IN SITU CHARACTERIZATION OF SOIL PROPERTIES USING VISIBLE NEAR-INFRARED DIFFUSE REFLECTANCE SPECTROSCOPY

A Thesis

by

TRAVIS HEATH WAISER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2006

Major Subject: Soil Science

IN SITU CHARACTERIZATION OF SOIL PROPERTIES USING VISIBLE NEAR-INFRARED DIFFUSE REFLECTANCE SPECTROSCOPY

A Thesis

by

TRAVIS HEATH WAISER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Cristine L. S. Morgan
Committee Members,	C. T. Hallmark
	Jerry Stuth
Head of Department,	C. Wayne Smith

May 2006

Major Subject: Soil Science

ABSTRACT

In Situ Characterization of Soil Properties Using Visible Near-Infrared Diffuse Reflectance Spectroscopy. (May 2006)

Travis Heath Waiser, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Cristine Morgan

Diffuse reflectance spectroscopy (DRS) is a rapid proximal-sensing method that is being used more and more in laboratory settings to measure soil properties. Diffuse reflectance spectroscopy research that has been completed in laboratories shows promising results, but very little has been reported on how DRS will work in a field setting on soils scanned *in situ*. Seventy-two soil cores were obtained from six fields in Erath and Comanche County, Texas. Each soil core was scanned with a visible near-infrared (VNIR) spectrometer with a spectral range of 350-2500 nm in four different combinations of moisture content and pre-treatment: field-moist *in situ*, air-dried *in situ*, field-moist smeared *in situ*, and air-dried ground. Water potential was measured for the field-moist *in situ* scans. The VNIR spectra were used to predict total and fine clay content, water potential, organic C, and inorganic C of the soil using partial least squares (PLS) regression. The PLS model was validated with data 30% of the original soil cores that were randomly selected and not used in the calibration model. The root mean squared deviation (RMSD) of the air-dry ground samples were within the *in situ* RMSD and comparable to literature values for each soil property. The validation data set had a total clay content root mean squared deviation (RMSD) of 61 g kg^{-1} and 41 g kg^{-1} for the

field-moist and air-dried *in situ* cores, respectively. The organic C validation data set had a RMSD of 5.8 g kg^{-1} and 4.6 g kg^{-1} for the field-moist and air-dried *in situ* cores, respectively. The RMSD values for inorganic C were 10.1 g kg^{-1} and 8.3 g kg^{-1} for the field moist and air-dried *in situ* scans, respectively. Smearing the samples increased the uncertainty of the predictions for clay content, organic C, and inorganic C. Water potential did not improve model predictions, nor did it correlate with the VNIR spectra; r^2-values were below 0.31. These results show that DRS is an acceptable technique to measure selected soil properties *in-situ* at varying water contents and from different parent materials.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Morgan, and my committee members, Dr. Hallmark and Dr. Stuth for their support and guidance throughout my research.

I would also like to thank my friends, fellow graduate students, faculty, and staff at Texas A&M University for their guidance and continual support. Appreciation goes to the Texas Water Research Institute for supporting my research and the Texas Agricultural Experiment Station for my assistantship. I appreciate the laboratory work which Dr. Drees and Donna Prochaska completed on clay mineralogy, clay content, and total carbon. I would also like to thank all the contributors of scholarships that I have received during the time at Texas A&M University. County extension agents, Joe Pope and Robert Hamilton, were instrumental in contacting land owners and getting permission for us to take soil samples, and I am grateful to the land owners for cooperating and allowing access to their land.

Finally, thanks to my mother, father, and parents-in-law for their steady encouragement and support and to my wife and daughter for their patience and love.

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION

Precision management of land at the sub-field (management zones) and smaller 10 m grid-sized scale has increased over the past 10 years because of inexpensive personal computers and innovations such as global positioning system (GPS) and geographical information system (GIS) technologies. In production agriculture, these technologies have led to yield monitors, tractors with GPS guidance systems, and GIS systems with the capability of monitoring daily farming activities in real-time. Precision farming advances have compelled the development of models based on high spatial and temporal resolution inputs that aid farm management and model water movement and solute transport across agricultural landscapes. Both precision agriculture and precision management are limited by our ability to quantify the spatial variability of SOIL. The soil survey has traditionally been the resource for quantifying soil variability; however, soil surveys have proven to be at spatial resolutions that are too low for in-field modeling (Packepsky et al., 2001; Ellert et al., 2002; Morgan et al., 2004; Sadler, 2004). The research presented in this study investigates a new tool that may provide rapid and reliable quantification of soil constituents across the landscape in greater detail, leading to mapping soils at higher resolutions.

The overall goal of this research is to evaluate the feasibility of visible near-infrared diffuse reflectance spectroscopy (VNIR-DRS) for *in situ* characterization of soil

---

This thesis follows the style of Soil Science Society of America Journal.

properties. Specifically, this research addresses the following objectives: 1) Evaluate the precision of 350 to 2500 nm (VNIR region) soil reflectance measurements in quantifying soil clay content, organic C, inorganic C, and water potential of *in situ* soils at field-moist and air-dry water contents; 2) Quantify any change in measurement error or prediction accuracy for *in situ*, field-moist soil with a smeared surface; and 3) Quantify any change in prediction accuracies using regional calibration models. The results of this research will help determine the feasibility and limitations of using a portable VNIR-DRS in the field.

Literature Review

VNIR Spectroscopy

Spectroscopy is the study of light as a function of wavelength that has been emitted, reflected, or scattered from a solid, liquid, or gas (Clark, 1999). When light strikes a material, light is absorbed, reflected, or transmitted, and spectral measurements can quantify the amount of light reflected or transmitted (Fig. 1.1) (Workman and Shenk, 2004). Diffuse reflectance spectroscopy measures the scattering of light reflected at all angles from a surface. When the diffuse reflectance of a material is measured, the absorbance bands provide information about the material's molecular composition. Absorbance peaks are viewed as valleys of the spectral signature when presented as reflectance (Fig. 1.1). The three key parameters in a spectrum that are important are the following: 1) the wavelength at which peaks occur, 2) the amplitude of the peak compared with a 100% reflected or transmitted standard, and 3) the bandwidth, which refers to the broadness of the peak (Workman and Shenk, 2004).

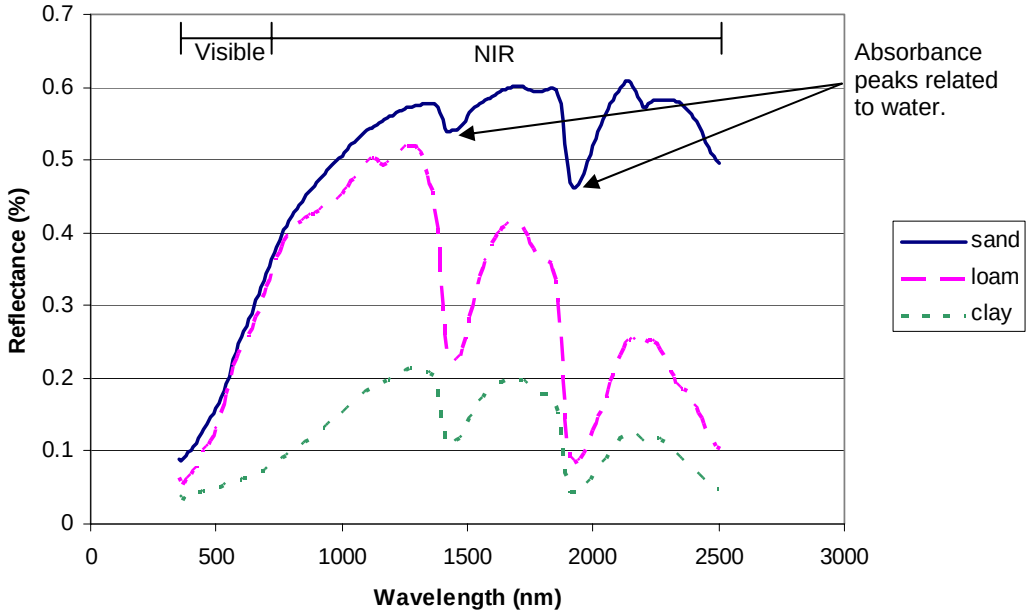

Fig.1.1. Soil reflectance of three soils from Erath County, Texas.

Visible to mid-infrared (MIR) spectroscopy has been used to quantify soil properties with varying accuracies and results. The visible, near-infrared (NIR), and MIR spectral ranges are 350 to 700 nm, 700 to 2500 nm, and 2500 to 25000 nm, respectively. McCarty et al. (2002) have shown that when measuring organic and inorganic C, the MIR region produced higher r^2-values and lower root mean squared deviations than NIR. Mid-infrared spectroscopy works better because the fundamental absorptions of interest in soils exist in the MIR range (McCarty et al., 2002). However, though MIR has been proven to give better prediction accuracies, MIR is less feasible for field and laboratory studies because of cost, portability, and required sample preparation. Hence, the majority of spectroscopy research in soils has occurred in the VNIR regions. The visible region gives information on iron oxides like hematite (Gaffey et al., 1993) while the NIR region is dominated by vibration overtones of SO_4^{2-}, CO_3^{2-}, and OH^- and combination bands of H_2O and CO_2 (Clark, 1999; McCarty et al., 2002). Though the overtones and combination bands in the NIR spectrum are more indirect signatures of key soil constituents, VNIR spectrometers are smaller, cheaper, and field portable as compared to MIR units (Janik et al., 1998).

Spectra

Before the reflectance measurements can be used for model calibration, the spectral data require pretreatment, which includes transformations, averaging, splicing, and smoothing. Transformations are used to condition the data for better model fits or to meet model assumptions such as normality. First and second derivatives of transformations are used to reduce albedo effects. For air-dried, ground samples, the

first derivative generally performs best (Reeves III et al., 1999). Splicing is required when the spectrometer contains several detectors across the wavelength spectra to eliminate interruptions in reflectance where detectors overlap. Smoothing the data reduces noise caused by changes in atmospheric pressure and temperature and light intensity variations (Williams, 1987; Workman, 2004).

Statistical Analysis of Spectra

Once spectral data have been processed, several techniques are available to extract information. Regression techniques, such as step-up, step-wise, partial least squares (PLS), and boosted trees, as well as principal component analysis have been used previously. Confalonieri et al. (2001) compared step-up regression, step-wise regression, and modified PLS regression methods and found that modified PLS regression created the calibrations with the best fit to the data. Partial least squares regression is the most common in soil science literature (Janik et al., 1998; Reeves III, 1999; Reeves III and McCarty, 2001; Dunn et al., 2002; Lee et al., 2003). Partial least squares regression is an orthogonal data compression method that allows researchers to look at two dimensional spectra in multidimensional space. The advantage of looking at the data in multidimensional space is that some of the redundancy is removed, patterns can be described from the center of the spectra, and the distances between peaks can be quantified (Workman and Shenk, 2004).

Partial least squares regression is used to construct a predictive model with many factors (also called predictors or X-variables) that are highly collinear (Tobias, 1995; Wold et al., 2001). One advantage PLS has over multiple linear regression and principal

components is that PLS is more robust, meaning that the calibration model changes little with new calibration samples (Geladi and Kowalski, 1986). There are several assumptions with PLS. The first assumption is that the model is built from a small number of latent variables (Wold et al., 2001). These latent variables are the most important points to the model and carry more weight in determining the predicting property. The concept of latent variables allows the assumption that the X and Y variables are not independent, allowing for a few spectrum versus the spectra to predict some property of a given material (Wold et al., 2001). The second assumption is that a multidimensional function F(**u**,**v**) is created from the X and Y data (Wold et al., 2001). The **u**-vector describes changes in the observations and the **v**-vector describes changes within the spectra variables (Wold et al., 2001). A third assumption is made concerning homogeneity of samples, meaning that the parameters that influence X on Y stay the same (Wold et al., 2001).

Partial least squares regression starts by converting the X-variables (X, spectral data) into two vectors called X-scores (T) and X-loadings (P') (Fig. 1.2), which makes PLS similar to principal component analysis (Geladi and Kowalski, 1986). The Y-variables (Y, soil laboratory data) are treated the same way by creating Y-score (Q') and Y-loading (U) vectors. The following formulas express these two outer relationships,

$$X = TP' + E \text{ and} \quad (1.1)$$

$$Y = UQ' + F^*. \quad (1.2)$$

The E and F* values are errors or the residuals. An inner relationship which links the X and Y blocks together is,

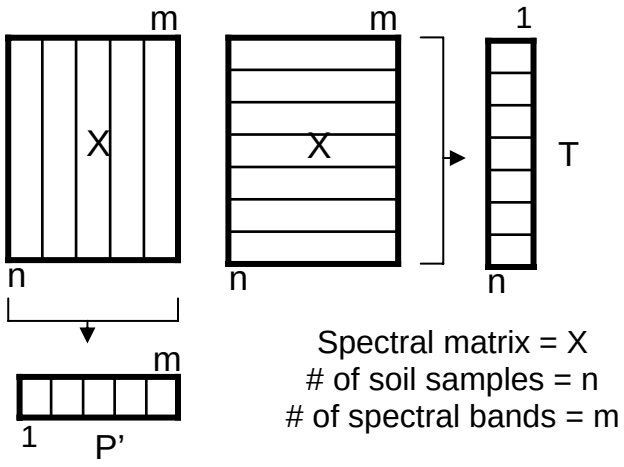

Fig. 1.2. Diagram showing how X-loadings (P') are converted from columns in the X-variable matrix and the X-scores (T) from the rows in the X-variable matrix in partial least squares regression. The same technique is used with the Y-matrix as well. (Figure from Geladi and Kowalski, 1986)

$$U = bT, \quad (1.3)$$

where b is the regression coefficient. These three equations are used in principal component analysis, but in PLS the P' value is replaced by weights (W'). Weights have to be used in PLS because the order of operations is changed, and otherwise the orthogonal t-values would not get calculated (Geladi and Kowalski, 1986). Equations (1.2) and (1.3) are combined to give the mixed relationship in PLS where U' is a row vector and B is b_1b_2 (Geladi and Kowalski, 1986),

$$W' = U'X/U'U \text{ and} \quad (1.4)$$

$$Y = TBQ' + F. \quad (1.5)$$

Soil Properties in VNIR Spectrum

Visible near-infrared DRS scans of air-dried, ground soil have been used to directly measure soil properties such as mineralogy, clay content, organic C, inorganic C, and water content (Ben-Dor et al., 1999). These soil properties are considered direct measurements because each has absorption bands in the VNIR region (Ben-Dor et al., 1999). Figure 1.3 shows the reflectance of four clay minerals commonly found in soil. Other soil properties like cation exchange capacity, potassium, phosphorus, pH, sodium (salt), electrical conductivity, and extractable bases have also been predicted using VNIR, but are considered indirect measurements, sometimes leading to lower prediction accuracies (Malley et al., 2004). These measurements are considered indirect because the soil properties do not have direct absorbance bands in the VNIR region, but are related to soil properties with absorptions in the VNIR region.

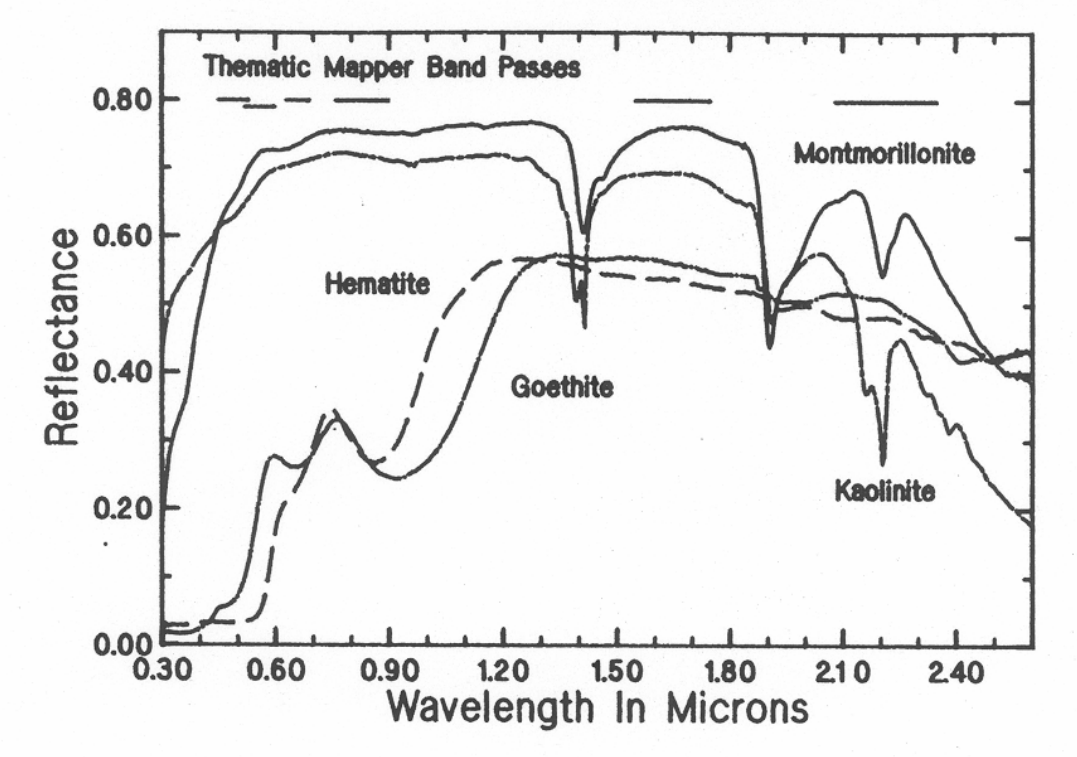

Fig. 1.3. Comparison of visible near-infrared reflectance of four common soil minerals (Figure from Mustard and Sunshine, 1999).

Clay content, organic C, inorganic C, and water potential will be discussed in proceeding chapters, but the air-dried, ground soil sample r^2-values for clay content, organic C, and inorganic C have ranged from 0.56 to 0.79, 0.49 to 0.96, and 0.69 to 0.96, respectively. Soil water correlations with VNIR reflectance have not been previously made where water is expressed as potential but have been studied in terms of gravimetric and volumetric water content. The VNIR-DRS literature demonstrates the use of VNIR to quantify soil clay content, organic C, and inorganic C on air-dried, ground samples, which suggest that VNIR can quantify these same properties on *in situ* soil samples, with some loss in prediction accuracy. Sudduth and Hummel's (1993) research predicted organic C in the field (*in situ*) with r^2-values of 0.85 and 0.89, exhibiting the potential exists for VNIR-DRS to be used *in situ* to quantify soil properties in real-time.

CHAPTER II

IN SITU CHARACTERIZATION OF SOIL CLAY CONTENT

Synopsis

Diffuse reflectance spectroscopy (DRS) is a rapid proximal sensing method that has proven useful in quantifying constituents of dried and ground soil samples. However, very little is known about how DRS will perform in a field setting on soils scanned *in-situ*. The overall goal of this research was to evaluate the feasibility of visible near-infrared (VNIR) DRS for *in situ* quantification of clay content of soil from a variety of parent materials. Seventy-two soil cores were obtained from six fields in Erath and Comanche Counties, Texas. Each soil core was scanned with a visible near-infrared spectrometer, with a spectral range of 350-2500 nm, at four different combinations of moisture content and pre-treatment: field-moist *in situ*, air-dried *in situ*, field-moist smeared *in situ*, and air-dried ground. The VNIR spectra were used to predict total and fine clay content of the soil using partial least squares (PLS) regression. The PLS model was validated with 30% of the original soil cores that were randomly selected and not used in the calibration model. The validation data set had a root mean squared deviation (RMSD) of 61 g kg^{-1} and 41 g kg^{-1} for the field-moist and air-dried *in situ* cores, respectively. The RMSD of the air-dry ground samples was between the two *in situ* RMSDs and comparable to values in the literature. Smearing the samples increased the field-moist *in situ* RMSD to 74 g kg^{-1}. Measured water potential did not improve model predictions. Whole-field holdout validation results showed that soils from all parent materials need to be represented in the calibration samples for maximum predictability.

In summary, DRS is an acceptable technique to measure soil clay content *in situ* at variable water contents and different parent materials.

Introduction

The resolution of a 1:24,000 scale soil survey map is too coarse to capture soil variability within soil mapping units and transitions between soil mapping units when used as inputs for some agriculture or environmental uses. However, soil maps that capture soil variability at a 10-50-m scale are necessary for resource management such as, precision agriculture, non-point source pollution modeling, and resource use planning (Packepsky et al., 2001; Ellert et al., 2002). A typical method for creating such a high resolution soil map includes some sort of survey technique to capture the spatial variability followed by collecting soil cores and laboratory analyses of those cores. Researchers have developed methods for mapping the variability of soil properties across landscapes, including remote sensing, proximal landscape sensors (GPS, electromagnetic), and terrain modeling (Moore et al., 1993; Sudduth et al., 1997; Zhu et al., 1997). Each of these methods has the advantage of creating a high-resolution map of horizontal soil variability, but is limited in the ability to get high-resolution, vertical soil information. Collecting and analyzing soil cores to capture that vertical variability is time consuming and cost prohibitive. For example, current laboratory soil analyses can take several weeks to months for each pedon and costs upwards of $2,000. The lack of soil sensors that can rapidly quantify soil profile information demonstrates the need to find new methods for soil mapping. One such method is visible and near-infrared diffuse reflectance spectroscopy (VNIR-DRS). Recent research has shown the

effectiveness of VNIR-DRS in providing a non-destructive rapid prediction of soil physical, chemical, and biological properties of air-dried ground soil samples in the laboratory (Shepard and Walsh, 2002). However, it is uncertain how VNIR-DRS will handle *in situ* soil analysis in the field (Sudduth and Hummel, 1993).

Sudduth and Hummel (1993) tested a portable near-infrared (NIR) spectrometer to measure soil properties *in situ.* Their research concluded that VNIR-DRS could be used to accurately predict cation exchange capacity and moisture content in the laboratory, but the technique was not sufficiently accurate at quantifying soil organic matter in a furrow. They concluded that movement of the sample across the sensor while scanning was the main contributor in poor estimates of organic matter. Most spectroscopy work has focused on air-dried, ground soil samples (Slaughter et al., 2001; Brown et al., 2005b; Sorensen and Dalsgaard, 2005). Air-drying the sample reduces the intensity of bands that are related to water so the signals associated with other soil properties are not masked or hidden. Additionally, particle size affects accuracy of the spectral scan. Smaller particles increase reflectance scatter which reduces the absorption peak height (Workman and Shenk, 2004). The most common sample preparation has been drying the soil and grinding the sample into a fine powder (< 0.6 mm) for VNIR scanning (Ben-Dor and Banin 1995; Reeves III et al., 1999; Confalonieri et al., 2001; Reeves III and McCarty, 2001; Dunn et al., 2002; Lee et al., 2003).

Soil Properties in VNIR Spectrum

Measuring soil water content can be made with numerous kinds of equipment. In the VNIR region, water has four absorption features associated with the OH group (Ben-

Dor et al., 1999). The absorption peaks associated with water are 1900, 1400, 1200, and 950 nm, ordered from very strong signal to a very weak signal (Ben-Dor et al., 1999). Slaughter et al. (2001) and Islam et al. (2003) predicted soil gravimetric moisture with r^2-values ranging from 0.81 to 0.98 using the VNIR spectrum. Lobell and Asner (2002) showed that reflectance decreased with increasing water content and suggested that the longer wavelengths in the VNIR region were better at predicting volumetric water content.

In the VNIR spectrum, absorptions by water bonds associated with clay content and other bonding associated with clay type provide the opportunity to use VNIR for quantifying clay information in soil. Previous research on air-dried, ground soil samples has shown VNIR predictions of soil clay content with r^2-values ranging from 0.56 to 0.91 and RMSD's ranging from 23 g kg^{-1} to 10.8 g kg^{-1} (Ben-Dor and Banin, 1995; Janik et al., 1998, Shepherd and Walsh, 2002; Islam et al., 2003; Brown et al., 2005b). Clay mineralogy can be inferred in VNIR reflectance measurements by looking at overtones and combination bands that occur from chemical bonds within the minerals (Brown et al., 2005b). Kaolinite [$Al_2Si_2O_5(OH)_4$] is an aluminum silicate with two very strong hydroxyl bands near 1400 and 2200 nm (Hunt and Salisbury, 1970; Clark et al., 1990). Smectite (montmorillonite) has two very strong water bands around 1400 and 1900 nm due to molecular water and an Al-OH band at 2200 nm (Hunt and Salisbury, 1970; Goetz et al., 2001). Brown et al. (2005b) compared X-ray diffraction to VNIR ordinal "peak intensity" values and showed that 96% and 88% of the time kaolinite and montmorillonite, respectively, were within one ordinal unit from agreeing. Goetz et al.

(2001) investigated using NIR spectroscopy to measure smectite content in selected Colorado soils. Their prediction models had a correlation coefficient (r) of 0.83. Muscovite [$KAl_3Si_3O_{10}(OH)_2$] displays hydroxyl bands at 1400 nm and between 2200 to 2600 nm (Hunt and Salisbury, 1970). The iron oxides, hematite and goethite, have absorptions in the visible range of the spectrum. Hematite (Fe_2O_3) has absorption bands at 860, 630, and 450 nm (Gaffey et al., 1993). Goethite (α-FeOOH) has absorption bands at 940, 660, and 493 nm (Scheinost and Schwertmann, 1997).

In the field, DRS measurements pose problems that have not been addressed by research on air-dried ground soils. Problems that need to be addressed in research for *in situ* analysis are varying amounts of soil moisture, varying particle size and aggregation, smearing of soil surfaces, small scale heterogeneity (mottles, accumulations, and redox features), and regionality of calibration models. Sudduth and Hummel (1993) indicated that local buried residue and roughness of the soil surface varying from moisture content caused a reduction in their prediction accuracies. Smearing of the soil causes the reflectance properties to change, possibly altering prediction accuracies. The smearing problem was likely visited by Sudduth and Hummel (1993) who noted differences with changes of soil surface roughness.

Researchers have studied how the regionality of soils affects prediction accuracies. Ben-Dor and Banin (1995) used soils with an organic matter range of 0.09% to 13.23% and obtained poor predictions compared to Dalal and Henry (1986). Dalal and Henry (1986) used three soils which had a smaller range of organic C. Ben-Dor and Banin (1995) speculate that the samples Dalal and Henry (1986) used were in similar

decomposition stages, which indicates a regionality effect, resulting in high r^2-values. Sudduth and Hummel (1996) concluded that the geographic range of the soil samples affects the prediction accuracies of NIR.

Research Objectives

The overall goal of this research is to evaluate the feasibility of VNIR-DRS for *in situ* quantification of clay content of soil profiles from a variety of parent materials. Specifically this research addresses the following objectives: 1) Evaluate the precision of 350 to 2500 nm (VNIR region) soil reflectance measurements in quantifying soil clay content and water potential of *in situ* soils at field-moist and air-dry water content; 2) Quantify any change in prediction accuracies with consideration to inter-field regionality; and 3) Quantify any change in measurement error or prediction accuracy for *in situ*, field-moist soil with a smeared surface.

The results of this research will allow a quantitative determination of the feasibility of using a VNIR spectrometer in the field as a proximal sensor to aid soil mapping activities. For VNIR-DRS to be considered useful in the field, it must be reliable, meaning that the errors should be equal to or less than current field techniques and it must be rapid, meaning that it should quantify clay much faster than any current techniques. There is no question that VNIR-DRS is faster than hand texturing or lab particle size methods (i.e. pipette and hydrometer). Therefore the accuracy of VNIR-DRS is considered adequate if VNIR-DRS can quantify clay content accurate enough for precision management and watershed modeling needs.

Materials and Methods

Soil Coring

Seventy-two soil cores were collected using a Giddings hydraulic soil sampler (Windsor, CO) attached to a truck from six fields in Erath County (fields 1, 2, 3, and 6) and Comanche County (fields 4 and 5), Texas in May 2004. Each soil core was collected to a maximum depth of 105 cm or to the depth of a coring-restrictive horizon. The soil cores were contained in a 6.0-cm diameter plastic sleeve that was capped on both ends. The soil cores that were collected were chosen to represent the large variability in soil properties over these two counties. For example, 21 soil series were mapped by the Natural Resources Conversation Service in these fields (Table 2.1), and the parent material of these soils included alluvium, sandstone, shale, and limestone.

VNIR-DRS Scanning

An ASD "FieldSpec® Pro FR" VNIR spectroradiometer (Analytical Spectral Devices, Boulder, CO), with a spectral range of 350-2500 nm, 2-nm sampling resolution and spectral resolution of 3 nm at 700 nm and 10 nm at 1400 and 2100 nm and equipped with a contact probe, was used to scan the soil cores. The contact probe has a viewing area defined by a 2-cm diameter circle and its own light source. A Spectralon® panel with 99% reflectance was used to optimize the spectrometer each day; the same panel was used as a white reference before scanning each core. The 72 soil cores were prepared for the first scan by first slicing lengthwise the plastic sleeve surrounding each core with a utility knife, and then halving the soil core lengthwise (surface to subsoil) with a piano wire. One-half of the core was smeared using a stainless steel spatula to

Table 2.1. Series and taxonomic names of the soils mapped by the Natural Resources Conservation Service within the fields which were used for VNIR-DRS clay content predictions (USDA, 1973; USDA, 1977; USDA, 2005).

Soil series	Taxonomic name	Field number
Abilene	Fine, mixed, superactive, thermic Pachic Argiustolls	4
Altoga	Fine-silty, carbonatic, thermic Udic Haplustepts	6
Blanket	Fine, mixed, superactive, thermic Pachic Argiustolls	6
Bolar	Fine-loamy, carbonatic, thermic Udic Calciustolls	1,5,6
Bosque	Fine-loamy, mixed, superactive, thermic Cumulic Haplustolls	5
Brackett	Loamy, carbonatic, thermic, shallow Typic Haplustepts	5
Bunyan	Fine-loamy, mixed, active, nonacid, thermic Typic Ustifluvents	2
Chaney	Fine, mixed, active, thermic Oxyaquic Paleustalfs	4
Cisco	Fine-loamy, siliceous, superactive, thermic Typic Haplustalfs	4
Denton	Fine-silty, carbonatic, thermic Udic Calciustolls	1,6
Frio	Fine, smectitic, thermic Cumulic Haplustolls	1,4,5
Houston Black	Fine, smectitic, thermic Udic Haplusterts	1
Karnes	Coarse-loamy, carbonatic, thermic Typic Calciustepts	5
Lewisville	Fine-silty, mixed, thermic Udic Calciustolls	5,6
Maloterre	Loamy, carbonatic, thermic Lithic Ustorthents	1
Nimrod	Loamy, siliceous, active, thermic Aquic Arenic Paleustalfs	3,4
Pedernales	Fine, mixed, superactive, thermic Typic Paleustalfs	4
Purves	Clayey, smectitic, thermic Lithic Calciustolls	1,5,6
Selden	Fine-loamy, siliceous, active, thermic Aquic Paleustalfs	3
Venus	Fine-loamy, mixed, thermic Udic Calciustolls	5
Windthorst	Fine, mixed, active, thermic Udic Paleustalfs	2,3,6

mimic possible smearing by a soil probe and the other half was left unsmeared. A wire grid was used to identify two columns and 3-cm wide rows within each core half (Fig. 2.1). Each row within each column was scanned twice with the FieldSpec® Pro FR with a 90º rotation of the contact probe between scans. Both halves of each core, smeared and unsmeared, were scanned at field-moist water content. The water potential was measured at each soil horizon from each core, with a maximum of six samples per core, using a SC-10 thermocouple psychrometer (Decagon, Pullman, WA) (Rawlins and Campbell, 1986). The cores were then placed in a drier at 44°C for two days for air-dried *in situ* scans. Before air-dried scans, the cores were removed from the drier and left on the countertop in the laboratory to equilibrate to room temperature. The soil core half that was scanned in unsmeared condition was then rescanned, *in situ*, at air-dry moisture content.

Each row of the unsmeared soil core was ground and passed through a 2-mm sieve, and re-scanned as air-dried, ground. The air-dried, ground soils were scanned with a mug lamp connected to the FieldSpec® Pro FR. The same Spectralon® 99% reflectance panel was used to calibrate the spectrometer each day, as well as being used as the white reference to set reflectance to 100%. Approximately 28 g of ground soil was placed into a borosilicate glass "puck". Each sample was scanned twice with a 90º rotation between scans.

Sampling for Laboratory Analysis

The soil cores were divided by row for laboratory analysis (Fig. 2.1). The entire row from each half of the core was used for particle size analysis. Particle size

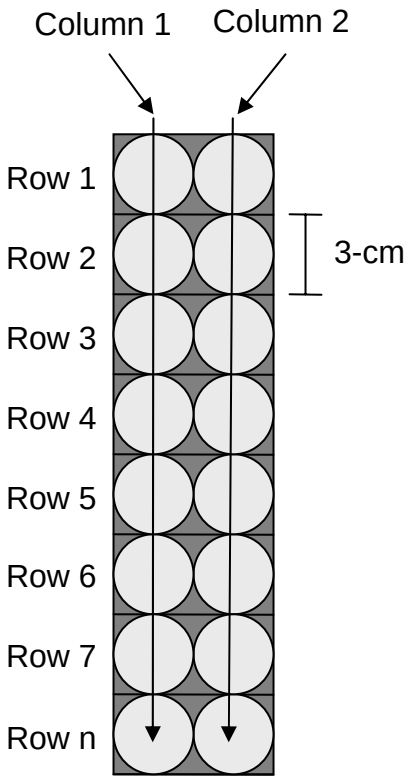

Fig. 2.1. Schematic of a vertically sliced soil core used for clay content. Columns and rows indicate locations scanned using the contact probe for *in situ* scans.

distribution was determined in the laboratory using the pipette method with an error of ±1% clay (Steele and Bradfield, 1934; Kilmer and Alexander, 1949; Gee and Or, 2002). Eighteen samples were selected for determination of clay mineralogy. Clay mineralogy by X-ray diffraction was completed following techniques described in Hallmark et al. (1986); no pretreatment was done. The mineralogy technique used was to determine the absence or presence of clay minerals, existing in quantities of greater than 5 % by volume of clay fraction.

Pretreatment of Data

Pretreatment of the spectral data included, splicing, averaging, and taking the 1st and 2nd derivatives. The spectral data were spliced where the three detectors overlapped. Results of the two scans, 0º and 90º, were averaged (mean). The mean and 1st and 2nd derivatives were taken on 10-nm intervals from 360-2490 nm after a cubic smoothing spline, implemented in the R "smooth spline" function (R Development Core Team, 2004), was fit to each raw spectral curve.

Model Building and Validation

Four different models were built to help determine how *in situ* scans and water content affect prediction accuracies for clay content. The input for the models were from the air-dried ground, air-dried *in situ*, field-moist *in situ*, and smeared field-moist *in situ* scans. Water potential models were built using the field-moist *in situ* scans. The models for clay content and water potential were produced using 70% of the soil cores randomly chosen as the calibration samples. Whole cores were used to maintain independence between the calibration and validation data (Brown et al., 2005a). Two

prediction models for clay content were built using soil reflectance, one model used the 1st derivative and the other used the 2nd derivative of soil reflectance. The prediction models were built using 1/25th cross validation PLS method in Unscrambler 9.0 (CAMO Tech, Woodbridge, NJ). The remaining 30% of the cores were used to validate the model. Negative clay content predictions were changed to zero clay content before comparison of measured to predicted clay. Measured versus predicted values of the validation samples were compared using simple regression. The coefficient of determination (r^2), root mean squared deviation (RMSD), ratio of standard deviation (SD) to RMSD (RPD) and bias were calculated to compare the accuracy of different PLS models. Statistical formulas to calculate RMSD, RPD, and bias follow Gauch et al. (2003), Brown et al. (2005a) and Chang et al. (2005):

$$\text{RMSD} = \sqrt{\sum_{n} \left(Y_{pred} - Y_{meas}\right)^2 / N}\,, \tag{2.1}$$

$$\text{RPD} = \text{SD/RMSD, and} \tag{2.2}$$

$$\text{Bias} = \sum_{n} \left(Y_{pred} - Y_{meas}\right) / N; \tag{2.3}$$

where Y_{pred} are predicted values of the validation set using the PLS model, and Y_{meas} are the laboratory measurements of the validation set, and N is the total number of samples in the validation data.

Model calibration and validation sets were also completed on whole-field holdouts as a means to compare to the findings of Brown et al. (2005a). Whole-field holdouts were achieved by calibrating a model using PLS with five of the six fields. The

sixth field was held out as the validation samples. Six models were created so all six fields were represented as a validation set.

Significant wavelengths were chosen by Unscrambler 9.0 using an uncertainty test. Regression coefficients for each wavelength were calculated with a t-test and uncertainty limits that did not cross the zero line were significant. Uncertainty limits corresponded to two standard deviations in an ideal case.

Results and Discussion

Sample Descriptions

From the 72 soil cores, 270 soil samples were analyzed for clay content and water potential. Of these samples, 188 were used in the calibration model. The other 82 samples were used to validate the PLS model. Though selected randomly, the calibration and validation data sets were similar (Table 2.2). A t-test proved the means of the calibration and validation sets were not dissimilar with $p>0.1$. Clay content range of the 270 soil samples was 12 g kg^{-1} to 578 g kg^{-1}. The mean and median for the calibration and validation data were similar, indicating that the samples were evenly split above and below the mean, and that the prediction model produced should not be skewed. The minimum water potential measured, -5.8 MPa, in the calibration samples was from a tilled surface (0-3 cm) of a loamy-textured field. The maximum water potentials, 0 MPa, occurred in deep horizons in several of the irrigated fields. The samples represented a range of clay mineralogy. The minerals found in the clay fractions were: smectite, kaolinite, mica, quartz, calcite, and feldspar (Table 2.3). The

Table 2.2. Clay content and water potential summary statistics for calibration and validation datasets.

	N	min.	max.	mean	SD†	median
calibration samples						
Total clay (g kg^{-1})	188	12	525	255	139	262
Fine clay (g kg^{-1})	188	3	380	138	85	140
Water potential (MPa)	188	-5.8	0.0	-0.49	0.60	-0.35
validation samples						
Total clay (g kg^{-1})	82	28	578	271	144	247
Fine clay (g kg^{-1})	82	18	362	152	79	147
Water potential (MPa)	82	-2.2	0.0	-0.48	0.47	-0.38

† standard deviation

Table 2.3. Clay minerals present (>5% of clay fraction) in soil cores from each field. Mineralogy is presented as a summary by field and not all cores are represented. Cores selected for mineralogy were chosen to represent clay minerals within a field.

	Clay Minerals					
Field no.	Vermiculite	Smectite	Kaolinite	Mica	Calcite	Quartz	Feldspar
1		XXX†	X§		XX‡	X§	
2		XX‡	X§	X§		X§	
3	t¶	X§	X§	X§		X§	t¶
4		X§	X§	X§		X§	
5		X§	X§	X§	X§	X§	
6		X§	X§	t¶	X§	X§	

† greater than 50% of mineral present
‡ 20-50% of mineral present
§ 5-20% of mineral present
¶ trace amounts of mineral present

soils in field 1 were notably high in clay content and were dark colored Vertisols; field 3 had very few clay-sized particles.

Model Validation

Reeves III et al. (1999) stated that the 1st derivative of the VNIR spectra generally created the best prediction model data transformation, but did not explain why the 1st derivative performed better than reflectance or the 2nd derivative. The results of this study showed that the 1st derivative of the reflectance did, in fact, produce the best model, but the reflectance model performed well also (Table 2.4). The 2nd derivative model performed the worst; therefore, we conclude that the 2nd derivative is least useful for predicting clay content in soils (Table 2.4). The 1st derivative of the soil spectra probably works well because the 1st derivative has all the information of the reflectance, but removes albedo effects associated with soil moisture or soil condition. Models to predict water potential were not very successful (Table 2.5) and using water potential to predict clay content did not improve clay content prediction (Table 2.4).

Because 1st derivative PLS models performed the best for field-moist *in situ,* and the 1st derivative worked the best in other VNIR work (Reeves et al., 1999; Reeves and McCarty, 2001; Brown et al., 2005b), the remaining PLS models reported in this paper all use the 1st derivative of the VNIR spectra between 350 and 2500 nm. Three of the four models used to predict clay content were very similar, and the fourth, field-moist *in situ* smeared, had the largest prediction error and scatter about the validation regression. Table 2.6 summarizes the prediction accuracies between the four VNIR models. The air-dried ground model was expected to produce the most accurate prediction model

Table 2.4. Results for VNIR models predicting clay content of field-moist *in situ* soils using the VNIR reflectance (R), 1st derivative (D), and 2nd derivative (DD) as the predictors.

	r^2	RMSD	Bias
		----------g kg^{-1}----------	
Clay (R)	0.78	71	8.2
Clay (D)	0.83	61	3.0
Clay (DD)	0.58	95	5.2
Clay+water pot. (D)	0.82	64	4.8

Table 2.5. Results for the log of water potential models on field-moist *in situ* scans using the VNIR reflectance (R), 1st derivative (D), and 2nd derivative (DD) as the predictors.

	r^2	RMSD	Bias
		----------MPa----------	
Water Potential (R)	0.31	1.6	0.13
Water Potential (D)	0.28	1.7	0.12
Water Potential (DD)	0.15	1.8	0.11

Table 2.6. Prediction accuracies of total clay and fine clay content models using the 1st derivative of VNIR spectra.

	r^2	RMSD	RPD	Bias
total clay		g kg^{-1}		g kg^{-1}
Air-dried ground	0.84	62	2.32	-16.3
Air-dried *in situ*	0.92	41	3.51	-2.0
Field-moist *in situ*	0.83	61	2.36	3.0
Field-moist smeared	0.75	74	1.95	7.4
fine clay				
Air-dried ground	0.81	34	2.32	0.65
Air-dried *in situ*	0.85	31	2.55	4.2
Field-moist *in situ*	0.84	32	2.47	4.4
Field-moist smeared	0.75	41	1.93	11

because the sample was reduced to a uniform size, moisture was uniform, and grinding homogenized the soil, but the air-dried *in situ* model did slightly better (Fig. 2.2). There are two possible explanations for why the air-dried ground model did not perform as well as the air-dried *in situ* model. One explanation could be just random error. The second possible explanation is that *in situ* soil has a higher bulk density than dried ground soil; and therefore the *in situ* soil may have a stronger reflectance signal.
The air-dried ground prediction had a greater bias (-16.3), than all other predictions. The large bias associated with the air-dried ground prediction was attributed to 60% of the samples being under predicted in clay content and large discrepancies between measured and predicted clay content of soils greater than 450 g kg^{-1} clay. Because the air-dried *in situ* model slightly outperformed the air-dried ground model using the 1^{st} derivative of VNIR reflectance spectra, we conclude natural soil heterogeneity (clay films etc…) and varying particle size associated with aggregation had little effect on model predictions. Gaffey (1986) found the same results when looking at different carbonate sizes. His results showed a difference in reflectance but the number of bands, band positions and width, and relative band intensities were unchanged.

Another significant finding was that the field-moist *in situ* model slightly decreased the prediction accuracy from the air-dried *in situ* model but did just as well when compared to the air-dried ground model (Fig. 2.2). These results indicate that the amount of water in the soil sample did not change the prediction accuracy compared to air-dried laboratory situations. The soil VNIR literature discusses the affect of water and aggregation on hindering prediction accuracy. Soil water may reduce accuracy

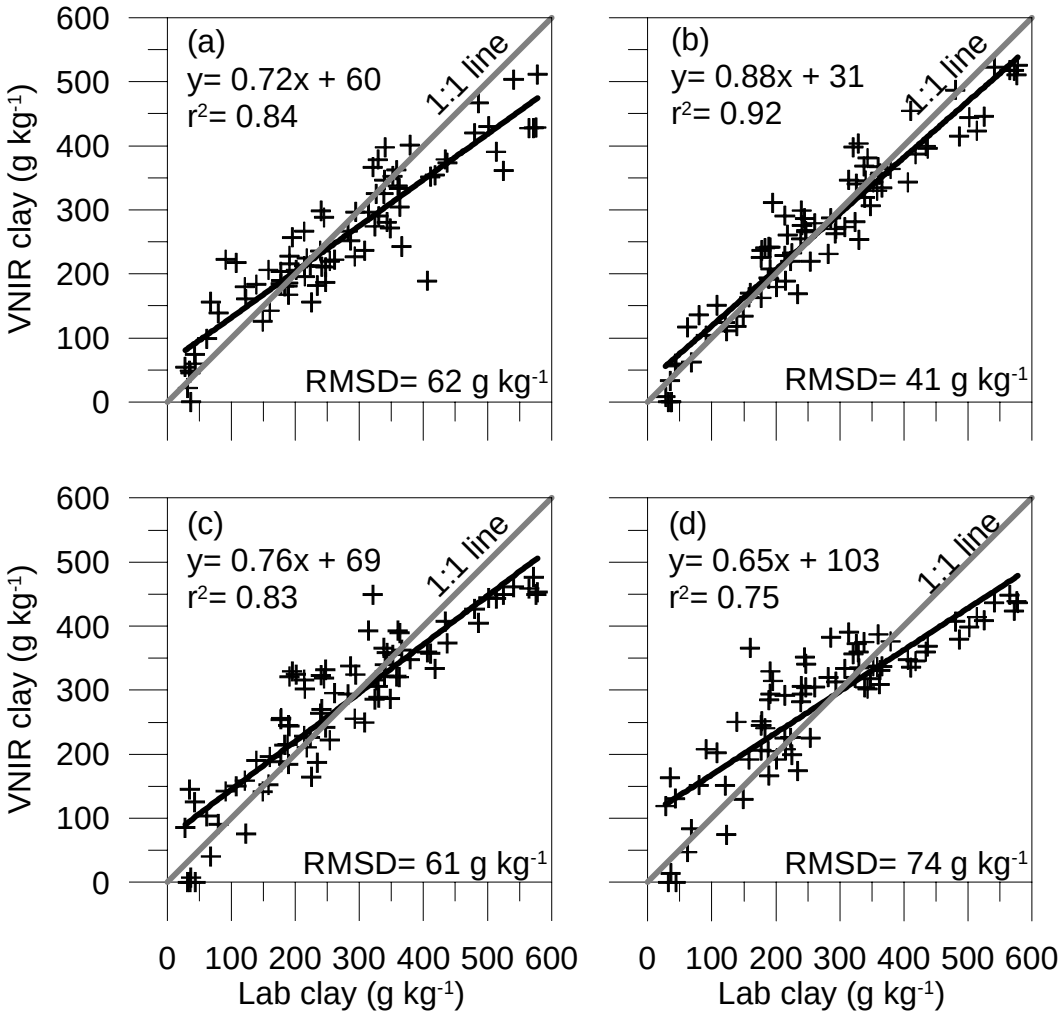

Fig. 2.2. Predicted vs. measured clay content of the validation data set for (a) air-dried ground (b) air-dried *in situ* (c) field-moist *in situ* (d) field-moist *in situ* smeared. Clay content predictions were executed from models built with PLS using the 1st derivative of VNIR reflectance spectra.

somewhat (increased RMSD by 20 g kg^{-1}), but this reduction in accuracy was no more than air-drying, grinding, and scanning the soil through a borosilicate dish (Fig. 2.2). Chang et al. (2005) showed that r^2-values decreased slightly from 0.79 to 0.76 from an air-dried soil to a moist soil, respectively, which agrees with other findings. Scanning soils *in situ* with VNIR-DRS allows soil scientist to predict clay content in a small area (<7-cm^2) within a profile (Fig. 2.3).

Smearing of the field-moist *in situ* cores reduced the accuracy of the prediction model (Fig. 2.2). The decrease in the prediction accuracy of the smeared cores was probably due to the change in reflectance properties of the soil. Smearing causes the soil surface to become shiny, which changes the reflectance from diffuse to diffuse specular reflectance. Diffuse specular reflectance has a specular component which carries less information than diffuse reflectance alone (Coates, 1998). When comparing the significant wavelengths between the field-moist *in situ* and field-moist *in situ* smeared cores, the number of wavelengths found in the visible portion of the spectra are absent from the smeared core. The absence of these wavelengths may be the source of loss in prediction accuracy. The smearing test was performed to quantify any reduction in accuracy that might occur if a soil probe, pushed into the ground smeared the sides of the hole. In this laboratory exercise, smearing moist, high clay content soil required multiple passes with the spatula. Therefore, this test was a worst-case scenario and the soil was smeared to its maximum capacity. In a field situation, the loss of prediction accuracy due to smearing might be less.

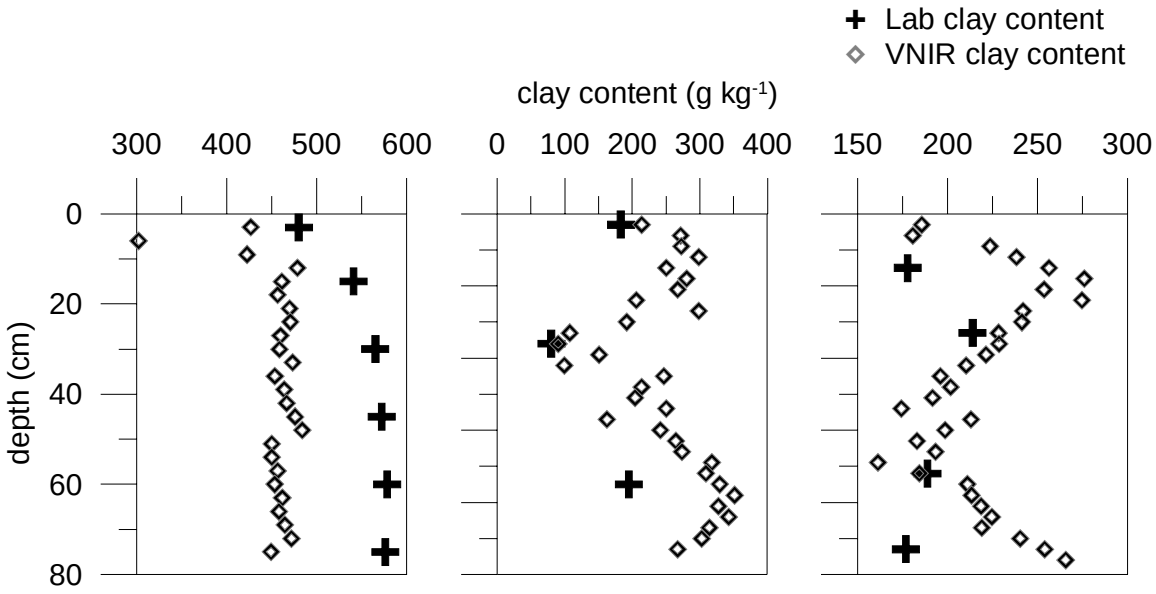

Fig. 2.3. Laboratory measured and visible near-infrared (VNIR) predicted clay content for three soil cores with depth. Laboratory measurements were taken to represent soil horizons, and VNIR predictions were taken every 3-cm depth.

The same set of VNIR models was created to predict fine clay. The amount of fine clay in a soil is of interest because fine clay is associated with a soil's shrink-swell potential (Reid-Soukup and Ulery, 2002). Fine clay models have lower RMSD values than total clay models but the RPD values are similar to the total clay models (Table 2.6). The similar RPD values indicate that the smaller RMSD values for fine clay models was due to smaller standard deviations (smaller range); so the total clay and fine clay models performed similarly.

Whole-field holdouts showed that field 1 and 3 impacted the clay content field-moist *in situ* RMSD of the validation samples the most (Table 2.7). Results of field 1 and 3 whole-field holdouts showed the greatest decrease in prediction accuracies. The highest RMSD was accompanied by a bias of -128.1 g kg^{-1} (Table 2.7). The RMSD whole-field holdouts were increased for each field but only two were of any great significance, field 1 and 3. Both of these field's samples were at the extremes for clay content. Field 1 consisted of high clay soils that were dark in color through the entire length of the soil core, which was not represented in the calibration model. Field 3 was also not represented by other fields in the calibration model since this field contained the soil cores low in clay content. We conclude from these results that the calibration model should have samples that resemble the soils to be scanned with VNIR-DRS.

The lack of correlation of VNIR spectra to water potential was not expected (Table 2.5). Even though these results are negative, two other studies (Slaughter et al., 2001; Islam et al., 2003) have shown good correlations between gravimetric water and VNIR reflectance. Water potential is composed of three separate potentials:

Table 2.7. Prediction accuracies of clay content using whole-field holdouts. Models were created using PLS with five fields and then validated using the sixth field. Six separate models were calibrated and validated.

Validation field	RMSD	Bias
	----------g kg^{-1}----------	
1	143	-128.1
2	64	4.6
3	97	9.2
4	89	44.0
5	72	31.7
6	68	30.0

gravitational, matric, and osmotic. Each potential contributes to the total water potential of the soil with matric having the greatest influence. Matric potential varies with water content and soil texture; so VNIR spectroscopy might not be able to correlate these properties together to give reliable water potential measurements.

Significant Wavelengths

Assigning peaks to chemical bonds or certain minerals in the VNIR spectroscopy can be done when working with pure minerals or pure mixtures, but when working with a natural soil mixture, it is difficult. This difficulty arises from overlapping of peaks and shifts that can occur by multiple interactions of bonds within the soil. Some significant wavelengths in the clay prediction models built with the *in situ* and ground soil scans corresponded to wavelengths which have been associated with absorption bands of pure clay minerals (Table 2.8). In particular, most significant wavelengths were found in the NIR range of the spectrum with the most significant being at 1400 nm, 2200 nm, and 2300 nm. Wavelengths in the visible region corresponded to iron oxides near 500 nm, 650 nm, and 810 nm. The NIR range had the greatest effect on the prediction accuracy for clay content.

The significant wavelengths in each model were plotted to help determine what portions of the spectra were important for clay content predictions. Figure 2.4 shows the significant wavelengths for the field-moist *in situ*, air-dried *in situ*, and air-dried ground models. Each model had a slightly different number of significant wavelengths. The air-dried *in situ* model had 78 significant wavelengths and the field-moist *in situ* and air-dried ground models had 102 and 83 significant wavelengths, respectively. Most

Table 2.8. Clay minerals with their corresponding absorptions.

Minerals	Wavelength			References
	----------nm----------			
Smectite	1400	1900	2200	Hunt and Salisbury, 1970; Goetz et al., 2001
Mica	1400	2200-2600		Hunt and Salisbury, 1970; Clark et al., 1990
Kaolinite	1400	2200		Hunt and Salisbury, 1970
Hematite	450	630	860	Gaffey et al., 1993
Goethite	493	660	940	Scheinost and Schwertmann, 1997

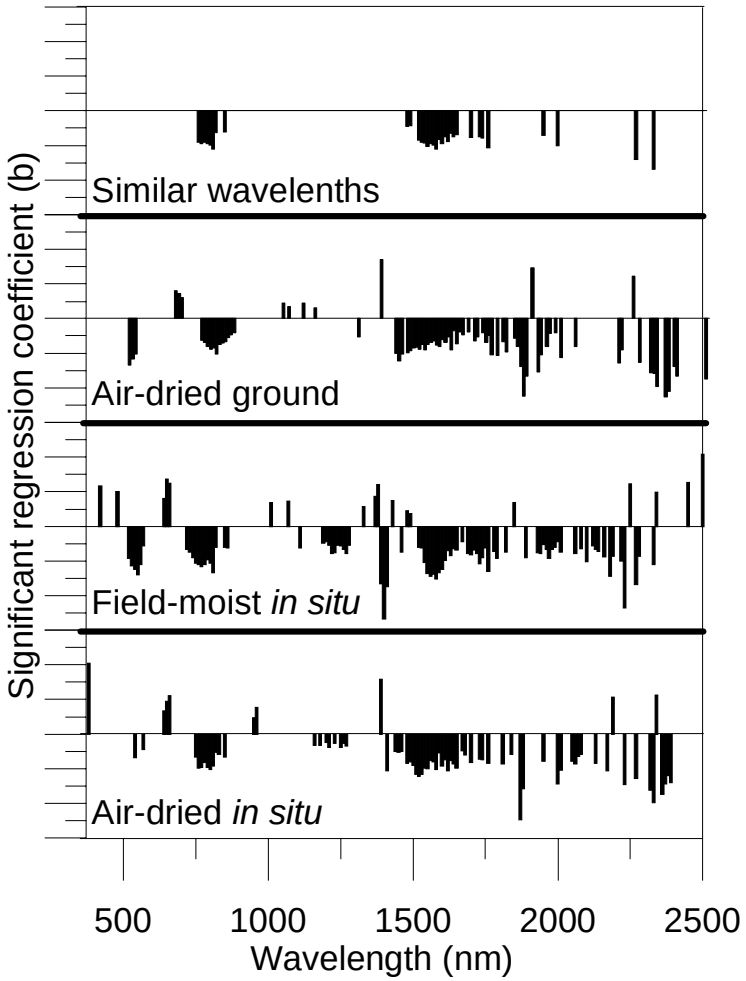

Fig. 2.4. The wavelengths that contributed to significant regression coefficients in the prediction of clay content are shown for field-moist and air-dry *in situ* and air-dry ground models. The relative magnitude of each regression coefficient indicates the strength of the correlation. The top plot shows the mean of the regression coefficients common in all three models. All plots are on the same x-axis. Values of the y-axis are not shown, but all y-axes are on the same scale.

importantly, Fig. 2.4 shows the 32 significant wavelengths that were common in all three models. Scans between the field-moist *in situ*, unsmeared and smeared, cores did have some differences in significant wavelengths (Fig. 2.5). The smeared core model had 70 significant wavelengths and the unsmeared core model had 102, with 50 of these wavelengths being common in both models. The smeared core model had fewer significant wavelengths in the visible range of the spectrum, which may have lead to an increase in the RMSD.

Conclusions

In this study of 72 soil cores from Central Texas, there was a strong relationship between VNIR reflectance and clay content for the air-dried *in situ* (RMSD= 41 g kg^{-1}), field-moist *in situ* (RMSD= 61 g kg^{-1}), and air-dried ground (RMSD= 62 g kg^{-1}) scans. Visible near-infrared DRS was capable of predicting soil clay content *in-situ* at varying water contents. There was a slight decrease in predictive power for the smeared field-moist *in situ* (RMSD= 74 g kg^{-1}) scans. Soil aggregates did not appear to change the prediction accuracies of VNIR reflectance when compared to air-dried ground samples. Variable water contents did reduce prediction accuracy evidenced by comparing air-dried *in situ* and field-moist *in situ* models, but the field-moist *in situ* prediction accuracies were the same as the air-dried ground predictions. Total clay models and fine clay models performed similarly even though the RMSD values were lower for the fine clay models. Water potential was found to have a poor relationship with VNIR reflectance in this study. Soil water should be analyzed as gravimetric or volumetric water content rather than water potential.

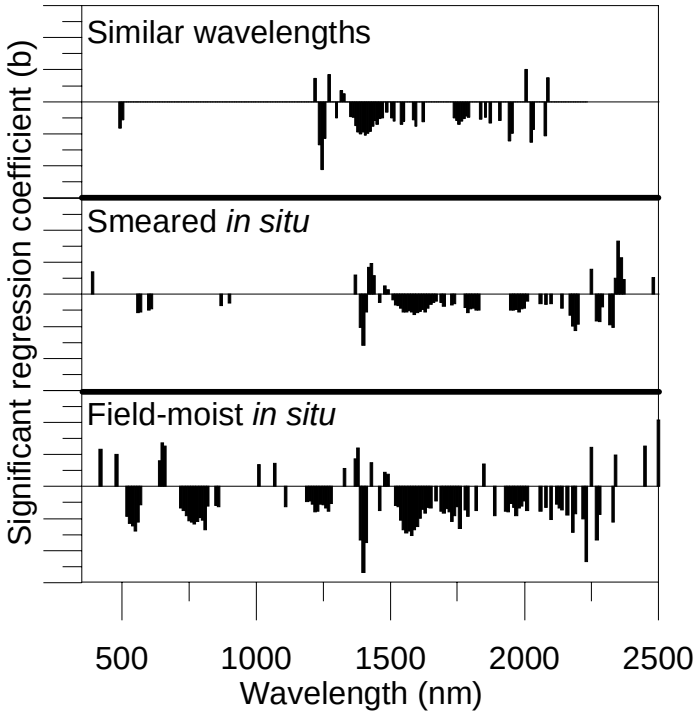

Fig. 2.5. The wavelengths that contributed to significant regression coefficients in the prediction of clay content are shown for field-moist, smeared and unsmeared, models. The relative magnitude of each regression coefficient indicates the strength of the correlation. The top plot shows the, mean of the regression coefficients common in both models.. All plots are on the same x-axis. Values of the y-axis are not shown, but all y-axes are on the same scale.

These results indicate that VNIR-DRS may be useful as a proximal soil sensor in field conditions. Visible near-infrared DRS can be beneficial by giving the soil scientist the ability to map soils across the landscape at higher resolutions, and by providing a quick method for quantifying soil properties such as clay content. Additionally, VNIR-DRS provides the ability to show how soils slowly change across a landscape or continuously within a profile (3-cm). Continued research to study *in situ* scans of larger geographical ranges and other soil properties is needed. Also, data analysis methods that help the user interpret reflectance information are needed.

CHAPTER III

IN SITU CHARACTERIZATION OF SOIL ORGANIC AND INORGANIC CARBON

Synopsis

Diffuse reflectance spectroscopy (DRS) is a rapid proximal sensing method that is being developed in laboratory settings to measure soil properties. Diffuse reflectance spectroscopy research has been completed in laboratories with promising results, but very little has been reported on how DRS will work in a field setting. Seventy-two soil cores, representing 21 soil series and four parent materials, were excavated from six fields in Central Texas. Each soil core was scanned with a visible near-infrared (VNIR) spectrometer with a spectral range of 350-2500 nm in four combinations of moisture content and pre-treatment, including field-moist *in situ*, air-dried *in situ*, field-moist smeared *in situ*, and air-dried ground. Visible near-infrared spectra were then used to predict soil organic and inorganic C using partial least squares (PLS) regression. The PLS model was validated with 30% of the original soil cores that were randomly selected and withheld from the calibration model. The organic C validation data set had a root mean squared deviation (RMSD) of 5.8 g kg^{-1} and 4.6 g kg^{-1} for the field-moist and air-dried *in situ* cores, respectively. The RMSD values for inorganic C were 10.1 g kg^{-1} and 8.3 g kg^{-1} for the field-moist and air-dried *in situ* scans, respectively. Smearing the samples had minimal affects on prediction accuracies for organic and inorganic C. Soil moisture did reduce prediction accuracies. These results show that DRS could be an

acceptable technique to measure organic and inorganic C *in situ* at varying water contents and among different parent materials.

Introduction

Quantifying organic C and inorganic C content in the soil is of great interest because of carbon sequestration studies. A method to rapidly quantify soil samples for organic C is needed because many carbon sequestration studies need to quantify soil organic C over large areas and through time as management changes (Bronson et at., 2004; Russell et al., 2005; Wright and Hons, 2005). Visible near-infrared diffuse reflectance spectroscopy (VNIR-DRS) might be a possible tool to use in such studies because the scanning is rapid, costs are fixed, and it has proven useful in plant analysis (Batten, 1998). Calcite abundance in soil is important because of its effects on plant growth and land management decisions. Soils with carbonates have alkaline pH values and may inhibit plant growth if high carbonate amounts are at the surface. Calcite has the ability to reduce availability of phosphorus; therefore, in areas with manure applications, the potential exists for some phosphorus to bond with the calcite reducing phosphorus loss to surface water (Doner and Grossl, 2002). Development of a rapid and reliable tool for quantifying organic C and inorganic C in the field could be very useful to precision agriculture, carbon sequestration studies, precision management, and other applications requiring spatially- or temporally-intensive data collection.

Soil Properties in VNIR Spectrum

Organic C amounts are very low in most mineral soils. However, the amount of organic C in a soil is important because organic C has a strong influence on soil

chemical and physical properties. Increases of organic C in the soil can increase water movement, increase water holding capacity, increase cation exchange capacity, and reduce bulk density. The active bonds in organic matter in the visible near-infrared (VNIR) region are O-H, C-N, N-H, and C=O groups (Malley et al., 2002). Henderson et al. (1992) reported wavelengths important to organic C predictions in three categories, wavelengths related to organic C content, wavelengths related to other soil properties and organic C content, and wavelengths related to other soil properties that mask organic C content. Models created from VNIR-DRS spectra have predicted organic C from air-dried, ground soil samples with r^2-values ranging from 0.49 to 0.96 and RMSD's ranging from 1.1 g kg^{-1} to 12.7 g kg^{-1} (Den-Dor and Banin, 1995; Janik et al., 1998; Reeves III et al., 1999; Confalonieri et al., 2001; Reeves III and McCarty, 2001; Chang and Laird, 2002; Dunn et al., 2002; McCarty et al., 2002; Shepherd and Walsh, 2002; Islam et al., 2003; Lee et al., 2003; Brown et al., 2005b; Chang et al., 2005). Sudduth and Hummel's (1993) research predicted organic C of the soil surface *in-situ* with r^2-values of 0.85 and 0.89 on four test sites in Illinois.

Soil inorganic C consists of small fractions of C in numerous minerals, but the two most significant minerals are calcite ($CaCO_3$) and dolomite [$CaMg(CO_3)_2$]. The observed absorption bands for calcite are due to the planar CO_3^{-2} ion. The strongest absorption bands occur in the mid-infrared region at 1063 cm^{-1} (9.407 μm), 879 cm^{-1} (11.4 μm), 1415 cm^{-1} (7.067 μm), and 680 cm^{-1} (14.7 μm), but there are several overtone and combination bands in the near-infrared (NIR) region (Clark et al., 1990). The bands at 2500 to 2550 nm and 2300 to 2350 nm are the two strongest with weaker bands near

2120 to 2160 nm, 1970 to 2000 nm, and 1850 to 1870 nm (Clark et al., 1990; Hunt and Salisbury, 1971). Inorganic C has been measured in the laboratory on air-dried, ground soil samples with r^2-values ranging from 0.69 to 0.96 and RMSD's ranging from 1.4 g kg^{-1} to 7.1 g kg^{-1} (Ben-Dor and Banin, 1995; Chang and Laird, 2002; McCarty et al., 2002, Brown et al., 2005b; Chang et al., 2005). Gaffey (1986) looked at different particle sizes of carbonates and did find a difference in reflectance but the number of bands, band positions and widths, and relative band intensities were unchanged.

Most spectroscopy work has focused on air-dried, ground soil samples (Slaughter et al., 2001; Brown et al., 2005b, Sorensen and Dalsgaard, 2005), but Sudduth and Hummel (1993) tested a portable NIR spectrometer to measure soil properties *in-situ*. Air-drying the sample reduces the intensity of bands that are related to water; therefore, the signals associated with other soil properties are not masked or hidden. Additionally, the particle size affects accuracy of the spectral scan. Smaller particles increase reflectance scatter which reduces the absorption peak height (Workman and Shenk, 2004). The most popular sample preparation has been grinding the soil into a fine powder (< 0.6 mm) for VNIR scanning (Ben-Dor and Banin 1995; Reeves III et al., 1999; Confalonieri et al., 2001; Reeves III and McCarty, 2001; Dunn et al., 2002; Lee et al., 2003).

In the field, diffuse reflectance spectroscopy (DRS) measurements may be more problematic as wide ranges in water contents, particle size and aggregation, smearing of soil, local and field heterogeneity, and regionality of the calibration model exist. To date, studies that address these possible limitations are rare. Chang et al. (2005) showed

that small increases in moisture content can considerably change the reflectance baseline and increase the peak intensities at 1400 nm and 1900 nm. Chang et al. (2005) and Slaughter et al. (2001) achieved good predictions of water content in similar soils over a large range in water contents. Few studies have been conducted on unground or minimally ground soil samples. Slaughter et al. (2001) showed that unground samples increased the overall absorbance of the soil. Sudduth and Hummel (1993) used DRS to measure organic C of the soil surface in the field. They had problems with local buried residue and soil surface roughness in wetter soils reducing prediction accuracies.

If the spectrometer's optical sensor were lowered down a hole in the soil created by a probe, the soil on the sides may be smeared. Smearing causes the reflectance properties to change which may alter prediction accuracies. Sudduth and Hummel (1993) suggested that their organic C predictions were less then optimal because of changes in soil roughness.

Researchers have worked to determine how the regionality of soils affects prediction accuracies. Ben-Dor and Banin (1995) used soil reflectance to predict organic matter with a range of 0.09% to 13.23% which resulted in poor predictions when compared to Dalal and Henry (1986). Dalal and Henry used three similar soils with a small range of organic C. Ben-Dor and Banin speculated that similar organic matter decomposition stages led to the high r^2-values for Dalal and Henry. Sudduth and Hummel (1996) concluded that the geographic range, especially more than one temperature/moisture regime, of the soil samples affects the prediction accuracies of NIR. Those conclusions indicate that organic matter decomposition along with

regionality affect prediction accuracy when using VNIR-DRS. Brown et al. (2005a) concluded that regional models could predict organic C.

Research Objectives

The overall goal of this research is to evaluate the feasibility of VNIR-DRS for *in situ* characterization of soil profiles. Specifically this research addresses the following objectives: 1) Evaluate the precision of 350 to 2500 nm (VNIR region) soil reflectance measurements in quantifying soil organic and inorganic C of *in situ* soils at field-moist and air-dry water content, 2) Quantify any change in measurement error or prediction accuracy for *in situ*, field-moist soil with a smeared surface, and 3) Quantify how the heterogeneity within the soil and between fields changes predictions. The results of this research will quantify errors associated with using VNIR-DRS to predict soil properties *in situ* and determine the feasibility of mounting a fiber-optic cable attached to a VNIR spectrometer into a soil probe.

Materials and Methods

Soil Coring

Seventy-two soil cores were collected from six fields in Erath County (fields 1, 2, 3, and 6) and Comanche County (fields 4 and 5), Texas in May 2004 to represent soil textures from clay to sand. Each soil core was collected to a maximum depth of 105 cm or until a shallower restrictive horizon was reached, using a Giddings hydraulic soil sampler (Windsor, CO) attached to a truck. The soil cores were contained in a 6.0 cm diameter plastic sleeve that was capped on both ends. Soil cores were collected to represent the large variability in soil properties over these two counties. For example, 21 soil series

were mapped by the Natural Resources Conversation Service in these fields (Table 3.1), and the parent material of these soils included loamy alluvium, sandstone, shale, and limestone. The management practices between and within fields were variable. All fields had dairy waste applied as fertilizer, compost or irrigation water. Fields consisted of improved pasture used for grazing or hay production, or the fields were tilled and planted in hay grazer or wheat.

VNIR-DRS Scanning

An ASD "FieldSpec® Pro FR" VNIR spectroradiometer (Analytical Spectral Devices, Boulder, CO), with a spectral range of 350-2500 nm, 2 nm sampling resolution and spectral resolution of 3 nm at 700 nm and 10 nm at 1400 and 2100 nm was used. The *in situ* soil cores were scanned using a contact probe with a 3.14-cm^2 diameter viewing area and its own light source. A Spectralon® panel with 99% reflectance was used to optimize the spectrometer each day; the same panel was used as a white reference between scanning each core. The 72 cores were cut using a utility knife to slice the plastic sleeve surrounding each core; then piano wire was used to half the soil core vertically, surface to subsoil. The exposed face of one of the split core halves was smeared using a stainless steel spatula to resemble smearing by a soil probe and the other half was left unsmeared. A wire grid was used to identify two columns and 3-cm thick rows within each core half (Fig. 3.1). Each row within each column was scanned twice with the FieldSpec® Pro FR with a 90º rotation of the contact probe between scans. Both halves of each core, smeared and unsmeared, were scanned at field-moist water content. The water potential was measured on each soil horizon from each core, with a

Table 3.1. Series descriptions mapped by the Natural Resources Conservation Service within the fields which were used for VNIR-DRS organic and inorganic C predictions (USDA, 1973; USDA, 1977; USDA, 2005).

Soil series	Taxonomic name	Field number
Abilene	Fine, mixed, superactive, thermic Pachic Argiustolls	4
Altoga	Fine-silty, carbonatic, thermic Udic Haplustepts	6
Blanket	Fine, mixed, superactive, thermic Pachic Argiustolls	6
Bolar	Fine-loamy, carbonatic, thermic Udic Calciustolls	1,5,6
Bosque	Fine-loamy, mixed, superactive, thermic Cumulic Haplustolls	5
Brackett	Loamy, carbonatic, thermic, shallow Typic Haplustepts	5
Bunyan	Fine-loamy, mixed, active, nonacid, thermic Typic Ustifluvents	2
Chaney	Fine, mixed, active, thermic Oxyaquic Paleustalfs	4
Cisco	Fine-loamy, siliceous, superactive, thermic Typic Haplustalfs	4
Denton	Fine-silty, carbonatic, thermic Udic Calciustolls	1,6
Frio	Fine, smectitic, thermic Cumulic Haplustolls	1,4,5
Houston Black	Fine, smectitic, thermic Udic Haplusterts	1
Karnes	Coarse-loamy, carbonatic, thermic Typic Calciustepts	5
Lewisville	Fine-silty, mixed, thermic Udic Calciustolls	5,6
Maloterre	Loamy, carbonatic, thermic Lithic Ustorthents	1
Nimrod	Loamy, siliceous, active, thermic Aquic Arenic Paleustalfs	3,4
Pedernales	Fine, mixed, superactive, thermic Typic Paleustalfs	4
Purves	Clayey, smectitic, thermic Lithic Calciustolls	1,5,6
Selden	Fine-loamy, siliceous, active, thermic Aquic Paleustalfs	3
Venus	Fine-loamy, mixed, thermic Udic Calciustolls	5
Windthorst	Fine, mixed, active, thermic Udic Paleustalfs	2,3,6

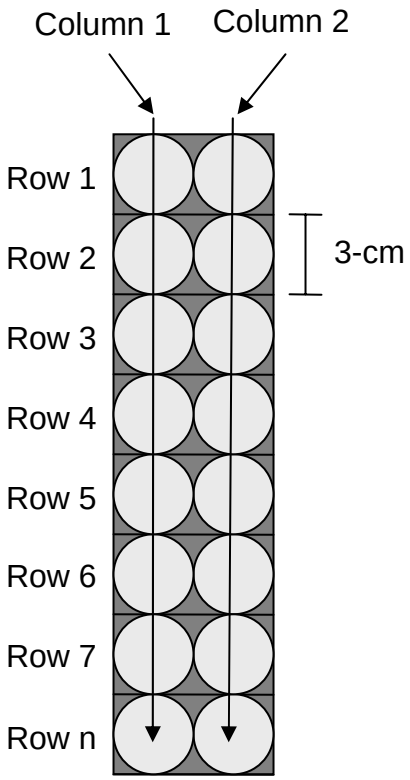

Fig. 3.1. Schematic of a vertically sliced soil core used for organic and inorganic C. Columns and rows indicate locations scanned using the contact probe.

maximum of six samples per core, using a SC-10 thermocouple psychrometer (Decagon, Pullman, WA) (Rawlins and Campbell, 1986). The cores were then placed in a drier at 44C for two days. Before rescanning, the cores were removed from the drier and left on the countertop in the laboratory to equilibrate to room temperature. The core half that was scanned in unsmeared condition was then rescanned *in situ* at air-dry moisture content.

Each row of the unsmeared soil core was ground and passed through a 2-mm sieve, and rescanned as air-dry, ground. The air-dried, ground soils were scanned with a mug lamp connected to the FieldSpec® Pro FR. The same Spectralon® 99% reflectance panel was used to calibrate the spectrometer each day, as well as being used as the white reference to set reflectance to 100%. Approximately 28 g of ground soil was placed into a borosilicate glass "puck". Each sample was scanned twice with a 90º rotation between scans.

Sampling for Laboratory Analysis

Two small 2-cm diameter by 1.5-cm deep soil samples were removed from each row from the non-smeared core half and placed in an envelope for laboratory analysis before the whole row was sampled for air-dried ground scans. Each sample was taken from the exact spot that the contact probe scanned the soil. Total C was measured using the dry combustion method (Soil Survey Staff, 1996; Nelson and Sommers, 1982) and inorganic C was measured using the modified pressure calcimeter method (Sherrod et al., 2002). Sherrod et al. (2002) reported the pressure calcimeter method detection limit was 0.42 g kg^{-1}. Organic C was calculated by subtracting the inorganic C from the total

C. Total C standard deviation of 0.31 was achieved using 191 observations of a calcium carbonate standard (Soil Survey Staff, 1996). Laboratory organic C errors would be a combination of inorganic C and total C errors.

Pretreatment of Data

Pretreatment of the spectral data included, splicing, averaging, and taking the 1st and 2nd derivative. The spectral data were spliced where the three detectors overlapped. The two scans, 0º and 90º, were averaged. The average and 1st and 2nd derivatives were taken on 10-nm intervals from 360-2490 nm after a cubic smoothing spline, implemented in the R "smooth spline" function (R Development Core Team, 2004), was fit to each raw spectral curve.

Model Building and Validation

Four different models were built to help determine how *in situ* scans and water content affect prediction accuracies. Air-dried ground, air-dried *in situ*, field-moist *in situ*, and smeared field-moist *in situ* were the four models that were created for organic and inorganic C. The models for organic and inorganic C were produced using 70% randomly chosen cores, as the calibration samples. Whole cores were used to maintain independence between the calibration and validation data (Brown et al., 2005a). The prediction models were built using 1/25th cross validation PLS method in Unscrambler 9.0 (CAMO Tech, Woodbridge, NJ). The remaining 30% of the cores were used to validate the model. Negative organic and inorganic C predictions were manually changed to zero before comparison of measured to predicted values were made. Measured versus predicted values of the validation samples were compared using simple

regression. The coefficient of determination (r^2), root mean squared deviation (RMSD), ratio of standard deviation (SD) to RMSD (RPD) and bias were calculated to compare the accuracy of different PLS models. Statistical formulas to calculate RMSD, RPD, and bias following Gauch et al. (2003), Brown et al. (2005b) and Chang et al. (2005),

$$\text{RMSD} = \sqrt{\sum_{n} \left(Y_{pred} - Y_{meas}\right)^2 / N}\ , \tag{3.1}$$

$$\text{RPD} = \text{SD} / \text{RMSD}\ , \text{and} \tag{3.2}$$

$$\text{Bias} = \sum_{n} \left(Y_{pred} - Y_{meas}\right) / N; \tag{3.3}$$

where Y_{pred} are predicted values of the validation set using the PLS model, and Y_{meas} are the laboratory measurements of the validation set, and N is the total number of samples in the validation data. Models for organic and inorganic C were considered satisfactory with $r^2 > 0.70$ and RMSD > 5.0 g kg^{-1}.

Model calibration and validation sets were also completed on whole-field holdouts to determine how field to field heterogeneity affected prediction accuracies (Brown el al., 2005a). Whole-field holdouts were achieved by calibrating a model using PLS with five of the six fields. The sixth field was held out as the validation samples. Six models were created so all six fields could be represented as a validation set. To determine how core heterogeneity affected prediction accuracies, column one scans were used to calibrate an organic and inorganic C model. The models were then cross-validated with column one and column two and RMSD's compared.

Significant wavelengths in each model were plotted to help determine what portions of the wavelength were important for organic and inorganic C predictions.

Significant wavelengths were chosen by Unscrambler 9.0 with an uncertainty limits of two standard deviations in an ideal case. Regression coefficients for each wavelength were calculated with a t-test and uncertainty limits that did not cross the zero line were significant.

Results and Discussion

Sample Descriptions

From the 72 soil cores, 540 soil samples were analyzed for organic and inorganic C. Of these samples, 376 were used in the calibration model. The remaining 164 samples were used to validate the PLS model. The organic C range of all the data was 0 g kg^{-1} to 74.4 g kg^{-1} for surface and subsurface horizons. The mean for organic C was 7.3 g kg^{-1} and 8.4 g kg^{-1} for the calibration and validation samples, respectively, indicating that the majority of the samples are lower in organic C. The means of calibration and validation data sets were statistically similar. The means were tested using the nonparametric Mann-Whitney test, SPSS (Chicago, IL), $p>0.1$ (Table 3.2). Inorganic C samples covered a wide range as well, 0 g kg^{-1} to 111.4 g kg^{-1}. Though selected randomly from all six fields, the means of the inorganic C calibration and validation data sets were significantly different as tested by the Mann-Whitney test. Because both data sets were randomly selected and independent, no assumptions for PLS were broken, but significantly different means will affect the accuracy of the calibration model in predicting the validation data values Inorganic C had low means for the calibration and validation samples, 4.9 g kg^{-1} and 2.0 g kg^{-1}, respectively, indicating the majority of samples were well below the median. Organic and inorganic C samples

Table 3.2. Organic and inorganic C statistics for calibration and validation samples along with core water potential results.

	N	min.	max.	mean	SD	median
calibration samples						
organic C (g kg^{-1})	376	0	74.4	11.0	11.7	7.3
inorganic C (g kg^{-1})	376	0	111.4	18.7	23.2	4.9
water potential (MPa)	188	-5.80	0	-0.49	0.60	-0.35
validation samples						
organic C (g kg^{-1})	164	0	40.0	9.7	8.0	8.4
Inorganic C (g kg^{-1})	164	0	82.3	10.8	17.5	2.0
water potential (MPa)	82	-2.15	0	-0.48	0.47	-0.38

were not normally distributed and no transformation was successful at reaching a normal transformation. Water potential of the field-moist scans ranged from saturated (0 MPa) to -5.8 MPa. The extremely low water potential sample came from a tilled surface (0-3 cm) of a loamy textured field. The maximum water potentials occurred in samples deeper in the profile from several irrigated fields.

Model Validation

Reeves III et al. (1999) stated that the 1st derivative of the VNIR spectra generally created the best prediction model. The field-moist *in situ* model for organic C performed the best when the 1st derivative of the reflectance was used (Table 3.3). Inorganic C models performed well using the reflectance or the 1st derivative for all soil conditions, but the 2nd derivative of reflectance did a poor job (Table 3.3). Because 1st derivative organic C PLS models performed the best for field-moist *in situ* and the 1st derivative worked the best in other VNIR work (Reeves et al., 1999; Reeves and McCarty, 2001; Brown et al., 2005b), all the remaining organic C model reported in this paper use the 1st derivative of the VNIR spectra between 350 and 2500 nm. Inorganic C models had better prediction accuracies using the reflectance versus the 1st derivative (Table 3.3). The inorganic C PLS models performed the best using the reflectance data for the field-moist *in situ* scans; therefore, all PLS models for inorganic C were created with the reflectance. The 1st derivative prediction accuracies were satisfactory; so 1st derivatives of the reflectance could be used to predict inorganic C. The 2nd derivative results were unsatisfactory and does not show promise as possible transformations for organic and inorganic C predictions (Table 3.3).

Table 3.3. Results for organic and inorganic C models on field-moist *in situ* scans using different reflectance transformations.

	r^2	RMSD	Bias
organic C		----------g kg^{-1}----------	
Reflectance	0.58	5.8	1.4
1st derivative	0.62	5.8	1.2
2nd derivative	0.43	6.6	1.1
Inorganic C			
Reflectance	0.74	9.4	2.5
1st derivative	0.73	10.1	3.3
2nd derivative	0.34	14.9	4.2

Air-dried ground and air-dried *in situ* models for organic C were satisfactory, while field-moist *in situ* models, smeared and unsmeared, were less accurate. Table 3.4 summarizes the prediction accuracies between the four calibration models. The air-dried ground model was expected to produce the best model since the sample was reduced to a uniform size, moisture was reduced, and any natural heterogeneity of the soil was homogenized. Air-dried ground scans for organic C gave the best fit line (Fig. 3.2) because the prediction accuracies of the air-dried *in situ* model was similar to the air-dried ground model, we conclude that soil heterogeneity did not greatly affect VNIR effectiveness for predicting organic C. Air-dry *in situ* cores had better prediction accuracies than field-moist *in situ*, meaning water content reduced prediction accuracy. The field-moist *in situ* smeared core scans produced the poorest fit for organic C (Fig. 3.2). Undecomposed organic matter, loose soil structure, and roots at the soil surface apparently increased the prediction error for organic C. Eleven samples in the validation data set had organic C values predicted greater than 10 g kg^{-1} from the laboratory measured organic C and nine of those samples were from the top 6 cm of the soil Variable soil parent materials also contribute to worse prediction accuracies (Henderson et al., 1992).

Inorganic C results were similar to the organic C with the air-dried ground model performing the best followed by air-dried *in situ*, and field-moist *in situ*, smeared and unsmeared, having the lowest r^2 values (Fig. 3.3). No inorganic C model performed satisfactory but the results indicated that some correlation between VNIR and soil inorganic C did exists. Table 3.4 summarizes the inorganic C results using the

Table 3.4. Prediction accuracies of organic and inorganic C content models using the reflectance or 1st derivative of VNIR-DRS scans.

	r^2	RMSD	RPD	Bias
		g kg^{-1}		g kg^{-1}
organic C (1st derivative)				
Air-dried ground	0.73	4.5	1.78	0.9
Air-dried *in situ*	0.72	4.6	1.74	1.3
Field-moist *in situ*	0.62	5.8	1.38	1.2
Field-moist smeared	0.47	5.7	1.40	1.6
inorganic C (Reflectance)				
Air-dried ground	0.88	6.6	2.30	0.85
Air-dried *in situ*	0.83	7.9	2.22	0.89
Field-moist *in situ*	0.74	9.4	1.86	2.50
Field-moist smeared	0.70	9.0	1.94	2.90
inorganic C (1st derivative)				
Air-dried ground	0.87	7.0	2.24	2.4
Air-dried *in situ*	0.82	8.3	2.11	3.8
Field-moist *in situ*	0.73	10.1	1.73	3.3
Field-moist smeared	0.64	9.9	1.77	3.4

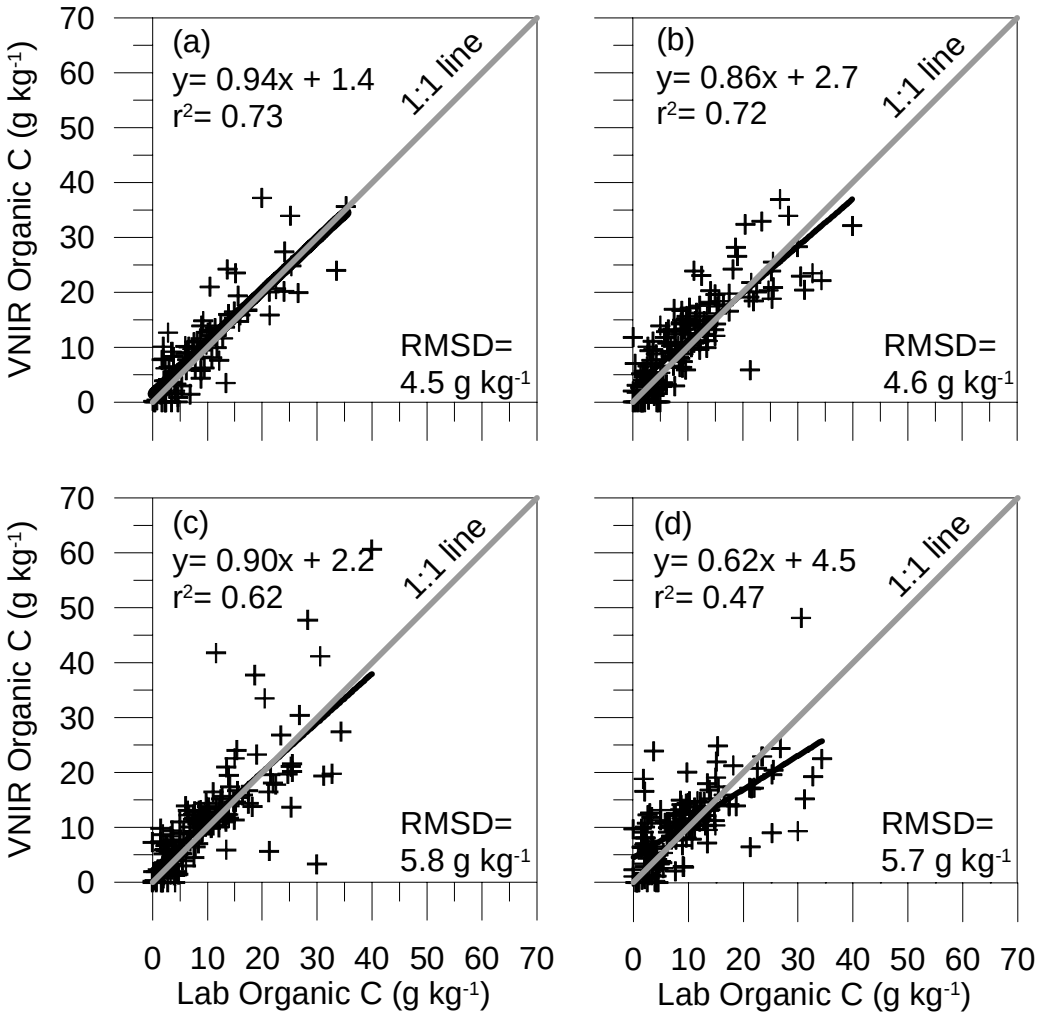

Fig. 3.2. Predicted vs. measured organic C of the validation data set for (a) air-dried ground (b) air-dried *in situ* (c) field-moist *in situ* (d) field-moist *in situ* smeared. Organic C content predictions were executed from models built with PLS using the 1st derivative of VNIR reflectance spectra.

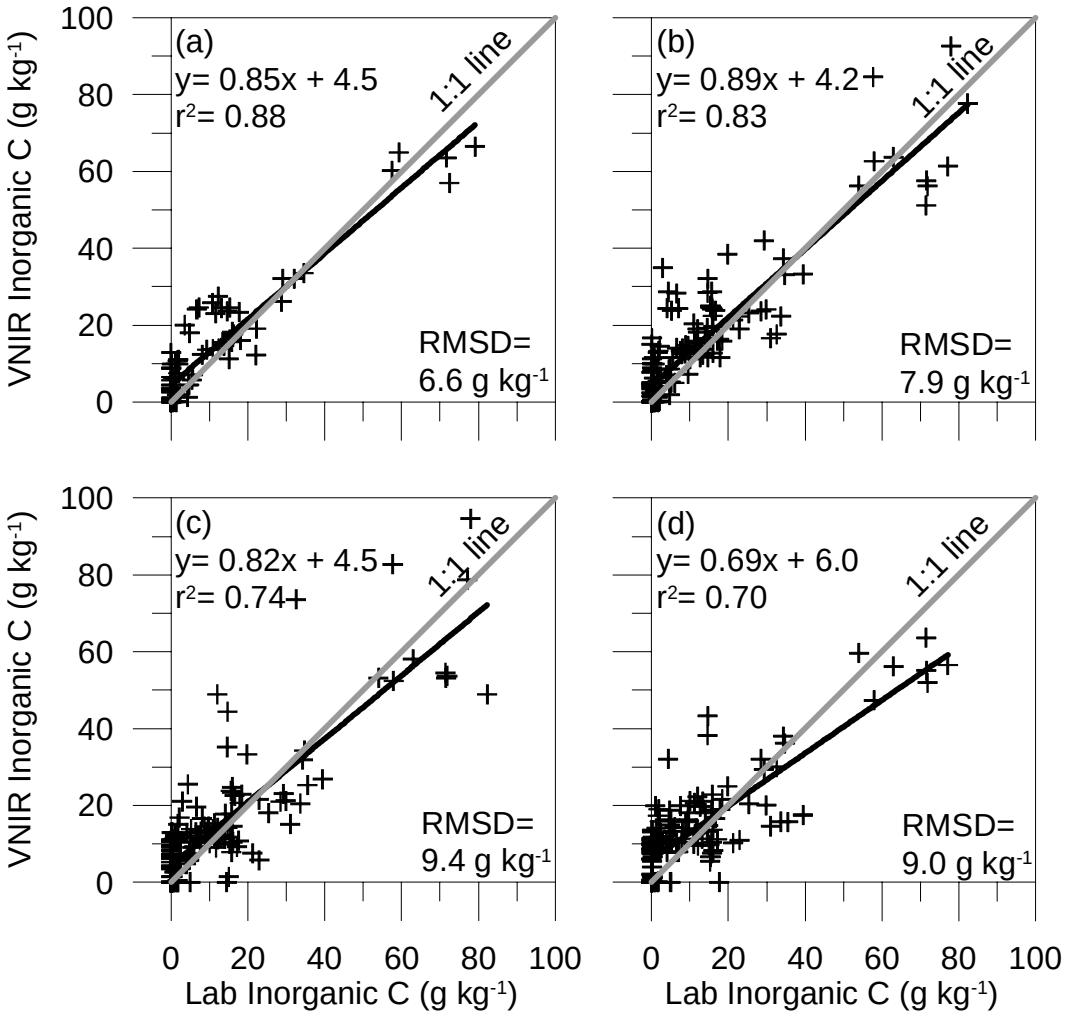

Fig. 3.3. Predicted vs. measured inorganic C of the validation data set for (a) air-dried ground (b) air-dried *in situ* (c) field-moist *in situ* (d) field-moist *in situ* smeared. Inorganic C content predictions were executed from models built with PLS using the VNIR spectral reflectance.

reflectance and 1st derivative with respect to the reflectance. Prediction accuracies of the inorganic C models were affected by not grinding and not air-drying the samples. Air-dried ground inorganic C models had the best prediction accuracy, which was expected, because of air-drying and sample homogenization. The effect of homogenizing the sample for inorganic C was more significant than with organic C because the *in situ* scans dealt with secondary $CaCO_3$ concentrations. Since air-dried *in situ* and field-moist *in situ* inorganic C models increased RMSD, water content and soil heterogeneity both affected the prediction of inorganic C.

Smearing the field-moist cores did not change prediction accuracies but did reduce the r^2 value (Table 3.4). The decrease in the r^2 values of the smeared cores was probably due to the change in reflectance properties of the soil. Smearing causes the soil surface to become shiny, which changes the reflectance from diffuse to diffuse specular reflectance. Diffuse specular reflectance has a specular component which carries less information than diffuse reflectance (Coates, 1998). The decline in prediction accuracies could also be linked to the fact that cores with high clay content smeared better than other cores, leading to greater variability. A similar phenomenon may have been experienced by Sudduth and Hummel (1996) when they reported prediction errors because some soil was rough and other soil were smooth.

When looking at the fields individually, conclusions can be drawn from which fields negatively affected prediction accuracy for inorganic C. Heterogeneity within a single scan worsened prediction accuracies for fields 1 and 5, because numerous $CaCO_3$ concentrations were present. The majority of the secondary $CaCO_3$ concentrations in

field 1 was present as small, hard nodules. In field 5, the secondary $CaCO_3$ concentrations were large, soft masses. In fields 4 and 6, $CaCO_3$ was diffused throughout the soil profile, and fields 2 and 3 were free of almost all $CaCO_3$ except for some from lime applications.

To determine how within-core heterogeneity affected prediction accuracy, one column was used to calibrate the model and the second column was used in validation. Within-core heterogeneity showed that heterogeneity within a scanning area, 2 cm in diameter, slightly affected the prediction. Column one RMSD cross-validation of organic and inorganic C was 5.4 g kg^{-1} and 9.9 g kg^{-1}, respectively. When column two scans were used to validate column one for organic C and inorganic C the RMSD values were 6.0 g kg^{-1} and 9.8 g kg^{-1}, respectively. These results indicate that when working in soils with $CaCO_3$ concentrations that are throughout a soil horizon, *in situ* scans should still be effective. Soil samples used in this comparison had a large range of $CaCO_3$ concentrations which also varied from hard to soft. Concentrations also varied in size from very small (threads and filaments) to large concentrations, encompassing the whole scanning area (2-cm diameter).

Holding out whole fields reduced prediction accuracies in cases where the soils of the field held out were not represented by another field (Table 3.5). Validation with field 1 had the highest increase in RMSD for organic and inorganic C compared to the 30 % random validation (Table 3.5). Field 1 had Vertisols, which were dark in color, high in smectitic clays, and contained hard $CaCO_3$ nodules rather than soft concentrations. Holding out the other fields from the calibration resulted in little

Table 3.5. Comparison between validation with individual-core holdouts versus validation with whole-field holdouts. The change in RMSD for organic and inorganic C are reported between the two validation sets, which was done by subtracted RMSD from the validation of whole-field holdouts from the validation of individual-core holdouts.

	RMSD	
	organic C	inorganic C
	----------g kg^{-1}----------	
Field 1	4.5	15.1
Field 2	0.9	-2.3
Field 3	0.0	0.5
Field 4	0.1	-0.2
Field 5	-0.4	2.6
Field 6	2.7	2.3

significant change in RMSD for those fields. Field 6 had an increase in the RMSD of organic C but all other predictions changed by less than 1.0 g kg^{-1} (Table 3.5).

Significant Wavelengths

Visible near-infrared models for organic C performed similar for air-dried ground and air-dried *in situ* scans, but the number of significant wavelengths in each model was different (Fig. 3.4). The dried ground model had more significant wavelengths (114) than the dried *in situ* (69), but the ratio of wavelengths in the NIR region to the visible region is about the same (Fig. 3.4). The two air dried models had 56 wavelengths in common. Field-moist *in situ* significant wavelengths numbered 110 with 43 of those wavelengths common with either the air-dried ground or air-dried *in situ* models (Fig. 3.4). The field-moist *in situ* model had a greater portion of significant wavelengths in the visible region. This may indicate that these parts of the visible region are not contributing to a more accurate prediction. The wavelengths in the visible range might be linked to the soil changing color at varying water contents, which is noted in soil surveys as a moist and dry color.

The inorganic C models built using the reflectance or 1st derivative of reflectance performed similarly and numbers of significant wavelengths in the field-moist models were similar, 94 and 89, respectively. The air-dried ground reflectance model had 51 significant wavelengths and the field-moist *in situ* and air-dried *in situ* reflectance models had 94 and 54 significant wavelengths, respectively (Fig. 3.5). Only, nine significant wavelengths were common in all three models (Fig. 3.5).

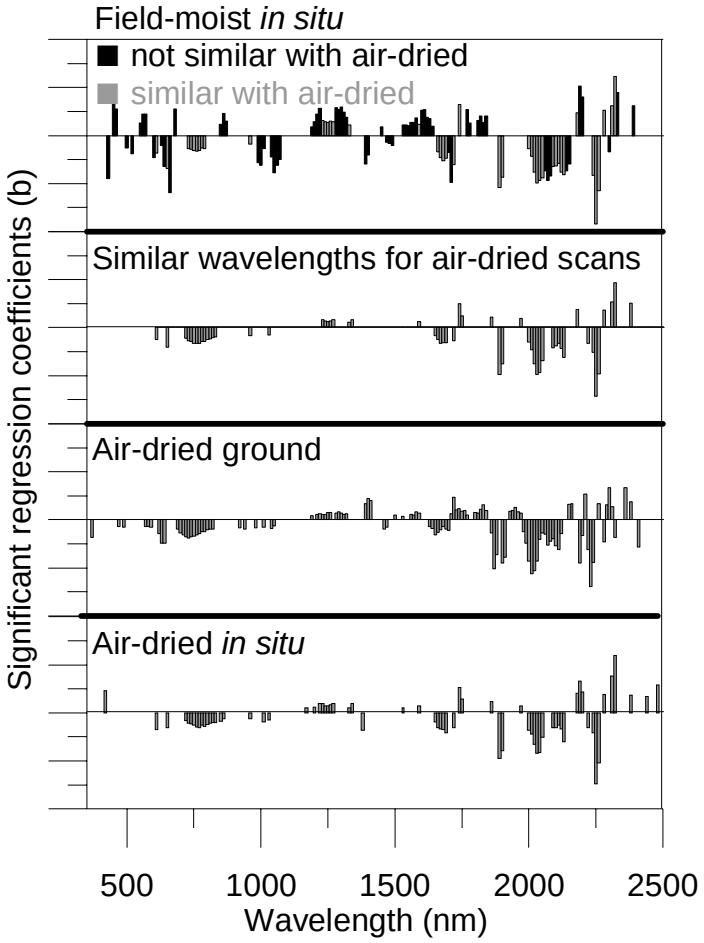

Fig. 3.4. The wavelengths that significantly contributed to the prediction of organic C are shown for field-moist and air-dry *in situ* and air-dry ground models. The third plot from the bottom shows the average of regression coefficients common in air-dried models. All plots are on the same x-axis and y-axis scale is the same. Values of the y-axis are not shown because the existence of the significant wavelength and relative values to other wavelengths within one model is what is being shown. The greater amplitude, negative or positive, means the wavelength held more significance to the model.

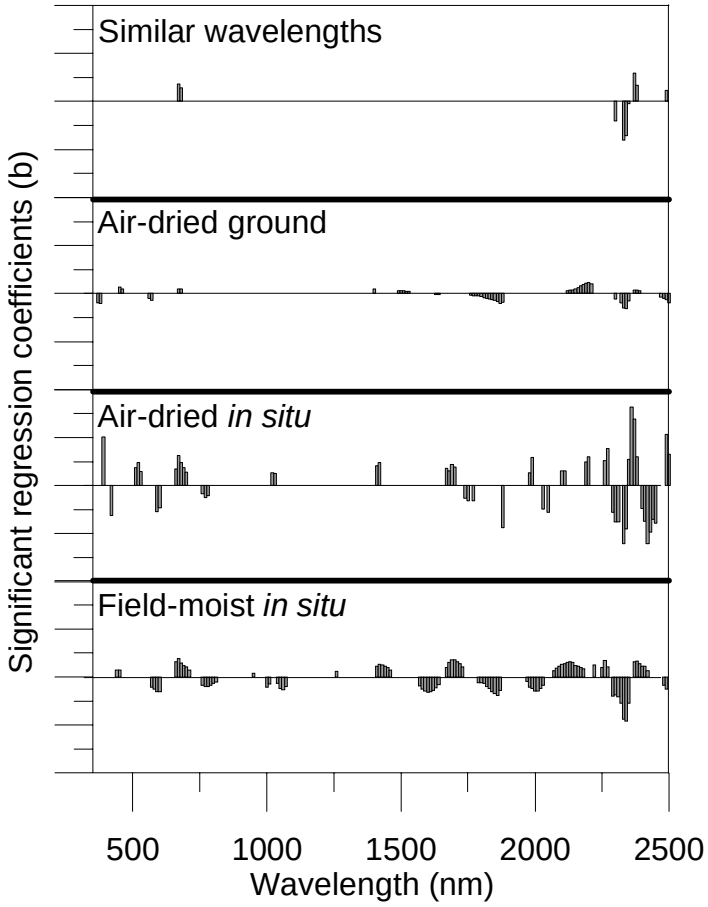

Fig. 3.5. The wavelengths that significantly contributed to the prediction of inorganic C are shown for field-moist and air-dry *in situ* and air-dry ground models. The top plot shows the average of regression coefficients common in all three models. All plots are on the same x-axis and y-axis scale is the same. Values of the y-axis are not shown because the existence of the significant wavelength and relative values to other wavelengths within one model is what is being shown. The greater amplitude, negative or positive, means the wavelength held more significance to each model.

Wavelengths which were significant for inorganic C models may explain the decrease in prediction accuracy from air-dried ground, air-dried *in situ*, to field-moist *in situ.* The air-dried ground model used several wavelengths at the longer wavelengths of the NIR region, and the air-dried *in situ* model still used numerous wavelengths in the NIR region, but there were not as many significant wavelengths at the high NIR region (Fig. 3.5). The field-moist *in situ* model had even fewer wavelengths in the NIR region and had more in the visible region (Fig. 3.5). *In situ* inorganic C predictions had worse RMSD's than air-dried ground probably due to the lack of significant wavelengths in the longer wavelengths of the NIR spectrum since the optimum range for inorganic C predictions includes the mid-infrared region.

Conclusion

In this study of 72 soil cores from Central Texas, a strong correlation existed between the 1st derivative with respect to the VNIR reflectance and organic C for the air-dried ground (RMSD= 4.5 g kg^{-1}) and air-dried *in situ* (RMSD= 4.6 g kg^{-1}) scans. A slight decrease in predictive power was observed for the field-moist unsmeared *in situ* (RMSD= 5.8 g kg^{-1}) and smeared field-moist *in situ* (RMSD= 5.7 g kg^{-1}) scans. The correlation between inorganic C and VNIR reflectance was weaker, with air-dried ground (RMSD= 6.6 g kg^{-1}) and air-dried *in situ* (RMSD= 7.9 g kg^{-1}) scans. Field-moist, smeared and unsmeared, decreased the prediction accuracies with RMSD values of 9.4 g kg^{-1} and 9.0 g kg^{-1}, respectively. Variable soil water contents did decrease the prediction accuracies of organic and inorganic C models. The soil heterogeneity did not affect the prediction accuracy for the organic C models when comparing the air-dried

ground and air-dried *in situ* but did affect the inorganic C models. When predicting organic and inorganic C, the calibration model should have soil cores that represent the unknown soil cores otherwise calibration samples should be added. Predictions will not vary greatly when working in soils with heterogeneity within the scanning area, such as redox features, secondary accumulations, and mottles.

These results indicate that precision agriculture and precision management will be able to benefit from proximal sensing using VNIR-DRS. Visible near-infrared DRS can give soil scientists the ability to map soils across the landscape at high spatial resolution for precision agriculture and management. The ability to show how soils slowly change across a landscape or within a profile (3-cm) is possible with VNIR-DRS. More research should be continued to include larger geographical ranges and other soil properties, but the future of VNIR reflectance in field measurements is promising.

CHAPTER IV

SUMMARY AND CONCLUSIONS

Visible near-infrared diffuse reflectance spectroscopy (VNIR-DRS) was capable of predicting clay content, organic C, and inorganic C *in situ*. Soil heterogeneity did not affect prediction accuracies for clay content or organic C. Inorganic C air-dried *in situ* models did decrease the prediction accuracy of VNIR-DRS. Variable moisture content of field-moist *in situ* scans did decrease prediction accuracies for clay content, organic C, and inorganic C, but clay content field-moist *in situ* scans did not hinder predictions when compared to the air-dried ground models. Smearing the field-moist *in situ* cores reduced the prediction accuracy for clay content. Organic C and inorganic C field-moist *in situ* smeared models did not show a decrease in prediction accuracies but did reduce the fit between laboratory analysis and VNIR predictions. Water potential did not produce successful prediction models nor did water potential increase predictions of clay content.

These results indicate that precision management would benefit from VNIR-DRS. Visible near-infrared diffuse reflectance spectroscopy will be beneficial by giving soil scientists the ability to rapidly and reliably map soils across the landscape at a smaller scale for the use in precision agriculture and management. The ability to show how soils slowly change across a landscape or within a profile (3-cm) is also available with VNIR-DRS. Clay content, organic C, and inorganic C models were not changed for air-dried *in situ* cores which indicates better predictions would be expected during drier periods of the year. Variable water contents negatively affected all models to some

degree. Scanning at drier water contents also decreases the possibility of smearing which increased RMSD compared to the air-dried scans. Visible near-infrared diffuse reflectance spectroscopy was capable of predictions across parent materials.

In situ analysis of soil with VNIR-DRS is in an infancy stage. More research should be continued to include the following:

1. Samples that represent larger geographical ranges;
2. Determine what geographical information (longitude, latitude, etc…) can improve predictions;
3. Determine prediction accuracies of cation exchange capacity, base saturation, pH, nitrogen, potassium, phosphorus, and other soil properties/nutrients;
4. Look at new data mining techniques to improve predictions;
5. Determine prediction accuracy of soil tilth or quality; and
6. Investigate new statistical approaches to determine significant wavelengths in the predictions.

The full capabilities of VNIR-DRS have not been exploited at this time and those capabilities need to be determined.

REFERENCES

Batten, G.D. 1998. Plant analysis using near infrared reflectance spectroscopy: the potential and the limitations. Aust. J. Exp. Agric. 38:697-706.

Ben-Dor, E. and A. Banin. 1995. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. Soil Sci. Soc. Am. J. 59:364-372.

Ben-Dor, E., J.R. Irons, and G.F. Epema. 1999. Spectroscopy of rocks and minerals, and principles of spectroscopy. p. 111-188. *In* A.N. Rencz (ed.) Remote sensing for the earth sciences: Manual of remote sensing. John Wiley & Sons, New York.

Bronson, K.F., T.M. Zobeck, T.T. Chua, V. Acosta-Martinez, R.S. van Pelt, and J.D. Booker. 2004. Carbon and nitrogen pools of Southern High Plains cropland and grassland soils. Soil Sci. Soc. Am. J. 68:1695-1704.

Brown D.J., R.S. Bricklemeyer, and P.R. Miller. 2005a. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. Geoderma 129:251-267.

Brown, D.J., K.D. Shepherd, M.G. Walsh, M.D. Mays, and T.G. Reinsch. 2005b. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma (Available on-line at http://www.sciencedirect.com).

Chang, C.W. and D.A. Laird. 2002. Near-infrared reflectance spectroscopic analysis of soil C and N. Soil Sci. 167:110-116.

Chang, C.W., D.A. Laird, and C.R. Hurburgh, Jr. 2005. Influence of soil moisture on near-infrared reflectance spectroscopic measurement of soil properties. Soil Sci. 170:244-255.

Clark, R.N. 1999. Spectroscopy of rocks and minerals, and principles of spectroscopy. p. 3-52. *In* A.N. Rencz (ed.) Remote sensing for the earth sciences: Manual of remote sensing. John Wiley & Sons, New York.

Clark, R.N., T.V.V. King, M. Klejwa, and G.A. Swaze. 1990. High spectral resolution reflectance spectroscopy of minerals. J. Geophys. Res. 95:12653-12680.

Coates, J. 1998. A review of sampling methods for infrared spectroscopy. p. 49-91. *In* J. Workman and A.W. Springsteen (ed.) Applied spectroscopy, a compact reference for practioners. Academic Press, San Diego, CA.

Confalonieri, M., F. Fornasier, A. Ursino, F. Boccardi, B. Pintus, and M. Odoardi. 2001. The potential of near infrared reflectance spectroscopy as a tool for the chemical characterization of agricultural soils. J. Near Infrared Spectrosc. 9:123-131.

Dala, R.C. and R.J. Henry. 1986. Simultaneous determinations of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry. Soil Sci. Soc. Am. J. 50:120-123.

Doner, H.E. and P.R. Grossl. 2002. Carbonates and evaporites. p. 199-228. *In* J.B. Dixon and D.G. Schulze (ed.) Soil mineralogy with environmental applications. ASA, CSSA, and SSSA, Madison, WI.

Dunn, B .W., H. G. Beecher, G. D. Batten, and S. Ciavarella. 2002. The potential of near-infrared reflectance spectroscopy for soil analysis – a case study from the Riverine Plain of south-eastern Australia. Aust. J. Exp. Agric. 42:607-614.

Ellert, B.H., H.H. Janzen, and T. Entz. 2002. Assessment of a method to measure temporal change in soil carbon storage. Soil Sci. Soc. Am. J. 66:1687-1695.

Gaffey, S.J. 1986. Spectral reflectance of calcite minerals in the visible and near infrared (0.35-2.55 microns): Calcite, aragonite, and dolomite. Am. Mineral. 71:151-162.

Gaffey, S.J., L.A. McFadden, D. Nash, and C.M. Pieters. 1993. Ultraviolet, visible, and near-infrared reflectance spectroscopy: Laboratory spectra of geologic materials. p. 43-73. *In* C.M. Pieters and P.A.J. Englert (ed.) Remote geochemical analysis elemental and mineralogical composition. University of Cambridge Press, Cambridge.

Gauch, H.G., Jr., J.T.G. Hwang, and G.W. Fick. 2003. Model evaluation by comparison of model-based predictions and measured values. Agron. J. 95:1442-1446.

Gee, G.W. and D. Or. 2002. Particle-size analysis. p. 255-293 *In* J.H. Dane and G.C. Topp (ed.) Methods of soil analysis. Part 4. ASA, CSSA, and SSSA, Madison WI.

Geladi, P. and B.R. Kowalski. 1986. Partial least-squares regression: A tutorial. Anal. Chim. Acta. 185:1-17.

Goetz, A.F., S. Chabrillat, and Z. Lu. 2001. Field reflectance spectroscopy for detection of swelling clays at construction sites. Field Anal. Chem. Technol. 5:143-155.

Hallmark, C.T., L.T. West, L.P. Wilding, and L.R. Drees. 1986. Characterization data for selected Texas soils. Texas Agricultural Experiment Station. College Station, TX.

Henderson, T.L., M.F. Baumgardner, D.P. Franzmeier, D.E. Stott, and D.C. Coster. 1992. High dimensional reflectance analysis of soil organic matter. Soil Sci. Soc. Am. J. 56:865-872.

Hunt, G.R. and J.W. Salisbury. 1970. Visible and near-infrared spectra of minerals and rocks: I. Silicate minerals. Modern Geol. 1:283-300.

Hunt, G.R. and J.W. Salisbury. 1971. Visible and near-infrared spectra of minerals and rocks: II. Carbonates. Modern Geol. 2:23-30.

Islam, K., B. Singh, and A. McBratney. 2003. Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. Aust. J. Soil Res. 41:1101-1114.

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

Kilmer, V. H., and L. Z. Alexander. 1949. Methods for making mechanical analyses of soil. Soil Sci. 68:15-24.

Lee, S.W., J. F. Sanchez, R.S. Mylavarapu, and J. S. Choe. 2003. Estimating chemical properties of Florida soils using spectral reflectance. Trans. ASAE. 46:1443-1453.

Lobell, D.B. and G.P. Asner. 2002. Moisture effects on soil reflectance. Soil Sci. Soc. Am. J. 66:722-727.

Malley, D.F., L. Yesmin, and R.G. Eilers. 2002. Rapid analysis of hog manure and manure-amended soils using near-infrared spectroscopy. Soil Sci. Soc. Am. J. 66:1677-1686.

Malley D.F., P.D. Martin, and E. Ben-Dor. 2004. Application in analysis of soils. p. 729-784. *In* C.A. Roberts, J. Workman Jr., and J.B. Reeves III (ed.) Near-infrared spectroscopy in agriculture. ASA, CSSA, and SSSA, Madison, WI.

McCarty, G.W., L.B. Reeves, V.B. Reeves, R.F. Follett, and J.M. Kimble. 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil C measurement. Soil Sci. Soc. Am. J. 66:640-646.

Moore I.D., P.E. Gessler, G.A. Nielsen, and G.A. Peterson. 1993. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 57:443-452.

Morgan, C.L.S., J.M. Norman, C.C. Molling, K. McSweeney, and B. Lowery. 2003. Evaluating soil data from several sources using a landscape model. p. 243-260. *In* Y. Pachepsky, D.E. Radcliffe, and H.M. Selim (ed.) Scaling methods in soil physics. CRC Press, Boca Raton, FL.

Mustard, J.F. and J.M. Sunshine. 1999. Spectral analysis for earth science: Investigations using remote sensing data. p. 251-306. *In* A.N. Rencz (ed.) Remote sensing for the Earth sciences: Manual of remote sensing. John Wiley & Sons, New York.

Nelson, D.W. and L.E. Sommers. 1982. Total C, organic C and organic matter. *In* A.L. Page (ed.) Method of soil analysis. Part II. (2nd edition) Agronomy 9:539-580.

Pachepsky, Ya.A., D.J. Timlin, and W.J. Rawls. 2001. Soil water retention as related to topographic variables. Soil Sci. Soc. Am. J. 65:1787-1795.

R Development Core Team. 2004. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rawlins, S.L. and G.S. Campbell. 1986. Water potential: Thermocouple psychrometry. p. 597-662. *In* A. Klute (ed.) Methods of soil analysis. Part I. ASA, SSSA, Madison, WI.

Reeves III, J. B. and G. W. McCarty. 2001. Quantitative analysis of agricultural soils using near infrared reflectance spectroscopy and a fibre-optic probe. J. Near Infrared Spectrosc. 9:25-34.

Reeves III, J. B., G. W. McCarty, and J. J. Meisinger. 1999. Near infrared reflectance spectroscopy for the analysis of agricultural soils. J. Near Infrared Spectrosc. 7:179-193.

Reid-Soukup, D.A. and A.L. Ulery. 2002. Smectites. p. 467-499. *In* J.B. Dixon and D.G. Schulze (ed.) Soil mineralogy with environmental applications. ASA, CSSA, and SSSA, Madison, WI.

Russell, A.E., D.A. Laird, T.B. Parkin, and A.P. Mallarino. 2005. Impact of nitrogen fertilization and cropping system on carbon sequestration in Midwestern mollisols. Soil Sci. Soc. Am. J. 69:413-422.

Sadler, E.J., B.K. Gerwig, D.E. Evans, J.A. Millen, P.J. Bauer, and W.J. Busscher. 1999. Site-specificity of CERES-Maize model parameters: A case study in the Southeastern US Coastal Plain. p. 551-560. *In* J.V. Stafford (ed.) Precision agriculture '99. Proc. 2nd European Conf. Sheffield Academic Press, UK.

Scheinost, A.C. and U. Schwertmann. 1997. VIS-NIR reflectance spectra of goethite (α-FeOOH) as a function of particle size, unit-cell size, and cation substitution. Lunar and Planetary Science XXVIII. Lunar and Planetary Inst. Houston, TX.

Shepherd, K.D. and M.G. Walsh. 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Sci. Soc. Am. J. 66:988-998.

Sherrod, L.A., G. Dunn, G.A. Peterson, and R.L. Kolberg. 2002. Inorganic C analysis by modified pressure-calcimeter method. Soil Sci. Soc. Am. J. 66:299-305.

Slaughter, D. C., M. G. Pelletier, and S. K. Upadhyaya. 2001. Sensing soil moisture using NIR spectroscopy. Appl. Eng. Agric. 17:241-247.

Soil Survey Staff. 1996. Soil survey laboratory methods manual. U.S. Dept. Agri. Natural Resources Conservation Service. Soil survey investigations report, no. 42. U.S. Gov. Print. Office, Washington, DC.

Sorensen, L. K. and S. Dalsgaard. 2005. Determination of clay and other soil properties by near infrared spectroscopy. Soil Sci. Soc. Am. J. 69:159-167.

Steele, J. G. and R. Bradfield. 1934. The significance of size distribution in the clay fraction. Report of the 14th annual meeting. Am. Soil Sci. Assn., Bull. 15:88-93.

Sudduth, K.A. and J.W. Hummel. 1993. Soil organic matter, CEC and moisture sensing with a portable NIR spectrophotometer. Trans. ASAE. 36:1571-1582.

Sudduth, K.A. and J.W. Hummel. 1996. Geographic operating range evaluation of a NIR soil sensor. Trans. ASAE. 39:1599-1604.

Sudduth, K.A., J.W. Hummel, and S.J. Birrell. 1997. Sensors for site specific management for agriculture. p. 183-210 *In* F. J. Pierce and E. J. Sadler (ed.) The state of site specific management for agriculture. ASA, CSSA, and SSSA, Madison WI.

Tobias, R. 1995. An introduction to partial least squares regression. Proceedings of the Twentieth Annual SAS Users Group International Conference. Cary, NC: SAS Institute Inc., 1250-1257.

U.S. Department of Agriculture. 1973. Soil survey of Erath County, Texas. Soil Conservation Service. U.S. Gov. Printing Office, Washington, DC.

U.S. Department of Agriculture. 1977. Soil survey of Comanche County, Texas. Soil Conservation Service. U.S. Gov. Printing Office, Washington, DC.

U.S. Department of Agriculture. 2005. Official soil series descriptions [Online]. Available at http://soils.usda.gov/technical/classification/osd/index.html.

Williams, P.C. 1987. Variables affecting near-infrared reflectance spectroscopy analysis. p. 143-167. *In* P.C. Williams and K. Norris (ed.) Near-infrared technology in the agricultural and food industries. Am. Assoc. Cereal Chem., St. Paul, MN.

Wold, S., M. Sjostrom, and L. Eriksson. 2001. PLS-regression: A basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58:109-130.

Workman, J., Jr. 2004. Near-infrared spectroscopy. p. 11-31. *In* C.A. Roberts, J. Workman Jr., and J.B. Reeves III (ed.) Near-infrared spectroscopy in agriculture. ASA, CSSA, and SSSA, Madison, WI.

Workman, J., Jr. and J. Shenk. 2004. Understanding and using the near-infrared spectrum as an analytical method. p. 3-10. *In* C.A. Roberts, J. Workman, Jr., and J.B. Reeves III (ed.) Near-infrared spectroscopy in agriculture. ASA, CSSA, and SSSA, Madison, WI.

Wright, A.L. and F.M. Hons. 2005. Tillage impacts on soil aggregation and carbon and nitrogen sequestration under wheat cropping sequences. Soil Tillage Res. 84:67-75.

Zhu, A-Xing, L. Band, R. Vertessy, and B. Dutton. 1997. Derivation of soil properties using a soil land inference model (SoLIM). Soil Sci. Soc. Am. J. 61:523-533.

VITA

Name: Travis Heath Waiser

Address: Department of Soil and Crop Sciences, 2474 TAMU, College Station, TX 77845-2474

Email Address: twaiser@ag.tamu.edu

Education: B.S., Agronomy and Rangeland Ecology & Management, double major, Texas A&M University at College Station, Magna cum laude, 2003
M.S., Soil Science, Texas A&M University at College Station, 2005