5	Prediction plots for soil bulk density modeling with (a) insertion force F and insertion depth D , and (b) insertion force F , insertion depth D , as well as gravimetric moisture content and clay content estimated from <i>in situ</i> VisNIR data corrected with external parameter orthogonalization.	79
Figure 4.6	High resolution predictions of bulk density, clay, gravimetric moisture content, Total Carbon, and Total Nitrogen along the soil profile for one selected sampling location.	80

LIST OF TABLES

Table 1.1 Fundamental mid-IR absorptions of soil constituents and their overtones and combinations in the VNIR (Source: Viscarra Rossel and Behrens (2010)).	4
Table 2.1 Summary statistics of soil organic carbon and inorganic carbon in the three subsets (S^0 , S^1 and S^2) in this study.	27
Table 2.2 Performance of five different moisture correction approaches to predict soil organic carbon and inorganic carbon.	33
Table 2.3 Performance of external parameter orthogonalization (EPO), moisture-explicit direct standardization, direction standardization (DS), and global moisture modeling (GMM) across the eight moisture contents for soil organic C prediction.	39
Table 2.4 Performance of external parameter orthogonalization (EPO), moisture-explicit direct standardization, direction standardization (DS), and global moisture modeling (GMM) across the eight moisture contents for soil inorganic C prediction.	41
Table 3.1 Summary statistics of the soil properties for the library and the local set (Groups A and B) used in the lab testing of the VisNIR probe.	51
Table 3.2 Prediction performance of different models with different validation sets for Organic Carbon (OC), Total Carbon (TC), and Total Nitrogen (TN).	58
Table 4.1 A summary of the test fields, soils, soil cores, and soil samples analyzed from each field.	70
Table 4.2 Summary statistics of the soil properties for the VisNIR dry-ground library from USDA-NRCS and the field samples.	72

Table 4.3 Prediction model statistics for soil total carbon (TC) and total nitrogen (TN).	76
---	----

CHAPTER 1

GENERAL INTRODUCTION

1.1 VISNIR TECHNOLOGY FOR SOIL SENSING

1.1.1 What is VisNIR spectroscopy

In the electromagnetic (EM) spectrum, which is the range of all possible frequencies of EM radiation, visible light lies between 350–700 nm range while near infrared (NIR) lies in 700–2500 nm range (Figure 1.1). Hence the visible and near infrared (VisNIR) region is generally considered from 350 to 2500 nm.

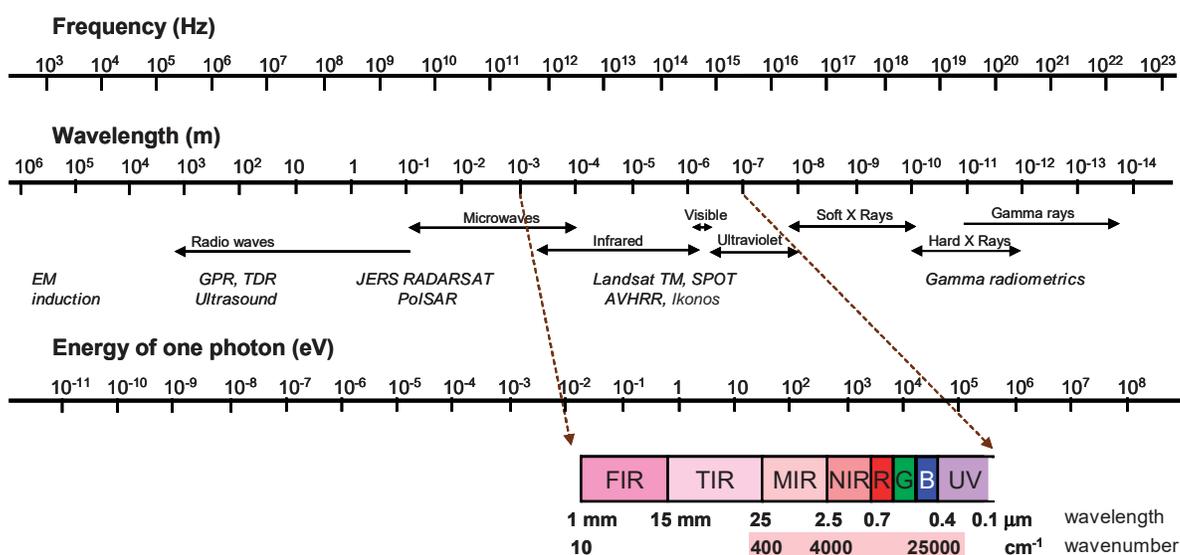


Figure 1.1. Electromagnetic spectrum (Source: Viscarra Rossel, Walvoort, McBratney, Janik, and Skjemstad (2006)).

Various chemical substances absorb radiation of different wavelengths, which correspond to the chemical bonds in the compounds, giving rise to characteristic signatures in the spectra. Though most of the fundamental vibrations occur in the mid infrared (MIR) region, their overtones and combinations can be observed in the NIR

region (Burns & Ciurczak, 2007). However, these overtones are complex in nature and not easily distinguishable, which requires the use of more advanced multivariate calibration techniques to develop models to detect different compounds. VisNIR spectroscopy where VisNIR spectral signatures are used to detect/estimate chemical compounds, is an inexpensive, non-destructive and powerful tool and is used for quality control and process monitoring in industrial settings. This technology has developed as a tool for proximal sensing in natural resources with the evolution of technology and research interest during the past few decades.

1.1.2 VisNIR spectral signatures of soil constituents

EM radiation interacts with soil causing the individual molecules to absorb energy and vibrate, either by bending or stretching. The energy absorbed is related to the energy quantum corresponding to different energy levels of the bonds. The resulting absorbance spectrum has characteristic patterns (signatures) which can be used to identify different properties and constituents of soil (Miller, 2001). Though the majority of the fundamental spectral signatures of soil constituents occur in MIR region, discernible overtones of the primary signatures can be observed in the VisNIR region (Figure 1.2), which can be used to identify different soil properties. For instance, iron-containing mineral absorptions occur in the visible region (Sherman & Waite, 1985). Soil organic matter signatures in the NIR region are characterized by the overtones and combinational absorptions of O–H, C–H and N–H bonds (Clark, 1999; Clark, King, Klejwa, Swayze, & Vergo, 1990). Clay mineral absorption overtones are attributed to the spectral signatures of OH, H₂O and CO₃, observed in longer wavelengths (Stenberg, Viscarra Rossel, Mouazen, & Wetterlind, 2010). Moisture absorption bands occur near 1400 and 1900 nm (Bowers & Hanks, 1965; Dalal & Henry, 1986).

Some of these observed spectral signatures for different soil constituents are shown in Table 1.1. Literature provides ample evidence in the use of VisNIR spectroscopy

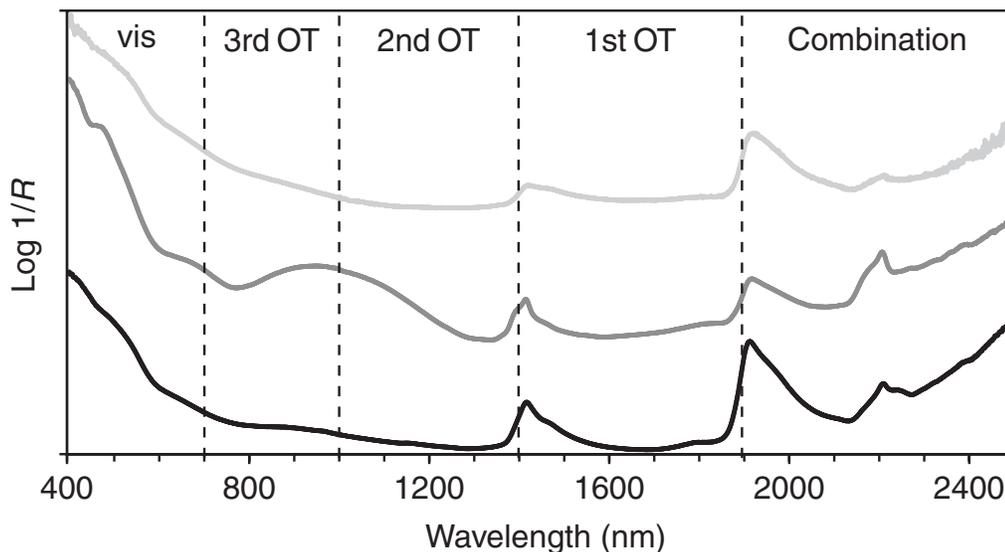


Figure 1.2. Soil VisNIR spectra showing approximate occurrence of the combination, first, second, and third overtone (OT) vibrations (Source: Stenberg, Viscarra Rossel, Mouazen, and Wetterlind (2010)).

to detect different soil properties: moisture (Ben-Dor, Heller, & Chudnovsky, 2008; Chang, Laird, & Hurburgh, 2005; Hummel, Sudduth, & Hollinger, 2001), organic carbon (Chang, Laird, Mausbach, & Hurburgh, 2001; Islam, Singh, & McBratney, 2003; Shepherd & Walsh, 2002), texture (Brown, Shepherd, Walsh, Dewayne Mays, & Reinsch, 2006; Ge, Morgan, & Ackerson, 2014; Stenberg, Jonsson, & Börjesson, 2002) and plant nutrients (Stenberg et al., 2010). Soil organic carbon (OC) and clay are two soil properties which have been extensively researched and proved for their potential to be modeled from VisNIR spectra due to their unique spectral signatures (Figure 1.3). Still, these soil properties do not show clear distinct spectral signatures so that one can easily model for target characteristics.

1.1.3 Model calibrations

Due to the overlapping nature of absorption bands, spectral signatures of different soil properties are non-specific which is further confounded by scattering effects caused by soil structure or constituents such as quartz. The resulting absorption patterns are complex

Table 1.1. Fundamental mid-IR absorptions of soil constituents and their overtones and combinations in the VNIR (Source: Viscarra Rossel and Behrens (2010)).

Soil constituent	Fundamental (cm^{-1})	VNIR wavelength (nm)
Fe oxides		
Geothite		434, 480, 650, 920
Haematite		404, 444, 529, 650, 884
Water	ν_1 O–H 3278 ν_2 H–O–H 1645 ν_3 O–H 3484	1915 1455 1380, 1135, 940
Hydroxyl	ν_1 O–H 3575	1400, 930, 700
Clay minerals		
Kaolin doublet	ν_{1a} O–H 3695	1395
	ν_{1b} O–H 3620	1415
	δ Al–OH 915	2160, 2208
Smectite	ν_1 O–H 3620	2206
	δ_a Al–OH 915	2230
	δ_b AlFe–OH 885	
Illite	ν_1 O–H 3620	2206, 2340, 2450
Carbonate	ν_3 CO_3^{2-} 1415	2336
Organics		
Aromatics	ν_1 C–H 3030	1650, 1100, 825
Amine	δ N–H 1610	2060
	ν_1 N–H 3330	150, 1000, 751
Alkyl asymmetric symmetric doublet	ν_3 C–H 2930	1706
	ν_1 C–H 2850	1754, 1138, 1170, 853, 877
Carboxylic acids	ν_1 C=O 1725	1930, 1449
Amides	ν_1 C=O 1640	2033, 1524
Aliphatics	ν_1 C–H 1465	2275, 1706
Methyls	ν_1 C–H 1445–1350	2307–2469, 1730–1852
Phenolics	ν_1 C–OH 1275	1961
Polysaccharides	ν_1 C–O 1170	2137
Carbohydrates	ν_1 C–O 1050	2381

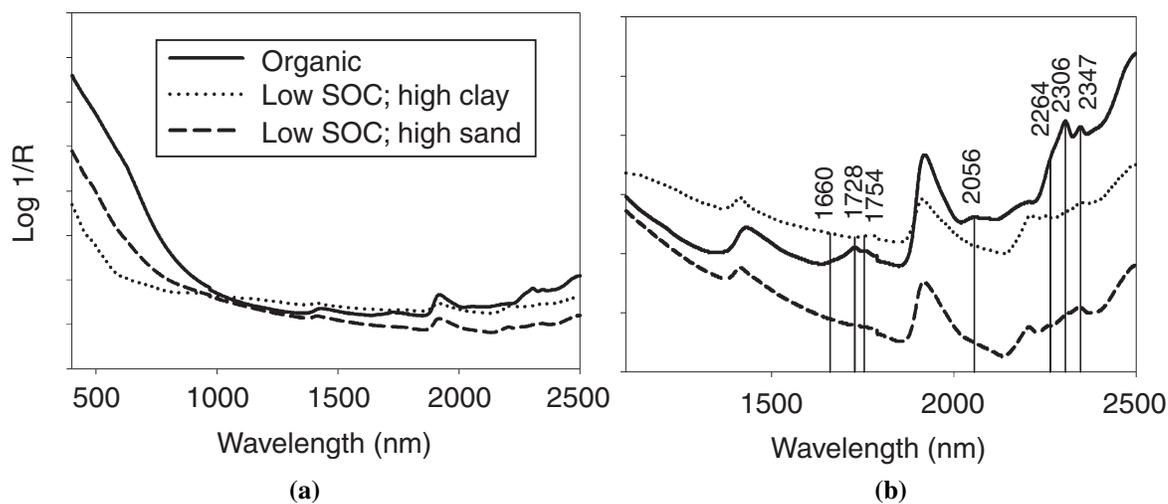


Figure 1.3. (A) Soil VisNIR spectra and (B) the region 1100–2500 nm showing the spectra of three soils: organic agricultural soil with 40% SOC, 1% SOC with 87% sand and 4% clay, 1% SOC with 12% sand and 44% clay (Source: Stenberg, Viscarra Rossel, Mouazen, and Wetterlind (2010)).

which cannot be used to derive models using simple correlation techniques and require more advanced multivariate calibration techniques (Martens & Naes, 1992). There are many linear and non-linear techniques to calibrate models such as partial least squares regression (PLSR) (Wold, Martens, & Wold, 1983), step-wise multiple linear regression (Ben-Dor & Banin, 1995; Dalal & Henry, 1986), principal component regression, artificial neural networks (ANN) (Daniel, Tripathi, & Honda, 2003), boosted regression trees (Brown et al., 2006), random forests (RF) (Viscarra Rossel & Behrens, 2010) and support vector regression (SVR) (Stevens et al., 2008; Viscarra Rossel & Behrens, 2010) to obtain robust models. Out of these, PLSR is the most commonly used linear regression method whereas ANN, RF and SVR are considered non-linear modeling techniques. These modeling techniques have tuning parameters which are changed iteratively until the lowest error is obtained through cross-validation. In this dissertation work, PLSR and SVR modeling techniques were selected as linear and non-linear modeling techniques.

PLSR is a modeling technique used to build models for highly collinear data.

PLSR uses an algorithm similar to principal component analysis to reduce the number of

dimensions into several latent variables. It considers the response variable simultaneously when constructing the latent variables as well. Then, a linear model is fitted between the latent variables and the target response (Helland, 2004). PLSR is often preferred due to its ability to explain the response variable with a reduced number of predictor variables, making it more interpretable and reduce computational demand (Stenberg et al., 2010). The tuning parameter for PLSR is the number of latent variables (n_{LV}) used for regression.

Support Vector Regression (SVR) focuses on building an optimal hyperplane in the higher dimensional feature space (Vapnik, 2013). In classification setting, a boundary which has the smallest distance from the hyperplane to the observations, is calculated as the decision margin (Hastie, Tibshirani, & Friedman, 2001). In the regression, a linear regression function is calculated in the higher dimensional feature space for which the input data is mapped using a kernel function. This technique tries to minimize the generalization error, instead of reducing the observed training error (Basak, Pal, & Patranabis, 2007). This modeling technique is effective in higher dimensional modeling of NIR spectra (Thissen, Pepers, Üstün, Melssen, & Buydens, 2004). Viscarra Rossel and Behrens (2010) showed that SVR can result in smaller root mean squared errors compared to many other modeling techniques. The tuning parameter for the SVR used in this work is 'C', which determines the number and severity of the violations to the boundary (simply, it is the budget for how many observations can violate the margin) (James, Witten, Hastie, & Tibshirani, 2013).

1.2 VISNIR BASED SOIL SENSORS

Unlike other sensing technologies, VisNIR spectroscopy can derive multiple soil properties non-destructively, rapidly and at a low cost due to the presence of spectral signatures for different constituents in the same spectrum. VisNIR has the potential to complement expensive laboratory analysis and enable *in-situ* sensing (Kodaira & Shibusawa, 2013). For several decades, researchers have been trying to use this optical

technology to develop sensors to measure different surface soil properties.

1.2.1 Horizontal VisNIR based soil sensors

Griffis (1985) developed a simple soil carbon sensor which consisted of an incandescent light source and a silicon phototransistor. The sensor setup was mounted inside a light-proof housing. The laboratory testing of the sensor with 18 air dried Arkansas soils resulted an R^2 of 0.75 for soil carbon. The same sensor was reported on a system consisting of an elevating chain and horizontal belt used to convey soil past the sensor to mimic on-the-go *in situ* sensing. The laboratory evaluation with sieved air-dried soils showed successful in locating a step change in soil type (Kocher & Griffis, 1989). Shonk, Gaultney, Schulze, and Van Scoyoc (1991) tested a prototype real-time soil surface organic matter sensor probe which measured the light reflectance using a photodiode and red light emitting diodes as light sources (Figure 1.4). They conducted six field tests and observed good correlations between sensor output and soil organic matter in the 1–6% range for fine and medium textured soils.

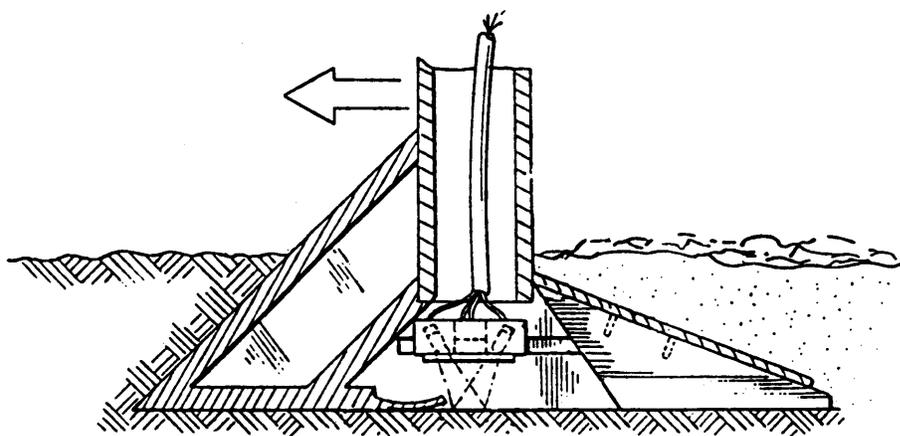


Figure 1.4. Cross-sectional view of the organic matter sensor and probe developed and tested by Shonk, Gaultney, Schulze, and Van Scoyoc (1991).

A portable spectrometer producing a continuous spectrum with a bandpass of 60 nm or less, was introduced by Sudduth and Hummel (1993) to estimate soil organic matter (Figure 1.5). The laboratory testing with 30 representative Illinois soils resulted in a R^2 of

0.89 and standard error of prediction (SEP) of 0.23%. However, the limited field tensing in-furrow did not provide accurate estimates due to movement of the sample relative to the sensor during scanning. The same sensor was updated later to have faster data collection. The updated sensor was used to predict soil organic matter and moisture contents of surface and subsurface soils using three soil cores collected at 16 sites in US Corn Belt. The results showed a SEP of 0.62 and 5.31% for organic matter and moisture respectively (Hummel et al., 2001).

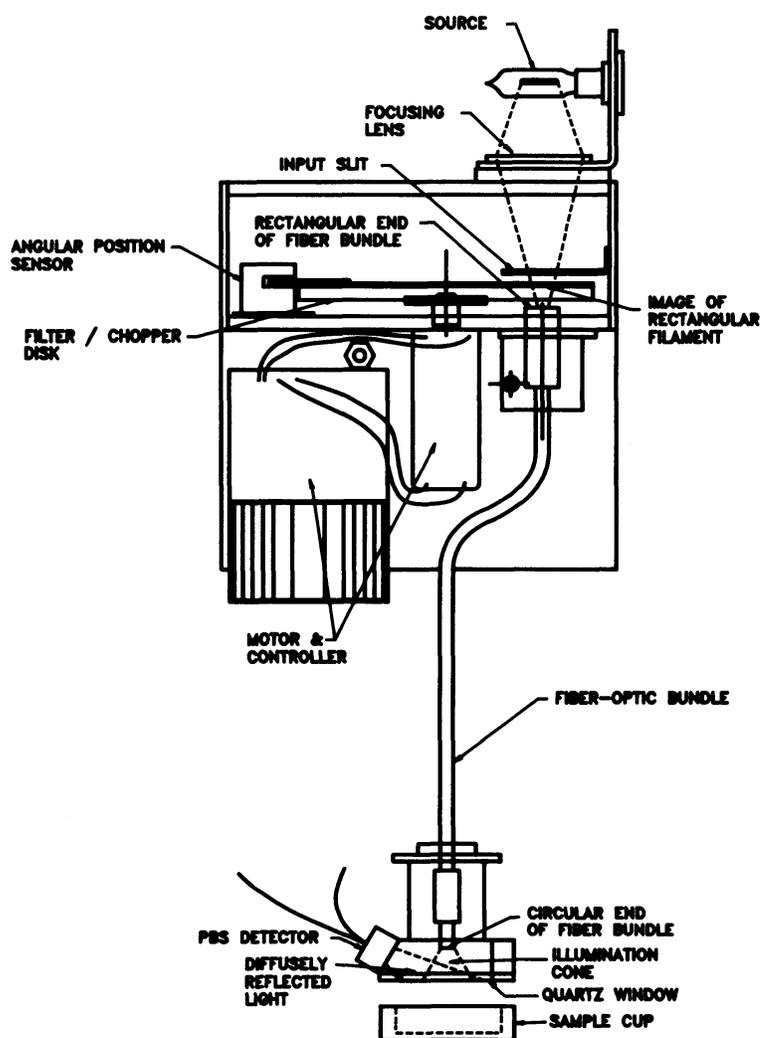


Figure 1.5. Prototype portable spectrometer developed by Sudduth and Hummel (1993) to estimate soil organic matter.

Mouazen, De Baerdemaeker, and Ramon (2005) used a fiber-type visible and

near-infrared spectrophotometer, with a light reflectance measurement range of 306.5–1710.9 nm to measure soil moisture. The optical unit was attached to a subsoiler chisel backside to scan soil surface and obtain reflectance spectra (Figure 1.6). The system was calibrated under stationary laboratory conditions on sample collected from an Arenic Cambisol field with different soil textures. The laboratory testing showed a root mean squared error (RMSE) of $0.0175 \text{ kg kg}^{-1}$ and correlation of 0.978 for cross-validation. The field testing of the sensing system showed a RMSE of 0.025 kg kg^{-1} . The same system was later used for predicting carbon, pH and phosphorus, showing its capacity to measure multiple soil properties on-line (Mouazen, Maleki, De Baerdemaeker, & Ramon, 2007). Maleki, Mouazen, Ramon, and De Baerdemaeker (2007) and (2008) demonstrated the ability of this sensor to be used for variable rate phosphorus application in the field (Figure 1.7).



Figure 1.6. Subsoiler-optical unit set up developed by Mouazen, De Baerdemaeker, and Ramon (2005).

An on-the-go in situ spectrophotometer based sensor system was developed by

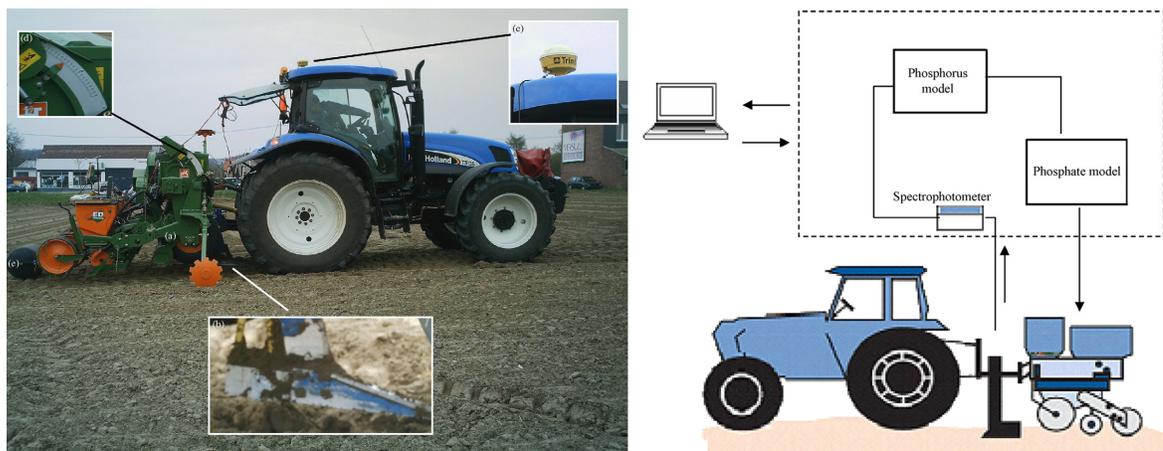


Figure 1.7. Field testing of variable rate phosphorus applicator (VR) based on the soil sensor developed by Mouazen, De Baerdemaeker, and Ramon (2005). (a) planter and fertilizer applicator (AMAZONE, ED302); (b) sensor and subsoiler; (c) DGPS antenna; (d) electrical actuator and (e) roller for closing the trench made by subsoiler. Right: schematic diagram of the soil sensor-based VR applicator.

Christy (2008) for real-time measurements of soil attributes. Figure 1.8 shows the schematics of the basic design of the sensor system which was evaluated using soil samples from eight fields in Kansas. One-field out validation results showed an RMSE of 0.52% for soil organic matter with an R^2 of 0.67. Bricklemeyer and Brown (2010) employed a new commercially available on-the-go VisNIR sensor (Veris[®] Technologies Inc., Salina, KS, USA) which was built into an agricultural shank mounted on a toolbar to be pulled behind a tractor, to obtain soil spectra and predict soil properties. The field evaluation of this sensor in Montana showed degraded accuracy compared to laboratory testing due to soil heterogeneity and moisture variation. However, Knadel, Thomsen, Schelde, and Greve (2015) demonstrated that the new Veris[®] mobile sensor platform can be used to estimate soil organic carbon with VisNIR spectra and fusion of sensor data.

A real-time surface soil sensor attached to the tractor was developed by Kodaira and Shibusawa (2013) to predict multiple attributes: moisture, organic matter, pH, electrical conductivity, cation exchange capacity, total carbon, ammonium nitrogen, hot water extractable nitrogen, nitrate nitrogen, total nitrogen, available phosphorus, and phosphorus absorptive coefficient (Figure 1.9). One hundred and forty-four soil spectra

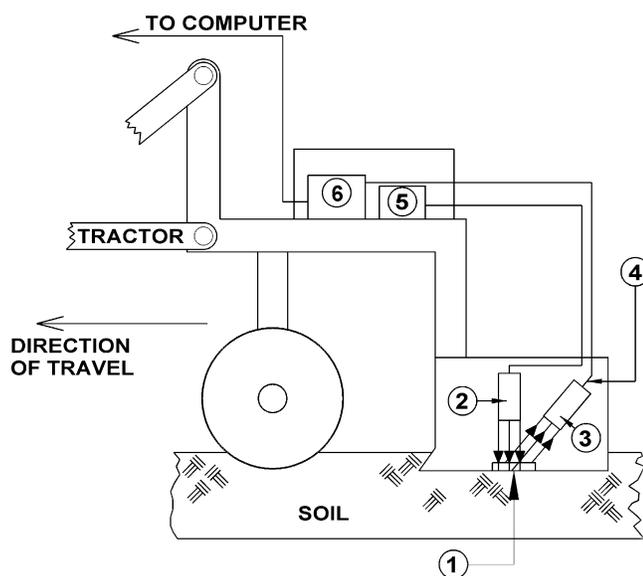


Figure 1.8. The shank-based spectrometer demonstrated by Christy (2008) to obtain soil NIR reflectance spectra. (1) Sapphire window; (2) halogen lamp; (3) collection optic; (4) fiber optic; (5) spectrometer; (6) power supply.

were collected using this sensor in the field to build PLSR model and results showed that the system can accurately estimate some of the soil properties tested and comparable to lab-based results. Later, the same sensor system was used to develop soil property maps as well (Aliah Baharom, Shibusawa, Kodaira, & Kanda, 2015).

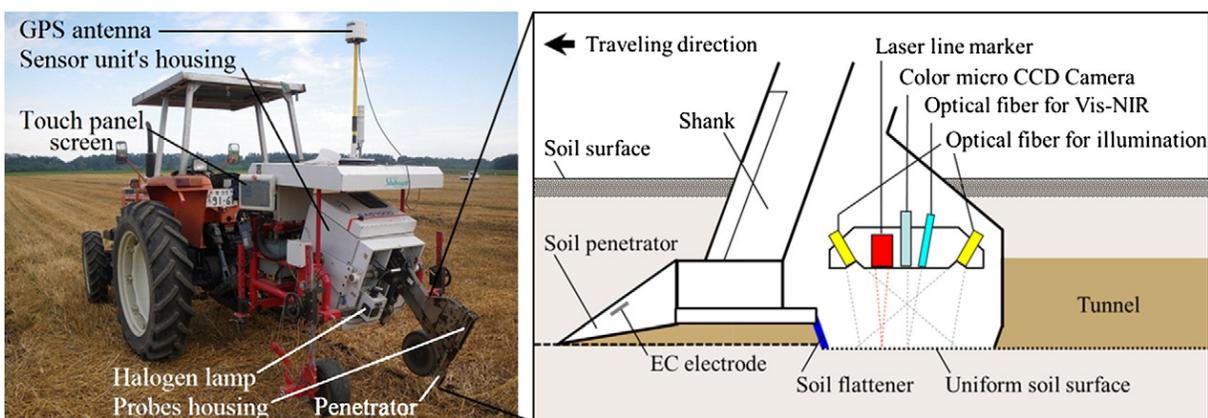


Figure 1.9. The real-time soil surface sensor developed by Kodaira and Shibusawa (2013). Left: sensor mounted on the tractor. Right: cross-sectional view of the soil penetrator and probe housing.

Rodionov et al. (2015) developed a tractor-driven measuring chamber which obtain VisNIR spectra of surface soil to be used to estimate soil organic carbon (Figure

1.10). The field testing of the system was done on two long-term experiments and the results showed an R^2 of 0.65 and standard error of 1.12 g kg^{-1} . They concluded that the system could be used as a reliable on-the-go sensor with further improvements.

1.2.2 Vertical VisNIR based soil sensors

Unlike horizontal sensors which are intended to estimate surface soil attributes, the vertical sensors based of VisNIR technology are sporadically reported in the literature. Ben-Dor et al. (2008) constructed a housing assembly which could be adapted to any portable spectrometer to obtain subsoil spectral readings (Figure 1.11). They tested the assembly with an attached field spectrometer on drilled holes, trenches, and soil banks to estimate soil moisture, organic matter, carbonates, free iron oxides and specific surface area which yielded successful results. Kusumo, Hedley, Tuohy, Hedley, and Arnold (2010) also developed a modified soil probe for a portable ASD spectroradiometer to scan the extracted soil cores.

A multi-sensor penetrometer to measure bulk density of surface soils was developed by Quraishi and Mouazen (2013). The penetrometer consisted of a spectrophotometer to obtain the VisNIR reflectance spectra of the soil to estimate moisture, organic matter, and clay content, and a load cell to measure the penetration resistance. The sensor data was then used to model bulk density. The system was evaluated with 471 samples collected from various fields across four European countries and the results showed a validation R^2 of 0.84 and RMSE of 0.08 Mg m^{-1} . Poggio, Brown, and Brickleyer (2015) evaluated the optical performance of a new soil penetrometer VisNIR foreoptic under the laboratory conditions (Figure 1.12). Their design consisted of a halogen lamp to produce broad-spectrum light and mirrors to direct light to soil surface and to fiber optic. The second design consisted of a bifurcated fiber optic and a mirror to direct light which was acquired from ASD (Analytical Spectral Devices formerly, now Malvern Panalytical Inc., Boulder, CO). They compared these

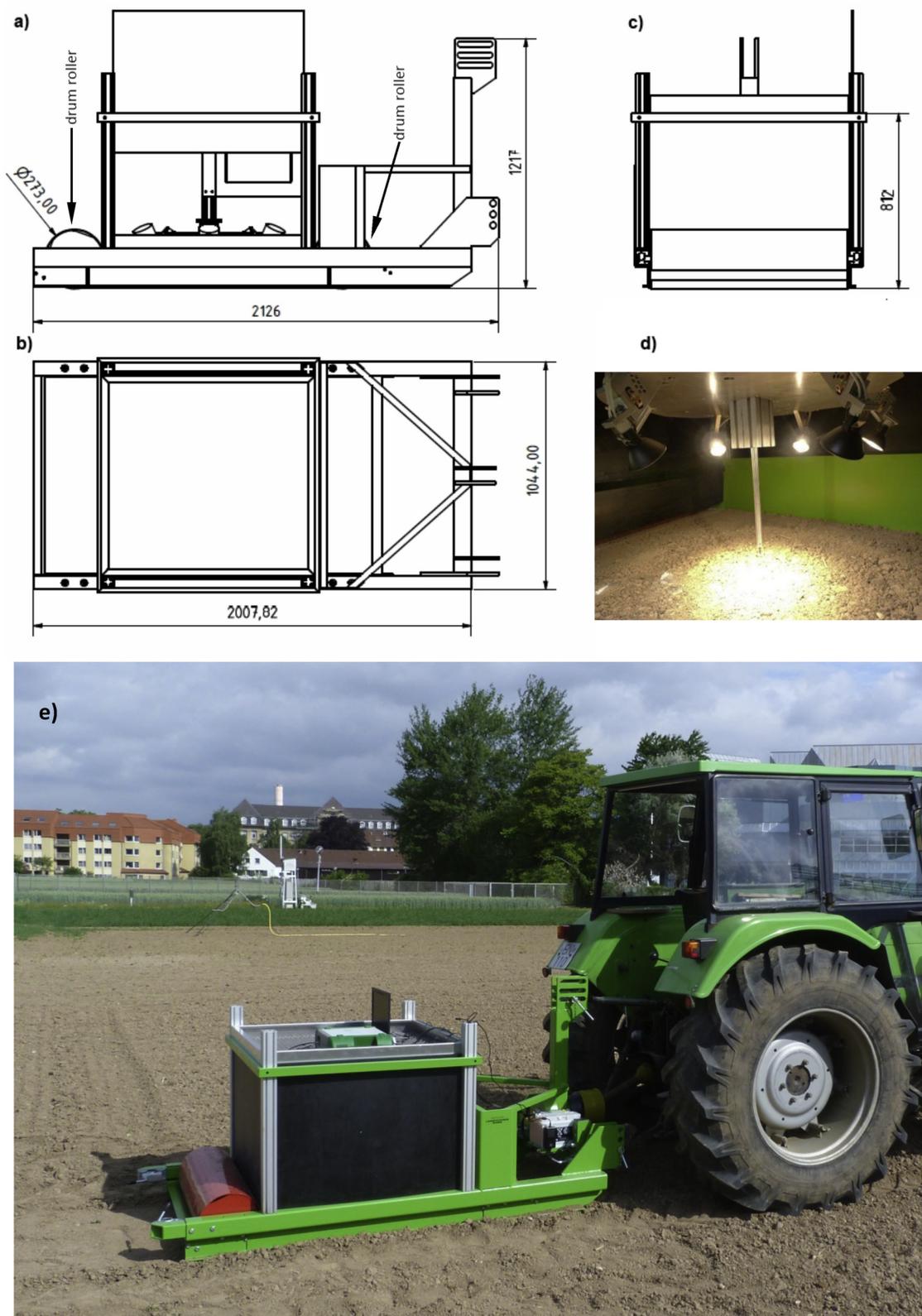


Figure 1.10. Tractor-drive chamber developed by Rodionov et al. (2015). (a) lateral view with opened side wall, (b) top view, and (c) rear view with dimensions (mm). (d) The lamp holder comprises six adjustable halogen lamps and the sensor head with the optical fiber. (e) Measurement setup during the field VisNIR spectra acquisitions.

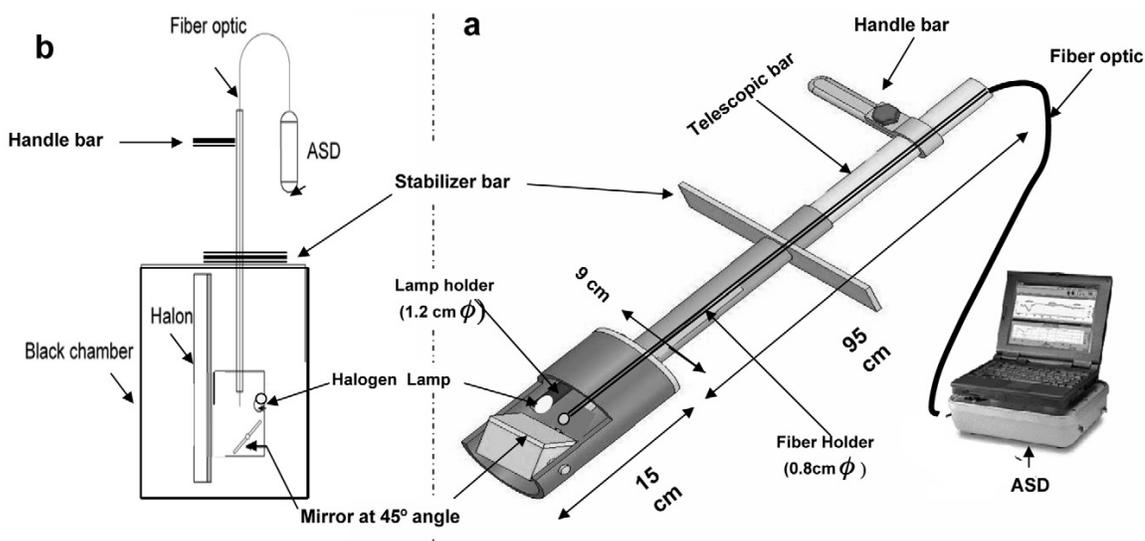


Figure 1.11. Housing assembly developed by Ben-Dor, Heller, and Chudnovsky (2008). (a) sketch showing the all parts. (b) schematic configuration of the assembly.

designs with ASD contact probe and observed that their design performed well in terms of spectral quality and chemometric models.

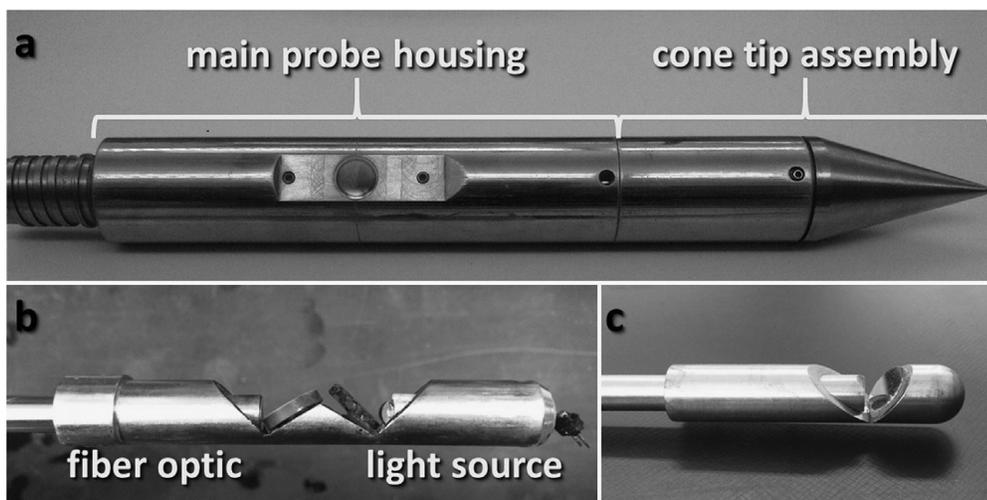


Figure 1.12. Optical probe designs tested by Poggio, Brown, and Bricklemyer (2015). (a) main probe housing and cone tip assembly, (b) current optic assembly and (c) first ASD optic assembly.

Viscarra Rossel, Lobsey, Sharman, Flick, and McLachlan (2017) developed a new soil profile sensing system called as ‘Soil Condition Analyses System’ (SCANS) (Figure 1.13). This system integrated an automated soil core sensing system with statistical

analytics and modeling to enable soil characterization at finer depth resolutions. It consisted of a gamma-ray densitometer to measure bulk density, digital cameras to image the measured soil, and a VisNIR spectrometer to measure iron oxides, clay mineralogy and other soil attributes. Their intention was to provide rapid, precise, quantitative and spatially explicit information on the properties of soil profiles with a level of detail which is difficult to obtain using other means. This system was further successfully tested by Poggio, Roudier, Blaschek, and Hedley (2018).

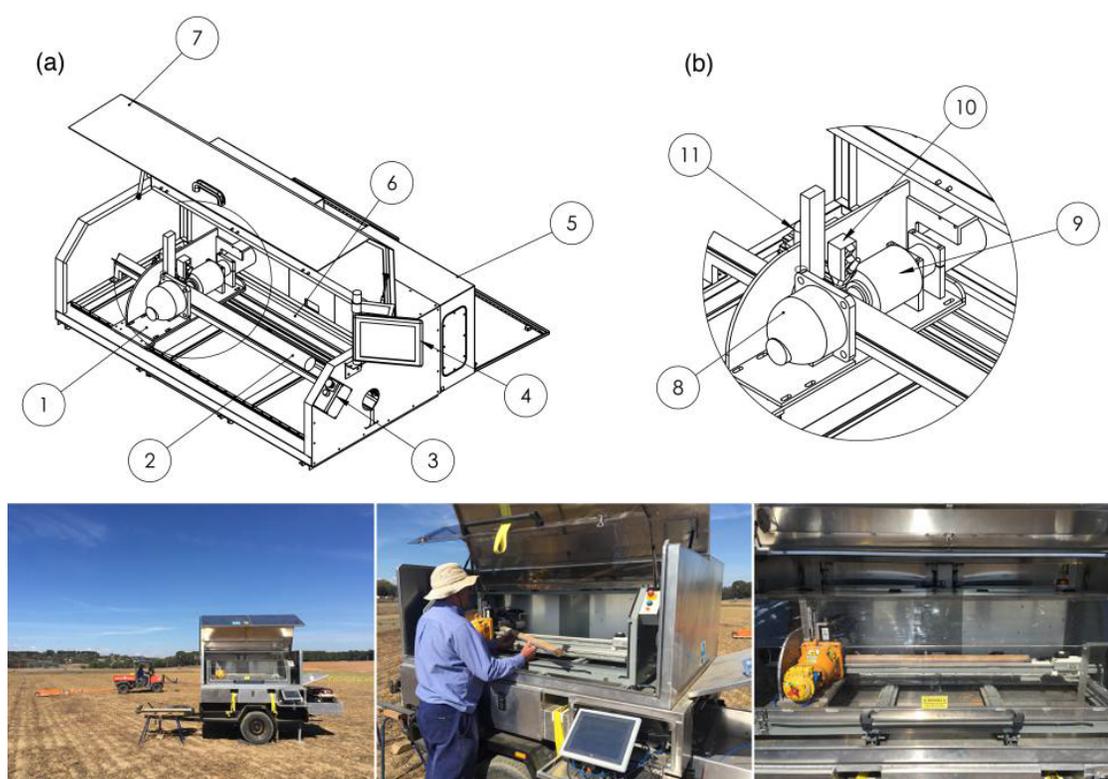


Figure 1.13. Soil Condition Analyses System developed by Viscarra Rossel, Lobsey, Sharman, Flick, and McLachlan (2017). Top row: Schematic of the SCANS core sensing system: (a) (1) sensor head (2) soil core (3) emergency stop and reset buttons (4) touch screen PC (5) electronics boxes (6) linear actuator (7) polycarbonate hood with safety sensors. (b) (8) gamma-ray source (9) gamma-ray detector (10) spectrometer contact probe attachment (11) cameras. Bottom row: field-deployable core sensing system's operation in the field.

There is a commercially available system developed by Veris[®] Technologies (P4000) for soil profile sensing which can estimate soil properties such as EC, force of

penetration, and carbon. Wetterlind, Piikki, Stenberg, and Söderström (2015) demonstrated the use of this sensor on soil texture and organic matter estimation using field specific calibration. Veum, Parker, Sudduth, and Holan (2018) demonstrated the ability to use this system to predict soil OC, TN and texture fractions with a large, regional dataset and External parameter orthogonalization (EPO) correction (Figure 1.14). They used field samples and scanned under the air dry conditions using the same system to implement EPO correction for in situ spectra.



Figure 1.14. Left: Veris[®] P4000 used by Veum, Parker, Sudduth, and Holan (2018) in the field. Right: close-up view of P4000 probe tip adapted for acquiring air dry spectra.

1.3 USE OF VISNIR SPECTRAL LIBRARIES IN SOIL SENSING

1.3.1 VisNIR spectral libraries

All the aforementioned sensors require calibration soil samples to be obtained from the same field or area to build models to accurately predict soil attributes. However, this is costly and time consuming. As an alternative, there has been an increasing interest within the soil community to setup large spectral libraries to be used for model calibration

(Brown et al., 2006; Shepherd & Walsh, 2002). Recently, there have been a rapid growth of spectral libraries due to ease of spectral measurement, and decrease in cost per measurement with technological advances. These spectral libraries can readily provide calibration samples for a specific model calibration. An Australian spectral library with >20,000 samples (Viscarra Rossel & Webster, 2012), a Chinese spectral library with 3993 samples (Ji et al., 2016), a global VisNIR spectra library (Viscarra Rossel & Bouma, 2016), the LUCAS database with >20,000 samples (Montanarella, Tóth, & Jones, 2011), and Rapid Carbon Assessment (RaCA) spectral library (Wills et al., 2014) in US are some examples of such large spectral libraries available.

The Rapid Carbon Assessment (RaCA) project was initiated in 2010 by the Soil Science Division of USDA-NRCS with the objective of capturing the baseline soil carbon stocks across the conterminous U.S. (CONUS). It used a multi-hierarchical design to obtain samples evenly distributed across regions based on major land resources areas (MLRA) and land use land cover classes (LULC). Samples were obtained from 6,148 sites across CONUS, described 32,084 pedons in the field, and collected 144,833 samples. A fraction of the samples (~20,000) were extracted and measured for Total Carbon, Total Nitrogen, Total Sulfur, Carbonates, and Organic Carbon. The intension was to use these measured properties to calibrate VisNIR models and use to predict for the remaining samples (Wills et al., 2014).

1.3.2 Challenges of using VisNIR libraries for soil sensors

Some of the spectral libraries are available freely or by request for the users. This enables the common users such as farmers, government agencies or researchers to use these sample to build their own models without spending money or time on calibration sampling. The true potential of VisNIR based sensors is that they can be used with soil spectral libraries to build models which can immensely reduce cost and time (Brown et al., 2006; Ge et al., 2014). However, the use of these spectral libraries for VisNIR based

sensor applications can pose a new challenge.

Spectral libraries are often constructed from dry ground soils whereas the field samples are quite different in terms of moisture, aggregation, and temperature. Among these factors, moisture has the most pronounced effect on VisNIR spectra (Bricklemyer & Brown, 2010; Kuang & Mouazen, 2013; Minasny, McBratney, Pichon, Sun, & Short, 2009; Sudduth & Hummel, 1993). There is an obvious decrease of performance when models calibrated on dry ground spectra applied to field moist soils (Bricklemyer & Brown, 2010; Minasny et al., 2009). Due to the need of large numbers of field samples, the use of field samples for model calibration is neither practical nor economical.

To bridge the gap between dry ground and field spectra, researchers are investigating different techniques such as external parameter orthogonalization (Ge et al., 2014; Minasny et al., 2011), direct standardization (Ji, Viscarra Rossel, & Shi, 2015), spiking (Brown, 2007; Sankey, Brown, Bernard, & Lawrence, 2008; Wetterlind & Stenberg, 2010), slope bias correction (Osborne & Fearn, 1983), and selective wavelength modeling (Wu, Jacobson, Laba, & Baveye, 2009).

1.4 REQUIREMENTS FOR A ROBUST VISNIR BASED SOIL SENSORS

There are four essential components to produce a robust real-time VisNIR-based proximal soil sensing system: (i) a soil (dry ground) spectral library that can be used to predict the target soil properties (such as texture and carbon forms) with a satisfactory accuracy, (ii) robust designs that can acquire high quality VisNIR spectra of soil, (iii) an effective method that can link field moist scans to the dry ground scans in the library, (iv) a complete software package with an algorithm to do real-time acquisition of spectra, do the processing, correction and display estimated soil properties on-the-go.

In actual field implementation, there are two significant sources of variation that the VisNIR sensor system should be able to account for. The first source comes from *in situ* soils themselves including moisture, surface roughness, natural aggregation, and

temperature. The second is from the field operational conditions of fluctuating temperature and humidity that cause varying responses of the optical/electronic components and the spectrometers. Since the sensor collects soil VisNIR spectra on-the-go under natural, field moist conditions, the software should be equipped with two major components: (i) dry ground spectral libraries and calibration models for targeted soil properties and (ii) a suitable correction approach (EPO or DS) to account for the field *in situ* condition. The sensor software can then convert soil VisNIR spectra to the targeted soil properties in real time.

Overall, the development of a VisNIR soil sensor involves a good design of both hardware and software. The hardware should ensure the acquisition of high-quality reflectance spectra of soil while the software should account for the spectral shifts caused by various factors including field intactness. Being able to leverage the use of external soil spectral libraries is of immense importance to make this sensor economically viable under the practical setting. The following chapters of this dissertation address some of these challenges leading to develop a VisNIR based multi-sensing penetrometer.

CHAPTER 2

MOVING FROM LAB TO FIELD: CHALLENGE OF USING DRY GROUND VISNIR LIBRARY TO PREDICT MOIST SOILS

2.1 INTRODUCTION

2.1.1 VisNIR for soil sensing

Visible and near infrared reflectance spectroscopy (VisNIR) is widely used as a rapid and cost-effective method to quantitatively infer soil properties (Chang et al., 2001; Stenberg, 2010; Viscarra Rossel, Walvoort, McBratney, Janik, & Skjemstad, 2006). Numerous studies have shown VisNIR soil spectra to successfully predict a wide array of soil properties including soil carbon (Brown et al., 2006; Minasny et al., 2011; Nocita, Stevens, Noon, & van Wesemael, 2013; Sarkhot, Grunwald, Ge, & Morgan, 2011), texture (Brown et al., 2006; Sørensen & Dalsgaard, 2005; Waiser, Morgan, Brown, & Hallmark, 2007), moisture (Ben-Dor et al., 2008; Mouazen et al., 2005; Zhu et al., 2010), and plant macro- and micro-nutrients (Ge, Thomasson, Morgan, & Searcy, 2007; Shepherd & Walsh, 2002).

There have been increasing demands for dense and spatially explicit soil data for many applications. A good example of such applications is precision agriculture, where high-resolution (meter and sub-meter scale) soil property maps are generated and used to enable variable-rate application and the delineation of crop management zones (Adamchuk, Viscarra Rossel, Hartemink, & McBratney, 2010; Wetterlind, Stenberg, & Söderström, 2008). Another good example is the Rapid Carbon Assessment (RaCA) project undertaken by USDA-NRCS Soil Science Division (Wills et al., 2014). The RaCA's goal is to estimate soil carbon (C) stocks of the USA, and more than 140,000 georeferenced soil samples were taken across the nation for this purpose. Currently, there

is an intense focus of research on soil C dynamics and inventory that requires high-resolution soil data because of the importance of soil C in global climate change, nutrient cycling, crop production and ecosystem function (Lal, 2004). Hydrological modeling, remote sensing, mass and energy transport between the atmosphere and lithosphere, and micro-climate modeling are other applications that would benefit greatly from high-resolution and spatially explicit soil data (Lagacherie, McBratney, & Voltz, 2006).

Conventional soil sampling and laboratory chemical analysis is time consuming and labor intensive and will not meet this increasing demand for soil data by many disciplines. Consequently, *in situ* proximal soil sensors are being more widely used (Ben-Dor et al., 2008; Christy, 2008; Kodaira & Shibusawa, 2013; Mouazen et al., 2007), and the cost for obtaining each soil data point is being greatly reduced. Among the many sensing principles from which *in situ* soil sensors can be developed (Adamchuk, Hummel, Morgan, & Upadhyaya, 2004), VisNIR is the most promising. It is rapid, non-destructive and, most importantly, several soil attributes can be inferred from a single scan (Brown et al., 2006).

2.1.2 Use of VisNIR libraries

The true potential of *in situ* VisNIR-based soil sensors is that they can be coupled with the soil spectral libraries and calibration models that are under development and are being tested at the regional, national and continental scales (Brown et al., 2006; Ge et al., 2014). This would minimize the need to obtain new samples for model calibration, which is the most costly part for VisNIR.

While conceptually attractive, a primary challenge for using soil spectral libraries in field applications is the wide range of soil moisture that will be encountered. Spectral libraries are most likely constructed from dry and ground soils; but field samples will be in quite different conditions in terms of moisture, small-scale heterogeneity, and

temperature. Among these three factors, moisture has the most pronounced effects on VisNIR spectra (Brickleyer & Brown, 2010; Kuang & Mouazen, 2013; Minasny et al., 2009; Sudduth & Hummel, 1993). The decrease of VisNIR model performance, when dealing with field moist soil samples, is also documented in the literature (Brickleyer & Brown, 2010; Minasny et al., 2009). Several authors used VisNIR spectra of intact soils and demonstrated successful prediction of intact soils (Gomez, Viscarra Rossel, & McBratney, 2008; Morgan, Waiser, Brown, & Hallmark, 2009; Waiser et al., 2007). But it is neither practical nor economical to develop both field moist spectral libraries and dry ground libraries due to the large number of field samples needed.

To address this challenge, researchers are investigating methods to link the VisNIR dry ground models to field moist in situ soil scans. Two approaches that have been put forward quite recently are external parameter orthogonalization (Ge et al., 2014; Minasny et al., 2011) and direct standardization (Ji et al., 2015). There are three other approaches in the literature that have aroused little attention, but could also be useful. These approaches require statistical treatments of VisNIR spectra, or calibration models or predicted variables. The following section provides the background information for each approach.

2.1.3 Approaches to link dry ground VisNIR models to moist spectra

2.1.3.1 External parameter orthogonalization (EPO)

This approach was initially introduced by Roger, Chauchard, and Bellon-Maurel (2003) to remove the effect of temperature on the prediction of Brix (sugar content of an aqueous solution) of intact apples from their NIR (near infra-red) spectra. Minasny et al. (2011) were the first to apply EPO to minimize the effect of soil moisture on OC prediction. In EPO, a VisNIR spectrum is decomposed into two orthogonal components: a useful component that has a direct relation with the response variable and a parasitic component that is affected by an external variable (soil moisture in our case). The key step in EPO is to find the EPO transformation matrix \mathbf{P} that is used to transform both the dry ground and

moist spectra. The original dry ground model needs to be recalibrated with the EPO-transformed spectra. It is a spectrum-based approach that transforms both dry ground scans and moist scans to the common useful space. The VisNIR models developed on EPO-transformed spectra appear to be insensitive to the variation in soil moisture. More recently, EPO was tested by Ge et al. (2014), Ji et al. (2015) and Ackerson, Demattê, and Morgan (2015), all with quite positive results. A detailed account of the EPO algorithm can be found in Roger et al. (2003) and Minasny et al. (2011).

2.1.3.2 Direct standardization (DS)

This was introduced first as a calibration transfer approach that allows the model calibrated on a primary instrument to be applied to spectra acquired by a secondary instrument (Feudale et al., 2002; Wang, Velkamp, & Kowalski, 1991). Ge, Morgan, Grunwald, Brown, and Sarkhot (2011) demonstrated the value of DS for the transfer of soil VisNIR models among multiple spectrometers. Ji et al. (2015) proposed that DS could potentially remove or minimize the effect of soil moisture on VisNIR modeling. The rationale is that the set of dry ground spectra is from a ‘virtual’ primary instrument, whereas the set of moist spectra is from a ‘virtual’ secondary instrument. The key step in DS is to find the transformation matrix \mathbf{F} that transforms secondary spectra to primary spectra. Direct standardization is also a spectrum-based approach, but unlike EPO, model recalibration is not required. The original dry ground model can be applied directly to DS-transformed spectra. Ji et al. (2015) demonstrated that DS could predict OC from field soil spectra with the model from laboratory dry ground scans.

2.1.3.3 Global moisture modelling (GMM)

Multivariate chemometric modeling methods such as partial least squares regression, are known for their ability to adapt to the intrinsic variation of data. If the model of a primary variable is sensitive to the variation of a secondary variable, we can create variation

intentionally in the secondary variable for the samples in model calibration. By doing so, the model is still satisfactory for the primary variable, but is more robust to variation in the secondary variable (Feudale et al., 2002; Roger et al., 2003). This is similar to the situation where an existing library (which lacks variation in the secondary variable) can be augmented (or spiked) by a set of local samples (with large variation in the secondary variable) to improve performance (Brown, 2007; Sankey et al., 2008; Wetterlind & Stenberg, 2010). The GMM approach, therefore, involves recalibration of the initial dry ground model by including more calibration samples at different moisture contents. Kawano, Abe, and Iwamoto (1995) showed the potential of this method to improve the accuracy of prediction for the Brix value of peaches with temperature as the external variable that affected the spectra.

2.1.3.4 Slope bias correction (SB)

This approach also appeared first in NIR calibration transfer (Osborne & Fearn, 1983). The assumption for the SB method is that moisture will cause a systematic error in prediction that can be corrected for through a linear, univariate slope and bias correction. Slope bias correction is different from EPO and DS because it is a response variable-based approach where correction is done on the predicted response variable.

2.1.3.5 Selective wavelength modelling (SWM)

This approach is based on research by Wu et al. (2009). The first derivatives of certain wavelength ranges, 800–1400, 1600–1700, 2100–2200 and 2300–2500 nm, are affected minimally by moisture content. The authors suggested that VisNIR calibration models based on these wavelength ranges could be applied across different moisture contents. This approach is different from the others because no spectral transformation, response variable correction or model recalibration is involved.

2.1.4 Objective

The objective of this study was to investigate and compare the usefulness of the aforementioned five approaches, namely EPO, DS, GMM, SB and SWM, in enabling VisNIR dry ground models to be applied directly to moist soil spectra to predict selected soil properties. Soil OC and inorganic carbon (IC) are the focus of this research, which is important for the development of in situ VisNIR-based soil sensors.

2.2 MATERIALS AND METHODS

2.2.1 Soil samples, VisNIR spectral measurement and sample rewetting

Soil samples used in this study were in an air-dry ground condition (<2.0 mm) and obtained from the soil archive at the Kellogg Soil Survey Laboratory of USDA-NRCS. Three hundred and fifty-two samples were selected considering two criteria. First, the samples should originate from Nebraska, USA. Second, they should represent the full range of soil OC in the archive from a 20-stratum stratified random sampling. The selected samples were then divided into two subsets: the dry ground set (n=185) and the rewetting set (n=167). Samples in the dry ground set were scanned once only in the dry ground state. An ASD LabSpec[®] spectrometer with a mug light (Analytical Spectral Devices, Boulder, Colorado, USA) was used for soil scanning. The spectrometer has a spectral range from 350 to 2500 nm and a spectral sampling interval of 1 nm. Samples were placed in petri dishes for scanning. A Spectralon[®] panel (Labsphere Inc., North Sutton, NH, USA) was also set up in a petri dish and used as the white reference. Each spectrum was obtained from an average of 100 instantaneous internal scans. The spectral range from 350 to 499 nm was removed from data analysis because of the considerable noise in these wavebands.

Samples in the rewetting set were scanned at eight different moisture contents and in the dry ground state. First, a soil rewetting experiment was carried out on all rewetting samples; the procedure was as follows. The samples were divided into four individual

batches consisting of 35, 46, 46 and 40 samples for easy handling. For each sample in batch one, approximately 10 g of soil was placed in a petri dish and the VisNIR spectrum was obtained (this is the dry ground scan). This was followed by adding a known amount of deionized water to achieve a moisture content of $\sim 33\%$ (dry basis, gravimetric). The petri dish was then covered with a glass lid for ~ 24 hours to allow homogenization of the moisture within the sample. Then the sample was weighed and scanned, and the petri dish was uncovered to facilitate moisture loss. When the sample reached the next desired moisture content, the petri dish was again covered with the lid for 24 hours (for moisture homogenization within samples), followed by scanning and weighing for the second moisture content. This step was repeated to obtain scans at eight different moisture contents. The same procedure was applied to other batches, except that 15 g of soil was used for each sample for better control of the moisture content. Almost the same moisture contents were maintained for all batches; the average moisture contents were 33, 29, 25, 21, 17, 13, 10, 8% and air dry. All moisture contents were calculated as dry basis adjusted by the ADOD (air-dried weight/oven-dried weight) value of each sample. For the rewetted samples, ADOD values ranged from 1.002 to 1.082, with an average of 1.027.

2.2.2 Data subsets and analysis

The entire dataset was divided into three subsets for data analysis as follows.

2.2.2.1 Dry ground data set (S^0)

This set consisted of 185 samples that were scanned in their dry ground state. It was used to develop the initial dry ground VNIR model.

The rewetted samples (n=167) were divided into two subsets.

2.2.2.2 Transformation data set (S^1)

This set consisted of 100 rewetted samples. The number of spectra in S^1 is 900 (water contents 33, 29, 25, 21, 17, 13, 10, 8% and air dry). This set was used to develop transformations for the five different approaches (e.g. EPO transformation matrix \mathbf{P} or DS transformation matrix \mathbf{F}).

2.2.2.3 Test data set (S^2)

This set consisted of the remaining 67 rewetted samples. The number of spectra in S^2 is 603. This set was used for independent validation of the five approaches under study. The summary statistics of soil OC and IC in the different subsets are given in Table 2.1.

Table 2.1. Summary statistics of soil organic carbon and inorganic carbon in the three subsets (S^0 , S^1 and S^2) in this study.

		S^0	S^1	S^2
Soil Property	Number of samples (n)	185	100	67
	Number of scans (N)	185	900	603
Organic C (%)	Minimum	0	0.1	0
	Median	0.8	1.4	1.1
	Maximum	6.2	6.5	5.7
	Interquartile range	1.2	1.3	1.5
	Kurtosis	7.2	4.9	7.6
	Skewness	1.9	1	2
	Shapiro-Wilk statistic	1.1×10^{-14}	1.6×10^{-7}	2.8×10^{-6}
Inorganic C (%)	Minimum	0	0	0
	Median	0.1	0.1	0.1
	Maximum	3.8	1.9	3.9
	Interquartile range	0.3	0.2	0.8
	Kurtosis	11.5	14.1	5.1
	Skewness	2.8	3.1	1.9
	Shapiro-Wilk statistic	2.2×10^{-16}	1.4×10^{-9}	8.6×10^{-7}

2.2.3 Analyses to link dry ground and moist spectra

Below we describe the procedures of the five approaches (EPO, DS, GMM, SB and SWM) mentioned above; their flow charts are given in Figure 2.1.

First, the dry ground dataset \mathbf{S}^0 was used to calibrate a partial least squares regression (PLSR) model (correlated soil OC and IC with the dry ground spectra), which was used to predict the test dataset (\mathbf{S}^2) at different moisture contents (dry ground prediction). Because no effort was made to correct for the effect of moisture, dry ground prediction is expected to be a weak predictor and it was used as a baseline against which other predictions were compared.

For EPO, the transformation set \mathbf{S}^1 was used to develop an EPO transformation matrix \mathbf{P} . The dry ground spectra in \mathbf{S}^0 were then transformed with the matrix \mathbf{P} , and the transformed spectra were used to calibrate an EPO model. Matrix \mathbf{P} was also applied to the moist spectra in \mathbf{S}^2 , which has eight different moisture contents. The EPO model was then applied to EPO-transformed \mathbf{S}^2 spectra to predict soil OC and IC at the different moisture contents. Figure 2.1(a) shows the EPO modeling and validation procedure. Due to its complexity and importance, a summary of the mathematical procedure (Figure 2.2) is given below. The details can be found in Roger et al. (2003).

External Parameter Orthogonalization assumes that spectra matrix \mathbf{X} can be decomposed into two systematic components: a useful component \mathbf{XP} and a parasitic component \mathbf{XQ} , as indicated in Eq. 2.1. \mathbf{R} is the noise component originated from lack of fitting.

$$\mathbf{X} = \mathbf{XP} + \mathbf{XQ} + \mathbf{R} \quad (2.1)$$

The procedure to find \mathbf{XP} is through spectra matrix \mathbf{D} , which is the difference between the spectra matrix with and without external influence. \mathbf{Q} is estimated through singular value decomposition of \mathbf{D} , and \mathbf{XP} is then calculated as $\mathbf{X}(\mathbf{I}-\mathbf{Q})$; \mathbf{I} is the identity

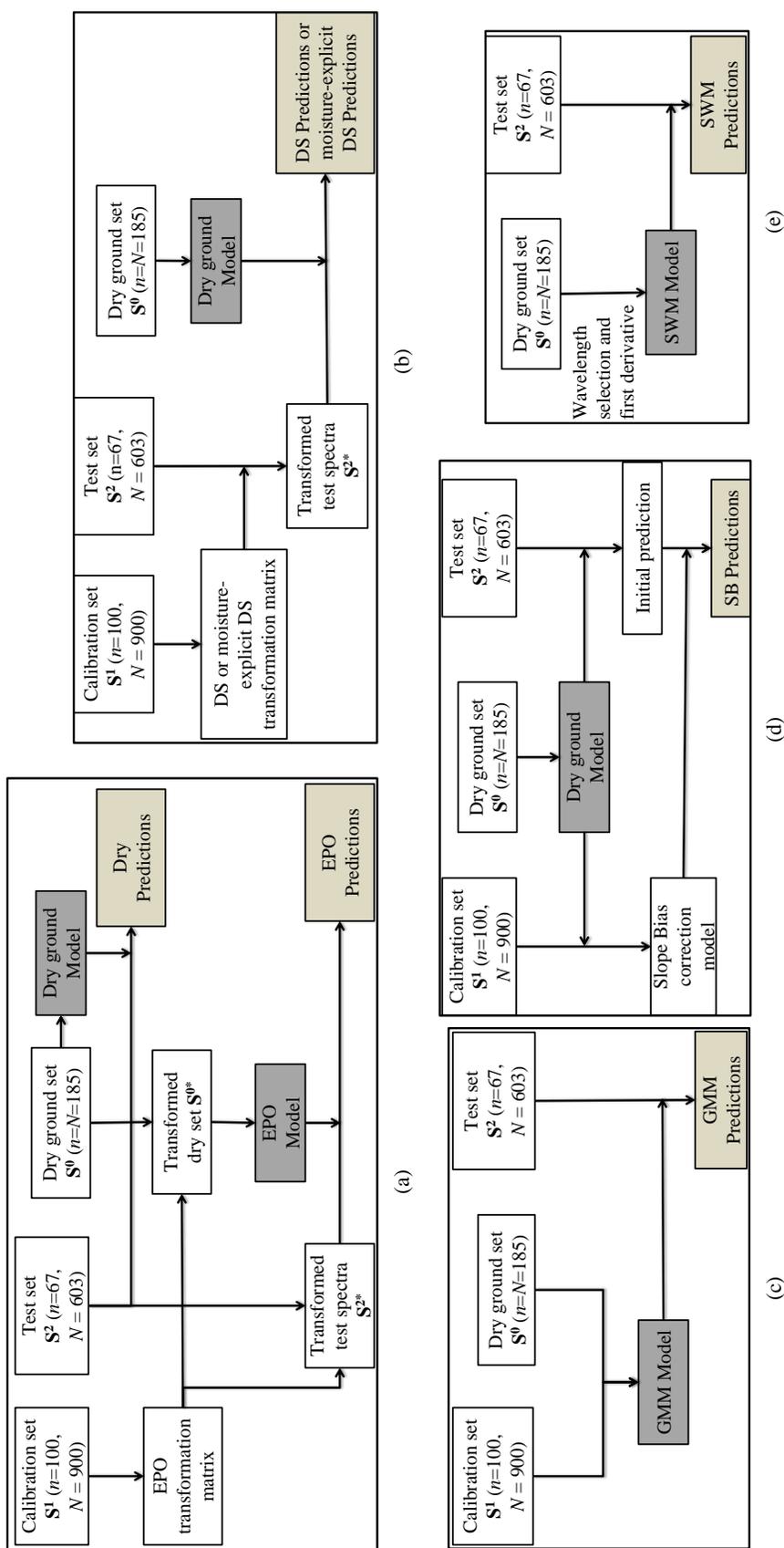


Figure 2.1. Flow charts of the five approaches for moisture in soil VisNIR reflectance: (a) external parameter orthogonalization (EPO), (b) moisture-explicit direct standardization (moisture-explicit DS) and direct standardization (DS), (c) global moisture modeling (GMM), (d) slope bias correction (SB) and (e) selective wavelength modeling (SWM).

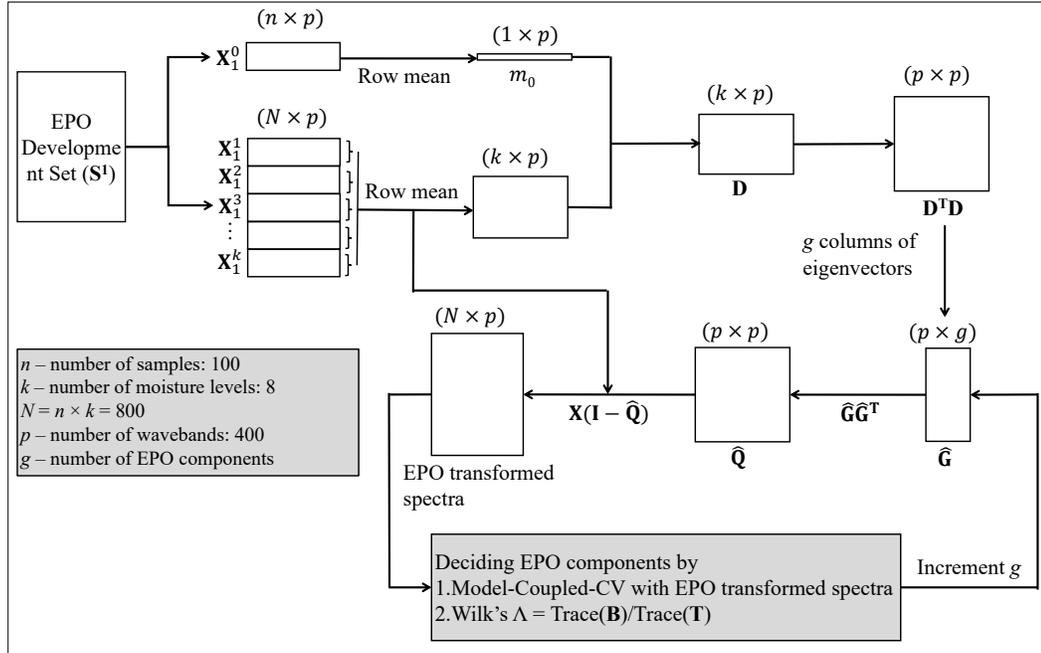


Figure 2.2. Implementation of external parameter orthogonalization (EPO) transformation with the model-coupled-Cross Validation (model-coupled-CV) and Wilk's Λ method. The matrix symbols drawn in the figure are for the understanding of matrix operations in the EPO procedure.

matrix.

One of the most important parameters to be determined during EPO development is the number of EPO components g (the same notation as in Roger et al. (2003) and Minasny et al. (2011); Ge et al. (2014) used c for the same meaning). Roger et al. (2003) suggested two methods to determine g : (1) Cross validation of PLS calibration on transformed spectra \mathbf{S}^{1*} (PLS-CV); and (2) calculating Wilk's Λ of the transformed spectra \mathbf{S}^{1*} as:

$$\text{Wilk's } \Lambda = \frac{\text{Trace}(\mathbf{B})}{\text{Trace}(\mathbf{T})} \quad (2.2)$$

where \mathbf{T} is the variance-covariance matrix of the EPO transformed spectra \mathbf{S}^{1*} , and \mathbf{B} is the variance-covariance matrix of \mathbf{S}^{1*} aggregated by sample (i.e., averaging across all moisture levels for each sample).

For DS, two different implementations were carried out. In the first (referred to as

moisture-explicit DS), the dataset S^1 was used to develop eight DS transformation matrices that corresponded to each moisture content. The moist spectra of S^2 were then transformed to a dry ground state by the corresponding transformation matrix. The dry ground model was then applied to the DS-transformed spectra to predict OC and IC of the S^2 samples at each moisture content. In the second implementation (simply referred to as DS), we created a new set of data ($n=100$, $N=100$) by randomly selecting one moisture content from each sample in S^1 . This set of scans was used to compute an overall DS transformation matrix, which was then used to transform all moist spectra in S^2 . The rationale for the moisture-explicit DS is that the spectra at different moisture contents should be treated as different ‘secondary’ states. For field samples, however, soil moisture content will not be known explicitly. This renders moisture-explicit DS impractical. Therefore, an investigation of overall DS transformation with all moisture contents involved is needed and more practically valuable. In Ji et al. (2015), DS was implemented in the second way. Figure 2.1(b) shows the procedure for moisture-explicit DS and DS.

The GMM does not involve spectral transformation as in EPO and DS. Instead, the S^1 and S^0 datasets were combined and a global model (in terms of the variation in moisture) was calibrated with the combined set to account for the variation in moisture in the spectra. The global model was then applied directly to S^2 to predict OC and IC at different moisture contents. Figure 2.1(c) shows the flow chart of GMM modeling and validation.

The SB method does not involve spectral transformation either. Instead, the predictions are corrected. First, the dry ground model was used to predict OC and IC of the S^1 spectra at different moisture contents. Eight different SB corrections that represent the linear relation between the predictions in the dry state and each moisture state were developed. The dry ground model was then used to predict OC and IC for S^2 at different moisture contents and the predictions were corrected with the SB corrections at each moisture content. Figure 2.1(d) shows the procedure used for SB correction.

For SWM, first we calculated the first derivative of soil spectra at the selected wavelength ranges: 800–1400, 1600–1700, 2100–2200 and 2300–2500 nm (Wu et al., 2009). An SWM model was then calibrated with the \mathbf{S}^0 samples, and then applied to all moist spectra in \mathbf{S}^2 . Dataset \mathbf{S}^1 was not used in this approach. Figure 2.1(e) shows the modeling procedure.

Partial least squares regression (PLSR) was used for the calibration of all models. The sizes of the PLSR models were determined by selecting the number of PLSR factors that gave the first local minimum in the root mean squared error of cross-validation (RMSE_{CV}) with 25 random segments. The coefficient of determination (R^2), root mean squared error of prediction (RMSE_{P}), ratio of performance to interquartile range (RPIQ, Bellon-Maurel, Fernandez-Ahumada, Palagos, Roger, and McBratney, 2010) and bias were used to compare the performance of the five moisture correction approaches. The RPIQ was used instead of the ratio of performance to deviation (RPD) because the data subsets had non-normal distributions (Table 2.1). Data analysis was carried out in the R statistical environment (R Core Team, 2018) with the pls (Mevik, Wehrens, & Liland, 2013), gnm (Turner & Firth, 2015), psych (Revelle, 2015) and ggplot2 (Wickham, 2009) packages.

2.3 RESULTS AND DISCUSSION

2.3.1 Comparison of the five approaches to account for moisture effect

The initial dry ground model developed from \mathbf{S}^0 for OC had a cross-validation R^2 of 0.59 and RMSE_{CV} of 0.74%. For IC, the model had an R^2 of 0.71 and RMSE_{CV} of 0.33%.

Table 2.2 summarizes the accuracy of prediction of soil OC and IC for the five moisture correction approaches. The result for no moisture correction is also given in Table 2.2. The statistics were calculated by combining all eight moisture contents. Prediction for dry ground spectra is not included in this comparison because the purpose is to compare the performance of each approach for moist spectra prediction.

Table 2.2. Performance of five different moisture correction approaches to predict soil organic carbon and inorganic carbon.

Moisture correction approach	Organic C					Inorganic C						
	R ²	RMSEP ^a (%)	RPIQ ^b	Bias (%)	R ²	RMSEP ^a (%)	RPIQ ^b	Bias (%)	R ²	RMSEP ^a (%)	RPIQ ^b	Bias (%)
No Correction	0.01	3.97	0.38	3.25	0.62	1.52	0.53	-1.26				
External parameter orthogonalization (EPO)	0.55	0.87	1.73	0.02	0.63	0.66	1.21	-0.23				
Moisture-explicit direct standardization	0.58	0.84	1.8	-0.02	0.69	0.5	1.61	-0.01				
Direct standardization (DS)	0.55	0.88	1.7	-0.26	0.77	0.63	1.27	0.12				
Global moisture modelling (GMM)	0.61	0.78	1.91	-0.11	0.6	0.57	1.41	-0.08				
Slope bias correction (SB)	0.02	1.25	1.2	-0.08	0.65	0.71	1.12	-0.19				
Selective Wavelength Modeling (SWM)	0	4.17	0.36	3.4	0.55	0.97	0.82	0.67				

^aRoot Mean Squared Error of Prediction; ^bRatio of Performance to Inter-quartile range

For soil OC, EPO, moisture-explicit DS, DS and GMM increase R^2 from almost 0 to greater than 0.55 and increase RPIQ from 0.38 to more than 1.7 (Table 2.2). The largest accuracy was obtained by GMM with an R^2 of 0.61 and RPIQ of 1.91. The model improvements for these four approaches are related mainly to two factors. First, there is a large increase in the linear correlation between model-predicted and laboratory-measured OC (increase in R^2). Second, there is an appreciable decrease in the systematic bias of prediction that is caused by the presence of moisture (decrease in $RMSE_P$ and bias). The improvements produced by SB for the prediction of OC are marginal only. There appears to be no improvement in the linear correlation between the predicted and measured variables (R^2 remains small, 0.02) (Table 2.2). However, both bias and $RMSE_P$ decrease substantially and RPIQ increases, which suggest that SB does partially account for the systematic prediction bias caused by soil moisture. The SWM method fails completely; its R^2 , $RMSE_P$, RPIQ and bias are similar to those for no correction.

For soil IC, the dry ground model applied directly to the moist scans achieved an R^2 value of 0.62, which suggests a strong linear relation between predicted and measured IC (Table 2.2). The small RPIQ value (0.53), however, indicates poor model performance, which arises from the systematic bias of prediction caused by moisture. For this reason, the improvement in R^2 after EPO, moisture-explicit DS, DS and GMM correction is quite marginal (R^2 for GMM even decreases slightly). However, all of these approaches improve the $RMSE_P$, bias and RPIQ, which suggests that the systematic bias in prediction was effectively removed from the original dry ground model. Table 2.2 shows that the prediction of soil IC by the SB and SWM approaches is better than that of OC, but to a lesser extent than for the other four approaches.

Figure 2.3 shows the scatterplot of measured against predicted OC with no correction and with GMM correction for the eight moisture contents. Similarly, Figure 2.4 shows the scatterplot of measured against predicted IC with no correction and moisture-explicit DS correction. These plots illustrate how moisture correction improves

prediction at all moisture contents in terms of R^2 , $RMSE_p$, RPIQ and bias.

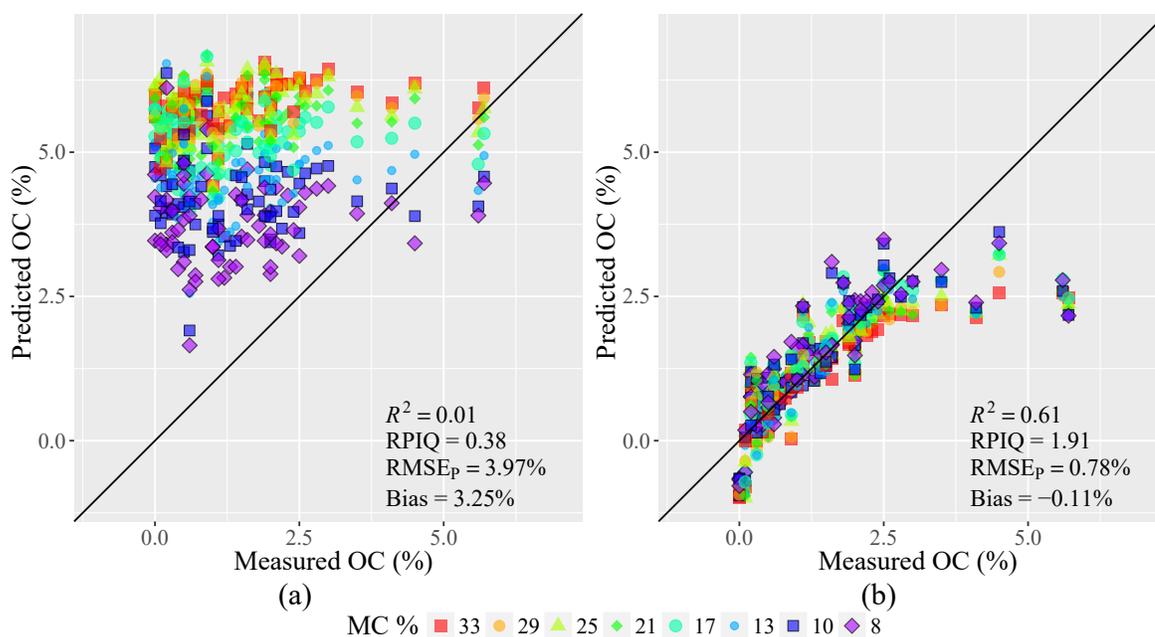


Figure 2.3. Scatterplots of soil organic carbon (OC) prediction at eight different moisture contents: (a) no correction and (b) global moisture modeling (GMM). The solid line is the 1:1 line.

Overall, EPO, moisture-explicit DS, DS and GMM improved predictions of both soil OC and IC at various moisture contents. Slope bias correction provides a marginal improvement in prediction for both soil properties, and there is also a marginal improvement for the prediction of soil IC with SWM but it fails for soil OC prediction. This result is not surprising because, as stated above, the effect of moisture on soil VisNIR spectra is nonlinear and complex.

External parameter orthogonalization, moisture-explicit DS, DS and GMM involve some form of transformation of spectral matrices or model recalibration, or both. Therefore, the complex effect of moisture on soil spectra and modeling can be accounted for better. Slope bias correction, however, uses only simple linear corrections of predicted variables that are clearly not adequate to account for the complex and nonlinear effect of soil moisture. Our results accord with those of Minasny et al. (2011) and Ji et al. (2015),

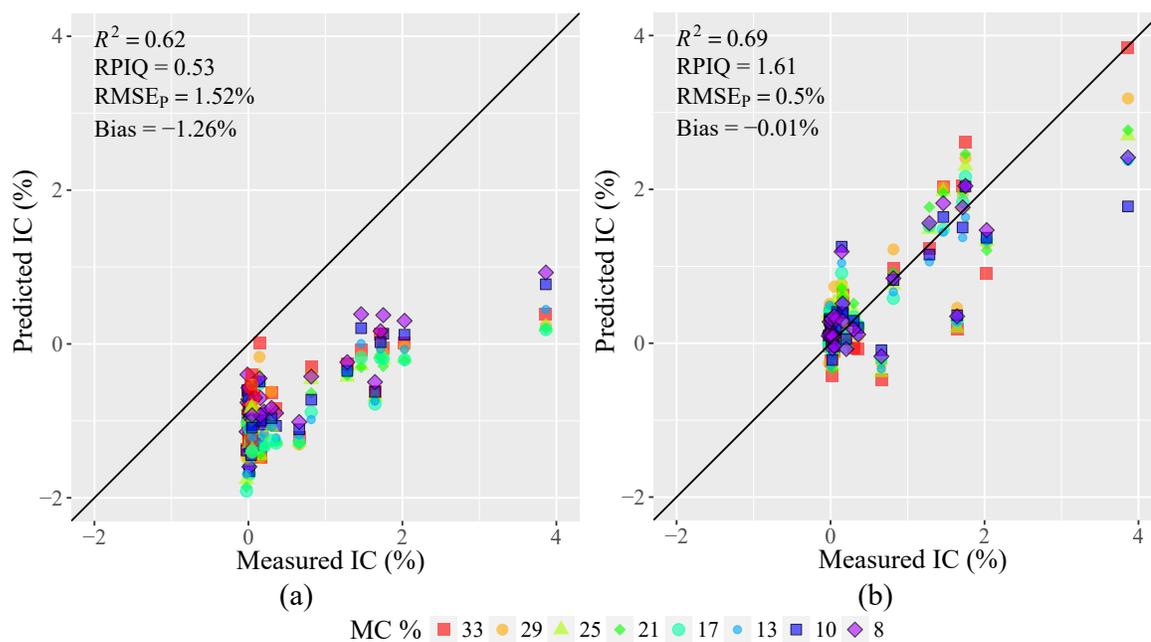


Figure 2.4. Scatterplots of soil inorganic carbon (IC) prediction at eight different moisture contents: (a) no correction and (b) moisture-explicit direct standardization (moisture-explicit DS). The solid line is the 1:1 line.

who applied EPO and DS, respectively, to predict soil OC in the presence of soil moisture.

2.3.2 Comparison between EPO, moisture-explicit DS and DS

Ji et al. (2015) was the only paper in the literature that compared DS with EPO, and they found that DS was marginally superior to EPO for soil organic matter. Our result also shows that for OC, EPO and DS perform similarly well; for IC, DS is slightly better than EPO. Originally, we thought that moisture-explicit DS would outperform EPO and DS because eight different transformation matrices were developed for this method, each dedicated to correcting for one moisture content, whereas in EPO and DS only one transformation matrix is developed for all moisture contents. To our surprise, however, moisture-explicit DS performed the same for OC and was just slightly superior for IC (in terms of RMSEP, RPIQ and bias). It seems that this method does not require moisture-specific correction.

Figure 2.5 illustrates transformations by EPO, moisture-explicit DS and DS. The first row shows the spectra of a randomly selected soil sample scanned at different moisture contents. The middle row shows the transformation matrices and the bottom row the transformed spectra. Note that there are eight moisture-explicit DS matrices; only one example of the transformation matrix (for 21% moisture content) is shown in Figure 2.5.

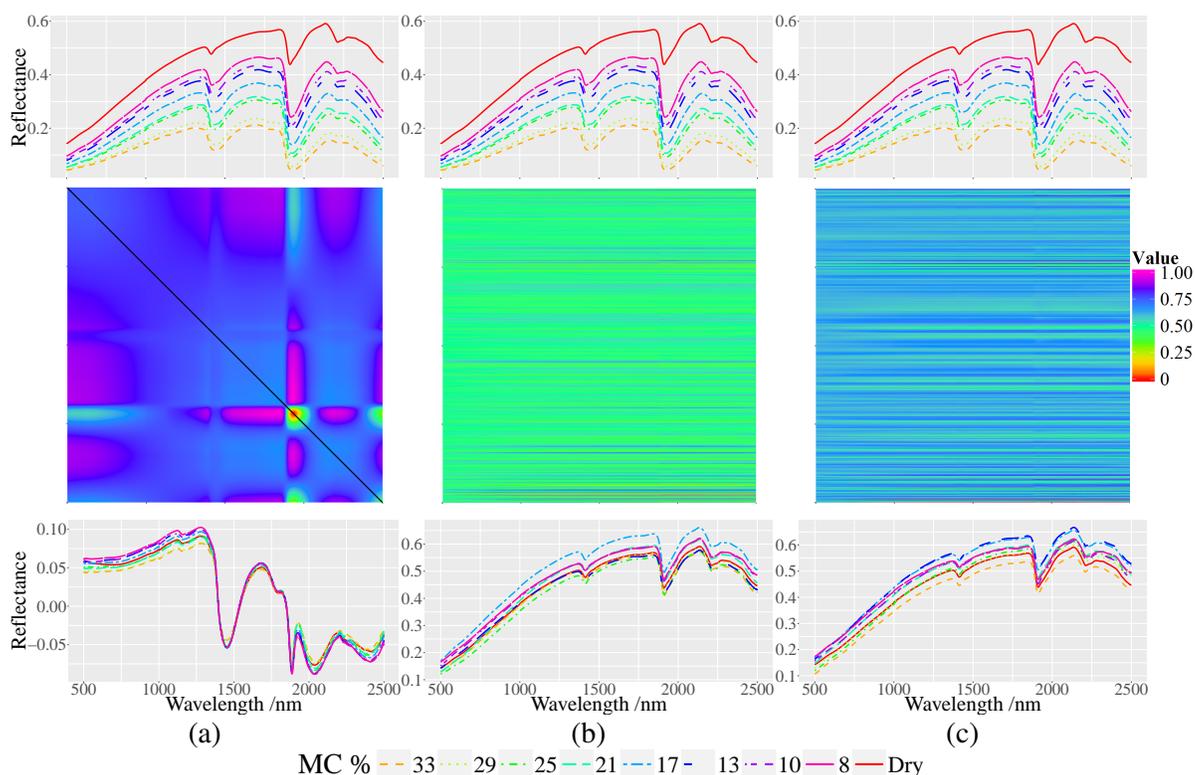


Figure 2.5. Spectral transformation at different moisture contents by: (a) external parameter orthogonalization (EPO), (b) moisture-explicit direct standardization (moisture-explicit DS) and (c) direct standardization (DS). There are eight moisture-explicit DS transformations, one for each moisture content; the transformation matrix at 21% moisture content is shown in (b). The first row indicates the spectra at different moisture contents before spectral transformation. The second row shows the transformation matrices. The third row illustrates the spectra after transformation at different moisture contents.

There are a few important points to note from Figure 2.5. First, EPO, moisture-explicit DS and DS all reduce within-sample variation in VisNIR spectra caused by differences in moisture content. This can be seen in the transformed spectra, which have much less variation than the original spectra. Second, EPO, moisture-explicit DS and

DS use different methods for moisture correction. The EPO transformation matrix (Figure 2.5a) appears to be quite different from the moisture-explicit DS (Figure 2.5b) and DS transformation matrices (Figure 2.5c). The EPO-transformed spectra have a different spectral shape, and the magnitude of reflectance values has decreased. This is why a model-recalibration step is needed for EPO. Moisture-explicit DS and DS, on the contrary, strive to match the moist spectra at different moisture contents to the dry ground spectrum. Therefore, the original dry ground model is still valid, and no model recalibration is needed. Finally, the transformation matrices of moisture-explicit DS and DS appear to be similar, but they have quite different values for the matrix elements, and their transformed spectra deviate little from the original dry ground spectra. This supports the result in Table 2.2 that DS can be effective in moisture correction and provide substantial improvements in the prediction of soil OC and IC.

2.3.3 Comparison of the results of moisture correction at different moisture contents

In Table 2.2, EPO, moisture-explicit DS, DS and GMM show the potential to be able to correct for the moisture effect on soil VisNIR spectra and improve the prediction of soil OC and IC. A further question that arises is, do these moisture correction approaches perform equally well for different moisture contents? In other words, for two groups of soil samples, one with large and one with small moisture content, can these approaches make the correction equally well and predict with comparable accuracy? This is important because in the field the moisture content of the soil will vary. A satisfactory approach to moisture correction should show no bias in relation to the soil's moisture content. Table 2.3 summarizes the performance of EPO, moisture-explicit DS, DS and GMM at the eight moisture contents for the prediction of OC.

Table 2.3 shows that for EPO there is a consistent decrease in the accuracy of prediction for OC with the increase in moisture content in the samples (expressed by a continuous decrease in R^2 and RPIQ, and an increase in $RMSE_P$). The accuracy of

Table 2.3. Performance of external parameter orthogonalization (EPO), moisture-explicit direct standardization, direction standardization (DS), and global moisture modeling (GMM) across the eight moisture contents for soil organic C prediction.

Moisture correction approach	Parameter	Moisture content (%)										
		8	10	13	17	21	25	29	33			
EPO	R ²	0.71	0.7	0.65	0.57	0.52	0.49	0.45	0.4			
	RMSE _P ^a (%)	0.73	0.73	0.78	0.86	0.9	0.93	0.97	1.01			
	Bias (%)	0.18	0.12	0.03	-0.05	-0.08	-0.06	0.02	0.11			
	RPIQ ^b	2.05	2.05	1.92	1.74	1.67	1.61	1.55	1.49			
Moisture-explicit DS	R ²	0.58	0.68	0.56	0.53	0.62	0.58	0.48	0.62			
	RMSE _P ^a (%)	0.84	0.74	0.84	0.88	0.79	0.85	0.92	0.8			
	Bias (%)	-0.02	-0.02	-0.02	-0.03	-0.02	-0.03	-0.02	-0.03			
	RPIQ ^b	1.79	2.03	1.79	1.7	1.9	1.76	1.63	1.88			
DS	R ²	0.57	0.62	0.62	0.53	0.56	0.56	0.56	0.48			
	RMSE _P ^a (%)	0.83	0.78	0.78	0.92	0.89	0.96	0.9	0.99			
	Bias (%)	-0.10	-0.16	-0.13	-0.28	-0.28	-0.44	-0.32	-0.34			
	RPIQ ^b	1.81	1.92	1.92	1.63	1.69	1.56	1.67	1.52			
GMM	R ²	0.61	0.6	0.61	0.63	0.61	0.63	0.61	0.67			
	RMSE _P ^a (%)	0.78	0.79	0.79	0.76	0.79	0.78	0.8	0.78			
	Bias (%)	0.01	-0.04	-0.13	-0.05	-0.15	-0.16	-0.19	-0.26			
	RPIQ ^b	1.92	1.9	1.9	1.97	1.9	1.92	1.88	1.92			

^aRoot Mean Squared Error of Prediction; ^bRatio of Performance to Inter-quartile range

prediction with DS also depends in general on moisture content. The accuracy decreases appreciably when moisture content is larger than 17%. On the other hand, moisture-explicit DS shows no clear trend in the accuracy of prediction for various moisture contents; the model statistics fluctuate with the increase in moisture content. This is not surprising because again moisture-explicit DS develops one transformation matrix for each moisture content. Finally, GMM seems to be quite insensitive to the moisture content of the sample sets; R^2 , $RMSEP$ and RPIQ remain stable with the increase in moisture content, but there is a slight increase in bias (absolute value).

Table 2.4 summarizes the accuracy of prediction of soil IC by EPO, moisture-explicit DS, DS and GMM at the eight moisture contents. Again, prediction accuracy of EPO depends strongly on moisture content; there is a consistent decrease in R^2 and RPIQ and increase in $RMSEP$ and bias (absolute value) as moisture content increases.

Moisture-explicit DS does not show a clear trend in the accuracy of prediction with moisture content. Direct standardization differs from other approaches. The R^2 values remain quite stable and there is a substantial decrease in $RMSEP$ and bias with the increase in moisture. This leads to a substantial increase in RPIQ and therefore in the model's prediction performance. This is the only example in the study where there is an increase in the accuracy of prediction with the increase in moisture. Lastly, GMM shows a consistent but only small decrease in the accuracy of prediction as moisture content increases.

For prediction of both soil OC and IC, moisture-explicit DS and GMM appear to be the best approaches for moisture correction because their predictions depend little on moisture content. The correction and accuracy of prediction by EPO for both soil properties was superior to the previous two methods at smaller moisture contents (8 and 10%), but they decline when the moisture content increases. This suggests that EPO might not bring adequate correction for spectra with large moisture contents.

Table 2.4. Performance of external parameter orthogonalization (EPO), moisture-explicit direct standardization, direction standardization (DS), and global moisture modeling (GMM) across the eight moisture contents for soil inorganic C prediction.

Moisture Correction Approach	Parameter	Moisture content (%)										
		8	10	13	17	21	25	29	33			
EPO	R ²	0.81	0.82	0.76	0.66	0.63	0.58	0.56	0.51			
	RMSE _P ^a (%)	0.46	0.48	0.56	0.69	0.72	0.75	0.74	0.78			
	Bias (%)	-0.15	-0.17	-0.20	-0.28	-0.31	-0.33	-0.28	-0.31			
	RPIQ ^b	1.73	1.68	1.43	1.16	1.1	1.06	1.08	1.03			
Moisture-explicit DS	R ²	0.7	0.62	0.71	0.72	0.69	0.68	0.74	0.72			
	RMSE _P ^a (%)	0.81	0.92	0.84	0.79	0.83	0.84	0.76	0.83			
	Bias (%)	0	-0.05	-0.09	-0.04	0.06	0.03	0.04	-0.01			
	RPIQ ^b	0.99	0.87	0.96	1.01	0.97	0.95	1.05	0.96			
DS	R ²	0.85	0.83	0.84	0.79	0.8	0.79	0.8	0.79			
	RMSE _P ^a (%)	0.95	0.86	0.66	0.57	0.48	0.46	0.41	0.42			
	Bias (%)	0.44	0.38	0.25	0.17	0.1	0.05	0.05	-0.05			
	RPIQ ^b	0.85	0.93	1.22	1.42	1.65	1.74	1.94	1.9			
GMM	R ²	0.63	0.65	0.63	0.63	0.62	0.59	0.57	0.57			
	RMSE _P ^a (%)	0.55	0.53	0.54	0.55	0.57	0.58	0.6	0.6			
	Bias (%)	-0.01	-0.04	-0.09	-0.12	-0.12	-0.10	-0.11	-0.09			
	RPIQ ^b	1.45	1.5	1.47	1.44	1.42	1.38	1.33	1.34			

^aRoot Mean Squared Error of Prediction; ^bRatio of Performance to Inter-quartile range

2.3.4 Practical considerations: field implementation of the moisture correction approaches

There are three essential components for a viable VisNIR-based proximal soil sensing system. First, high quality soil (dry ground) spectral libraries are required that can be used to predict an array of soil properties (such as texture and carbon forms) accurately. Second, robust and flexible designs are needed that can acquire high quality VisNIR spectra of soil surfaces (or along soil profiles). Third, there must be effective methods that can link field moist scans to the dry ground spectral libraries in the laboratory. Although there has been considerable advancement in the literature on the first two aspects above, effective moisture correction approaches for soil VisNIR modeling have not been investigated thoroughly. This is a major stumbling block for *in situ* VisNIR technology, and this study fills that gap by comparing various moisture correction approaches to predict soil OC and IC at different moisture contents.

We now consider how a VisNIR-based *in situ* soil sensor would work in the field. The sensor will collect soil VisNIR spectra on-the-go in natural, field moist conditions. The sensor's software will be equipped with two major components: (i) dry ground spectral libraries and calibration models for targeted soil properties and (ii) a suitable moisture correction approach (EPO, DS or GMM). The sensor software can then convert soil VisNIR spectra to the targeted soil properties in real time.

In this context, the moisture content of field samples is not known a priori. This would preclude the moisture-explicit DS approach because we do not know which DS transformation should be selected. On the other hand, EPO, DS and GMM would be useful because they use one transformation or model recalibration for all (unknown) moisture contents.

Another practical concern is the effort and labor needed to acquire soil samples at different moisture contents for the implementation. We created our moist set by rewetting archive soil samples to obtain a few predetermined moisture contents. Alternatively,

others have used moist samples from the field; there is evidence that the correction done in this way also works for EPO (Ge et al., 2014) and DS (Ji et al., 2015). There are differences between using the archive and actual field samples, both with their respective advantages and disadvantages. In addition to moisture, natural aggregation also affects soil VisNIR spectra and the prediction performance of dry ground models (Ge et al., 2014). The use of actual field samples might provide the opportunity to correct for this factor also (possibly implicitly through EPO or DS). The acquisition of VisNIR spectra for field samples (either by transporting a VisNIR spectrometer to fields or by carefully preserving the natural state of samples to scan in the laboratory) may, however, be challenging logistically. Rewetted archive samples do not involve fieldwork, and the researchers can have more control over the targeted moisture contents and soil properties to improve the performance of moisture correction algorithms.

Regardless of the methods being used, there are two critical questions that need to be addressed in future research. The first is: are the moisture correction approaches investigated in this study transferable? Can transformations developed in one scenario be used in another, for example a different field, geographic region or soil type? This is the key for the development of a universally functional moisture correction algorithm for the VisNIR-based soil proximal sensor. The second question is: what is the optimal proportion of moist samples required in relation to the dry ground library to achieve good moist correction? In our study, the correction set has 100 samples (with 8 moisture levels) in relation to a dry ground sample set of 185. An answer to this question would minimize the effort and time needed to develop adequate moisture correction transformations.

2.4 CONCLUSIONS

Moisture correction is an essential component of the development of VisNIR-based proximal soil sensors if the existing dry ground soil spectral libraries and calibration models are to be used. We found that EPO, moisture-explicit DS, DS and GMM corrected

for the moisture effect successfully and increased substantially the accuracy of prediction for both soil OC and IC in this research, whereas SB and SWM were not successful. Global moisture modeling and moisture-explicit DS showed a consistent correction performance across different moisture contents, which is desirable for field samples. Although moisture-explicit DS had larger prediction accuracy overall, its use in the field would be limited because the moisture content of field samples is not known a priori. Any of the other three approaches can be integrated with the VisNIR-based soil proximal sensor for real-time moisture correction and soil property determination. Two future research directions include (i) the investigation of a universally functional moisture correction algorithm for the VisNIR-based soil proximal sensor and (ii) the optimal proportion of moist samples required in relation to the dry ground library to achieve good moisture correction.

CHAPTER 3

DESIGN, DEVELOPMENT, AND LAB TESTING OF A VISNIR OPTICAL PROBE FOR SOIL SENSING

3.1 INTRODUCTION

3.1.1 Soil sensing technologies

The demands for spatially explicit, high resolution 3D soil data are increasing for many disciplines such as precision agriculture, hydrological and ecological modeling, climate modeling, and land resource management. Digital soil mapping is a successful sub-discipline of soil science where traditional and modern soil measurement and inference methods are used in a holistic framework to provide 3D quantitative soil data across different scales (Minasny & McBratney, 2016).

Several modern techniques are available that can characterize the horizontal variation of soils effectively. These techniques include aerial and satellite remote sensing (Barnes et al., 2003; Ben-Dor, 2002), on-the-go (*in-situ*) soil sensors (Adamchuk et al., 2004), apparent electrical conductivity/resistivity (Corwin & Lesch, 2003), ground penetrating radar (Davis & Annan, 1989), and γ -ray sensors (Triantafilis, Gibbs, & Earl, 2013). They are now widely used to produce high resolution soil maps that capture lateral variability. On the other hand, technologies to characterize soil profiles without pulling a soil core or opening a soil pit are still limited (Hartemink & Minasny, 2014). While technologies like soil apparent electrical conductivity do sense soil to a certain depth, they return a signal that is a weighted integration of the profile and are not capable of providing layered information directly.

3.1.2 VisNIR technology for vertical soil sensing

Visible and near infrared reflectance spectroscopy (VisNIR) has a few advantages that make it a viable technology for *in situ* soil sensing in general and high resolution vertical soil sensing in particular: (1) the non-contact nature of interactions between VisNIR electromagnetic energy and soils, (2) ease of miniaturization for a VisNIR sensing probe, and (3) the ability to interrogate multiple soil properties from one scan. There have already been a number of reports on the development and application of *in situ* VisNIR sensors for surface (lateral) soil mapping (Bricklemyer & Brown, 2010; Christy, 2008; Kodaira & Shibusawa, 2013; Mouazen et al., 2007).

Several authors conducted studies where soil cores were collected and scanned in the lab at fine vertical resolution to obtain “simulated” *in situ* soil profile sensing. Hummel et al. (2001) collected 48 cores from Illinois and soil samples (at 2.5 cm increment from the cores) were scanned at six moisture levels to predict soil organic matter and moisture. Waiser et al. (2007) and Morgan et al. (2009) conducted simulated profile characterization of clay and soil organic and inorganic carbon, respectively, with the 72 soil cores collected from Texas. Doetterl, Stevens, Van Oost, and van Wesemael (2013) collected 151 soil cores from central Belgium and demonstrated the simulated high vertical resolution (3 cm) soil organic carbon assessment with VisNIR. While these are not true *in situ* sensing, they establish the technical feasibility and demonstrate the potential of a VisNIR-based vertical sensing system.

Demonstrations of *in situ* vertical soil sensing are only sporadically reported in the literature. Ben-Dor et al. (2008) described a custom-made spectral head device that can penetrate into the subsoil to measure VisNIR reflectance along the profile after a hole is drilled. Wetterlind et al. (2015) and Veum et al. (2018) tested a commercial Veris[®] P4000 probe (Veris[®] technologies Inc., Salina, Kansas, USA) for the measurement of soil texture and organic matter along the soil profile. One limitation of this system is that it employs a spectrometer with lower spectral resolution, which makes it incompatible with soil spectra

and libraries generated elsewhere. Poggio et al. (2015) evaluated the optical performance of a newly developed VisNIR foreoptic in the lab. They scanned 389 milled and pressed surface and subsurface samples with the new VisNIR foreoptic and an ASD's contact probe (Analytical Spectral Devices formerly, now Malvern Panalytical Inc., Boulder, CO). Their main objective was to quantify the potential performance degradation of the VisNIR foreoptic compared to an ASD contact probe.

Different from other direct sensing mechanisms, VisNIR relies on multivariate calibration models to convert spectral measurements to target soil properties. In almost all previous studies, a significant portion of the field samples (usually more than 50%) were lab analyzed for model calibration. While this is justifiable for research, in real applications analysis of field samples should be kept at minimum. A more practical and economical approach is to use the calibration models from the existing VisNIR soil spectral libraries. This brings two issues that should be carefully evaluated. Because *in situ* VisNIR sensors will have different measurement geometry and operation condition compared to a lab-based system (for example using a bench-top spectrometer and scanning accessories), the first question is how the two sets of spectra (even for identical samples) are different from each other. The second question is whether it is possible to correct for this apparent difference between them so that VisNIR models calibrated on the lab-based spectral libraries can be applied to *in situ* scans to predict target soil properties.

3.1.3 Objective

We developed a new optical probe that is designed to measure VisNIR reflectance from soil profiles. Our long-term goal was to develop a VisNIR equipped soil penetrometer system for *in situ* high resolution vertical soil sensing. The specific objective of this study was to evaluate the performance of this new VisNIR probe in terms of its spectral quality and predictive power using an external spectral library. Our particular effort was to investigate whether the prediction performance of the VisNIR probe can be improved via

two strategies: calibration transfer (direct standardization) and sample spiking (i.e. adding probe scans to model calibration dataset).

3.2 MATERIALS AND METHODS

3.2.1 VisNIR optical probe

Two VisNIR probes were developed to be tested in this study to find the design with minimum noise and internal scattering. First probe (D1) consisted of a bifurcated cable with an angled ferrule and the light source outside; while the second probe (D2) consisted of a broad-spectrum light source (a halogen lamp) and a 45° parabolic reflective mirror inside (Figure 3.1). The probe had an outside diameter of 31.8 mm (1.25 in.). The wall of the penetrometer has a 12.8 mm (1/2 in.) aperture sealed with a 1-mm thick fused silica window. A custom designed cone tip is connected to the bottom of the penetrometer; and the top end was machined with threads so that it can be connected to a hydraulic press. The VisNIR probe was designed with the goal of coupling with a commercial spectrometer (ASD Labspec, Analytical Spectral Devices formerly, now Malvern Panalytical Inc., Boulder, CO) by which majority of soil spectra are collected, but can also be coupled with spectrometers from other manufacturers.

3.2.2 Soil samples and datasets

The spectral library we used to build calibration models for assessing the performance of the VNIR probe is from the Rapid Carbon Assessment (RaCA) Project (Wills et al., 2014). RaCA is a national-wide effort to capture the baseline soil carbon stocks across the conterminous U.S. A subset of RaCA samples from RaCA Region 5 was used. This gave us a total of 1595 samples, all having the full characterization of Organic Carbon (OC), Total Carbon (TC), and Total Nitrogen (TN), as well as the VNIR reflectance spectrum. The VisNIR spectra were collected with an ASD LabSpec[®] spectrometer and its MugLite[®] accessory in the lab setting.

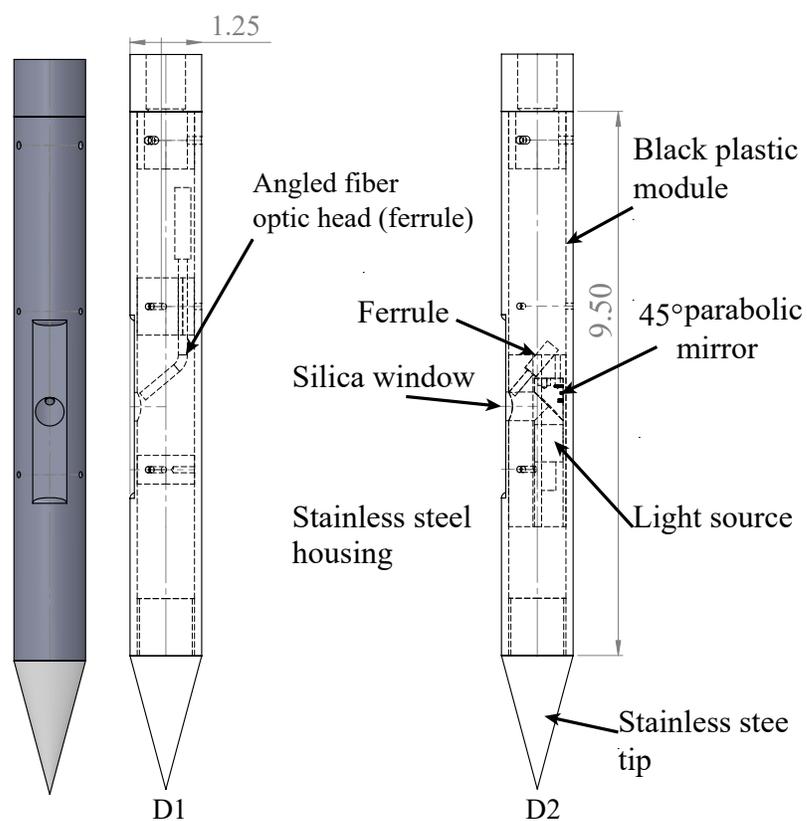


Figure 3.1. Designed VisNIR optical probes. The first design (D1) consisted of a bifurcated fiber optic cable with an angled ferrule, the second design (D2) had a straight ferrule, 45° mirror and a halogen lamp inside.

One hundred and fifty air-dried, ground and sieved (<2 mm) soil samples archived at the National Soil Survey Center of USDA-NRCS were extracted and used for the performance assessment of the VisNIR probe. Three criteria were used for sample selection from the archive. First, the geographic region of the samples is Nebraska, USA. This matches RaCA Region 5 of the soil library we are using. Second, the samples have the measurement of OC, TC, and TN (again, to be consistent with the library). Since OC is an emphasis property, the third criterion is to select the samples representing the full range of OC in the archive through a 20-stratum stratified random sampling.

These 150 samples were further divided into two sets using the Kennard-Stone algorithm (Kennard & Stone, 1969): Group A and B. Group A has 50 samples and was used for two different purposes: (1) to develop a spectral transfer by direct standardization (DS) to match the spectra of the VNIR probe with those of the ASD MugLite[®] (Analytical Spectral Devices formerly, now Malvern Panalytical Inc., Boulder, CO), and (2) to spike the library for calibration model development. Group B has 100 samples and was used for model validation and performance assessment. Table 3.1 gives the summary statistics of OC, TC, and TN for the library, Group A, and Group B, as well as the spectral measurements acquired for each set.

3.2.3 VNIR probe testing and modeling

The testing of VisNIR probes was conducted at the Kellogg Soil Survey Laboratory of USDA-NRCS and Figure 3.2 shows the experimental setup. Probes were mounted on a table with the viewing aperture pointing upwards using a vise clamp to hold their position and orientation tightly. Since D2 used a halogen lamp as the light source, a DC power supply was used to provide a constant 1.0 ampere electric current with approximate 3.8 V supply voltage. For D1, SLS201 - compact stabilized broadband light source (Thorlabs Inc., Newton, New Jersey, USA) was used as the light source. Reflected light energy from the probe was acquired by an ASD LabSpec[®] spectrometer (Analytical Spectral Devices

Table 3.1. Summary statistics of the soil properties for the library and the local set (Groups A and B) used in the lab testing of the VisNIR probe.

Dataset	No. of samples	Spectra measured by	Soil property	Min.	Median	Mean	Max.
Library ^a	1595	MugLite [®]	OC ^b (%)	0	0.93	1.48	39
			TC ^c (%)	0	1.21	1.84	39
			TN ^d (%)	0	0.12	0.17	3.09
Group A	50	MugLite [®] & VisNIR probe	OC ^b (%)	0	0.55	1.00	5.5
			TC ^c (%)	0	1.06	1.32	5.52
			TN ^d (%)	0	0.09	0.14	0.73
Group B	100	VisNIR probe	OC ^b (%)	0	0.94	1.17	9.01
			TC ^c (%)	0	1.03	1.30	10.7
			TN ^d (%)	0	0.12	0.14	0.77

^aThe library contains the samples from RaCA Region 5 of the Rapid Carbon Assessment Project; ^bRoot Mean Squared Error of Prediction; ^cRatio of Performance to Deviation; ^dPartial Least Squares Regression

formerly, now Malvern Panalytical Inc., Boulder, CO) through the fiber optical bundle. Pucks with soil samples were then placed on the top of the viewing aperture of the probe for scanning. This configuration is similar to standard soil spectra acquire by ASD MugLite[®] (Analytical Spectral Devices formerly, now Malvern Panalytical Inc., Boulder, CO) attachment where light goes through two layers of fused silica windows between the samples and the receiving fiber optics. However, this is slightly different from the intended field application since there will only be one layer of silica window (only the window of the probe).

The selected 150 soil samples (groups A and B in table 3.1) were scanned by both D1 and D2 probes while the samples from group B (table 3.1) were scanned with the standard MugLite[®] attachment as well. A standard Spectralon[®] panel with 99% reflectance (Labsphere Inc., North Sutton, NH, USA) was used as the white reference in 10-minute intervals to ensure proper calibration of the spectrometer. Additionally, a standard dark reference panel with 2% reflectance (Labsphere Inc., North Sutton, NH, USA) was used to acquire the internal light back scattering of the probes. All acquired

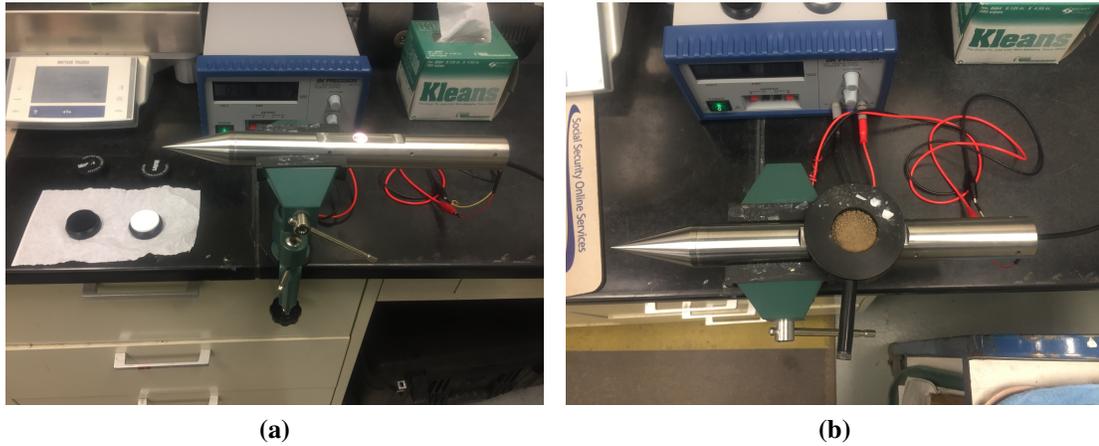


Figure 3.2. Laboratory experimental setup for the testing of the VisNIR probes. (a) shows the optical probe attached to the clamp and (b) shows the puck on the optical probe being scanned.

spectra from D1 and D2 were corrected as shown in equation 3.1.

$$R_{DRC} = \frac{E_s - E_{DR}}{E_{WR} - E_{DR}} \quad (3.1)$$

where R_{DRC} is the dark reference corrected (DRC) reflectance of the sample, E_s is the reflected energy by the sample, E_{DR} is the reflected energy by the dark reference panel and E_{WR} is the reflected energy by the white reference panel.

We used only the 400–2500 nm wavelength range to avoid noise observed in the 350–399 nm range and the spectra were preprocessed by Savitzky-Golay smoothing with 3rd order polynomial and 11 nm window prior to the data analysis. Three types of VisNIR models were calibrated. The first type of models was built from the library samples only ($n = 1595$). The second type of models were built from the library samples spiked by the VisNIR probe scans of the samples in Group A ($n = 1595 + 50$). The third type of models was similar to the second type except that extra weight was given to the spiking sample in Group A. In our case, the samples in Group A were replicated 31 times and added to the library for model calibration ($n = 1595 + 50 \times 31$). Guerrero et al. (2014) showed that putting extra weight on the local spiking samples (that is, to add copies of the spiking set

to achieve 50-50% composition of the final library) can improve the prediction of soil OC.

To test DS, we first developed a spectral transfer matrix \mathbf{F} using the two spectral sets of Group A by the VisNIR probe and the MugLite[®] using the equation 3.2.

$$\mathbf{F} = (\mathbf{R}_{GroupA,VNIRP})^{-1} \times (\mathbf{R}_{GroupA,MugLite^{\circledR}}) \quad (3.2)$$

where $\mathbf{R}_{GroupA,VNIRP}$ is the reflectance spectra of Group A acquired by the VisNIR probe; $\mathbf{R}_{GroupA,MugLite^{\circledR}}$ is the reflectance spectra of Group A by the MugLite[®]. $(\cdot)^{-1}$ denotes matrix inversion.

In principle, this transfer matrix \mathbf{F} characterizes and can account for the systematic differences of the spectra acquired by the two systems. \mathbf{F} was then used to transfer the VisNIR probe scans of Group B with equation 3.3.

$$\mathbf{R}_{GroupB,DS} = \mathbf{R}_{GroupB,VNIRP} \times \mathbf{F} \quad (3.3)$$

where $\mathbf{R}_{GroupB,VNIRP}$ is the reflectance spectra of Group B acquired with the VisNIR probe.

The first type model developed with the library samples only was applied to the transferred spectra ($\mathbf{R}_{GroupB,DS}$) for prediction. DS is a standard NIR calibration transfer method (Feudale et al., 2002) and has been successfully tested on soils (Ge et al., 2011). All models were evaluated with the test set group B (n = 100), which was scanned by the VNIR probe only.

Partial least squares regression (PLSR) is used for model calibration. The number of latent factors (n_{LV}) was allowed to vary from 1 to 30, and the size of a model was selected for the n_{LV} that gave the minimum RMSE_{CV} (Root Mean Squared Error of Cross Validation) with 25 random segment cross-validation. Statistics including R^2 (Coefficient of Determination), Bias, RMSE_P (Root Mean Squared Error of Prediction), SEP (Standard Error of Prediction), and RPD (Ratio of Performance to Deviation) were calculated to

assess different types of models for different soil properties.

All data analyses were performed in the R environment (R Core Team, 2018) with the pls package for PLSR modeling (Mevik et al., 2013), caret as the modeling wrapper (Max et al., 2015), doParallel for parallel processing (Revolution Analytics & Weston, 2015), ggplot2 package for graph generation (Wickham, 2009) and gnm for calculating Moore-Penrose generalized inverse (Turner & Firth, 2015)).

3.3 RESULTS AND DISCUSSION

3.3.1 Comparison of D1 and D2

Figure 3.3 shows the spectrum of a sample obtained from D1, D2 and MugLite[®].

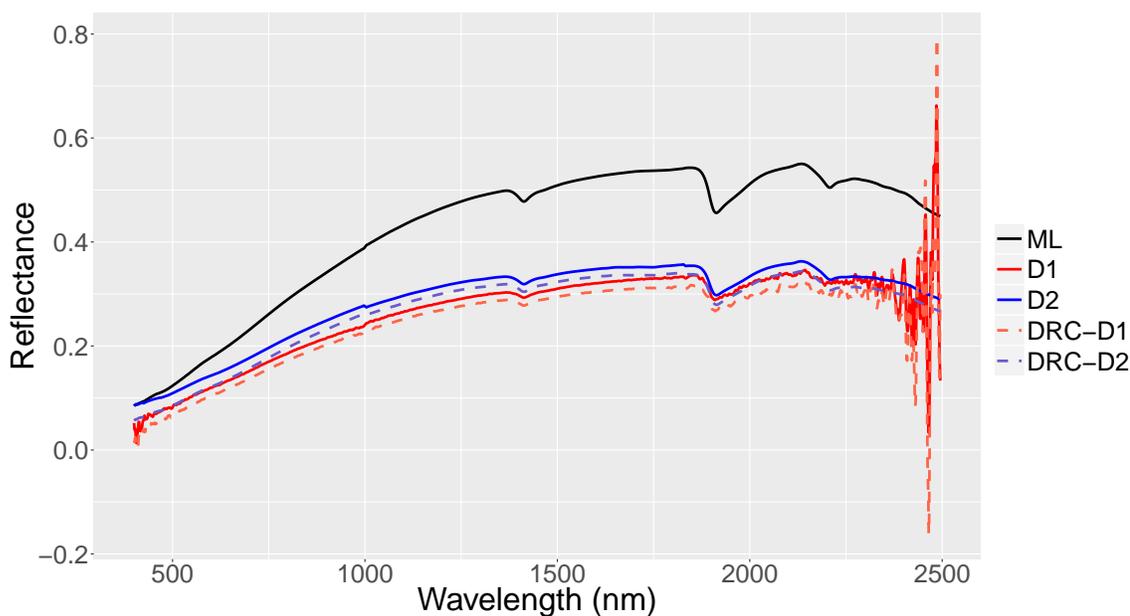


Figure 3.3. Spectra obtained from VisNIR probe designs D1, D2, dark reference corrected D1 (DRC-D1), dark reference corrected D2 and MugLite[®] (ML).

It can be seen in figure 3.3 that both D1 and D2 had systematic lower reflectance throughout the entire wavelength range. This could be mainly attributed to the difference in optical configurations between the MugLite[®] and the VisNIR probes. If the soil surface acts as perfect Lambertian surface, both MugLite[®] and VisNIR probes should receive the

same amount of energy at any viewing angle. However, since these soil surfaces are not Lambertian and the viewing angles of the MugLite[®] and VisNIR probes are different, VisNIR probes could intercept lower reflected energy compared to MugLite[®]. This type of systematic deviations of spectra can be corrected using spectral treatment methods such as direct standardization and piecewise direct standardization (Fearn, 2001; Feudale et al., 2002; Ji et al., 2015; Wang et al., 1991).

Dark reference correction did not affect the spectra significantly, suggesting lower internal scattering (<2%) in the VisNIR probes. According to Figure 3.3, it was evident that D1 had higher noises at the beginning (400–499 nm) and end (2301–2500 nm) of the wavelength range. This could be due to the optical characteristics of the custom-made bifurcation cable with the angled ferrule used in D1 restricting sufficient energy throughput at lower and higher wavelength regions. Unlike D1, D2 showed smooth spectra analogous to MugLite[®] spectra throughout the whole wavelength range. This suggests satisfactory optical configuration of D2 in comparison with MugLite[®]. Due to the obvious high-quality spectra, only the spectra obtained from D2 were used for the proceeding analysis. The D2 is hereinafter called as ‘VisNIR probe’ or ‘VNIRP’.

3.3.2 Qualitative Assessment of soil spectra acquired by the VNIR probe

We compared group B’s MugLite[®] scans, VisNIR probe scans, and VisNIR probe scans transformed with the DS matrix **F** and the result is shown in Figure 3.4. Soil spectra acquired by the VisNIR probe was very smooth compared to the MugLite[®] scans, indicating that the VisNIR probe has a satisfactory optical configuration. The signal to noise ratio of spectral readings at all wavelengths was quite high.

The VisNIR probe scans showed an overall reduction in reflectance across all wavelengths. This was likely due to the difference in optical design and configurations between the MugLite[®] and the VisNIR probe. Similar to the reduced amplitude, the spectral variation of the VisNIR probe scans (spectral bounding box in Figure 3.4) was

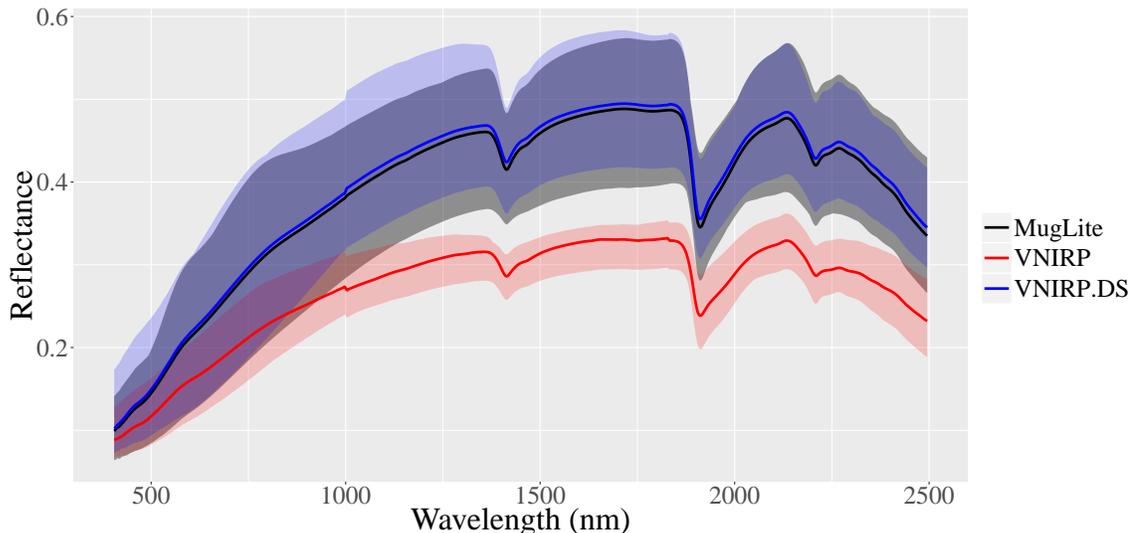


Figure 3.4. Qualitative assessment of the soil spectra of 15 samples from Group B scanned by the MugLite[®] and the VisNIR probe (VNIRP), and the VNIR probe scans transformed by direct standardization (VNIRP.DS). The bounding boxes are the maximum and minimum reflectance at each wavelength for the three sets.

also smaller than that of the MugLite[®] scans. This systematic shift in spectral measurement can lead to poor prediction if the spectral library acquired by the MugLite[®] is directly used for prediction. In this work we tried to use two approaches to account for this systematic error. DS was the first approach, where a spectral transfer matrix \mathbf{F} was derived to account for this shift explicitly before applying the model. As can be seen in Figure 3.4, the DS-transferred VisNIR probe spectra matched with the MugLite[®] spectra very well (with the mean spectrum almost identical and the spectral variance windows overlap significantly). The convex hull in the first two principal component space gave even better visualization of how DS account for this spectral difference (Figure 3.5). The second approach was sample spiking, where a portion of the samples scanned by the VisNIR probe was added to the library for model calibration. In this approach, we relied on multivariate modeling (PLSR in this study) to account for this spectral shift implicitly.

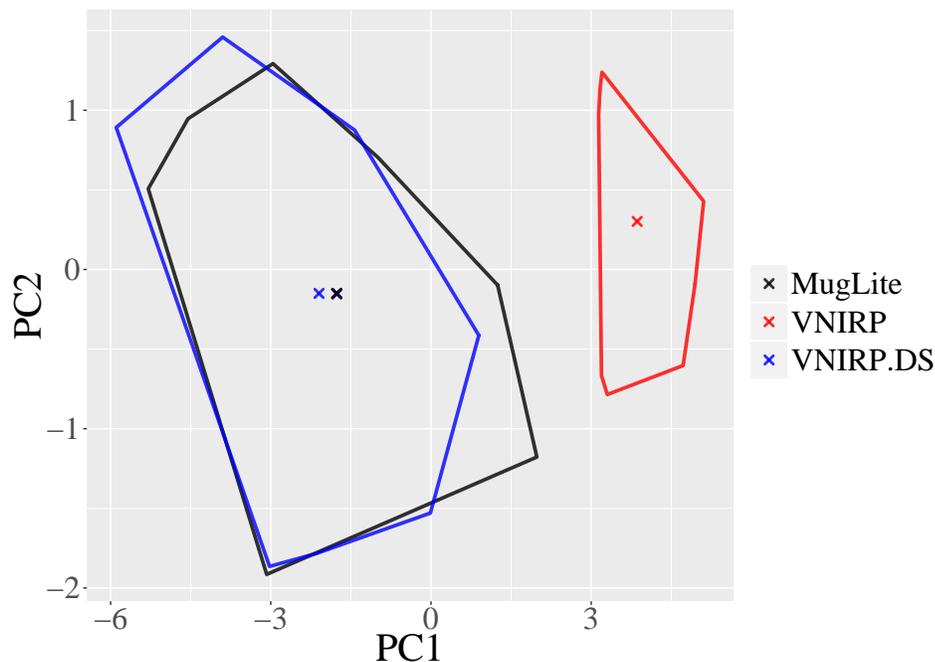


Figure 3.5. The convex hull of the three spectral sets in the first two principal component space. The black convex hull encloses the scans by the ASD MugLite®; the red convex hull encloses the scans by the VisNIR probe (VNIRP); the blue convex hull encloses the VisNIR probe scans after Direct Standardization transformation (VNIRP.DS).

3.3.3 Prediction performance of the spectra acquired by the VisNIR probe

Table 3.2 shows the prediction performances of the three VisNIR models (library only, library & group A, and library & Group A with extra weight) tested on the scan sets in Group B.

Per table 3.2, cross validation R^2 of all three models for all three soil properties were quite similar, indicating similar model performance. For OC and TC, the models calibrated with library & Group A with extra weight showed significantly lower $RMSEP$. This was because the ranges of these soil properties are much lower in group A than the library. When extra weight was applied to combine with the library, that lowers the range of the combined calibration set and therefore the lower model $RMSEP$ (given the model performance remains the same).

When the model calibrated with the library was directly applied to the VisNIR

Table 3.2. Prediction performance of different models with different validation sets for Organic Carbon (OC), Total Carbon (TC), and Total Nitrogen (TN).

Soil property	Model calibrated from	Cross-Validation			Validation				
		R ²	RMSE _{Cv} ^d	Test spectral set in Group B	R ²	RMSE _P ^e	Bias	SEP ^f	RPD ^g
OC ^a	Library	0.74	1.17	VNIRP ^h	0.81	1.1	-0.93	0.58	1.16
	Library & Group A	0.74	1.18	VNIRP.DS ⁱ	0.84	0.66	0.24	0.61	1.95
	Library & Group A with extra weight	0.74	0.98	VNIRP ^h	0.82	0.73	-0.46	0.56	1.76
TC ^b	Library	0.76	1.2	VNIRP ^h	0.79	1.37	-1.22	0.63	1
	Library & Group A	0.75	1.22	VNIRP.DS ⁱ	0.8	0.83	0.41	0.72	1.67
	Library & Group A with extra weight	0.75	1.01	VNIRP ^h	0.83	0.58	-0.01	0.58	2.37
TN ^c	Library	0.74	0.11	VNIRP ^h	0.74	0.08	-0.04	0.07	1.57
	Library & Group A	0.7	0.11	VNIRP.DS ⁱ	0.79	0.07	0.04	0.06	1.65
	Library & Group A with extra weight	0.72	0.1	VNIRP ^h	0.76	0.06	-0.02	0.06	1.91

^aOrganic carbon ^bTotal carbon ^cTotal nitrogen ^dRoot Mean Squared Error of cross-validation ^eRoot Mean Squared Error of validation ^fStandard Error of Prediction ^gRatio of prediction deviation ^hscans from VisNIR probe ⁱdirect standardization transformed VisNIR spectra

probe scans, the predictions were not quite satisfactory. The prediction R^2 was high (ranging from 0.74 to 0.81), meaning the linear association between the predicted and measured values was strong. The main problem was the high prediction biases, which was clearly linked to the systematic differences of the spectra acquired by the VisNIR probe versus the MugLite[®] (Figures 3.4 and 3.5).

Direct standardization transform improved the prediction accuracy consistently for all three soil properties. The improvement was substantial for OC (RDP from 1.16 to 1.95) and TC (RPD from 1.00 to 1.67) and marginal for TN (RPD from 1.57 to 1.65). In comparing the model statistics, it was clear that DS removes the systematic prediction biases, which were caused by the systematic differences of the spectra between the VisNIR probe and the MugLite[®] (Figure 3.4). DS can partially remove this difference and therefore reduce the prediction bias.

The results in Table 3.2 also suggested that spiking samples scanned with the VisNIR probe was an effective approach to improve prediction accuracy. Even without extra weight (50 Group A samples combined with 1595 library samples), the model performance was greatly improved for all three soil properties (RPD from 1.16 to 1.76 for OC, from 1.10 to 1.89 for TC, and from 1.57 to 1.91). Similar to DS, the improvement primarily came from the decrease in prediction bias whereas the R^2 values (indicating the linear association between measured and predicted values) remain quite constant. Giving extra weight to the spiking samples can further improve the prediction accuracy (RPD of 2.38, 2.37, and 2.15 for OC, TC, and TN, respectively). This is to be expected since extra weighing incorporates the effect of foreoptic differences into the calibration dataset more effectively allowing the modeling technique to rectify for it. Augmenting the local samples with extra weight makes the local sample to carry almost the same weight as the library samples (given the fact that the number of samples in a library is much high than the spiking set). In our case, we replicated the spiking set 31 times to achieve a nearly 1:1 ratio. The result showed this almost completely removes the prediction bias (-0.06, -0.01

and 0.01% for OC, TC, and TN, respective) and thus improved the prediction substantially.

Figure 3.6 shows the prediction plots of OC for Group B samples with the different models. This figure provides readers with the information on how the measured vs. predicted OC points scattered around the 1:1 line. Although not shown here, TC and TN showed similar patterns.

When comparing the different methods (DS, spiking, and spiking with extra weight), the results indicated that spiking with extra weight is most effective in accounting for the differences between the sets and improve the prediction. Two sources of differences existed between the library and Group B. First, there was difference in the soils of these two sets because they were created and compiled under different situations (i.e. under different projects in different times). Second, there is difference in the spectra as the library was scanned by the MugLite® and Group B was scanned with the VisNIR probe. Note that all previous studies on sample spiking deal with the first source of difference.

It is certain that DS would only account for the second source of difference because the DS matrix was developed from Group A (part of the local set). On the other hand, spiking would account for both sources of difference by combining Group A with the library in PLSR modeling. As pointed out by several authors (Gogé, Gomez, Jolivet, & Joffre, 2014; Guerrero et al., 2014; Wetterlind & Stenberg, 2010), spiking ensures the calibration set to contain samples with similar characteristics to the local set and enables the models to capture the spectral variations present in the local set. With a small spiking set and a large spectral library (in our case, 50 spiking samples versus 1595 library samples), the effect of spiking is diluted. Therefore, putting extra weight would be effective because the calibration set achieve a balance between the numbers of local and library samples (Guerrero et al., 2016). Our results clearly indicate that spiking with extra weight gives the best prediction for the VisNIR probe scans of Group B.

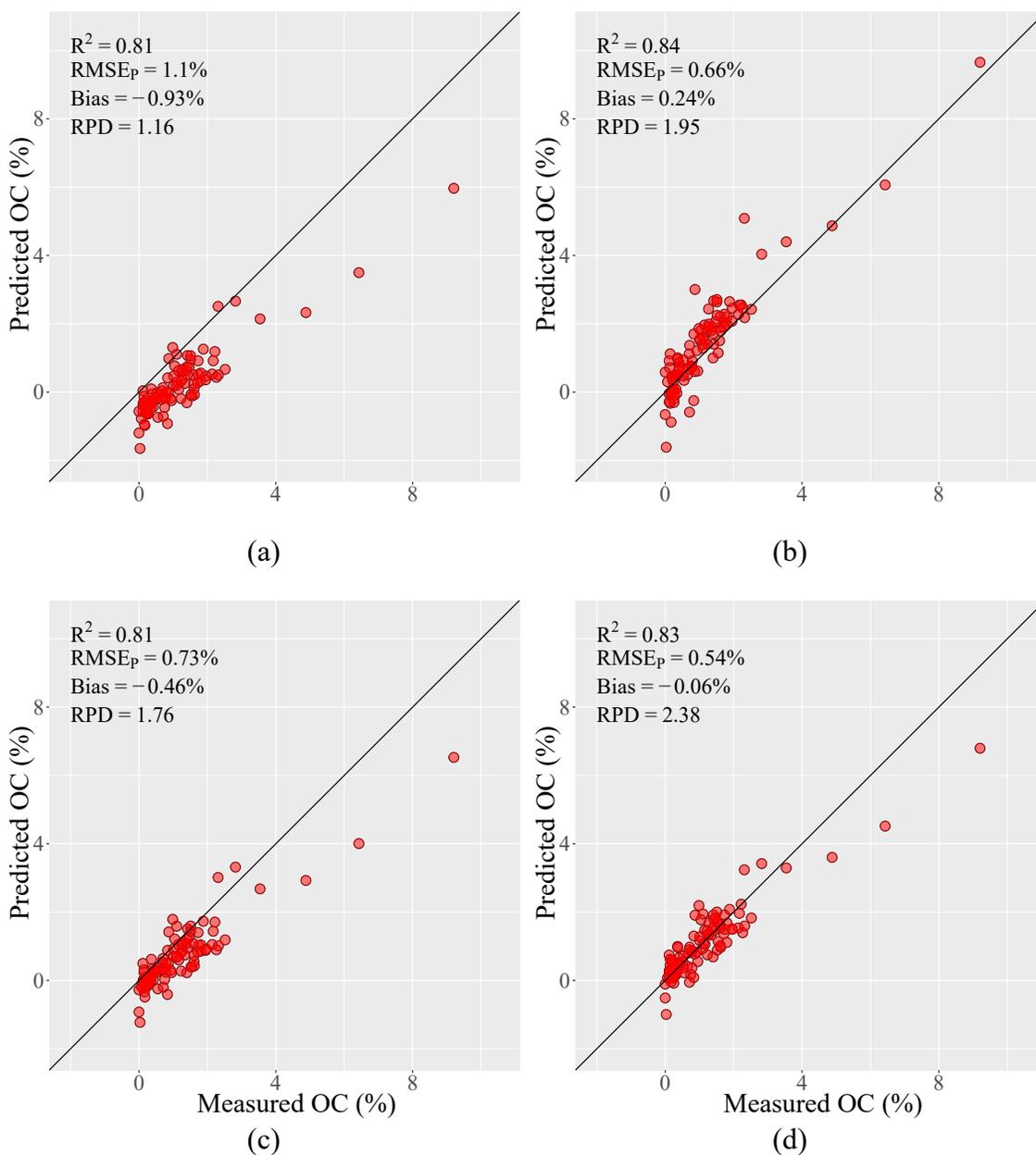


Figure 3.6. Prediction plot of Organic Carbon (OC) in Group B: (a) Library only model on VisNIR probe scans; (b) Library only model on DS transformed scans; (c) spiked model on VisNIR probe scans; and (d) spiked and extra weighted model on VisNIR probe scans.

3.3.4 Development and field implementation of the VisNIR based penetrometer: Challenges

It should be noted that this study only examined the spectral difference and subsequently the predictive performance due to the inherent optical difference between the VisNIR probe and the ASD MugLite[®]. In actual field implementation, there are other significant sources of variation that the VisNIR penetrometer system should be able to account for. These variations can be classified into two categories. The first category comes from *in situ* soils themselves including moisture, surface roughness, natural aggregation, and temperature. The second category comes from the field operational conditions of fluctuating temperature and humidity that cause varying responses of the optical/electronic components and the spectrometers.

Although there have been a number of reports on VisNIR-based *in situ* soil sensors for lateral surface layer soil sensing (Christy, 2008; Hummel et al., 2001; Kodaira & Shibusawa, 2013; Maleki, Mouazen, De Ketelaere, Ramon, & De Baerdemaeker, 2008; Mouazen et al., 2007), they all used so called internal validation scheme where the field samples were used for model calibration. This makes these sources of spectral variations not critical when it comes to model calibration and validation. However, if external dry ground soil spectral libraries are to be utilized for model calibration, the development of algorithms to account for these spectral variations is of paramount importance.

Several studies have already examined the correction of soil moisture effect on VisNIR modeling and demonstrated that spectral transformation algorithms like external parameter orthogonalization and DS are quite useful (Ackerson et al., 2015; Ge et al., 2014; Ji et al., 2015; Minasny et al., 2011). In this study, we also demonstrated that DS is useful to account for the difference in optical configuration between the VisNIR probe and the ASD MugLite[®]. It remains a research question whether multiple transformations should be developed and employed, or one “catch-all” transformation could be sufficiently constructed, to account for these different sources of spectral variations. In the case of

sample spiking, no spectral transformation is needed. But it does require lab analysis of a small fraction of field samples (to obtain Y measurement) to be merged with the external library for chemometric modeling.

Taken together, the development of a VisNIR penetrometer entails good design of both hardware and software. The hardware ensures the acquisition of high-quality reflectance spectra and the software ensures the compensation of spectral shifts caused by various factors. Being able to leverage the external soil spectral libraries is critical to make this technology economically viable in the practical setting. Addressing some of the challenges discussed in this section are presented in the subsequent chapter.

3.4 CONCLUSIONS

In this study, we conducted a lab-based performance evaluation of a new VisNIR optical probe for high resolution vertical soil mapping. We used an external soil spectral library for VisNIR model calibration, and investigated two strategies, namely direct standardization and sample spiking, to improve the prediction of soil Organic Carbon, Total Carbon, and Total Nitrogen. We found that, although the spectra acquired by the VisNIR probe are of high quality (low in noise), there exist systematic differences (in both amplitude and variation) when they are compared to the spectra by the MugLite[®]. We also found that both direct standardization and sample spiking are viable approaches to account for the systematic differences and substantially improve the prediction accuracy. The improvement is mainly from a reduction in prediction Bias. Sample spiking with extra weight showed the highest prediction accuracy, with Bias lower than 0.06% and RPD greater 2.15 for all three soil properties. We conclude that the current design of the VisNIR optical probe, together with spectral libraries and the spectral compensation algorithms, can lead to a rapid, accurate and cost-effective penetrometer system for *in situ* high resolution vertical soil sensing.

CHAPTER 4

DEVELOPMENT, AND FIELD TESTING A VISNIR INTEGRATED MULTI-SENSING SOIL PENETROMETER

4.1 INTRODUCTION

4.1.1 Soil sensing

The demands for spatially explicit, high resolution soil data in three dimensions (i.e., laterally and vertically) are increasing for a number of crosscutting disciplines such as precision agriculture, soil process and crop modeling, and digital soil mapping (Minasny & McBratney, 2016; Vereecken et al., 2016; Viscarra Rossel & Bouma, 2016). Several modern technologies are available to effectively characterize the lateral variation of soils. These techniques include aerial and satellite remote sensing (Barnes et al., 2003; Ben-Dor, 2002), on-the-go soil sensors (Adamchuk et al., 2004), apparent electrical conductivity/resistivity (Corwin & Lesch, 2003), ground penetrating radar (Lunt, Hubbard, & Rubin, 2005), and the γ -ray sensor (Triantafilis et al., 2013). These technologies are commonly used to produce high resolution soil maps to capture lateral variability. On the other hand, technologies to characterize soil profiles without pulling a soil core or opening a soil pit are limited (Hartemink & Minasny, 2014). While technologies like soil apparent electrical conductivity do sense soil to a certain depth, they return a signal that is a weighted integration of the profile and are not capable of providing layered information directly.

4.1.2 VisNIR for *in situ* soil sensing

Visible and near infrared reflectance spectroscopy (VisNIR) has several advantages that make it viable for *in situ* soil sensing and, in particular high resolution vertical soil

sensing: (1) the non-contact nature of interactions between VisNIR electromagnetic energy and soils, (2) ease of miniaturization for a VisNIR sensing probe using optical fibers, and (3) the ability to infer multiple soil properties from a single VisNIR scan.

There have been a number of reports on the development and application of in situ VisNIR sensors for surface (lateral) soil mapping (Bricklemyer & Brown, 2010; Christy, 2008; Kodaira & Shibusawa, 2013; Mouazen et al., 2007).

Several authors conducted studies where soil cores were collected and scanned in the lab at fine vertical resolution to obtain “simulated” *in situ* soil profile sensing. Hummel et al. (2001) collected 48 cores from Illinois and soil samples (at 2.5 cm increment from the cores) were scanned at six moisture levels to predict soil organic matter and moisture. Waiser et al. (2007) and Morgan et al. (2009) conducted simulated profile characterization of clay and soil organic and inorganic carbon, respectively, with the 72 soil cores collected from Texas. Doetterl et al. (2013) collected 151 soil cores from central Belgium and demonstrated the simulated high vertical resolution (3 cm) soil organic carbon assessment with VisNIR. While these were not true *in situ* sensing, they established the technical feasibility and demonstrated the potential of a high-resolution vertical soil sensing system based on VisNIR.

Demonstrations of *in situ* vertical soil sensing are sporadically reported in the literature. Ben-Dor et al. (2008) described a custom-made spectral head device that was inserted into the subsoil to measure VisNIR reflectance along the profile after a hole was drilled. Wetterlind et al. (2015) tested a commercial Veris P4000 probe (Veris® Technologies Inc., Salina, Kansas, USA) for the measurement of soil texture and organic matter along the soil profile. Cho, Sheridan, Sudduth, and Veum (2017) tested the same Veris probe in Missouri USA to measure total carbon, bulk density, and clay content. Poggio et al. (2015) evaluated the performance of a new VisNIR foreoptic in the lab. They scanned 389 milled and pressed surface and subsurface samples with the VisNIR foreoptic and an ASD contact probe (Malvern Panalytical Company, Longmont, CO, USA). Their

main objective was to quantify the potential performance degradation of the VisNIR foreoptic compared to an ASD contact probe. More recently, Ackerson, Morgan, and Ge (2017) tested a custom-developed VisNIR penetrometer for vertical soil sensing in Texas USA. A Texas Soil Spectral Library was used to develop VisNIR calibration models together with External Parameter Orthogonalization (EPO) for clay content prediction. EPO (Roger et al., 2003) was a method designed to minimize the effect of moisture on soil VisNIR spectra and allow the VisNIR models calibrated on dry-ground soil samples to predict soil properties of field moist samples (Ge et al., 2014; Minasny et al., 2011).

4.1.3 Objectives

The objectives of this work were to develop a fully integrated, multi-sensing penetrometer system for high resolution vertical soil sensing and field test the penetrometer to evaluate its performance. The developed penetrometer incorporated an improved VisNIR sensing probe (coupled to an ASD spectrometer), a load cell, an ultrasonic depth sensor, and a GPS receiver. A software program was also developed in LabVIEW (National Instruments, Austin, TX, USA) and run on a controlling computer for automated data collection and storage. This fully automated system was tested in 11 agricultural fields in Nebraska (NE), Illinois (IL), Iowa (IA), and South Dakota (SD).

4.2 MATERIALS AND METHODS

4.2.1 VisNIR integrated multi-sensing soil penetrometer system

The key components of the penetrometer system included a VisNIR probe based on the design discussed in chapter 3, an ultrasonic distance sensor (ToughSonic14, Senix Corp., Hinesburg, VT, USA), and a miniature load cell (LC202, Omega Engineering, Norwalk, CT, USA) (Figure 4.1). The VisNIR probe comprised of a broad-spectrum halogen light source (MR4-188, International Light Technologies, Peabody, MA, USA), a 45°parabolic reflective mirror (MPD019-G01, Thorlabs Inc, Newton, NJ, USA), and a fiber optical

cable (ASD, a Malvern Panalytical Company, Longmont, CO, USA) to collect light reflected from soil surface. The other end of the fiber optical cable was connected to an ASD LabSpec[®] spectroradiometer (ASD, a Malvern Panalytical Company, Longmont, CO, USA). The VisNIR measurement was made through an aperture (12.7 mm in diameter) on the wall of the penetrometer sealed with a 1-mm thick UV-fused silica broadband window (WG40530, Thorlabs Inc, Newton, NJ, USA).

The load cell measured insertion force F as the penetrometer was inserted into the soil. The ultrasonic distance sensor measured insertion depth D . A GPS receiver (BU-353S4, GlobalSat WorldCom Corporation, Taiwan) was also included to record the geographic coordinates of each sampling location. The penetrometer housing (which housed the VisNIR probe and the load cell) was 31.75 mm in diameter and connected to a 28.58 mm diameter extension shaft. The housing of the penetrometer and the extension shaft were made from stainless steel. A custom designed cone tip (tip half angle = 15°) was connected to the bottom of the penetrometer. During field sampling, the whole penetrometer assembly was attached to a Giddings probe truck (Giddings Machine Company, Windsor, CO, USA). At each sampling location, the penetrometer was hydraulically driven into the soil, while *in situ* soil VisNIR reflectance spectra, insertion force F , insertion depth D , and GPS coordinates were continuously and automatically measured and logged. Figure 4.1 shows the cross-sectional view and three-dimensional solid rendering of the multi-sensing penetrometer system.

A program was written in LabVIEW (National Instrument) to control the sensors on the penetrometer, retrieve, and store measurements on the controlling computer's hard drive. The program's graphic user interface displayed real-time measurements including *in situ* soil VisNIR spectra, insertion force, insertion depth, and GPS coordinates. An instrument box was constructed to house the ASD spectroradiometer, the GPS module, the datalogger (LabJack U6, LabJack Corporation, CO, USA), and additional circuits to monitor and control the voltage supply for the light source (Figure 4.2). The instrument

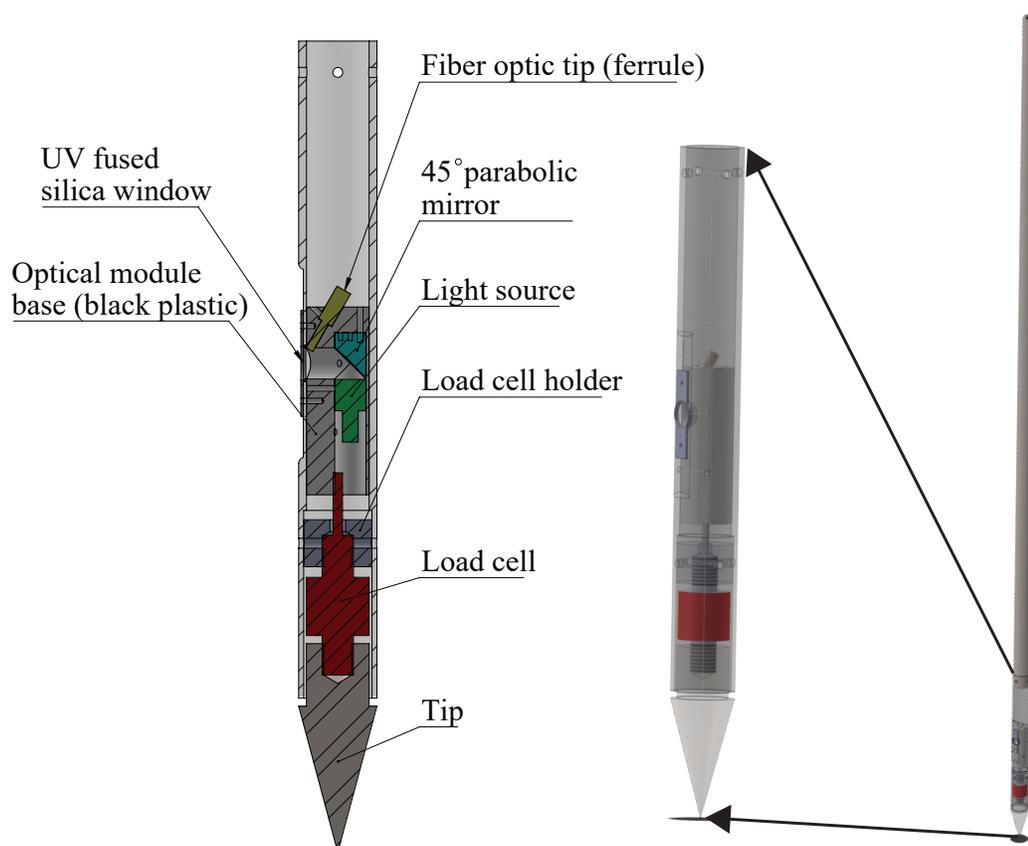


Figure 4.1. Cross-sectional view of the multi-sensing penetrometer to show the key system components (left) and the solid rendering of the penetrometer with the extension shaft (right).

box was also equipped with a temperature sensor to monitor the temperature inside the box and fans to regulate the temperature when needed.

4.2.2 Field experiment to test the integrated VisNIR multi-sensing penetrometer

4.2.2.1 Field testing

The integrated VisNIR multi-sensing penetrometer system was tested in 11 agricultural fields of four states (NE, IL, IA, and SD) in the U.S.A. The location, predominant soil type, and number of soil cores collected in each field are summarized in Table 4.1.

At each sampling location, the penetrometer was first mounted to a Giddings probe truck. Before each insertion, a white reference panel (99% reflectance, Labsphere,

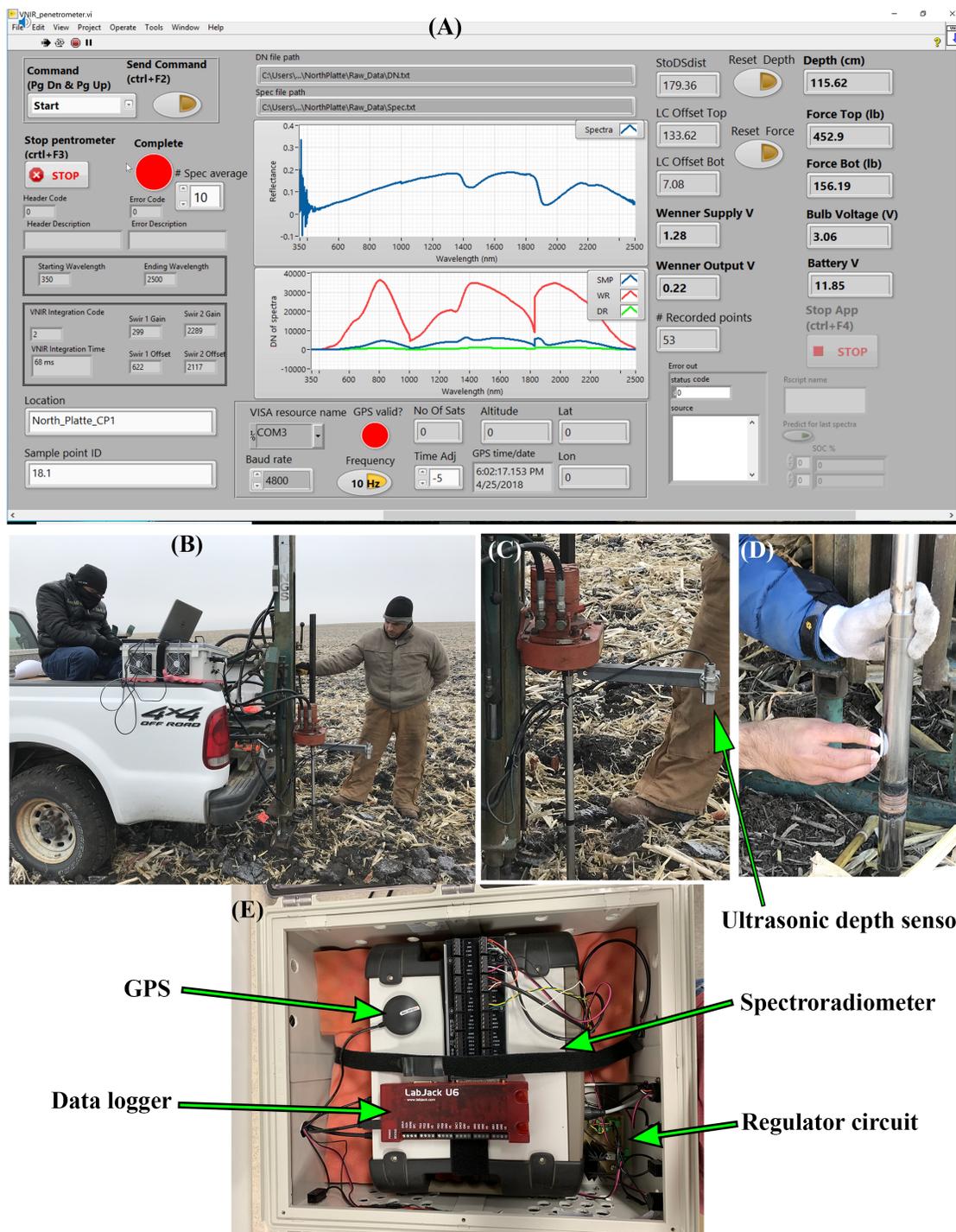


Figure 4.2. The graphical user interface of the LabVIEW program showing data collection from the VisNIR multi-sensing penetrometer system. **B:** The penetrometer was mounted on a Giddings probe truck for the field test. **C:** Close-up photo of the ultrasonic depth sensor for measuring insertion depth. **D:** A white reference panel was pressed on the penetrometer's aperture before each insertion. **E:** The instrument box to hold the spectroradiometer, data logger and other system components is shown.

Table 4.1. A summary of the test fields, soils, soil cores, and soil samples analyzed from each field.

State	Field	GPS coordinates	Dominant soil series	Number of soil cores	Number of soil samples		
					TC ^b and TN ^c	BD ^d	MC ^e
NE	HL	40° 51' 31" N, 96° 36' 50" W	Crete silt loam	3	42		
	RL	40° 50' 39" N, 96° 28' 12" W	Aksarben silty clay loam	3	42		
IL	CIL-1	39° 18' 14" N 87° 51' 20" W	Hoyleton silt loam	6	26		
	CIL-2	39° 13' 46" N, 87° 48' 24" W	Cisne-Huey silt loams	3	12		NA ^a
	CIL-3	39° 12' 46" N, 87° 48' 21" W	Newberry silt loam	4	16		
	SIL-1	39° 42' 48" N, 89° 55' 49" W	Sable silty clay loam	3	13		
	SIL-2	39° 45' 50" N, 89° 46' 48" W	Ipava silt loam	6	28		
	SIL-3	39° 58' 59" N 89° 59' 06" W	Ipava silt loam	3	14		
IA	IA-1	42° 27' 31" N, 94° 13' 42" W	Clarion loam	6	30		60
SD	SD-1	43° 55' 24" N, 96° 59' 47" W	Egan silty clay loam	4	20		42
	SD-2	43° 57' 28" N, 96° 52' 34" W	Egan silty clay loam	4	18		36

^aNot available; ^bTotal carbon; ^cTotal nitrogen; ^dBulk density; ^eMoisture content

North Sutton, NH, US) was used to cover the penetrometer's aperture to standardize the ASD spectroradiometer. The penetrometer was inserted into the soil to a nominal depth of 1 m. The insertion was maintained at a constant speed of approximately 0.5 cm s^{-1} . This slow and constant insertion speed was beneficial to improve the quality of *in situ* VisNIR and insertion force measurements. Two insertions were made at each sampling location, with the insertion points roughly 20 to 30 cm apart. The two sets of *in situ* soil VisNIR spectra were averaged depth-wise for later analysis.

Two 76.2 mm diameter soil cores (for gravimetric moisture content and bulk density determination) and one 50.8 mm diameter core (for soil property determination) were extracted at each location to validate the VisNIR penetrometer measurement. Note that gravimetric moisture content and bulk density cores were only collected from the fields in IA and SD (Table 4.1). To measure bulk density and soil moisture, cores were cut in the field at 10 cm intervals (e.g., 0 to 10 cm, 10 to 20 cm, etc.) to ensure known volumes. The samples were weighed after collection, while moist, then oven-dried at 105°C , and weighted again. Values of moisture content and bulk density from the two cores were averaged. The cores collected to measure soil properties were described for basic soil morphology, such as horizon nomenclature, in the lab and cut according to horizon, except for the fields in NE. For the NE fields, the soil property cores were segmented by depth to match with the *in situ* VisNIR measurements. For instance, if 5 consecutive spectra were acquired along 10 cm in the beginning and 12 cm afterwards, the core was segmented at 10 and 22 cm depths. All samples from the soil property cores were air dried at 40°C for 3 to 7 days, ground and sieved to pass a 2 mm screen. The air-dry ground and sieved samples were scanned in the lab with the ASD spectroradiometer to obtain dry-ground VisNIR spectra. The soil samples were analyzed in the lab for total carbon (TC) and total nitrogen (TN) by dry combustion and clay content by the pipette method (Soil Survey Staff, 2014).

A soil VisNIR spectral library from USDA-NRCS-Kellogg Soil Survey Lab

(Lincoln, NE) was used as an independent, external library for spectral modeling. The whole library had ~20,000 soil samples from the U.S.; and 3603 samples, which originated from NE (1815), IL (725), IA (589) and SD (474) were extracted. Table 4.2 gives the summary statistics of the soil properties of the extracted library and the soil samples collected from the field campaign. More details on the USDA-NRCS spectral library can be found in Wijewardane, Ge, Wills, and Loecke (2016) and Wills et al. (2014).

Table 4.2. Summary statistics of the soil properties for the VisNIR dry-ground library from USDA-NRCS and the field samples.

Dataset	Property	Summary statistics					
		Min.	Max.	Mean	Median	Skewness	Kurtosis
Library	TC ^a (%)	0.01	9.57	1.59	1.26	2.02	8.72
	TN ^b (%)	0.00	0.89	0.15	0.13	1.86	7.72
Field samples	TC ^a (%)	0.12	5.52	1.35	1.29	1.53	7.15
	TN ^b (%)	0.01	0.45	0.11	0.09	1.25	5.57
	Clay (%)	9.80	54.00	33.34	33.70	0.13	2.38
	MC ^c (%)	8.20	53.81	24.14	24.54	0.51	4.68
	BD ^d (g cm ⁻³)	0.84	2.09	1.39	1.38	0.13	4.04

^aTotal carbon; ^bTotal nitrogen; ^cBulk density; ^dMoisture content

4.2.2.2 Modeling of TC and TN

The goal of this analysis was to predict soil TC and TN from the *in situ* soil spectra collected by the VisNIR penetrometer. The field samples had two sets of VisNIR spectral measurements: (1) *in situ* spectra measured by the VisNIR penetrometer and (2) dry-ground spectra measured in the lab. A total of 261 cores samples were analyzed for TC and TN. They were randomly divided into an EPO calibration transfer set (30%) and a validation set (70%). Firstly, the dry-ground spectral library was used to calibrate a TC and TN model using support vector regression (SVR). The dry-ground models were applied to the dry-ground spectra of the field samples (the validation set) to assess the model performance.

Next, an EPO transformation matrix was calculated using the matched *in situ* and dry-ground spectra of the field samples (the EPO calibration transfer set). The optimum number of EPO components (eigenvectors) was decided on the lowest cross validation RMSE (Root Mean Squared Error) of partial least squares regression (PLSR). The EPO transformation matrix was then applied to both the dry-ground spectral library and the *in situ* spectra of field samples (the validation set). EPO was employed to correct for the moisture effect on *in situ* soil VisNIR spectra and allow the models calibrated on the dry-ground library to predict with *in situ* spectra. Readers are referred to Roger et al. (2003) for a theoretical basis of EPO, and Minasny et al. (2011), Ge et al. (2014) and Wijewardane, Ge, and Morgan (2016a) for implementing EPO on soil spectral data.

After the EPO transformation, a second TC and TN model was calibrated using the EPO-transformed spectral library (using SVR) and validated with the EPO-transformed spectra of the field samples.

Support vector regression is a nonlinear modeling technique that constructs an optimal hyperplane in a higher dimensional feature space (Vapnik, 2013). A linear regression function is then computed in the higher dimensional feature space for the input data, which are mapped through a kernel function (Basak et al., 2007). This modeling technique was known to effectively model VisNIR spectral data with low errors (Thissen et al., 2004; Viscarra Rossel & Behrens, 2010). All models were tuned using ten random segment cross-validation and C (soft margin cost function) as a tuning parameter, which was the degree of penalty for the violations to the margin (James et al., 2013). Models were assessed and compared with RMSE, R^2 , Bias, RPD (Ratio of Performance to Deviation) and RPIQ (Ratio of Performance to Inter Quartile Range).

4.2.2.3 Modeling and prediction of soil bulk density

The goal of this analysis was to predict soil bulk density using *in situ* measurements of insertion force F , insertion depth D , and clay content and soil moisture content by the

multi-sensing penetrometer. Literature suggests that these four parameters are important to model bulk density (Bennie & Burger, 1988; Henderson, Levett, & Lisle, 1988; Hernanz, Peixoto, Cerisola, & Sánchez-Girón, 2000; Lin, Sun, & Schulze Lammers, 2014; Mirreh & Ketcheson, 1972). Two models suggested by Hernanz et al. (2000) were explored;

$$\rho_b = K_1 \cdot F^{K_2} \cdot D^{K_3} \quad (4.1)$$

$$\rho_b = K_1 \cdot F^{K_2} \cdot D^{K_3} \cdot MC^{K_4} \cdot CL^{K_5} \quad (4.2)$$

In Equations 4.1 and 4.2, ρ_b is bulk density (g cm^{-3}); F is insertion force (N); D is insertion depth (cm); MC is gravimetric soil moisture content (%); CL is clay content (%); and K 's are regression coefficients.

Note that F and D were directly measured by the multi-sensing penetrometer, whereas soil moisture and clay content were modeled from the *in situ* soil spectra measured by the VisNIR penetrometer. The total number of field samples available for bulk density analysis was 138 (see Table 4.1), which was divided into calibration (70%) and validation (30%) sets. The calibration set was used for two purposes; (1) developing PLSR models to predict soil moisture and clay content from *in situ* soil spectra, and (2) calibrating the regression coefficients K 's to predict bulk density. The calibrated models were then applied to the validation set to evaluate the model performance.

4.3 RESULTS AND DISCUSSION

4.3.1 Performance of the VisNIR penetrometer during the field test

The penetrometer system was tested in 11 agriculture fields in four states during different seasons and years. In all fields, the penetrometer performed satisfactorily with minor adjustments as needed. The system was first supplied by 12V 9Ah lead-acid battery, which could run the system for ~ 2 hours before recharge. Later, a 12V deep cycle battery was used to supply the whole system (including the laptop), which allowed the operation

for the whole day without recharge. In the first two NE fields, we also used a dark reference panel (2% reflectance) to characterize the internal scattering of the penetrometer. This practice was not continued in other fields as it did not improve the spectral quality or modeling performance significantly.

During insertion of the penetrometer, soil smearing or scratching on the fused quartz window of the optical aperture was not observed, which was important for the acquisition of genuine soil spectra at the registered depths. Compared to the dry-ground soil spectra, *in situ* soil spectra did show higher noise at 350 to 400 and 2450 to 2500 nm spectral ranges, which was due to energy attenuation by the long fiber optic cable used to carry reflected energy from the penetrometer head to the spectrometer. One key challenge of operation was to maintain a constant speed of penetration at $\sim 0.5 \text{ cm s}^{-1}$. At this speed, the LabVIEW program could obtain and record 10 *in situ* spectra corresponding to an insertion depth interval of 2 to 3 cm. This constant insertion speed also improved insertion force measurement by removing fluctuations due to acceleration or deceleration. One insertion of the penetrometer took 5 to 8 min. for completion (i.e. insert in and out). Overall, the system was capable of acquiring high quality, *in situ* soil spectra with low noise at 400 to 2450 nm, as well as the insertion force measurements.

4.3.2 Using *in situ* soil spectra from the VisNIR penetrometer to predict soil TC and TN

Table 4.3 gives prediction performance of soil TC and TN using the dry-ground library models on dry ground spectra, *in situ* spectra, and after EPO correction. When the models were applied to the dry-ground soil samples, the prediction performance was satisfactory: TC had a R^2 of 0.81, RPIQ of 2.73, and RMSE of 0.43%; and TN had a R^2 of 0.87, RPIQ of 2.17, and RMSE of 0.05%. However, when the dry-ground library models were applied directly to the *in situ* soil spectra from the penetrometer, their prediction capability degraded substantially. TC was predicted with R^2 of 0.29, RPIQ of 0.38, and RMSE of

3.06; and TN was predicted with R^2 of 0.51, RPIQ of 0.33, and RMSE of 0.36%.

Table 4.3. Prediction model statistics for soil total carbon (TC) and total nitrogen (TN). Models were calibrated on the USDA-NRCS dry-ground spectral library and applied to the dry-ground spectra and *in situ* (from the VisNIR multi-sensing penetrometer) of the field samples.

Property	Spectra source	RMSE ^f	R ²	Bias	RPD ^g	RPIQ ^h
TC ^a (%)	DG ^c	0.43	0.81	-0.14	2.07	2.73
	VNIRP ^d	3.06	0.29	2.22	0.29	0.38
	EPO ^e	0.79	0.5	-0.11	1.12	1.48
TN ^b (%)	DG ^c	0.05	0.87	0.04	1.75	2.17
	VNIRP ^d	0.36	0.51	0.35	0.26	0.33
	EPO ^e	0.06	0.62	-0.01	1.49	1.86

^aTotal carbon; ^bTotal nitrogen; ^cDry ground spectral; ^dPenetrometer spectra; ^eExternal Parameter Orthogonalization; ^fRoot mean squared error; ^gRatio of performance to deviation; ^hRatio of performance to inter-quartile range

These results indicated that dry-ground models were able to predict for the dry-ground spectra of the field samples but failed for *in situ* spectra that were not corrected using EPO. This was due to the differences in dry-ground and *in situ* conditions. The dry-ground model was calibrated on the VisNIR spectra of air-dried sieved samples; whereas the *in situ* spectra were obtained in the *in situ* field moist intact condition. These two conditions differed because of moisture content, temperature, and inhomogeneity. Among these factors, soil moisture content was known to alter soil VisNIR spectra significantly and has been investigated by several authors (Ji et al., 2015; Wijewardane, Ge, & Morgan, 2016b). Figure 4.3 depicts the differences between dry ground and *in situ* spectra for field samples. This discrepancy highlights the need to employ EPO, which allows models calibrated on the dry-ground spectra to predict samples scanned *in situ*.

The EPO correction improved predictions from the *in situ* VisNIR spectra for both TC and TN (Table 4.3). For TC, prediction RMSE reduced from 3.06% to 0.79%; bias reduced from 2.22 to -0.11; and R^2 increased from 0.29 to 0.50. Similarly, for TN, RMSE

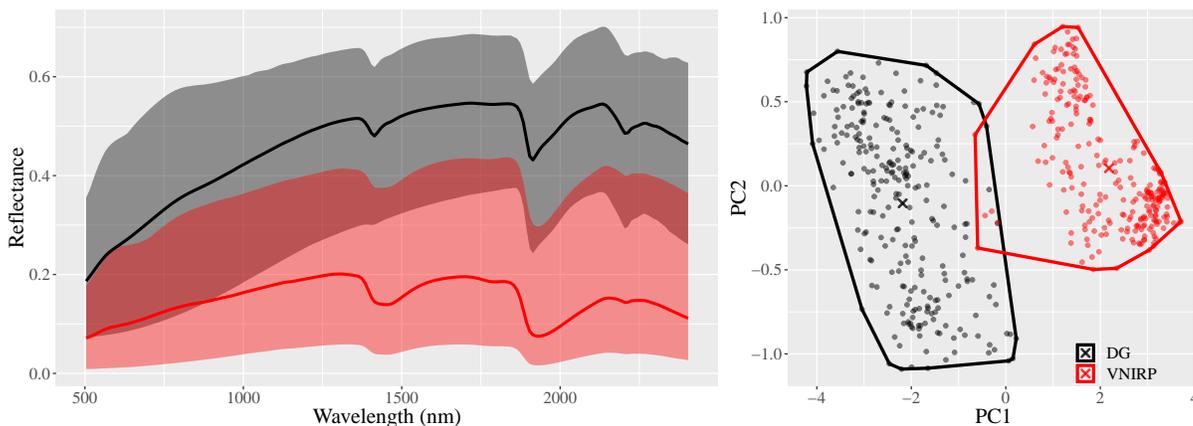


Figure 4.3. Difference in soil VisNIR spectra acquired in the lab (dry ground) and by the VisNIR multi-sensing penetrometer (*in situ*).

reduced from 0.36 to 0.06%; bias reduced from 0.35 to -0.01%; and R^2 increased from 0.51 to 0.62. Scatterplots of the predictions are given in Figure 4.4 to provide the visual assessment on how EPO improved the prediction of *in situ* soil spectra. Without EPO, both TC and TN were predicted with a higher bias than the lab-measured values. This over prediction was attributed to the difference between the dry-ground and *in situ* spectra, primarily from soil moisture effects (Figure 4.3). EPO successfully removed this over prediction and made the points closer to the 1:1 line. The linear association between lab-measured values and EPO-predicted values also became stronger. However, predictions after the EPO correction with *in situ* spectra did not exceed the accuracy achieved by dry-ground predictions. This result was also expected, because EPO (or any other moisture correction methods for that matter) was designed to minimize the effect of external parameters on prediction but would not reduce prediction bias or uncertainty that were already present in modeling dry-ground spectra. The potential of EPO to improve *in situ* spectral predictions was also reported in Ackerson et al. (2017).

4.3.3 Modeling soil bulk density

Penetration force (or cone index), moisture content, texture, and depth have been reported in literature as key soil physical properties contributing to bulk density (Bennie & Burger,

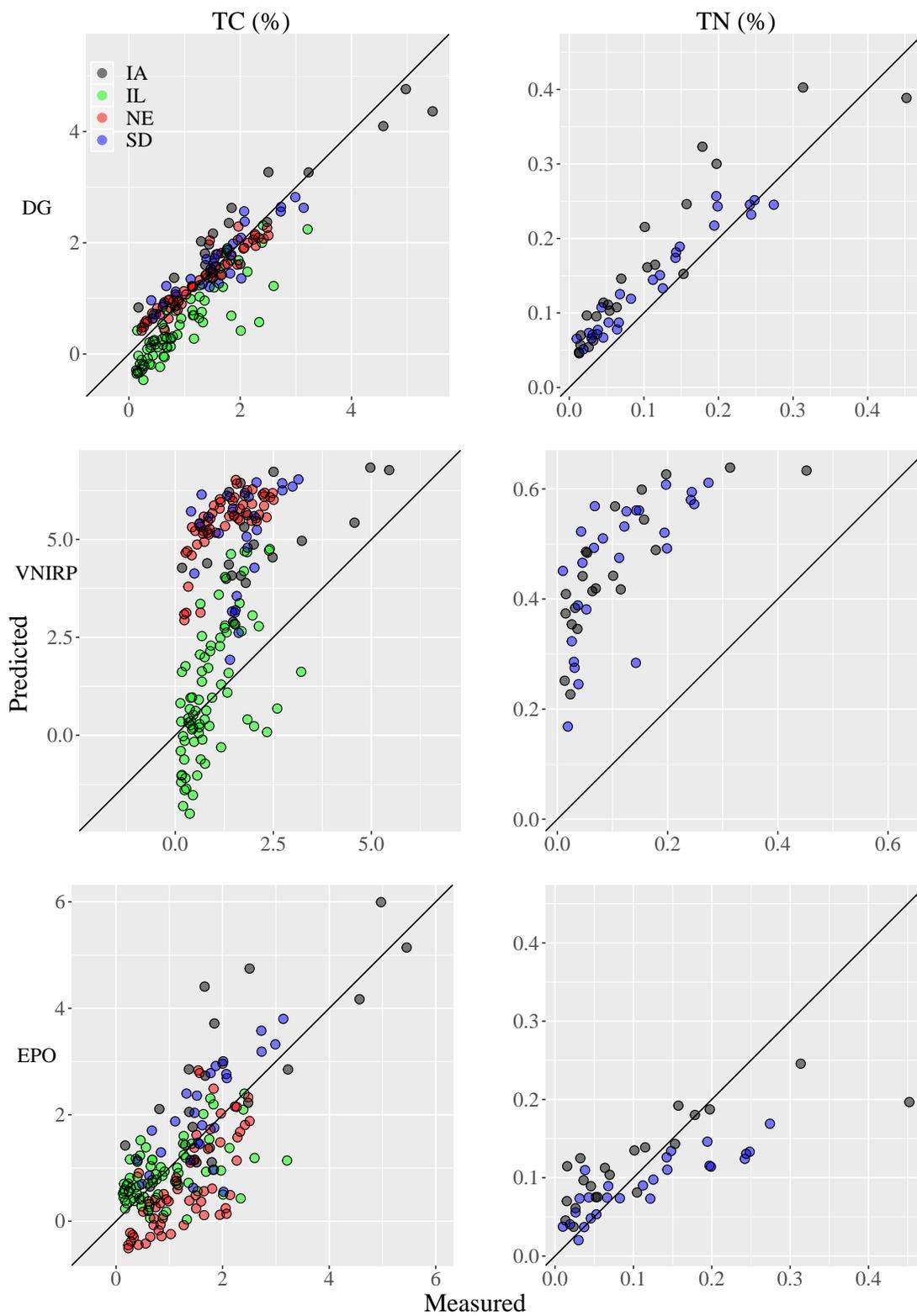


Figure 4.4. Prediction plots for dry-ground (DG) spectra, uncorrected *in situ* (VNIRP) spectra, and external parameter orthogonalization corrected (EPO) *in situ* spectra.

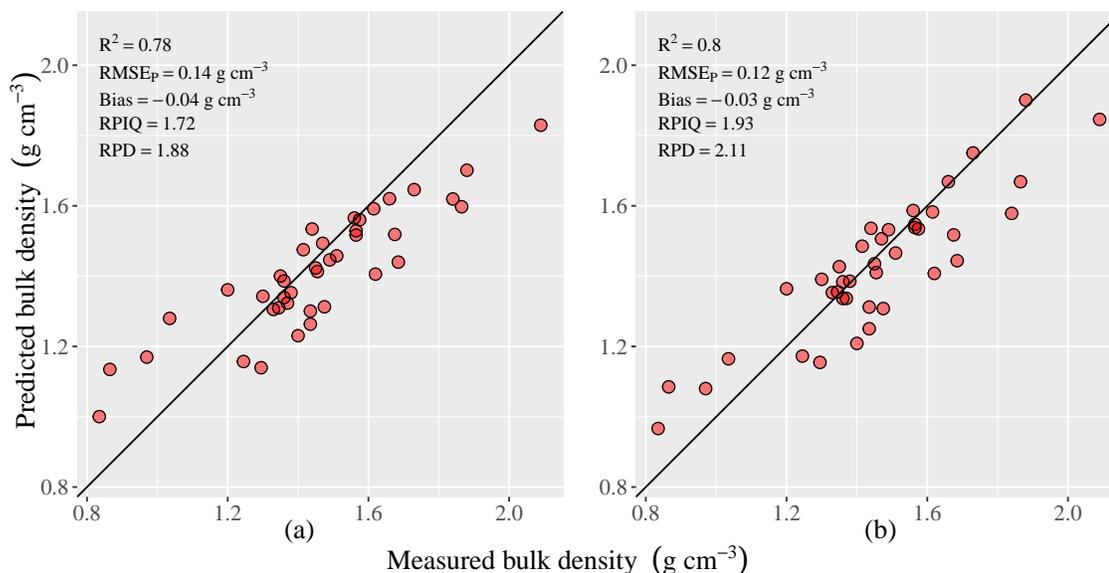


Figure 4.5. Prediction plots for soil bulk density modeling with (a) insertion force F and insertion depth D , and (b) insertion force F , insertion depth D , as well as gravimetric moisture content and clay content estimated from *in situ* VisNIR data corrected with external parameter orthogonalization.

1988; Ehlers, Köpke, Hesse, & Böhm, 1983; Elbanna & Witney, 1987; Hernanz et al., 2000; Lin et al., 2014). Figure 4.5a shows the validation result using Eq. 4.1, or only insertion force F and insertion depth D in the model. The model had moderate accuracy with an R^2 of 0.78 and RMSE of 0.14 g cm^{-3} . This model underestimated bulk density of soil depths with high bulk density. Including VisNIR-predicted clay and moisture contents (Equation 4.2) improved the model performance, with R^2 of 0.80 and RMSE of 0.12 g cm^{-3} (Figure 4.5b). Moreover, all points were evenly distributed around the 1:1 line, with no obvious prediction bias for high bulk density samples.

4.3.4 VisNIR integrated multi-sensing penetrometer

Since the multi-sensing penetrometer obtained soil VisNIR data and other measurements at high vertical resolution ($\sim 2.5 \text{ cm}$), these data could be used to estimate soil properties along the profile at the same resolution. An example showing the variation of the five soil properties (TC, TN, bulk density, gravimetric soil moisture, and clay) along the soil profile

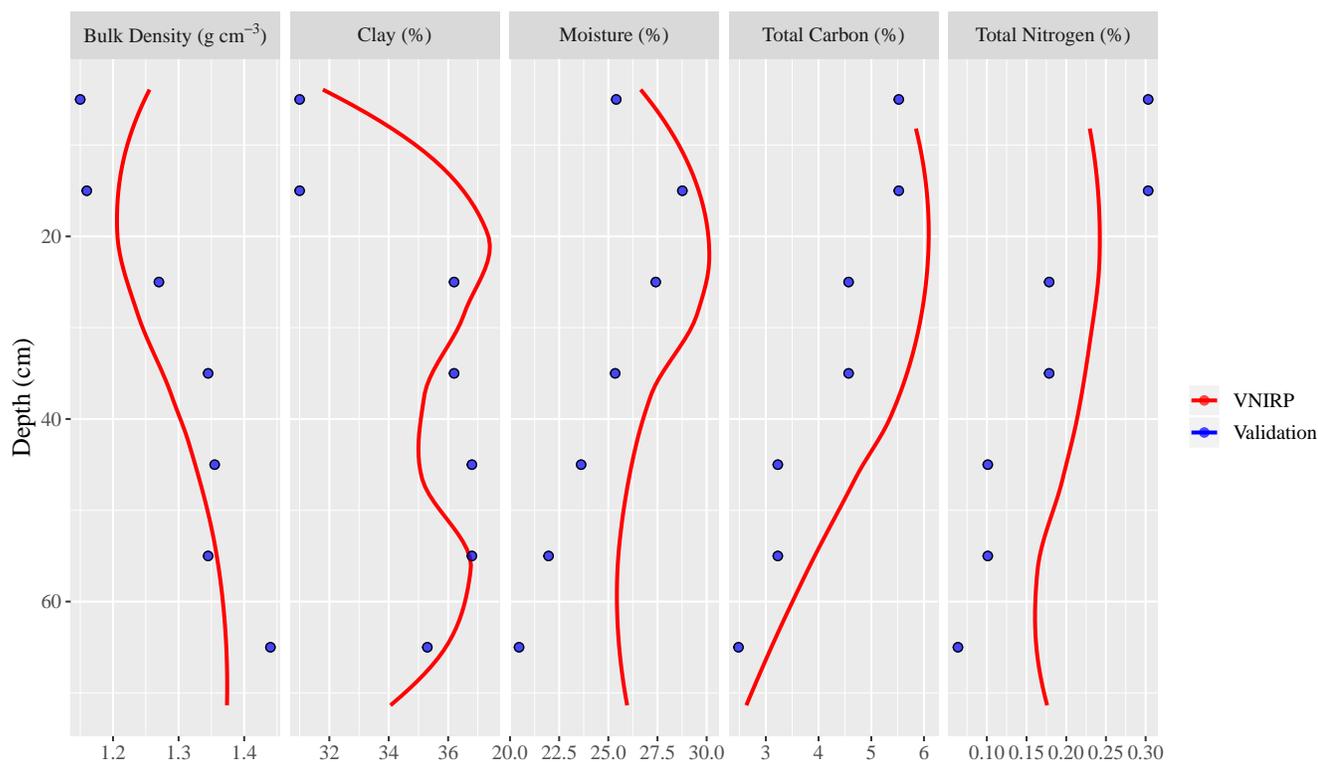


Figure 4.6. High resolution predictions of bulk density, clay, gravimetric moisture content, Total Carbon, and Total Nitrogen along the soil profile for one selected sampling location. The blue points represented the lab-analyzed values and the red lines represented the high-resolution predictions from the penetrometer data.

from our data is given in Figure 4.6. The continuous curves represent the high-resolution estimate from the penetrometer data, whereas the regularly spaced points represented the lab-analyzed values at each depth. Although the curves did not fall exactly on the points, they clearly captured the trends of the soil properties along the profile. Soil bulk density showed an increase as a function of depth, whereas TC, TN and gravimetric soil moisture decrease. For clay, there was an abrupt transition between A and B horizons (approximately at 20 cm). This transition was also captured in the clay curve, more or less in a continuous fashion.

Once again, we would like to emphasize that obtaining high resolution, vertical distribution of soil properties, like in Figure 4.6, was rapid, compared to needing subsequent lab measurements. We achieved this by inserting our multi-sensing

penetrometer into the soil and continuously logging *in situ* soil VisNIR spectra along with insertion force F and insertion depth D . These *in situ* measurements were then used to predict targeted soil properties without the need to pull soil cores or open soil pits. The multi-sensing soil penetrometer therefore fills a current technological gap for rapid and low-cost direct sensing of subsoils at high vertical resolution. This kind of soil data would benefit a number of disciplines such as precision agriculture and site-specific management of irrigation and fertilization, soil carbon inventory and change detection, and digital mapping of soil properties across various scales.

When VisNIR-based, on-the-go soil sensors were reported in the literature, more than half of the field samples were lab-analyzed with the reference methods for model calibration (Aliah Baharom et al., 2015; Kodaira & Shibusawa, 2013; Maleki et al., 2007; Mouazen et al., 2007; Veum et al., 2018; Zhang, Biswas, Ji, & Adamchuk, 2017). While this approach was acceptable in the research setting, it would not be viable in real applications, because it was neither practical nor economical to lab-analyze more than half of the field samples for each and every sampling campaign. Using an external, independent soil spectral library would eliminate this need by providing calibration samples with known soil properties. However, this approach caused another problem. The library samples were scanned in the dry-ground state, whereas field samples were scanned *in situ* with varying moisture contents and aggregation. An EPO (or other spectral correction methods) became an essential piece of the system to bridge the dry-ground spectral library and *in situ* scans of the field samples.

Our results demonstrate that EPO consistently improved soil TC and TN prediction (improved R^2 and reduced bias and RMSE), even though the prediction accuracy was still quite low compared to direct dry-ground prediction (Table 4.3 and Fig. 4.4). In addition to EPO, other methods including direct standardization, piece-wise direct standardization and spiking were also investigated in the literature for spectral correction (Ji et al., 2016; Ji et al., 2015; Wijewardane, Ge, & Morgan, 2016b). Future research is

needed to comprehensively evaluate these methods and identify the ones that are high in performance and logistically favorable to implement for *in situ* soil spectral data.

We will continue to develop the VisNIR integrated multi-sensing penetrometer in the following aspects. We will incorporate a Wenner electrode array (either ring-type or button-type) into the penetrometer. The Wenner electrode array, when in good contact with the soils, can measure bulk soil electrical conductivity (Adamchuk et al., 2010; Corwin & Lesch, 2005) and infer moisture content of non-saline soils. This measurement could provide independently measured soil moisture to substantiate EPO and spectral correction, as well as improve modeling for soil bulk density.

Currently, predictions of soil properties were carried out in two steps. At first, *in situ* soil VisNIR spectra and other sensor data along the soil profile were obtained during the field campaign. Then, EPO correction and VisNIR modeling were conducted in the office for the prediction of soil properties. Therefore, another improvement is to incorporate the dry-ground VisNIR spectral library (and calibrated models) and the EPO implementation into system's software program. By doing so we will realize true *in situ* prediction of soil properties while the penetrometer is taking soil measurements.

4.4 CONCLUSIONS

In this chapter, we reported a new VisNIR-integrated multi-sensing penetrometer for *in situ*, high resolution vertical soil sensing. Field testing of this penetrometer demonstrated its potential to quantify several targeted soil properties including total carbon, total nitrogen, and bulk density. The use of an external, independent soil spectral library for model calibration, and the use of the EPO algorithm to correct for the spectral difference between the dry-ground spectra in the library and the *in situ* spectra acquired by the penetrometer, was demonstrated. These steps were essential for the viability of the penetrometer in real applications. Finally, we showed the example of high-resolution vertical distribution of the soil properties along a profile. It was concluded that the new

soil penetrometer could greatly enhance our ability to measure properties of subsoil, and benefit disciplines such as precision agriculture and digital soil mapping where high-resolution vertical soil data are currently lacking.

CHAPTER 5

GENERAL CONCLUSIONS AND FUTURE WORK

The overall goal of the work presented in this dissertation was to design, develop, and test a VisNIR integrated multi-sensing penetrometer to estimate soil properties in vertical profile. This goal was further expanded to several specific objectives.

Soil VisNIR spectra is affected by intactness, i.e. moisture, temperature, and aggregation. If a VisNIR based sensor needs to use a dry ground spectral library to obtain calibration samples, this “dry ground–field intactness” gap has to be closed. To that end, the first objective was devised to investigate and compare the usefulness of five approaches: External Parameter Orthogonalization (EPO), Direct Standardization (DS), Global Moisture Modeling (GMM), Slope Bias Correction (SB) and Selective Wavelength Modeling (SWM), to enable VisNIR dry ground models to be applied directly to moist soil spectra to predict soil organic carbon (OC) and inorganic carbon (IC).

To test the applicability of these techniques, we conducted a rewetting experiment with selected soil samples from a soil archive at the USDA-NRCS-KSSL (Kellogg’s Soil Survey Laboratory). The results of this study revealed that EPO, moisture-explicit DS, DS and GMM can correct for the moisture effect successfully and increase the accuracy of prediction substantially for both soil OC and IC, whereas SB and SWM showed otherwise. GMM and moisture-explicit DS demonstrated a consistent correction performance across different moisture levels, which is desirable for field samples. However, the use of moisture-explicit DS in the field would be limited because the moisture content of field samples is not known *a priori*.

Integration of VisNIR technology to a soil penetrometer demands the optical components arranged neatly which can acquire quality spectra under field conditions. To

this end, the second objective was to design new VisNIR probes and test them in terms of spectral quality and predictive power using an external spectral library under laboratory conditions. Under this objective, two VisNIR probes: one with a bifurcated cable with an angled ferrule and the light source outside and the second with a broad-spectrum light source (a halogen lamp) and a 45° parabolic reflective mirror inside, were developed and tested with 150 dry ground soil samples. A spectral library was used for calibrating models for OC, Total Carbon (TC) and Total Nitrogen (TN).

The results showed that second design was superior in terms of spectral quality. In spite of the high spectral quality acquired by the VisNIR probe, there was still a systematic difference when compared with the standard MugLite® spectra. Then direct standardization and sample spiking were identified as approaches to account for this difference and substantially improve the prediction accuracy by reducing the bias. Spiking with extra weight showed the highest accuracy with bias lower than 0.06% and Ratio of Performance to Deviation (RPD) greater than 2.15 for all three soil properties.

With the experience gained through the first two objectives, the third and final objective were devised to develop the fully integrated, multi-sensing penetrometer system for high resolution vertical soil sensing and field test the penetrometer to evaluate its performance. The developed penetrometer incorporated an improved VisNIR sensing probe from second objective (coupled to an ASD spectrometer), a load cell, an ultrasonic depth sensor, a GPS receiver, and a LabVIEW software program for automated data collection and storage. This system was tested in 11 agricultural fields in Nebraska (NE), Illinois (IL), Iowa (IA), and South Dakota (SD).

Field testing of this VisNIR integrated multi-sensing penetrometer showed its potential to quantify several targeted soil properties including TC, TN, and bulk density (BD). An independent dry ground soil spectral library was used for model calibration and the EPO algorithm was successfully used to account for the spectral discrepancies due to field intactness under *in situ* conditions. Estimating soil properties in vertical soil profile at

high resolution was also demonstrated using the penetrometer. This newly developed soil VisNIR integrated multi-sensing penetrometer could enhance the measurement of subsoil properties, and benefit disciplines such as precision agriculture, digital soil mapping, and other environmental studies where high-resolution vertical soil data are required.

Through the work presented in this dissertation, we developed, and field tested a VisNIR integrated multi-sensing penetrometer which can acquire high quality spectra under the field conditions and be used to map soil profile attributes such as total carbon, organic carbon, total nitrogen and bulk density, at a higher resolution. The real strength of this system was originated from the use of already available dry ground spectral library to calibrate models instead of field sample spectra which reduces the cost. However, all these analyses were conducted after the field campaigns. In the future, we want to incorporate the whole process to the developed LabVIEW program so that it can acquire *in situ* field spectra, use in-built model to predict soil properties with EPO correction, and show on the screen in real time. In addition, we expect to incorporate a Wenner electrode array (either ring-type or button-type) into the penetrometer which can measure bulk soil electrical conductivity and infer moisture content of non-saline soils. This can provide independently measured soil moisture to substantiate EPO and spectral correction, as well as improve modeling for soil bulk density. With all these modifications and upgrades, more field testing under diverse conditions and geographic regions has to be conducted to further ensure its robustness in the field.

REFERENCES

- Ackerson, J. P., Demattê, J. A. M., & Morgan, C. L. S. (2015). Predicting clay content on field-moist intact tropical soils using a dried, ground VisNIR library with external parameter orthogonalization. *Geoderma*, 259–260(0), 196–204. doi:<http://dx.doi.org/10.1016/j.geoderma.2015.06.002>
- Ackerson, J. P., Morgan, C. L. S., & Ge, Y. (2017). Penetrometer-mounted VisNIR spectroscopy: Application of EPO-PLS to in situ VisNIR spectra. *Geoderma*, 286, 131–138. doi:<https://doi.org/10.1016/j.geoderma.2016.10.018>
- Adamchuk, V. I., Hummel, J. W., Morgan, M. T., & Upadhyaya, S. K. (2004). On-the-go soil sensors for precision agriculture. *Computers and Electronics in Agriculture*, 44(1), 71–91. doi:<http://dx.doi.org/10.1016/j.compag.2004.03.002>
- Adamchuk, V. I., Viscarra Rossel, R. A., Hartemink, A. E., & McBratney, A. B. (2010). Development of on-the-go proximal soil sensor systems. In R. A. Viscarra Rossel, A. B. McBratney, & B. Minasny (Eds.), *Proximal soil sensing* (pp. 15–28). doi:10.1007/978-90-481-8859-8
- Aliah Baharom, S. N., Shibusawa, S., Kodaira, M., & Kanda, R. (2015). Multiple-depth mapping of soil properties using a visible and near infrared real-time soil sensor for a paddy field. *Engineering in Agriculture, Environment and Food*, 8(1), 13–17. doi:<http://dx.doi.org/10.1016/j.eaef.2015.01.002>
- Barnes, E. M., Sudduth, K. A., Hummel, J. W., Lesch, S. M., Corwin, D. L., Yang, C., . . . Bausch, W. C. (2003). Remote- and Ground-Based Sensor Techniques to Map Soil Properties. *Photogrammetric Engineering & Remote Sensing*, 69(6), 619–630. doi:10.14358/PERS.69.6.619
- Basak, D., Pal, S., & Patranabis, D. C. (2007). Support vector regression. *Neural Information Processing-Letters and Reviews*, 11(10), 203–224.
- Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.-M., & McBratney, A. B. (2010). Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *TrAC Trends in Analytical Chemistry*, 29(9), 1073–1081. doi:<http://dx.doi.org/10.1016/j.trac.2010.05.006>
- Ben-Dor, E. (2002). Quantitative remote sensing of soil properties. *Advances in agronomy*, Volume 75, 173–243. doi:[http://dx.doi.org/10.1016/S0065-2113\(02\)75005-0](http://dx.doi.org/10.1016/S0065-2113(02)75005-0)

- Ben-Dor, E., & Banin, A. (1995). Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Science Society of America Journal*, 59(2), 364–372.
- Ben-Dor, E., Heller, D., & Chudnovsky, A. (2008). A novel method of classifying soil profiles in the field using optical means. *Soil Sci. Soc. Am. J.* 72(4), 1113–1123. doi:10.2136/sssaj2006.0059
- Bennie, A. T. P., & Burger, R. d. T. (1988). Penetration resistance of fine sandy apedal soils as affected by relative bulk density, water content and texture. *South African Journal of Plant and Soil*, 5(1), 5–10. Retrieved from https://journals.co.za/content/plantsoil/5/1/AJA02571862%7B%5C_%7D303
- Bowers, S. A., & Hanks, R. J. (1965). Reflection of radiant energy from soils. *Soil Science*, 100(2), 130–138.
- Bricklemeyer, R. S., & Brown, D. J. (2010). On-the-go VisNIR: Potential and limitations for mapping soil clay and organic carbon. *Computers and Electronics in Agriculture*, 70(1), 209–216. doi:<https://doi.org/10.1016/j.compag.2009.10.006>
- Brown, D. J. (2007). Using a global VNIR soil-spectral library for local soil characterization and landscape modeling in a 2nd-order Uganda watershed. *Geoderma*, 140(4), 444–453. doi:<http://dx.doi.org/10.1016/j.geoderma.2007.04.021>
- Brown, D. J., Shepherd, K. D., Walsh, M. G., Dewayne Mays, M., & Reinsch, T. G. (2006). Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma*, 132(3–4), 273–290. doi:<http://dx.doi.org/10.1016/j.geoderma.2005.04.025>
- Burns, D. A., & Ciurczak, E. W. (2007). *Handbook of near-infrared analysis*. CRC press.
- Chang, C.-W., Laird, D. A., & Hurburgh, C. R. J. (2005). Influence of soil moisture on near-infrared reflectance spectroscopic measurement of soil properties. *Soil Science*, 170(4), 244–255.
- Chang, C.-W., Laird, D. A., Mausbach, M. J., & Hurburgh, C. R. (2001). Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. *Soil Science Society of America Journal*, 65(2), 480–490.
- Cho, Y., Sheridan, A. H., Sudduth, K. A., & Veum, K. S. (2017). Comparison of Field and Laboratory VNIR Spectroscopy for Profile Soil Property Estimation. *Transactions of the ASABE*, 60(5), 1503–1510. doi:<https://doi.org/10.13031/trans.12299>

- Christy, C. D. (2008). Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. *Computers and Electronics in Agriculture*, *61*(1), 10–19. doi:<http://dx.doi.org/10.1016/j.compag.2007.02.010>
- Clark, R. N. (1999). Spectroscopy of rocks and minerals, and principles of spectroscopy. *Manual of remote sensing*, *3*, 3–58.
- Clark, R. N., King, T. V. V., Klejwa, M., Swayze, G. A., & Vergo, N. (1990). High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysical Research: Solid Earth (1978–2012)*, *95*(B8), 12653–12680.
- Corwin, D. L., & Lesch, S. M. (2003). Application of Soil Electrical Conductivity to Precision Agriculture. *Agronomy Journal*, *95*(3), 455–471. doi:[10.2134/agronj2003.4550](https://doi.org/10.2134/agronj2003.4550)
- Corwin, D. L., & Lesch, S. M. (2005). Apparent soil electrical conductivity measurements in agriculture. *Computers and Electronics in Agriculture*, *46*(1), 11–43. doi:<https://doi.org/10.1016/j.compag.2004.10.005>
- Dalal, R. C., & Henry, R. J. (1986). Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry. *Soil Science Society of America Journal*, *50*(1), 120–123.
- Daniel, K. W., Tripathi, N. K., & Honda, K. (2003). Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Soil Research*, *41*(1), 47–59.
- Davis, J. L., & Annan, A. P. (1989). Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy1. *Geophysical Prospecting*, *37*(5), 531–551. doi:[10.1111/j.1365-2478.1989.tb02221.x](https://doi.org/10.1111/j.1365-2478.1989.tb02221.x)
- Doetterl, S., Stevens, A., Van Oost, K., & van Wesemael, B. (2013). Soil Organic Carbon Assessment at High Vertical Resolution using Closed-Tube Sampling and Vis-NIR Spectroscopy. *Soil Science Society of America Journal*, *77*(4). doi:[10.2136/sssaj2012.0410n](https://doi.org/10.2136/sssaj2012.0410n)
- Ehlers, W., Köpke, U., Hesse, F., & Böhm, W. (1983). Penetration resistance and root growth of oats in tilled and untilled loess soil. *Soil and Tillage Research*, *3*(3), 261–275. doi:[https://doi.org/10.1016/0167-1987\(83\)90027-2](https://doi.org/10.1016/0167-1987(83)90027-2)
- Elbanna, E. B., & Witney, B. D. (1987). Cone penetration resistance equation as a function of the clay ratio, soil moisture content and specific weight. *Journal of Terramechanics*, *24*(1), 41–56. doi:[https://doi.org/10.1016/0022-4898\(87\)90058-9](https://doi.org/10.1016/0022-4898(87)90058-9)

- Fearn, T. (2001). Standardisation and calibration transfer for near infrared instruments: A review. *Journal of near Infrared Spectroscopy*, 9(4), 229–244.
- Feudale, R. N., Woody, N. A., Tan, H., Myles, A. J., Brown, S. D., & Ferré, J. (2002). Transfer of multivariate calibration models: a review. *Chemometrics and Intelligent Laboratory Systems*, 64(2), 181–192.
doi:http://dx.doi.org/10.1016/S0169-7439(02)00085-0
- Ge, Y., Morgan, C. L. S., & Ackerson, J. P. (2014). VisNIR spectra of dried ground soils predict properties of soils scanned moist and intact. *Geoderma*, 213(0), 61–69.
doi:http://dx.doi.org/10.1016/j.geoderma.2014.01.011
- Ge, Y., Morgan, C. L. S., Grunwald, S., Brown, D. J., & Sarkhot, D. V. (2011). Comparison of soil reflectance spectra and calibration models obtained using multiple spectrometers. *Geoderma*, 161(3), 202–211.
doi:http://dx.doi.org/10.1016/j.geoderma.2010.12.020
- Ge, Y., Thomasson, J. A., Morgan, C. L., & Searcy, S. W. (2007). VNIR diffuse reflectance spectroscopy for agricultural soil property determination based on regression-kriging. *Transactions of the ASABE*, 50(3), 1081–1092.
- Gogé, F., Gomez, C., Jolivet, C., & Joffre, R. (2014). Which strategy is best to predict soil properties of a local site from a national Vis–NIR database? *Geoderma*, 213, 1–9.
doi:http://dx.doi.org/10.1016/j.geoderma.2013.07.016
- Gomez, C., Viscarra Rossel, R. A., & McBratney, A. B. (2008). Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma*, 146(3–4), 403–411.
doi:http://dx.doi.org/10.1016/j.geoderma.2008.06.011
- Griffis, C. L. [Carl L]. (1985). Electronic Sensing of Soil Organic Matter. *Transactions of the ASAE*, 28(3), 703–705. doi:https://doi.org/10.13031/2013.32323
- Guerrero, C., Stenberg, B., Wetterlind, J., Viscarra Rossel, R. A., Maestre, F. T., Mouazen, A. M., . . . Kuang, B. (2014). Assessment of soil organic carbon at local scale with spiked NIR calibrations: effects of selection and extra-weighting on the spiking subset. *European Journal of Soil Science*, 65(2), 248–263.
doi:10.1111/ejss.12129
- Guerrero, C., Wetterlind, J., Stenberg, B., Mouazen, A. M., Gabarrón-Galeote, M. A., Ruiz-Sinoga, J. D., . . . Viscarra Rossel, R. A. (2016). Do we really need large spectral libraries for local scale SOC assessment with NIR spectroscopy? *Soil and Tillage Research*, 155, 501–509. doi:http://dx.doi.org/10.1016/j.still.2015.07.008

- Hartemink, A. E., & Minasny, B. (2014). Towards digital soil morphometrics. *Geoderma*, 230–231, 305–317. doi:<http://dx.doi.org/10.1016/j.geoderma.2014.03.008>
- Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning*. Springer.
- Helland, I. (2004). Partial Least Squares Regression. In *Encyclopedia of statistical sciences*. doi:10.1002/0471667196.ess6004.pub2
- Henderson, C., Levett, A., & Lisle, D. (1988). The effects of soil water content and bulk density on the compactibility and soil penetration resistance of some Western Australian sandy soils. *Soil Research*, 26(2), 391–400. Retrieved from <https://doi.org/10.1071/SR9880391>
- Hernanz, J. L., Peixoto, H., Cerisola, C., & Sánchez-Girón, V. (2000). An empirical model to predict soil bulk density profiles in field conditions using penetration resistance, moisture content and soil depth. *Journal of Terramechanics*, 37(4), 167–184. doi:[https://doi.org/10.1016/S0022-4898\(99\)00020-8](https://doi.org/10.1016/S0022-4898(99)00020-8)
- Hummel, J. W., Sudduth, K. A., & Hollinger, S. E. (2001). Soil moisture and organic matter prediction of surface and subsurface soils using an NIR soil sensor. *Computers and Electronics in Agriculture*, 32(2), 149–165.
- Islam, K., Singh, B., & McBratney, A. B. (2003). Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. *Soil Research*, 41(6), 1101–1114.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.
- Ji, W., Li, S., Chen, S., Shi, Z., Viscarra Rossel, R. A., & Mouazen, A. M. (2016). Prediction of soil attributes using the Chinese soil spectral library and standardized spectra recorded at field conditions. *Soil and Tillage Research*, 155, 492–500. doi:<http://dx.doi.org/10.1016/j.still.2015.06.004>
- Ji, W., Viscarra Rossel, R. A., & Shi, Z. (2015). Accounting for the effects of water and the environment on proximally sensed vis–NIR soil spectra and their calibrations. *European Journal of Soil Science*, 66(3), 555–565. doi:10.1111/ejss.12239
- Kawano, S., Abe, H., & Iwamoto, M. (1995). Development of a calibration equation with temperature compensation for determining the Brix value in intact peaches. *Journal of near Infrared Spectroscopy*, 3(4), 211–218. Retrieved from <http://dx.doi.org/10.1255/jnirs.71>

- Kennard, R. W., & Stone, L. A. (1969). Computer aided design of experiments. *Technometrics*, *11*(1), 137–148. doi:10.1080/00401706.1969.10490666
- Knadel, M., Thomsen, A., Schelde, K., & Greve, M. H. (2015). Soil organic carbon and particle sizes mapping using vis–NIR, EC and temperature mobile sensor platform. *Computers and Electronics in Agriculture*, *114*, 134–144. doi:https://doi.org/10.1016/j.compag.2015.03.013
- Kocher, M. F., & Griffis, C. L. [C L]. (1989). Laboratory performance of a sampler for on-the-go in-situ soil carbon content measurement. *Paper-American Society of Agricultural Engineers (USA)*.
- Kodaira, M., & Shibusawa, S. (2013). Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping. *Geoderma*, *199*(0), 64–79. doi:http://dx.doi.org/10.1016/j.geoderma.2012.09.007
- Kuang, B., & Mouazen, A. M. (2013). Non-biased prediction of soil organic carbon and total nitrogen with vis–NIR spectroscopy, as affected by soil moisture content and texture. *Biosystems Engineering*, *114*(3), 249–258. doi:http://dx.doi.org/10.1016/j.biosystemseng.2013.01.005
- Kusumo, B. H., Hedley, M. J., Tuohy, M. P., Hedley, C. B., & Arnold, G. C. (2010). Predicting Soil Carbon and Nitrogen Concentrations and Pasture Root Densities from Proximally Sensed Soil Spectral Reflectance. In R. A. Viscarra Rossel, A. B. McBratney, & B. Minasny (Eds.), *Proximal soil sensing* (Chap. 15, pp. 177–190). doi:10.1007/978-90-481-8859-8_15
- Lagacherie, P., McBratney, A. B., & Voltz, M. (2006). *Digital soil mapping: An introductory perspective*. Elsevier.
- Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. *Science*, *304*(5677), 1623–1627. doi:10.1126/science.1097396
- Lin, J., Sun, Y., & Schulze Lammers, P. (2014). Evaluating model-based relationship of cone index, soil water content and bulk density using dual-sensor penetrometer data. *Soil and Tillage Research*, *138*, 9–16. doi:https://doi.org/10.1016/j.still.2013.12.004
- Lunt, I. A., Hubbard, S. S., & Rubin, Y. (2005). Soil moisture content estimation using ground-penetrating radar reflection data. *Journal of Hydrology*, *307*(1), 254–269. doi:https://doi.org/10.1016/j.jhydrol.2004.10.014
- Maleki, M. R., Mouazen, A. M., De Ketelaere, B., Ramon, H., & De Baerdemaeker, J. (2008). On-the-go variable-rate phosphorus fertilisation based on a visible and

- near-infrared soil sensor. *Biosystems Engineering*, 99(1), 35–46.
doi:<http://dx.doi.org/10.1016/j.biosystemseng.2007.09.007>
- Maleki, M. R., Mouazen, A. M., Ramon, H., & De Baerdemaeker, J. (2007). Optimisation of soil VIS–NIR sensor-based variable rate application system of soil phosphorus. *Soil and Tillage Research*, 94(1), 239–250.
doi:<http://dx.doi.org/10.1016/j.still.2006.07.016>
- Martens, H., & Naes, T. (1992). *Multivariate calibration*. John Wiley & Sons.
- Max, K., Jed, W., Weston, S., Williams, A., Keefer, C., Engelhardt, A., . . . Scrucca, L. (2015). Caret: Classification and regression training. Retrieved from <http://cran.r-project.org/package=caret>
- Mevik, B., Wehrens, R., & Liland, K. H. (2013). Pls: Partial least squares and principal component regression. Retrieved from <http://cran.r-project.org/package=pls>
- Miller, C. E. (2001). Chemical principles of near-infrared technology. *Near-infrared technology in the agricultural and food industries*, 2.
- Minasny, B., & McBratney, A. B. (2016). Digital soil mapping: A brief history and some lessons. *Geoderma*, 264, Part, 301–311.
doi:<http://dx.doi.org/10.1016/j.geoderma.2015.07.017>
- Minasny, B., McBratney, A. B., Bellon-Maurel, V., Roger, J.-M., Gobrecht, A., Ferrand, L., & Joalland, S. (2011). Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. *Geoderma*, 167–168(0), 118–124. doi:<http://dx.doi.org/10.1016/j.geoderma.2011.09.008>
- Minasny, B., McBratney, A. B., Pichon, L., Sun, W., & Short, M. G. (2009). Evaluating near infrared spectroscopy for field prediction of soil properties. *Soil Research*, 47(7), 664–673.
- Mirreh, H. F., & Ketcheson, J. W. (1972). Influence of soil bulk density and matric pressure on soil resistance to penetration. *Canadian Journal of Soil Science*, 52(3), 477–483. doi:10.4141/cjss72-059
- Montanarella, L., Tóth, G., & Jones, A. (2011). Soil components in the 2009 LUCAS survey. *Land Quality and Land Use Information–in the European Union*, 209–220.
- Morgan, C. L. S., Waiser, T. H., Brown, D. J., & Hallmark, C. T. (2009). Simulated in situ characterization of soil organic and inorganic carbon with visible near-infrared diffuse reflectance spectroscopy. *Geoderma*, 151(3–4), 249–256.
doi:<http://dx.doi.org/10.1016/j.geoderma.2009.04.010>

- Mouazen, A. M., De Baerdemaeker, J., & Ramon, H. (2005). Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer. *Soil and Tillage Research*, *80*(1), 171–183.
- Mouazen, A. M., Maleki, M. R., De Baerdemaeker, J., & Ramon, H. (2007). On-line measurement of some selected soil properties using a VIS–NIR sensor. *Soil and Tillage Research*, *93*(1), 13–27. doi:<http://dx.doi.org/10.1016/j.still.2006.03.009>
- Nocita, M., Stevens, A., Noon, C., & van Wesemael, B. (2013). Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy. *Geoderma*, *199*(0), 37–42. doi:<http://dx.doi.org/10.1016/j.geoderma.2012.07.020>
- Osborne, B. G., & Fearn, T. (1983). Collaborative evaluation of universal calibrations for the measurement of protein and moisture in flour by near infrared reflectance. *International Journal of Food Science & Technology*, *18*(4), 453–460. doi:[10.1111/j.1365-2621.1983.tb00287.x](http://dx.doi.org/10.1111/j.1365-2621.1983.tb00287.x)
- Poggio, M., Brown, D. J., & Brickley, R. S. (2015). Laboratory-based evaluation of optical performance for a new soil penetrometer visible and near-infrared (VisNIR) foreoptic. *Computers and Electronics in Agriculture*, *115*(0), 12–20. doi:<http://dx.doi.org/10.1016/j.compag.2015.05.002>
- Poggio, M., Roudier, P., Blaschek, M., & Hedley, C. (2018). Integration of NIR on a multi-sensor platform to improve soil resource assessments. *NIR news*, *29*(5), 15–18. doi:[10.1177/0960336018782035](http://dx.doi.org/10.1177/0960336018782035)
- Quraishi, M. Z., & Mouazen, A. M. (2013). A prototype sensor for the assessment of soil bulk density. *Soil and Tillage Research*, *134*(0), 97–110. doi:<http://dx.doi.org/10.1016/j.still.2013.07.011>
- R Core Team. (2018). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.r-project.org/>
- Revelle, W. (2015). Psych: Procedures for psychological, psychometric, and personality research. Evanston, Illinois: Northwestern University. Retrieved from <http://cran.r-project.org/package=psych>
- Revolution Analytics, & Weston, S. (2015). doParallel: Foreach Parallel Adaptor for the 'parallel' Package. CRAN. Retrieved from <https://cran.r-project.org/package=doParallel>
- Rodionov, A., Welp, G., Damerow, L., Berg, T., Amelung, W., & Pätzold, S. (2015). Towards on-the-go field assessment of soil organic carbon using Vis–NIR diffuse

- reflectance spectroscopy: Developing and testing a novel tractor-driven measuring chamber. *Soil and Tillage Research*, 145(0), 93–102.
doi:<http://dx.doi.org/10.1016/j.still.2014.08.007>
- Roger, J.-M., Chauchard, F., & Bellon-Maurel, V. (2003). EPO–PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. *Chemometrics and Intelligent Laboratory Systems*, 66(2), 191–204. doi:[http://dx.doi.org/10.1016/S0169-7439\(03\)00051-0](http://dx.doi.org/10.1016/S0169-7439(03)00051-0)
- Sankey, J. B., Brown, D. J., Bernard, M. L., & Lawrence, R. L. (2008). Comparing local vs. global visible and near-infrared (VisNIR) diffuse reflectance spectroscopy (DRS) calibrations for the prediction of soil clay, organic C and inorganic C. *Geoderma*, 148(2), 149–158. doi:<http://dx.doi.org/10.1016/j.geoderma.2008.09.019>
- Sarkhot, D. V., Grunwald, S., Ge, Y., & Morgan, C. L. S. (2011). Comparison and detection of total and available soil carbon fractions using visible/near infrared diffuse reflectance spectroscopy. *Geoderma*, 164(1–2), 22–32.
doi:<http://dx.doi.org/10.1016/j.geoderma.2011.05.006>
- Shepherd, K. D., & Walsh, M. G. (2002). Development of reflectance spectral libraries for characterization of soil properties. *Soil Sci. Soc. Am. J.* 66(3), 988–998.
doi:10.2136/sssaj2002.9880
- Sherman, D. M., & Waite, T. D. (1985). Electronic spectra of Fe³⁺ oxides and oxide hydroxides in the near IR to near UV. *American Mineralogist*, 70(11-12), 1262–1269.
- Shonk, J. L., Gaultney, L. D., Schulze, D. G., & Van Scoyoc, G. E. (1991). Spectroscopic sensing of soil organic matter content. *Transactions of the ASAE (USA)*.
- Soil Survey Staff. (2014). *Kellogg Soil Survey Laboratory Methods Manual. Soil Survey Investigations Report No. 42, Version 5.0*. U.S. Department of Agriculture, Natural Resources Conservation Service. Retrieved from http://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/ref/?cid=nrcs142p2%7B%5C_%7D054247
- Sørensen, L. K., & Dalsgaard, S. (2005). Determination of clay and other soil properties by near infrared spectroscopy. *Soil Science Society of America Journal*, 69(1), 159–167.
- Stenberg, B. (2010). Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. *Geoderma*, 158(1–2), 15–22. doi:<http://dx.doi.org/10.1016/j.geoderma.2010.04.008>

- Stenberg, B., Jonsson, A., & Börjesson, T. (2002). Near infrared technology for soil analysis with implications for precision agriculture. In *Near infrared spectroscopy: Proceedings of the 10th international conference, kyongju s. korea. nir publications, chichester, UK* (pp. 279–284).
- Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M., & Wetterlind, J. (2010). Visible and near infrared spectroscopy in soil science. *Advances in agronomy*, *107*, 163–215.
- Stevens, A., van Wesemael, B., Bartholomeus, H., Rosillon, D., Tychon, B., & Ben-Dor, E. (2008). Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. *Geoderma*, *144*(1–2), 395–404. doi:<http://dx.doi.org/10.1016/j.geoderma.2007.12.009>
- Sudduth, K. A., & Hummel, J. W. (1993). Soil organic matter, CEC, and moisture sensing with a portable NIR spectrophotometer. *Transactions of the ASAE (USA)*.
- Thissen, U., Pepers, M., Üstün, B., Melssen, W. J., & Buydens, L. M. C. (2004). Comparing support vector machines to PLS for spectral regression applications. *Chemometrics and Intelligent Laboratory Systems*, *73*(2), 169–179. doi:<http://dx.doi.org/10.1016/j.chemolab.2004.01.002>
- Triantafilis, J., Gibbs, I., & Earl, N. (2013). Digital soil pattern recognition in the lower Namoi valley using numerical clustering of gamma-ray spectrometry data. *Geoderma*, *192*, 407–421. doi:<http://dx.doi.org/10.1016/j.geoderma.2012.08.021>
- Turner, H., & Firth, D. (2015). Generalized nonlinear models in R. Retrieved from <http://cran.r-project.org/package=gnm>
- Vapnik, V. (2013). *The nature of statistical learning theory*. Springer Science & Business Media.
- Vereecken, H., Schnepf, A., Hopmans, J. W., Javaux, M., Or, D., Roose, T., . . . Young, I. M. (2016). Modeling Soil Processes: Review, Key Challenges, and New Perspectives. *Vadose Zone Journal*, *15*(5). Retrieved from <http://dx.doi.org/10.2136/vzj2015.09.0131>
- Veum, K. S., Parker, P. A., Sudduth, K. A., & Holan, S. H. (2018). Predicting Profile Soil Properties with Reflectance Spectra via Bayesian Covariate-Assisted External Parameter Orthogonalization. doi:10.3390/s18113869
- Viscarra Rossel, R. A., & Behrens, T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, *158*(1–2), 46–54. doi:<http://dx.doi.org/10.1016/j.geoderma.2009.12.025>

- Viscarra Rossel, R. A., & Bouma, J. (2016). Soil sensing: A new paradigm for agriculture. *Agricultural Systems*, 148, 71–74. doi:10.1016/j.agsy.2016.07.001
- Viscarra Rossel, R. A., Lobsey, C. R., Sharman, C., Flick, P., & McLachlan, G. (2017). Novel soil profile sensing to monitor organic C stocks and condition. *Environmental Science & Technology*. doi:10.1021/acs.est.7b00889
- Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., & Skjemstad, J. O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131(1–2), 59–75. doi:http://dx.doi.org/10.1016/j.geoderma.2005.03.007
- Viscarra Rossel, R. A., & Webster, R. (2012). Predicting soil properties from the Australian soil visible–near infrared spectroscopic database. *European Journal of Soil Science*, 63(6), 848–860.
- Waiser, T. H., Morgan, C. L. S., Brown, D. J., & Hallmark, C. T. (2007). In situ characterization of soil clay content with visible near-infrared diffuse reflectance spectroscopy. *Soil Science Society of America Journal*, 71(2), 389. doi:10.2136/sssaj2006.0211
- Wang, Y., Veltkamp, D. J., & Kowalski, B. R. (1991). Multivariate instrument standardization. *Analytical Chemistry*, 63(23), 2750–2756. doi:10.1021/ac00023a016
- Wetterlind, J., Piikki, K., Stenberg, B., & Söderström, M. (2015). Exploring the predictability of soil texture and organic matter content with a commercial integrated soil profiling tool. *European Journal of Soil Science*, 66(4), 631–638. doi:10.1111/ejss.12228
- Wetterlind, J., & Stenberg, B. (2010). Near-infrared spectroscopy for within-field soil characterization: small local calibrations compared with national libraries spiked with local samples. *European Journal of Soil Science*, 61(6), 823–843. doi:10.1111/j.1365-2389.2010.01283.x
- Wetterlind, J., Stenberg, B., & Söderström, M. (2008). The use of near infrared (NIR) spectroscopy to improve soil mapping at the farm scale. *Precision agriculture*, 9(1-2), 57–69. doi:10.1007/s11119-007-9051-z
- Wickham, H. (2009). ggplot2: Elegant graphics for data analysis.
- Wijewardane, N. K., Ge, Y., & Morgan, C. (2016a). Moisture insensitive prediction of soil properties from VNIR reflectance spectra based on external parameter

- orthogonalization. *Geoderma*, 267, 92–101.
doi:<http://dx.doi.org/10.1016/j.geoderma.2015.12.014>
- Wijewardane, N. K., Ge, Y., & Morgan, C. (2016b). Prediction of soil organic and inorganic carbon at different moisture contents with dry ground VNIR: a comparative study of different approaches. *European Journal of Soil Science*, 67(5), 605–615. doi:10.1111/ejss.12362
- Wijewardane, N. K., Ge, Y., Wills, S., & Loecke, T. (2016). Prediction of Soil Carbon in the Conterminous United States: Visible and Near Infrared Reflectance Spectroscopy Analysis of the Rapid Carbon Assessment Project. *Soil Science Society of America Journal*, 80(4), 973–982. doi:10.2136/sssaj2016.02.0052
- Wills, S., Loecke, T., Sequeira, C., Teachman, G., Grunwald, S., & West, L. (2014). Overview of the U.S. Rapid Carbon Assessment project: sampling design, initial summary and uncertainty estimates. In A. E. Hartemink & K. McSweeney (Eds.), *Soil carbon* (Chap. 10, pp. 95–104). doi:10.1007/978-3-319-04084-4_10
- Wold, S., Martens, H., & Wold, H. (1983). The multivariate calibration problem in chemistry solved by the PLS method. In *Matrix pencils* (pp. 286–293). Springer.
- Wu, C.-Y., Jacobson, A. R., Laba, M., & Baveye, P. C. (2009). Alleviating moisture content effects on the visible near-infrared diffuse-reflectance sensing of soils. *Soil Science*, 174(8), 456–465. doi:10.1097/SS.0b013e3181b21491
- Zhang, Y., Biswas, A., Ji, W., & Adamchuk, V. I. (2017). Depth-Specific Prediction of Soil Properties In Situ using vis-NIR Spectroscopy. *Soil Science Society of America Journal*, 81, 993–1004. doi:10.2136/sssaj2016.08.0253
- Zhu, Y., Weindorf, D. C., Chakraborty, S., Haggard, B., Johnson, S., & Bakr, N. (2010). Characterizing surface soil water with field portable diffuse reflectance spectroscopy. *Journal of Hydrology*, 391(1–2), 133–140.
doi:<http://dx.doi.org/10.1016/j.jhydrol.2010.07.014>