

**CHARACTERIZATION OF SOIL SHRINK-SWELL POTENTIAL
USING THE TEXAS VNIR DIFFUSE REFLECTANCE
SPECTROSCOPY LIBRARY**

A Senior Scholars Thesis

by

KATRINA HUTCHISON

Submitted to the Office of Undergraduate Research
Texas A&M University
in partial fulfillment of the requirements for the designation as

UNDERGRADUATE RESEARCH SCHOLAR

April 2008

Major: Bioenvironmental Sciences

**CHARACTERIZATION OF SOIL SHRINK-SWELL POTENTIAL
USING THE TEXAS VNIR DIFFUSE REFLECTANCE
SPECTROSCOPY LIBRARY**

A Senior Scholars Thesis

by

KATRINA HUTCHISON

Submitted to the Office of Undergraduate Research
Texas A&M University
in partial fulfillment of the requirements for the designation as

UNDERGRADUATE RESEARCH SCHOLAR

Approved by:

Research Advisor:

Associate Dean for Undergraduate Research:

Cristine Morgan

Robert C. Webb

April 2008

Major: Bioenvironmental Science

ABSTRACT

Characterization of Soil Shrink-Swell Potential Using the Texas VNIR Diffuse
Reflectance Spectroscopy Library
(April 2008)

Katrina Hutchison
Department of Bioenvironmental Sciences
Texas A&M University

Research Advisor: Dr. Cristine Morgan
Department of Soil and Crop Sciences

Shrinking and swelling soils cause extensive infrastructure and economic damage worldwide. Shrink-swell soils are of great concern in Texas for two reasons, 1) Texas has the most acreage of shrink-swell soils in the United States, and 2) yearly evapotranspiration rates exceed those of precipitation creating optimal conditions for soil wetting and drying cycles. This study was conducted to determine if visible near infrared diffuse reflectance spectroscopy (VNIR-DRS) can be used to predict the coefficient of linear extensibility (COLE) of soils. If successful, VNIR-DRS would provide a means to rapidly and inexpensively quantify a soil's shrink-swell potential real-time. Using soils that have been previously analyzed and archived in the Texas Agrilife Research Soil Characterization Laboratory, our objectives were to: 1) predict the coefficient of linear extractability (COLE) using spectroscopy, 2) predict COLE using measurements of total clay and cation exchange capacity (CEC), and 3) compare the two models.

A total of 2454 soil samples were scanned to create the Texas spectral library. Of these samples, 1296 had COLE measurements. Seventy percent of the COLE samples were randomly selected to build a calibration model using partial least squares regression. The remaining thirty percent were used to validate the calibration model. The coefficient of determination (R^2), root mean square deviation (RMSD), and relative percent difference (RPD) were calculated to assess the prediction models. The COLE prediction using spectroscopy had an R^2 , RMSD, and RPD of 0.61, 0.028, and 1.6, respectively. Using stepwise regression and backward elimination, we determined that CEC and total clay together were the best predictors of COLE with R^2 , RMSD, and RPD of 0.82, 0.019, and 2.3, respectively. According to the RPD, using spectroscopy to predict COLE has some predictive value, while using CEC and total clay is more effective and stable. However, spectroscopy data collection is more rapid and has fixed costs.

NOMENCLATURE

COLE	Coefficient of Linear Extensibility
VNIR-DRS	Visible-Near Infrared Diffuse Reflectance Spectroscopy
CEC	Cation Exchange Capacity
RPD	Relative Percent Difference (Standard Deviation /RMSD)
RMSD	Root Mean Square Deviation

TABLE OF CONTENTS

	Page
ABSTRACT	iii
NOMENCLATURE.....	v
TABLE OF CONTENTS	vi
LIST OF FIGURES.....	vii
LIST OF TABLES	viii
 CHAPTER	
I INTRODUCTION.....	1
Literature Review	4
II METHOD.....	12
Laboratory Analysis	12
Data Preprocessing.....	13
Spectral Predictions.....	13
Pedotransfer Functions	14
III RESULTS AND DISCUSSION	16
Sample Description	16
Predicting COLE Using Total Clay and CEC	19
Predicting COLE Using Spectroscopy	23
IV SUMMARY AND CONCLUSIONS.....	27
REFERENCES	29
CONTACT INFORMATION	31

LIST OF FIGURES

FIGURE	Page
3.1 Predicted vs. measured COLE values of the validation set for (a) clay content and (b) clay content and CEC.	22
3.2 Predicted vs. measured COLE content using spectroscopy for the validation data.	25

LIST OF TABLES

TABLE	Page
3.1 Summary statistics for 1296 soils in the Texas spectral library	17
3.2 Summary statistics for calibration and validation datasets	18
3.3 Results of COLE predictions using spectroscopy, clay content and cation exchange capacity (CEC), and clay content.....	21
3.4 Shrink-swell classes and the corresponding COLE values	26

CHAPTER I

INTRODUCTION

Shrinking and swelling soils are known to cause extensive infrastructure and economic damage worldwide. Shrink-swell soils, which are high in clay content, have the potential to shrink or swell with changing moisture. Some are known to increase 150 times their size. Shrink-swell soils are of great concern in Texas for two reasons, 1) Texas has the most acreage of shrink-swell soils in the United States, and 2) yearly evapotranspiration rates exceeds those of precipitation creating optimal conditions for soil wetting and drying cycles (Godfrey et. al., 1973). If you have ever been to central Texas and have seen the cracks in the ceilings and bumps in the interstate you have seen the effect of these soils.

Whether the purpose of land use is agricultural or development purposes, the key to timely abatement and management is knowing the shrink-swell potential. Current methods of quantifying soil shrink-swell potential are time consuming and expensive. To calculate the coefficient of linear extensibility (COLE), which is the shrink-swell potential of soil natural fabric, an intact soil core must be collected in triplicate and the volume change of the core must be measured at moist and drying conditions. The

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

engineering equivalent of COLE uses a dried ground soil sample and measures the change in length after drying. Visible near infrared diffuse reflectance spectroscopy (VNIR-DRS) can be used to scan soils rapidly and in situ, and possibly quantify a soil's shrink-swell potential real-time, saving time, and fixing costs for developers and resource managers in shrink-swell areas.

Soil shrink-swell potential is primarily a function on two soil properties, clay mineralogy, particle size distribution or percent clay and fine clay particles (Wilding, 1998). Soil scientists have proven the ability of visible and near infrared diffuse reflectance spectroscopy (VNIR-DRS) for *in situ* and lab characterization of soil mineralogy and clay content (Waiser, 2007; Brown et al., 2005a; Brown et al., 2005b; Chabrilat et al., 2000). Soil minerals have distinct spectral signatures which are identified as distinct spectral absorbencies at different wavelengths.

In addition to quantifying clay mineralogy, clay content, and other soil properties that effect COLE research shows encouraging results for quantifying shrink-swell using VNIR-DRS. One study has used VNIR- DRS to categorize a geographically limited group of dried and ground soils into high, medium, and low potential for shrinking and swelling (Kariuki et al, 2004). Goetz et al. (2001) was able to predict the smectite content of soils with an r^2 value of 0.83 and place the soils in a shrink-swell class based on mineralogy. Research using reflectance spectroscopy has mainly focused on soils in localized areas. Kariuki et al. (2004) collected 198 samples in southern Spain and Goetz

et al. (2001) collected 178 samples from the Colorado Front Range. There has not been any research using VNIR spectral data to directly quantify a soils COLE value.

That is why we want to evaluate soils from all over Texas. Texas is a large state and the soils have formed from many different parent materials. Additionally a wide range of climate and vegetation can affect soil reflectance. We are fortunate to have access to a soil characterization lab here at Texas A&M University which has been collecting and characterizing samples since the 1970's. After they are analyzed, they are stored in a warehouse. They are now available for us to start scanning.

The overall objective of this research is to create a Texas spectral library that will expand the regional extent of available soil spectra. The spectral library can be used by anyone interested in rapidly quantifying soil properties using VNIR-DRS. The specific objectives of this research are to test the predictive ability of the spectrometer by performing the following tasks:

- 1) Create a VNIR-DRS spectral library from archived Texas soils,
- 2) Provide a summary and descriptive statistics of the soils in the spectral library, and
- 3) Create predictor models of COLE, clay content, and CEC that might affect COLE using the VNIR-DRS spectrometer.

Literature Review

Shrinking and Swelling Soils

Shrinking and swelling soils are soils that are high in clay content and have the potential to shrink or swell under changing moisture conditions (Nelson, 1992). The shrinking and swelling action is due to the clay's ability to attract and absorb water. Some shrink-swell soils are known to increase 150 times their size. Shrink-swell soils are very sticky when wet and usually crack under drying conditions. Therefore during drying seasons cracking of the soil surface can be indicative of shrink-swell soils. These cracks can be very large and can cause damage to buildings located on top of these shrink-swell soils. Shrinking and swelling soils are known to cause extensive infrastructure and economic damage worldwide. In fact in the United States, shrink-swell soils rank second to insect damages in economic losses. Shrink-swell soils are accountable for more damage to structures, such as buildings and pavement, than any other natural disaster including hurricanes and floods (Nelson, 1992). On top of causing extensive infrastructure damage the cracking of these soils allow for preferential flow. Due to the characteristic shrinking and swelling of shrink-swell soils, the hydrology of a soils landscape can be dramatically changed. Under ponded conditions the water channels through the cracks. As a result shrink-swell soils have the potential to transport water and pollutants such as pesticides and herbicides from agricultural practices and hazardous waste sites (Harris et al., 1994; Kelly and Pomes, 1998). This creates significant consequences for ground-water quality because contaminants can flow through the cracks directly into the groundwater.

Shrink-swell soils are of great concern in Texas for two reasons, 1) Texas has the most acreage of shrink-swell soils in the United States, and 2) yearly evapotranspiration rates exceed those of precipitation creating optimal conditions for soil wetting and drying cycles (Godfrey et. al., 1973). There are 12 million ha of shrink-swell soils in the US; Texas contains 6.5 million ha (Coulombe et al, 1996). These soils can be found in the most populated areas of Texas including the Coastal Plains (Houston) and the Blackland Prairies of central Texas (Dallas-Fort Worth, Austin, and San Antonio). As well as being highly urbanized with three of the top ten largest cities in the US, Texas is subjected to intensive and extensive agriculture practices. There is a greater possibility for infrastructure damages to the large cities as well groundwater contamination from the heavy agricultural practices.

Shrink-swell soils are the most difficult soils to manage because their physical and chemical properties vary dramatically over space and time. The spatial variation of soil properties is due to self mixing caused by soil displacement in between the cracks (Coulombe et al, 1996). Temporally shrink-swell soils are difficult to manage because the swelling cycle of these soils can take 5 to 8 years and are not immediately noticeable (Nelson, 1992). Management of shrink-swell soils is also difficult because the swelling is not determined by one factor but is the outcome of many interacting factors. Amount of clay, type of clay mineral, cation exchange capacity (CEC), organic matter, and the soil moisture all play a significant role to swelling (Kariuki et al, 2004). The swelling of

these soils is due to the presence of clay minerals with high surface areas that have the potential to accumulate water. These minerals can be categorized in to three main groups: smectite, illite, and kaolinites, listed in order of highest to lowest potential for swelling. However some studies show that in general soil shrink-swell potential increases with increased clay content, regardless of mineralogy (Schafer and Singer, 1976a). When land is being used for development purposes, the key to successful management and planning is knowing the shrink-swell potential of the soil before construction. When shrink-swell soils are abundant and land use is in its natural conditions, such as in the case of contamination clean up or agriculture practices, the key to understanding hydrology and solute transport is knowing the spatial extent and variability of soil shrink-swell potential.

Coefficient of Linear Extensibility

The most common way to quantify a soil's shrink-swell potential is calculating its coefficient of linear extensibility (COLE). The coefficient of linear extensibility (COLE) is the shrink-swell potential of soil natural fabric. A soil natural fabric describes a soil sample which has not been disturbed from its natural state. To measure COLE an intact soil core must be collected in triplicate and the volume change of the core must be measured at moist and drying conditions. When it is not possible to extract intact soils samples the $COLE_{rod}$ method is often used. Vaught et al (2006) has found that COLE is highly correlated with $COLE_{rod}$ with an r^2 value of 0.88. The $COLE_{rod}$ is an engineering method which uses a dried ground soil sample molded into the form of a rod and

measures the change in length of this rod after drying. Both of these techniques can take weeks to complete in the laboratory and can be very expensive if many samples need to be analyzed. Therefore there is a need to develop quick and inexpensive techniques that can preserve the integrity of the soil system and save time and money for those needing to quantify a soil's shrink-swell potential (Islam et al, 2003). Proximal spectroscopy is a rapid sensing technique that is built on already collected COLE data. Using this technique you can get more COLE estimations faster. This will have the advantage of providing a means of mapping shrink-swell potential at a high spatial resolution. One possible use is making a high resolution 3D model of shrink-swell potential. This has applications to planning land development and modeling landscape hydrology.

Spectroscopy

Visible near infrared diffuse reflectance spectroscopy (VNIR-DRS) can be used to scan soils rapidly and in situ, and therefore quantify a soil's shrink-swell potential real-time and at a spatial resolution of 5-10 m. Spectroscopy is a study of light as a function of wavelength that has been emitted, reflected, absorbed, or scattered from a solid, liquid, or gas. Clay minerals have distinct spectral signatures because of overtones and combination bands from chemical bonds within the soil minerals (Clark, 1999). For example kaolinite, smectite, and illite occur in the clay fraction of soil and have distinct spectral absorption features. Kaolinite $[Al_2Si_2O_5(OH)_4]$ has two hydroxyl bands near 1400nm and 2200nm (Hunt and Salisbury, 1970). Smectite $[(Na,Ca,Mg,K)_{0.3}Al_{2.7}Si_{3.3}O_{10}(OH)_2]$ has two strong water bands around 1400, 1900, and

2200nm (Goetz et al, 2001). Illite [$\text{KA}l_2(\text{AlSi}_3\text{O}_{10})(\text{F},\text{OH})_2$] has hydroxyl bands at 1400nm and between 2200 and 2600nm. (Hunt and Salisbury 1970). Therefore the clay content measured by VNIR is due to the recognition of these distinct spectral signatures of the common clay minerals.

Recent research has not only proven the ability of spectroscopy to predict mineralogy but also a soil's total clay content. Sheperd, (2002) were able to predict clay content of 1000 air dried ground samples from east and South Africa with an R^2 value of 0.80. Islam et al, (2003) predicted clay content of 161 air dried ground sample from Australia with an R^2 and RPD value of 0.73 and 1.9 respectively. Brown et al. (2005b) created a global spectral library from 3768 air dried ground soil samples from all 50 states and in Europe. They were able to predict clay content with an RMSD value of $54 \text{ g}\cdot\text{kg}^{-1}$. There has been some research predicting clay content of insitu soils samples. Questions have been raised as to how the heterogeneity of soil and soil water content affect spectral predictions. Chang et al., (2001) used 802 soil samples to create natural product cells and predicted clay content with an R^2 and RPD value of 0.67 and 1.71 respectively. Waiser (2007) found that water and heterogeneity of the soil does not significantly effect clay prediction. Waiser (2007) compared clay content prediction of 72 Texas soil samples for both field moist insitu, field moist smeared insitu, air dried insitu and air dried ground. They obtained R^2 values of 0.83, 0.75, 0.92, and 0.84 respectively.

Soil shrink-swell potential is primarily a function on two soil properties, clay mineralogy and clay percentage (Wilding, 1998). Vaught et al (2006) found COLE to be highly correlated with total clay $r^2=0.88$. Because soils shrink-swell potential is directly related to clay mineralogy and clay percentage and soil scientists have proven the ability of VNIR-DRS for *in situ* and lab characterization of soil mineralogy and clay content (Brown et al, 2005; Waiser, 2007) we expect spectroscopy to be able to predict COLE values. A soils COLE value is also influenced by other soil properties including cation exchange capacity (CEC), organic matter, and soil moisture (Kauriuki et al, 2003). Using stepwise regression and backward elimination, we determined that CEC alone was the best predictor of COLE with R^2 , RMSD, and RPD of 0.79, 0.021, and 2.2, respectively. Research has proven the ability of spectroscopy to predict not only clay content but these other soil properties that are highly correlated with COLE such as CEC. Chang et al. (2001) and Shepherd (2002) were able to predict CEC both with an R^2 of 0.81. Dunn et al. (2002) predicted CEC in the topsoil and subsoil of 550 air dried ground samples with an R^2 value of 0.90 from the topsoil and 0.80 from the subsoil.

In addition to quantifying clay mineralogy, clay content, and other soil properties that effect COLE research shows encouraging results for directly quantifying a soils shrink-swell potential using VNIR-DRS. One study has used VNIR- DRS to categorize a geographically limited group of dried and ground soils into high, medium, and low potential for shrinking and swelling based on clay mineralogy (Kariuki et al, 2004). Goetz et al. (2001) was able to predict the smectite content of soils with an R^2 value of

0.83 and place the soils in a shrink-swell class again based on mineralogy. Research using reflectance spectroscopy has mainly focused on soils in localized areas. Kariuki et al. (2004) collected 198 samples in southern Spain and Goetz et al. (2001) collected 178 samples from the Colorado Front Range. Research has used mineralogy to predict a soils shrink-swell potential class, there has not been research performed using VNIR spectral data to directly quantify a soils COLE value.

To more rigorously test the capability of VNIR DRS models that predict soil shrink-swell potential we would need soil samples that spanned the range of soil composition. Islam et al. (2003) found that a limited number of soil samples from similar parent material and land use can limit the robustness of the calibration. A large soil spectral libraries (10^4 - 10^5) which will represent the samples used for validation is needed to create calibration models (Brown et al, 2005b; Islam et al, 2003). Brown et al 2005b has expressed a need to expand soil-spectra libraries by scanning state and national soil archives. Models constructed using soils with greater geographic diversity would make a more robust model (Waiser, 2007; Brown et al 2005b). That is why we want to evaluate soils from all over Texas. Texas is a large state, and the soils have formed from many different parent materials. Additionally Texas has a wide range of climate and vegetation which can affect soil reflectance. We are fortunate to have access to a soil characterization lab here at Texas A&M which has been collecting and characterizing samples since the 1970's. They have been stored and are available for scanning into a spectral library.

The *overall goal* of this project is to determine if VNIR-DRS is an effective tool for directly quantifying COLE. A Texas spectral library needs to be built that will expand the regional extent of available soil spectra associated with soil characterization data. The spectral library can be used by anyone interested in investigating the association of soil properties and VNIR spectra. If successful VNIR-DRS can be used to scan soils in situ to map soil shrink-swell potential in the field. The specific objectives of this research project are the following: 1) Create a VNIR-DRS spectral library from archived Texas soils; 2) Provide a summary and descriptive statistics of the soils in the spectral library; 3) Create predictor models of COLE, clay content, and CEC that might affect COLE using the VNIR-DRS spectrometer.

CHAPTER II

METHODS

Laboratory Analysis

The Texas VNIR-DRS spectral library was created with 2454 soil samples, archived by the Texas Agrilife Research Soil Characterization Lab. The characterization lab has been analyzing and archiving soil pedons from all across Texas since 1978. At the lab particle size distribution was measured using the pipette method (Kilmer and Alexander, 1949); cation exchange capacity (CEC) was measured by a modified procedure of USDA Handbook 60 (U.S. Salinity Laboratory Staff, 1969); and COLE was measured using the procedure described in the Soil Survey Laboratory Methods Manual, 1996. All soil analysis was performed two times to check for accuracy and any possible errors. The oven dried, 2 mm ground soil samples were stored in the warehouse after analysis. For this project, 2454 archives soil samples were transferred to 20 ml vials for easy storage and transfer. The soils were scanned from below with a mug lamp connected to an AgSpec® Pro(Analytical Spectral Devices, Inc) with a spectral range of 350-2500 nm. A Spectralon® panel with 99% reflectance was used to calibrate the spectrometer before soils samples were scanned; the same panel was used as a white reference to set reflectance to 100% before each scanning session (Waiser, 2007). Additionally for quality control and documentation purposes, seven calibrated standards were scanned before and after each soil scanning session. The reflectance values for each standard

were 99%, 80%, 60%, 40%, 10%, 5%, and 2%. Approximately 28 g of ground soil was placed into a borosilicate glass “puck” prior to scanning. Each soil sample was scanned twice with a 90° rotation between scans.

Data Preprocessing

The spectral data were pretreated by splicing, averaging, and taking the 1st and 2nd derivatives. The spectral data were spliced to produce seamless spectra where the three detectors overlapped across the wavelength spectra. Results of the two scans at 0° and 90° were averaged (mean). The 1st and 2nd derivatives were taken at 10 nm intervals to remove albedo. The mean and 1st and 2nd derivatives were taken after a cubic smoothing spline, implemented in the R “smooth spline” function (R Development Core Team, 2004) was fit to each raw spectral curve (Waiser, 2007).

Spectral Predictions

Out of the 2454 soil samples, only 1296 had COLE, total clay content, and CEC values. These 1296 were divided into a calibration and validation set. Seventy percent of the soil samples were used to create calibration models; while the remaining thirty percent were used for model validation. Selection criteria for the validation and calibration samples were created to insure independence between the validation and calibration data. Complete soil pedons (A-C horizons) were randomly selected so that a single pedon was

not split between the calibration and validation datasets. Using only the calibration data set models were built with the first derivative, 10 nm averaged spectral data to predict COLE. The prediction model was built using 1/25th cross validation partial least squares (PLS) regression in Unscrambler 9.0 (CAMO Tech, Woodbridge, NJ). The remaining thirty percent of the soil samples were used to validate the model. Negative COLE values were changed to zero before comparison of predicted COLE values to measured COLE values. Measured vs. predicted values of the validation samples were compared using simple regression. The coefficient of determination (R^2), root mean squared deviation (RMSD), and relative percent difference (RPD) were calculated to compare the accuracy of different PLS models. Statistical formulas to calculate RMSD, RPD and bias follow Gauch et al. (2003), Brown et al. (2005a) and Chang et. al. (2001).

Pedotransfer Functions

To assess the usefulness of a VNIR-based COLE model, simpler, multiple regression equations were created using other laboratory data. The lab data for the soil samples include partial to complete pedon analysis such as, CEC, fine clay, total clay, fine sand, total sand, total silt, organic carbon, calcium carbonate, bulk density, COLE, and pH. Backward elimination using only the calibration dataset was run using the R program to predict the correlation and regression equation between COLE and other soil properties. First backward elimination was run with all the soil properties. The p-value was assessed for each variable and the variable with the largest p-value was pulled out until all the p-

values were below 0.05. Four variables showed the most significance, fine clay, CEC, organic carbon, and pH. Based on the literature clay content is also highly correlated with COLE (Vaught et al, 2006). Different combinations of these five variables were used to predict COLE. Diagnostics to choose the best models(s) included p-value, residuals plots, R^2 values, and simplicity. The validation data were used to validate these models. After using the calibration model on the validation data, negative COLE values were changed to zero before comparison of predicted COLE values to measured COLE values. Measured vs. predicted values of the validation samples were compared using simple regression. . The coefficient of determination (R^2), root mean squared deviation (RMSD), and relative percent difference (RPD) were calculated to compare the accuracy of different PLS models. Statistical formulas to calculate RMSD, RPD and bias follow Gauch et al. (2003), Brown et al. (2005a) and Chang et. al. (2001).

CHAPTER III

RESULTS AND DISCUSSION

Sample Description

A total of 2454 soil samples were scanned to create the Texas spectral library. These soil samples have been archived by the Texas Agrilife Research Soil Characterization Lab. The characterization lab has been analyzing and archiving soils from all over Texas since 1978. Texas has a wide range of geologies, annual temperatures and annual precipitation; therefore the soil data base that was scanned is extremely variable in its parent material, mineralogy, and other soil formation factors (Godfrey et al., 1973). Of the 2454 scanned samples, 1296 had COLE measurements. The samples had ranges of COLE, total clay content, and CEC from 0.001 to 0.24 cm cm⁻¹, 0.07 to 84.2 %, and 0.7 to 105 cmol(+)kg⁻¹, respectively (Table 3.1). The calibration and validation data had very similar ranges and averages of soil properties (Table 3.2). The mean COLE value, CEC and total clay content for the calibration set were 0.05 cm cm⁻¹, 16.4 cmol(+) kg⁻¹, and 27.8 %, respectively. The mean COLE value, CEC and total clay content for the calibration set were 0.044 cm cm⁻¹, 15.7 cmol(+) kg⁻¹, and 27.1 % respectively.

Table 3.1. Summary statistics for 1296 soils in the Texas spectral library

Soil property	Units	Mean	Max.	Min.	CV
COLE [†]	cm cm ⁻¹	0.048	0.240	0.001	0.881
Clay	%	27.67	84.1	0.7	0.699
Cation exchange capacity	cmol(+) kg ⁻¹	16.33	105	0.7	0.839
pH		6.4	9.5	3.3	0.247
Base saturation	%	74.00	100	3	0.438
CaCO ₃ equivalent	%	19.3	86.7	0.1	1.048
Organic carbon	%	0.53	7.66	0.01	1.242

[†] COLE is Coefficient of Linear Extensibility.

Table 3.2. Summary statistics for calibration and validation datasets

Soil property	Units	Mean	Standard deviation
<i>calibration samples, n = 1031</i>			
COLE [†]	cm cm ⁻¹	0.05	0.042
CEC [‡]	cmol(+) kg ⁻¹	16.4	12.9
Clay	%	27.8	18.9
<i>validation samples, n= 265</i>			
COLE	cm cm ⁻¹	0.044	0.045
CEC	cmol(+) kg ⁻¹	15.7	15.9
Clay	%	27.1	21.0

[†] COLE is Coefficient of Linear Extensibility.

[‡] CEC is cation exchange capacity.

Predicting COLE Using Multiple Regression Total Clay and CEC

To compare spectroscopy predictions of COLE to more general prediction using pedotransfer functions, we first predicted COLE using total clay and CEC. Using the calibration data (n=1031) and backward elimination, fine clay content, CEC, organic carbon and pH were the best predictors of COLE ($R^2=0.79$). Replacing fine clay with total clay only slightly reduced the R^2 . Subsequently different combinations of these variables were tried; R^2 residuals and model simplicity were evaluated.

After evaluating selection criteria, the multiple regression of clay content plus CEC, gave the best overall prediction to COLE for our Texas soils, $R^2 = 0.77$. The residuals were homoscadastic, and the overall regression looked good. Though clay content was not the best predictor of COLE, soil clay content is relatively easy to measure compared to CEC. Hence clay content is a less expensive alternative for estimating COLE. Total clay alone as a predictor had an R^2 value of 0.60; while CEC alone had an R^2 of 0.76; however, adding total clay to the CEC regression improved residuals.

Using the validation data (n=265), total clay alone predicted COLE with an RMSD, r^2 , and RPD value of 0.029, 0.57, and 1.5 respectively. Total clay and CEC together predicted COLE with an RMSD, R^2 , and RPD value of 0.019, 0.82, and 2.3 respectively (Table 3.3). The measured laboratory values of COLE were graphed against the predicted values of COLE using total clay content alone and total clay content with CEC

(Figure 3.1). Total clay with CEC model predicted COLE better than total clay alone. The clay alone model has some bias in under predicting COLE ($-5.6e-06$). Including CEC in the model probably improved the estimation of COLE because clay mineral type has been associated with soil shrink-swell potential, and clay minerals with a higher CEC values, such as smectite, are known to have high shrink-swell potentials (Wilding, 1998). The CEC value may be further differentiating between smectitic and kaolinitic soils. Even though CEC and total clay content were able to predict COLE with an R^2 value of 0.82 there is much laboratory work which has to go into determining both the CEC and total clay content of a soil. Both procedures for predicting CEC and total clay content can be time consuming and expensive.

Table 3.3. Results of COLE predictions using spectroscopy, clay content and cation exchange capacity (CEC), and clay content.

	Spectroscopy	Prediction model used	
		Clay content & CEC	Clay content
RMSD, cm cm^{-1}	0.028	0.019	0.029
R^2	0.61	0.82	0.57
RPD	1.6	2.3	1.5

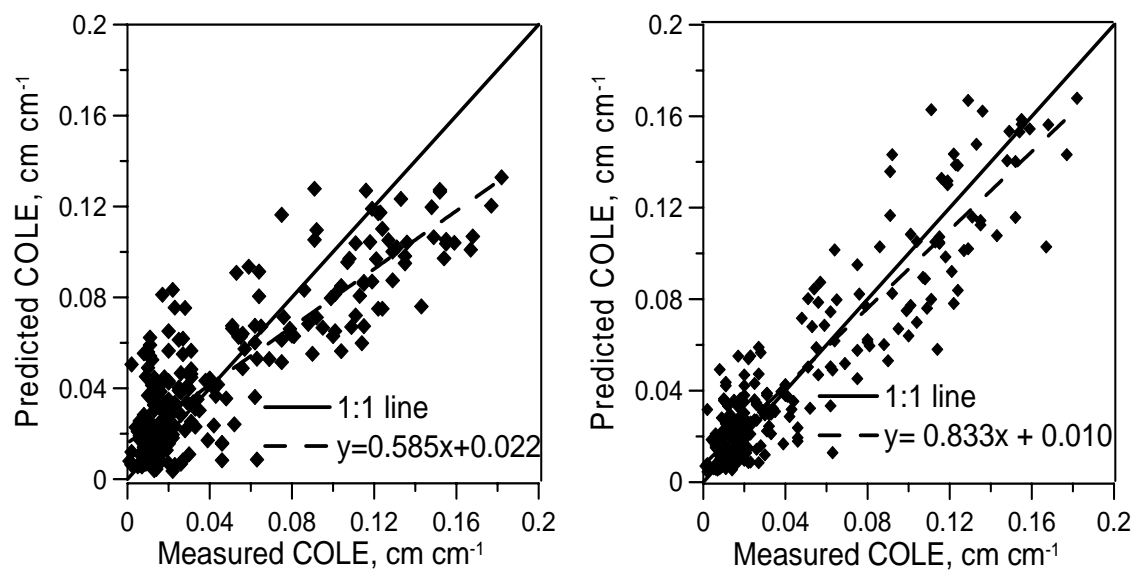


Fig. 3.1. Predicted vs. measured COLE values of the validation set for (a) clay content and (b) clay content and CEC.

Predicting COLE Using Spectroscopy

The results of the PLS prediction using spectroscopy are shown in Figure 3.2.

Spectroscopy was able to predict COLE with an R^2 , RMSD, and RPD value of 0.61, 0.028, and 1.6 respectively (Table 3.3). Spectroscopy predicted COLE better than predicting COLE with total clay content alone. However, clay content and CEC predicted COLE better than spectroscopy. Though spectroscopy was not as good as the clay plus CEC prediction, spectroscopy is very fast at scanning the soil compared to lab measurements of CEC and the costs of spectroscopy are fixed.

One useful way to interpret the spectroscopy prediction results is to look at how the prediction errors translate into predicting shrink-swell classes. According to the USDA NRCS, soils are classified into five shrink-swell classes, from very low to very high (Kariuki et. al. 2003). Given the prediction errors of spectroscopy, the results were still useful for classification purposes (Table 3.4). The spectroscopy prediction error was an RMSD of 0.028 cm cm^{-1} . In other words, the spectroscopy predictions will be within 2.8% of the actual COLE value, 66% of the time. The separation between the moderate, high and very high shrink-swell classes is greater than 3%. Therefore spectroscopy can correctly classify soils into these three shrink-swell classes. Total clay content and CEC prediction of COLE had an RMSD of 0.19 cm cm^{-1} . The CEC and total clay content predictions can be used to predict COLE within 1.9% of the actual COLE value as compared to the 2.8% of the spectroscopy predictions. Taking into consideration the

price and size of a project, this difference between the two predictions may not be practically significant.

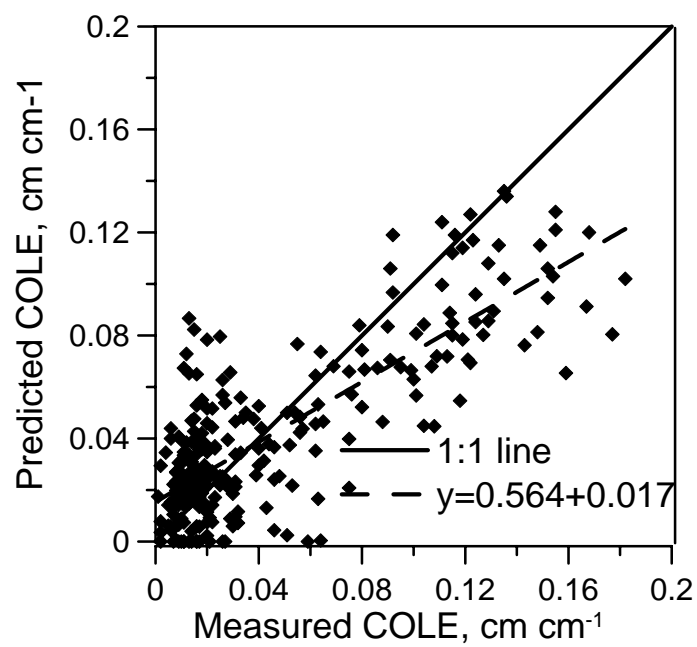


Fig. 3.2. Predicted vs. measured COLE content using spectroscopy for the validation data.

Table 3.4. Shrink-swell classes and the corresponding COLE values

Shrink-Swell Class	V. Low	Low	Moderate	High	V. High
COLE	0-.01	0.01-0.03	0.03-0.06	0.06-0.1	>0.1

CHAPTER IV

SUMMARY AND CONCLUSIONS

Visible near infrared DRS spectroscopy was used on dried ground (2 mm) soil samples to predict soil COLE values. Spectroscopy, combined with PLS regression predicted COLE with R^2 , RMSD, and RPD values of 0.61, 0.028, and 1.6 respectively. As a comparison to spectroscopy results, multiple regression analysis was used to predict COLE using total clay content alone and total clay content and CEC together. According to the RPD, using spectroscopy to predict COLE has some predictive value, while using CEC and total clay is more effective and stable. Though total clay did not predict the calibration data COLE values as well as the clay and CEC model, the clay-alone prediction was kept because soil clay is relatively cheaper, easier to measure, and more available in databases than CEC measurements. Total clay alone predicted COLE with an RMSD, R^2 , and RPD value of 0.029, 0.57, and 1.5 respectively. Total clay and CEC was the best of all three COLE prediction models together with a RMSD, R^2 , and RPD value of 0.019, 0.82, and 2.3 respectively.

Spectroscopy was able to predict COLE better than predicting COLE with total clay content alone. On the other hand total clay content and CEC predicted COLE better than spectroscopy. When using the prediction model results to classify the soils based on USDA NRCS shrink-swell potential classifications, the spectroscopy-based and clay and CEC- based prediction models both worked well. According to the NRCS classification

system, the separation between the moderate, high and very high shrink-swell classes is greater than 3%. Therefore spectroscopy should be able to distinguish between these three shrink-swell classes. CEC and total clay content together can be used to predict COLE within 1.9% of the actual COLE value as compared to the 2.8% of the spectroscopy predictions. When considering the amount of soils needing to be classified and the difference in cost for the two methods, the reduced accuracy in the spectroscopy-based prediction is tolerable.

Our results indicate that VNIR-DRS may be useful in predicting a soils shrink-swell potential. We envision using spectroscopy for in situ characterization of soils for greater spatial and vertical densities than is practical with conventional soil characterization techniques. To make this vision a reality, continued research is needed on in situ VNIR-DRS applications. These in situ studies should be careful to include a wide range of soil diversity and field conditions.

REFERENCES

- Brown, D.J., R.S. Bricklemyer, and P.R. Miller. 2005a. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. *Geoderma*. 129:251-267.
- Brown, D.J., K.D. Shepherd., G. Walsh., M.D. Mays., and T.G. Reinsch. 2005b. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma*. 132:273-290.
- Chabrilat, S., A.F.H. Goetz, L. Krosley, and H.W. Olsen. 2002. Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution. *Remote Sensing of Environment*. 82:431-445.
- Chang, C-W., D.A. Laird, M.J. Maushbach, and C.R. Hurburgh Jr. 2001. Near-infrared reflectance spectroscopy-principal components analyses of soil properties. *Soil Sci. Soc. Am. J.* 65:480-490
- Coulombe, C.E., L.P. Wilding, and J. B. Dizon. 1996. Overview of Vertisols: Characteristics and impacts on society. p. 289-375. *In* D. Sparks et al. (ed) *Advances in agronomy*. Academic Press, New York
- Dunn, B.W., H.G. Beecher, G.D. Batten, and S. Ciavarella. 2002. The potential of near-infrared reflectance spectroscopy for soil analysis-a case study from the Riverine Plain of south-eastern Australia. *Australian Journal of Experimental Agriculture*. 42:607-614.
- Godfrey, C.L., S. McKee, H. Oakes. 1973. General soil map of Texas. Department of Agriculture Communications. Texas A&M University, College Station, TX.
- Goetz, A.F.H., S. Chabrilat, and Z. Lu. 2001. Field reflectance spectrometry for detection of swelling clays at construction sites. *Field Analytical Chemistry and Technology*. 5:143-155.
- Islam, K., B. Singh, and A. McBratney. 2003. Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. *Australian Journal of Soil Research*. 41:1101-1114.
- Kariuki, P.C., F. Van Der Meer. and P.N.W. Verhoef. 2003. Cation Exchange Capacity (CEC) determination from spectroscopy. *Int. J. Remote Sensing*. 24:161-167.

- Kariuki, P.C., T. Woldai., and F. Van Der Meer. 2004. Effectiveness of spectroscopy in identification of swelling indicator clay minerals. *Int. J. Remote Sensing*. 25:455-469.
- Nelson, J.D. and D.J. Miller, 1992. *Expansive soils: problem and practice in foundation and pavement engineering*. Wiley, New York.
- 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.
- Vaught, R., K. R. Brye, and D. M. Miller. 2006. Relationships among coefficient of linear extensibility and clay fractions in expansive, stoney soils. *Soil Sci. Soc. Am. J.* 70:1983-1990.
- Waiser, T., C.L.S. Morgan, D.J. Brown, and C.T. Hallmark. 2007. In situ characterization of soil clay content with VNIR diffuse reflectance spectroscopy. *Soil Sci. Soc. Am. J.* 71:389-396.
- Wilding, L.P., and D. Tessier. 1998. Genesis of Vertisols: shrink-swell phenomena. p. 55-79. *In* L.P. Wilding et al. (ed) *Vertisols: Their distribution, properties, classification, and management*. Texas A&M University, College Station, TX.

VITA

Name: Katrina Hutchison

Professional Address: c/o Dr. Cristine Morgan
Department of Soil and Crop Sciences
Texas A&M University
College Station, TX 77843-2474

Email Address: katrinahutchison@hotmail.com

Education: B.S., Bioenvironmental Sciences; Minor, Environmental
Soil Sciences, Texas A&M University, May 2008
Undergraduate Research Scholar