SOIL PROPERTY ESTIMATION USING

VISIBLE AND NEAR-INFRARED REFLECTANCE

A Thesis

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

---

by

ALEXANDER H. SHERIDAN

Dr. Kenneth A. Sudduth, Thesis Supervisor

MAY 2012

The undersigned, appointed by the dean of the Graduate School, have examined the

thesis entitled

SOIL PROPERTY ESTIMATION USING

VISIBLE AND NEAR-INFRARED REFLECTANCE

Presented by Alexander H. Sheridan,

A candidate for the degree of Master of Science,

And hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Kenneth A. Sudduth, Biological Engineering

Dr. Jinglu Tan, Biological Engineering

Dr. Newell R. Kitchen, Soil, Environmental, and
Atmospheric Sciences

Dr. Robert J. Kremer, Soil, Environmental, and
Atmospheric Sciences

This thesis is dedicated to my family, past and present, your support and guidance has been invaluable.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SOIL PROPERTY ESTIMATION USING

VISIBLE AND NEAR-INFRARED REFLECTANCE

Alexander H. Sheridan

Dr. Kenneth A. Sudduth, Thesis Advisor

ABSTRACT

Site-specific management of soil physical and chemical characteristics for maintaining high yields within production systems while preserving soil quality has grown in use in recent years. This management includes measurement and mapping of soil variability across fields, which often requires a large number of soil samples. Additionally, traditional soil tests can be slow and/or expensive to complete, leading many farmers to skip soil tests completely. Recently, however, sensor-based approaches including reflectance spectroscopy have been proposed as quicker, easier alternatives. Therefore, the objectives of this study were 1) to investigate the capabilities of reflectance spectroscopy for the determination of both full profile and surface soil properties important in soil quality and 2) to determine if reflectance sensing at a limited number of wavelengths is a viable tool for measuring in-field variation of surface soil properties for site-specific management.

To address the first objective, soil samples were obtained from a research site which contained 32 plots with a wide variation in topsoil depth and assumed differences in soil quality. Soil samples from the 32 plots were scanned with a laboratory spectrometer in both field-moist and oven-dried conditions. Statistical calibrations were

developed relating reflectance data to conventional lab analysis. Total carbon (C) and nitrogen (N) showed high estimation accuracy with an R^2 of 0.97 and 0.91, respectively. Estimation accuracy of other soil variables was much more variable. Magnesium (Mg) and cation exchange capacity (CEC) exhibited good estimations ($R^2 = 0.83$ and 0.74, respectively), while pH and saturated hydraulic conductivity (K_s) did not ($R^2 = 0.07$ and 0.37, respectively). This research showed that spectroscopic analysis of field-moist soil is a viable option for estimating several important soil quality factors. This is advantageous because sample preparation time is further reduced and because sensing field-moist samples is a step toward successful in-situ sensing.

To address the second objective, data were obtained in two central Missouri production fields with a two-band sensor capable of capturing mobile measurements. Based on variability observed in field mapping with the two-band sensor, points were selected for soil sample collection. Collected soils were scanned with both the two-band sensor and a spectrometer in the laboratory. Of the different data sources, the best prediction accuracy ($R^2 = 0.75$) for soil organic carbon (SOC) was from in-field sensing with the two-band sensor. However, high accuracy was not maintained when calibration equations were applied in full-field mapping. In one field SOC was well-estimated with a root mean squared error (RMSE) of 6.9 g kg^{-1} but the other field had a RMSE of 18.4 g kg^{-1}. This difference could be attributed to the areas selected for soil collection not being indicative of the overall variation within the entire field. Various combinations of wavelengths from the spectrometer were examined to determine optimum two-band indices for SOC estimation. There was high correlation with SOC in the near-infrared

range, specifically when combining wavelengths greater than 1800 nm with wavelengths near 1400 and 1900 nm. When only considering combinations of wavelengths from <1000 nm, the highest correlation to SOC was when both were from the visible wavelength range. Additional data collection and analysis would be needed to determine if these same wavelengths were best for a wider range of soils.

CHAPTER 1
INTRODUCTION

Soil quality is a term that has been used by many researchers to explain the fitness and functionality of the soil (Karlen et al., 1997; Lal, 1993; Arshad and Coen, 1992; Larson and Pierce, 1991). Karlen et al. (1997) established a well-accepted version of the definition of soil quality which is the 'ability of soil to function'. Soil is a matrix of inorganic and organic matter, water, and air and in order to function properly these constituents must be in proper balance (Karlen et al., 1997). Actions such as agricultural management, building construction, and erosion can disrupt the natural processes of soil and change soil quality. Global increase in grain production has slowed and it is most likely attributed to inadequate soil and water management (Steer, 1998) leading to the reduction of soil quality.

Soil tests are needed to map changes in soil properties within a field and over time, but due to the time and cost associated with traditional testing, these tests are often not completed. Traditional soil testing may also not be a viable method due to the amount of samples needed to effectively map spatial variation seen in a field (Jung et al., 2006). But without testing, the quality of the soil cannot be determined.

Soil organic carbon (SOC) is an important soil property in soil management because SOC has a great effect on different components within the soil matrix. Water holding capacity (Hudson, 1994), infiltration (MacRae and Mehuys, 1985; Boyle et al., 1989), and bulk density (MacRae and Mehuys, 1985; Ekwue and Stone 1995) are just a few of the soil quality factors which researchers have found to be affected by soil SOC.

These researchers found that with increased amounts of SOC, other soil quality factors increased as well. Therefore, by increasing the amount of SOC, it may be possible to increase soil quality and fertility of the soil. While SOC is an important factor it is not the only soil property of importance when considering soil quality. Also, Sojka and Upchurch (1999) point out that there is a limit to the benefits of SOC.

Site-specific management for agriculture has been growing in use due to recent developments in measurement technologies and application techniques that attempt to boost yield while reducing input of chemicals and preserving soil quality. One measurement technology that has gained use in the past few decades is reflectance spectroscopy (Janik et al, 1998; McCarty et al., 2002; Viscarra Rossel et al., 2006; La et al., 2008; Lee et al., 2009). Light from visible (VIS; 350 to 700 nm) and near infrared (NIR; 700 to 2500 nm) wavelengths is directed at the soil surface and interacts with molecules causing energy shifts. Due to energy shifts of molecules within the soil, chemical and physical properties can be estimated (Murray and Williams, 1987). This method of measurement is much faster and cheaper than traditional soil tests. Soil spectroscopy is also a nondestructive test and does not require the use of hazardous chemicals. Many soil properties can be estimated with a single scan, including SOC (McCarty et al., 2002; Chang and Laird, 2002), nitrogen (N) (Janik et al., 1998; Dalal and Henry, 1986), and pH (Viscarra Rossel et al., 2006; Lee et al., 2009). Not only is SOC an important soil quality indicator but it is also well-estimated by reflectance spectroscopy. Several researchers (Lee et al., 2010; Reeves et al., 1999; Chang and Laird, 2002) have shown SOC to be very well predicted ($R^2 > 0.90$) from reflectance data.

One issue that has been discussed for reflectance spectroscopy as a site specific tool is the ability to detect within-field variation. Sudduth et al. (2010) found that spectroscopy gave poor estimates within a field that contained a small range in soil organic matter (SOM). Gomez et al. (2008) found similar results within fields that had varied amounts of SOC. Fields which contained a small range of SOC (5.4 to 10.0 g kg^{-1}) were poorly correlated with spectral values ($R^2 = 0.16$). In other fields which had greater ranges of SOC (10.8 to 51.0 g kg^{-1} and 0.02 to 51.0 g kg^{-1}) the R^2 was higher, 0.42 and 0.71, respectively. Conversely McCarty and Reeves (2006) found both C (3.6 to 29.2 g kg^{-1}) and N (0.3 to 2.2 g kg^{-1}) to be well estimated within a 20 ha field. These conflicting results leave the question of whether reflectance spectroscopy can be used for site specific management unresolved.

Most of the past soil spectroscopy studies have dealt with soil which has been dried and sieved (Viscarra Rossel et al., 2006; La et al., 2008; Sudduth et al., 2010), but this process adds to the overall analysis time. Collecting data on soils which remain in field-moist condition would not only improve efficiency but also be a necessary step towards in-situ measurements. Morgan et al. (2009) and Waiser et al. (2007) both studied reflectance spectroscopy with soils that had been subjected to different processing techniques. These studies included cores left intact, both moist and dried, and dried and sieved soil. They found that both types of dried samples provided similar prediction accuracy and results with the field-moist samples were slightly less accurate.

Both Morgan et al. (2009) and Waiser et al. (2007) studied cores which had been collected to a depth of roughly 105 cm, but other studies have discussed the estimation of

surface soil properties with on-the-go spectral data collection. In one of the first studies with an on-the-go in-situ sensor, Shonk et al. (1991) found SOM could be estimated with R^2 as high as 0.95 with single-wavelength sensing at a depth of 10 cm. Sudduth and Hummel (1993a, 1993b) experimented with a prototype multi-wavelength on-the-go sensor for estimating SOM, cation exchange capacity (CEC), and moisture content. The results found in these studies were not as positive as those found by Shonk et al. (1991) and it was determined that high standard deviations were caused by the sensor moving in reference to the target soil during data collection. More recently, Roberts et al. (2011) used a two-band reflectance sensor, which has typically has been used to detect N deficiencies in crops, for SOM estimation. This sensor, different from the sensors used by Shonk et al. (1991) and Sudduth and Hummel (1993a, 1993b), was used in an above ground position not making contact with the soil and was able to predict SOM moderately well ($R^2 \geq 0.76$).

OBJECTIVES

Past research has shown that reflectance sensing can be a viable method for the estimation of soil properties. However, much of the research has studied a limited number of soil properties and the capabilities have not been truly tested. Also, many of these studies used soil samples from many fields over vast areas to calibrate their data to the soil properties in question. This leaves the question of whether reflectance sensing can be used as a within-field site specific management tool unresolved.

Therefore the overall objectives of this study were 1) to investigate the capabilities of reflectance spectroscopy for the determination of both full profile and

surface soil properties and 2) to determine if reflectance sensing at a limited number of wavelengths is a viable tool for measuring in-field variation of surface soil properties for site-specific management.

Specific objectives were to:

- Select appropriate mathematical and statistical methods for calibrating reflectance data to lab measured soil properties for both full profile and surface soil samples.

- Develop reflectance estimates of soil properties using field-moist soil and compare the results from the field-moist data to reflectance data obtained with methods from past research (dried and sieved samples).

- Evaluate the capabilities of a prototype two-band surface soil on-the-go sensor both in a lab and field setting for the soil properties of C, N, and moisture.

- Compare the lab results from the two-band soil sensor to results from a full-range spectrometer and evaluate whether the selected wavelengths within the two-band sensor also provided the best results with full spectrum sensing.

CHAPTER 2
REVIEW OF LITERATURE

DEFINING SOIL QUALITY

Soil is an integral part of nature consisting of water, air, organic, and inorganic matter in a complex matrix. It serves as a medium for plant growth, provides support for physical structures, serves as a filter for contaminants that leach into the environment, and has many other functions. Soil has also been considered as a possible sink for carbon (C) storage to reduce atmospheric C. Different soils are better suited for different functions based on their chemical (e.g. available plant nutrients), biological (e.g. organic matter; microbe content), and physical (e.g. sand, silt, and clay; bulk density) properties (Karlen et al., 1997). Many of these properties can be altered by management practices, thereby altering the soil's full ability to function. Due to increased levels of chemical run-off, erosion, and reduction of productivity, the concern of soil degradation has been raised (Karlen et al., 1997; Lerch et al., 2005, Rachman et al, 2003).

Conservation of soils dates back to passages written within the Bible, where the book of Leviticus speaks of continuous farming for six years then allowing land to fallow on the seventh year "For it is a year of rest unto the land" (Scofield, 1945). This passage points to the importance of soil quality for continuous farming. Nutrients cannot be continuously stripped from the land without the expectation of diminished yields and destruction of the soil. The Dust Bowl was a devastating disaster in American history caused by inappropriate land use (e.g. cultivation of prairie land) and exacerbated by years of drought conditions. Previous to the Dust Bowl, many farmers set out to start a

new life in an unfamiliar area. Lacking the knowledge of how to farm this new region, farmers caused vast changes to the plains and reduced the quality of the soil on which they farmed. It was not until after the Dust Bowl catastrophe that more emphasis was placed on what farmers should do to retain the quality of soil needed for annual cropping (Hurt, 1981).

Without a definition or idea of what constitutes soil quality, it is difficult to determine how to assess issues affecting soil quality. Therefore, Dr. L.P. Wilding, 1994 President of the Soil Science Society of America (SSSA), appointed a group of members within the SSSA to provide a concise concept of soil quality and the attributes which affect it. This committee defined soil quality as the ability of soil to function (Karlen et al., 1997). Along with this definition was also stated that the soil is functioning 'within natural or managed ecosystem boundaries, to sustain plant and animal productivity, maintain or enhance water and air quality, and support human health and habitation' (Karlen et al., 1997). Previous to the definition given by the SSSA, researchers defined soil quality similarly or have adopted this definition after its inception (Arshad and Martin, 2002; Lal, 1993; Arshad and Coen, 1992; Larson and Pierce, 1991). Arshad and Coen (1992) described soil quality as a way to understand and document how soil adapted to environmental and management changes.

SOIL QUALITY INDICATORS

The definition of soil quality is useless to farmers and researchers if it cannot be used in a practical application. Therefore, it is important that along with these definitions, one also understands the factors and practices which affect soil quality. This can be a

difficult task due the complexity of soil and its dynamic relationship with the surrounding

environment. In a production setting, soil quality is measured by the soil's ability to be

sustainably productive. Therefore, factors which affect quality are ones which affect the

ability of soil to be productive. Much research has been done (Stern et al., 1991; Doran et

al., 1990; Arshad and Coen 1992; Karlen et al., 1994a,b) to determine how different soil

properties affect soil processes (Table 2.1). Many soil properties have an interdependent

relationship with one another. This means, for example, that if there is a drop in

aggregation, soil organic carbon (SOC) could be affected (Chappell et al., 1999). It is

therefore important that key soil properties are maintained within healthy levels to ensure

that soil quality is maintained.

Table 2.1. Selected indicators of soil quality and some processes they impact (from
Karlen et al. 1997)

Measurement	Process affected
Organic matter	Nutrient cycling, pesticide and water retention, soil structure
Infiltration	Runoff and leaching potential, plant water use efficiency, erosion potential
Aggregation	Soil structure, erosion resistance, crop emergence, infiltration
pH	Nutrient availability, pesticide absorption and mobility
Microbial biomass	Biological activity, nutrient cycling, capacity to degrade pesticides
Forms of N	Availability to crops, leaching potential, mineralization
Bulk density	Plant root penetration, water- and air-filled pore space, biological activity
Topsoil depth	Rooting volume for crop production, water and nutrient availability
Conductivity	Water infiltration, crop growth, soil structure
Available nutrients	Capacity to support crop growth, environmental hazard

Tests to check for factors such as SOC, nitrogen (N), aggregation, and others are

done in a lab setting. When a field designated for research has been established, soil

samples are taken from different areas in the field to spatially map field variation. These

samples are taken from either the surface or a coring device is used to remove soil from

greater depths. Once samples are removed from the field, the desired analyses can be run.
For example, to analyze for aggregate stability, a portion of the sample is taken as a
subsample to measure gravimetric water content and another portion is air dried. The air
dried samples are then weighed and placed on a sieve which is attached to a 'dunking
machine'. The dunking machine repeatedly dunks the sieves into buckets of water to
wash away loose (weakly aggregated or non-aggregated) soil. After the soil has been
washed it is dried once more and then shaken in a solution so that only the aggregate
portion remains (Kemper and Rosenau, 1986; Angers et al., 1993). There are several
drawbacks to running a test such as this; chief among them is the significant time that
must be dedicated to running this test. Depending on the number of samples to be
evaluated, most soil labs could not run the entire set at once, meaning that several batches
would need to be run. With each batch taking 5 to 7 days, this test requires a large
amount of time to complete. Another drawback to this test is the multistep nature of it, as
within each step errors could occur and change the final result of the test. Many soil tests,
such as the test for aggregate stability, are often skipped due to the time, cost, and
multistep nature of the tests. Not testing for various soil properties (e.g. bulk density,
microbial biomass, etc.) limits the data available for soil quality assessment, thereby
limiting research capacity and decision assistance for land management.

The two main reasons to investigate soil properties are for research and for
management of production systems. By researching the different effects external actions
have on soil, production systems can be better managed to maintain soil quality and high
yield. In the area of research, the focus is primarily on the effect different natural and

agricultural actions have on soil. An example of this would be the effect of water or wind

erosion on soil quality. The second area of interest is providing information for

management of production systems. This includes the use of soil sampling to determine

soil input needs for plant growth. Both research and management of production systems

are important because they provide knowledge about the soil's reaction to external

factors. Soil C is of great interest in both cases. Within a production system, SOC can

greatly increase the productivity of the land (Larson and Pierce 1991; Cannell and

Hawes, 1994) as SOC can have significant effects on crop water availability and soil

erosion (Bruce et al., 1995). Many authors have discussed the benefits of SOC at length

and how increases in SOC relate to other soil quality parameters (Table 2.2). Nitrogen is

one soil property of great interest that is tied to SOC. Soil N content often parallels that of

C in a C:N ratio of roughly 10:1 (Paul and Clark, 1989). N is an essential nutrient needed

by all crop plants. The two main sources of available N in the soil are residual N and

mineralizable N contained in soil organic matter (SOM). With the breakdown of SOM

during the growing season, plants can uptake mineralizable N, reducing the need for

inputs of N by way of fertilizing (Follet, 2008). Holland and Schepers (2010) used this

idea by including SOM in a N application algorithm to reduce the amount of N applied to

the field. By examining the effects SOC and other parameters have on soil and its ability

to support plant growth, researchers can help farmers gain more knowledge for managing

their soil for maximum yield while at the same time protecting soil quality.

Table 2.2. Selected soil parameters which have been linked to soil organic carbon (SOC)

Soil Parameter affected by SOC	Reference
Plant available water capacity	Hudson, 1994
Infiltration	MacRae and Mehuys, 1985; Boyle et al., 1989
Aggregate formation and stability	MacRae and Mehuys, 1985: Boyle et al., 1989
Bulk density	MacRae and Mehuys, 1985; Ekwue and Stone 1995
Soil strength	Ekwue and Stone, 1995
Cation exchange capacity	Chan et al., 1992; Mahboubi et al., 1993
Soil enzyme activity	Dick, 1983
Earthworm activity	Hendrix et al., 1992

FUNDAMENTALS OF VISIBLE AND NEAR INFRARED REFLECTANCE

Over the past few decades a growing number of papers have focused on the use of reflectance spectroscopy for the estimation of soil properties (Janik et al, 1998; McCarty et al., 2002; Viscarra Rossel et al., 2006; La et al., 2008; Lee et al., 2009; Sudduth et al., 2010). Reflectance spectroscopy uses light from the visible (VIS; 400 to 700 nm) and infrared (700 nm to 1 mm) wavelength ranges. The VIS and infrared regions are roughly in the middle of the electromagnetic spectrum (1 pm to 100 Mm) between ultraviolet (10 to 400 nm) and microwave (1 mm to 1 m). The infrared range can be split in three divisions: near (NIR; 700 to 2,500 nm), mid (MIR; 2,500 to 25,000 nm), and far (FIR; 25,000 nm to 1 mm). Currently only the NIR and MIR divisions are used in soil spectroscopy.

In the 19th century, Sir Frederick William Herschel discovered NIR radiation when experimenting with the energy associated with each band of VIS light produced by the sun. When separating light with the use of a prism, he noticed that a thermometer placed just beyond the red region of light had a greater temperature than that of one in the red band. This experiment done by Herschel was the first insight into the energy related

to NIR radiation. With the use of spectroscopic technology, the energy that caused changes in Herschel's thermometer has now been used to estimate the chemical and physical properties of many items such as food and soil.

Much emphasis has been put on the importance of SOM. Soil organic matter is primarily made up of the elements C, N, oxygen, hydrogen, phosphorus and sulfur. These elements are bound together into molecules by either covalent or electrovalent bonds. At their ground state, all molecules vibrate at a specific frequency with vibrational and rotational energies. The sum of these energies is the total energy level of the molecule. Vibrational and rotational energy are abbreviated v and j respectively and at the ground state they are denoted as such. When energy is applied (e.g. NIR radiation), elements are excited to different levels of energy. The first is the fundamental absorption ($v_1 j_1$), but due to the fact that this absorption occurs within the mid-infrared range (2,500 to 15,000 nm) these absorptions will not be covered in this research. Absorption which does occur within the NIR range is in the first, second, and third overtone and combination bands. In most cases overtones above the third are ignored because elements typically do not gain enough energy to reach this level and therefore produce a weak signal. The weak signal above the third overtone can be masked by another molecule overtone which has the same frequency. These absorption bands occur at roughly one half, one third, and one fourth of the fundamental absorption wavelength with fewer molecules reaching increasing overtone levels.

Due to the nature of the bonding of molecules, they do not follow Hooke's law and do not exactly follow the spacing rule for overtones mentioned above, but are at

slightly longer wavelengths than expected. When the elements within the molecules become excited the bonds that hold them together undergo stretching and bending. Stretching occurs when the distance between the atoms changes and bending affects the angle of the bond. These bending motions are termed twisting, wagging, scissoring, and rocking and only the rhythmic changes in dipole moment of a molecule in both the bending and stretching motions can cause absorbance in the infrared (Murray and Williams, 1987). Reflected light from the object in question is captured by a spectrometer. The light is then split into its individual components for detection. The spectrometer output is reported as percent reflectance (R) which can be transformed to apparent absorbance (A): $A = \log(1/R)$ which is the amount of energy absorbed by the sample.

SOIL REFLECTANCE

As stated above, SOM is of much importance and the absorbance bands for SOM can be seen in the VIS range from absorption by chromophores and the darkened color of SOM. Darkened soil can be deceptive however, as water content plays a role in the color of a soil. Absorption in the VIS range is also associated with minerals containing iron (Mortimore et al., 2004). Within the NIR range there are several absorptions of interest for soils. Clay absorption occurs in the NIR due to metal-OH bending and O-H stretching vibrations (Viscarra Rossel et al., 2006). A strong water band is typically seen around 1400 and 1900 nm with weaker absorptions in other areas (Liu et al., 2002). Overtones and combinations of the groups of OH, SO_4, CO_2, and CH also absorb in the NIR range (Clark, 1999; Murray and Williams, 1987).

Not all data estimated by reflectance spectroscopy is correlated to a specific wavelength. While soil properties such as C (1100, 1600, 2000, and 2200 to 2400 nm; Dalal and Henry, 1986; Clark, 1999; Martin et al., 2002; Stenberg, 2010) and clay minerals such as illite (2300 nm; Clark et al., 1990) have specific wavelengths of absorption, many other soil properties are measured indirectly. Clay content is one such soil property. Clay content is estimated by the detection of the clay minerals it contains (Stenberg et al., 2010). Plant nutrients are another example of soil properties which can be estimated indirectly. Many researchers (Janik et al., 1998; Chang et al., 2001; Daniel et al., 2003; Cozzolino and Moron, 2003; Islam et al., 2003) have reported varying success for estimating soil plant nutrients such as K, P, Mg, and Ca ($R^2 = 0.55$ to 0.90). It is thought that successful estimation of these properties is due to the co-variation of plant nutrients with spectrally active soil constituents (Stenberg et al., 2010).

MULTIVARIATE CALIBRATIONS

Estimating soil properties with the use of reflectance spectroscopy is not as simple as determining which wavelengths have high absorbance and which do not to determine which molecules are present. Several factors cloud the water in terms of VIS and NIR data extraction. There are several areas of overlapping absorptions for soil properties and scattering effects caused by the soil structure adds to the issues with soil property estimation. These factors cause a low signal-to-noise ratio, which requires mathematical extraction of the absorption patterns to correlate to soil properties (Murray and Williams, 1987). Therefore, researchers have looked to multivariate calibrations to solve this issue (Marten and Naes, 1989, La et al., 2008; Sudduth et al., 2010; Stenberg, 2010; Lee et al.,

2010). The three main techniques that have been used in the past are partial least squares

regression (PLSR; La et al., 2008; Lee et al., 2010), principal components regression

(PCR; Chang et al., 2001), and stepwise multiple linear regression (SMLR; Dalal and

Henry 1986; Lee et al., 2010). Much of the work with SMLR has been to extract specific

wavelengths corresponding to the prediction of a particular soil property. PCR and PLSR,

both of which use data from all measured wavelengths, are better techniques for

estimating soil properties using reflectance spectroscopy due to their ability to cope with

data containing large numbers of predictor variables that are highly collinear. Among the

three mathematical techniques PLSR is used more by researchers due to its ability to

explain more variance in the response and predictor variables using fewer components.

Many other techniques are gaining acceptance, including neural networks (NN; Daniel et

al., 2003), boosted regression trees (Brown et al., 2006), and multivariate adaptive

regression splines (MARS; Shepherd and Walsh, 2002). These techniques are better

suited for handling data which is nonlinear in nature compared to PLSR, PCR, and

SMLR which are linear models. Some have even combined these methods (e.g. PLSR-

NN; Mouazen et al., 2010) to create a more robust prediction method.

VISIBLE AND NEAR INFRARED SPECTROSCOPY SENSORS

Visible and near-infrared spectroscopy makes use of the energy associated with

VIS and NIR radiation and advances in technology to estimate soil properties. This is a

nondestructive method which requires no chemicals, measurement time is very short

compared to traditional lab analysis, and several soil properties can be estimated from a

single scan. Since this sampling method is cheaper and less time consuming, more soil samples can be spatially quantified for use in research or precision agriculture.

Many variations of sensors currently exist that provide either a selected or full range of the VIS and NIR wavelengths. The sensors which provide a full range of VIS and NIR are much more versatile and powerful because of their ability to detect many more molecules at different overtones than the selected-wavelength sensors. Because of this, much of the past research has been focused on full-spectrum sensors. However, sensors providing a full spectrum do have their limitations. One of the main limitations currently is the cost of the sensor, which limits the use of full-spectrum sensors to research facilities with little to no farmer use. There is also a size limitation for full range VIS and NIR sensors. These sensors require a large unit not only to produce light from 350 to 2500 nm but also to contain the several detectors needed. Because of this, these sensors are cumbersome as an in-field sensing device. Since several detectors are used within these devices, external temperature during sensing, both hot and cold, can affect their accuracy (Markham et al., 1995). To detect the full range of data, many sensors contain three separate detectors and drifts in the overlapping areas between these sensors can be seen when temperature variation becomes an issue. Because of the drifting issue, sensor use could be limited in months of excessive heat or cold. While sensor drifts are made worse by extreme temperatures, the drifts are still present in normal lab conditions. To rectify this situation a reference of known reflectance is used as a correction. A common reflective reference used is a Spectralon panel (Labsphere, North Hutton, NH) which is 95 to 99% reflective.

Another limitation in several of the past studies is the area being sensed, or the sensor's optical footprint. Most past studies have used active light sensing (Waiser et al., 2007; Morgan et al., 2009; Tekin et al., 2012) meaning that the device illuminates the sample with light from the VIS and NIR spectrum. In this setting, the sensor's power output may not be great enough to create a large sensing area (e.g. optical footprint of the ASD Field Spec Pro is ~2 cm diameter when using its high intensity source probe; ASDi, 2002). This reduces not only the amount of soil that can be sensed per scan, but the soil variation that can be seen as well. There are alternatives to using active light sensing. The first is to use passive light; in this method, light from the sun is reflected from the surface of the object in question and captured by the spectrometer. Very little research has been done with the use of passive light for in-situ spectrometry, most likely due to the sensitivity of this sensing method to variable cloud conditions and requirement of specific sun angles during sensing. The other technique, also passive, is remote sensing which employs the use of aerial imagery by way of plane or satellite to provide images for analysis. Researchers have attempted this type of analysis for many soil factors such as soil C (Chen et al., 2000, 2005). These studies used color aerial imagery provided by the USDA Farm Service Agency and digital images from the National Aeronautics and Space Administration (NASA) ATLAS satellite sensor. The ATLAS sensor collects data at wavelengths ranging from 452 to 12,019 nm. Both of these studies found that by excluding trouble areas, such as shaded areas or areas of dense vegetation, SOC prediction accuracy was quite high with $R^2 = 0.98$. Despite this relationship, remote sensing is not often used on a routine basis due to issues of variability in physical features

of the soil (e.g. moisture content, surface roughness, vegetation) and cost of this sensing method.

SAMPLE PREPARATION

Much of the past research has been with soil samples which have been cored, sieved and dried (Sudduth et al., 2010; Stenberg, 2010; La et al., 2008; Lee et al., 2010; Shepherd and Walsh, 2002). The reason for this is variable soil moisture can have a large effect on prediction accuracy. Morgan et al. (2009) and Waiser et al. (2007) found that when comparing results from cores which had been dried to the results of cores which were in field-moist condition, the R^2 dropped by roughly 0.1 using cores in field-moist condition. Tekin et al. (2012) compared the prediction accuracy for SOM in soil that had various moisture contents. They too found that soil which was dried has the highest prediction accuracy, but the decrease in accuracy of the wetted soil was not as large as in the studies done by Morgan et al (2009) and Waiser et al (2007). Sudduth and Hummel (1993b) experimented with the prediction accuracy of VIS and NIR reflectance on soils which had been manually set to the moisture tension levels of 0.033, 0.33, and 1.5 MPa. Soils which had remained oven dried had a slightly better R^2 (0.92) and standard error of prediction which is the standard error of the estimate in the validation dataset (SEP; 2.0 g kg^{-1}) compared to the soils which had been rewetted (R^2= 0.89; SEP= 2.3 g kg^{-1}) for analysis of SOC.

When comparing intact to ground core samples, Morgan et al. (2009) and Waiser et al. (2007) found that cores which were left intact while scanning had better prediction accuracy. They concluded that the heterogeneity of the intact core played no role in the

ability of reflectance spectroscopy to predict soil properties and that the higher bulk density of the intact core provided a stronger reflectance signal. Therefore, soils which have been cored and dried but not sieved may provide better prediction accuracy. Although results from researchers have varied, drying and sieving soils may not be necessary to attain accurate estimation of soil constituents. Estimations could occur more quickly with the removal of this time consuming step.

SOIL CARBON

When considering which soil properties to test with the use of VIS and NIR spectroscopy, it is important to note which have led to good prediction accuracy in the past and which are most important for soil quality. Because SOC has such a large impact on other soil properties, a large amount of research has been devoted to it and many studies have shown that C, be it organic or inorganic, can be well estimated by spectroscopy. The absorptions of C occur at 2500 to 2550 nm, 2300 to 2350 nm, 2120 to 2160 nm, 1970 to 2000 nm, and 1850 to 1870 nm (Clark et al., 1990). Lee et al. (2010) found for soils which had been collected in the Midwest US that SOC prediction had an R^2 as high as 0.97 in one subset of soils with an average of 0.87 for all soils combined. Shepherd and Walsh (2002) also found SOC to be well estimated ($R^2 = 0.91$) with soil that had been dried and sieved. Both of these studies featured a wide range of C values for calibration. But when testing samples from single fields the range of C might be limited.

Gomez et al. (2008) tested several fields with varying SOC amounts. They found that in fields which contained limited amount of SOC variation (5.4 to 10.0 g kg^{-1}) that

prediction results were quite low ($R^2 = 0.16$). The study also investigated two other sets

of soils, one with a range of SOC of 10.8 to 51.0 g kg^{-1} and the other having a range of

0.02 to 51.0 g kg^{-1}%. In both of these sets of soil the range of data was much greater than

that of the first and both had higher R ^{2} values as well, 0.42 and 0.71, respectively. While

the data set containing a small range of C did not have a high R^2 the root mean square

error of prediction (RMSEP) was quite low (0.1%). Similarly Sudduth et al. (2010) found

that in a field which contained a limited range in SOM that the R^2 was low (0.31) but the

RMSEP was as good (0.28%) as that of published data with a better R^2 (RMSEP ~0.3%,

Viscarra Rossel et al., 2006). This study points to the fact that if the set of soils which is

to be tested contains a small range of data, that the calibration of the data may not be

robust and soil spectroscopy may not be a viable method for measurement.

SENSORS: SOIL NUTRIENTS

Nitrogen (N) is a soil property which has often been tied to SOC. Not only is the

amount of soil available N thought to be related to SOC (Table 2.1), some researchers

believe that the prediction of N with the use of VIS and NIR spectroscopy is correlated to

the prediction of SOC. N absorption occurs at roughly 751, 1000, 1500, and 2060 nm

(Viscarra Rossel and Behrens, 2010). Chang and Laird (2002) artificially changed soils

with the addition of $CaCO_3$, various composts, and humic acid. By doing so they created

C:N ratios which would typically not be found in soils. They found that the analyses of

total N and C:N ratio were much more accurate than what could be explained by

autocorrelation of total N with organic, inorganic, or total C.

While both C and N have been successfully estimated by spectroscopy, results with other soil properties have been varied (Janik et al., 1998; Chang et al., 2001; Cozzolino and Moron, 2003; Daniel et al., 2003; Islam et al., 2003). Nutrients such as potassium (K), phosphorus (P), magnesium (Mg), and calcium (Ca) are important for the support of plant growth and have been identified as important soil components for soil quality. These researchers reported highly variable coefficients of determination (R^2; in parentheses) for K (0.55 to 0.80), P (0.07 to 0.81), Mg (0.63 to 0.90), and Ca (0.67 to 0.90). These variations in predictability could be attributed to covariation within spectrally active constituents, the type of soil variable measured (extractable, available), or the corresponding laboratory method used. For example, Chang et al. (2001) compared Mehlich III (Mehlich, 1984) for extracting cations to NH_4OAc extraction and found better spectral prediction of those which were extracted by Mehlich III.

SENSORS: ON-THE-GO SOIL REFLECTANCE

A small number of studies have been devoted to on-the-go spectral soil sensing. Many of these sensors use only a few wavelengths instead of a full spectrum from the VIS and NIR for quickness and economy of measurement. By using a sensor with a select number of wavelengths, field-sized soil maps can be acquired much quicker. However, by reducing the number of wavelengths used, there is often a reduction of the soil properties that can be sensed effectively. Shonk et al. (1991) developed and tested a sensor which contained red light emitting diodes (LEDs; 660 nm). When attached to the rear of tractor and pulled at a depth of 10 cm in a field where SOM ranged from ~2 to 6%, accurate estimates ($R^2 = 0.95$) were obtained.

Sudduth and Hummel (1993a, 1993b) developed and tested a portable sensor for analyzing SOM, cation exchange capacity (CEC), and moisture on-the-go. This sensor collected data at wavelengths from 1700 to 2420 nm, which were then averaged to either a 40 or 60 nm spacing. In the field, SOM was not well predicted with this sensor. They suggested this was likely due to the movement of the sensor, which acquired data at the different wavelengths sequentially, in reference to the targeted soil during acquisition of spectral scans.

Christy (2008) used a commercial device mounted to an agricultural shank and pulled behind a tractor. This device acquired data at a rate of 20 scans/s between 350-2224 nm with an 8-nm resolution. The shank allowed the spectrometer to penetrate the soil surface and by doing so remove some variation due to water content and surface roughness. This study examined eight fields within Kansas totaling 292 ha. In testing for SOC and SOM the sensor had high prediction accuracy with an R^2 of 0.92 and 0.80, respectively. Bricklemyer and Brown (2010) used the same device and found very poor accuracy when predicting SOC ($R^2 = 0.00$). They suggested that low prediction accuracy was most likely attributed to low variation in SOC within the Montana soils that were tested as SEP values (3.5 g kg^{-1}) were similar to those (3.0 g kg^{-1}) obtained by Christy (2008).

The sensors described above were all designed to contact the soil, which could cause issues, such as smearing of soil onto the lens of the sensor, with the ability of sensor to predict the desired soil properties. As an alternative Roberts et al. (2011) examined a 2-band sensor developed to sense crop canopy reflectance to estimate the

required amount of N (Raun et al., 2005; Solari et al., 2008; Kitchen et al., 2010). The two bands in this sensor were split between the VIS (590 ± 5.5 nm) and NIR (880 ± 10 nm). For soil sensing, it was mounted to an ATV which followed behind a planter. Due to the soil surface disruption caused by the planter, soil moisture was less variable. The researchers used a linear prediction model to estimate SOM that incorporated high, medium and low values of lab-measured SOM and the sensor output measured in the field. Three different models were used, a universal model capable of predicting SOM on all fields, another which was field-specific and a final model which combined the slope from the universal equation with a field-specific intercept. In their investigation of the three estimation models, both the completely field-specific and field-specific intercept models had high prediction accuracies (R^2 = 0.88 and 0.85, respectively).

CHAPTER 3
ESTIMATION OF SOIL QUALITY CHARACTERISTICS USING REFLECTANCE SPECTROSCOPY

ABSTRACT

Site-specific management is an effective method for optimizing crop production while maintaining soil quality. Management of soil quality requires measurement and mapping of observed variability across fields. One drawback is that understanding soil variability requires a large number of traditional soil tests which can be slow and/or expensive. Recently, however, sensor-based approaches including reflectance spectroscopy have been proposed as quicker, easier alternatives. This study tested the ability of visible and near-infrared reflectance spectroscopy to estimate total carbon (C), total nitrogen (N), particulate organic matter (POM)-C, POM-N, and other soil quality factors. Samples were taken from 32 plots in Columbia, MO, with a wide variation in topsoil depth and assumed differences in soil quality. Soil samples were scanned with a laboratory spectrometer in both field-moist and oven-dried conditions. Statistical calibrations were developed relating reflectance data to conventional lab analyses. Results showed that total C and N models were highly predictive, with R^2 as high as 0.97 for C and 0.91 for N. Estimations of other soil quality factors were not as accurate, but estimates of some properties, including Mg and CEC were of good accuracy, with R^2 of 0.83 and 0.74, respectively. This research showed that spectroscopic analysis of field-moist soil is a viable option for estimating several important soil quality factors. This is

advantageous because sample prep-time is further reduced and it indicates that in-situ

sensing could be possible with this sensor.

INTRODUCTION

The importance of soil conservation was raised by Bennett (1939) who stated that

"…The decrease, or failure to increase [corn yield], has been restricted principally to soil

worn lean by erosion." Due to erosion, organic matter can be removed along with the

fertility and productivity of the soil (Andrews et al., 2004). To keep up with an increasing

demand for crops, it is important that the soil retains its quality or "capacity of the soil to

function" (Karlen et al., 1997). Soil conservation techniques used in farming are

important for retaining soil quality and the ability of the soil to produce various crops.

Not only do farmers have to feed the world but developing renewable energy markets

have taken part of the production that was typically used for food and feed production

and devoted it to biofuel production. The increase in consumption, be it for food or fuel,

has not been matched with increased production. In recent years production increases

have fallen from a 3% increase per year to 1.3% and it is thought that soil quality may

play a role in this decrease (Steer, 1998).

Even when changes are made in farming practices, methods for the determination

of improvement of soil quality have lagged (Arshad and Martin, 2002). A contributing

factor to this is the time and cost associated with traditional laboratory testing methods.

Soil samples are cored from the field, either manually or with a hydraulic coring device,

dried, and crushed to a proper size consistency, and finally tested for the particular

characteristic in question. A secondary issue in determining soil quality is deciding which

indices should be used. Factors such as productivity, relationship of soil and biotic health, erodibility, and others have been suggested, as well as a system that weights each factor based on degree of importance (Doran and Parkin, 1994). The final variable for soil quality is land use and depending on planned land use there are various factors that differ in importance. In a farming operation there are five key factors that influence soil quality as discussed by Arshad and Martin (2002). These factors are aggregation, infiltration, bulk density, microbial biomass, and available nutrients. These indicators are most affected by organic matter (OM), aggregation, pH, and the nutrients important for the capacity of the soil to support plant growth, such as N, P, and K.

To alleviate the problem of tedious, costly lab testing, soil spectroscopy could be a candidate method for use in place of traditional laboratory tests. Reflectance spectroscopy for soil analysis is a non-destructive technique which primarily uses light from the visible (VIS; 400-700 nm) and infrared (IR; 700 nm-1 mm) wavelength ranges (Janik et al, 1998; McCarty et al., 2002; Viscarra Rossel et al., 2006; La et al., 2008; Lee et al., 2009). The infrared range can be split into three divisions: near (NIR; 700-2500 nm), mid (MIR; 2500- 25000 nm), and far (FIR; 25000 nm-1 mm). Currently only the NIR and MIR divisions are used in soil spectroscopy.

Once soils have been scanned, the same spectrum can be used to analyze a range of data. Information such as pH (Viscarra Rossel et al., 2006; Lee et al., 2009), C/N ratio (Ludwig et al., 2002; Kweon et al., 2009), texture (Lee et al., 2009; Cozzolino and Moron, 2006), and others can all be obtained from the same spectra. Many of the past analyses have been done with soils that were dried and sieved (Lee et al., 2010; Viscarra

Rossel et al., 2006). However, there has been some research towards using soil in a field-moist condition. Sudduth and Hummel (1993b) found that organic C and moisture could be well estimated in the laboratory with an R^2 of 0.82 and 0.94, respectively, within soil moisture values that ranged from field capacity to wilting point. More recently, Kweon et al. (2009) found that total C and N could be estimated with an R^2 of 0.91 and 0.89, respectively, using in-field sensing across a range of various soil conditions. Morgan et al. (2009) studied the effect field-moist intact cores had on accuracy as compared to dried and sieved soils. These researchers found only a slight reduction in accuracy of soil organic C (SOC) from dried and sieved to field-moist soils (R^2 = 0.73 and 0.64, respectively). Spectroscopy can be a powerful tool for soil analysis and by using spectroscopy, soil characteristics can be quantified at a rate with which conventional soil testing cannot compete. Due to time savings, tests could be run more often to monitor soils on a yearly or shorter basis, to track the effectiveness of different soil conservation techniques.

OBJECTIVES

The goal of this paper was to determine the accuracy of field-moist soil spectroscopy for the estimation of chemical and physical soil properties which are considered to affect soil quality.

Specific objectives were:

- Select appropriate spectral pretreatment and validation methods

- Develop reflectance estimates of soil properties using field-moist soil

- Compare field-moist estimates to those obtained with dried soil samples

MATERIALS AND METHODS

Soil samples were taken from a research site at South Farm of the University of Missouri. The site contains 32 blocks (22.9 by 9.1 m) created to simulate eroded, neutral, and deposition areas with designed topsoil depths varying from 0 to 38 cm (Figure 3.1). This site was created in 1982 to study the effect of topsoil depth on crop production of a Mexico silt loam soil (mesic Vertic Epiaqualfs) (Gantzer and McCarty, 1987). Research was conducted on the site for about a decade until it was allowed to go fallow in the early 1990s.

Figure 3.1. Layout of the 32 plots at University of Missouri South Farm.

In the spring of 2009 a new project was initiated on this site to study the relationship of topsoil depth to crop grain and switchgrass biomass production for renewable energy. As part of this project initiation, a baseline soil sampling was conducted. The blocks were split into thirds (7.6 by 9.1 m) and samples were taken on a diagonal transect starting in the northeast corner of the east third and ending in the southwest corner of the west third for a total of three samples per block. A Giddings hydraulic soil sampling device with a 3.8 cm diameter and 120 cm long tube was used for coring. The three cores from each block were separated into sections by pedogenic

horizon in the field by a soil scientist and each section was combined with the same

sections from the other cores from that block. The samples were then pushed through a

coarse sieve (5 mm) and separated, allowing half for standard laboratory soil analysis and

the other half for spectral analysis. The samples were stored in a cooler at 4°C to

maintain their chemical and physical characteristics before analysis. There were a total of

252 samples within the set. The number of horizons varied based on sampling location,

but there were an average of 7 profile samples per block. Each of the 32 blocks was also

assessed for soil fertility properties by traditional sampling and lab analysis. Eight

dispersed cores were taken using a 1.8-cm diameter hand sampler to a depth of

approximately 18 to 20 cm from the same block as the hydraulic cores and these eight

cores were combined by block, for a total of 32 shallow samples. From two separate and

random locations within each block, two 17.5 cm diameter cores were taken at two depth

increments (0 to 10 and 10 to 20 cm) for saturated hydraulic conductivity (K_s) analysis.

Spectral Data Collection

Soil spectra were collected first using samples in a field-moist state as removed

from the field (field-moist analysis). A second set of spectra were obtained after the

samples were oven-dried (dry analysis). Spectral analysis procedures were similar for

both sets of samples. Spectra were obtained from the deep core samples described above,

but were not available for the fertility or K_s samples. Spectral data from the deep core

samples were matched block-by-block with lab data from fertility and K_s samples for

analysis as described below.

Spectral data were collected using the FieldSpec Pro FR spectrometer (Analytical

Spectral Devices, Boulder, CO) between 350 and 2500 nm at every 1 nm. Three detection

systems within the instrument included: 1) a silicon photodiode array for 350-1000 nm,

2) an InGaAs detector for 1001-1830 nm, and an enhanced InGaAs detector for 1831-

2500 nm. Instrument setup and data collection procedures were adapted from a protocol

used by Viscarra Rossel et al. (2009) to provide consistent procedures for data placed into

an international data base (see Appendix for details). Sample reflectance varied with

depth as shown for an example core in Figure 3.2.
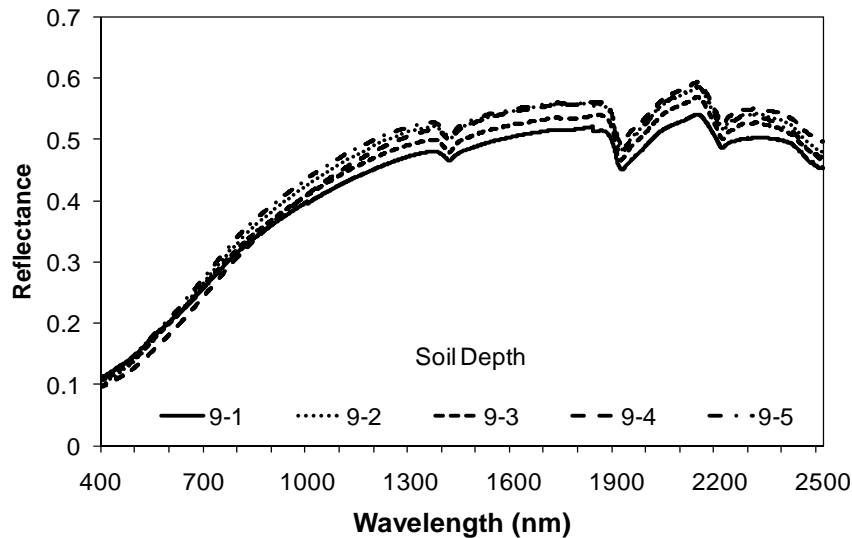
Figure 3.2. Dry soil reflectance spectra for five horizons at sampling site nine.

Illumination was provided by an ASD high intensity source probe (mug lamp).

This is a device that provides an illuminated stage for reflectance data collection from

small samples in the laboratory. In Figure 3.3 the mug lamp is the black cylinder on the

left side of the image. Connected the to the mug lamp are a power cable and a fiber optic

cable that transmits the reflected light to the spectrometer on the right. Once in the

spectrometer on the right, the light is split and directed toward the three detectors, each of which sends light intensity data to the computer where the data is analyzed and converted to reflectance.

Dry samples were illuminated through the bottom of a glass-bottomed dish 33 mm in diameter. For field-moist sample analysis, the mug lamp was inverted to view the top of the sample (Figure 3.3). This was done due to the fact that packing soil into the glass-bottomed dish could not be done in a manner which would allow for a consistent viewable area for each different soil type (see Results and Discussion for more on this topic). There could also be a residue left on the viewing glass by the field-moist soil, which could skew results for the next sample.
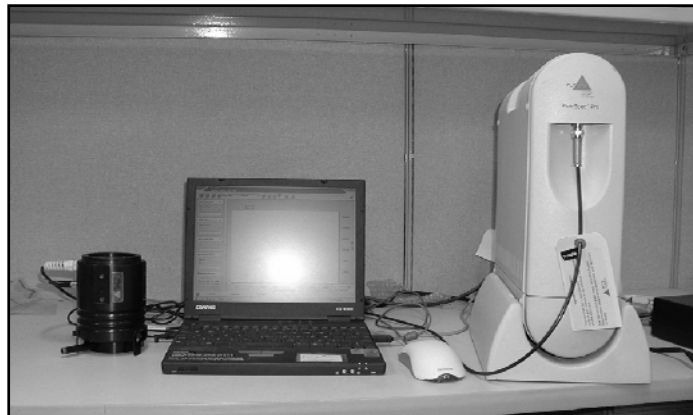
Figure 3.3. Spectral acquisition set-up, including mug lamp (left) and FieldSpec Pro (right).

Data Analysis

Unscrambler v9.8 (Camo Software, Oslo, Norway) was used to view and analyze the collected data. This program is able to take the data from the FieldSpec Pro and create graphs of output statistics, modify the data using mathematical equations, and complete

statistical analysis. After viewing the data, the wavelengths from 350-399 nm were deleted due to excessive noise and an average across 5 nm was done to reduce the number of variables for analysis. The final dataset contained 420 reflectance values for each of 252 samples.

Several preprocessing methods were screened to find the method giving the best estimation fit for the dependent variables. Results from previous research (La et al., 2008) were used to provide a list of candidate preprocessing methods.

1. Raw reflectance
2. Mean Centering (MC)
3. Mean Centering + 1st Derivative
4. Mean Centering + Detrending
5. Mean Centering + Standard Normal Variate (SNV)
6. Mean Centering + SNV + Detrending
7. Absorbance + MC
8. Absorbance + MC + 1st Derivative
9. Absorbance + MC + Detrending
10. Absorbance + MC + SNV

Each method and combination was chosen due to the ability of the preprocessing method to transform the data for better estimation of the desired soil property. Raw data was analyzed to get a baseline reading of the original data to compare to the different transformations used. Mean centering enhances the data by dividing each row of the data matrix by its average so that the influence of hidden factors can be removed. The Savitzky-Golay first derivative performs a least squares regression fit of the polynomial around each point in the spectrum. Then the derivatives are fitted to the polynomial at each point. Standard Normal Variate (SNV) centers and scales each spectrum to remove

the vertical shift from spectrum to spectrum. Detrending linearizes the data by using polynomial regression analysis to remove interactions across the NIR range. Absorbance (ABS) changes the units of reflectance to absorbance, which is simply the logarithm of one over reflectance (Camo, 2006).

Statistical analysis was done using partial least squares (PLS) regression within Unscrambler. This method uses information from all wavelengths in the spectrum to develop a calibration model. It is similar to the principal components regression method, but instead of regressing the dependent variables on the principal components of the independent variables, it creates a linear regression model by projecting predicted and observed variables in a new space. There are principal components (PC) or factors that are created to explain the variation of predicted and response variables. A key part of the PLS analysis is choosing how many PCs to use. Unscrambler chooses this number to minimize the errors in a cross-validation dataset (Lee et al., 2009).

Our basic PLS regression analysis used cross-validation with a split of ten randomly selected subsets. In cross-validation, a subset is removed from the dataset and a calibration on the rest of the data creates a prediction residual for the removed data set. This continues until each data subset has been removed and data prediction residuals are calculated for each subset. From this analysis the root mean squared error (RMSE) and R^2 were obtained. The ratio of prediction error to standard deviation (RPD), which is useful to compare results among datasets with different levels of variability, was then calculated by dividing data standard deviation by the RMSE (Lee et al., 2009). To confirm the cross-validation results, a separate PLS analysis was done with the data set

split into two parts – one-third for validation or prediction and two-thirds for generating the calibration. The remainder of the analysis procedure was similar to the method used for the cross-validation analysis. Within the output of the PLS analysis, beta-coefficients, which are the regression coefficients for the estimated soil property at each particular wavelength are calculated. Those coefficients which were significant (α=0.05) were determined for both the field-moist and dry C analysis.

Laboratory Analysis

After spectral analysis of the wet soil, the entire sample was dried at 105°C for determining gravimetric water content. The samples were then pushed through a 2 mm sieve and roughly 1.0 g was taken as a subsample and ground a using mortar and pestle. The subsample was used for total C and N analysis with a LECO Tru-Spec C/N Analyzer (LECO Corp., St. Joseph, Mich.). Laboratory analysis of POM followed methods of Cambardella and Elliott (1992) and Marriott and Wander (2006). In this method, 25 g dry soil, dispensed in 50-mL centrifuge tubes was shaken with 30 mL of 5% sodium hexametaphosphate for 2 h on a reciprocal shaker at approximately 180 oscillations per min. The suspension was passed through a 53-μm sieve and rinsed with distilled water until a clear solution was obtained. Material retained on the sieve, a mixture of sand and POM, was carefully transferred into flasks and dried at 60°C. The dried material was weighed, ground with mortar and pestle, and analyzed for C (POM-C) and N (POM-N) again using the LECO Tru-Spec C/N Analyzer.

The separate, manually-collected topsoil samples were analyzed for OM, pH, Bray I phosphorous, calcium, magnesium, potassium, and cation exchange capacity

(CEC) by the Soil Testing Laboratory at the University of Missouri (Nathan et al., 2006). K_s analysis was done according to standard procedures described by Reynolds et al. (2002) for constant and falling head measurements.

RESULTS AND DISCUSSION

Descriptive statistics of the laboratory analysis data are given in Table 3.1. The samples were split into two depths of sampling. Laboratory analyses included profile samples collected from the 120 cm hydraulic probe and surface samples collected manually from topsoil probes. The lab data for the top 20 cm of the profile samples was also included with the surface samples. The amount of variation varied widely among the different soil properties. For example pH varied little (CV=0.02) while others such as K_s (CV=2.17) varied greatly.

Table 3.1. Mean, standard deviation, coefficient of variation, and ranges of laboratory analysis data for profile (0 to 120 cm) and surface (0 to 20 cm) soil samples.

		Profile Samples		
Soil property	Mean	Standard Deviation	CV[a]	Range
Soil moisture (%)	26.9	4.2	0.15	22.3
POM-C (g kg^{-1})	20.4	23.7	1.16	133
POM-N (g kg^{-1})	1.08	1.4	1.30	8.80
Total C (g kg^{-1})	10.5	5.8	0.55	23.1
Total N (g kg^{-1})	0.95	0.54	0.57	2.41
		Surface Samples		
Soil property	Mean	Standard Deviation	CV	Range
Ca (kg ha^{-1})	3940	1330	0.34	4726
CEC (meq 100g^{-1})	13.1	5.21	0.40	18.0
K (kg ha^{-1})	210	109	0.52	371
K$_s$ (mm hr^{-1})	59.4	129	2.17	731
Mg (kg ha^{-1})	631	409	0.65	1470
OM (%)	2.9	0.32	0.11	1.7
P (kg ha^{-1})	25.1	15.7	0.63	65.3
pH	5.9	0.13	0.02	0.50
POM-C (g kg^{-1})	48.5	32.1	0.67	125
POM-N (g kg^{-1})	2.8	1.9	0.69	8.03
Total C (g kg^{-1})	18.4	5.63	0.31	16.3
Total N (g kg^{-1})	0.165	0.425	0.26	1.40

[a] CV = coefficient of variation

Evaluation of Preprocessing Methods

The guidelines given by Lee et al. (2009) were used to classify RPD and R^2 into categories describing accuracy of PLS prediction. These accuracy categories were defined as: RPD values of less than 1.5 are poor, values between 1.5 and 2.0 are good, and values above 2.0 are excellent. For the same categories, the R^2 values are less than 0.6, between 0.6 and 0.8 and above 0.8, respectively. Using these guidelines, the analysis of profile soil properties was examined to determine which single preprocessing method performed the best overall. As an example, RPD and R^2 values for field-moist and dry

-36-

analysis of total C were in the excellent category for each preprocessing method (Table

3.2).

Table 3.2. Partial least squares cross-validation statistics for total carbon (C; g kg^{-1}) analyses of field-moist and dry soil samples

Processing Method	Field-moist analysis					
	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation	RPD [a]	PC [b]
Raw	0.85	0.82	2.24	2.45	2.37	10
MC [c]	0.90	0.87	1.82	2.09	2.78	10
ABS [d]	0.89	0.89	1.89	2.11	2.75	10
MC+Deriv [e]	0.92	0.89	1.61	1.92	3.03	7
MC+Dtrend [f]	0.90	0.87	1.83	2.14	2.72	9
MC+SNV [g]	0.88	0.85	1.99	2.24	2.60	10
ABS+MC	0.89	0.86	1.93	2.17	2.68	10
ABS+MC+Deriv	0.91	0.87	1.77	2.35	2.48	8
ABS+MC+Dtrend	0.84	0.80	2.34	2.59	2.25	8
ABS+MC+SNV	0.88	0.85	1.98	2.23	2.61	10
Processing Method	Dry analysis					
	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation	RPD	PC
Raw	0.95	0.94	1.35	1.44	4.02	9
MC	0.96	0.95	1.15	1.23	4.72	9
ABS	0.95	0.94	1.34	1.42	4.08	9
MC+Deriv	0.95	0.95	1.25	1.33	4.36	5
MC+Dtrend	0.97	0.97	1.03	1.10	5.28	8
MC+SNV	0.95	0.94	1.30	1.41	4.14	9
ABS+MC	0.95	0.94	1.33	1.43	4.08	10
ABS+MC+Deriv	0.95	0.94	1.25	1.45	4.00	8
ABS+MC+Dtrend	0.94	0.93	1.40	1.51	3.85	9
ABS+MC+SNV	0.95	0.94	1.36	1.47	3.95	10

[a] RPD = Ratio of predicted deviation
[b] PC= number of principal components used in model
[c] MC = mean centering
[d] ABS = absorbance
[e] Deriv = 1st derivative
[f] Dtrend = detrending
[g] SNV = standard normal variate

The best preprocessing method for field-moist samples was mean centering (MC)

+ derivative and for dried samples was MC + detrending with RPD values of 3.03 and

5.28 respectively. Results from the total C data were representative of what was seen across all soil properties and therefore the results from MC + detrending and MC + derivative were chosen for further investigation with separate validation analysis.

Cross-validation vs. Separate Validation

To affirm that the method of cross-validation was computing reliable predictions, a separate validation analysis was done. For this analysis one-third of the data were randomly selected for validation and the rest were used for calibration. Brown et al. (2005) pointed out that using PLS cross-validation can lead to the over-fitting of data. They also stated that the utility of regression models is for aiding in prediction of future data and that most investigators do not consider this but focus solely on calibrating the data to a single analysis. When comparing the results from PLS cross-validation of the dry data (Table 3.2) against the results of the separate validation set (Table 3.3) there were very few differences. In particular, for MC + detrending, the same number of components was extracted and the difference in RMSE of validation between the two models was only 0.08 g kg^{-1}. Based on this, it can be assumed that there was little over fitting of this data set.

Table 3.3. Partial least squares two-thirds calibration/one-third validation for total carbon (g kg^{-1}) analysis of dry samples.

Processing Method	Carbon					
	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation	RPD [a]	PC [b]
Raw	0.95	0.93	1.27	1.56	3.73	9
MC [c]	0.98	0.96	0.93	1.15	5.03	10
ABS [d]	0.96	0.94	1.22	1.45	4.00	9
MC+Deriv [e]	0.95	0.95	1.28	1.25	4.67	5
MC+Dtrend [f]	0.97	0.96	0.96	1.18	4.91	8
MC+SNV [g]	0.96	0.95	1.19	1.23	4.73	10
ABS+MC	0.97	0.97	1.00	0.99	5.89	6
ABS+MC+Deriv	0.93	0.93	1.34	1.25	4.66	6
ABS+MC+Dtrend	0.96	0.94	1.15	1.26	4.63	7

[a] RPD = Ratio of predicted deviation
[b] PC = number of principal components used in model
[c] MC = mean centering
[d] ABS = absorbance
[e] Deriv = 1st derivative
[f] Dtrend = detrending
[g] SNV = standard normal variate

Although results were slightly less accurate than the cross-validation results above, the MC + derivative and the MC + detrending pretreatments were within the top few methods based on RPD values (Table 3.3). Combining the results from the cross-validation and the independent validation analyses, we concluded that the best overall data preprocessing method was MC + detrending and therefore this pretreatment was used for the remainder of the analyses.

Profile Soil Analysis

Full profile soil analysis examined the soil properties of POM-C, POM-N, total C, total N, and moisture (Table 3.4). All of these soil properties showed RPD and R^2 values at or above the range of "good" estimation. In particular, the accuracies of total C and total N estimations were quite high. Total C had R^2 values of 0.90 and 0.97 of for field-

moist and dry analysis respectively, while total N had values of 0.82 and 0.91 of R^2 for

field-moist and dry analysis, respectively. Accuracy values of total C ($R^2 = 0.90$ and 0.97

for field-moist and oven dry analysis, respectively) were much better than those found by

Morgan et al. (2009) who also compared the accuracy of field-moist and dry soil property

estimation and found $R^2 = 0.77$ and 0.85 for field-moist and dry analysis, respectively.

Moisture and POM-C and N also remained above the good estimation bench mark ($R^2 =$

0.83 moisture, $0.75 \leq R^2 \leq 0.80$ POM-C, and $0.77 \leq R^2 \leq 0.80$ POM-N). It is likely that

these soils have high accuracies of estimation due to the high variability of soil properties

within the data set (Table 3.1) and large number of soil samples (252).

Table 3.4. Partial least squares cross-validation statistics for analyses of field-moist and dry whole profile soil samples for moisture, particulate organic matter (POM)- carbon (C), POM-nitrogen (N), total carbon (C) and total nitrogen (N).

	Field-moist Analysis					
Soil Property	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation	RPD [a]	PC [b]
Moisture (%)	0.83	0.81	1.70	1.80	2.30	7
POM-C (g kg^{-1})	0.75	0.70	11.9	13.0	1.83	8
POM-N (g kg^{-1})	0.77	0.72	0.67	0.73	1.90	9
Total C (g kg^{-1})	0.90	0.87	1.83	2.14	2.72	9
Total N (g kg^{-1})	0.82	0.78	0.23	0.26	2.11	8
	Dry Analysis					
Soil Property	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation	RPD	PC
POM-C (g kg^{-1})	0.81	0.79	10.3	10.9	2.16	10
POM-N (g kg^{-1})	0.80	0.77	0.62	0.66	2.11	10
Total C (g kg^{-1})	0.97	0.97	1.03	1.10	5.28	8
Total N (g kg^{-1})	0.91	0.89	0.17	0.18	3.04	8

[a] RPD = ratio of predicted deviation
[b] PC = number of principal components used in model

The dry samples had higher R^2 and RPD values than the field-moist samples, most likely due to the fact that the dry samples were much more consistent in aggregate size causing less noise in the reflectance data. While samples used for the field-moist analysis were sieved to attain consistent size distribution, the problem faced was that soil from greater depths contained more clay, leading to aggregates with gaps, which in some cases required pressing into a flat, smooth surface. However, soil from the shallower depths was generally low in clay and did not form large consolidated aggregates. This difference (Figure 3.4) could have led to less consistent spectral scans of the wet samples, and less reliable results.
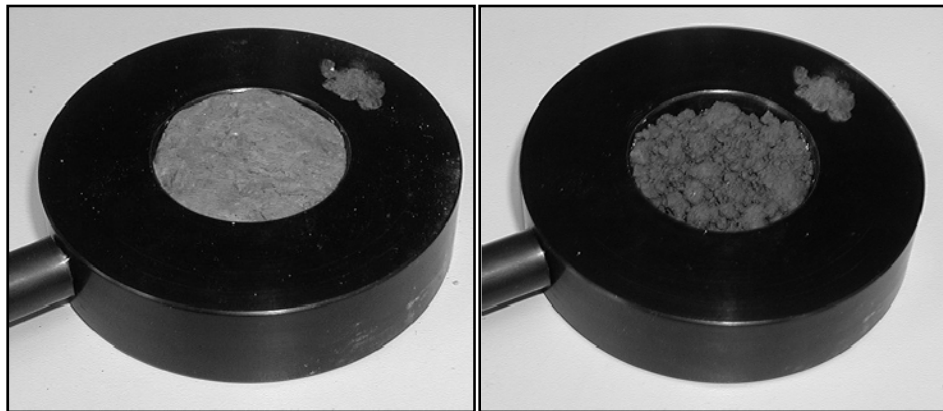
Figure 3.4. Glass bottomed dish containing different prepared soils. Image on the left is typical of soil from greater depths and on the right is of soil from shallow depths.

The MC+detrending cross-validation results for soil moisture, POM-C, and total

C and from the Unscrambler PLS analysis are shown in scatter plots (Figure 3.5). The

graph illustrates very little variation between the calibration points and the validation

points. The goal of each calibration technique is to fit the data without over-fitting. The

validation points are those points which are left out of calibration for a check to ensure

that the data has not been over-fit. Other than a few outliers, the calibration and

validation points match up well, meaning there was little over-fitting of the data during

PLS calibration. In the lab analysis of POM-C there were points that had low values,

some almost zero, which may have been the reason some predictions of POM-C were in

the negative range.
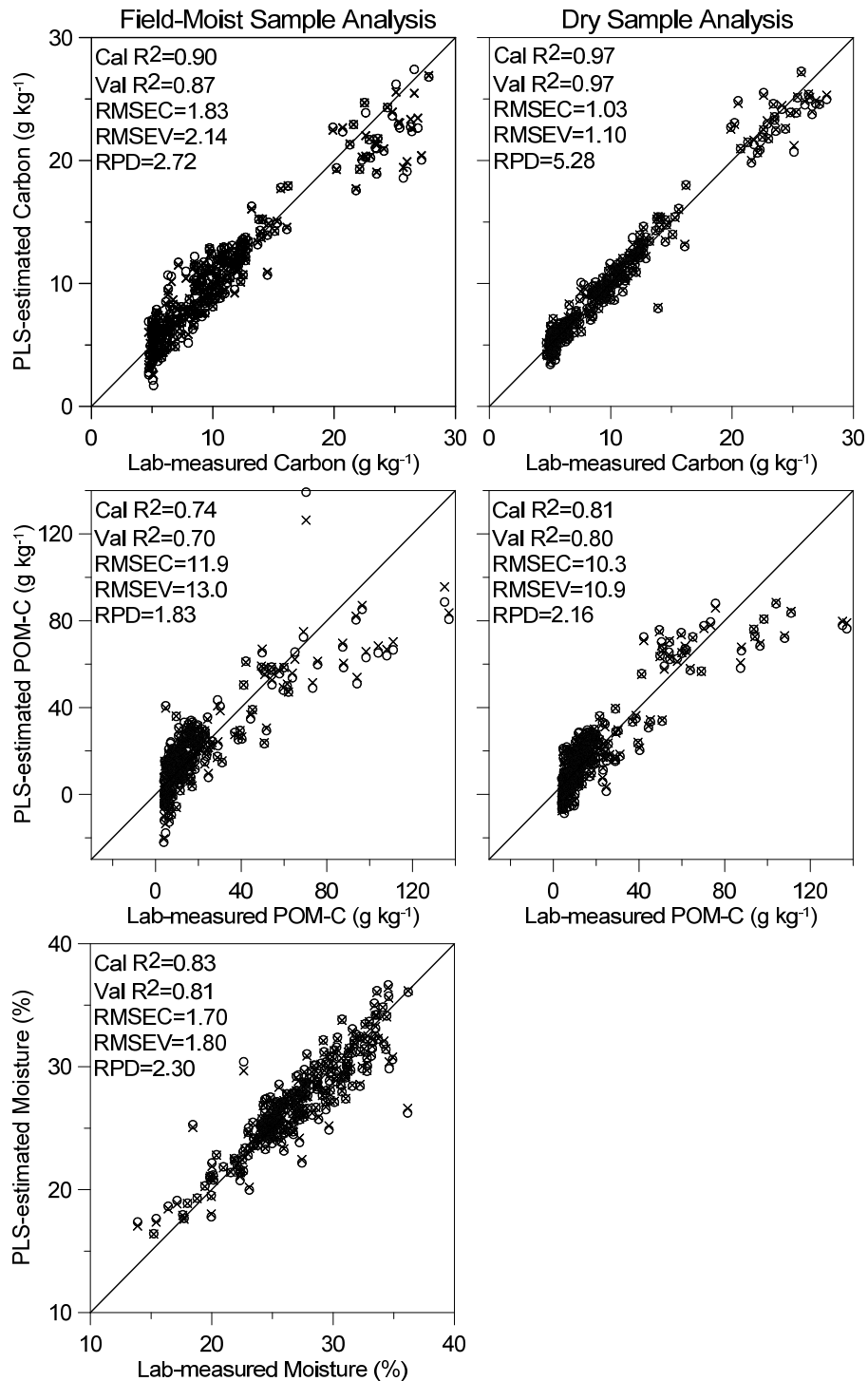
Figure 3.5. Soil property estimates of field-moist (left) and dry (right) samples by partial least squares (PLS) analysis of spectral data as compared to those obtained by standard laboratory analysis. Both calibration (Cal; x) and validation (Val; o) results are included as well as ratio of predicted deviation (RPD) for carbon, (C) particulate organic matter-C (POM-C), and soil moisture.
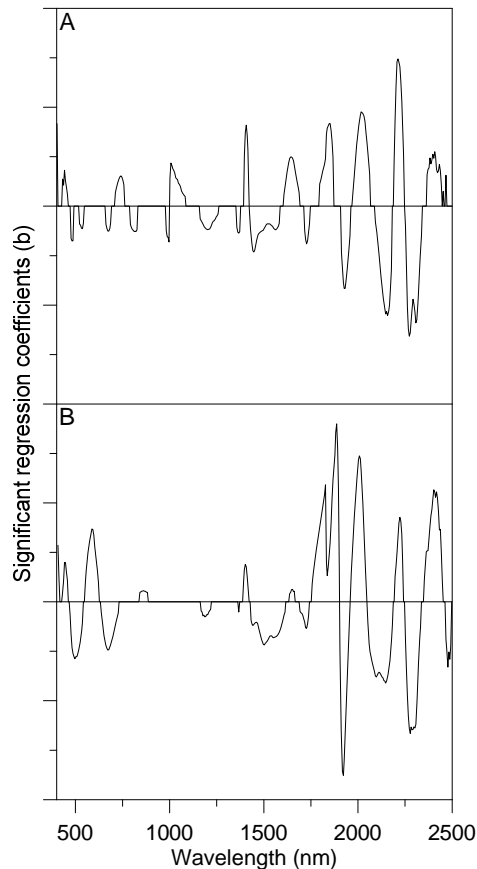
Figure 3.6. Wavelengths that contributed to significant regression coefficients (P ≤ 0.05) in the prediction of carbon are shown for field-moist (A) and dry (B) analysis. The magnitude of coefficients indicates the relative importance of that wavelength. Values on the y-axis are not shown but both graphs are on the same scale.

The beta-coefficients determined in PLS analysis show the relative importance of each wavelength in estimating the variable in question. From the field-moist analysis of total C (Figure 3.6A), the NIR range was more important for estimating total C than the VIS range. Clark et al. (1990) found that the wavelengths 2120 to 2160 nm, 1970 to 2000 nm, and 1850 to 1870 nm were useful in the prediction of total C. These wavelength ranges were included in the wavelengths and found to be significant in this analysis. Morgan et al. (2009) performed a similar analysis, finding that the NIR region had greater

significant wavelengths with total C than the VIS region, specifically at wavelengths around 2250 nm. Also, Lee et al. (2009) determined that wavelengths of 1877, 2138, 2333, 2381, and 2474 contributed to significant regression coefficients in prediction of total C. Larger peaks can be found in the dry as compared to field-moist analysis similar to what was found by Tekin et al. (2012). When comparing the field-moist analysis (Figure 3.6A) to the dry analysis (Figure 3.6B), there were 18 more wavelengths which were significant for the estimation of total C in the dry analysis, a factor which could help explain the increase in accuracy seen from field-moist to dry analysis.

Surface Soil Analysis

The PLS estimations for the surface soils were not quite as accurate as for the profile soils. Only a few soil properties were at or above the "good" estimation range for RPD and R^2. The most accurately estimated properties were CEC ($0.57 \leq R^2 \leq 0.74$ and $1.59 \leq RPD \leq 1.72$) and Mg ($0.74 \leq R^2 \leq 0.83$ and $1.79 \leq RPD \leq 2.10$) (Table 3.5). On the other hand, soil properties such as pH and OM were estimated quite poorly. This is in contrast to many other studies where pH and OM have been well-estimated by reflectance spectroscopy. However, in this study, the variability in pH and OM was quite small (Table 3.1), which likely contributed to the poor estimation.

Calibrations were developed for POM-C and N, total C, and total N using the soil spectral and lab data from only the top 20 cm for comparison to the surface data for a total of 64 data points. Although estimates were not as accurate as the full profile analysis, they were still better than for the other surface soil properties. Within this data set total C was the best estimated with values of 0.89 for R^2 and 2.45 for RPD in the dry

analysis (Table 3.5). POM-C and POM-N also had good estimation accuracy with an R^2 of 0.82 and 0.80, respectively. Compared to Sudduth et al. (2010; $R^2 = 0.31$ and RMSE = 2.8 g kg^{-1}) our estimation of POM-C was much better as measured by R^2 but produced a similar RMSE (R^2=0.82 and RMSE=3.4 g kg^{-1}, respectively). A factor contributing to the better results for these four properties may be that the same samples were used for spectral analysis and laboratory analysis, while lab data and spectral analysis for the other surface properties were obtained from separate samples.

Table 3.5. Partial least squares cross-validation statistics for analyses of field-moist and dry surface soils <20 cm for Calcium (Ca), cation exchange capacity (CEC), potassium (K), saturated hydraulic conductivity (K_s), Magnesium (Mg), organic matter (OM), phosphorus (P), pH, particulate organic matter (POM)- carbon (C), POM-nitrogen (N), and total C and N.

| | Field-moist Analysis | | | | | |
Soil Property	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation	RPD [a]	PC [b]
Ca (kg ha^{-1})	0.54	0.52	885	910	1.46	1
CEC (meq 100g^{-1})	0.57	0.54	3.77	3.90	1.50	2
K (kg ha^{-1})	0.49	0.41	76.6	82.7	1.32	3
K_s [c] (mm hr^{-1})	0.28	0.21	1.04	1.10	1.08	6
Mg (kg ha^{-1})	0.74	0.68	206	228	1.79	3
OM (%)	0.03	0.05	0.39	0.40	0.88	1
P (kg ha^{-1})	0.48	0.40	11.1	11.9	1.31	1
pH	<0.01	<0.01	0.30	0.30	0.44	1
POM-C (g kg-1)	0.71	0.58	17.1	2.10	1.53	6
POM-N (g kg-1)	0.70	0.56	1.04	1.27	1.50	6
Total C (g kg-1)	0.72	0.62	2.96	3.45	1.63	5
Total N (g kg-1)	0.59	0.49	0.28	0.31	1.36	4
	Dry Analysis					
Soil Property	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation	RPD	PC
Ca (kg ha^{-1})	0.42	0.38	1010	1040	1.28	1
CEC (meq 100g^{-1})	0.74	0.65	2.95	3.40	1.72	4
K (kg ha^{-1})	0.73	0.63	55.9	65.3	1.67	5
K_s [c] (mm hr^{-1})	0.37	0.29	0.95	1.02	1.17	7
Mg (kg ha^{-1})	0.83	0.77	164	193	2.10	5
OM (%)	0.28	0.24	0.30	0.32	1.13	2
P (kg ha^{-1})	0.27	0.19	14.1	15.0	1.04	3
pH	0.07	0.08	0.43	0.45	0.30	2
POM-C (g kg^{-1})	0.82	0.75	13.5	16.2	1.99	6
POM-N (g kg^{-1})	0.80	0.66	0.84	1.01	1.88	6
Total C (g kg^{-1})	0.89	0.83	1.83	2.30	2.45	6
Total N (g kg^{-1})	0.69	0.61	0.24	0.27	1.55	5

[a] RPD = ratio of predicted deviation
[b] PC = number of principal components used in model
[c] PLS analysis statistics for K_s were based on the logarithm of raw values (i.e. log K_s)

Surface pH had a CV of 0.022 (Table 3.1), the smallest variation of all properties tested. It also had the smallest R^2 and RPD (Table 3.5) and the data does not follow the 1:1 fit line (Figure 3.7). Although pH was poorly estimated, the RMSE of 0.45 was quite low. Estimates of pH reported by Chang et al. (2002) showed much greater R^2 than these data. However, the RMSE reported (0.56) was greater than the error we found. By comparing these results we surmise that our pH estimation was limited by the small range present in the data. On the other hand, surface CEC was estimated quite well, and much better than pH (Figure 3.7). The trend of CEC generally followed the 1:1 fit line and there were only a few outliers.

Figure 3.7. Soil property estimates of field-moist (left) and dry (right) samples by partial least squares (PLS) analysis of spectral data as compared to those obtained by standard laboratory analysis. Both calibration (Cal; x) and validation (Val; o) points are included as well as ratio of predicted deviation (RPD) for pH and cation exchange capacity (CEC).

Saturated hydraulic conductivity (K_s) values from lab analysis exhibited a very large range (731 mm h^{-1}; Table 3.1) and a highly skewed distribution (Figure 3.8). On similar soils Jiang et al. (2007) and Mudgal et al. (2010) found high values of K_s (125.6 and 316 mm hr^{-1}, respectively) when testing soils from conservation reserve program (CRP) and prairie land. Due to the fallow period at this location it would likely behave similarly to CRP management and also exhibit high K_s values. Extremely high K_s values

might also be attributed to macropores not being plugged for laboratory analysis because they were not present at the soil surface or because they were hidden by soil aggregates. To reduce the skewness of the distribution, the logarithm of the K_s values was taken prior to analysis (Figure 3.8). Most of the raw values of K_s were below 100 with a few extending beyond whereas the log (K_s) data are much more normally distributed. This point is reinforced by the reduction in CV from raw to logarithmic transformed data (2.17 vs, 1.61). However, even after taking the logarithm estimations of K_s PLS analysis still showed poor accuracy for both field-moist and dry analysis ($R^2 = 0.21$ and 0.29, respectively; Table 3.5).
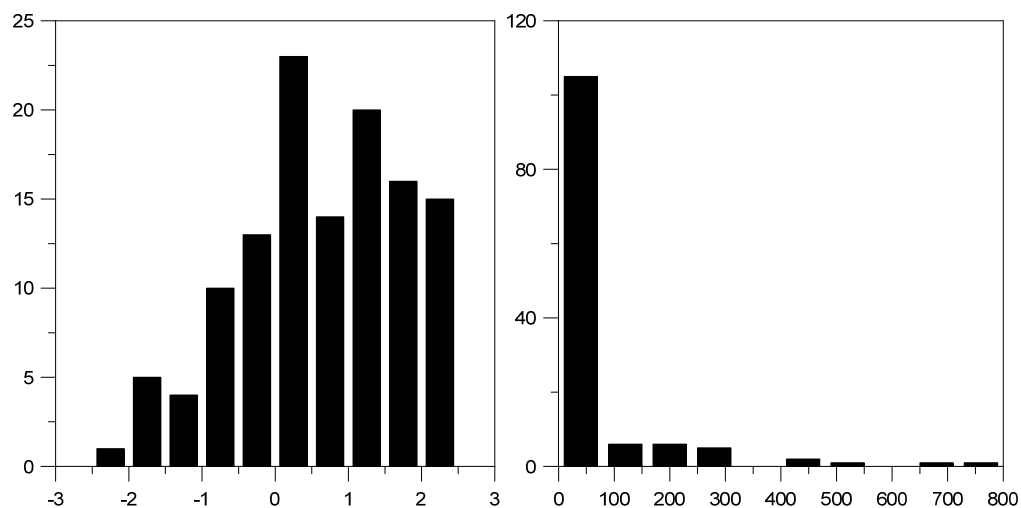
Figure 3.8. Histogram of the log of saturated hydraulic conductivity ($K_{s)}$ (left) and raw values for K_s (right)

Few researchers have attempted to estimate K_s with reflectance data, likely due to the absence of a direct correlation. Although accuracies of K_s data are low, soil texture (sand, silt, and clay), which has been estimated by several researchers with varying success (Sudduth et al., 2010; Lee et al., 2010; Janik et al., 1998), has been shown to have

an impact on K_s (Kutilek, 2004; Rawls et al., 1993). Like the other soil properties

estimated in this study, the dry analysis had greater accuracy than field-moist analysis.

Even after the logarithmic transformation was applied to K_s data, the prediction points

were not well correlated with the lab-measured points (Figure 3.9). There are two

potential causes for this, the first being that samples analyzed for K_s were not the same

samples used for spectral analysis. Secondly, since K_s does not have direct correlation to

reflectance data, any relationship would be based upon collinear relationships.

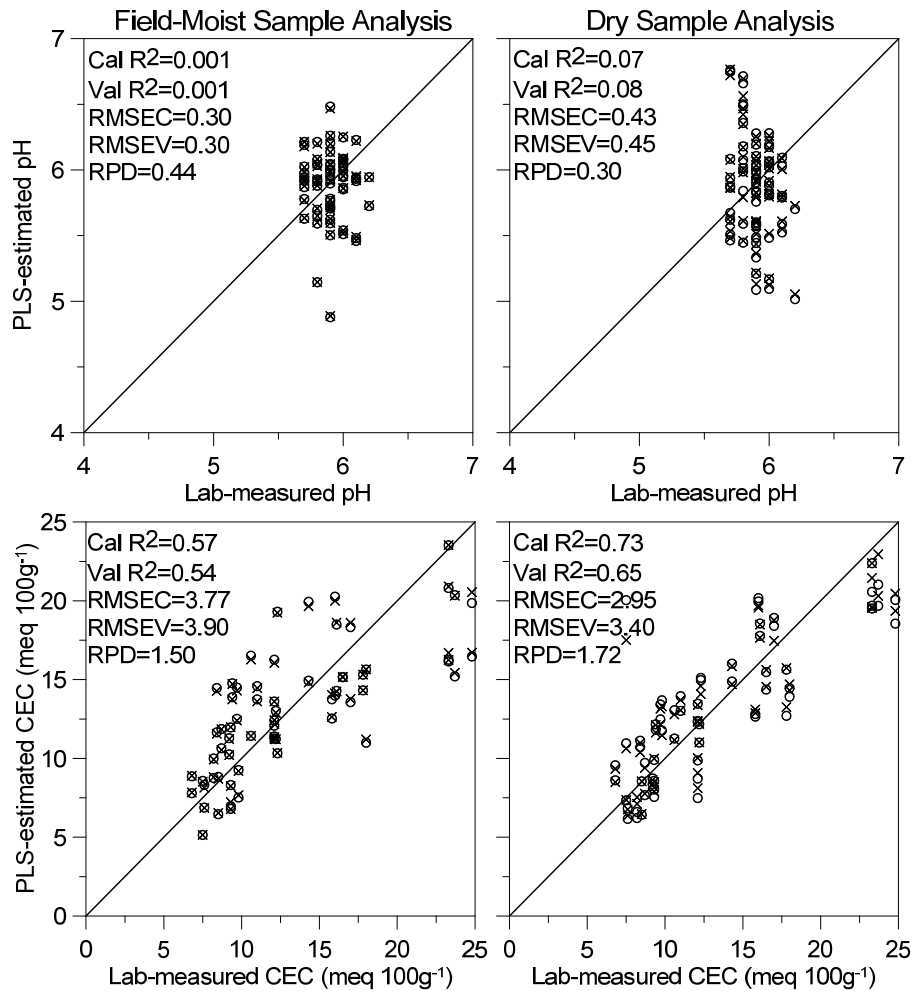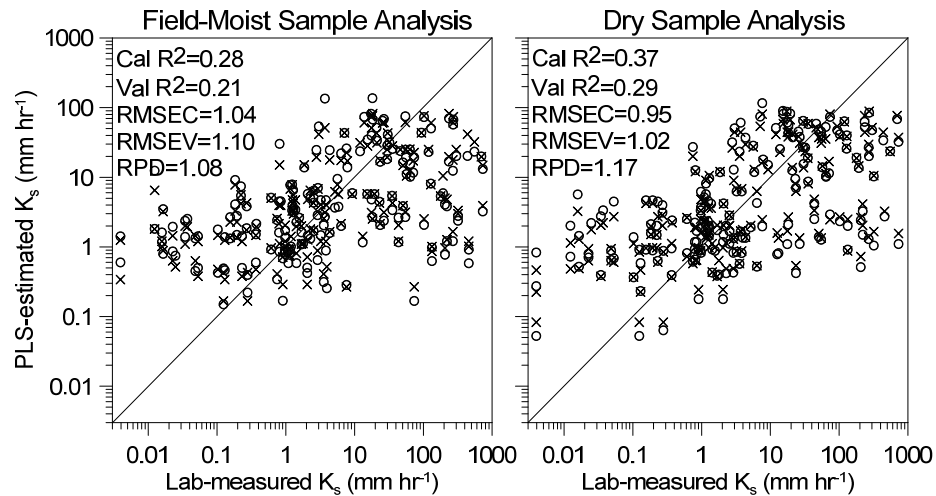Figure 3.9. Saturated hydraulic conductivity (K_s) estimates for field-moist (left) and dry (right) samples by partial least squares (PLS) analysis of spectral data as compared to those obtained by standard laboratory analysis. Prediction points of both calibration (Cal; x) and validation (Val; o) are included as well as ratio of predicted deviation (RPD) and other estimation statistics.

For some soil properties, reflectance spectroscopy proved to be a good tool for the

soil characterization testing needed in soil management decisions for improving or

maintaining soil quality. Several soil quality parameters were well-estimated including

total C, total, N, and moisture for both field-moist and dried full-profile samples and were

somewhat less accurate when considering only surface soil samples. The field-moist

samples proved to be slightly more difficult to analyze because of problems such as variation in aggregate size of different samples. Also, special care was needed in handling the samples to maintain their physical and chemical properties, such as keeping the samples cool and in air-tight containers. Another factor was that when placing the samples into the glass bottomed dish to be viewed by the ASD spectrometer, each soil sample and soil type required different techniques in preparing a consistent surface. These factors limit the speed and effectiveness of spectral analysis in a field-moist condition. In future research it would be necessary to resolve these issues so that spectral analysis of field-moist soil better correlates with the lab data, thereby removing the need for dry analysis.

CONCLUSION

A full range spectrometer was evaluated for the prediction of both surface and subsurface soil properties. Data from the spectrometer required mathematical preprocessing to improve prediction results. Considering the results of several tests comparing pretreatment combinations, the most accurate preprocessing method was mean centering + detrending. From our analysis, the results for total C were the best overall, with R^2 for both field-moist and dried samples at or above 0.90 for full profile analysis. From the surface analysis and leaving out POM-C and N, and total C and N, Mg had the best estimation with a value for R^2 of 0.83 in dry analysis. The NIR range was found to have greater significance for the prediction of total C in both field-moist and dry analysis, similar to what past researchers have found.

When comparing the field-moist sample analysis to dry sample analysis, reduced accuracy was noted with the field-moist samples, but for most of the samples only a slight reduction was observed. Based on this research, drying and sieving soil for spectral analysis may not be necessary if a relatively small decrease in accuracy of the estimation of soil physical and chemical properties can be tolerated. Particularly in situations where more samples were taken and there was more variability within the data than seen in the surface soils of this study, reflectance spectroscopy could provide a valuable tool to assist in soil quality assessment.

CHAPTER 4
SURFACE SOIL ORGANIC CARBON ESTIMATION USING AN ACTIVE
LIGHT REFLECTANCE SENSOR

ABSTRACT

The use of reflectance sensing as a tool for site specific estimation of soil properties, specifically soil organic carbon (SOC), has shown varied results. The objective of this research was to evaluate a prototype two-band reflectance sensor both within field and lab settings and compare the results to those from a spectrometer. Two sites were chosen for study based on within-field visual texture variations and measured organic matter (OM) variation maps. During field measurement, 20 field locations were chosen for soil sample collection and sensor measurements with residue completely raked away. Lab reflectance measurements, taken with the spectrometer and two-band sensor, used the field-collected soils. Of the different data sources, the best prediction accuracy for SOC was with the two-band sensor ($R^2 = 0.75$) using data taken within the field. While the field-based SOC estimates were the most accurate, when the prediction equation was applied to full-field sensor measurements, results were varied. One field was well predicted with a root mean squared error (RMSE) of 6.9 g kg^{-1} but the other field had a RMSE of 18.4 g kg^{-1}. This difference could be attributed to the raked areas not being indicative of the overall variation seen within the entire field. To determine optimum two-band indices for SOC estimation, various combinations of bands from the spectrometer analysis were examined. There was high correlation with SOC in the near infrared range, specifically when combining wavelengths greater than 1800 nm with

wavelengths near 1400 and 1900 nm. When only considering combinations of wavelengths from <1000 nm, the highest correlation to SOC was when both were from the visible wavelength range. Some combinations were consecutive wavelengths, though it is unlikely that these combinations would be chosen to create a two-band sensor.

INTRODUCTION

Maintaining soil quality is important for achieving high yields within an agricultural system. A secondary benefit to maintaining soil quality is that it may also help to maintain water and air quality (Kennedy and Papendick, 1995). Several farming practices said to improve soil quality, such as no-till or conservation tillage, crop rotation, and reduced chemical inputs, exist but methods for determining if soil quality has improved when farming practices are changed have lagged (Arshad and Martin, 2002). A major contributing factor to this problem is the time and cost associated with traditional soil tests. Due to the high spatial variation that can be seen in soil properties such as soil organic matter (SOM) (Jung et al., 2006), many soil samples would be required to adequately describe variability. Therefore, the amount of sampling and analysis required is prohibitive for most production fields.

One method that has been introduced to estimate soil properties more rapidly is reflectance spectroscopy. Light from the visible (VIS; 400 to 700 nm) and near infrared (NIR; 700 to 2500 nm) wavelength ranges is directed at the soil in question to estimate soil properties (Janik et al, 1998; McCarty et al., 2002; Viscarra Rossel et al., 2006; La et al., 2008; Lee et al., 2009; Sudduth et al., 2010). Energy from these wavelengths excites molecules within the soil constituents and is detected by the reflectance spectroscopy

sensor (Murray and Williams, 1987). Much research has been done on the estimation of soil carbon (C). Within the VIS range, absorption occurs by chromophores and the darkened color of SOM. Within the NIR range, many researchers have shown different wavelengths to be useful for prediction of soil C (1100, 1600, 2000, and 2200 to 2400 nm; Dalal and Henry, 1986; Clark, 1999; Martin et al., 2002; Stenberg, 2010).

One of the drawbacks to reflectance spectroscopy is that its ability to be used as a site specific management tool is somewhat unproven. While the use of reflectance spectroscopy for estimating soil organic carbon (SOC) in wide geographic areas is well documented by the review done by Viscarra Rossel et al. (2006), results from field-scale research have varied. Udelhoven et al. (2003) found that SOC was not well predicted in soil that was obtained from a field scale study. Similarly, Sudduth et al. (2010) found that when the SOM range was limited (CV=16.4%) prediction accuracy was poor ($R^2 = 0.31$) even with an RMSE (2.8 g kg^{-1}) which was lower than studies which boasted better R^2 (RMSEP 3.0 g kg^{-1}, Viscarra Rossel et al., 2006). Conversely McCarty and Reeves (2006) found good estimation of SOC ($R^2 = 0.95$) and nitrogen (N; $R^2 = 0.93$) within a field scale study. Based on past research, the ability of reflectance spectroscopy to accurately quantify within-field soil variation as a site specific management tool remains unresolved.

Much of the past reflectance sensing work has been with sensors which measured reflectance across a full spectrum of light in either the VIS or NIR or both. However a few studies have used a limited spectrum for the purpose of in-situ soil property estimation. Sudduth and Hummel (1993a, 1993b) developed and tested a portable sensor

for analyzing SOM, cation exchange capcity (CEC), and moisture on-the-go using wavelengths from 1700 to 2420 nm. When testing this sensor in the lab, they found the best accuracy ($R^2 = 0.89$) when 100 scans were averaged together per reading. In the field though, SOM was not well predicted due to the movement of the sensor in reference to the targeted soil during acquisition of spectral scans.

Christy (2008) and Bricklemyer and Brown (2010) used a commercial device (Veris VIS-NIR shank; Veris Technologies Inc., Salina KS) mounted to an agricultural shank and pulled behind a tractor. This device acquired data at a rate of 20 scans/s between 350-2224 nm with an 8-nm resolution. Christy (2008) found high prediction accuracy with an R^2 of 0.92 and 0.80 when testing for C and SOM, respectively. Bricklemyer and Brown (2010) found very poor accuracy when predicting SOC ($R^2 =$ 0.00). They suggested that low prediction accuracy was most likely attributed to low variation in SOC within their test soils as standard error of prediction (SEP) values (3.5 g kg^{-1}) were similar to those of Christy (2008) (3.0 g kg^{-1}).

Roberts et al. (2011) examined a 2-band sensor originally developed to sense crop canopy reflectance to guide nitrogen (N) application (Raun et al., 2005; Solari et al., 2008; Kitchen et al., 2010). The two bands in this sensor were 590 ± 5.5 nm and 880 ± 10 nm. For soil sensing, it was mounted to an ATV which followed behind a planter. To estimate SOM, the researchers used a linear prediction model that incorporated high, medium and low values of lab measured SOM and the sensor output measured in the field. These variables were used in equations for site-specific and universal SOM

estimation models. They found prediction accuracies as high as $R^2 = 0.88$ with a model which was site-specific.

OBJECTIVES

The overall objective of this study was to test a two-band reflectance sensor and determine its viability as a tool for measuring in-field variation of surface soil properties for site-specific management.

Specific objectives were to:

- Evaluate the capabilities of the two-band sensor both in a lab and field setting for predicting SOC, soil N, and soil moisture.

- Compare the lab results from the two-band soil sensor to those from a full-range spectrometer.

- Create field maps of predicted SOC.

- Using data from full spectrum spectroscopy, evaluate whether the wavelengths used in the prototype sensor were those which provided the best correlation with SOC.

MATERIALS AND METHODS

Test Sites

Two field sites were chosen for study based on visual and tactile texture differences and variation in measured SOM as shown on 1 ha grid soil maps. The two sites chosen were located on flood plain soils in Chariton County, Missouri (Table 4.1).

Table 4.1. Site information including location, area, and soil data.

Field	Location	Area (ha)	Predominant Soil
1	39° 61' N 92° 71' W	119	Norborne Loam; mesic Typic Agriudolls
2	39° 30' N 92° 92' W	23	Tice Silt Loam; mesic Fluvaquentic Hapludolls

Field Reflectance Sensing

      For field use, a rig was designed to carry a prototype non-contact two-band sensor developed by research collaborators through the field. Knowing that sensor measurements of soil could be greatly affected by residue, a row cleaner was placed ahead of the sensor. After testing this configuration, it was observed that during windy conditions residue could be blown back into the path of the sensor, and also that the row cleaner left tooth marks in the soil which created an uneven surface. Therefore side fenders and a drag chain were added to remedy these two issues (Figure 4.1). Both fields had soybean residue remaining from the previous year's crop. This was advantageous because in initial tests it was found that corn residue was not easily cleared by a row cleaner for creating a clear view of soil for the sensor.

Figure 4.1. Image of field sensing including two-band sensor (left) and sensor rig (right).

Field data were collected in January and February of 2012 using the two-band reflectance sensor. This sensor was designed to collect data at 660 and 850 nm from a nadir view at 15 to 25 cm above the soil surface. During field data collection the sensor was roughly 18 cm above the soil surface which provided sensed area of about 21 cm^2. Data was logged at a rate of 10 Hz with a portable data logger which collected both reflectance and GPS data. GPS position was obtained from a Raven Invicta 210 (Raven Industries, Sioux Falls, SD) capable of sub-meter accuracy. Distance between consecutive passes was about 10 m. Initial collected data was viewed and soil sensing and soil collection points within the field were designated by viewing the variation in reflectance data across the field. In these areas a 1 m^2 section was completely raked of

residue and the sensor was driven slowly to provide more than 120 readings per raked

area. At each site roughly 3 L of soil was collected from the top 3 cm of soil in the 1 m^2

raked area, to be analyzed by traditional lab tests and lab spectrometer sensing. A total of

20 samples were collected. Soil contained in air-tight zip-top bags was stored at 4°C to

maintain chemical and physical characteristics before analysis.

Lab Reflectance Sensing

Spectral and reflectance data were collected in a lab setting using the two-band

sensor described above and the FieldSpec Pro FR spectrometer (Analytical Spectral

Devices, Boulder, CO). Field-moist soil was placed onto trays which were 14x10 cm to a

depth of >1 cm. Three measurements were taken per sample in three different areas of the

tray with both the two-band sensor and spectrometer. Reflectance measurements with the

two-band sensor were taken at a rate of 10 Hz for 20 s, providing 200 measurements.

These measurements were taken at a height of 18 cm (Figure 4.2, top), similar to the

height during field sensing. Immediately after measurements were taken with the two-

band sensor, spectrometer measurements were taken in the same locations on the tray.

The spectrometer captured reflectance data between 350 and 2500 nm at every 1 nm.

Three detection systems within the instrument included: 1) a silicon photodiode array for

350-1000 nm, 2) an InGaAs detector for 1001-1830 nm, and an enhanced InGaAs

detector for 1831-2500 nm. Illumination was provided by an ASD high intensity source

probe (contact probe; Figure 4.2, bottom) which has an optical footprint diameter of 2

cm. At the beginning and after every third measurement, the spectrometer was optimized

with a Spectralon panel (Labsphere, North Hutton, NH), which is 99% reflective. This

panel was also used as a white reference before scanning each soil. The contact probe was placed directly onto the soil surface and by doing so, ambient light was blocked from entering during measurement.

Figure 4.2. Lab set-up of two-band sensor (top) and spectrometer (bottom).

After field moist soil reflectance was collected, sub-samples were then dried at 105ºC and pushed through a 2 mm sieve in preparation for additional reflectance analysis. Dry samples were illuminated by an ASD high intensity source probe (mug lamp) with a 2 cm diameter optical footprint through the bottom of a glass-bottomed dish. Instrument setup and data collection procedures were adapted from a protocol used by Viscarra Rossel et al. (2009) to provide consistent procedures for data placed into an international data base (see Appendix for details).

LABORATORY ANALYSIS

During preparation for dry spectral analysis roughly 1.0 g was taken as a subsample and ground using a mortar and pestle. The subsample was used for C and N analysis with a LECO Tru-Spec C/N Analyzer (LECO Corp., St. Joseph, Mich.). Soils were also weighed before and after drying to calculate gravimetric water content.

DATA ANALYSIS

The spectrometer data were viewed and analyzed using Unscrambler X v10.1 (Camo Software, Oslo, Norway). This program is able to take the data from the FieldSpec Pro and create graphs of output statistics, modify the data using mathematical equations, and develop and validate statistical models for estimation of the dependent variables. Two data sets were analyzed with partial least squares (PLS) regression. The first was the full-spectrum output and the second was limited to data within the first detector range (350 to 1000 nm). In both data sets the wavelengths from 350 to 450 nm were deleted due to excessive noise and data were averaged across each five nm to reduce the number of variables for analysis. Each of the data sets contained a different number of reflectance values per sample. Also dry reflectance measurements were repeated three times and not averaged together providing a total of 60 reflectance measurements. There were 111 reflectance values in the first detector analysis for 20 (field-moist) and 60 (dry) samples and 411 reflectance values in the full spectrum analysis for 20 (field-moist) and 60 (dry) samples.

Statistical analysis of the full spectrum and first detector data was done using partial least squares (PLS) regression within Unscrambler. This method uses information from all wavelengths in the chosen spectrum to develop a calibration algorithm (Lee et al.,

2009). It is similar to the principal components regression method, but instead of regressing the dependent variables on the principal components of the independent variables, it creates a linear regression model by projecting predicted and observed variables in a new space. There are principal components (PC) or factors that are created to explain the variation of predicted and response variables. A key part of the PLS analysis is choosing how many PCs to use. Unscrambler chooses this number to minimize the errors in a validation, or cross-validation, dataset (Lee et al., 2009).

The PLS regression analysis used cross-validation with a split of 10 randomly selected subsets. In cross-validation, a subset is removed from the dataset and a calibration on the rest of the data creates a prediction residual for the removed data set. This continues until each data subset has been removed and data prediction residuals are calculated for each subset. From this analysis the root mean squared error (RMSE) and coefficient of determination (R^2) were obtained. Ratio of prediction error to standard deviation (RPD), which is useful to compare results among datasets with different levels of variability, was then calculated by dividing data standard deviation by the RMSE (Lee et al., 2009). Values of RPD above 2.0 and R^2 above 0.8 have been considered of good accuracy in soil analysis (Chang et al., 2001).

The two-band data were analyzed using Proc STEPWISE in SAS v9.2 (SAS Institute, Cary, NC). This procedure creates a model for a chosen dependent variable based on given estimators. In each step a model is created with the best estimators to represent the chosen variable. The first model is created with one estimator which is assigned a slope and intercept, and this process continues until all estimators have been entered into the

model or a maximum R^2 has been reached. For this analysis the estimators were NIR (850 nm), VIS (660 nm), NIR^2. VIS^2, and NIR x VIS (NIRVIS) and the modeled variables were and SOC, N, and moisture. Two band analysis was applied to both active sensor data and spectrometer data averaged over the same wavelengths (660 ± 5.5 and 850 ± 5.5 nm). Models were chosen based on highest R^2 with all estimators significant at a level of $\alpha = 0.05$.

To determine if there were any combinations of two wavelengths that would have been more useful than those output by the two-band sensor, the correlation between reflectance and SOC at different wavelengths was computed. The first analysis determined the correlation of reflectance at a single wavelength with SOC. This analysis was not very useful as it produced a broad curve with no well-defined wavelengths with greater correlation to SOC than the others (Figure 4.3). Based on this result, further analysis was done using indices calculated from sensor data at two different wavelengths. Each of the two wavelengths was varied over two different ranges, from 450 to 1000 nm and from 450 to 2500 nm averaged in 10 nm increments. The two indices used, the difference of reflectance from wavelength A and wavelength B (A-B) and the ratio reflectance of A and B (A/B), were chosen from a group of indices based on having relatively high correlations and simplicity.
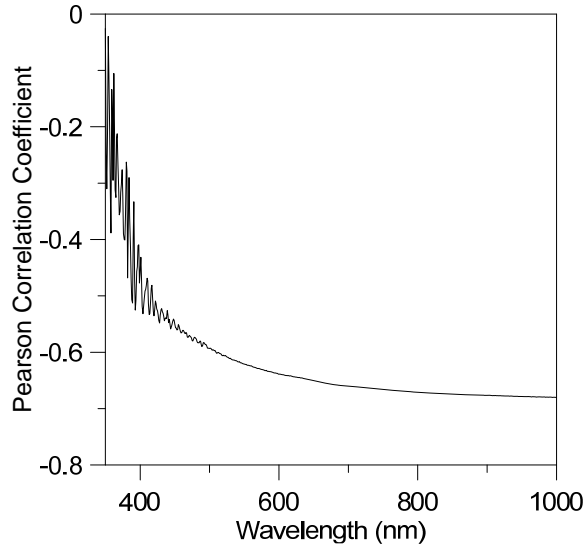
Figure 4.3. Pearson correlation coefficients between reflectance and SOC at different wavelengths.

RESULTS AND DISCUSSION

The 20 soil samples collected in the field were analyzed for SOC, total N and gravimetric moisture content (Table 4.2). The mean and range of total C and moisture were greater in field 1 than in field 2 and while mean total N was greater in field 1, total N varied slightly more in field 2 (CV = 0.20 and 0.26, respectively).

Table 4.2. Mean, standard deviation, coefficient of determination, maximum and minimum of laboratory analysis results of soil organic carbon (SOC), total nitrogen (N) and moisture.

Soil Property	Mean	Standard Deviation	CV [a]	Maximum	Minimum
Field 1 (n=9)					
SOC (g kg^{-1})	19.1	4.3	0.23	25.0	13.7
Total N (g kg^{-1})	1.81	0.36	0.20	2.35	1.35
Moisture (%)	27.8	5.5	0.20	35.2	18.3
Field 2 (n=11)					
SOC (g kg^{-1})	13.9	3.0	0.22	20.8	9.67
Total N (g kg^{-1})	1.22	0.32	0.26	1.86	0.89
Moisture (%)	16.8	4.9	0.29	25.7	10.8

[a] CV = coefficient of variation

Initial Lab Sensing Tests

Initial tests were done with the two-band sensor in the lab in a stationary setting. These tests were performed in an attempt to remove external variables, such as residue, dust, and movement in reference to the target, so that the best possible calibrations could be obtained with the sensor. Several initial tests were run before soil analysis began. Because little operational information was available for this sensor, one of the first tests completed was testing the range of reflectance outputs. Simulated targets were used to achieve a greater range of color differences than what may be typically seen in soils. We found reflectance readings of about 0.8 and 0.9 for NIR and VIS respectively for a black target and about 3.3 and 3.6 for NIR and VIS respectively, for a very light brown colored target.
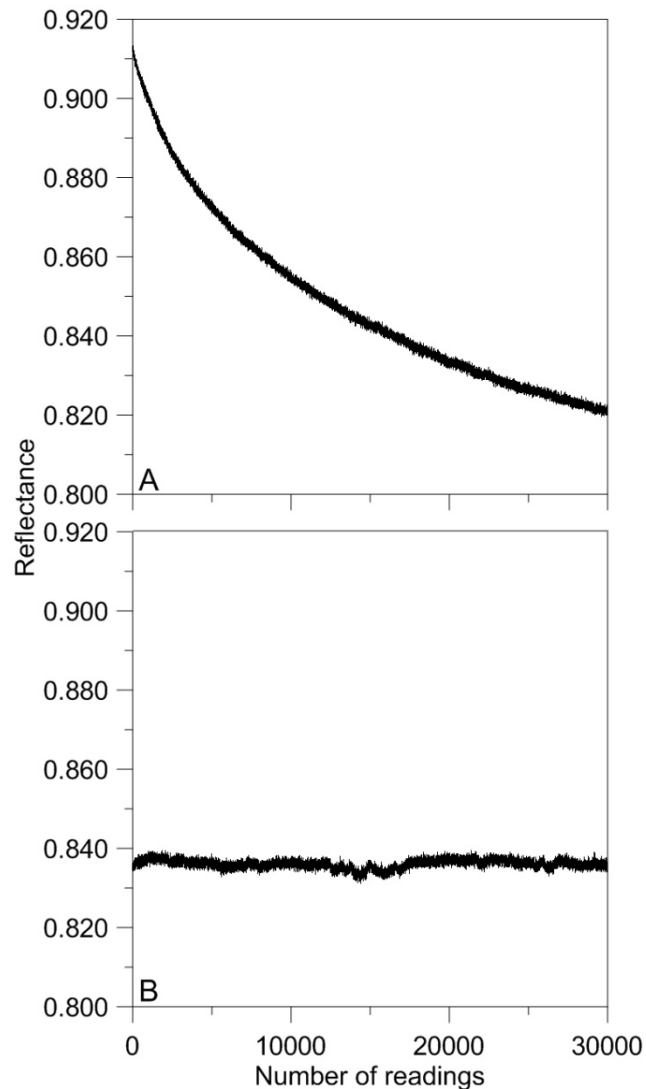
Figure 4.4. Raw output of NIR reflectance over an hour period varied much less in B compared to A. Readings in B were taken when the two-band sensor was attached to a DC power supply and warmed-up overnight, test A was battery powered and not warmed up.

Initially while testing the two-band sensor over a long time span, reflectance measurements decayed or increased considerably during a 5 min testing period. Therefore one hr measurements were taken to see if this trend continued. During the hour of testing, readings changed sharply and NIR sensor measurements decayed from 0.91 to 0.82 when

using a simulated black surface for reflectance (Figure 4.4, A). Based on the two-band

sensor developer's recommendation and since many full spectrum sensors suggest a

warm-up period before collecting data, the two-band sensor was left on overnight to

collect data the following day. Additionally, the sensor was connected to a DC power

supply to remove the possible decay in power output that could result from the

rechargeable batteries used in the previous test. When attached to a DC power supply and

warmed-up overnight sensor measurements were much more consistent, varying from

about 0.840 to 0.832 (Figure 4.4, B). These numbers are similar to those that were seen at

the end of the hour in the test with no warm-up suggesting that an hour warm-up may be

needed before measurements are taken with this prototype sensor. This warm-up

procedure was followed for the remainder of the lab tests.
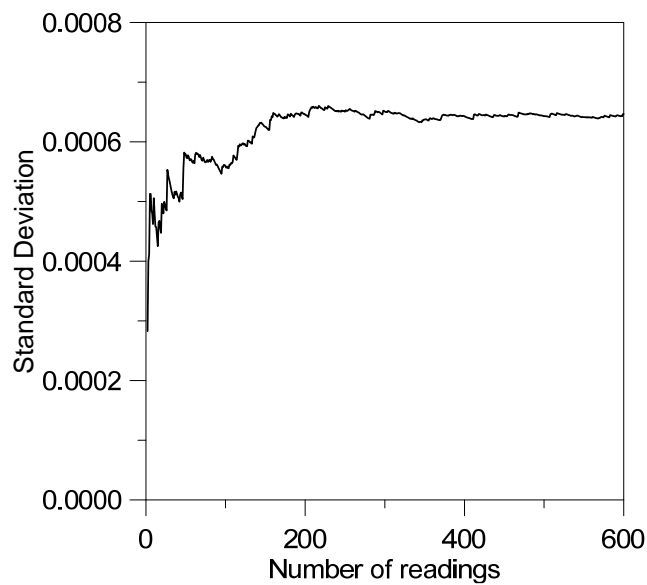
Figure 4.5. Standard deviation of reflectance readings from a black simulated surface was found to vary during the first 200 readings, after which fluctuations greatly decreased.

A second test was performed to determine the optimum amount of time to acquire

data for reflectance measurements within the lab. This was done by collecting stationary

measurements over a long time span (5 min.) to determine when the standard deviation of

the readings became less variable. After 200 readings the standard deviation fluctuated

much less and therefore this number of readings was chosen (Figure 4.5).

Full and Partial Spectrum Sensing

Several different preprocessing methods were applied to the data collected with

the full spectrum spectrometer. These preprocessing methods were tested on the full

spectrum (450 to 2500 nm) and first detector (450 to 1000 nm) data sets. The mean

centered (MC) model performed better than raw reflectance and combinations of MC,

first derivative, and detrending for the dried and sieved analysis; therefore the remaining

PLS models used the mean centering preprocessing method.

Based on the R^2 and RPD, the best overall model was for moisture calibrated with

data from the first detector (Table 4.3). This model was the only one that had good

accuracy (Chang et al., 2001) with an R^2 of 0.86 and RPD of 2.02. This was surprising

since past literature has stated that moisture is predicted in the NIR region around the

wavelengths of 1400 and 1900 nm (Liu et al., 2002). Data from the first detector was also

better able to predict N in a field-moist condition as compared to the dry counterpart. In

this restricted group of soils, darkness in the soil might have been the key contributing

factor to this high estimation accuracy. The full spectrum results were similar to those of

past research (Waiser et al., 2007; Tekin et al., 2012) and within the full spectrum; dry

analysis was better able to estimate the soil properties in question as compared to the

field-moist analysis. Other than moisture, accuracy was quite low with $R^2 \leq 0.65$.

Wetterlind et al. (2008) achieved much greater accuracy ($R^2 = 0.88$) for estimation of surface SOC but had a similar RMSE (3.3 g kg^{-1}). Reported SOC standard deviation was double what was seen in this study (8.7 and 4.3, respectively). In a study with a similar variation in SOC (standard deviation = 3.2) to this one, Bricklemyer and Brown (2010) had an R^2 of 0.36 with RPD of 1.1 using full spectrum sensing within the lab. Like what was found by Bricklemyer and Brown (2010), the low accuracies seen in this study may be attributed to the low range of SOC. This idea is supported by the fact that all of the models within this analysis required less than 4 PCs, and many required only two or three (Table 4.3). PCs are used to create new prediction variables as a linear combination of the original predictor variables and since variation within the data was small fewer PCs were needed.

Table 4.3. Partial least squares cross-validation statistics for soil organic carbon (SOC), nitrogen (N), and moisture content for the full spectrum and first detector analysis.

Soil Property	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation	RPD[a]	PC [b]
Full Spectrum (450 to 2500 nm)						
SOC						
Dried sieved	0.60	0.58	2.74[c]	2.90	1.53	3
Field-moist	0.41	0.29	3.35	3.71	1.20	2
Nitrogen						
Dried sieved	0.65	0.62	0.25	0.26	1.66	3
Field-moist	0.46	0.35	0.32	0.35	1.25	2
Moisture						
Field-moist	0.79	0.74	3.51	3.98	1.88	2
First Detector (450 to 1000 nm)						
SOC						
Dried sieved	0.45	0.39	3.23	3.41	1.31	3
Field-moist	0.43	0.31	3.28	3.64	1.23	2
Nitrogen						
Dried sieved	0.30	0.20	0.37	0.41	1.09	2
Field-moist	0.66	0.55	0.25	0.30	1.49	4
Moisture						
Field-moist	0.86	0.75	2.76	3.71	2.02	4

[a] RPD = ratio of predicted deviation (SD/RMSEV)
[a] PC = number of principal components used in model
[c] Units of RMSE are g kg^{-1} for SOC and N and % for moisture

Two-Band Reflectance Sensing

     In the analysis of two-band data, the best models generally used data from the NIR band, centered around 850 nm (Table 4.4). The most accurate models used the field data taken with the two-band reflectance sensor. This result was unexpected because it uses data taken with the most external variables (e.g. dust, variable sun conditions, and partial residue). Also, many of these areas had fewer than 200 data points, the number which was shown to reduce variation in standard deviation in the sensing. The field data with the two-band sensor was taken by driving slowly across areas about 1 m^2 that were

manually raked clear of residue, which provided a clear view of the soil for sensing. This result demonstrates the two-band sensor's ability to perform within its design environment. Comparing the estimation accuracy of soil variables, SOC estimation (R^2 = 0.75) was more accurate than the other soil variables followed by moisture (R^2 = 0.65) and N (R^2 = 0.57). This was also true for the prediction models with the two-band sensor in the lab as well as the field-moist and dry spectrometer analysis. No model for field-moist spectrometer sensing of N was significant (α=0.05) and therefore no equation was shown. Similar to the above PLS analysis (Table 4.2), RMSE was relatively low for each of the models within this analysis (Table 4.4). This again suggests that a low range of data and a small number of data points limited the model's ability to predict the soil properties. Models produced by the two-band sensor were slightly more accurate than those using averaged wavelengths from the spectrometer. This result could be due to the fact that the size of the optical footprint of the spectrometer was much less and therefore had less spatial averaging. Also possible was contamination of the spectrometer readings by ambient light. Although care was taken to press the spectrometer probe firmly against the soil surface a small amount of ambient light might have entered the detection system changing the reflectance readings. The R^2 and RMSE of SOC prediction for the spectrometer field-moist and dry two-band analysis (Table 4.4) and the first detector analysis (Table 4.3) are very similar. This shows that few wavelengths are needed for estimation of SOC when the spectrum is limited to 450 to 1000 nm and when soils are similar in nature (fields from same general area).

Table 4.4. Equations which best estimate soil organic carbon (SOC), nitrogen, and moisture in two-band analysis using the reflectance sensor and averaged wavelengths from the spectrometer from both the visible (VIS; 660 nm) and near-infrared (NIR; 850 nm) wavelengths.

Equation	R^2	RMSE
Two-band Field		
$SOC = 13.4 \times NIR^2 - 51.8 \times NIR + 60.6$	0.76	2.3[a]
$Nitrogen = -1.26 \times NIR + 3.25$	0.57	0.30
$Moisture = -23.2 \times NIR + 52.8$	0.65	4.65
Two-band Lab Field-moist		
$SOC = -15.0 \times NIR + 29.0$	0.49	3.2
$Nitrogen = -1.36 \times NIR + 2.67$	0.37	0.37
$Moisture = -12.6 \times NIR^2 + 30.8$	0.39	6.04
Spectrometer Field-moist		
$SOC = -65.0 \times NIR + 24.9$	0.43	3.4
Nitrogen not significant at a level of $\alpha = 0.05$		
$Moisture = -109 \times NIR + 36.4$	0.43	5.86
Spectrometer Dry		
$SOC = -21300 \times VIS^2 + 15900 \times NIRVIS + 4080 \times VIS - 3130 \times$		
$NIR + 65.8$	0.47	3.3
$Nitrogen = 3170 \times NIR^2 - 1830 \times NIR + 27.5$	0.26	1.3

[a] Units of RMSE are g kg^{-1} for SOC and N and % for moisture

Using the prediction equations for SOC obtained for each of the different data sources, graphs were made depicting the lab-measured SOC compared to the estimated SOC (Figure 4.6). R^2 is relatively low for all but the field two-band sensor field analysis. Based on results from this test, the two-band sensor is best suited to perform within the field taking moving measurements. The two-band sensor lab analysis (Figure 4.6, B) and field-moist spectrometer two-band average (Figure 4.6, C) which were taken at similar times show a similar distribution of data. Dry spectrometer two-band average (Figure 4.6, D) analysis was slightly better than field-moist ($R^2 = 0.47$ and 0.43, respectively).
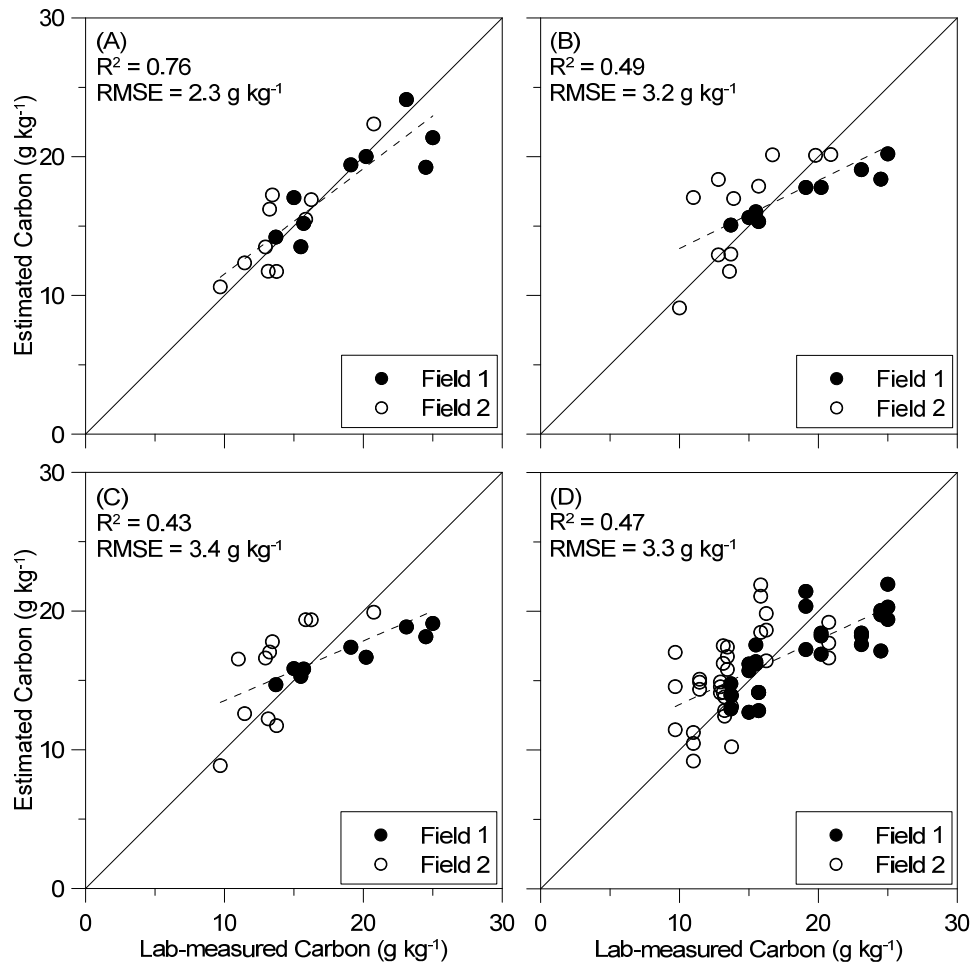
Figure 4.6. Sensor estimated vs. measured SOC content from two-band analysis. Data is from the two-band sensor in the field (A) and lab (B) and two-band averages from the spectrometer with field-moist (C) and dry (D) soils.

Similar to the field data models containing all within-field points from both fields, models were created for each individual field. These model equations were then applied to both the fields (Figure 4.7). This shows not only the ability of the sensor to estimate SOC for a single field but also its ability to estimate SOC in a different, but similar area. Both fields had an $R^2 \geq 0.69$ when their respective equations were applied to them. When the field equations were applied to the opposite field, the results were not as successful. If

a greater number of points and wider range of SOC were available to develop the

estimation model, the models may have been more robust for SOC prediction.
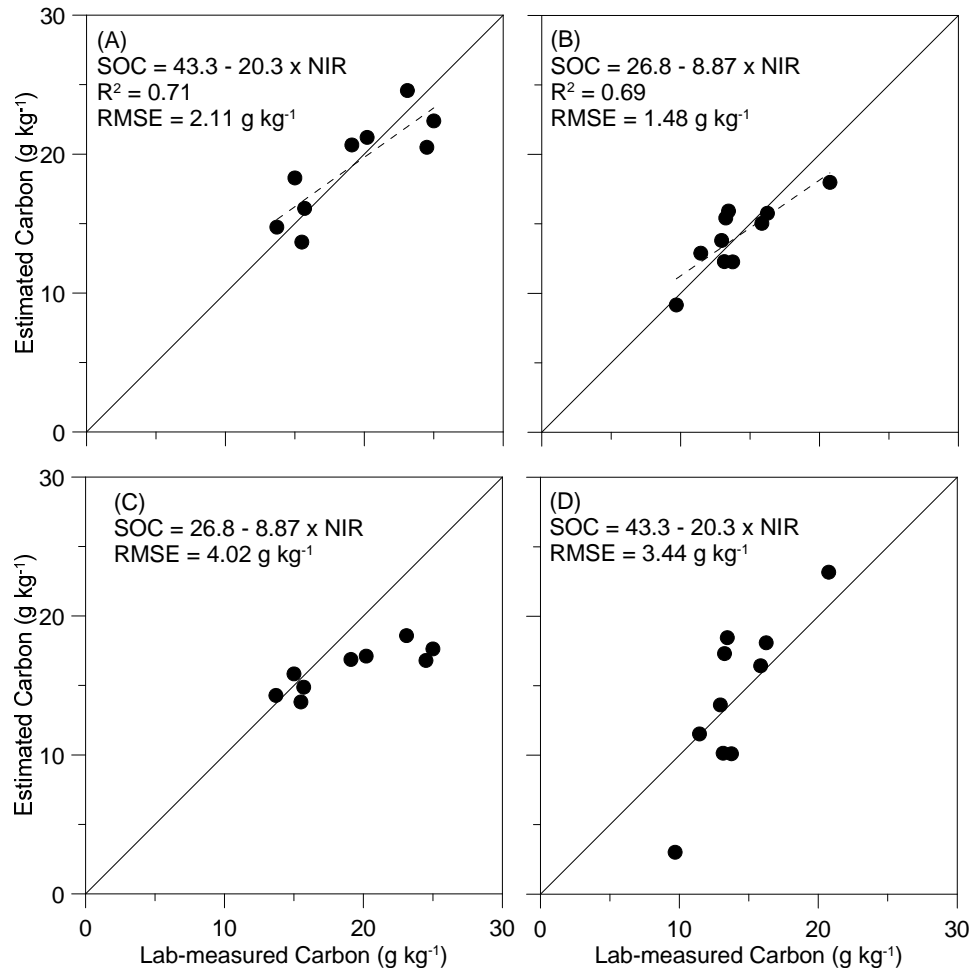
Figure 4.7. Sensor-estimated vs. measured soil organic carbon (SOC) content from two-band analysis. Models for estimating SOC were produced with field collected data. Each model included only field data from that particular field. These models were applied to the field for which they were created (A and B) and then applied to the opposite field (C and D).

Field Mapping
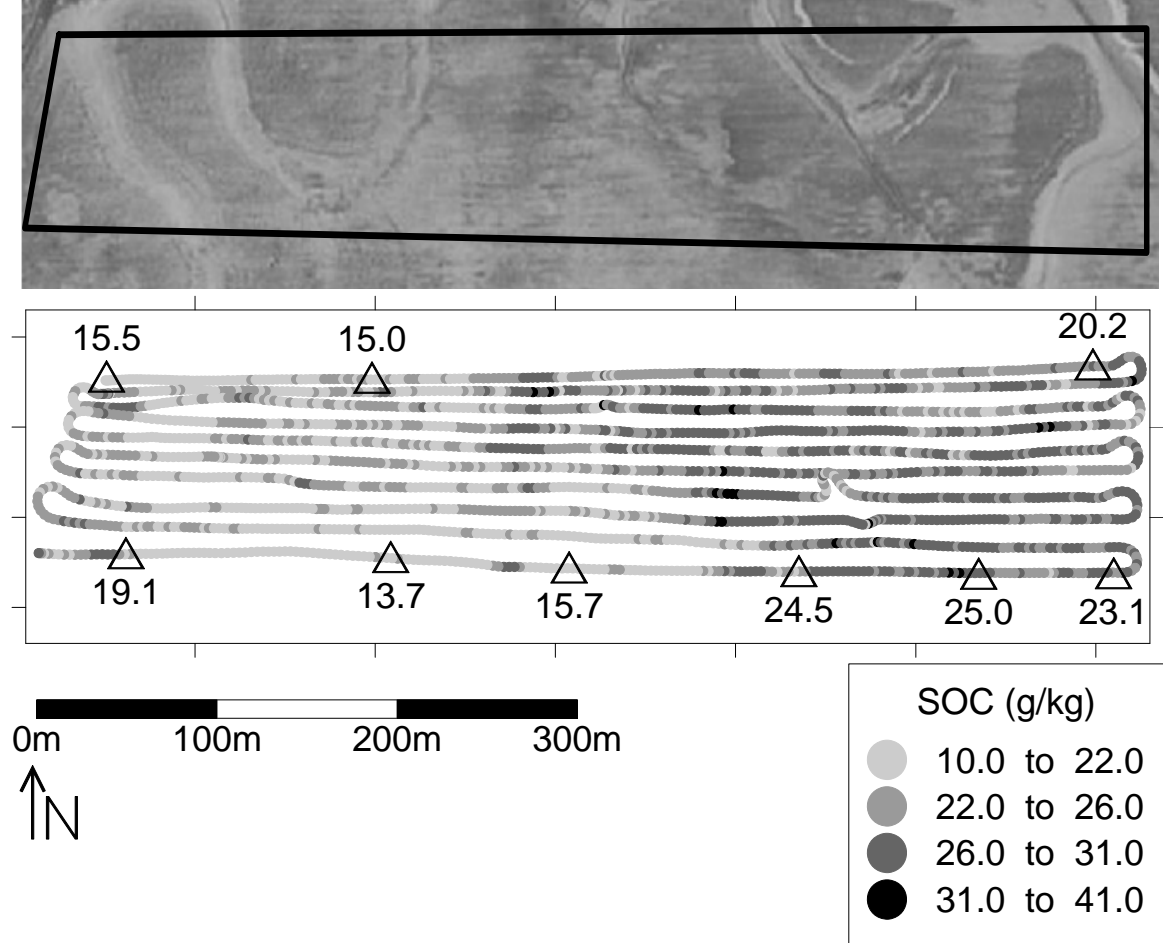
Figure 4.8. Map of predicted soil organic carbon (SOC) for field one using equation developed from two-band analysis of data from the raked areas.

Creation of field prediction SOC maps (Figure 4.8, 4.10, and 4.11) was based on the equation from two-band field analysis (Table 4.4). SOC was generally well predicted for field one (Figure 4.8). Higher values for the lab measured SOC tend towards the eastern portion of the field and this area had higher predicted SOC. When comparing the prediction map to the aerial image the soil lacks noticeable variation with only a slightly darker color in the western portion of the field. This is similar to was seen in the field

during data collection. The average predicted SOC using points from within three meters of the soil collection point was 22.8 g kg^{-1} with a standard deviation of 6.05 g kg^{-1}. The lab average (Table 4.2) was 19.1 g kg^{-1}, resulting in a RMSE of 6.9 g kg^{-1}. When comparing lab values marked on the prediction map to the points of the prediction map, the lab data corresponds well with predicted values.

Figure 4.9. Overview of field two, identifying the two data collection areas.

Figure 4.10. Map of predicted soil organic carbon (SOC) for the western portion of field two using equation developed from two-band analysis of data from the raked areas.

Soil organic carbon prediction for the western portion of field two (Figure 4.10) was not as successful as for field one, even though the aerial image and visual inspection during data collection showed much more variation than field one (Figure 4.10). The northern and southern portions of the area sensed were sandy with the most sand in the most northern soil sample location. The middle portion of the area sensed was loamier. Average predicted SOC using points from within three meters of the soil collection point 33.2 g kg^{-1} with a standard deviation of 5.1 g kg^{-1} whereas the average lab measured SOC

was 13.9 g kg^{-1} with a standard deviation of 3.0 g kg^{-1} (Table 4.2). This led to a RMSE of 18.4 g kg^{-1}, which is roughly triple the RMSE of field one. Much of the higher predicted SOC is in the central portion of the sensed area where a soil sample failed to be collected. Unlike field one, many of the lab measured values do not fall within the range of predicted SOC represented on the map leaving the question of whether samples taken within the area of high predictions would have reduced error.

Like the western portion of field two, the eastern portion (Figure 4.11) was not well estimated. This area had many prediction values ranging from 31 to 41 g kg^{-1} and no lab measured SOC value was this high. Average predicted SOC using points within three meters of the soil collection point was 24.3 g kg^{-1} with a standard deviation of 4.8 g kg^{-1} whereas the average lab measured SOC was 13.9 g kg^{-1} with a standard deviation of 3.0 g kg^{-1} (Table 4.2). This led to a RMSE of 13.0 g kg^{-1} in this area, similar to the RMSE in the western portion of field two. Even though none of the lab measured SOC values are >20.8 g kg^{-1}, the points fall in areas where the predicted SOC is >31.0 g kg^{-1}. This could have been due to the fact that this area was most unlike the other areas as it contained the most clay. In both field two measurement areas, raked areas were chosen at locations which showed visual differences in texture. It is possible that the locations chosen while visually different, were not indicative of the overall variation seen in the field.
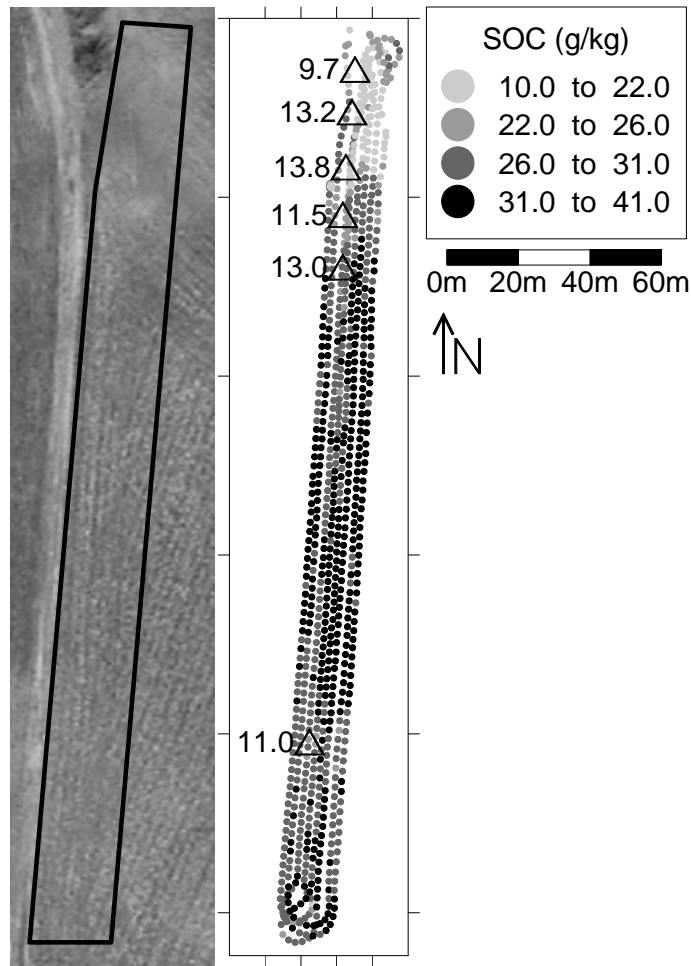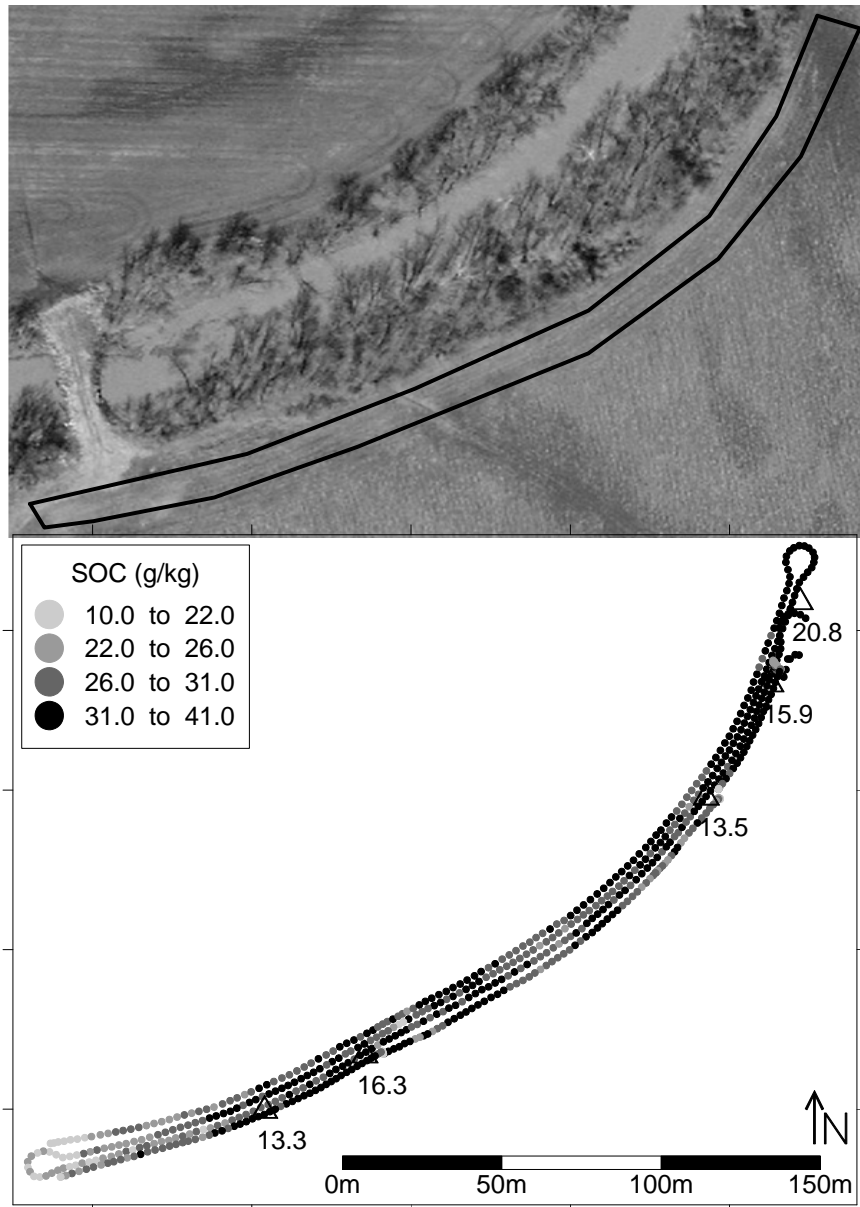
Figure 4.11. Map of predicted soil organic carbon (SOC) for the eastern portion of field two using equation developed from two-band analysis of data from the raked areas.

Optimization of Reflectance Sensor Wavelengths

Soil indices, both ratio and difference, were calculated with all possible

combinations of wavelengths within the full spectrum and the spectrum from 450 to 1000

nm. Correlation maps (Figure 4.12) were then used to visually examine trends in correlation of the wavelength indices to measured SOC. Ranges within each correlation map were selected so that the correlation trends could be evaluated more easily. When the full spectrum was used, generally the greatest correlation to SOC was seen in the NIR range (Figure 4.12, top). The combination of one point from the 2200 to 2400 nm wavelength range along with a second point from either 1400 or 1900 nm resulted in broad areas of high correlation within the full spectrum. The range of 2200 to 2400 is similar to the wavelengths used by other researchers for the prediction of SOC (Lee et al., 2009). Also interesting are the high correlation areas located at 1400 and 1900 nm as these wavelengths have been shown to have strong correlation with the prediction of moisture within soils (Liu et al., 2002). Within the ratio correlation map, many points with high correlation are located above 1800 nm. The difference correlation map though, shows high correlations spread throughout different areas across the entire wavelength range.
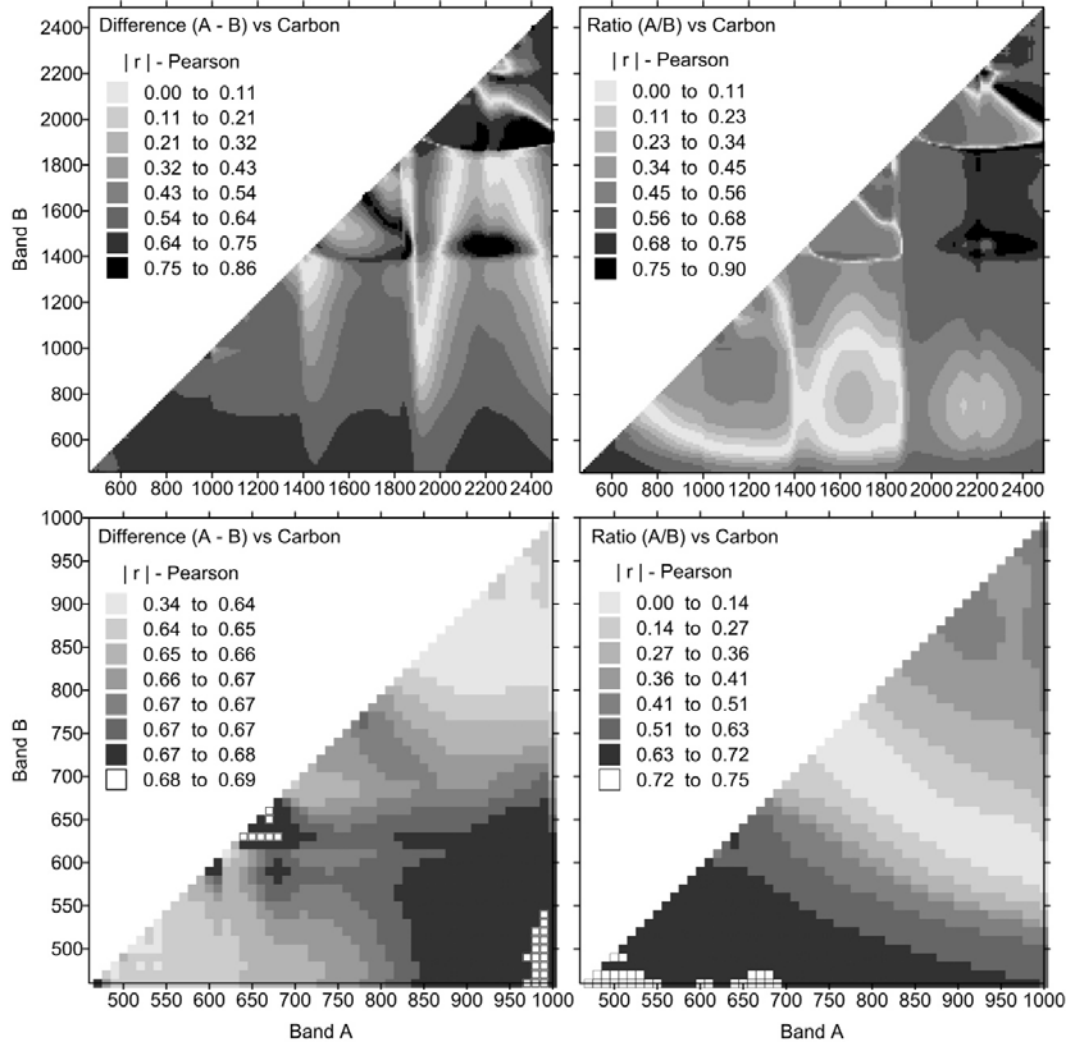
Figure 4.12. Correlation coefficients obtained between carbon and difference and ratio indices, using 450 to 2500 nm (top) and 450 to 1000 nm (bottom) wavelength ranges.

More closely examining correlations below 1000 nm is useful because these wavelengths are detected by less expensive technology such as that used in the two-band sensor (Figure 4.12, bottom). In this analysis both indices showed higher correlation values corresponding with low wavelengths on the y-axis but showed high correlation on opposite ends of the x-axis with the ratio tending toward the low wavelengths and difference toward high wavelengths.

Table 4.5. Wavelength combinations with the highest absolute Pearson correlation coefficients (r) within each plot of figure 11. Also included in the first detector analysis are the correlations for the wavelength combination used in the two-band sensor (marked by *).

Wavelength Combinations (nm)	|r|
Full Spectrum (450 to 2500 nm)	
Difference	
2460, 1900	0.86
2400, 1890	0.86
2190, 1450	0.85
2210, 1440	0.84
1830, 1530	0.82
Ratio	
2400, 1890	0.90
2350, 1800	0.87
2340, 1450	0.81
2350, 1460	0.81
2220, 2210	0.80
First Detector (450 to 1000 nm)	
Difference	
640, 630	0.69
990, 510	0.68
990, 460	0.68
670, 650	0.68
850, 660*	0.67
Ratio	
480, 470	0.75
670, 470	0.72
610, 460	0.72
730, 460	0.71
850, 660*	0.01

All of the full spectrum wavelength combinations showing highest correlation

with SOC had both wavelengths within the NIR spectrum. While the selected NIR

wavelengths in this analysis (Table 4.5) are much greater than the NIR wavelength in the two-band senor, this result reaffirms the results seen in the two-band reflectance equations (Table 4.4) in which most models for prediction of SOC included NIR data. This result is also similar to results in both the previous chapter (Figure 3.6) and Morgan et al. (2009) in which they found wavelengths most correlated to SOC to be above 2000 nm. Although the ratio index had the highest correlation in this analysis (r = 0.90) overall it showed very poor correlation when using the wavelengths contained in the two-band sensor (850 and 660 nm). Within the first detector using the ratio index, the best combination contained two bands from the VIS (480 and 470 nm). This suggests that estimation was made possible by different degrees of soil darkness relating to higher amounts of C (Stenberg, 2010).

Correlation to SOC with the bands used in the two-band sensor was much greater using the difference index (r = 0.67). There were also many wavelength combinations within this index which like the two-band sensor contained one band from the VIS spectrum and one from the NIR. Similarly to the ratio index, the highest correlation (r = 0.69) within the difference index contained two bands from the VIS spectrum (480, 470 and 640, 630, respectively). While these wavelength combinations had the best correlation, it is very unlikely that adjacent bands would be selected for use within a two-band reflectance sensor. These combinations show high correlation for this data but typically using adjacent bands would reduce estimation power.

CONCLUSION

Two types of sensors, a two-band on-the-go reflectance sensor and a full spectrum spectrometer, were evaluated for the prediction of surface soil properties. After initial tests with the two-band sensor it was determined that more stable results were achieved when it was attached to a DC power supply and allowed to warm-up for at least an hour. Based on this set-up, the range of NIR channel variation over an hour period on a simulated target was reduced by 167% with warm-up. Also during this test it was determined that the standard deviation fluctuated much less after 200 readings were taken on a static target.

When testing the spectrometer in the lab on field-collected soils, the best result was obtained for moisture ($R^2 = 0.93$) using wavelengths from 450 to 1000 nm. Other than this result, dry data using the full spectrum generally provided the best prediction results. Ranges of SOC and N within the collected soil were low (CV = 0.23 and 0.20, respectively) and similar to past research on soils with low variation (Bricklemyer and Brown, 2010; Sudduth et al., 2010) accuracies (measured as R^2) from this study were also low. Low R^2 did not lead to high RMSE (C = 3.35 g kg^{-1}, field-moist) values. Many values of RMSEV were as low or lower than past studies with a greater R^2 (C = 0.65 g kg^{-1}; Chang and Laird, 2002)

Two-band analysis led to the conclusion that the two-band sensor was much better suited to predict soil properties than averaged data for the same two bands obtained with the full spectrum sensor. This is most likely due to the fact that full spectrum spectrometer sensed a smaller area than the two-band sensor causing less variation to be averaged into each measurement. Also, spectrometer measurements are very sensitive to

ambient light entering the detector during measurement and although steps were taken to prevent this, it is possible that accuracy was affected by this problem. The highest overall accuracy was for SOC estimation from data collected in the field from raked areas with an R^2 of 0.76. This result, while encouraging for future use of the sensor in a field setting, was surprising because external variables (e.g. dust and movement of sensor in reference to target) were most prevalent during within-field sensing. Single-field equations used to predict SOC within a different field did not perform as well as the multi-field model. When the field calibration equation was applied to the whole field, results were mixed. SOC variations in field one were better estimated than field two with only a few points reaching values above the greatest value of measured data. Compared to field two, this field lacked visual variation in soil texture and color.

In a two-band correlation analysis, wavelengths within the NIR were most correlated to SOC. This tied back to the two-band SOC estimation equations in which most equations included NIR data for SOC estimation. Within the full spectrum using both the difference and ratio index, high correlations with SOC were found at wavelengths of 2200 to 2400 nm. The highest correlations ($r = 0.75$) within the first detector range were those that contained two wavelengths from the VIS range. This can most likely be attributed to soil darkness correlating to SOC. Combinations which did include both the VIS and NIR range were slightly less correlated to SOC ($r = 0.71$).

CHAPTER 5
CONCLUSIONS

Two sensors were tested for the estimation of surface and subsurface soil properties. Conclusions based on the results reported in this study are:

- All data taken with the spectrometer required some form of mathematical preprocessing to improve prediction accuracies. Two different processing equations were chosen to improve prediction, mean centering and a combination of mean centering + detrending. Similar to past studies a universal preprocessing equation could not be used on all data (Viscarra Rossel et al., 2006; Lee et al., 2009; Sudduth et al., 2010) and data taken within different fields required different preprocessing methods.

- Analysis of field-moist samples was generally less accurate than analysis of dried and sieved samples, but accuracy reduction was minimal in most cases. Based on this research, drying and sieving soil for spectral analysis may not be necessary if a relatively small decrease in estimation accuracy for soil physical and chemical properties can be tolerated. Of the variables analyzed, carbon (C) exhibited the highest overall accuracy ($R^2 = 0.97$) when estimated from dry soils obtained to a depth of 120 cm and using full-spectrum analysis. Other soil properties had mixed results. Magnesium (Mg), for example, which was sampled from soils <20 cm deep, showed good accuracy ($R^2 = 0.83$) whereas pH and saturated hydraulic conductivity (K_s) did not ($R^2 = 0.07$ and 0.37, respectively). Soil properties such as Mg, pH, and K_s do not have direct

spectral absorption features and successful estimation is thought to be attributed to locally present co-variation in spectrally active constituents (Stenberg et al., 2010). Based on the accuracy of these results, reflectance analysis could have potential use as an indicator within a soil quality assessment model.

- The near-infrared (NIR) range was found to have greater significance than the visible (VIS) range for estimations of C in both field-moist and dry analysis, similar to what other researchers have found (Clark et al., 1990; Morgan et al., 2009). Wavelengths above 1800 nm showed particularly high correlation with C. When the spectrum was limited to wavelengths under 1000 nm, wavelengths in the visible were most strongly related to C, which can most likely be attributed to soil darkness correlating to C. While combinations of VIS wavelengths showed the highest correlation for the narrow range of soils in this analysis, these combinations might not be the best for application to a wide range of soils.

- In preliminary tests with a prototype two-band sensor, output varied greatly when measurements were taken on a simulated black target. When a warm-up period was allowed and the sensor was attached to a DC power supply output variation diminished greatly (167%). Also during this test it was determined that the standard deviation fluctuated much less after 200 readings taken in a single location. Therefore, during lab sensing with the two-band sensor, the

sensor was warmed-up, connected to the DC power supply and 200 readings were taken per measurement.

- When field soil organic carbon (SOC) maps were made with the two-band sensor, results were varied. One field chosen for study showed a RMSE of 6.9 g kg^{-1} between field predicted and measured values but the other had a RMSE of 18.4 g kg^{-1}. This difference can most likely be attributed to the variation captured during calibration measurements on raked areas in field two not being indicative of the overall variation seen in the entire field.

SUGGESTIONS FOR FUTURE STUDY

Based on the results obtained from this study, the recommendations for future work are:

- Soil properties, such as POM-C, obtained to a depth of <20 cm provided good accuracy but many others did not (e.g. pH, potassium (K), phosphorus (P), etc.). With the increased testing of soils for site specific management the use of spectroscopy could increase testing speed and reduce cost. Therefore more analysis should be done with surface soils for various soil properties and specifically those that have been shown to affect soil fertility (e.g. pH, P, and K).

- Gather more measurements with the prototype two-band reflectance sensor in fields with varying amounts of surface SOC. Similar to past research we found that low SOC variation of tested soils can lead to low estimation accuracy for R^2 but good values of RMSE. Since literature is limited on this sensor, further testing is need to determine if the low SOC variation of the

tested soils caused low R^2 accuracy or if the accuracy seen in this research is what can be expected with this sensor.

- Our research showed that SOC could be estimated within the field where areas were manually cleared for sensing (by removing residue) but in full field tests, results varied. Therefore further testing is needed to determine why full field estimation accuracy did not match that of the manually cleared areas. In the initial tests with the two-band sensor it was noticed that residue greatly affected sensor output but these results were not quantified. To fully understand the effect residue has on sensor measurements, it would be useful to test known ratios of soil to residue with the two-band sensor. Also, the residue clearing device used in this study for field sensing did not provide a consistently cleared sensing path and tests were only done in fields containing soybean residue. Further testing should be done to determine a better residue clearing device that is not only more consistent but is capable of clearing both corn and soybean residue.

REFERENCES

Analytical Spectral Devices Inc. FieldSpec® Pro: user's guide. Boulder: Analytical Spectral Devices (ASD), 2002. 136 p.

Andrews, S.S., D.L. Karlen, and C.A. Cambardella. 2004. The soil management assessment framework: A quantitative soil quality evaluation method. S*oil Sci. Soc. Am. J*. 68: 1945-1962.

Angers, D.A., M.S. Bullock, and G.R. Mehuys. 1993. Chapter 6: *Aggregate stability to water. In Soil Sampling and Methods of Analysis*, 651-657. M.R. Carter ed. Boca Raton, FL: Lewis Publishers.

Arshad, M.A., and G.M. Coen. 1992. Characterization of soil quality: physical and chemical criteria. *Am. J. Alt. Agric*. 7: 25-31.

Arshad, M.A., and S. Martin. 2002. Identifying critical limits for soil quality indicators in agro-ecosystems. *Agric., Ecosys. Environ*. 88: 153-160.

Bennett, H.H. 1939. *Soil Conservation*. New York, NY: McGraw-Hill Book Company, Inc.

Boyle, M., W.T. Frankenberger Jr., and L.H. Stolzy. 1989. The influence of organic matter on soil aggregation and water infiltration. *J. Prod. Agric*. 2: 290-299.

Bricklemyer, R.S. and D.J. Brown. 2010. On-the-go VisNIR: Potential and limitations for mapping soil clay and organic carbon. *Comput. Electron. Agric*. 70: 209-216.

Brown, D.J., R.S. Bricklemyer, and P.R. Miller. 2005. 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. 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma* 132: 273-290.

Bruce, R.R., G.W. Langdale, L.T. West, and W.P. Miller. 1995. Surface soil degradation and soil productivity restoration and maintenance. *Soil Sci. Soc. Am. J*. 59: 652-660.

Cambardella, C.A., and E.T. Elliott. 1992. Particulate soil organic matter changes across a grassland cultivation sequence. *Soil Sci. Soc. Am. J*. 56:777-783.

Camo. 2006. The Unscrambler User Manual. Oslo, Norway: CAMO Software AS.

Cannell, R.Q., and J.D. Hawes. 1994. Trends in tillage practices in relation to sustainable crop production with special reference to temperate climates. *Soil Tillage Res.* 30: 245-282.

Chan, K.Y., W.P. Roberts, and D.P. Heenan. 1992. Organic carbon and associated soil properties of red earth after 10 years of rotation under different stubble and tillage practices. *Aust. J. Soil Res.* 30: 71-83.

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

Chappell, N.A., J.L. Ternan, and K. Bidin. 1999. Correlation of physiochemical properties and sub-erosional landforms with aggregate stability variations in a tropic Ultisol disturbed by forestry operations. *Soil Tillage Res.* 50: 55-71.Chang, C.W., D.A.

Laird, D.A. M.J. Mausbach, and C.R. Hurbergh Jr. 2001. Near-infrared reflectance spectroscopy – principal components regression analyses of soil properties. *Soil Sci. Soc. Am. J.* 65: 480-490.

Chen, F., D.E. Kissel, L.T. West, and W. Adkins. 2000. Field-scale mapping of surface soil organic carbon using remotely sensed imagery. *Soil Sci. Soc. Am. J.* 64: 746-753.

Chen, F., D.E. Kissel, L.T. West, D. Rickman, J.C. Luvall, and W. Adkins. 2005. Mapping surface soil organic carbon for crop fields with remote sensing. *J. SoilWater Conserv.* 60: 51-57.

Christy, C.D. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. *Comput. Electron. Agric.* 61: 10-19.

Clark, R.N. 1999. Chapter 1: Spectroscopy of rocks and minerals and principles of spectroscopy. In Manual of Remote Sensing, Volume 3, Remote Sensing for the Earth Sciences, 3- 58. A.N. Rencz, ed. New York, NY: John Wiley and Sons.

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.

Cozzolino D., and A. Moron. 2003. Potential of near-infrared reflectance spectroscopy and chemometrics to predict soil organic carbon fractions. *Soil Tillage Res.* 85:78-85

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

Daniel, K.W., N.K. Tripathi, and K. Honda. 2003. Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Aust. J. Soil. Res.* 41:47-59.

Dick, W.A. 1983. Organic carbon, nitrogen and phosphorus concentrations and pH in soil profiles as affected by tillage intensity. *Soil Sci. Soc. Am. J*. 47: 102-107.

Doran, J.W., and T.B. Parkin. 1994. Chapter 1: Defining and assessing soil quality. In *Defining Soil Quality for a Sustainable Environment*, 3-21. J.W. Doran et al. ed. Madison, WI: SSSA, Inc.

Doran, J.W. and M.R. Werner. 1990. Chapter 3: Management and soil biology. In *Sustainable Agriculture in Temperate Zones*, 205-230. C.A. Francis et al. ed. New York, NY: John Wiley & Sons.

Ekwue, E.I. and R.J. Stone. 1995. Organic matter effects on the strength properties of compacted agricultural soils. *Trans. ASAE* 38: 357-365.

Follet, R.T. 2008. Chapter 2: Transformation and transport processes of nitrogen in agricultural systems, 19-50. In *Nitrogen in the Environment: Sources, Problems, and Management*. J.L Hatfield and R.F. Follet eds. San Diego, CA: Academic Press, Inc.

Gantzer, C.J., and T.R. McCarty. 1987. Predicting corn yields on a claypan soil using a soil productivity index. *Trans. ASAE* 30: 1347-1352.

Gomez, C., R.A. Viscarra Rossel, and A.B. McBratney. 2008. Soil organic carbon prediction by hyperspectral remote sensing and field VIS-NIR spectroscopy: an Australian case study. *Geoderma* 146: 403-411.

Hendrix, R.F., B.R. Mueller, R.R. Bruce, F.W. Langdale, and R.W. Parmelee. 1992. Abundance and distribution of earthworms in relation to landscape factors on the Georgia Piedmont, USA. *Soil Biol. Biochem*. 24: 1357-1361.

Holland, K.H., and J.S. Schepers. 2011. Active-crop sensor calibration using the virtual-reference concept. Available at: http://hollandscientific.com/. Accessed 9 May 2012.

Hudson, B.D. 1994. Soil organic matter and available water capacity. *J. Soil Water Conserv*. 49: 189-194.

Hurt, D.R. 1981. *The Dust Bowl: an agricultural and social history*. Chicago, IL: Nelson-Hall Inc., Publishers.

Islam K., B. Singh, and A.B. McBratney. 2003. Simultaneous estimation of various 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. Skjemstand. 1998. Can mid infrared diffuse reflectance analysis replace soil extractions? *Aust. J. Exp. Agric*. 38: 681-696.

Jiang, P. S.H. Anderson, N.R. Kitchen, E.J. Sadler, and K.A. Sudduth. 2007. Landscape and conservation management effects on hydraulic properties of a claypan-soil toposequence. *Soil Sci. Soc. Am J*. 71: 803-811.

Jung, W.K., N.R. Kitchen, K.A. Sudduth, and S.H. Anderson. 2006. Spatial characteristics of claypan soil properties in an agricultural field. *Soil Sci. Soc. Am. J*. 70: 960-971.

Karlen, D.L., M.J. Mausbach, J.W. Doran, R.G. Cline, R.F. Harris, and G.E. Schuman. 1997. Soil quality: a concept, definition, and framework for evaluation. *Soil Sci. Soc. Am. J*. 61: 4–10.

Karlen, D.L. N.C. Wollenhaupt, D.C. Erbach, E.C. Berry, J.B. Swan, N.S. Eash, and J.L. Jordahl. 1994a. Crop residue effects on soil quality following 10-years of no-till corn. *Soil Tillage Res*. 31: 149-167.

Karlen, D.L. N.C. Wollenhaupt, D.C. Erbach, E.C. Berry, J.B. Swan, N.S. Eash, and J.L. Jordahl. 1994b. Long-term tillage effects on soil quality. *Soil Tillage Res*. 32: 313-327.

Kemper, W.D., and R.C. Rosenau. 1986. Aggregate stability and size distribution. In *Methods of Soil Analysis— Part 1: Physical and Chemical Properties*, 425-442. A Klute ed. Madison WI: SSSA, Inc.

Kennedy, A.C., and R.I. Papendick. 1995. Microbial characteristics of soil quality. *J. Soil Water Conserv*. 50: 243-252.

Kitchen, N.R., K.A. Sudduth, S.T. Drummond, P.C. Scharf, H.L. Palm, D.F. Roberts, and E.D. Vories. 2010. Ground-based canopy reflectance sensing for variable-rate nitrogen corn fertilization. *Agron. J*. 102: 71-84.

Kutilek, M. 2004. Soil hydraulic properties as related to soil structure. *Soil Tillage Res*. 79: 175- 184.

Kweon, G.Y., E. Lund, C. Maxton, P. Drummond, and K. Jensen. 2009. Soil profile measurement of carbon contents using a probe-type VIS-NIR spectrometer. *J. Biosys. Eng*. 34: 382-389.

La, W.J., K.A. Sudduth, S.O. Chung, and H.J. Kim. 2008. Preprocessing and calibration of optical diffuse reflectance signal for estimation of soil physical and chemical properties in the central USA. *J. Biosys. Eng*. 33: 430-437.

Lal, R. 1993. Tillage effects on soil degradation, soil, resilience, soil quality, and sustainability. *Soil Tillage Res*. 27: 1-8.

Larson, W.E., and F.J. Pierce. 1991. Conservation and enhancement of soil quality. In *Proc. Evaluation for Sustainable Land Management in the Developing World*, 175-203. Bangkok, Thailand: Int. Board for Soil Res. and Mgmt.

Lee, K.S., D.H. Lee, K.A. Sudduth, S.O. Chung, N.R. Kitchen, and S.T. Drummond. 2009. Wavelength identification and diffuse reflectance estimation for surface and profile soil properties. *Trans. ASABE* 52: 683-695.

Lee, K.S., K.A. Sudduth, S.T. Drummond, D.H. Lee, N.R. Kitchen, and S.O. Chung. 2010. Calibration methods for soil property estimation using reflectance spectroscopy. *Trans. ASABE* 53: 675-684.

Lerch, R.N., N.R. Kitchen, R.J. Kremer, W.W. Donald, E.E. Alberts, E.J. Sadler, K.A. Sudduth, D.B. Myers, and F. Ghidey. 2005. Development of a conservation-oriented precision agriculture system: water and soil quality assessment. *J. Soil Water Conserv.* 60: 411-420.

Liu, W.D, F. Baret, W.F. Gu, Q.X. Tong, L.F. Zheng, and B. Zhang. 2002. Relating soil surface moisture to reflectance. *Remote Sens. Environ.* 81: 238-246.

Ludwig, B., P. K. Khanna, J. Bauhus, and P. Hopmans. 2002. Near infrared spectroscopy of forest soils to determine chemical and biological properties related to soil sustainability. *Forest Ecology Mgmt.* 171: 121-132.

Marriott, E.E., and M.M. Wander. 2006. Total and labile soil organic matter in organic and conventional farming systems. *Soil Sci. Soc. Am. J.* 70:950-959.

MacRae, R.J., and G.R. Mehuys. 1985. The effects of green manuring on the physical properties of temperate-area soils. *Adv. Soil Sci.* 3: 71-94.

Mahboubi, A.A., R. Lal, and N.R. Fausey. 1993. Twenty-eight years of tillage effects on two soils in Ohio. *Soil Sci. Soc. Am. J.* 506-512.

Markham, B.L., D.L. Williams, J.R. Schafer, F. Wood, and M.S. Kim. 1995. Radiometric characterization of diode-array field spectroradiometers. *Remote Sens. Environ.* 51: 317−330.

Martens, H., and T. Naes. 1989. *Multivariate Calibration*. Chichester, UK: John Wiley & Sons.

Martin, P.D., D.F. Malley, G. Manning, and L. Fuller. 2002. Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. *Can. J. Soil. Sci.* 82: 413-422.

McCarty, G. W. and J.B. Reeves III. 2006. Comparison of near infrared and mid infrared diffuse reflectance spectroscopy for field-scale measurements of soil fertility parameters. *Soil Sci*. 171: 94-102.

McCarty, G. W., J.B. Reeves III, V.B. Reeves, R.F. Follet, and J.M. Kimble. 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Sci. Soc. Am. J*. 66(3-4): 640-646.

Mehlich, A. 1984. Mehlich-3 soil test extractant—A modification of Mehlich 2 extractant. *Commun. Soil. Sci. Plant Anal*. 15: 1409-1416.

Morgan, C.L.S., T.H. Waiser, D.J. Brown, and C.T. Hallmark. 2009. Simulated in situ characterization of soil organic and inorganic carbon with visible near-infrared diffuse reflectance spectroscopy. *Geoderma* 151: 249-256.

Mortimore, J.L., L.J. Marshall, M.H. Almond, P. Hooins, and W. Jatthes. 2004. Analysis of red and yellow ochre samples from Clearwell Caves and Catalhoyuk by vibrational spectroscopy and other techniques. *Biomol. Spectrosc*. 60: 1179-1188.

Mouazen, A.M., B. Kuang, J. De Baerdemaeker, and H. Ramon. 2010. Comparison between principal component, partial least squares, and artificial neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. *Geoderma* 158: 23-31.

Mudgal, A., S.H. Anderson, C. Baffaut, N.R. Kitchen, and E.J. Sadler. 2010. Effects of long-term soil and crop management on soil hydraulic properties for claypan soils. *J. Soil Water Conserv. 65: 393-403*

Murray, I., and P.C. Williams. 1987. Chapter 2: Chemical principles of near-infrared technology. In *Near-Infrared Technology in the Agricultural and Food Industries*, 17-34. P.C. Williams and K.H. Norris, eds. St. Paul, MN: Am. Assoc. of Cereal Chemists, Inc.

Nathan, M., J. Stecker, and Y. Sun. 2006. Soil testing in Missouri: A guide for conducting soil tests in Missouri. Extension Circular 923, 4/06 revision. Columbia, MO: Missouri Cooperative Extension Service. (available online at: http://soilplantlab.missouri.edu/soil/ soiltestinginmissouri.pdf).

Paul, E.A. and F.E. Clark. 1989. *Soil Microbiology and Biochemistry*. San Diego, C.A.: Academic Press, Inc.

Rachman, A., S.H. Anderson, C.J. Gantzer, and A.L. Thompson. 2003. Influence of long-term cropping systems on soil physical properties related to soil erodibility. *Soil Sci. Soc. Am. J*. 67: 637-644.

Raun, W.R., J.B. Solie, M.L. Stone, K.L. Martin, K.W. Freeman, R.W. Mullen, H. Zhang, J.S. Schepers, and G.V. Johnson. 2005. Optical sensor-based algorithm for crop nitrogen fertilization. *Commun. Soil Sci. Plant Anal*. 36:2759-2781.

Rawls, W.J., D.L. Brakensiek, and S.D. Logsdon. 1993. Predicting saturated hydraulic conductivity utilizing fractal principles. *Soil Sci. Soc. Am. J*. 57:1193-1197.

Reeves, J.B. III, G.W. McCarty, and J.M. Kimble. 2006. Can near or mid-infrared diffuse reflectance spectroscopy be used to determine soil carbon pools? *Commun. Soil Sci. Plant Anal*. 37: 2307-2325.

Reynolds, W.D., D.E. Elrick, E.G. Youngs, A. Amoozegar, H.W.G. Booltink, and J. Bouma. 2003. Chapter 3.4: Saturated and field-saturated water flow parameters. In *Methods of Soil Analysis – Part 4: Physical Methods*, 797-815. J.H. Dane and G.C. Topp eds. Madison, WI: SSSA, Inc.

Roberts D.F., V.I. Adamchuk, J.F. Shanahan, R.B. Ferguson, and J.S. Schepers. 2011. Estimation of surface soil organic matter using a ground-based active sensor and aerial imagery. *Precision Agric*. 12: 82-102

Scofield, C.I. (Ed.) 1945. *The Holy Bible (The Scofield Reference Bible)*. New York, NY: Oxford Uni. Press.

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.

Shonk, J.L., L.D. Gaultney, D.G. Schulze, and G.E. Van Scoyoc. 1991. Spectroscopic sensing of soil organic matter content. *Trans ASAE* 34: 1978-1984.

Sojka, R.E. and D.R. Upchurch. 1999. Reservations regarding the soil quality concept. *Soil Sci. Soc. Am. J*. 63: 1039-1054.

Solari, F., J. Shanahan, R. Ferguson, J. Schepers, and A. Gitelson. 2008. Active sensor reflectance measurements of corn nitrogen status and yield potential. *Agron. J*. 100: 571–579.

Steer, A., 1998. Making development sustainable. *Adv.Geoecology* 31: 857–865.

Stern, R., M. Ben-hur, and I. Shainberg. 1991. Clay mineralogy effect on rain infiltration, seal formation and soil losses. *Soil Sci*. 152: 455-468.

Stenberg, B. 2010. Effects of soil sample pretreatments and standardized rewetting as interacted with sand classes on VIS-NIR predictions of clay and soil organic carbon. *Geoderma* 158: 15-22.

Stenberg, B., R.A. Viscarra Rossel, A.M. Mouazen, and J. Wetterlind. 2010. Visible and near infrared spectroscopy in soil science. *Adv. Agron*. 107: 163-215.

Sudduth, K.A., and J.W Hummel. 1993a. Portable, near-infrared spectrophotometer for rapid soil analysis. *Trans. ASAE* 36: 185-193.

Sudduth, K.A., and J.W Hummel. 1993b. Soil organic matter, CEC, and moisture sensing with a portable NIR spectrophotometer. *Trans. ASAE* 36: 185-193.

Sudduth, K.A., N.R. Kitchen, E.J. Sadler, S.T. Drummond, and D.B. Myers. 2010. Chapter 13: VNIR spectroscopy estimates of within-field variability in soil properties. In *Proximal Soil Sensing, Progress in Soil Science*, 153-163. R.A. Viscarra Rossel et al. eds. Dordrecht, Netherlands: Springer.

Tekin, Y., Z. Tumsavas, and A.M. Mouzen. 2012. Effect of moisture content on prediction of organic carbon and pH using visible and near-infrared spectroscopy. *Soil Sci. Soc. Am. J*. 76: 188-198.

Udelhoven, T., C. Emmerling, and T. Jarmer. 2003. Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression: A feasibility study. *Plant Soil* 251: 319-329.

Viscarra Rossel, R. 2009. The Soil Spectroscopy Group and the development of a global soil spectral library. *NIR N*ews 20: 17-18.

Viscarra Rossel, R.A., and T. Behrens. 2010. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma* 158: 46-54.

Viscarra Rossel, R.A., R.N. McGlynn, and A.B. McBratney. 2006. Determining the composition of mineral-organic mixes using UV-VIS-NIR diffuse reflectance spectroscopy. *Geoderma* 137: 70-82.

Waiser, T.H., C.L.S. Morgan, D.J. Brown, and C.T. Hallmark. 2007. In-situ characterization of soil clay content with visible near-infrared diffuse reflectance spectroscopy. *Soil Sci. Soc. Am. J*. 71: 389-396.

Wetterlind, J., B. Stenberg, and M. Soderstrom. 2008. The use of near infrared (NIR) spectroscopy to improve soil mapping at the farm scale. *Precision Agric*. 9: 57-69.

APPENDIX

A. DETAILED PROCEDURES FOR ASD SOIL ANALYSIS

INSTRUMENT SET-UP:

- Instrument and lamp warmed up for minimum of 1 hour before measurements

- Configuration of instrument settings:

  - bare foreoptic

  - number of spectrum samples: 30

  - number of dark current samples: 50

  - number of white reference samples: 50

  - wavelength interval: 1 nm

SAMPLE DATA COLLECTION:

Wet Samples

- Load samples into glass-bottomed sample dish, filling sample to the top of the dish. The top of the sample must be flush with top of the dish for consistent data collection. Most samples were relatively granular and could be prepared by simply filling the glass bottom dish flush with the top. Deeper soil samples which had greater clay content could not be done in the same way. These samples were flattened into the glass bottom viewing dish with a laboratory spatula to produce a flat viewing surface.

- Place Spectralon (Labsphere, Inc., North Sutton, N.H.) white reference in one dedicated dish of the same type. Click the "WR" button to reference each soil scan to this baseline reading.

- Place soil sample dish under optical window; rotate sample dish 30 degrees between each reading for a total of three readings.

Dry Samples

- Load samples into glass-bottomed sample dish filling sample to the top of the dish.

- Place white reference in one dedicated dish of the same type. Click the "WR" button to reference each soil scan to this base line reading.

- Place sample dish on optical window; rotate the sample dish 30 degrees between each reading obtaining a total of three readings.

Guidelines to follow while taking measurements

- Use a Halon white reference (Spectralon is the commercial product) to: (i) optimize (Opt) and (ii) take the white reference (WR) measurement. It is important to use the same geometry for these measurements as for the soil measurements

- Assess the white reference spectrum. It should be a straight line at 1 reflectance with no more than approximately 0.03 reflectance units of noise. Noise will mostly occur in the ends at 350 nm and at 2500 nm.

- Measure the reflectance of soil samples. Make three measurements per sample, rotating the sample container after each measurement so that each time measures a different 'spot' of soil.

- During measurement sessions, after every 5 samples (i.e. 15 measurements) optimize the instrument, take the white reference and assess it as in step 2 above, before continuing.

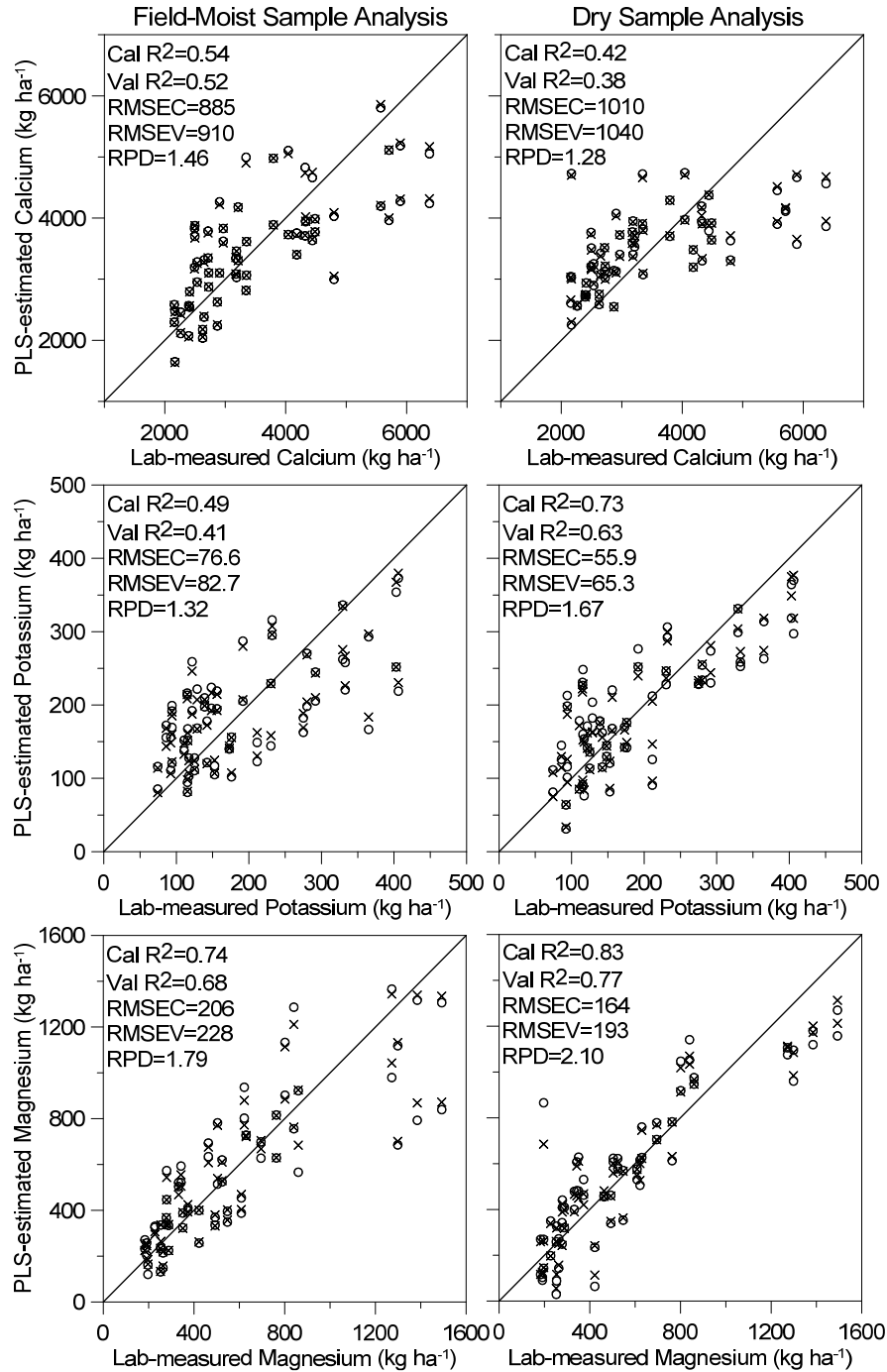## B. PARTIAL LEAST SQUARES CROSS-VALIDATION SCATTER PLOTS

Figure B.1. Surface soil property estimates of field-moist (left) and dry (right) samples by partial least squares (PLS) analysis of spectral data as compared to those obtained by standard laboratory analysis. Both calibration (Cal; x) and validation (Val; o) points are included as well as ratio of predicted deviation (RPD).
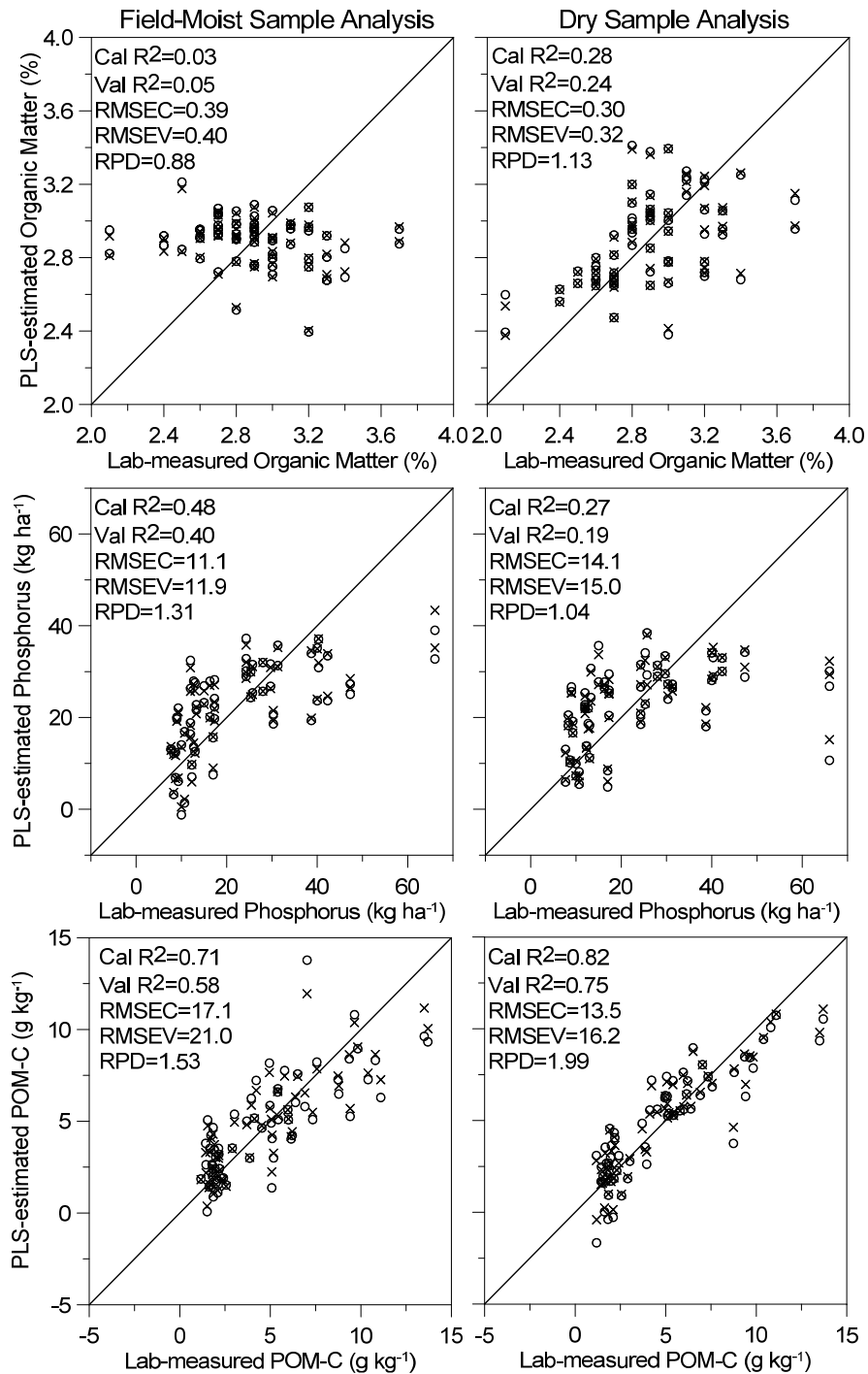
Figure B.2. Surface soil property estimates of field-moist (left) and dry (right) samples by partial least squares (PLS) analysis of spectral data as compared to those obtained by standard laboratory analysis. Both calibration (Cal; x) and validation (Val; o) points are included as well as ratio of predicted deviation (RPD) for organic matter, phosphorus, and particulate organic matter-carbon (POM-C).
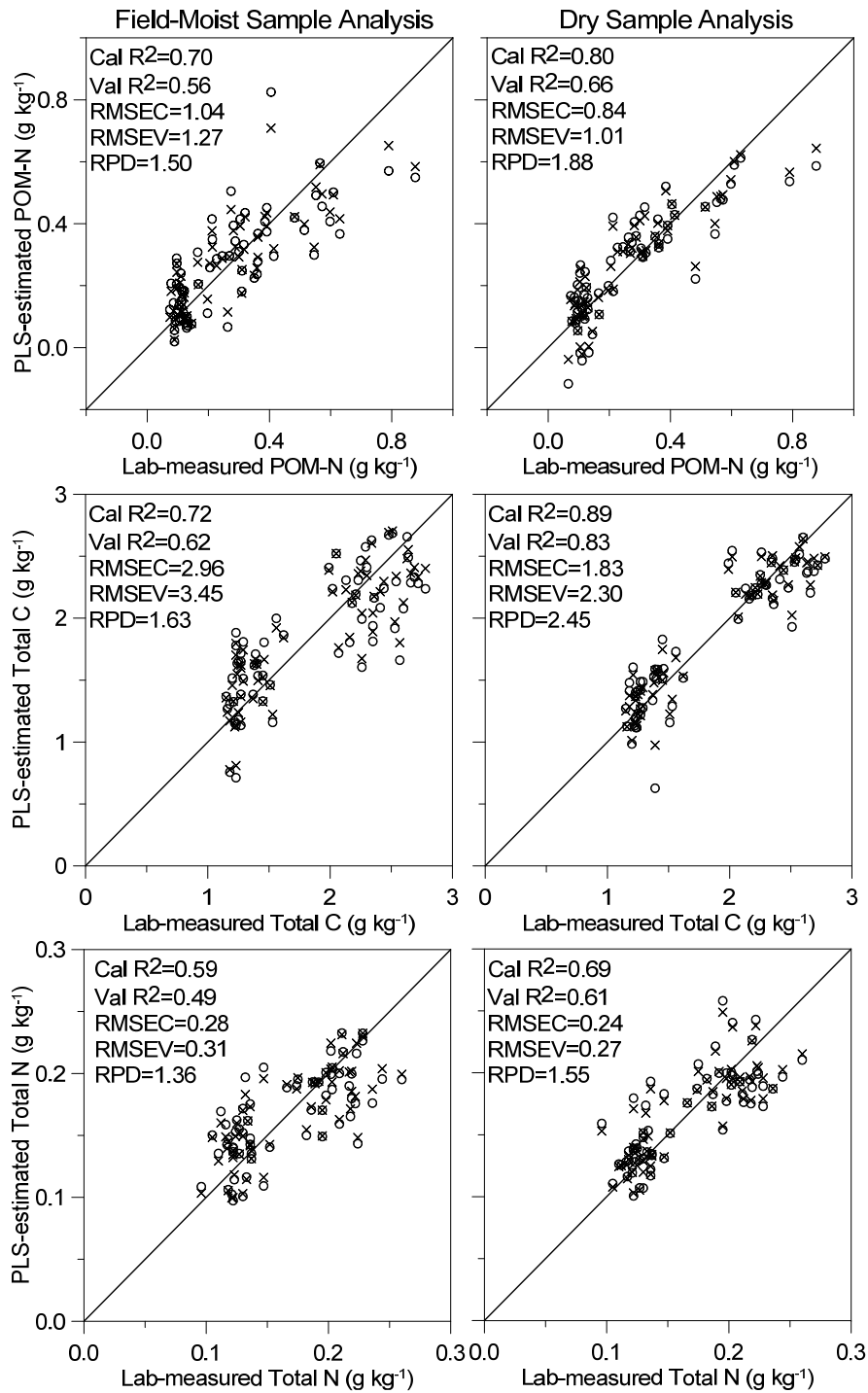
Figure B.3. Surface soil property estimates of field-moist (left) and dry (right) samples by partial least squares (PLS) analysis of spectral data as compared to those obtained by standard laboratory analysis. Both calibration (Cal; x) and validation (Val; o) points are included as well as ratio of predicted deviation (RPD) for particulate organic matter-nitrogen (POM-N), total carbon (C), and total N.
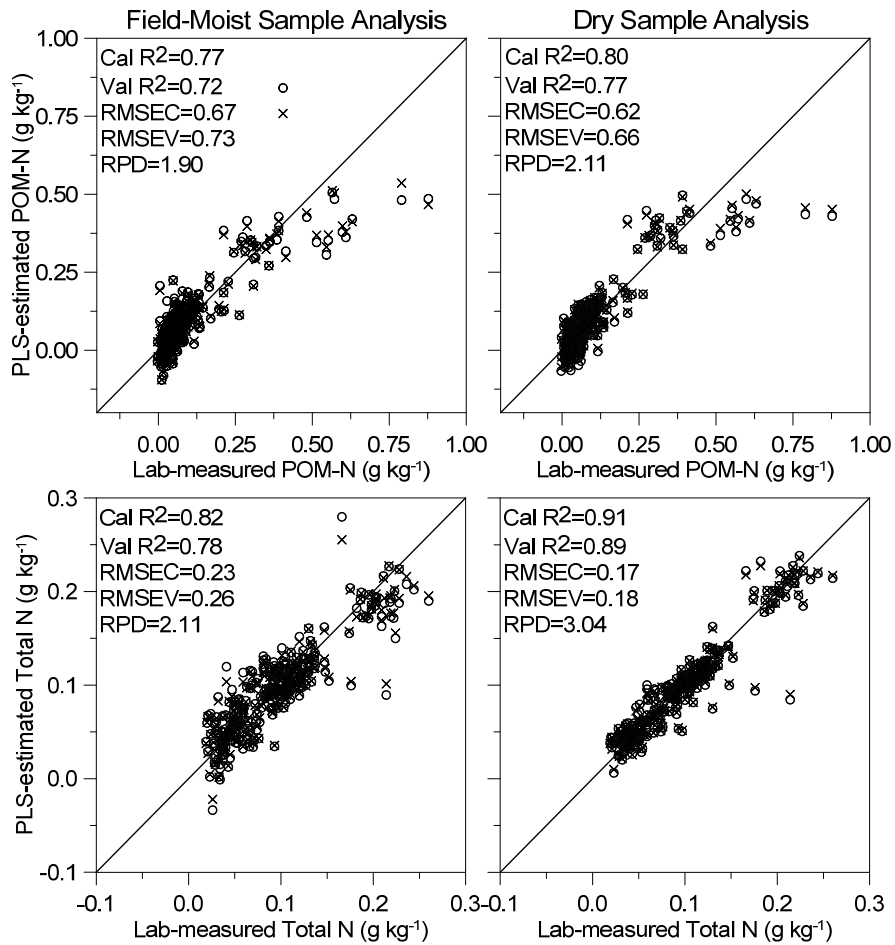
Figure B.4. Profile soil property estimates of field-moist (left) and dry (right) samples by partial least squares (PLS) analysis of spectral data as compared to those obtained by standard laboratory analysis. Both calibration (Cal; x) and validation (Val; o) points are included as well as ratio of predicted deviation (RPD) for particulate organic matter-nitrogen (POM-N) and total N.