METHODS FOR
HIGH-RESOLUTION SOIL-LANDSCAPE MODELING
IN MIDWEST UPLAND LANDSCAPES

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

DAVID BRENTON MYERS

Dr. Randall J. Miles and Dr. Newell R. Kitchen

Dissertation Supervisors

August 2008

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

METHODS FOR

HIGH-RESOLUTION SOIL-LANDSCAPE MODELING

IN MIDWEST UPLAND SOILS

presented by David Brenton Myers,

a candidate for the degree of doctor of philosophy,

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

Randall J. Miles

Newell R. Kitchen

Kenneth A. Sudduth

E. John Sadler

Sabine Grunwald

DEDICATION

To Melissa, David, and Luke.

ACKNOWLEDGEMENTS

If I have made any accomplishments in the pursuit of this degree they are because of the tremendous support provided by those around me. The members of my dissertation committee are exemplary of that support. Randy Miles started me down the path of Soil Science. It was through his introductory course that I realized that soil was the precise blend of physics, chemistry, and biology that I could really sink my teeth into. Beyond this introduction, Randy passed on to me a respect for the history, the mystique, the character, and the very nature of what it means to be a soil scientist. Thanks Randy for that, and for all of the beers. To Newell Kitchen I owe most of my success so far. Newell provided me an ecosystem of opportunities to creatively pursue my ideas. I cannot account for the unconditional confidence and support he has shown me. Newell has been a boss, a mentor, and at the same time a friend and I will not forget it. Together, Newell and Ken Sudduth somehow found a way to provide me the financial support and data I needed to keep going all these years - truly remarkable. I think I should thank Ken mostly for putting up with me and Newell. Ken's scientific cunning, skeptical approach, and blunt presentation of reality have scared me more than a few times, but these are the qualities that I appreciate in him the most. These are the qualities that have pushed me to be a better scientist. Ken has also become a mentor to me. I want to thank John Sadler for fostering an environment for excellence, and for holding me to that standard as well. Things changed

dramatically for the better when he arrived at CSWQ. John, in concert with Newell and Ken, is greatly responsible for the nitty gritty details of business and management that made it possible for me to pursue my degree. I want to thank Sabine Grunwald for hosting that Pedometrics conference in Naples! That conference has had a greater impact on my scientific perspective than any other conference or workshop. From that conference I gained a vision of my future self as a scientist. Sabine is representative of that vision, quantitative, integrative, and innovative. I'm looking forward to pursuing those qualities with her in Gainesville. Thanks Sabine for serving on my committee from a distance, and for the new opportunity. To all of my dissertation committee members, thank you indeed.

Then there are the people that really make things happen. Four men deserve my special thanks, Scott Drummond, Bob Mahurin, Kurt Holiman, and Matt Volkmann. Every part of my research has been enabled by the tools, facilities, data, advice, and sweat that these men provided in the pursuit of their own work. Beyond their professional excellence they are each excellent men and role models as fathers, and friends. Specifically I'd like to thank Scott for his help with instruments and computers, especially the penetrometer and mini-probe interfaces, and for help using them in the field. Thanks to Bob for performing surveys of GPS and EC_a, soil moisture measurements, pH and EC, and probably more. Thanks to Kurt for helping me design and construct a soil grinder, and then for using it. I shouldn't forget that Ken and Newell both did EC surveys too. Thanks to Matt for being impeccable, and for making it possible to do field research in so many ways. Three other men, Dr. Kelley Nelson and Randall

Smoot at the Greenley Research Farm and Ranjith Udawatta from the Center for Agroforestry were generous with their data, facilities, and equipment.

Several students had important roles in getting the work done. Kristen Kallash was indispensible and participated in nearly every step of getting preparing and measuring samples. She did most of the mini-probe and diffuse reflectance measurements (wet and dry), took field descriptions, weighed samples, labeled samples, bagged samples, dried samples, entered data, and kept me from making all kinds of mistakes. Without her excellent help this work would not have been completed. Thanks to Newell and John for dedicating her time to my project. Scott Willingham ran the penetrometer, ground soil samples, and weighed samples in the lab. Matt Holtman ground soil samples too.

I would like to acknowledge the overall support of the USDA-ARS Cropping Systems and Water Quality Unit, the University of Missouri, and the Department Soil, Environmental and Atmospheric Sciences. These institutions gave me a top-notch place to work and learn.

I close my acknowledgements with recognition of my family. Melissa, it has been a difficult road to be parents, and to be each engaged in our own rigorous pursuit of our professions. Thanks for compromising where you did and I apologize for not being more accommodating myself. David, thanks for making it all worth it.

D. Brenton Myers

July 22, 2008

ABSTRACT

Traditional soil mapping concepts do not sufficiently address the spatial resolution of some soil management problems. Continuous models of soil profiles and landscapes are needed to move beyond the categorical paradigm of horizons and soil map units. This work proposes a strategy combining sensors and empirical functions of profile properties to develop high resolution 3-D models of soil landscapes. The strategy proceeds in three steps as follows: 1) estimate soil profile properties at high resolution with the combined use of a diffuse reflectance spectroscopy (DRS) sensor and several soil electrical conductivity (EC) sensors, 2) model measured or sensor predicted soil profile properties with nonlinear peak functions, and 3) map the parameters of peak functions across the landscape to produce a continuous numerical soil-landscape model. Coherent depth translation (CDT) was introduced as a method to transform and combine sparse soil profile data into a single dataset for improved modeling. These methods were tested in the upland landscapes of northern Missouri. Sensors, especially DRS, successfully estimated profile clay and organic carbon. Peak functions were valuable for modeling profile clay content and covariates of clay. Coherent depth translation enabled the modeling pedogenic trends in peak function parameters. Prototype numerical soil-landscape models were developed for a lithosequence and a toposequence of common soil series in northern Missouri.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

CHAPTER 3

LIST OF NOMENCLATURE

η	skew parameter
BD	bulk density
CDT	coherent depth translation
CEC	cation exchange capacity
CI	cone index
D_t	translated depth (cm)
$D_{t,\emptyset}$	the origin of the translated depth scale
DTC	depth to claypan
EC_a	proximal bulk apparent soil electrical conductivity
$EC_{a\ d\text{-}sh}$	DUALEM-2S shallow channel
$EC_{a\ d\text{-}dp}$	DUALEM-2S deep channel
$EC_{a\ em}$	Geonics EM-38 deep channel
$EC_{a\ v\text{-}sh}$	Veris 3150 shallow channel
$EC_{a\ v\text{-}dp}$	Veris 3150 deep channel
EC_m	Wenner mini-probe bulk apparent soil electrical conductivity
EC_p	cone penetrometer bulk apparent soil electrical conductivity
EMI	electromagnetic induction
$L(\cdot)$	landscape or regional-scale soil forming process
mini-probe	Wenner style miniature EC_a hand probe
MLRA	major land resource area
OC	organic carbon
$P(\cdot)$	profile scale soil forming processes
PLSR	partial least squares regression a.k.a. projection to latent structures regression
RMSE(C,E)	root mean squared error (of calibration, estimation)
RPD	ratio of predicted deviation
TPD	transition peak depth
VIP	variable importance to projection
VNIR-DRS	visible-to-near-infrared diffuse-reflectance spectroscopy

CHAPTER 1

Coherent Depth Translation and Asymmetric Nonlinear Functions for Continuous Soil-Landscape Modeling

INTRODUCTION

Soil horizons and soil map units are fundamental concepts of pedology. They are the result of a taxonomic methodology, an algorithm, which organizes the complex relationships of the soil-landscape into abridged categorical units. The taxonomic algorithm transforms a continuous input into a discrete result. There is great utility in this function for communication and for compact representation, but information is lost by discretization. Problems in high resolution soil management and quantitative pedology demand a more continuous view of the soil landscape. However, the need still arises to have a compact representation.

Mathematical or statistical functions applied to soil landscapes are compact in that they represent a continuous volume of soil by a simple statement of numerical formulae and parameters. Numerical methods are needed to provide this compact representation of soil-landscapes as well as a change in support from the discrete to the continuous basis. This paper proposes a method which hierarchically combines nonlinear profile depth functions and landscape scale functions of their parameters into a compact numerical formulation. The method is enabled by the coherent depth translation (CDT) technique. These methods provide the reconstruction of continuous soil-landscape models from discretely sampled soil profile data and are demonstrated in the context of the soil-landscapes of northern Missouri.

Soil-Landscapes of Northern Missouri

Many upland soil landscapes of northern Missouri are formed in loess over glacial till. Glacial outwash on the Missouri river flood plain provided silty sediments which accumulated in loess deposits across northern Missouri. These deposits are deepest on summits near the Missouri river (~20 m), and thin with distance from the loess source (~1.5 m) (Ruhe, 1969; Fehrenbacher, 1973; Young and Geller, 1995). The gradient of loess thickness occurs largely parallel to the cumulative Pleistocene and Holocene paleowinds (Muhs and Bettis, 2000). Norton (1928) described a 'maturity sequence' of soils within this loess gradient. Soils along this pedogenic gradient succeed from well drained Mollisols, with only moderate clay illuviation (~30 %) and leaching of bases to moderately acid, poorly drained Alfisols with extreme (~60 %) clay accumulations, and significant leaching of bases (Bray, 1935; Whiteside, 1944). Since the age of the parent materials are contemporaneous, this is more accurately called a lithosequence, and spatial trends in apparent maturity are due to a gradient of loess thickness and weathering intensity (Smith, 1942).

Regional Quantitative Pedology Models

The Central Midwestern USA served as a testing ground for early investigations in quantitative pedogenesis. Indeed, the study region for this work intersects the areas covered by Jenny in his seminal works (Jenny, 1930; Jenny and Leonard, 1934; Jenny, 1935) based on the ideas of Dukochaev. Later, Jenny (1941) popularized the five factor model for soil formation as a system for quantitative pedology. This model was tested in Midwest USA datasets for a

variety of soil properties. In an effort to isolate the effects of the climate factor, Jenny and Leonard (1934) looked at the gradient in rainfall which occurs in this region from west to east. They discovered strong increasing linear and logistic spatial trends in soil properties across the rainfall gradient which they modeled by regression. These models described an increasing trend in profile average clay content, and other trends in surface horizon properties including pH, cation exchange capacity, and nitrogen and clay content. Jenny and Leonard (1934) also found linear and exponential relationships between average profile clay content and mean annual temperature.

Later quantitative pedology research similar to Jenney's also used regression to isolate factor effects. Richardson and Edmonds (1987) found several strong linear relationships along factor gradients. Their findings included increasing cation exchange capacity of B horizons in a biosequence of increasing forest cover, a lithosequence of decreasing sand content with decreasing depositional energy in a flood plain, and a complex piecewise linear relationship of Fe-Mn concretions in a hydrosequence. The regression approach produces functional results which can be applied in digital soil-landscape models. However, up to now these results have been limited to point estimates, single horizons, or aggregate estimates of profile properties.

Also common in the quantitative pedology research are studies which differentiate soil profiles or explain some pedogenic trend by qualitative or categorical factors that correlate to latent state factor gradients. A few examples of the discrete variables used are landscape position (Brubaker et al., 1993;

Brubaker et al., 1994), maturity sequence (Norton, 1928; Bray, 1934; Runge, 1973; Smeck, 1973), and soil series (Richardson and Riecken, 1977). The results of these works, while useful to advance the science of Pedology, are not as useful for continuous digital modeling of soil landscapes.

The quantitative pedology research described above provided estimates of one or a few point depths or profile aggregated estimates such as surface nitrogen, average profile texture, or B-horizon depth. None of these studies have estimated entire soil profiles for a significant portion of the large suite of soil property data that are needed for management solutions. Profile depth functions can describe the continuous distribution of soil properties in whole sections of the profile. The goal of this research is to find pedogenic relationships in the parameters of depth functions. This approach combines the spatial continuity of the factor trends described above with the profile continuity of a depth function. We develop this approach using case studies from a lithosequence of loess deposition and soil maturity found in the upper tier of counties in northern Missouri, and a common toposequence of soils from the claypan landscapes of northeast Missouri.

Profile Reconstruction Methods

Several methods have been proposed to approximate soil property depth distributions with empirical functions including: linear and polynomial functions, exponential and logarithmic functions (Dwyer et al., 1988; Cook and Kelliher, 2006), and orthogonal polynomials (Colwell, 1970). Non parametric functions and procedures are very useful to interpolate or smooth soil profile

data. Some valuable non-parametric techniques are local regression (Cleveland and Devlin, 1988), generalized additive models (Hastie and Tibshirani, 1990), splines (Webster, 1978), and pycnophylactic (area preserving) splines (Ponce-Hernandez et al., 1986; Bishop et al., 1999). However these techniques lack the advantages of a parametric model.

Modeling soil property depth profiles with parametric functions is problematic, in part due to the sparse and discrete nature of soil profile data (Colwell, 1970). Downwardly vectored processes of soil genesis can produce depth profiles which are anisotropic compounding the data sparsity problem (Walker and Green, 1976). Few parametric functions are capable of smoothly fitting sparse and anisotropic soil profile data. However, function parameters can quantitatively describe critical structures in a depth profile such as a gradient, a relative minimum or maximum, or an inflection point. Quantified landscape relationships among these functional structures are empirical descriptions of pedogenic processes. For this research we describe a specific class of nonlinear depth functions called asymmetric peak functions which can address these problems in modeling soil profiles. Asymmetric peak functions have the flexibility needed to handle anisotropy yet retain the value of parameters for examining soil-landscape relationships.

Most profile modeling methods, including non-linear functions, can be biased when compared to true depth functions of soil properties (Cook and Kelliher, 2006). The bias is due to two issues. First, soil property data is intrinsically sparse when sampled by discrete intervals or horizons, and must be

represented by upper and lower boundries or an average depth. Neither alternative accurately indicates the location of critical values within the horizon such as relative or absolute extrema. These intra-horizon features are integrated by the sampling method. Average depth may bias the function away from these unseen critical points, even if the function can be made to match the average magnitude of the true function, as in pycnophylactic splines. Second, the smooth and often symmetric nature of functions causes them to take 'unnatural' forms, even while they fit the data points well. The second form of bias can occur due to the anisotropic distribution of soil properties. Because only a small number of points are available to model a single profile, large gradients and multiple inflections in anisotropic regions cannot be fitted. Sparsity and bias can be mitigated by combining multiple profiles datasets via CDT.

Coherent Depth Translation

Coherent depth translation of soil property data may be used when a ubiquitous and pedogenically controlling feature, such as a local maxima or minima in clay content, is present in a group of soil profiles. The CDT takes the depth of the controlling feature as the origin ($D_{t,\emptyset}$) of a linear translation of the depth scale (fig. 1). Different soil landscapes, with different controlling genetic features (e.g. spodic, fragic) may require the use of a different soil property for CDT. Application of CDT to an ensemble of similar soils from across a local or regional landscape results in a coherent dataset. Once data are translated, parametric and non-parametric functions can be fit to the ensemble, resulting in

a global depth profile model. Models fit to the coherent data have continuous support and are resistant to near surface modifications in depth profiles.

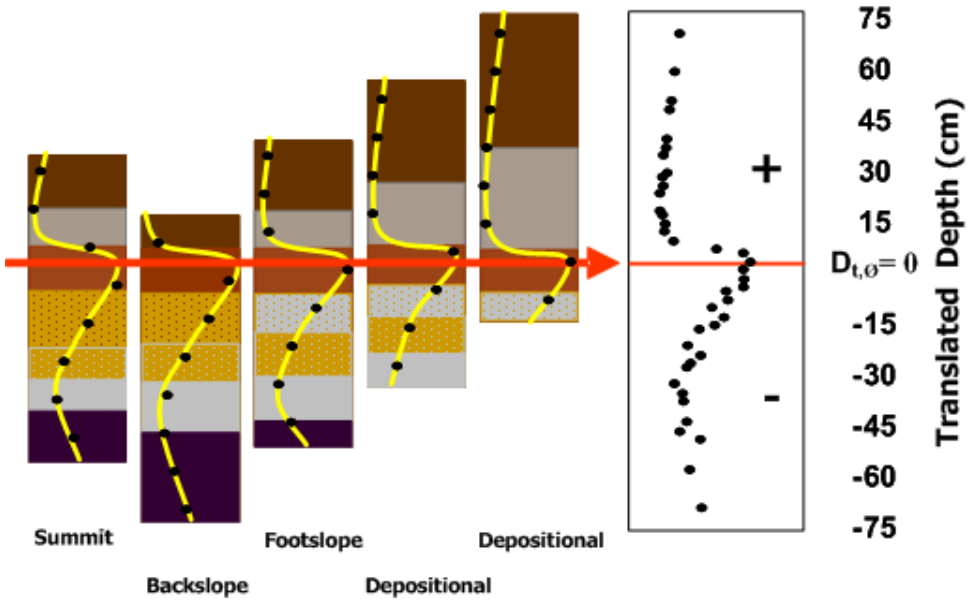

Figure 1. Conceptual diagram of coherent depth translation (CDT). The depth scale of each profile is individually translated such that the depth to a critical and controlling feature, here the depth to clay maximum, is the origin of the depth scale.

Comparing and aggregating related soil data require expressing properties on a universal scale with a common origin, usually depth below the surface. Unfortunately, surface elevation is a transient property due to post-settlement acceleration of erosion and deposition. Because of surface instability, soil properties viewed by depth from the surface can be noisy, and pedogenic trends can be hidden. Coherent depth translation can filter or quantify this noise from populations of soil profile data. Similar methods using the depth of an argillic horizon, peak gamma-ray attenuation, or radionucleide have been used by previous researchers to quantify soil erosion and redistribution (Lewis and Lepele, 1982; Olson et al., 1994). Another similar approach examined a

climosequence trend in depth to carbonates (Wenner et al., 1961). The CDT method was used in Myers et al. (2007) to explain the distribution of soybean roots in claypan soils. This paper formalizes the concept, presents methods for estimation of $D_{t,\emptyset}$ and extends the use of CDT to hierarchical modeling of soil profiles and soil landscapes.

Depth functions fit to depth translated data have properties of continuity, coherency, and resistance to altered surface effects (e.g. erosion/deposition). These properties allow better quantification of pedogenetic trends. Alfisols and Mollisols from northern Missouri upland landscapes have more or less prominent argillic horizons with a profile clay maximum (clay-max). Depth of the profile clay-max is used in this study as $D_{t,\emptyset}$ in order to demonstrate CDT in datasets from this landscape. The advantages of CDT for changing the support of discrete profile data and for examining surface modifications are demonstrated. Following the introduction of CDT, profile clay content is modeled for a common toposequence of claypan soils in northeast Missouri and the lithosequence of summit across northern Missouri. The Pearson IV (Pearson, 1895a) and asymmetric logistic (Tawn, 1988) nonlinear peak functions are used to model the clay distribution in soil profiles from the toposequence and lithosequence respectively. Finally, to illustrate the hierarchical method proposed, the parameters of the two peak function are modeled by linear functions of the pedogenic sequences to produce numerical soil-landscape models of clay content.

The overall objectives of this manuscript proceed in three integrated steps. First, CDT and its advantages are presented. Second, peak functions are introduced for modeling soil profiles. Third, parameters of peak functions are hierarchically modeled to build continuous soil-landscape models.

MATERIALS AND METHODS

Soil-Landscapes and Datasets

Three different datasets are presented in this study representing three scales: the field scale, a physiographic region (toposequence), and a sequence of physiographic regions (lithosequence). The lithosequence crosses four Missouri Major Land Resource Areas (MLRA): the Iowa and Missouri Deep Loess Hills (107), the Illinois and Iowa Deep Loess and Drift (108), the Iowa and Missouri Heavy Till Plain (109), and the Central Claypan Areas (113) (USDA-NRCS, 2006). The toposequence and field scale dataset come from upland soils in MLRA 113. Soil property data from nine soil series were used in these studies. Their taxonomic designation to the family level, general landscape position, and their order in the lithosequence and toposequence are listed in table 1. Only soil landscape clay distribution is considered in analysis of the toposequence and lithosequence datasets, while several more soil properties are considered from the field scale dataset.

The lithosequence clay content data comes from the primary interfluve divide and summit soil series distributed across the top tier of counties in Missouri (fig. 2). These stable landscape positions reflect variation mainly due to loess thickness and weathering intensity (as described above). The specific soils

chosen in order from west to east were: Marshall, Sharpsburg, Grundy, (Mollisols), Pershing, and Adco soils (Alfisols). Figure 2 depicts the generalized sequence of horizons that occurs along this genetic gradient. These five soil series are indexed by integers (1–5) in an ordinal categorical sequence representing their position in the pedogenic gradient (table 1). Measurements of clay content from these soil series were collected by the Missouri Cooperative Soil Survey over several decades and measured by the sieve pipette method (unpublished data).

Table 1. Soil series, taxonomy to the family level, and pedogenic index (I) which defines the ordinal sequence of soil series topology.

	Position	Order	Series	Taxonomy
Toposequence	summit	1	Putnam	Fine, smectitic, mesic Vertic Albaqualfs
	summit	2	Adco	Fine, smectitic, mesic Vertic Albaqualfs
	shoulder	3	Mexico	Fine, smectitic, mesic Vertic Epiaqualfs
	upper backslope	4	Leonard	Fine, smectitic, mesic Vertic Epiaqualfs
	lower backslope	5	Armstrong	Fine, smectitic, mesic Aquertic Hapludalfs
Lithosequence	West	1	Marshall	Fine-silty, mixed, superactive, mesic Typic Hapludolls
	:	2	Sharpsburg	Fine, smectitic, mesic Typic Argiudolls
	:	3	Grundy	Fine, smectitic, mesic Aquertic Argiudolls
	:	4	Pershing	Fine, smectitic, mesic Vertic Epiaqualfs
	East	5	Adco	Fine, smectitic, mesic Vertic Albaqualfs

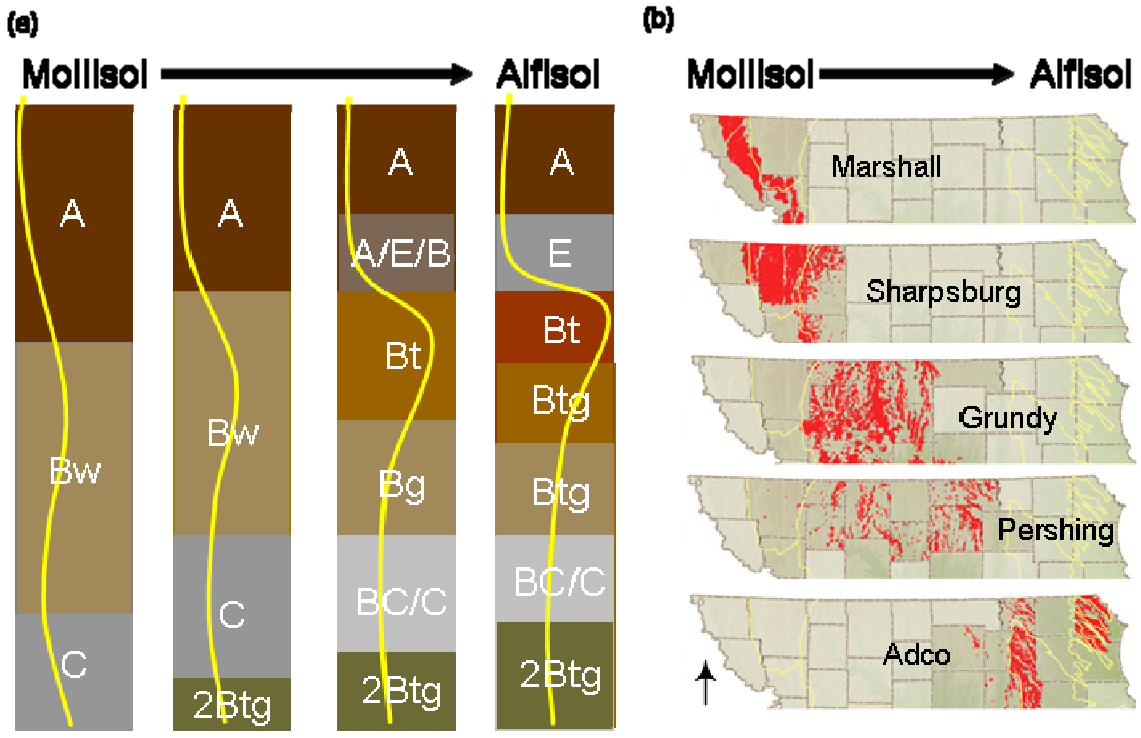

Figure 2. Lithosequence of loess thickness and soil profile development in summit soils from east to west in northern Missouri. (a) The common sequence of horizons moving from west to east (Mollisol to Alfisol). The yellow line is relative clay content from 10 to 65 (%). (b) Red areas show the distribution of the named soil series for the Northern tier of counties in Missouri. (Soil extent maps courtesy of www.soilsurvey.org)

The toposequence clay content data comes entirely from claypan soils in MLRA 113. The term 'claypan soil' is used here to describe the overall collection of soil series that occur in regular associations in these landscapes. More technically, claypan refers to an abrupt (within 0.5 < 2-cm) to very abrupt (within < 0.5-cm) textural change of greater than 100 % increase in clay content between horizons, most generally occurring in the first argillic horizon (Bt1/Btg1), coupled with a very slow permeability. Depth of the claypan ranges from 0 to 100 cm depending on landscape position and degree of erosion. Clay content at clay-max depth commonly ranges from 55 to 65 % and is composed of primarily smectite and hydroxy-interlayered smectite-illite.

Figure 3 shows two block diagrams indicating the order, parent materials, and landscape position of the claypan soil toposequence examined here. Soils in the toposequence were (from summit to lower backslope): Putnam, Adco, Mexico, Leonard, and Armstrong, all Alfisols. Pleistocene and Holocene erosion incised this landscape, reworked and redistributed loess and the surface of the glacial till (forming a pedisediment), and exposed the glacial till in some locations. These parent material processes give this toposequence properties of a lithosequence. The summit soils are formed mainly in loess while soils from within shoulder down to backslope positions have increasing influence of pedisediment and glacial till. As for the lithosequence above, the five soils in the toposequence are assigned an integer index (1-5) representing their position in the pedogenic gradient. The toposequence clay content data was also gathered for soil survey purposes and sampled by horizon. Soil texture analysis for both the lithosequence and toposequence datasets was performed using the sieve pipette method by the Missouri Cooperative Soil Survey Characterization Laboratory.

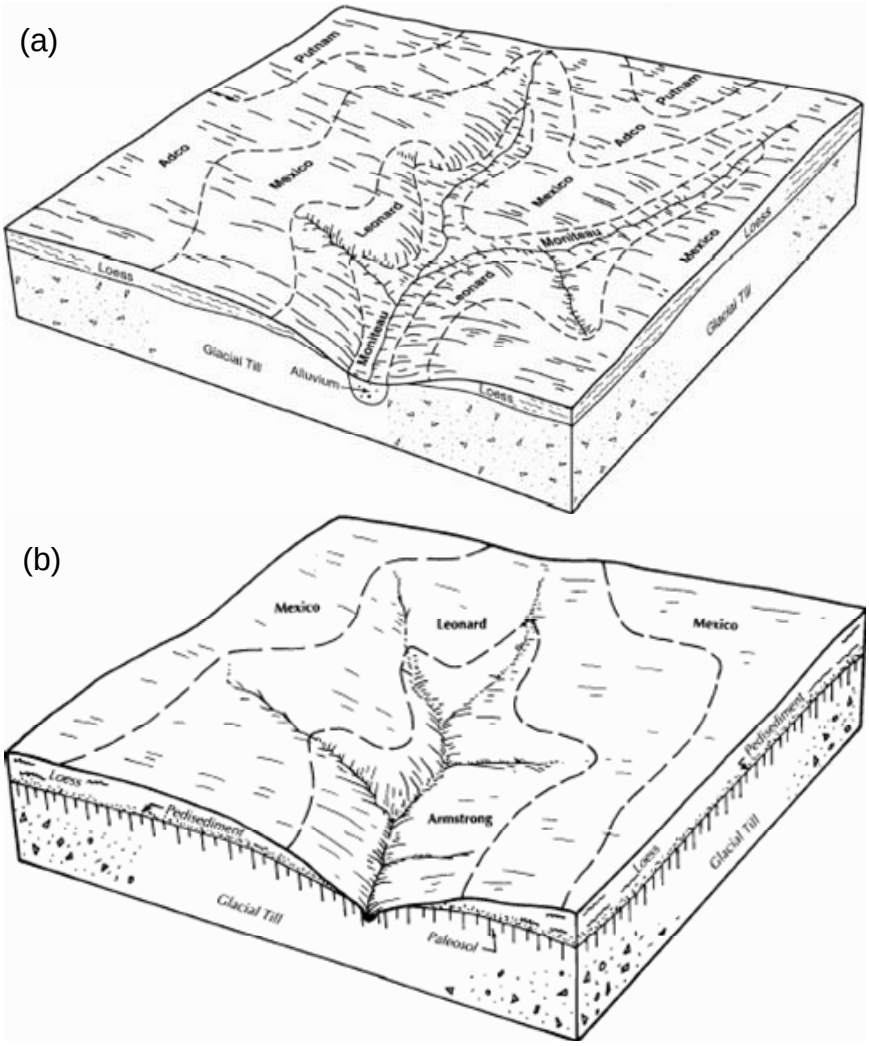

Figure 3. (a) Putnam, Adco, Mexico and Leonard soil association. (b) Mexico, Leonard, Armstrong soil association. (Young and Geller, 1995).

The field-scale dataset consists of several soil property measurements from prior research on five plots and fields located in the Missouri claypan soils of Audrain and Boone counties, within MLRA 113. Measurements used in this analysis included clay content by sieve pipette, organic carbon (OC) by combustion, Bray 1 P, Ca and K by emission spectroscopy, and pH in 1:2 soil:0.01 m NaCl solution. Pedons were sampled for these analyses between 1988 and 2001 either by horizon or by discrete intervals of varying thicknesses for research

and intensive soil survey purposes. Table 2 provides a listing of the sample support and references for more detailed description of methods used to develop each of the five field scale datasets.

Table 2. Field scale data sources and sampling support.

Study Name	Sample	Appx. Depth	Reference
Sanborn Field	horizon	1.2 m	(Miles and Hammer, 1989)
MSEA baseline characterization, Fields 1,2,3	horizon	1.2 m	(Kitchen et al., 1997; Sudduth et al., 2005)
Zones soil characterization Gvillo Field	horizon	1.2 m	(Sudduth et al., 2005)
USDA-CSWQ Field 1 Deep Fertility Plots	15 cm lifts	1.2 m	(Spautz, 1998)
Claypan Soil Quality	7.5, 15 cm lifts	30 cm	(Jung et al., 2005)

Coherent Depth Translation

Coherently translated depth, referred to subsequently as translated depth (D_t), of a sample from layer *i* was calculated by [1], where $D_{t,Ø}$ is the depth to profile clay-max for a soil profile and where *d* is the average depth (cm) of sampled layer *i*.

$$D_{t,i} = D_{t,Ø} - d_i \quad \textbf{[1]}$$

The results of this equation indicate the height that a specific soil layer *i* occurs above or below its profile's clay-max. Samples with positive D_t are above a clay-max, while samples with negative D_t are from below a clay-max. The D_t origin ($D_{t,Ø}$), and the clay-max for all of the translated profiles are at 0 cm D_t. Coherent depth translation could be applied in a soil-landscape with a different ubiquitous pedogenic feature by substitution for $D_{t,Ø}$.

Depth to Clay Maximum

Three methods are useful for measuring or estimating the depth to clay-max in the example landscape: direct field observation, prediction via bulk apparent electrical conductivity (EC_a), and anisotropic projection. Observation of depth to clay-max can be performed in the field on soil cores using manual techniques, it is simple and appropriate where subsequent use of CDT is known *a priori*. Depth to clay-max is obvious in claypan soils cores (± 1 cm) by examining the moist consistency of the soil, resistance to tool insertion, and gloss of the smeared core surface, and can be confirmed via the texture by feel method.

Bulk apparent soil electrical conductivity (EC_a) has been successfully used to predict depth to claypan in claypan soils (Doolittle et al., 1994), but was adopted here to predict depth to clay-max for most of the field scale datasets. A Geonics† EM-38 ($EC_{a\text{-}em}$) was used to map field sites on 10-m transects using an ATV and cart (Sudduth et al., 2005). Data was logged at 1 Hz and position data was logged by GPS. For the example field-scale sites, the following non-linear regression relationship was developed from field measured depth to clay-max and EC_a surveys (rmse = 11.61 cm), [2].

$$\mathbf{D_{t,\emptyset}=49054925.03 + EC_{a\text{-}em}^{-4.171}} \qquad \mathbf{[2]}$$

Estimation of clay-max depth by anisotropic projection is useful for archived soil characterization data where actual observations of depth to clay-

†Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Dept. of Agriculture, Univ. of Florida, or Univ. of Missouri.

max are not available. Anisotropic projection was applied to estimate $D_{t,\emptyset}$ for the lithosequence and toposequence data and for the Sanborn Field dataset. Anisotropic projection uses the physical arrangement of the two soil horizons above and below the clay maximum horizon, and the clay content of all three, to estimate $D_{t,\emptyset}$. The procedure is initialized by identifying the clay-max horizon and its two neighbors. A simple search algorithm locates these horizons within a specified depth window, excluding any secondary clay peaks in paleosol argillic horizons. Anisotropic projection parameterizes the anisotropy in the unknown depth function integrated by the three horizons. The method is represented geometrically by the intersection of lines projecting from the midpoint of the horizons above (c_1) and below (c_3) the clay-max horizon (fig. 4).

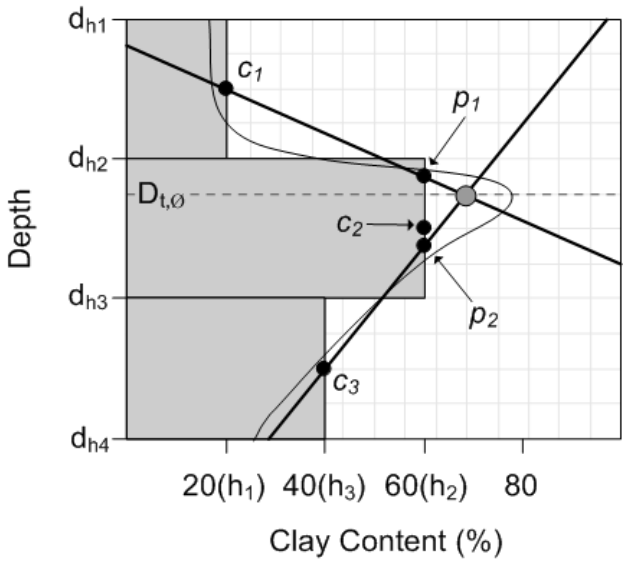

Figure 4. Anisotropic projection projection estimates the depth to profle clay maximum as the depth of intersection for two vectors. The two vectors project from the centroids of the horizons above (c_1) or below (c_3) the clay-max-horizon. The vectors are constrained by the points p_1 and p_2 whose positions are calculated based on the difference in clay content between the horizons (see equations 3,4, and 5). The method estimates the depth to clay-max, but does not estimate clay content at clay-max depth. The curved line represents the true continuous depth distribution of clay in the profile while the dashed line represents the depth to clay-max ($D_{t,\emptyset}$).

The lines are constrained by points p_1 and p_2 located at the leading edge of the clay-max horizon step function. The depths of these points are calculated with the aid of a skew parameter (η) [3],

$$\eta = \frac{h_1 - h_3}{\max((h_2 - h_1),(h_2 - h_3))} \quad \textbf{[3]}$$

where h_2, h_1, and h_3, are the clay content of the clay-max horizon, it's superjacent horizon, and it's subjacent horizon respectively. Anisotropy in the magnitudes of h_1, h_2, and h_3 control the sign and magnitude of η. Negative η shifts p_1 and p_2 upward relative to the midpoint of the clay max horizon while positive η shifts them downward. Projected $D_{t,\emptyset}$ is shifted in the direction of the larger h difference.

The depths of the constraining points p_1 and p_2 are calculated by the following equations [4, 5],

$$p_1 = \frac{\eta t_2}{4} + \left(\frac{t_2}{4} + d_{h2} \right) \quad \textbf{[4]}$$

$$p_2 = \frac{\eta t_2}{4} + \left(\frac{3t_2}{4} + d_{h2} \right) \quad \textbf{[5]}$$

where t_2 and d_{h2} are the thickness and depth of the clay-max horizon. The intersecting lines passing through the points $\{c_1, p_1\}$ and $\{c_2, p_2\}$ are given by the following linear functions [6,7]:

$$y_1 = m_1 x_1 + b_1 \quad \textbf{[6]}$$

$$y_2 = m_2 x_2 + b_2 \quad \textbf{[7]}$$

where m_1 and m_2, and b_1 and b_2 are the slopes and intercepts of the upper and lower lines respectively. These equations correspond to a linear system [6] who's solution includes $D_{t,\varnothing}$ the depth of their intersection.

$$D_{t,\varnothing} = \mathbf{m}x + \mathbf{b} \quad \textbf{[8]}$$

The purpose of anisotropic projection is not to estimate the clay concentration at clay-max (via *x* in equation 8) but only the depth of the clay-max. Clay concentration is estimated by profile depth functions fit to the depth transformed data. A strong covariate of clay-content such as a specific cation or cation exchange capacity (CEC) could be used for the values of *h* when clay content has not been measured.

Asymmetric Nonlinear Functions

A variety of peak functions and other nonlinear functions are potentially useful for modeling soil profile data. It is suggested that the reader consult function libraries to see comprehensive lists of candidate functions (Abramowitz, 1974; Systat, 2002). Many functions are asymmetric, exhibiting separate behavior on either side of a modal maxima or minima. This allows them to handle, for instance, a different shape for a soil property depth function above or below a clay-max. For this study the Pearson IV [9] and the asymmetric logistic

peak [6,7] functions were used to model depth distribution of clay. These and other peak functions can handle upright or inverted peaks and are able to model the distribution of clay content and its positive covariates such as Ca, Mg, K, cation exchange capacity (CEC), as well as negative covariates such as silt, pH, and P.

The Pearson IV [9] asymmetric probability distribution function can handle extensive tails with a minimal number of parameters (Pearson, 1895b; Heinrich, 2004). The parameters of the Pearson IV function are the intercept (*a*), amplitude (*b*), peak center (*c*), full width at half maximum (*d*), kurtosis (*e*), and skew (*f*). The last three parameters combine to give great flexibility in modeling a range of asymmetric peak shapes.

$$F(x) = a + \frac{b\left(1+\frac{\left(x-\frac{1}{2}\frac{df}{e}-c\right)^2}{d^2}\right)^{(-e)} \exp\left(-f\left(\arctan\left(\frac{x-\frac{1}{2}\frac{df}{e}-c}{d}\right)+\arctan\left(\frac{1}{2}\frac{f}{e}\right)\right)\right)}{\left(1+\frac{1}{4}\frac{f^2}{e^2}\right)^{(-e)}} \quad [9]$$

The parameters of the Pearson IV function describe physical structures of coherently translated depth profiles of clay content. The intercept, *a,* is the approximate surface clay content, *b* is the amplitude, *a*+*b* is the magnitude of the clay-max peak, *c* is the peak center, ($c=D_{t,\emptyset}=0$ by CDT), *d* controls the broadness of the peak, *e* controls the overall distinctness of the peak, and *f* controls the abruptness of the clay peak transition.

The asymmetric logistic function [10, 11] is another probability distribution developed to handle extreme value cases (Tawn, 1988). Again the *a* parameter is the intercept, *b* is the amplitude, *a*+*b* is the magnitude of the peak, *c*

is the peak center ($c=D_{t,\emptyset}=0$ by CDT), and the ratio of e to f controls the asymmetry of the peak.

$$\kappa(x) = \exp\left(-\frac{x + d\ln e - c}{d}\right) \quad [10]$$

$$y(x) = a + b\left(1+\kappa(x)\right)^{-e-1} e^{-e}(e+1)^{e+1}\kappa(x) \quad [11]$$

Hierarchical Soil-Landscape Models

A hierarchical model represents multiple processes that interact, but which occur on different scales. Two high-order processes can be used to describe the profile distribution of soil properties. They are regional or landscape scale processes, *L(·)*, and profile scale processes, *P(·)*. The hierarchical approach taken here was to first model a sequence of soil profiles by peak functions to represent *P(·)*., and then to model linear or quadratic functions of each of the peak function parameters to model *L(·)*. The profile scale process *P(·)* was modeled in i=1 to n pedogenically related soil series with n peak functions of soil property θ as a function f of depth d [12],

$$P(\theta, d)_{i=1}^{n} \begin{cases} \theta = f(d, \boldsymbol{\beta}_1) \\ \theta = f(d, \boldsymbol{\beta}_2) \\ \theta = f(d, \boldsymbol{\beta}_i) \\ \vdots \\ \theta = f(d, \boldsymbol{\beta}_n) \end{cases} \quad [12]$$

where $\boldsymbol{\beta}_i$ is a vector of m parameters $\beta_{i,j}$, j = 1 to m, fitted by nonlinear regression for each soil series i and $\mathbf{A}$ is a matrix with $\boldsymbol{\beta}_i$ as rows [13].

$$\mathbf{A} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \boldsymbol{\beta}_i \\ \vdots \\ \boldsymbol{\beta}_n \end{bmatrix} \quad [13]$$

The regional process $L(\cdot)$ is defined across a pedogenic gradient, but is represented by a sequence of soil series given the ordinal index I. Each of the $j = 1$ to m columns in the parameter matrix **A** are modeled as m components of $L(\cdot)$ by linear or quadratic functions g of **I**, [14].

$$L(\overset{m}{\underset{j=1}{\mathbf{A}}}, \mathbf{I}) \begin{cases} \mathbf{A}_{1\cdots n,1} = g(\mathbf{I}) \\ \mathbf{A}_{1\cdots n,2} = g(\mathbf{I}) \\ \mathbf{A}_{1\cdots n,j} = g(\mathbf{I}) \\ \vdots \\ \mathbf{A}_{1\cdots n,m} = g(\mathbf{I}) \end{cases} \quad [14]$$

Estimated parameters of the profile functions in P are generated by [14] and are used to calculate a continuous surface of θ as a function of d, **A,** and **I** [15].

$$\theta = f(d, L(\mathbf{A}, \mathbf{I})) \quad [15]$$

Analytical and Statistical Procedures

Different profile function fitting procedures were used for the field scale data and for the regional topo- and lithosequence data. The field scale dataset is used to demonstrate the properties of CDT and to illustrate some of its advantages. Thus, for simplicity, these profiles are modeled using locally weighted regression.

Surface soil truncation and locally weighted regressions were used to demonstrate the surface noise reduction effect of CDT. Depth translated P, K, Ca and buffer pH measurements from the field scale datasets were pooled and fitted with locally weighted regression to emphasize the profile distribution. Successive truncations of surface soil measurements were performed by removing samples from 0 to 15, 0 to 30, and then 0 to 45 cm. Local regression models were fit for each truncation to emphasize the differences in profile distribution at each level of truncation.

The regional scale dataset is used to demonstrate the peak functions and the hierarchical modeling approach. The Pearson IV and asymmetric logistic functions were fit using the Levenberg Marquardt algorithm on eigendecomposition denoised and depth translated clay profile data (Systat, 2002). Numerical soil-landscape models were constructed by modeling the peak function parameters as linear functions of the ordinal soil index.

RESULTS AND DISCUSSION

Change in Support: Discrete Horizons to Continuous Profiles

The technique of coherent depth translation is demonstrated here using the field scale dataset. The prominence of argillic horizons in these soils is visible in depth profiles of clay on the standard depth scale (fig. 5). The very sharp 100 to 200 % increase in clay content with depth gives these soils their ‘claypan’ designation. The piecewise step functions shown in figure 5a are the most accurate representation of this discretely sampled soil profile data. The discrete depth classes (horizons) is the sample support at which clay content was

measured. However, horizons are subjectively defined and coarsely integrate the continuity of soil properties.

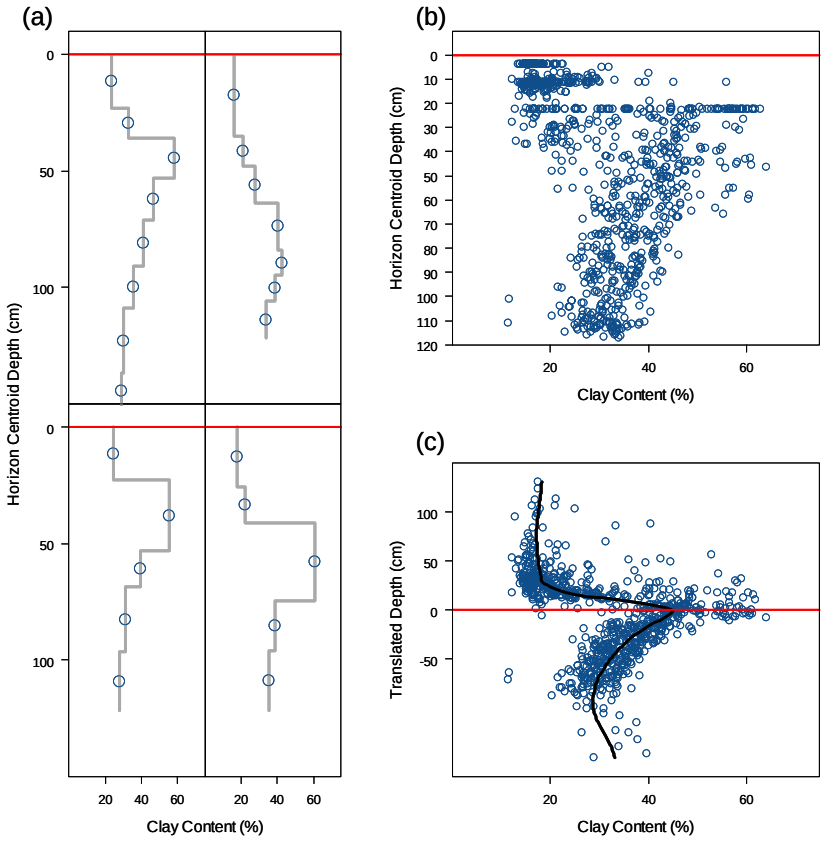

Figure 5. Coherent depth translation fundamentally alters soil profile data. A change of support from categorical to continuous occurs from panel a to panel c. (a) Soil profile data sampled by horizons is discrete and piecewise step functions most accurately represent horizon sampled data. (b) An ensemble of discretely sampled untranslated pedons displays the incoherence of the standard depth scale. (c) Translation of the depth scale coherently resolves the true profile relationship of clay content. A locally weighted regression function is plotted to emphasize this relationship.

Subjectivity in horizon designation can cause problems for use and interpretation of soil profile data. Two soil scientists may designate horizons with different levels of discrimination and bias. One may split soil horizons that are apparently uniform to another, or one may have a specific intent in making detailed horizon designations not shared by the other. The result is that some

samples are based on a smaller volume relative to others and ranges in clay content between adjacent sub-horizons may differ.

Compare the two upper and the two lower pedons in figure 5a. The upper pedons, though similar from a continuous perspective, were split into thinner sub-horizons. Furthermore, the sampling support (physical dimensions) of soil property data can be different based on the purpose for the measurement. For instance, the local scale dataset here are compiled from several different studies and have a mixed support of horizon and discrete depth intervals (see linear patterns at 3.75, 11.25 and 22.5 cm in fig. 5b). Because the supports of these composited data do not match comparison is problematic. The standard depth scale (i.e. below the surface) obscures the shape of the clay-max peak and is not sensible due to the spatial variability of surface modification and depth to clay-max present in the data. The coherently translated pedon ensemble provides a clearer view of clay distribution in these soils (fig. 5c). The abruptness of the clay-max peak is well defined and the discretely sampled data are deconvolved into the continuity of the natural profile.

Clay distribution dominates and reflects the soil genetic processes in MLRA 113. It also dominates and/or reflects the distribution of other soil properties. The extreme argillic horizon is a result of lessivage, and neoformation, during the Holocene temperate humid environment. Weathering of felsic minerals in the loess parent material provides reactants for aluminosilicate neoformation and a source for cation accumulation in argillic horizons (Norton, 1928; Bray, 1934). The abruptness of the argillic transition is due to ferrolysis and

eluviation acting on the upper boundary of the argillic horizons (Brinkman, 1970). These processes are enhanced by lateral flow of the water table that is perched on the claypan aquiclude.

The same pedogenic processes responsible for clay depth distribution impact many soil properties which, as a result, are covariate with clay distribution. This includes chemical properties such as pH and exchangeable cations (Bray, 1935), hydrologic properties including saturated hydraulic conductivity and plant available water holding capacity (Blanco-Canqui et al., 2002; Jiang et al., 2007), linear expansion (Jamison and Thompson, 1967; Baer and Anderson, 1997), and biological properties such as root distribution (Fraisse et al., 2001; Wang et al., 2003; Myers et al., 2007) or microbial activity (Jung et al., 2005).

Organic carbon (OC) distribution in claypan soils is also pedogenically covariate with clay content. For instance, argillic horizons of claypan soils have greater OC content due to greater concentrations of phylosillicates and sesqioxides (Culver and Gray, 1968). These mineral components are substrates and reactants in specific and non-specific OC adsorption processes including ligand exchange, cation bridging, van-der Walls forces, and complexation (Lutzow et al., 2006). Because of these mechanisms CDT with a clay-max origin can resolve landscape-scale depth relationships of OM (fig. 6c).

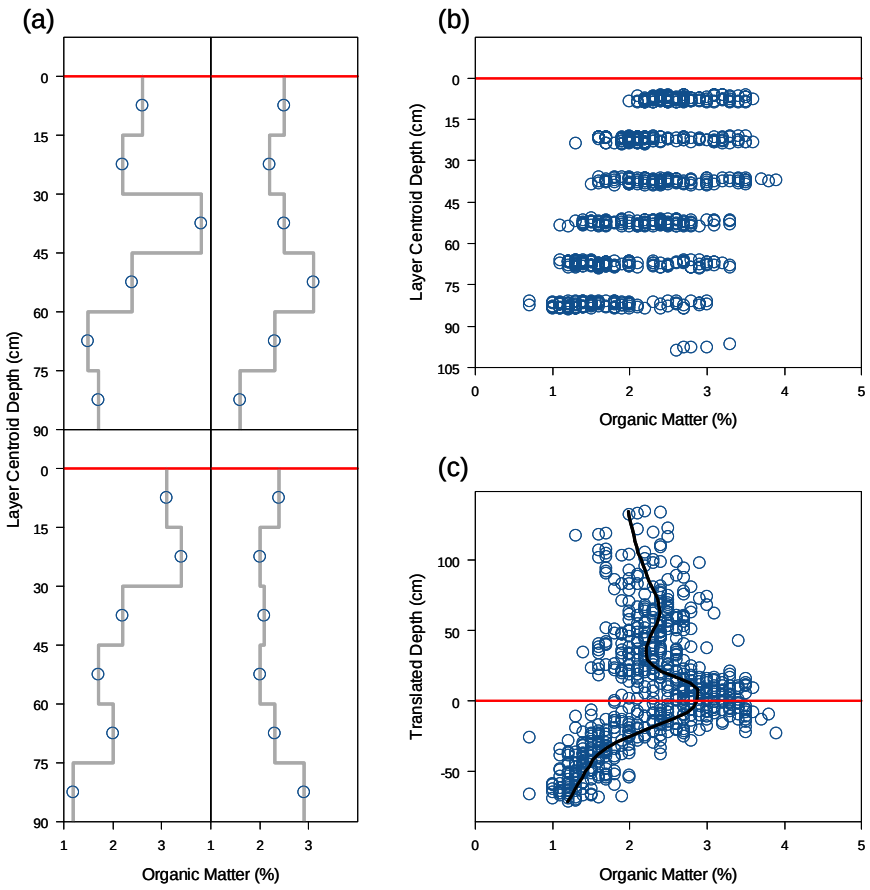

Figure 6. Coherent depth translation (CDT) of Organic Matter (OM). CDT was performed on an ensemble of pedons from from field scale datasets in claypan soils, sampled by discrete 15-cm increments. A change of support from categorical to continuous occurs from panel a to panel c. (a) Piecewise step functions most accurately represent the discrete interval support of the original data. (b) An ensemble of 108 untranslated pedons demonstrates the incoherence of the standard depth scale and shows the discrete nature of the data. (c) Translation of the depth scale coherently resolves the true profile relationship of OM into a continuous support. A locally weighted regression emphasizes the complex distribution of OM in the claypan landscape (black line).

The organic matter from the field scale datasets plotted in figure 6 was discretely sampled in 15-cm layers. This fact is emphasized in the uniform interval piecewise step functions (fig. 6a) and in the composited data (fig. 6b). Some random noise has been added to the data in panel b to more clearly show the character of the pooled untranslated data. Unless they are sufficiently small, discrete depth categories hide the complex detail of OM distribution, such as the

local OC minimum in the E horizon, and the local peak in OC in the claypan (around 0-cm D_t). Pooling an ensemble of coherently translated discrete data increases the chance that important features will be concentrated into single increments. This occurs because it is $D_{t,\emptyset}$ which best reflects the pedogenic control on the distribution of organic matter, not the sampled increment. The result is a deconvolution of the true form of the continuous depth profile of OM (fig. 6c).

Filtering Surface Profile Noise by Coherent Depth Translation

Coherent depth translation imparts the property of robustness against surface alterations to analyses of soil profile data, and can be used to reconstruct an approximation of the natural soil profile. Soil near the surface in most agricultural environments is extensively modified by tillage practices and by applications of manures, fertilizers, and liming materials. Claypan landscapes have been cultivated for more than one century (Jamison et al., 1968). The effects of management are imprinted on the physical and chemical properties of the upper portion of the soil profile. These land-use impacts add unnatural noise to pedogenic relationships. Coherent depth translation can expose this noise and the native properties of the near surface soil.

The panels in figure 7 demonstrate the variation in the profile distribution of P, K, Ca, and buffer pH, four soil properties that are commonly amended in agricultural soils by fertilizers and liming materials. These amendments solublize, react with the soil solution, and migrate in the soil profile at various rates controlled by well characterized reactions (Lindsay, 1979). For instance, P is

readily and strongly fixed in many soils as $Ca_x(PO_4)_{2x}$ and $Fe/AlPO_4$. Likewise, K^+ is readily and perhaps not as strongly fixed by exchange reactions with layer silicates. The by-products of $Ca(CO_3)_2$ and $Mg(CO_3)_2$ liming materials, including Ca^{2+} and Mg^{2+} cations and acid-neutralizing HCO_3^- and OH^- anions, are more mobile than P and K fertilizer reaction products.

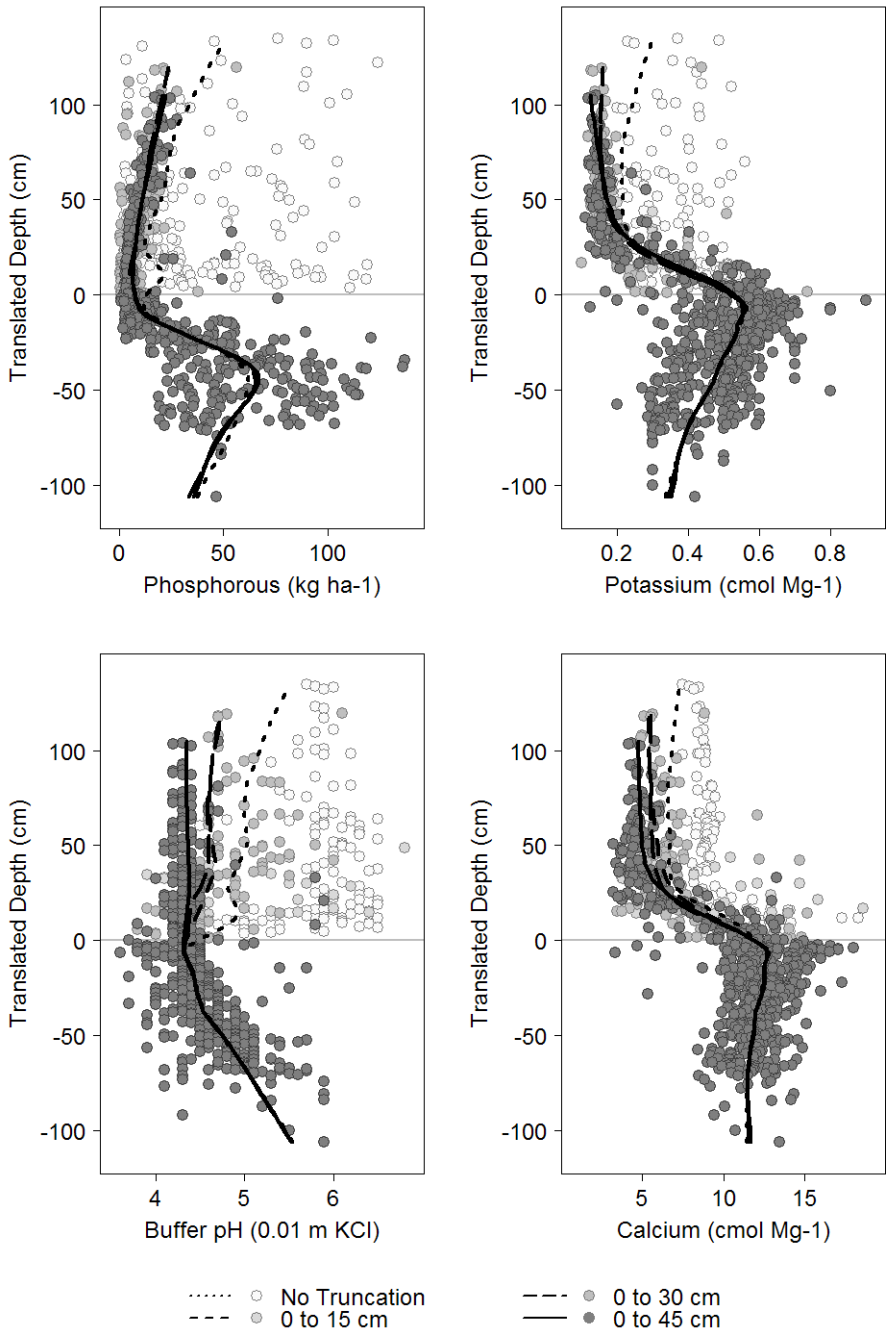

Figure 7. Penetration of historical fertilizer amendments. Locally weighted regression functions are fit to datasets with successive truncations of surface soils. Samples are truncated from the surface to incrementally larger depths by the following intervals: no truncation, 0 to 15 cm, 0 to 30 cm, and 0 to 45. Movement of the local regression line indicates the diminishing impact of surface modifications. The solid line represents the native distribution of unmodified soil properties.

Each panel of figure 7 plots the change in depth function for a soil chemical property by truncating an incremental depth of surface soils – 0 to 15, 0 to 30, and 0 to 45 cm. The incremental truncations are represented by the shading of the point and the regression line symbol. The cloud of lightest colored values in the upper right quadrant of each panel is measured from 0 to 15 cm, the plow layer. Soil property measurements in the plow layer are dramatically different than the native profile distribution approximated by the darkest points. The darkest points demonstrate the natural pedogenic distribution of soil properties in claypan soils without influence from land-use or land-cover. The difference between each light colored data point from the darker population is the amount of fertilizer or change in pH that has been imparted by amendments. Excluding first 15, then 30, then 45 cm of each soil ensemble then refitting the local regression illustrates the mobility of the amendments and their by-products in the profile.

Successive local regression depth functions of P and K beyond 15 cm are essentially unchanged. Difference occurs only by excluding the plow layer. Indeed, a century of amendments and manure containing P and K have not penetrated significantly beyond the plow layer. In contrast, the depth functions of pH and Ca concentration change with each increment of surface soil exclusion. These results suggest that liming can modify the pH and Ca concentration of claypan soils to a depth between to 30 to 45 cm, but the effects of liming materials also did not penetrate beyond the clay max.

This non-parametric analysis demonstrates three advantages of the CDT technique. First, surface modifications can be filtered from the native profile by an appropriate truncation. Second, the depth profile of soil properties can be clearly and continuously viewed. Third, subtleties of physical, chemical, and pedogenic processes can be elucidated. The next section develops the use of peak functions in order to quantitatively examine these last two points and extend them across the soil-landscape using depth translated data.

Hierarchical Modeling of a Lithosequence

The value of CDT is enhanced when it is combined with parametric functions fitted to a sequence of soils which are related by a soil genetic process. This approach models the fitted parameters of the peak function as a function of a pedogenic index. Figure 8 displays the asymmetric logistic function fit to the five soils in the west to east lithosequence of northern Missouri summit soils. The adjusted R^2 for these fits are 0.60, 0.38, 0.76, 0.93, and 0.87 for the Marshall, Sharpsburg, Grundy, Pershing, and Adco soils respectively. The parameters of the asymmetric logistic function were modeled with linear regression as a function of the toposequence index (1-5) (fig. 9). After modeling the landscape relationships in the parameters, the estimated parameters were used in the peak functions to calculate estimates of clay content for a continuous surface according to equation 9. The numerical soil-landscape model for clay content in the lithosequence is provided in figure 10.

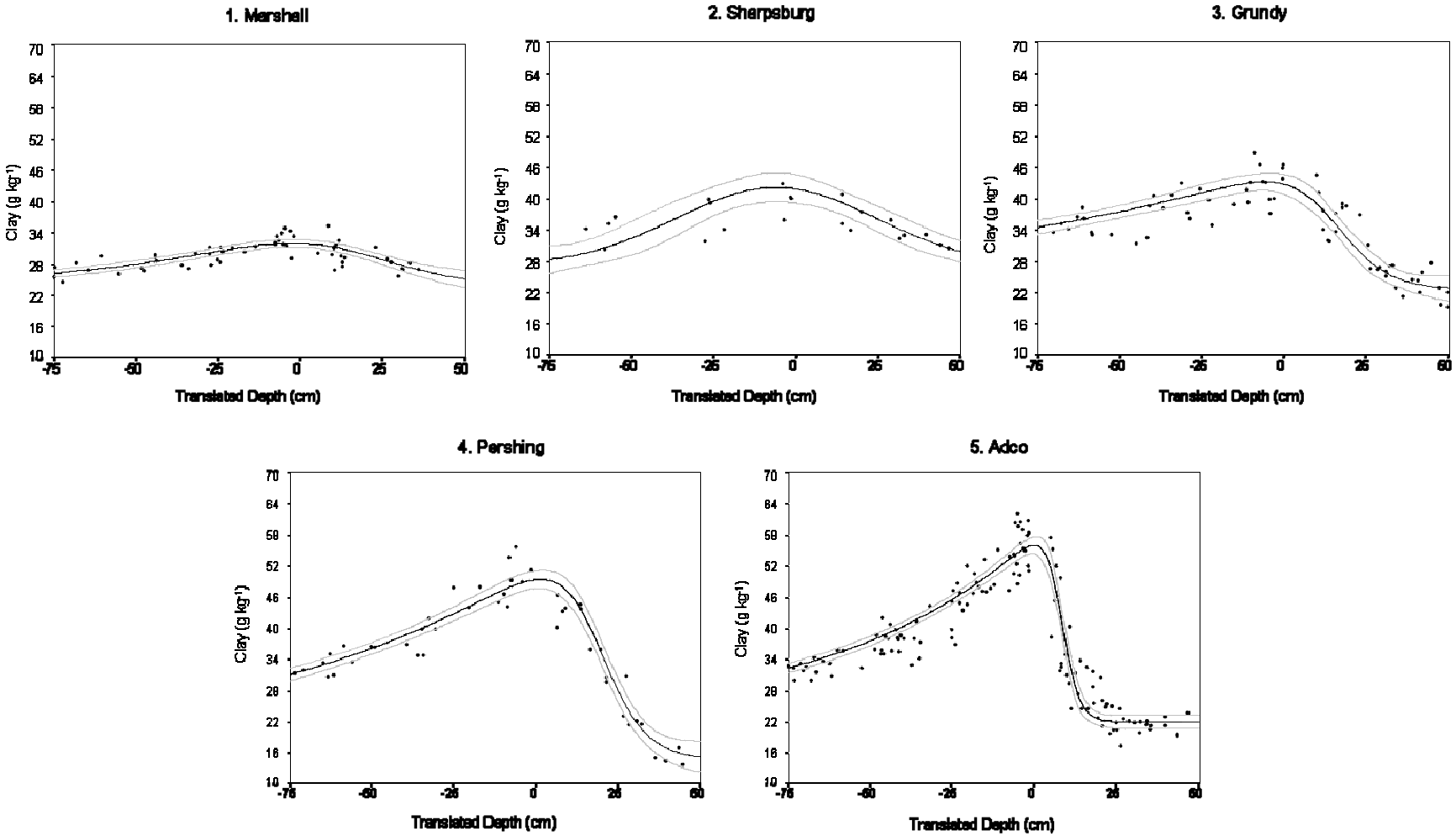

Figure 8. Asymmetric logistic function fit to a lithosequence of soils from northern Missouri.

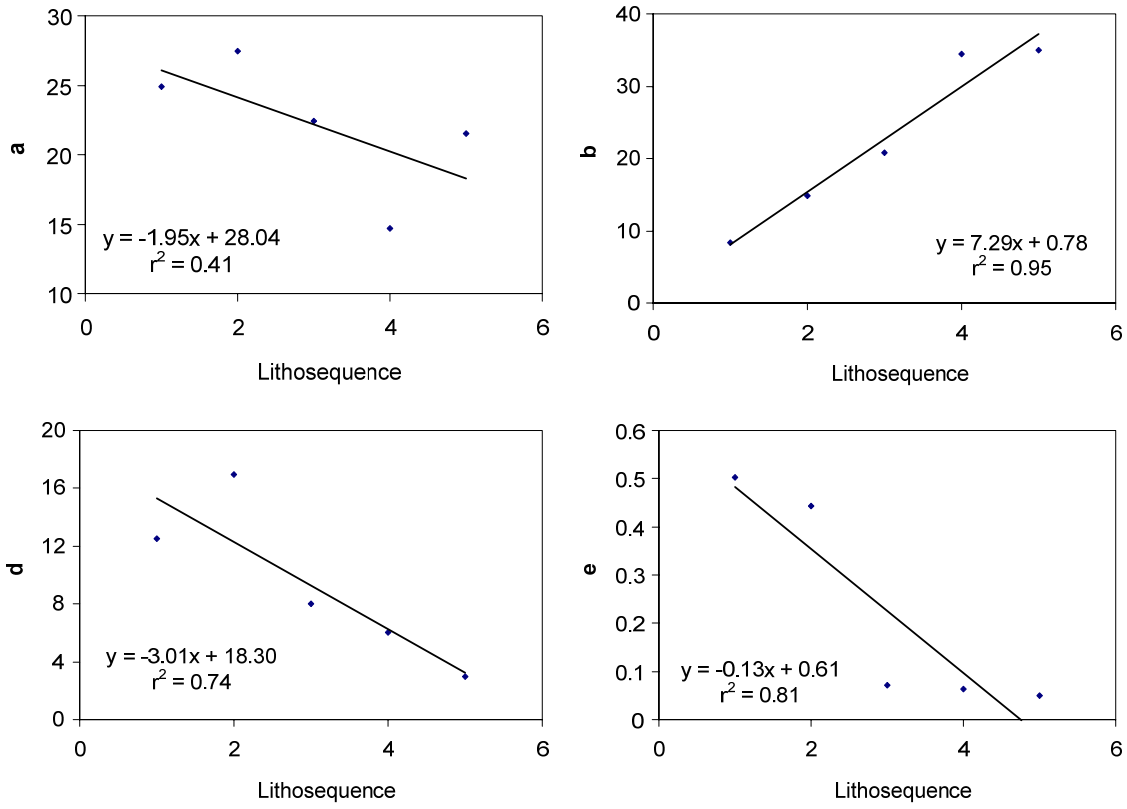

Figure 9. Parameters of the asymmetric logistic function as a function of a pedogenic index (1-5) for a lithosequence of summit soils in northern Missouri. Parameters correspond to the following peak function features: a – intercept, b – amplitude, c – zero by coherent depth translation, d – kurtosis, e – skew.

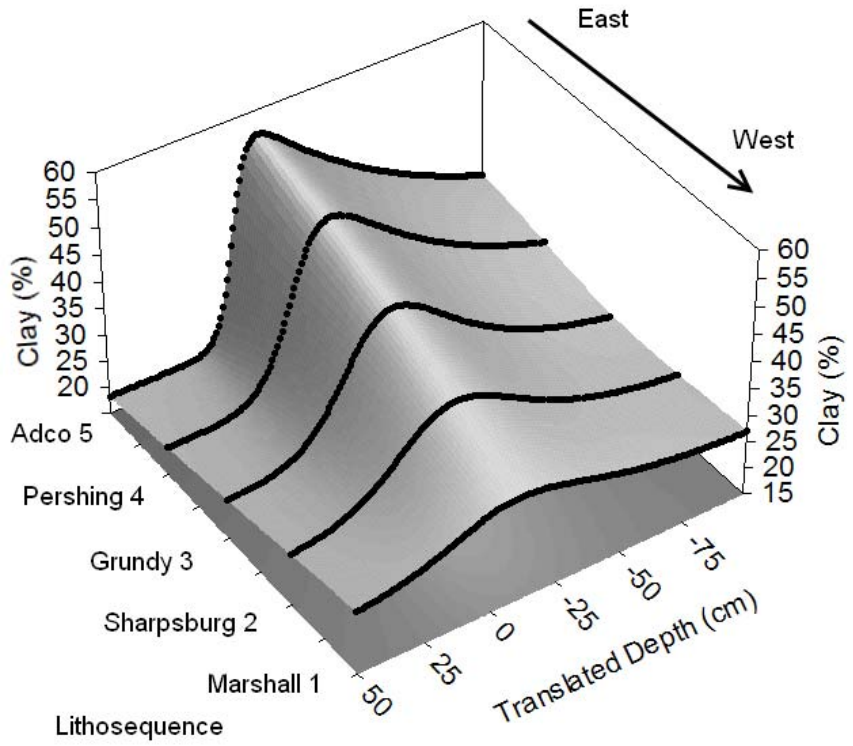

Figure 10. A hierarchical numerical soil-landscape model of a northern Missouri lithosequence.

Several patterns are evident in these figures. For instance, the peak of the clay distribution becomes progressively larger and more abrupt moving from the Mollisols in deep loess to the Alfisols on thinner loess (west to east). Overall clay content increases west to east as does clay content at the base of the translated depth profile. Conversely, clay content decreases in surface soils, west to east. These lithosequence trends have been attributed to increased weathering intensity (Smith, 1942; Goddard et al., 1973; Scrivner et al., 1973). Ferrolysis, eluviation, and lessivage dissolve and translocate clay minerals from the upper portion of the profile into the argillic zone. Annual rainfall and temperatures increase from west to east in the lithosequence as does the frequency and intensity of midsummer drought. The increased water percolation accelerates the

clay transformations and translocations while the increased temperatures and desiccation allow more frequent cycles for precipitation and neoformation of layer silicates. The coherency gained in applying CDT to a single soil series is augmented by the synergistic coherency gained from these pedogenic processes. This larger coherency is realized in the numerical model due to the pedogenic control and/or signature provided by the clay-max. These same processes are at work in the claypan toposequence.

Hierarchical Modeling of a Toposequence

The Pearson IV peak function was fit to depth translated clay content from the Putnam-Adco-Mexico-Leonard-Armstrong soil toposequence of MLRA 113 (summit to backslope) with an adjusted R^2 of 0.85, 0.92, 0.86, 0.63, and 0.69 respectively (fig. 11). The asymmetric morphology of the Pearson IV function handled the anisotropic profile data while fitting some soils better than others. The fit of the regression degraded sequentially for the downslope Mexico and Leonard map units as the influences of a paleosol and glacial till occur within -75 cm D_t. Figure 12 shows the parameters of the Pearson IV function modeled by linear regressions as functions of the toposequence index (1-5). As in the above lithosequence, regression estimates of the peak function parameters were used to calculate a continuous surface of clay content. Figure 13 shows the hierarchical numerical soil-landscape model.

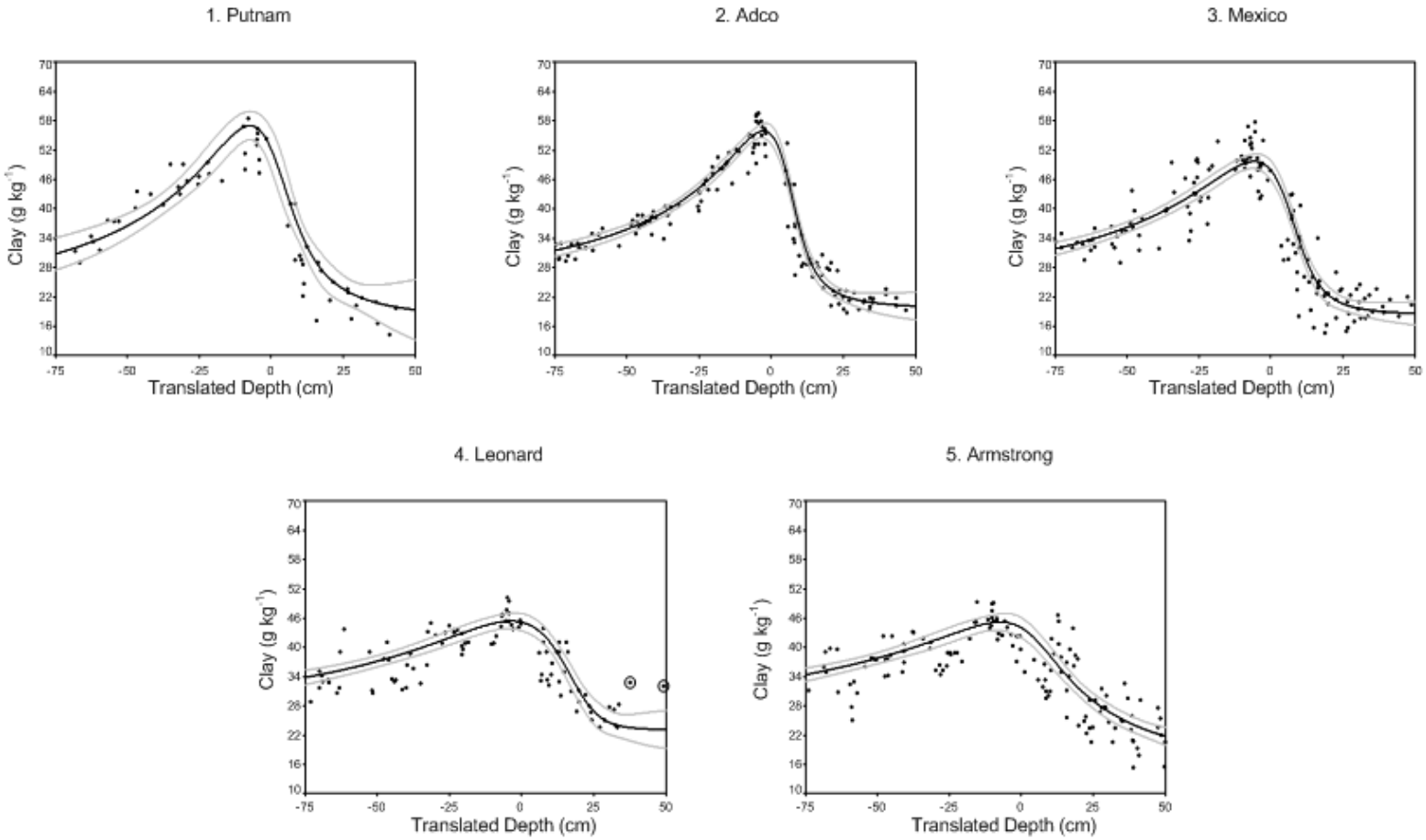

Figure 11. Pearson IV peak function fitted to a common claypan toposequence from northeast Missouri: Putnam-Adco-Mexico-Leonard-Armstrong. Putnam and Adco are on flat summits, Mexico on summits, shoulders, and upper backslopes, Leonard on upper to mid backslopes, and Armstrong on mid to lower backslopes. Circled data points are outliers.

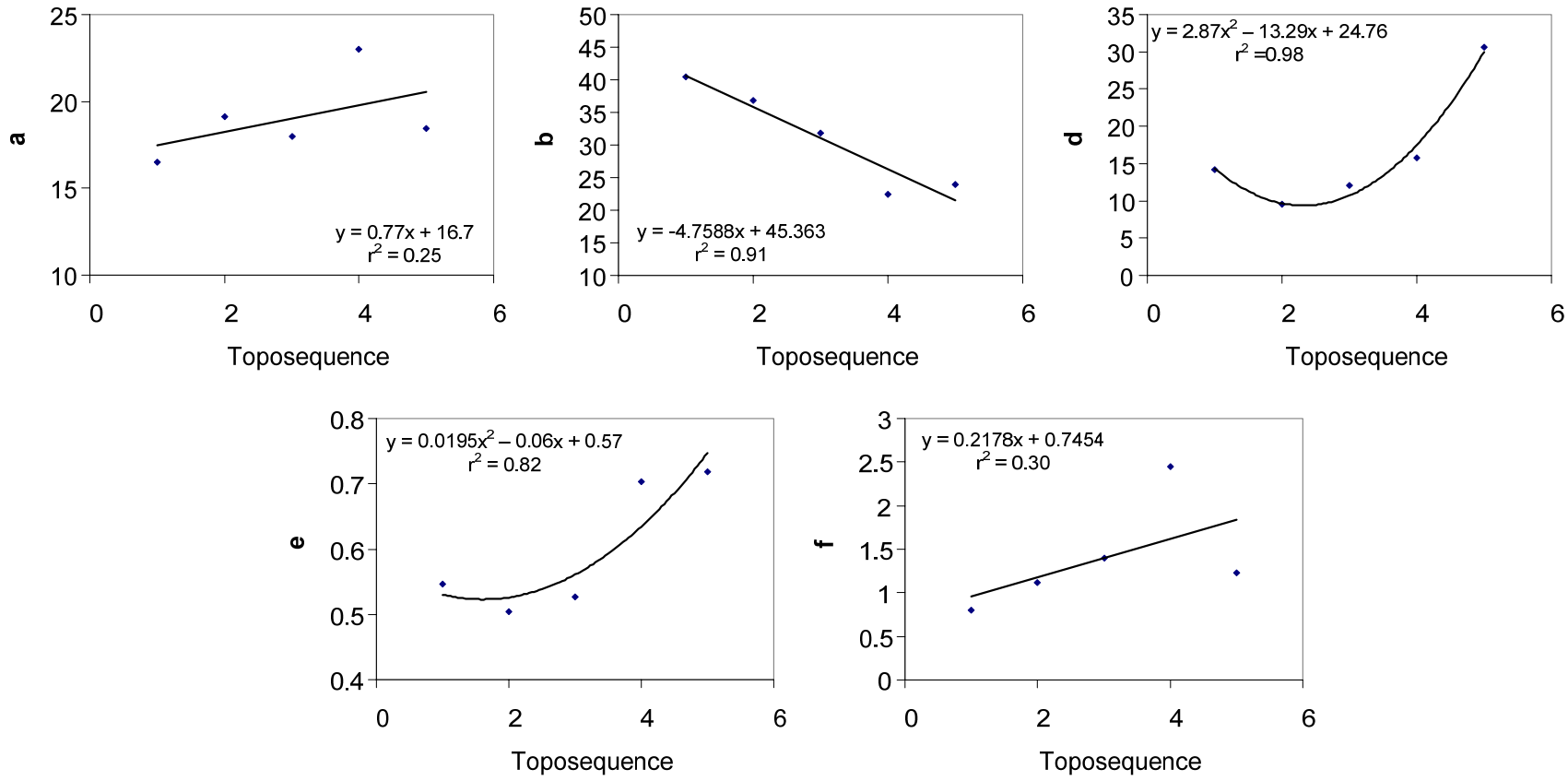

Figure 12. Pearson IV parameters as a function of ordinal toposequence index (1-5). Parameters correspond to the following peak function features: a – intercept, b – amplitude, c – zero by coherent depth translation, d – full width at half maximum, e – kurtosis, f - skew.

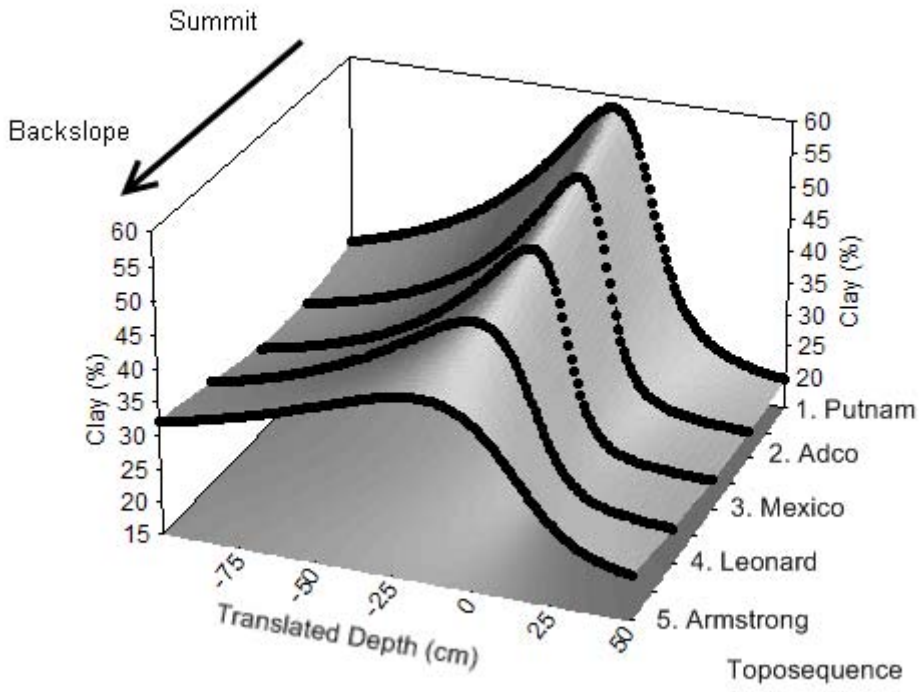

Figure 13. A hierarchical soil-landscape model of a northeast Missouri toposequence.

This toposequence encompasses several pedogenic gradients including slope, curvature, loess thickness, water movement, parent materials, and primary productivity. Quantitative estimates of clay content and useful observations about the genesis of claypan soils can be obtained from the results in figure 13. First, the clay content at $D_{t,\emptyset}$ decreases from 55 to 45 % in a continuous manner from summit to backslope (fig. 12). The $a + b$ parameters of the Pearson IV function quantify this relationship. Second, the abruptness of the argillic horizon transition decreases down the toposequence. This systematic decrease in argillic horizon abruptness implies a soil genetic process that may be quantified by the d, e, and f parameters of the hierarchical numerical soil model.

Abruptness tends to occur with the presence of an E horizon and a transitional E/Bt horizon boundary. The E horizon is thickest and its low chroma (2-3), low pH (~4.3), and high silt content (~70%) are at their extremes in Putnam and Adco soils on poorly drained and flat summit positions. The E horizon is a zone of stagnation and lateral flow due to water perched on the argillic aquiclude. Seasonal reduced conditions probably enhance ferrolysis and eluviation in this horizon. This leads to degradation of the upper limb of the argillic peak, increasing the abruptness of the argillic boundary. The E horizon shrinks to a thinner transitional layer in Mexico soils and is largely absent in Leonard soils. Greater runoff at shoulder and backslope landscape positions mitigates the E horizon formation process. Abruptness in the upper portion of the fitted peak is quantified by the e parameter which in turn quantifies the intensity of this soil genetic process (fig. 12)

Further Work on Hierarchical Modeling

The linear regression approach to linking translated soil series data provides a good starting point for further work on hierarchical modeling. Ideally, the landscape scale relationships in the parameters could be provided through pedogenic variables and functions such as loess thickness, rainfall, annual temperature, or metrics of seasonal dessication and recharge (Scrivner et al., 1973). This requires further work both in quantifying the pedogenic variables and functions and in locating a given soil series in the pedogenic gradient. Additionally, functions of the pedogenic gradients and nonlinear profile functions need to be modeled together. Fitting the parameter and profile models in a

hierarchical statistical framework could improve the stability of the nonlinear profile functions by borrowing strength from the landscape scale relationships. Importantly, this would allow the full propagation of errors, enabling the calculation of confidence and tolerance intervals for soil properties (Wilding et al., 1964). This could enable a quantitative and probabilistic definition of the 'range in characteristics' for a soil property as opposed to the expert knowledge that is currently relied upon.

Using CDT and hierarchical landscape modeling, the challenge of making continuous soil property maps can be reduced to mapping the genetic gradients and the translation origin. This is a much more tractable problem than the sampling of soil landscapes to the degree needed for direct geostatistical estimation. This approach could allow the use of more efficiently obtained spatially exhaustive information such as EC_a, digital elevation models, historical climate averages, and remote sensing products to build empirical or process based geospatial models. These models could be calibrated to spatially estimate the appropriate depth function parameters. This approach might allow a 3-D model of soil properties to be estimated from a relatively small number of sampled locations.

CONCLUSIONS

Traditional soil survey relies upon the tacit soil-landscape model of field soil scientists to populate a map of discrete categorical units. This approach has been a practical necessity given the immensity of area and spatial variation in the soil landscape. But, more continuous and quantitative approaches are needed as

the scale of problems in land-use and soil management exceeds the current discrete paradigm. Coherent depth translation allows the linking of sparse and discrete pedon data by the translation origin D_t. The technique allows confident functional fits that are robust against surface modifications due to land use. The linking mitigates data sparsity by putting each pedon ensemble in the continuous paradigm of the soil genetic sequence. The linkage of genetically related profiles is provided both by CDT, and by the choice of parametric functions. Asymmetric peak functions such as the Pearson IV were very successful at fitting the highly anisotropic properties of the Central Claypan Regions of Missouri. Hierarchical estimation of the parameters of the depth function resulted in a compact numerical soil-landscape model. Unlike the traditional discrete soil survey concept, the hierarchical model is continuous.

REFERENCES

Abramowitz, M., 1974. Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables. Dover Publications, Incorporated.

Baer, J.U. and Anderson, S.H., 1997. Landscape effects on desiccation cracking in an aqualf. Soil Science Society of America journal, 61(5): 1497.

Bishop, T.F.A., McBratney, A.B. and Laslett, G.M., 1999. Modelling soil attribute depth functions with equal-area quadratic smoothing splines. Geoderma, 91(1-2): 27.

Blanco-Canqui, H., Gantzer, C.J., Anderson, S.H., Alberts, E.E. and Ghidey, F., 2002. Saturated hydraulic conductivity and its impact on simulated runoff for claypan soils. Soil Science Society of America Journal, 66(5): 1596.

Bray, R.H., 1934. A chemical study of soil development in the Peorian loess region of Illinois. American Soil Survey Association Bulletin, 15: 58.

Bray, R.H., 1935. The origin of horizons in claypan soils. American Soil Survey Association Bulletin, 16: 70.

Brinkman, R., 1970. Ferrolysis, a hydromorphic soil forming process. Geoderma, 3(3): 199.

Brubaker, S.C., Jones, A.J., Frank, K., and Lewis, D.T., 1994. Regression models for estimating soil properties by landscape position. Soil Science Society of America Journal, 58: 1763.

Brubaker, S.C., Jones, A.J., Lewis, D.T. and Frank, K., 1993. Soil properties associated with landscape position. Soil Science Society of America Journal, 57(235-239).

Cleveland, W.S. and Devlin, S.J., 1988. Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. Journal of the American Statistical Association, 83(403): 596.
Colwell, J.D., 1970. A statistical-chemical characterisation of four great soil groups in southern New South Wales based on orthogonal polynomials. Australian journal of soil research, 8: 221.
Cook, F.J. and Kelliher, F.M., 2006. Determining vertical root and microbial biomass distributions from soil samples. Soil Sci Soc Am J, 70(3): 728.
Culver, J.R. and Gray, F., 1968. Morphology and Genesis of Some Grayish Claypan Soils in Oklahoma: I. Morphology, Chemical and Physical Measurements. Soil Science Society of America Journal, 32(6): 845.
Doolittle, J.A., Sudduth, K.A., Kitchen, N.R. and Indorante, S.J., 1994. Estimating depths to claypans using electromagnetic induction methods. Journal of Soil and Water Conservation, 49(6): 572.
Dwyer, L.M., Stewart, D.W. and Balchin, D., 1988. Root characteristics of corn, soybean and barley as a function of available water and soil physical characteristics. Canadian Journal of Soil Science, 68: 121.
Fehrenbacher, J.B., 1973. Loess stratigraphy, distribution and time of deposition in Illinois. Soil Science, 115(3): 176.
Fraisse, C.W., Sudduth, K.A. and Kitchen, N.R., 2001. Calibration of the CERES-Maize model for simulating site-specific crop development and yield on claypan soils. Applied Engineering in Agriculture, 17(4): 547.
Goddard, T.M., Runge, E.C.A. and Ray, B.W., 1973. The relationship between rainfall frequency and amount to the formation and profile distribution of clay particles. Soil Science Society of America Journal, 37: 299.
Hastie, T. and Tibshirani, R., 1990. Generalized Additive Models. Chapman & Hall/CRC.
Heinrich, J., 2004. A guide to the Pearson type IV distribution. 6820, Fermilab, Batavia, IL.
Jamison, V.C., Smith, D.D. and Thornton, J.F., 1968. Soil and Water Research on Claypan Soil, Tech. Bull. No. 1379. USDA-ARS, Washington, D. C.
Jamison, V.C. and Thompson, G.A., 1967. Layer Thickness Changes in a Clay-Rich Soil in Relation to Soil Water Content Changes. Soil Science Society of America Journal, 31(4): 441.
Jenny, H., 1930. A study on the influence of climate upon the nitrogen and organic matter content of the soil. Missouri Agriculture Experimentation Research Bulletin, 152: 1.
Jenny, H., 1935. The clay content of the soil as related to climatic factors, particularly temperature. Soil Science, 40: 111.
Jenny, H., 1941. Factors of soil formation. A system of quantitative pedology. McGraw-Hill, New York.
Jenny, H. and Leonard, C.D., 1934. Functional relationships between soil properties and rainfall. Soil Science, 38: 363.
Jiang, P., Anderson, S.H., Kitchen, N.R., Sudduth, K.A. and Sadler, E.J., 2007. Estimating Plant-Available Water Capacity for Claypan Landscapes Using Apparent Electrical Conductivity. Soil Science Society of America Journal, 71(6): 1902.
Jung, W.K., Kitchen, N.R., Sudduth, K.A., Kremer, R.J. and Motavalli, P.P., 2005. Relationship of apparent soil electrical conductivity to claypan soil properties. Soil Science Society of America Journal, 69(3): 883.

Kitchen, N.R., Lerch, R.N., Blanchard, P.E. and Hughes, D.E., 1997. Impact of historical and current farming systems on groundwater nitrate in northern Missouri. Journal of Soil and Water Conservation, 52(4): 272.
Lewis, D.T. and Lepele, M.J., 1982. Quantification of Soil Loss and Sediment Produced From Eroded Land. Soil Science Society of America Journal, 46(2): 312.
Lindsay, W.L., 1979. Chemical Equilibrium in Soils. John Wiley & Sons.
Lutzow, M.v. et al., 2006. Stabilization of organic matter in temperate soils: mechanisms and their relevance under different soil conditions - a review. European Journal of Soil Science, 57(4): 426.
Miles, R.J. and Hammer, R.D., 1989. One hundred years of Sanborn Field: soil baseline data, The Sanborn Field Centennial. Missouri Agriculture Experiment Station, Columbia, MO, pp. 100.
Muhs, D.R. and Bettis, E.A., 2000. Geochemical variations in Peoria loess of Western Iowa indicate paleowinds of midcontinental North America during last glaciation. Quaternary Research, 53(1): 49.
Myers, D.B., Kitchen, N.R., Sudduth, K.A., Sharp, R.E. and Miles, R.J., 2007. Soybean root distribution related to claypan soil properties and apparent soil electrical conductivity. Crop Science, 47(4): 1498.
Norton, E.A., 1928. Succession of soil profiles along the Mississippi River in Illinois. Proceedings of the American Soil Survey Association Bulletin, 12: 158.
Olson, K.R., Phillips, S.R. and Kitur, B.K., 1994. Identification of eroded phases of an Alfisol. Soil Science, 157(2): 108.
Pearson, K., 1895a. Contributions to the mathematical theory of evolution. II. skew variation in homogeneous material. Philosophical Transactions of the Royal Society of London. A, 186: 343.
Pearson, K., 1895b. Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. Philosophical Transactions of the Royal Society of London, 186(A): 343.
Ponce-Hernandez, R., Marriott, F.H.C. and Beckett, P.H.T., 1986. An improved method for reconstructing a soil profile from analyses of a small number of samples. Journal of SoilScience, 37: 455467.
Richardson, J.L. and Edmonds, W.J., 1987. Linear regression estimations of Jenny's relative effectiveness of state factors equation. 144(3): 203.
Richardson, J.L. and Riecken, F.F., 1977. Differences in exchangeable aluminum and soil acidity in loess soils of Iowa. Soil Science Society of America Journal, 41: 588.
Ruhe, R.V., 1969. Quaternary Landscapes in Iowa. University of Iowa Press, Ames, IA.
Runge, E.C.A., 1973. Soil development sequence and energy models. Soil Science, 115(3): 193.
Scrivner, C.L., Baker, J.C. and Brees, D.R., 1973. Combined daily climatic data and dilute solution chemistry in studies of soil profile formation. Soil Science, 115(1973): 213.
Smeck, N.E., 1973. Phosphorous: An indicator of pedogenetic weathering processes. Soil Science, 115(3): 199.
Smith, G.D., 1942. Illinois loess-variations in its properties and distribution. Illinois Agriculture Experimentation Bulletin 490, University of Illinois, Urbana-Champaign.
Spautz, R.E., 1998. Topsoil thickness influence on phosphorus and potassium availability and crop response. MS Thesis, University of Missouri, Columbia, MO, 127 pp.
Sudduth, K.A. et al., 2005. Relating apparent electrical conductivity to soil properties across the north-central USA. Computers and Electronics in Agriculture, 46(1-3): 263.

Systat, 2002. TableCurve 2D. Systat Software Inc., Port Richmond, CA.
Tawn, J.A., 1988. Bivariate extreme value theory: models and estimation. Biometrika, 75(3): 397.
USDA-NRCS, 2006. Land Resource Regions and Major Land Resource Areas of the United States, Agriculture Handbook #269. USDA-Natural Resource Conservation Service, Washington D.C.
Walker, P.H. and Green, P., 1976. Soil trends in two valley fill sequences. Aust. J. Soil Res., 14(3): 291.
Wang, F., Fraisse, C.W., Kitchen, N.R. and Sudduth, K.A., 2003. Site-specific evaluation of the CROPGRO-soybean model on Missouri claypan soils. Agricultural Systems, 76(3): 985.
Webster, R., 1978. Mathematical treatment of soil information, Transactions of the 11th International Congress of Soil Science, pp. 161.
Wenner, K.A., Holowaychuk, N. and Schafer, G.M., 1961. Changes in the Clay Content, Calcium Carbonate Equivalent, and Calcium/Magnesium Ratio With Depth in Parent Materials of Soils Derived from Calcareous Till of Wisconsin Age. Soil Science Society of America Journal, 25(4): 312.
Whiteside, E.P., 1944. Clay formation and movement in two claypan soils, the Putnam and Cowden. Soil Science Society of America Proceedings, 9: 211.
Wilding, L.P., Schafer, G.M. and Jones, R.B., 1964. Morley and Blount Soils: A Statistical Summary of Certain Physical and Chemical Properties of Some Selected Profiles from Ohio. Soil Science Society of America Journal, 28(5): 674.
Young, F.J. and Geller, A.W., 1995. Soil Survey of Audrain County. National Cooperative Soil Survey, Columbia, MO, 1 pp.

CHAPTER 2

Combining Proximal and Penetrating Soil Electrical Conductivity Sensors for High-Resolution Digital Soil Mapping

ABSTRACT

Proximal ground conductivity sensors produce high spatial resolution maps that integrate the bulk electrical conductivity (EC_a) of the soil profile. Variability in conductivity maps must either be inverted to profile conductivity, or be directly calibrated to profile properties for meaningful interpretation. Penetrating apparent electrical conductivity (EC_p) sensors produce high depth resolution data at relatively fewer spatial locations. The objectives of this research were to (i) investigate the profile source of EC_a in claypan soils via a detailed examination of EC_p profiles, (ii) examine the potential for feature detection with EC_p in claypan soils, and (iii) determine if EC_a sensors can be calibrated to EC_p features. Two study areas were chosen representing the claypan soils of northeast Missouri, USA. Profile conductivity was measured at high depth resolution on soil cores using a miniaturized Wenner conductivity probe and in the field using a conductivity penetrometer. Proximal ground conductivity was mapped with one direct contact sensor and two non-contact sensors providing 5 distinct coil/electrode geometries. Increasing EC_p was observed below the claypan correlating with decreasing clay and water content and increasing bulk density. Depth to the claypan was successfully calibrated to derivative peaks on EC_p profiles (r^2 0.72, $p<0.001$). Relationships between EC_a and EC_p features were poor (r^2 ~0.21) to good (r^2~0.87) on a field specific basis. Results show that EC_p can be used for calibration of EC_a to the depth to claypan.

INTRODUCTION

Proximal bulk apparent electrical conductivity (EC_a) sensors can be used to produce high spatial resolution maps that integrate soil profile EC_a variation by a depth response function. The conductivity data must either be inverted to approximate profile conductivity, or directly calibrated to profile properties for meaningful interpretation. Penetrating apparent electrical conductivity (EC_p) sensors measure EC from a smaller soil volume localized to their sensing electrode. Penetrating sensors measure at high depth resolution, but at sparse locations compared to proximal EC_p sensors. These two types of EC sensors have synergistic potential. We examined two avenues for their combined use with a case study in the claypan landscapes of northeast Missouri, USA. First we examined the potential for EC_p to identify soil morphological features. Next we examined the calibration of EC_p features to the spatially dense EC_a data. We focus on resolving the profile source of conductivity integrated by proximal EC_a sensors.

Three specific pathways of electrical conductance occur in soils: free water in large soil pores, hygroscopic or tightly interacting particle-water interfaces, and direct soil particle contact (Corwin and Lesch, 2005). As outlined by Corwin and Lesch (working in western USA soils formed in semi-arid to arid environments) the magnitude of EC_a is dependent mainly on soil salinity, Na^+ saturation percentage, water content, and bulk density (BD). The claypan soils of Missouri exist in a humid temperate environment. They are leached of salts and carbonates and have small concentrations of exchangeable Na^+ (< 2 cmol kg^{-1}).

These variables are unlikely to affect EC_a. The experiments described in this work allowed the examination of the remaining factors important for influencing proximal EC_a variation in claypan soils.

Previous studies in claypan soils discovered the relationship between EC_a and depth to claypan (DTC) (Dolittle et. al., 1994, Sudduth and Kitchen, 2006). These investigations speculated that depth to argillic horizon layer silicates was the primary cause of EC_a variation. Several properties of the smectitic clay mineralogy were considered to be important. Smectite and similar clay minerals might provide greater physical contact due to their size and platy structure, substantial interlayer water is usually present, and exchangeable cations are at very large concentration. Clay content decreases below the claypan and therefore if clay mineralogy were largely responsible for EC_a variation, then less conductivity response would be expected from there. However, greater below-claypan EC_p was detected during some of our early investigations with EC_p data (Sudduth et al., 2000). Confirmation of these observations on isolated samples is needed to understand the proximal EC_a response.

From our experiences with EC_a and EC_p data we suspected that profile conductivity features could be identified by penetrometer more objectively, at better depth resolution, and more quantitatively than by coring or auguring. Mapping subsoil EC_p features via their relationship to EC_a would be more efficient than grid survey. An EC_a-to-EC_p feature calibration should provide the spatial and depth resolution needed for high resolution soil mapping. We hypothesized that a large gradient in the first derivative of the EC_p profile could

be used to identify a claypan or a lithologic discontinuity. We examined the relationship between EC_p derivative peaks and observed depth to claypan in order to test this possibility. Further, we hypothesized that EC_a could predict the depth to EC_p first derivative peaks.

The specific objectives of this research were to:

i.) confirm increasing sub-claypan EC_p.

ii.) determine if EC_p data can be used to estimate depth to claypan.

iii.) determine if EC_a sensors can be calibrated to EC_p features.

MATERIALS AND METHODS

Soil Landscapes, Measurements, and Observations

Four agricultural fields in the claypan region of northeast Missouri were chosen for this study, three fields with a loess solum near Centralia, MO (Fields A, B, and C) (39° 13' 43" N., 92° 8' 20" W.) and a field with a loess-till solum, near Novelty, MO (Field D) (40° 1' 46.5" N., 92° 11' 19" W.). Fields A, B, and C are located near the southern limit of the claypan region while field D is at the northern limit, about 90 km away. Physical and chemical characterization data by horizon was available from 44 pedons with claypan features. Field descriptions and horizon designations for these pedons were made by experienced soil morphologists. Observed depth to claypan was determined as depth to the Bt1 or Bt2 horizon based on the field descriptions and lab data.

EC_p Measurement

Penetrometer EC_p and cone index (CI) were measured at the 44 claypan locations using a Veris† Profiler 3000 with an insulated shaft (Veris Technologies, Salina, KS, USA). Measurements of EC_p and CI were made on all fields in the late spring of 2007, and occurred within a few days of EC_a measurements on fields B and D. However, EC_p and CI were measured on fields A and C approximately 18 months after the EC_a surveys. Gravimetric soil moisture and BD determinations were made in 15-cm layers at the time of EC_p measurement.

Cone index and EC_p were measured to 92 cm, with 5 penetrations per site. Replicate EC_p soundings were pooled and fitted with locally weighted regression models. A Savitsky-Golay procedure was used to calculate the derivative of the fitted EC_p profiles. The upper claypan boundary was determined from derivative plots as the peak gradient in EC_p. This feature is referred to as the EC_p transition peak. Clay maximum depth translation was applied to each EC_p profile independently in order to explore the landscape relationship in sub-claypan EC. Translated depth (D_t) indicates the depth a measurement occurs either above or below the claypan. Translated depth profiles were pooled into a single dataset and again fitted with a locally weighted regression.

†Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Dept. of Agriculture, Univ. of Missouri, or Univ. of Florida.

EC_m Measurement

We developed a miniaturized Wenner array on a hand held probe (mini-probe) to measure EC (EC_m) on ex-situ soil cores to confirm EC_p observations. Wenner mini-probe apparent electrical conductivity was measured every 1.25 cm. The mini-probe had 5 mm electrode spacing, 5 mm insertion depth, and was operated using the electronics from a Veris EC_a sensor. Veris supplied custom software accounting for the probe geometry. Calibration standards were used to verify EC_m across the range of measurement and were found to agree within 10%. Measurements of EC_m were made on soil cores compressed into a linear steel die which formed the cores into equilateral triangular prisms. Die-pressing repaired breaks and extrusion cracks in soil cores, consolidated loose soil, and formed two uniform flat surfaces providing consistent contact of the mini-probe. Gravimetric soil moisture of these cores was measured on 2.54-cm intervals.

EC_a Measurement

Three conductivity sensors were used to measure EC_a with DGPS position on 10-m transects at 4 to 6-m intervals. Sensors were the DUALEM-2S electromagnetic induction (EMI) sensor (Dualem, Inc., Milton, Ontario, Canada) in horizontal coplanar mode ($EC_{a\ d\text{-}sh}$) and perpendicular plane mode ($EC_{a\ d\text{-}dp}$) (2-m coil spacing), the Geonics EM-38 EMI sensor (Geonics Limited, Mississauga, Ontario, Canada) in horizontal coplanar mode ($EC_{a\ em}$), 1-m coil spacing, and the Veris 3150 rolling coulter Wenner array (Veris Technologies, Salina, KS, USA) with 0.7 ($EC_{a\ v\text{-}sh}$) and 2.2-m ($EC_{a\ v\text{-}dp}$) electrode spacing. This combination of sensors provided 5 distinct coil/electrode geometries for EC_a

measurement. Fields B and D had all EC_a surveys made within a relatively narrow window of one month in the spring of 2007. Surveys of fields A and C were made within three days in the fall of 2005.

The five EC_a instrument geometries used for this study were unique, but their depth response functions were overlapping to some degree and their measurements are correlated (Sudduth and Kitchen, 2006). Partial least squares regression (PLSR) was used to model the EC_a relationship to EC_p in order to mitigate correlation in the predictors and to capitalize on any orthogonality in their response to EC_p.

RESULTS AND DISCUSSION

Soil Profile EC

The major morphological features found throughout the study fields were visible in the depth profile of EC_p or EC_m in a representative claypan site from field A (fig. 1a). First, the silty, granular structured, low-density surface had very small EC_p. The remaining A horizons had greater EC_p, but still relatively lesser EC_p compared to the claypan and lower solum. When an E horizon was present it appeared as a zone of minimum conductivity. Conductivity abruptly increased in the transition to the Bt1 horizon, the claypan feature. Conductivity continued to increase below the claypan to 90 cm and beyond, even as clay content decreased.

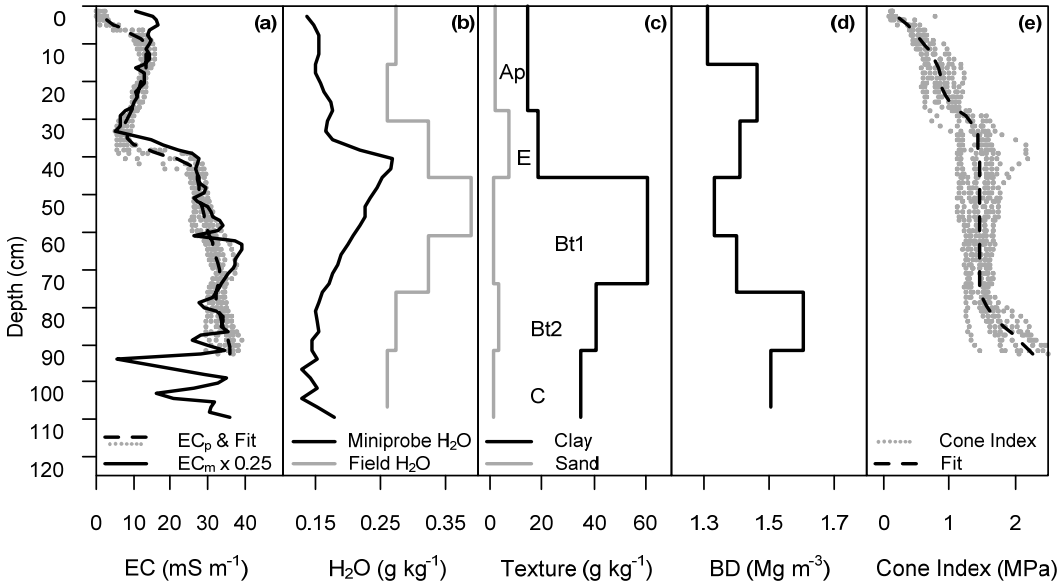

Figure 1. Data from a representative claypan soil on field A. Panels show: (a) penetrometer electrical conductivity (EC_p) and scaled Wenner mini-probe conductivity (EC_m*0.25), (b) field (taken with EC_p) and high resolution (taken with EC_m) gravimetric soil moisture, (c) percent clay and sand, (d) bulk density (BD), and (e) cone index.

Mean EC_p above and below the claypan for all 44 study locations was 20.9 and 47.4 mS m^{-1} with standard errors 0.31 and 0.26 mS m^{-1} respectively. These means were significantly different ($p<0.001$). This difference and the landscape trend in EC_p distribution were emphasized in pooled D_t profiles of EC_p (fig. 2). The depth translation procedure allowed comparison of pedons on a coherent depth scale. These results verified large and increasing sub-claypan EC_p and emphasized the similarity of these soils to the theoretical bilayered earth discussed in the geotechnical literature (McNeill 1980, Callegary et. al 2007).

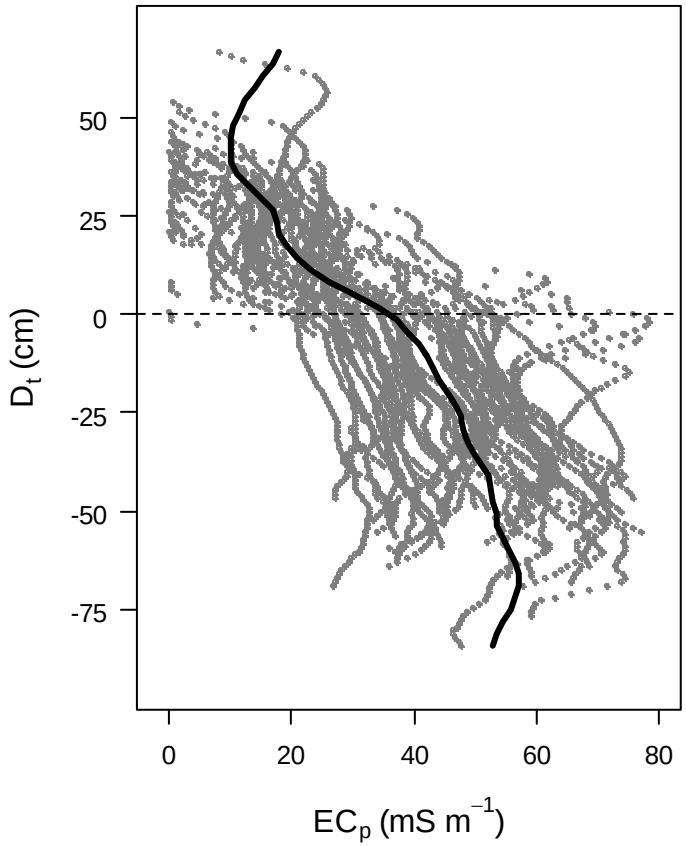

Figure 2. Clay-maximum translated depth (D_t) profiles of EC_p from 44 locations in 4 claypan fields. The depth scale is translated such that the profile clay maximum for each location is at 0 cm (dashed horizontal line). Translated depth is positive above the claypan and negative below it. Measurements of EC_a increase below the claypan (D_t<0) as emphasized by the locally weighted regression (solid black line).

Measurements of EC_m were highly correlated with EC_p measurements (Pearson correlation coefficient = 0.82), but were greater by a factor of 3.3 (fig 3). These results confirmed the EC_p sensor measurement and further indicated that as clay content decreases, profile EC_p increases – counter to the clay-source hypothesis. The greater magnitude of EC_m data relative to EC_p warrants further investigation but is likely due to the core pressing procedure used in the EC_m measurement.

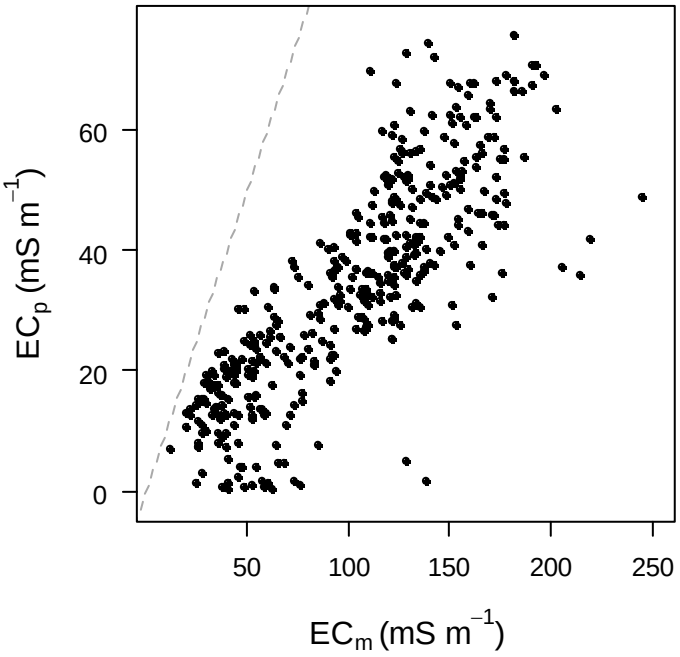

Figure 3. Scatterplot showing the correlation between EC_a in claypan soil profiles measured in-situ by penetrometer (EC_p) and ex-situ by a Wenner mini-probe (EC_m). The Pearson correlation coefficient between these sensor measurements is 0.82 and EC_m was proportional to EC_p by a factor of 3.3.

EC_p Predicted Depth to Claypan

A major objective of this research was to examine the potential for using EC_p to rapidly identify and map subsoil features. The claypan is a critical soil morphological feature because it impacts hydrology, plant available water capacity, water quality, subsoil fertility, root distribution, and crop yield. The claypan transition peak was clear on first derivative plots of EC_p (fig. 4a).

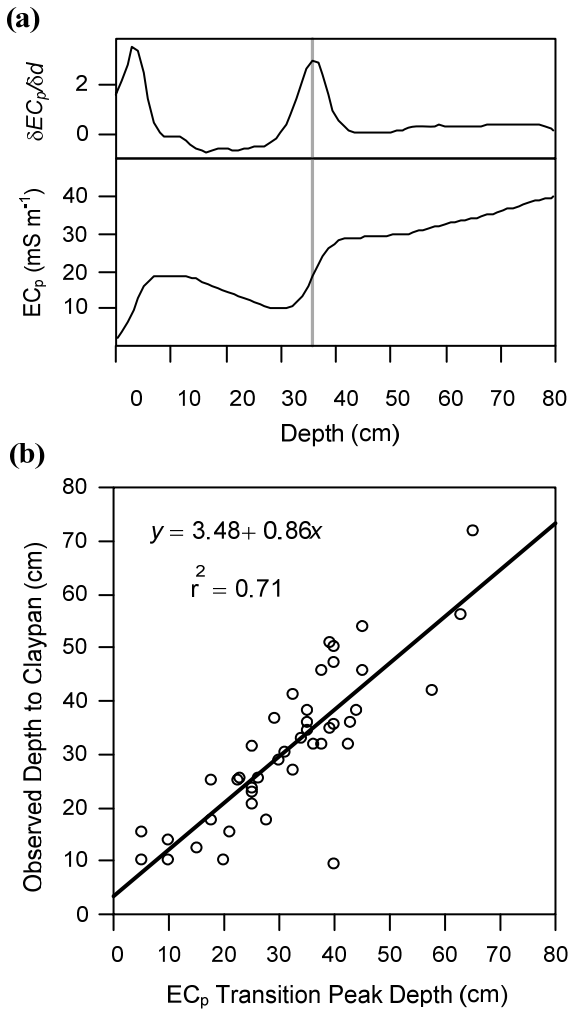

Figure 4. Penetrometer electrical conductivity (EC_p) and first derivative ($\delta EC_p/\delta d$) of a representative claypan soil profile. EC_p transistion peak is identified by a vertical grey line. (b) Depth to EC_p transition peak is analogous to depth to claypan (DTC) and is compared to observed DTC by linear regression.

Claypan transition peaks indicate the depth at which an experienced soil morphologist would describe the E-Bt boundary. The EC_p sensor allows an objective and quantitative determination of the claypan and provides a continuous representation. Depth to claypan was significantly related to EC_p transition peak depth (r^2=0.71, n=44, p<0.001) (fig. 4b). This result includes data from all four study sites which are at opposite ends of the Missouri claypan region. Based on these results, transition peak depth (TPD) might be used to

predict claypan depth anywhere within this area. The penetrometer can rapidly capture short-range spatial variability with multiple penetrations, and may be more consistent, quantitative, and time efficient than a morphologist could be using cores or pits. This type of relationship is very useful for densifying investigations along a transect or within an area (Drummond et al., 2005). Quantified EC_p feature observations can be efficiently collected at smaller intervals while more time-intensive coring or auguring can be performed at larger intervals.

Calibrating EC_a to EC_p Features

Partial least squares regression modeling of transition peak depth as a function of the five EC_a variables produced varying results, from no significant model for field A to a very good relationship for field C (table 1). The pedons from these four field sites were chosen, based on previous research needs, to represent the landscape variability present within each field. However, the field datasets differed in their realization of this goal. Fields C and D have greater relief and a wider distribution of landscape positions and thus a greater range of DTC than field A. Field B had a relatively wider range in DTC than A, but had a smaller number of pedon samples concentrated in a fairly narrow range of DTC. Transition peak depth was poorly related to EC_a for these fields. Pooled results showed a moderate relationship (table 1). A larger or better stratified sample of EC_p profiles from within each field would likely be needed to produce better results. The potential for within field mapping of EC_p features with EC_a data is shown in the site C results (table 1, fig. 5).

Table 1. Fit statistics and number of components for partial least squares regression models of EC_p transition peak as a function of five EC_a sensor measurements.

		Intercept	Comp. 1		Comp 1-2		Comp. 1-3		Comp 1-4	
Field	**n**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**
A	16	12.38	--	--	--	--	--	--	--	--
B	7	17.67	-0.01	15.20	0.21	13.45	--	--	--	--
C	9	13.67	0.80	5.40	0.76	5.90	0.87	4.45	--	--
D	11	16.00	0.64	8.75	--	--	--	--	--	--
Pooled	43	13.71	0.37	10.67	0.39	10.48	0.43	10.12	0.44	9.99

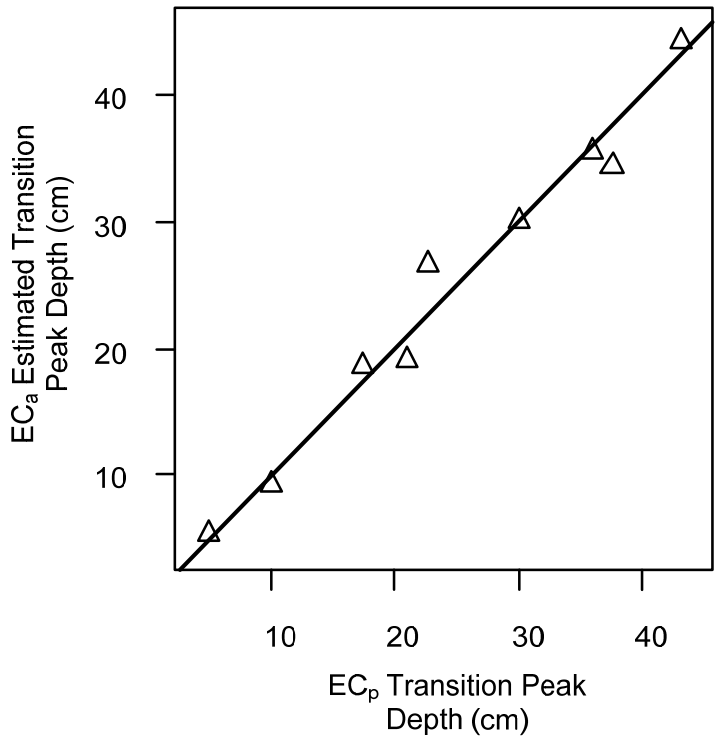

Figure 5. Partial least squares regression estimate of the transition peak depth (TPD) in EC_p depth profiles compared to measured TPD from site C.

Multi-component PLSR models provided only minor gains in R^2 or root mean squared error for transition peak depth. One or two components accounted for most of the variability within sites B, C, and D. This suggests that a single EC_a instrument with dual-simultaneous investigation depths is sufficient for mapping TPD. The pooled model included four components, due to additional orthogonal variability in EC_a potentially caused by temporal differences in temperature and

soil moisture between EC_a surveys. This asynchrony in survey conditions is known to cause bias between surveys of the same field (Abdu et al 2007).

Profile Sources of Proximal EC_a

According to Corwin and Lesch (2005), and discounting salinity and Na^+ saturation, the next most important factors determining profile conductivity are water content and BD. As mentioned previously, lesser EC_p in surface soils is probably due to granular structure and silty texture causing reduced particle contact and proximity (fig. 1a). Minimum conductivity in the strongly leached E horizons may have been due to the greater felsic mineral (e.g. quartz, feldspar) content and reduced contact of the silt-sized particles. The particle contact pathway of EC_a may be dominating the response above the claypan.

The large positive EC_p gradient at the transition peak coincides with the largest increase in clay and water content (fig. 1b,c). Conversely, elevated concentration of expanded smectite clays in the claypan cause a reduction in BD (fig. 1c,d). These relationships suggest that within the transition zone, EC_p is more sensitive to the clay-bound soil-solution conductivity pathway (perhaps enhanced by large cation saturation) rather than to the particle contact pathway. This is in contrast to what happens below the claypan where BD and CI are greater.

We found that clay and water content decreased below the claypan while BD and CI both increased. These relationships suggest that below the claypan, EC_p is more sensitive to the particle contact pathway than to the soil solution or solution-particle pathway. Structural units also tend to be larger in size with

depth. Profile distribution of clay, bulk density, structure, and water content are confounded by soil genesis. Integrated processes of soil formation, including loess deposition, eluviation, illuviation, and subsoil densification, vary systematically with landscape. This combination of effects is probably responsible for success in the calibration of proximal EC_a to DTC and TPD.

CONCLUSIONS

The spatial resolution of EC_a sensors and the depth resolution of EC_p sensors offer the potential to synergistically improve high resolution soil mapping. Claypan soils are successfully handled in this way because they are essentially bilayered with respect to EC_a. Direct calibration of EC_p depth profile features to soil profile features such as depth to claypan is effective, but global or regional hierarchical calibration of EC_a to EC_p features is complicated by field-to-field and temporal variability in EC_a measurements. In general, the multiple EC_a sensor geometries of the common commercially-available platforms studied here do provide at least two orthogonal vectors of EC_a information. Profile conductivity actually increases somewhat below the claypan, probably due to increased particle contact in denser soils. Response of EC_a to EC_p transition peak and DTC is due to the confounded processes of soil and landscape genesis rather than just depth to argillic horizon clay minerals. It is a combined effect of lesser EC_a near the surface, a profile minimum EC_a in E horizons, greater EC_a in the claypan, and even greater EC_a in the dense soil material below the claypan.

REFERENCES

Abdu, H., Robinson, D.A., and Jones, S.B. 2007. Comparing bulk soil electrical conductivity determination using the DUALEM-1S and EM38-DD electromagnetic induction instruments. Soil Sci. Soc. Am. J., 71:189-196.

Callegary, J.B., Ferre, T.P.A., and Groom, R.W. 2007. Vertical spatial sensitivity and exploration depth of low-induction-number electromagnetic-induction instruments. Vadose Zone J., 6:158-167.

Corwin, D.L, and Lesch, S.M. 2005 Apparent soil electrical conductivity measurements in agriculture. Comput. Electron. Agr., 46:11-43.

Drummond, P.E., Christy, C.D. and Lund, E.D. 2005. Using an automated penetrometer and soil EC probe to characterize the rooting zone. [CD-ROM]. *In:* P.C. Robert et al. (eds.) Proceedings of the 5th International Conference on Precision Agriculture, Minneapolis, MN, Jul. 16-19, 2000, ASA, CSSA, and SSSA, Madison, WI.

Dolittle, J.A., Sudduth, K.A., Kitchen, N.R., and Indorante S.J. 1994. Estimating depths to claypans using electromagnetic induction methods. J. Soil Water Cons. 49(6):572-575.

McNeill, J.D. 1980. Electromagnetic terrain conductivity measurments at low induction numbers. Geonics Technical Note TN-6. Mississauga, Ontario.

Sudduth, K.A., Hummel, J.W., Kitchen, N.R., and Drummond, S.T. 2000. Evaluation of a soil conductivity sensing penetrometer. 2000 ASAE International Meeting, Milwaukee, WI, July 9-12. Paper No. 001043: ASAE, St. Joseph, MI.

Sudduth, K.A., and Kitchen, N.R. 2006. Increasing information with multiple soil electrical conductivity datasets. 2006 ASABE International Meeting, Portland OR, July 9-12. Paper No. 061055: ASABE St. Joseph, MI

Pellerin, L and P.E. Wannamaker. 2005. Multi-dimensional electromagnetic modeling and inversion with application to near-surface earth investigations. Comput. Electron. Agr., 46:3, 71-102.

CHAPTER 3

Estimating Claypan Soil Profile Properties by Combined Proximal and Penetrating Sensors.

INTRODUCTION

The extensive mapping of the world's soils that has occurred up to now is a remarkable achievement by many nations; however, the decades long effort is currently outpaced by the scale of modern soil management problems (Bramley and Janik, 2005; Dobos et al., 2006). Additionally, there is an increasing demand for quantitative soil data, which is not well served by the traditional framework of soil mapping (McBratney, 1992; McSweeney et al., 1994). For instance, crop yield monitors have highlighted the tremendous (>150%) grain yield variability that can occur within and between soil map unit delineations in crop fields (Jaynes and Colvin, 1997; Kitchen et al., 1999; Kaspar et al., 2003). The problem of managing soil variation is critical due to the potential improvement of profitability and soil and water quality (Lerch et al., 2005). Biogeochemistry is receiving considerable attention in agroecosystem and ecosystem contexts as mankind struggles to understand human impact and sustainability (Post and Kwon, 2000). Intensive land use management and investigations in plant productivity, soil and water quality, and biogeochemistry need high resolution quantitative soil property information in order to meet their objectives.

Potential solutions for the problems listed above can be analyzed in four dimensional (4-D) models of the soil landscape. Hydrology, water quality, crop growth, biogeochemical, and other ecological process models require 3-D (spatial) soil property data in order to achieve the 4-D (spatiotemporal) result, from which management decisions and policy can be made. Gessler et al. (2000) identify the utility of geomorphic soil landscape models for high resolution

application of biogeochemistry models such as CENTURY. The suitability of available soil and landscape data for process modeling was reviewed by Wilson (1996), and significant need for reliable soil property data was identified. Ahuja et al. (2006) concluded that the future of process models in agricultural systems will be limited by the quantity of soil profile data available to populate the models. They suggest high resolution soil property data was needed to meet the requirements of modelers and advocated the development of innovative methods to predict soil properties.

This research proposes that soil sensors calibrated to accurately measure soil properties can provide high resolution and quantitative information to meet the needs described above. Soil property sensors on mobile platforms have the capability to make measurements *in-situ*, and if successfully calibrated, to estimate soil properties at many more locations than reference lab measurements. Further, this research hypothesizes that combining different types of soil sensors will be better for predicting soil properties than single sensor systems. This hypothesis was examined in the Central Claypan Region of northeast Missouri for specific practical combinations of commercially available soil sensors. The soil sensors were a two channel ground conductivity sensor, a penetrometer with a split cone electrode, and a spectral radiometer.

Several types of sensors measure the bulk apparent soil electrical conductivity (EC_a). The two most common methods are electromagnetic induction (McNeill, 1980), and soil contact electrodes (Lund et al., 1999). The measurement is volume integrated by a function which is dependent on the

geometry of the sensor. These sensors have been developed for surface based geophysical investigations (McNeill, 1980), and have more recently been adopted for precision agriculture applications (Kitchen et al., 1999; Lund et al., 2000), and environmental monitoring (Benson et al., 1997; Atekwana et al., 2000). For the purposes of this paper these sensors are designated 'proximal' EC_a sensors since they measure the entire depth weighted profile conductivity from the surface of the soil. Applications of these proximal soil sensors include mapping of soil salinity (Corwin and Lesch, 2005), topsoil depth (Brus et al., 1992; Doolittle et al., 1994), water content (Reedy and Scanlon, 2003), plant available water content (Jiang et al., 2007) and CEC and soil texture (Sudduth et al., 2005). Given that EC_a is related to soil morphological features that correlate with soil physical and chemical properties, there is a potential for EC_a to estimate these properties at specific depths in the profile.

Resistance of soil to insertion of a cone tipped rod or penetrometer has been used for quantitative investigation of soils for over 70 years (Richards, 1941). Measurements from cone penetrometers are expressed as cone index (CI) which is the force necessary for insertion divided by the surface area of the cone (Perumpral, 1987). Common uses of the penetrometer include using CI to estimate soil bulk density (Ayers and Perumpral, 1982), measure soil compaction (Evans et al., 1996), explain root growth and distribution (Taylor and Gardner, 1963; Thompson et al., 1987), and delineate parent materials (Grunwald et al., 2001). Modern adaptations of the penetrometer include extended depth profiles (Butler et al., 2002), multiple probes (Raper et al., 1999), computerized data logging equipment (Rooney and Lowery, 2000), and the integration of additional

sensors such as, electrical capacitance (Sun et al., 2006), dielectric permittivity (Vaz et al., 2001) electrical conductivity (Beck et al., 2000; Drummond et al., 2000; Schulmeister et al., 2003), and diffuse reflectance (Hummel et al., 2004) sensors. Recently, horizontal penetrometers have been employed to map soil properties (Mouazen et al., 2003; Adamchuk et al., 2004; Andrade-Sanchez et al., 2007; Chung et al., 2008) and also have been combined with multiple sensors to measure capacitance (Adamchuk et al., 2004; Sun et al., 2006; Richard et al., 2008) and radar (Richard et al., 2008).

Another proven technology for soil sensing is near infrared (NIR), visible to near infrared (VNIR), and mid infrared (MIR) diffuse reflectance spectroscopy (Sudduth and Hummel, 1991; Ben-Dor and Banin, 1995; Viscarra Rossel and McBratney, 1998). Recently the technique is beginning to take hold as a supplement to or replacement of lab measured soil properties by calibration of spectral libraries using *ex-situ* dry ground soils (Shepherd and Walsh, 2002; Brown et al., 2006; Viscara-Rossel et al., 2008). A better scenario for the use of diffuse reflectance sensors would be successful *in-situ* field deployment. Hummel et al. (2004) developed an NIR diffuse reflectance penetrometer intended for simultaneous measurement of CI and water content in the soil profile. Waiser et al. (2007) simulated *in-situ* measurements with a soil penetrometer by measuring moist soil cores in a smeared and unsmeared condition. A commercial shank type sensor for near-surface on-the-go VNIR diffuse reflectance mapping is currently available (Christy, 2008) and commercial penetrometer systems are likely to be developed.

The commercial availability of EC_a, EC_p, CI, and VNIR sensors bring the possibility that such sensors could be employed in practical combinations to estimate soil profile properties (S_i). The purpose of this study is to consider this possibility and test the value of combined sensor measurements to estimate claypan soil properties. The most basic combination would be a penetrometer with EC_p and CI sensors, [1]. Proximal conductivity sensors are becoming more common and therefore the second incremental combination pairs the conductivity penetrometer variables of EC_p and CI with one or more channels of EC_a, [2]. The third incremental combination replicates a penetrometer with EC_p, CI, and VNIR diffuse reflectance sensors, [3]. The fourth combination again adds one or more channels of EC_a, [4]. For comparative purposes, the fifth model represents a VNIR-DRS sensor alone, [5].

$$S_i = f(EC_p, CI) \quad [1]$$

$$S_i = f(EC_p, CI, EC_a) \quad [2]$$

$$S_i = f(EC_p, CI, VNIR) \quad [3]$$

$$S_i = f(EC_p, CI, EC_a, VNIR) \quad [4]$$

$$S_i = f(VNIR) \quad [5]$$

The primary goal of this research was to determine if the field deployable combinations of sensor measurements can improve estimation of organic carbon (OC), clay content (clay), and 1:1 soil:water pH, and electrical conductivity (pH and $EC_{1:1}$) over single sensor systems. A secondary goal was to examine the consequence of measuring VNIR diffuse reflectance on *ex-situ* soil cores while measuring EC_p on *in-situ* soils since the sensors are not measuring the same soil

volume. To achieve this goal, a miniaturized Wenner array on a hand probe was used to measure soil electrical conductivity (EC_m) at high resolution on ex-situ soil cores at the same locations as measurements of VNIR diffuse reflectance.

Specific objectives were to:

i. Determine best sensor combinations (1-5 above) for the estimate of OC, clay, $pH_{1:1}$, and $EC_{1:1}$.

ii. Examine the performance of *ex-situ* EC_m measurements versus *in-situ* EC_p measured from adjacent soil in combined sensor estimates of soil properties.

MATERIALS AND METHODS

Sites and Soils

Four agricultural fields in the claypan region of northeast Missouri were chosen for this study, three fields (A, B, C) with a loess solum near Centralia, MO (Lat. 39° 13' 43", Lon. 92° 8' 20") and a paired watershed research site (D) with a loess-till solum, near Novelty, MO (Lat. 40° 1' 46.5", Lon. 92° 11' 19"). The Centralia site is located on the southern end of MLRA 113 while the Novelty site is on the northern end, about 90 km away. These sites were chosen to represent the range of parent materials and landscape positions commonly seen in the uplands of the Central Claypan Area of northeast Missouri. The common layering of parent materials in these glacially formed landscapes is Wisconsinan loess (0.5-1.5 m), Illinoian pedisediment (0-0.5 m), and the underlying pre-Illinoian glacial till (Follmer, 1983; Ruhe, 1969).

Soil cores and penetrometer measurements were taken at a total of 75 locations in the four study fields (17, 10, 15, and 33 on fields A, B, C, and D respectively). These locations are coincidental with existing soil characterization data and observations from previous research projects described by (Sudduth et al., 2005) for fields A B C, and (Udawatta et al., 2002) for field D. These sites were selected by those researchers to span the landscape positions present within each field and thus are representative of the upland soils in the central claypan region as a whole. Three soil cores 1.2 m in length and 4.2 cm in diameter were sampled at each site within 1-2 days after penetrometer measurements. One core was cut and bagged by 15-cm increments in the field for determination of gravimetric water content and bulk density. Another core was retained as backup and for profile descriptions. The third core was retained for the benchtop sensor measurements described below.

In-Situ Soil Sensors

Bulk soil electrical conductivity was measured from the surface with the Veris† 3150 rolling coulter Wenner array (Veris Technologies, Salina, KS, USA) with 0.7 ($EC_{a\ v\text{-}sh}$) and 2.2-m ($EC_{a\ v\text{-}dp}$) electrode spacing. Global positioning was used to map EC_a at 1 Hz on 10-m transects. Ordinary kriging was used to estimate EC_a at the soil coring sites. Survey dates and field conditions were not well synchronized between field sites. Fields B and D had all EC_a surveys made within a relatively narrow window of one month in the spring of 2007, while surveys of

†Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Dept. of Agriculture, Univ. of Florida, or Univ. of Missouri.

fields A and C were made within three days in the fall of 2005. This asynchrony in survey conditions is known to cause bias between field EC_a measurements due to temperature and soil moisture variations (Abdu et al., 2007; Weller et al., 2007).

A Veris Profiler 3000 (Veris Technologies, Salina, KS, USA) was used to measure CI and EC_p to 92-cm depth, with 5 penetrations per site. Replicate EC_p soundings were pooled and fitted with locally weighted regression models (Cleveland and Devlin, 1988) for further analysis. Measurements were collected every 1.27 cm with depth. The Profiler 3000 uses a split cone electrode to measure conductivity. The penetrometer shaft is directly attached to the upper portion of the cone while the lower portion of the cone is separated by an insulating ring and wired through the hollow shaft. To prevent soil contact and potential electrical bridging between the shaft and the electrode, heavy duty shrink tubing was applied to the penetrometer shaft. Measurements of EC_p were made on all fields in the late spring of 2007. This occurred within a few days of EC_a measurements on fields B and D; however, EC_p on fields A and C were measured approximately 18 months after the EC_a surveys.

For clarity in this report EC_p is used to refer to penetrometer measured EC_a to differentiate it from proximally measured EC_a. Additionally, a miniaturized Wenner array on a hand-held probe (mini-probe) was constructed to perform more detailed investigations on *ex-situ* cores. The measurement provided by this sensor is denoted as EC_m.

Ex-Situ Soil Sensors

Die-Pressing of Soil Cores

Measurements of EC_m from the mini-probe and a VNIR diffuse reflectance sensor (both sensors described below) were made *ex-situ* on moist soil cores. A die-pressing procedure was used to control core quality and provide uniform sensor placement. Soil cores obtained from the claypan landscape by hydraulic sampler vary in uniformity and consolidation due to the clay content of the core and moisture conditions at the time of sampling. Argillic horizons of claypan soils can be tacky and can adhere to coring equipment, smearing and balling up on the surface of the core. Surface soils usually are friable, and possess granular structure which can disintegrate easily. Soil cores will often break at one or more places during sampling, transport, and handling. These factors can lead to cores with many unnatural fine and large cracks and irregular surfaces. Measurements made with the contact probe of the diffuse reflectance sensor or with the mini-probe would likely have been influenced by the quality of the core in the original condition.

A 1.5 m linear steel die was constructed from a 57 mm right-angle channel of 6.4 mm thickness (fig. 1). The V-shaped die was lined with plastic film into which the soil core was placed. The core was wrapped with the film and then covered with a wooden tray and a steel C-channel. A forklift was driven onto the die assembly until full closure achieved, ensuring consistent shaping and compression. The die formed a right triangular prism the length of the core which was then handled by the wooden tray. Die pressing conditioned the *ex-situ* soil

sample to emulate the *in-situ* soil seen by a penetrometer based sensor, repairing various previously discussed inconsistencies caused by core sampling.

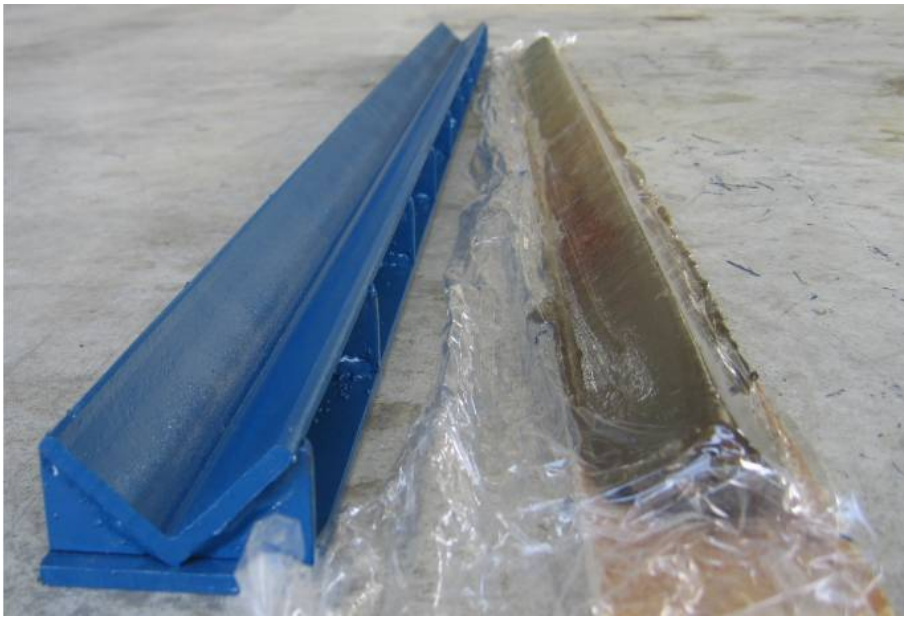

Figure 1. Ex-situ sensor measurements were performed on die-pressed soil cores. A pressed core is shown along side the forming die. Not shown is the C-channel which contains the forming die and receives the forming pressure applied by the wheels of a forklift.

Another important purpose for die pressing was to provide a flat smooth surface for a consistent depth of mini-probe insertion and consistent placement of the diffuse reflectance contact probe. Consistent placement of the mini-probe was important as measurements of EC_m varied by the volume of the soil under the mini-probe which changed depending on the location of insertion. The mini-probe fixture (fig. 2) and prismatic core combined to make this volume uniform among depths and cores. Insertion angle of the mini-probe was also controlled by the fixture. The flat surfaces and shape of the core allowed identical positioning of the mini-probe and diffuse reflectance measurements.

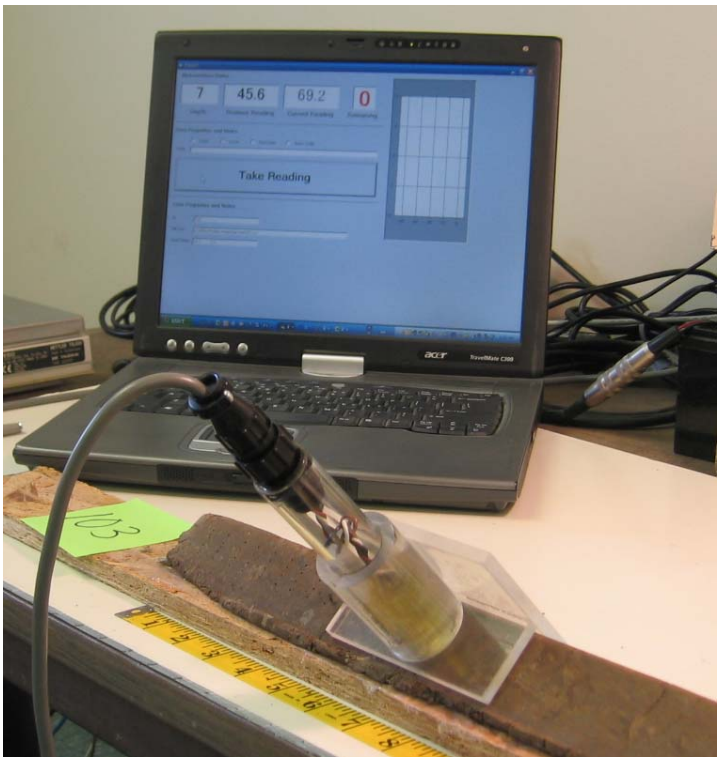

Figure 2. Miniaturized Wenner soil electrical conductivity probe (mini-probe). The probe fixture controlled insertion angle and position.

EX-SITU Visible to Near Infrared Diffuse Reflectance Spectroscopy

Diffuse reflectance of moist soil cores was measured with a FieldSpec Pro spectral radiometer (Analytical Spectral Devices Inc., Boulder, CO). Spectra were measured using a contact probe (ASDI model #A122300), with integrated light source and fiber-optic, 10 mm spot size. A polycarbonate fixture held the contact probe in a perpendicular axis from the die-pressed soil core, ensured shielding from stray light and provided consistent positioning of the sensor relative to the mini-probe sensor. The spectral radiometer measured diffuse reflectance spectra at 1 nm intervals from 350 to 2500 nm. Spectra were measured at 1.27-cm depth increments, producing 75 to 100 VNIR measurements along each soil core over about 10 to 15 minutes. A Spectralon® white reflectance standard was scanned at

the head and tail of each core. The average of these scans was used to calculate reflectance for the measurements of each core. All analysis were performed with reflectance from every fifth wavelength (430 wavelengths).

EX-SITU Wenner Mini-Probe Soil Electrical Conductivity

A Wenner mini-probe bulk apparent electrical conductivity (EC_m) sensor was constructed for this study (fig. 2). Four steel needles were embedded in epoxy within a polycarbonate 25 mm tube. The mini-probe had 5 mm electrode spacing, 5 mm insertion depth, and was operated using the electronics from a Veris 3150 mobile sensor platform. Veris supplied a custom software modification with a corrected scale parameter for theoretical calculations describing the electrode configuration.

Mini-probe measurements of EC_m were made at high resolution (every 1.27 cm) on the die pressed cores. Data logging software automated these measurements and allowed the notation of EC_m features such as Fe-Mn oxide mottles and stains. Calibration standards were used to verify EC_m across the range of measurement and were found to agree within 10%. Gravimetric soil moisture determinations were made from the die pressed cores on 2.54-cm intervals (coinciding with EC_m and VNIR measurements). The resulting samples were ground and retained for the lab measurements, described next.

Reference Soil Measurements

Four lab-measured soil properties were targeted for investigation: organic carbon (OC), clay content by hydrometer, and 1:1 (soil:water) suspension pH and

EC ($pH_{1:1}$, and $EC_{1:1}$). These measurements were chosen for their expediency and low relative cost given the large sample dataset and for their contrasting depth profiles. A subset of 32 coring locations was chosen from the original 75 locations for model development. These sample locations were chosen to span the landscape positions and parent materials of all four fields, and subsequent research will focus on model application at all 75 locations. Nine to ten sample segments (2.54 cm) from each of the 32 modeling subset were chosen to represent the characteristic parent materials, horizons, and horizon boundaries, with special emphasis on the clay content gradient between A, E, and Bt1 horizons. Sensor combination comparisons were made using the full model development dataset. These samples were further divided into independent calibration and test datasets, sixteen samples each, for final model development. Total sample size for the calibration and test dataset were 233 and 245 respectively. The clay dataset was half this size and divided into proportionally smaller calibration and test datasets.

Organic carbon was determined on 0.33 g subsamples by a LECO C144 induction furnace CO_2 analyzer (LECO Corporation, St Joseph, MI). Soil texture, including clay, was determined by the hydrometer method on 30-40 g samples. The small sample size and the precision of this method led to a theoretical measurement error of 2.5 %. Silt content in this dataset had large negative correlation with clay content (r=-0.75), essentially its inverse, while sand contents were minute and did not provide useful results. Only the clay measurements are included with this analysis. Soil suspension $pH_{1:1}$ was measured with a standard pH meter and ruggedized pH electrode, while $EC_{1:1}$ was

measured with a four electrode conductivity cell. The four electrode design of the conductivity cell allows operation in suspensions and slurries by automatically correcting for electrode polarization and changes in cell constant due to adherence of suspended particles (DuraProbe™, Thermo Fisher Scientific, Waltham, MA). One-to-one mixtures of soil and deionized water (10g:10 ml) were shaken for 1 hour and $pH_{1:1}$, then $EC_{1:1}$ measurements were taken while agitating the electrodes in the soil slurry.

Statistical Methods

Analysis of the multivariate soil and sensor dataset to determine optimal sensor combinations proceeded in four steps: 1) outlier detection, 2) modeling soil properties as a function of sensor combinations by partial least squares regression (PLSR) with cross validation, 3) selection and comparison of sensor combinations by model performance statistics, and 4) variable selection and comparison. Final model development with the optimal sensor combinations was performed with independent calibration and test datasets. The calibration and test datasets both spanned the parent materials and landscape positions found in the study fields, but did not include the same sampling locations. The R open source statistical programming environment was used for data handling, outlier removal procedures, PLSR modeling and model selection (R Core Development Team, 2008).

Partial least squares regression (Wold et al., 2001) was used as the primary analytical technique of this study because of the large number and correlated nature of the independent variables. Best results from calibration type

experiments using PLSR can be obtained when calibrating a sensor or measurement within a defined target population using a representative sample of measured data (as described above) while free of gross measurement errors. Spectral and sensor data outlier detection and measured data outlier detection were performed separately with a robust outlier detection procedure (Filzmoser et al., 2008). The procedure identified outliers on a robustly sphered principal components decomposition of the data space.

The NIPALS 2 algorithm was used to model the measured soil properties as a function of the sensor variables (Mevik and Wehrens, 2007). The model comparison and variable selection steps were performed on the merged calibration and test datasets (i.e. the whole dataset) using a jackknife cross-validation procedure with 10 segments. The optimal number of components for each PLSR model was selected by the Akaike Information Criterion which penalizes the root mean squared error (RMSE) of the model by its effective degrees of freedom. The root mean squared error of calibration (RMSEC), cross-validation adjusted coefficient of determination (R^2), and ratio of the standard deviation of the measured variable to the RMSE of the fit (RPD) were used as model comparison statistics. Good model estimates were determined as those having an $R^2>0.7$ and an RPD >2.0 (Chang et al., 2001). Best sensor combination models were then re-calibrated and tested with the fully independent calibration and test datasets.

The PLSR results allow the calculation of the variable importance to projection (VIP) statistic to examine predictors or spectral features that are

important for reducing error in the PLSR estimate (Wold et al., 2001; Chong and Jun, 2005). This statistic was used to identify four influential wavelengths from the full range VNIR models of OC, clay, $pH_{1:1}$, and $EC_{1:1}$ without combining EC_p, CI, or EC_a sensor data. These reduced variable models were used to compare the information content and usefulness of the various sensors to estimate soil properties.

RESULTS AND DISCUSSION

Basic Relationships among Sensor and Measured Variables

Descriptive statistics (table 1) and correlations among the soil and sensor measurements (table 2) provide an overview of the dataset and point to trends likely to be seen in the modeling results. A few selected correlations with VNIR wavelengths are shown based on their performance in the variable selection procedure (results below). Organic Carbon was negatively correlated with CI, EC_p, and visible range diffuse reflectance (425-1000 nm). This result is due to the roughly exponential decrease in OC with depth (fig. 3) and the general increase with depth in CI, EC_p, and visible diffuse reflectance.

Table 1. Basic statistics for the soil property and sensor data.

	Unit	N	Mean	S.D.	Min	Median	Max
OC	g kg-1	302	0.56	0.37	0.04	0.42	1.77
Clay	g kg-1	167	41.1	10.0	17.3	40.3	66.5
pH1:1	-log[H+]	314	5.8	0.7	4.6	5.7	7.8
EC1:1	mS cm-1	315	245	83	88	240	508
Water	g kg-1	301	0.20	0.04	0.11	0.20	0.36
ECp	mS m-1	236	35.9	15.5	1.1	35.6	74.1
CI	kPa	236	112.2	30.8	32.2	113.8	197.4
ECm	mS m-1	271	108.3	43.5	25.9	118.7	205.6
Vsh	mS m-1	33	12.1	4.9	4.3	10.9	24.6
Vdp	mS m-1	33	35.1	12.3	14.3	33.9	59.9

Quite different from OC, the correlations of clay and $EC_{1:1}$ with CI and EC_p were positive and moderate while the correlations with near infrared diffuse reflectance (1370-2500 nm) were negative and large. Correlations of $pH_{1:1}$ with EC_p and CI were smaller in general and opposite in sign to those seen for clay and $EC_{1:1}$. Indeed, clay and $pH_{1:1}$ have a large negative correlation (r = -0.78). Correlations of the measured soil properties to the two EC_a variables ($EC_{aV\text{-}sh}$, $EC_{aV\text{-}dp}$) were low or not significant. These results indicated poor potential for combining the proximal EC_a data with the penetrometer based data for estimations of OC, clay, and $pH_{1:1}$, but suggested that EC_p might provide useful information for combined sensor models.

Table 2. Correlation coefficients for the measured soil variables, the sensor variables, and diffuse reflectance at selected visible to near infrared wavelengths (VNIR). All entries are significant (p>0.001).

	Depth	OC	Clay	$pH_{1:1}$	$EC_{1:1}$	Water	EC_p	Force	EC_m	$EC_{aV\text{-}sh}$	$EC_{aV\text{-}dp}$	480 nm	560 nm	755 nm	960 nm	1370 nm	1415 nm	1870 nm	1640 nm	2225 nm
Depth	1																			
OC	-0.86	1																		
Clay	0.33	-0.15	1																	
$pH_{1:1}$	-0.40	0.25	-0.78	1																
$EC_{1:1}$	0.49	-0.34	0.58	-0.40	1															
Water	-0.24	0.37	0.54	-0.33	--	1														
EC_p	0.59	-0.58	0.42	-0.19	0.62	--	1													
Force	0.67	-0.58	0.28	-0.41	0.27	-0.18	0.32	1												
EC_m	0.67	-0.62	0.59	-0.41	0.77		0.88	0.42	1											
$EC_{aV\text{-}sh}$	--	-0.20	--	0.21	--	--	0.57	-0.18	0.36	1										
$EC_{aV\text{-}dp}$	--	-0.21	--	0.17	0.22	--	0.62	-0.15	0.43	0.96	1									
480 nm	0.61	-0.71	0.25	-0.34	0.33	-0.06	0.42	0.42	0.49	--	--	1								
560 nm	0.69	-0.77	0.39	-0.47	0.33	-0.01	0.46	0.50	0.56	--	--	0.93	1							
755 nm	0.56	-0.60	0.53	-0.64	0.24		0.27	0.49	0.41	--	--	0.72	0.88	1						
960 nm	0.30	-0.41	0.22	-0.43	--	0.21	--	0.28	--	-0.15	-0.17	0.64	0.71	0.86	1					
1370 nm	--	-0.14	-0.61	0.30	-0.55	-0.28	-0.42	--	-0.50	--	-0.21	0.12	0.10	0.12	0.50	1				
1415 nm	--	--	-0.86	0.59	-0.53	-0.50	-0.39	-0.23	-0.52	--	--	--	-0.19	-0.31	--	0.83	1			
1870 nm	--	--	-0.86	0.61	-0.49	-0.56	-0.33	-0.19	-0.46	--	--	--	--	-0.32	--	0.80	0.99	1		
1640 nm	--	-0.13	-0.77	0.49	-0.56	-0.45	-0.38	-0.13	-0.50	--	--	--	--	--	0.25	0.95	0.94	0.94	1	
2225 nm	--	--	-0.88	0.64	-0.50	-0.55	-0.32	-0.21	-0.47	--	--	--	-0.20	-0.37	--	0.77	0.99	1.00	0.92	1

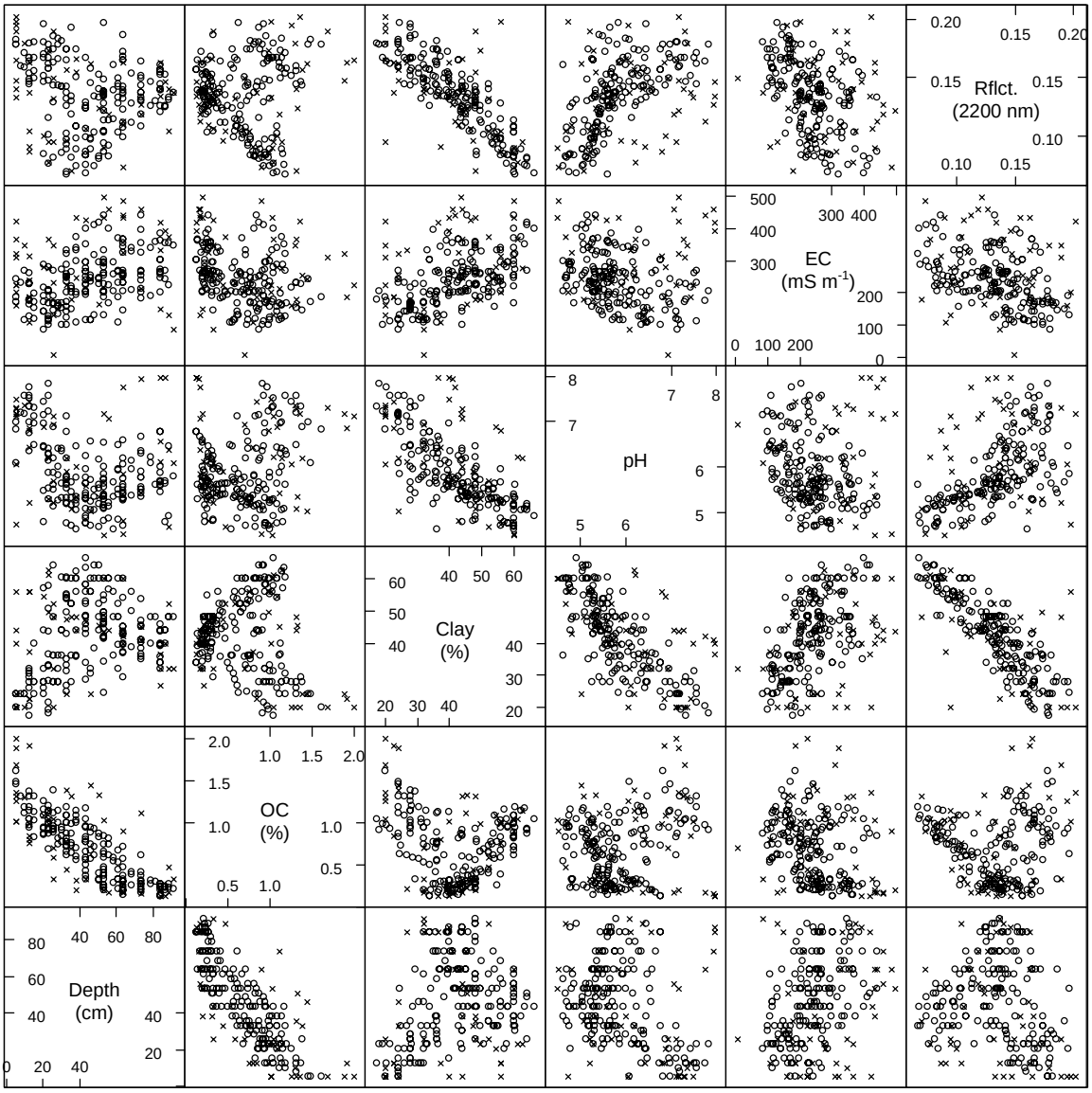

Figure 3. A scatterplot matrix showing the relationship between the measured soil properties of organic carbon (OC), clay content (clay), 1:1 (soil:water) suspension pH and EC ($pH_{1:1}$, and $EC_{1:1}$) and two important covariates, depth and diffuse reflectance at 2200 nm. Points marked with an X were removed by a robust outlier detection algorithm

Overall Combined Sensor Results

The uniform model testing framework produced a consistent ranking of the sensor combinations by the adjusted R^2, RMSE, and RPD of cross-validated PLSR models (fig. 4, table 2). Two trends are apparent in the result. First, the

best estimations were achieved for clay and OC relative to $pH_{1:1}$ or $EC_{1:1}$. Second, models including VNIR spectra are dramatically improved over models using just the EC_p, CI, and EC_a sensor data. The combination of these two trends was that only the OC and clay models which included VNIR spectra as predictors were successful by the RPD and R^2 criteria (>2.0 and >0.7) (table 3).

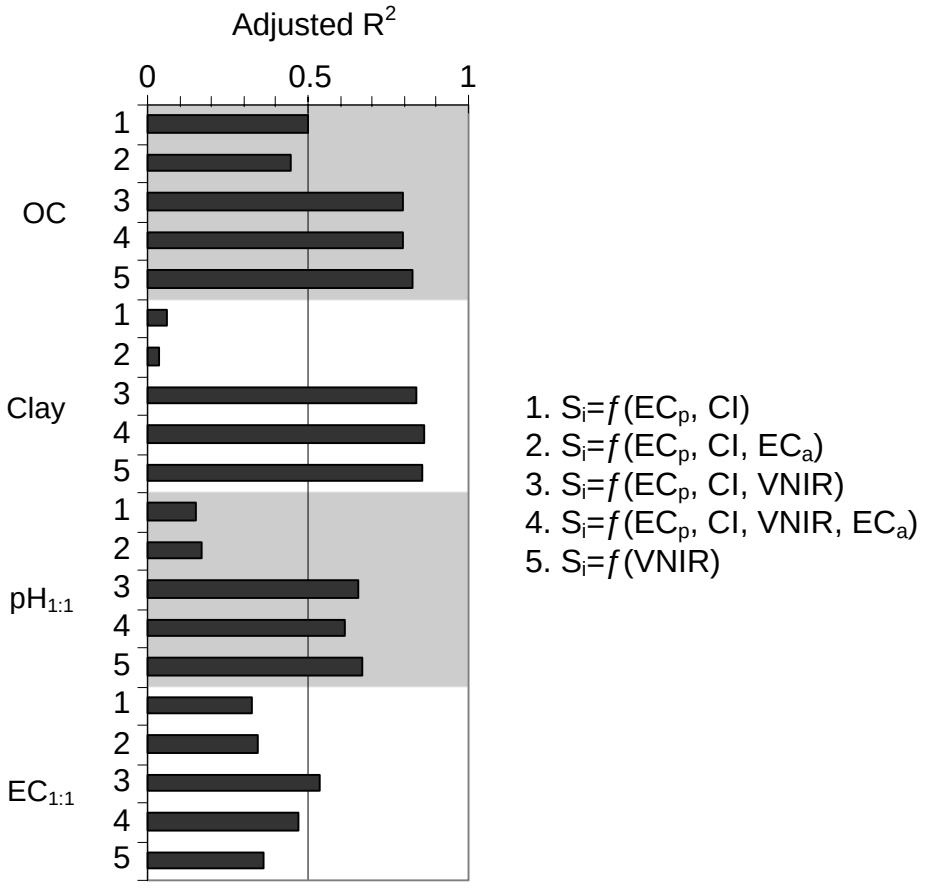

Figure 4. Overall comparison of PLSR model success for estimation of organic carbon (OC), clay content (Clay), 1:1 soil: water soil suspension pH ($pH_{1:1}$) and electrical conductivity ($EC_{1:1}$). Bars represent the R2 of estimates by incremental combinations of soil sensors.

Table 3. Overall model fit results of the combined sensor models based on cross validation results from the full modeling dataset. Shown are the dependent variables, the independent variables, the number of components in the model (n-comp), the cross-validation adjusted R_2 (Adj. R_2), root mean squared error of calibration (RMSEC), and the ratio of standard deviation of the measured variable to the RMSEC (RPD). Models with an R_2>0.7 and an RPD >2.0 were determined to be successful.

S_i	Model	Independent Variables	n-comp	Adj. R^2	RMSEC	RPD
	1	EC_p, CI	2	0.50	0.26	1.00
	2	EC_p, CI, $EC_{a\text{-}Vsh}$, $EC_{a\text{-}Vdp}$	2	0.45	0.27	0.95
OC	3	EC_p, CI, VNIR	10	0.80	0.16	2.12
	4	EC_p, CI, VNIR, $EC_{a\text{-}Vsh}$, $EC_{a\text{-}Vdp}$	10	0.80	0.16	2.12
	5	VNIR	19	0.83	0.15	2.37
	1	EC_p, CI	2	0.06	10.63	0.73
	2	EC_p, CI, $EC_{a\text{-}Vsh}$, $EC_{a\text{-}Vdp}$	2	0.03	10.76	0.72
clay	3	EC_p, CI, VNIR	7	0.84	4.31	2.36
	4	EC_p, CI, VNIR, $EC_{a\text{-}Vsh}$, $EC_{a\text{-}Vdp}$	10	0.86	4.09	2.56
	5	VNIR	11	0.85	3.77	2.53
	1	EC_p, CI	2	0.15	0.71	0.77
	2	EC_p, CI, $EC_{a\text{-}Vsh}$, $EC_{a\text{-}Vdp}$	2	0.17	0.70	0.78
$pH_{1:1}$	3	EC_p, CI, VNIR	14	0.66	0.44	1.67
	4	EC_p, CI, VNIR, $EC_{a\text{-}Vsh}$, $EC_{a\text{-}Vdp}$	6	0.61	0.48	1.47
	5	VNIR	13	0.67	0.42	1.67
	1	EC_p, CI	2	0.33	66.44	0.86
	2	EC_p, CI, $EC_{a\text{-}Vsh}$, $EC_{a\text{-}Vdp}$	3	0.34	65.59	1.01
$EC_{1:1}$	3	EC_p, CI, VNIR	12	0.54	55.05	1.41
	4	EC_p, CI, VNIR, $EC_{a\text{-}Vsh}$, $EC_{a\text{-}Vdp}$	6	0.47	58.99	1.26
	5	VNIR	18	0.36	64.92	1.24

Observed and estimated values from the best model of each target variable are shown in scatterplots (fig. 5) along with the linear fit of model performance. Multi-sensor estimations of OC and clay were also successful by comparison of RMSE with the reported lab error. Measurements of OC by the method used here are generally reported to the nearest 0.1% and the RMSEC of the best model was 0.16%, very near the instrument precision. The theoretical precision of the method used for clay content was less than 3% and the RMSEC of the best model was again similar at 4.1%. The RMSEC for $pH_{1:1}$ (0.4 $pH_{1:1}$) and $EC_{1:1}$ (53 mS m^{-1}) far exceeded the accepted lab measurement error. The value of knowing the

estimated depth profile distribution of these parameters at many sites may outweigh their poor estimates, though the modeled values are biased relative to the measured data (fig. 5). The values of RMSEC obtained for OC and clay indicate the utility of the sensor approach. The precision and accuracy of these estimates is made more useful by the high depth resolution obtainable with the sensors.

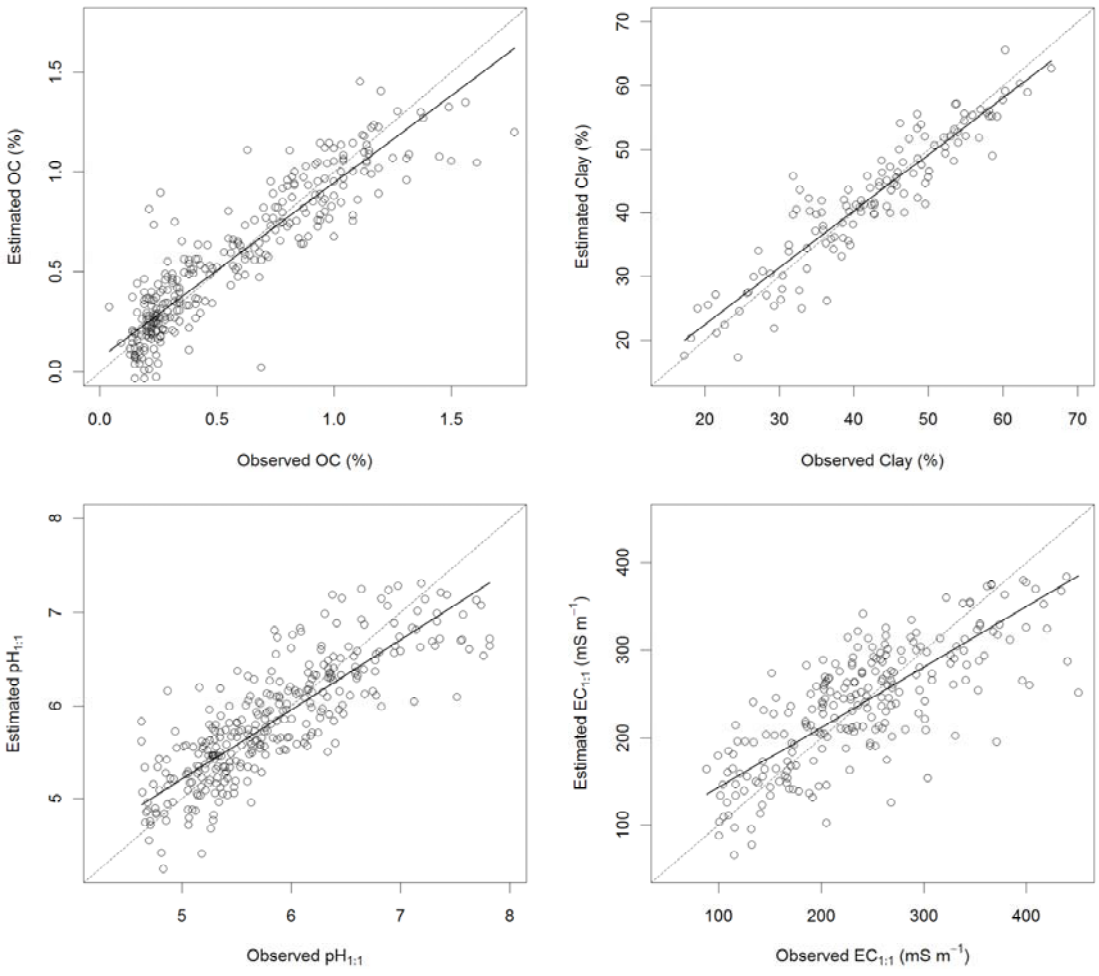

Figure 5. Best PLSR cross-validated models for OC=f(VNIR) (Adj. R^2 = 0.83), clay= f(EC_p, CI, VNIR, EC_a) (Adj. R^2 = 0.86), $pH_{1:1}$= f(VNIR) (Adj. R^2 = 0.67) and $EC_{1:1}$=f(EC_p, CI, VNIR) (Adj. R^2 = 0.54). Dotted line is 1:1 and solid line is the linear fit of the model result. These results are produced from the entire modeling dataset (32 cores).

The success of the combined sensor models to produce good estimates of OC and clay is primarily due to the influence of the VNIR spectra. In fact, VNIR spectra alone produced better predictions for OC than any other combination of sensors. Combining EC_a did not provide any useful improvements and this was true for sensor combinations with or without VNIR spectra. From these results three general conclusions can be drawn: 1) EC_a was not useful for combined sensor estimations of OC, clay, $pH_{1:1}$, and $EC_{1:1}$ at specific depths in the profile, 2) clay and OC are well estimated by models including VNIR diffuse reflectance spectra, 3) for estimating any of the target soil variables, information content of VNIR spectra overwhelms that in EC_p, CI, and EC_a sensors, and 4) models that included other independent variables in addition to VNIR slightly degraded results for all soil properties, except $EC_{1:1}$.

Sensor Combinations with Limited Spectral Features

The dominance of the VNIR measurements in PLSR estimates of soil properties might be explained by the larger amount of information in VNIR diffuse reflectance spectra (430 variables) as compared to the smaller amount of information provided by EC_p, CI, and EC_a measurements (4 variables). The information from these sensors may simply have had a poor signal to noise ratio relative to VNIR spectra, but smaller relative information content might be more important when used in combination with a reduced range reflectance sensor. Variable selection, modeling on the reduced variable space, and model comparison were used to test if EC_p, CI, and EC_a are more important in a reduced variable context. Results from the reduced wavelength models were compared to

results from models including the full VNIR spectral dataset to determine if reduced wavelength VNIR diffuse reflectance sensor would be as good as a full range sensor.

The VIP statistic provided a variable selection criterion. The statistic was calculated from the first PLSR component of the full range VNIR models of OC, clay, $pH_{1:1}$, and $EC_{1:1}$ - without combining EC_p, CI, or EC_a sensor data. Local maxima of VIP greater than or near 1 were used to identify four important wavelengths for each target variable (fig. 6). The VIP results also aided the identification of important spectral features and interpretation of potential physical causes. For instance, OC estimates were influenced by a broad range of mostly visible wavelengths (425 to 1000 nm). These wavelengths were negatively correlated with OC (table 2) indicating that soil darkness (melanization) in general was an important property for estimating OC. Wavelengths above 1200 nm were unimportant and, at least for these soils, a sensor optimized for estimation of OC might be made to capture only the 425 to 1000 nm waveband, probably at reduced expense.

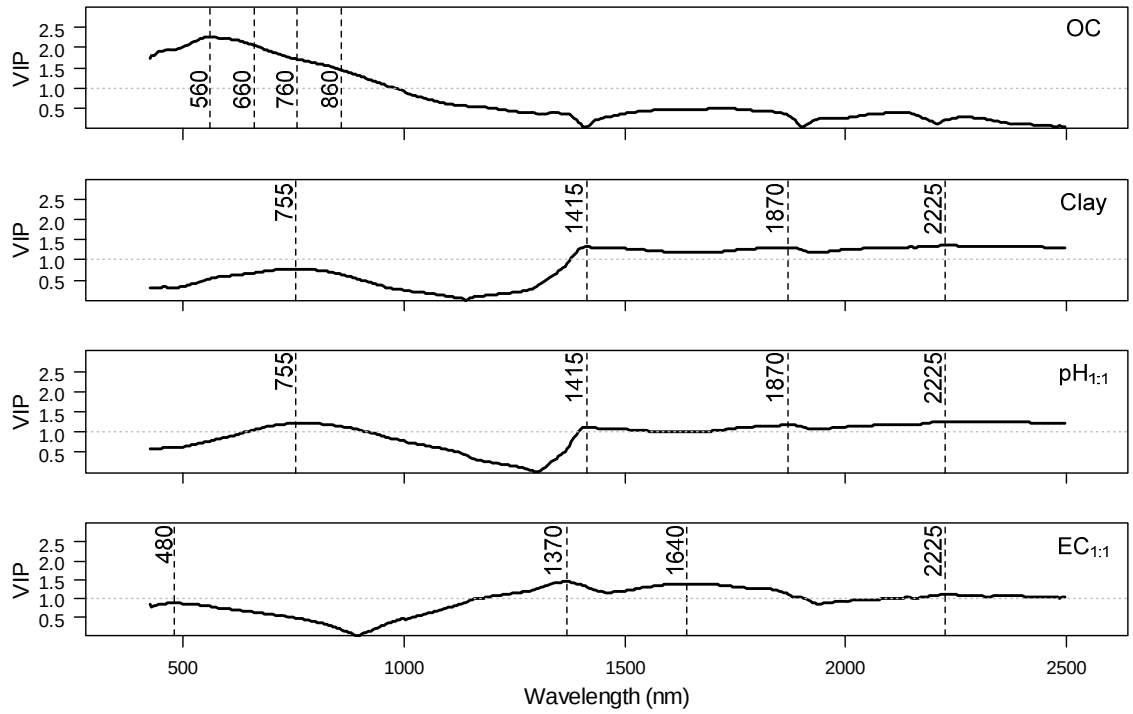

Figure 6. Variable importance to projection (VIP) of the first PLSR component for visible to near infrared (VNIR) diffuse reflectance. Panels (top to bottom) show VIP results for organic carbon (OC), clay content (clay), and 1:1 (soil:water) suspension pH and electrical conducitvity ($pH_{1:1}$ and $EC_{1:1}$). Regression models included only VNIR diffuse reflectance.

Clay and $pH_{1:1}$ regression models had similar VIP results, identifying the 755, 1415, 1870, and 2225 nm wavelengths. Water and X-OH absorption features around 1400, 1900, and 2200 nm were the likely physical bases. The common clay-minerals in these soils are smectite and hydroxy-interlayer illite-smectites. The crystalline structure of these layer silicates includes a large proportion of X-OH bonds with their large surface area and interstitial spaces contain significant water. Gravimetric water content of the argillic horizons (20-30 %) at the time of VNIR diffuse reflectance measurements were greater than all other horizons, and water content was positively correlated to clay content (r = 0.54).

As mentioned above, clay and pH have a large inverse correlation in these soils and the same wavelengths that influenced clay PLSR models also influenced pH models. However, the NIR water absorption features were secondary in importance to the broad VIP peak centered at 755 nm. Soil pH was negatively correlated to reflectance in the visible-red (620 - 780 nm) range covered by this VIP peak (table 2). Bright chroma mottles of 2.5 and 5 YR (reddish) hues were observed in the upper portion of the claypan from a large proportion of the locations. These mottles, due to Fe oxides such as goethite and hematite, coincided with maximum profile acidity (pH ~4.0 to 4.5) at the claypan boundary (Myers et al., 2007); this may be informing the relationship. The color of the mottles at the top of the claypan may also explain the relative maximum in VIP at 755 nm for PLS regressions of clay. Another likely contributing factor to the inverse behavior of pH and clay regression models is the neutralizeable acidity (H^+ ions) bound to cation exchange sites of clay minerals. An indirect relationship of $pH_{1:1}$ to diffuse reflectance (driven by the exchangeable acidity of clay minerals) may occur through spectral response to clay and clay-bound water. Additionally, darker surface soils with small clay content and which have larger pH (6.0-7.0) are potentially driving this relationship by their decreased reflectance in the visible range.

Reflectance at 480, 1370, 1640, and 2225 nm were most important for estimates of $EC_{1:1}$. These wavelengths and the VIP structure of $EC_{1:1}$ in general are similar to that for clay. A likely explanation is another indirect calibration through clay's contribution to the solid-liquid coupled phase pathway of electrical conductance (Corwin and Lesch, 2005). The coupled phase EC pathway may be

enhanced by the large surface area and interstitial water of clay in the soil:water suspension where $EC_{1:1}$ was measured. There are, however, differences in the VIP statistics of clay and $EC_{1:1}$ between 480 and 1370 nm. The PLSR procedure may simply not have found correct ordinations in the spectral data due to the noise in the independent variable. This same noise carries over to the VIP calculation as there is only a small proportion of variation explained in $EC_{1:1}$ to be partitioned into VIP for each variable.

Once the important spectral variables were identified, their VIP was re-examined in a reduced variable context. The selected wavelengths were combined with the EC_p, CI and EC_a sensor measurements in PLS regressions of the soil properties. The VIP statistic was re-calculated from these reduced variable models. Results (fig. 8) showed that even by using fewer wavelengths, the diffuse reflectance variables were more informative. The EC_a variables in particular did not reach the VIP threshold of 0.5 in most cases. This again indicates that proximal EC_a was relatively unimportant to estimates the target soil variables at any given depth in the profile.

Cone index and EC_p were mixed in their importance to the reduced variable regression models of the measured soil properties. They were important in tandem for OC, but unimportant in tandem for clay. This is in contrast to the considerable literature demonstrating relationships between CI and clay and highlights the redundancy between CI and VNIR spectral features (e.g., water absorption bands). Estimates of $EC_{1:1}$ were influenced by CI, but not by EC_p. The

opposite was true for $pH_{1:1}$. EC_p was the most important variable among all variables in the reduced sensor context, VNIR included.

The importance of EC_p and CI for estimating OC is likely due to the covariance of OC (fig. 3), CI, and EC_p with depth in these profiles (fig. 7). Organic carbon generally decreased with depth in these soils except for a secondary relative maximum in OC corresponding to increased root length density previously reported in argillic horizons (Myers et al., 2007). Beneath argillic horizons, root length density and OC decrease rapidly, and over all depths OC showed a moderate negative correlation with CI and EC_p ($r = -0.58$ for both, table 2). Profile CI increased with depth at all sites without exception, and specific features and parent materials such as pedisediment and till could be identified. Likewise, other than a relative minimum in E horizons, EC_p increased with depth for nearly all sampled locations. The same parent materials were also identifiable on EC_p profiles. Smallest OC (0.1-0.3 %) generally occurred 75 to 120 cm deep in the profile coinciding with silty (60-70 %) pedisediment of firm to very firm moist consistence. These materials exhibited the largest cone index (1.5-2 MPa) and EC_p (40-80 mS m^{-1}) values throughout the study area and are probably limiting root penetration and OC accumulation at these depths.

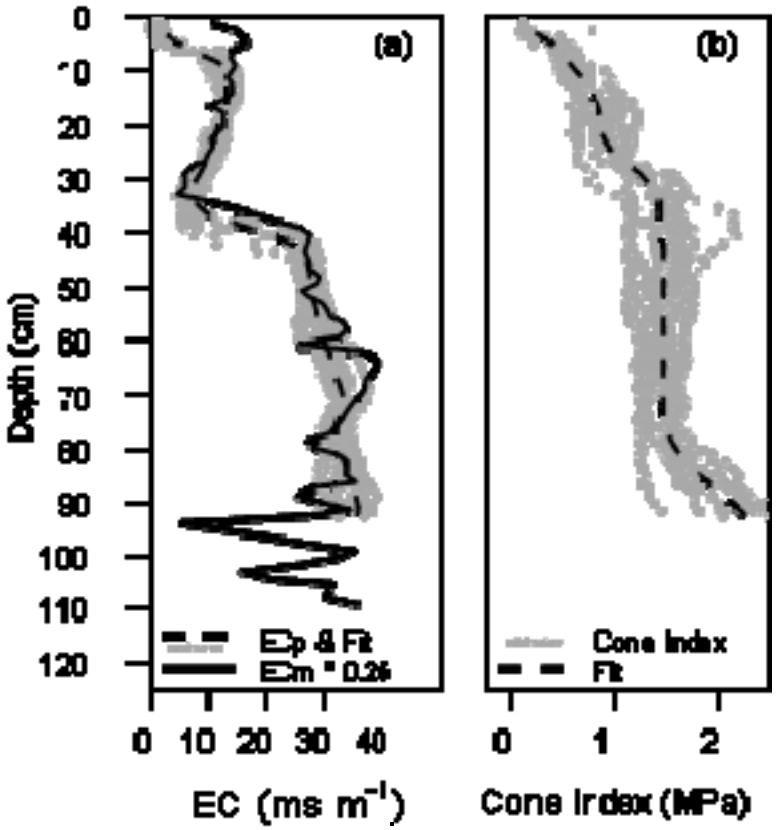

Figure 7. Penetrometer and Wenner-miniprobe measurements of (a) ECp and ECm, and (b) cone index from a representative claypan soil profile.

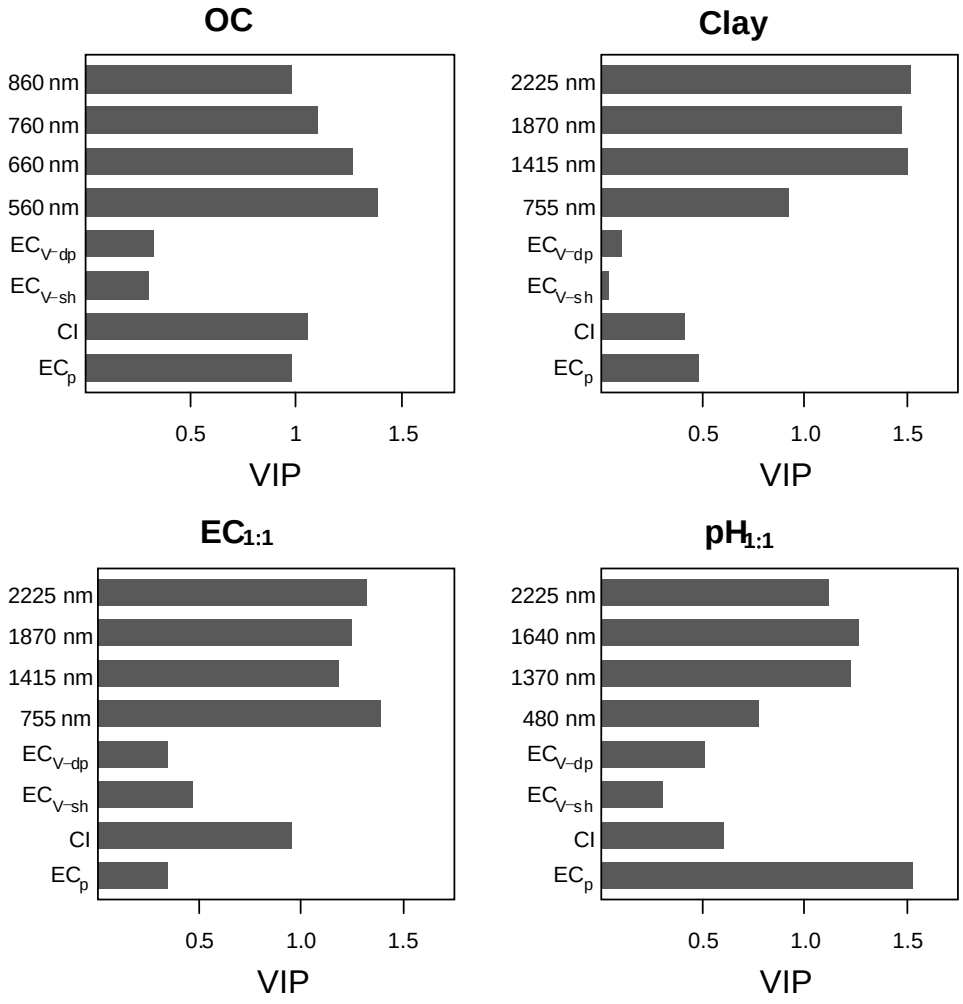

Figure 8. Barplots showing the variable importance to projection (VIP) for PLSR models with EC_p, CI, $EC_{aV\text{-}sh}$, $EC_{aV\text{-}sh}$, and selected VNIR wavelengths. Relative to the other sensor variables listed, VNIR wavelengths rank very highly.

The importance of CI for $EC_{1:1}$ estimates is not as straightforward in light of the small positive correlation of CI and $EC_{1:1}$ (r = 0.27, table 2). It is especially interesting that the large positive correlation of $EC_{1:1}$ and EC_p (r = 0.62) did not result in a large VIP for EC_p in the combined sensor model. The most likely explanation is that spectral variables carry the same information as EC_p, while CI contains information orthogonal to both. That orthogonal information could be

the contrasting response of CI and EC_p in E horizons. Silty E horizons had increasing CI, but relative minima in EC_p and $EC_{1:1}$ occurred there also. Increasing CI occurs in E horizons because their high silt content and platy structure resists cone insertion. Minimum $EC_{1:1}$ occurs there due to the larger proportion of residual quartz and other resistant primary minerals with low surface area and low particle conductivity.

The PLS regression of the reduced wavelength model for $pH_{1:1}$ resulted in very large VIP for EC_p, even larger than spectral variables. This is strongly contrasted by the results from the other soil properties which had larger VIP for spectral variables. The correlation between EC_p and pH is small ($r = -0.19$), but this is due to a non-linear structure in their relationship. The nonlinear structure is caused by E horizon properties related to the discussion above. Measurements of EC_p and pH decrease to a minimum in the E horizon, while in and below the claypan they both increase. Regardless of the success of EC_p in influencing PLS regressions of $pH_{1:1}$, diffuse reflectance measurements were still dominant in the reduced wavelength context overall. Examining the VIP results confirms the superiority of VNIR diffuse reflectance over the other sensor variables for estimation of the soil properties measured.

The potential of a reduced wavelength sensor was investigated by comparison of reduced VNIR spectra models and the full VNIR spectra models. Both cases included all of the other sensor variables. Results were slightly worse or unchanged for the reduced spectrum approach (table 4). Organic carbon in particular showed a 0.1 reduction in R^2 and the largest proportional increase in

RMSEC. The reduced model results for clay and $pH_{1:1}$ were only slightly degraded, while the reduced model for $EC_{1:1}$ was essentially unchanged. These minor decreases in estimation precision suggest that a sensor with specific tailored sensitivity may useful for these soils. However, the difference between the spectral variables important to OC (visible) versus clay, $EC_{1:1}$ and $pH_{1:1}$ (near infrared) means that a reduced wavelength sensor still needs to cover a wide range of important wavelengths to be effective for multiple soil properties.

Table 4. Comparison of reduced spectra PLSR models to full VNIR spectra models. Both the reduced and the full spectrum models included EC_p, CI, and a deep and shallow EC_a channel.

	Reduced spectra model			Full spectra model		
	n-comp	R^2	RMSEC	n-comp	R^2	RMSEC
OC	6	0.69	0.20	11	0.80	0.16
Clay	4	0.84	4.38	10	0.86	4.09
$pH_{1:1}$	6	0.56	0.51	6	0.61	0.48
$EC_{1:1}$	3	0.47	58.62	6	0.47	58.99

Best Models: Organic Carbon and Clay

These repeated model comparisons with various field deployable combinations of soil sensors used a cross-validation approach that was adequate for the model testing and selection process. Application of combined sensor systems in the field requires a more rigorous calibration and model testing approach with independent datasets. The best sensor combination identified above for each measured soil variable was modeled using this more rigorous strategy. Results are shown in scatterplots (fig. 9) of the model estimates from the location-wise independent test datasets. Fit statistics are provided in table (table 5) below as compared to the model selection cross-validation results.

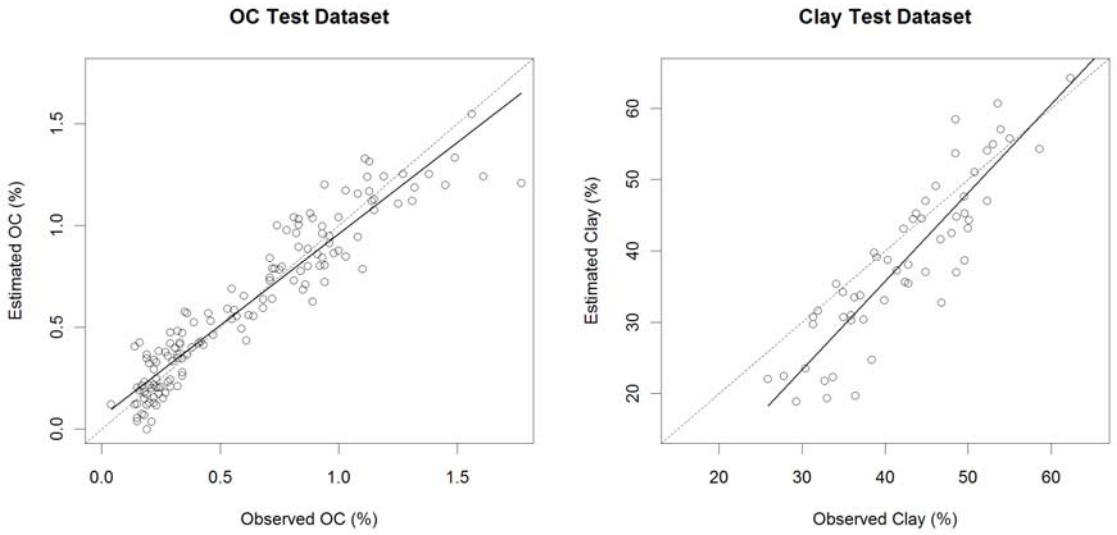

Figure 9. Best sensor combination PLSR models for organic carbon (OC) and clay content applied to the independent test dataset. The model used for OC included only VNIR diffuse reflectance spectra while the model for clay included EC_p, CI, and EC_a.

Table 5. Best sensor combination PLSR regression statistics. The coefficient of determination (R^2) shown corresponds to the linear fit of the cross-validation and independent datasets respectively. Root mean squared errors correspond to the cross-validation and test data estimate (RMSEC and RMSEE). RPD = standard deviation of the observed / RMSE.

		Crossvalidation			Independent Test Data		
	Model	**R^2**	**RMSEC**	**RPD**	**R^2**	**RMSEE**	**RPD**
OC	VNIR	0.90	0.15	2.37	0.89	0.13	3.03
Clay	EC_p + CI + VNIR+EC_a	0.91	4.09	2.56	0.82	6.60	1.33

Independent test dataset results compared to the original cross-validated model selection results indicate good repeatability in the PLSR modeling procedure and a minimum of overfitting. Organic carbon estimates actually improved slightly in regards to the RMSE while R^2 stayed relatively unchanged. This improvement in RMSE and concomitant improvement in RPD is likely due to the smaller sample size (~1/2) in the independent test dataset, containing fewer outliers. Clay estimates from the test dataset on the other hand are

somewhat worse than seen in the full dataset model selection procedure. The R^2 of the linear fit for clay is well above the 0.7 threshold, however RPD dips to 1.33. This may be due to the smaller sample size as well, but in the case of clay, fewer data points may leave the PLSR procedure unable to fully explore the spectral data space for orthogonal vectors. Only 10 calibration sites (n = 90) and 8 validation sites (n = 72) were available for clay as compared to the 16 and 16 (n = 233 and 245) respectively for OC.

Sensor Co-Location Issues

The EC_p measurements were taken as smooth local regression models of five penetrometer profiles measured about 15 cm apart and (collectively) as much as one meter away from the coring site. Thus, the soil measurements taken by the penetrometer and those measured by the diffuse reflectance sensor occurred on different soil. The mini-probe EC sensor was used to better emulate a combined EC_p and VNIR diffuse reflectance system with synchronized measurement site. A comparison was made between the performance of EC_m to predict the measured soil properties versus that of EC_p measured in adjacent soil. This was accomplished by two PLSR models for each measured soil property (S_i):

A. $S_i = EC_m + VNIR$

B. $S_i = EC_p + VNIR$

Measurements of EC_m were better correlated with clay, $pH_{1:1}$, and $EC_{1:1}$ than were EC_p measurements (r = 0.51 to 0.43, -0.27 to -0.15, and 0.70 to 0.59 respectively, table 2). No improvement in correlation was seen for OC. However,

regression models including EC_m as an independent variable showed small improvements in RMSE and R^2 over models with EC_p for OC, clay, and $EC_{1:1}$ (fig. 10). The co-location of the EC_m measurement with the soil property and measurements is probably the reason for this effect.

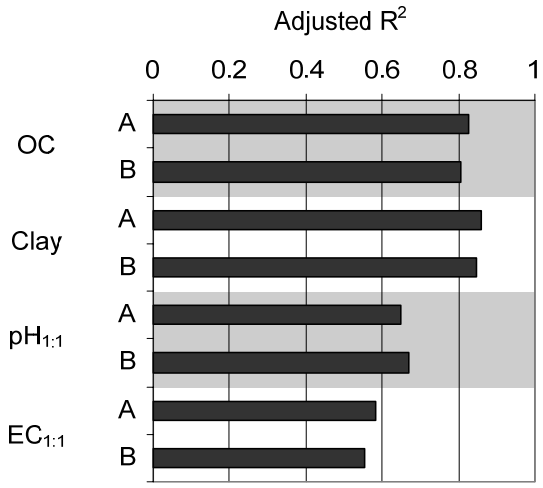

Figure 10. Comparison of models with EC_m (A) against those with EC_p (B).

Significant variation can occur between and within soil cores in this landscape. The three cores in Figure 11were taken within 1 meter radius of the same field location. Considerable variation occurs in the distribution of redoximorphic features. Additionally, large variation in EC_m occurred at small depth increments on soil cores, especially below the claypan (fig. 7a). Usually this was due to the difference in conductivity between the gleyed matrix of the subsoil (10YR 6/1), common large prominent yellow and orange mottles (10YR 6/8 and 7.5YR 5/8), and common large prominent dark grey mottles (10YR 3/1). These mottles were Fe and Fe-Mn oxide stains and concentrations that had reduced conductivity compared to the adjacent gleyed matrix. A combined EC_p and VNIR

diffuse reflectance sensor should be designed to synchronize the depth of these measurements and prevent this subtle source of variation.

Figure 11. Three claypan soil cores sampled < 1m apart. Black rectangle on inset shows location of the zoomed image. Considerable variation in redoximorphic features occurs within a very short distance indicating the importance of making localized sensor measurements.

CONCLUSIONS

Sensor combinations can be expected to improve estimates of a soil property so long as there is unique (orthogonal) information between their responses to the soil property. Procedures such as PLSR which orthogonalize the variable space should be able to capitalize on the unique information between sensors. The sensor combinations studied here are all individually responsive to the claypan soil-landscape, but when combined, did not provide a large improvement in soil property estimates over VNIR diffuse reflectance alone. This

was likely because clay, which controls most soil property distributions in claypan soils, has physical characteristics that produce a similar response profile in CI, EC_p, EC_m, and VNIR sensors. Because they are so similar, little orthogonality is present between these sensor response profiles, and VNIR diffuse reflectance spectra themselves contain the same orthogonal information supplied by CI and EC_p.

Proximal EC_a sensors contributed very little or no improvement in soil property estimates within the soil profile, but their spatial coverage and proven relationship to soil morphology is not diminished. A linkage between penetrating sensor estimates (as developed here) and spatial mapping provided by EC_a may still be accomplished by using EC_a to estimate depth functions of soil properties instead of estimating soil properties directly. This hybrid hierarchical approach would use penetrating sensors to estimate depth distribution of soil properties which can then be fitted to parametric depth functions. The parameters of the depth functions (Chapter 1, this dissertation) might then be mapped across the landscape by their relationships to EC_a. This hybrid sensor and depth function would require additional study.

Deployment of penetrating sensors will require careful control in localizing multiple sensor measurements and careful sampling of measurement sites for calibration development. Soil is heterogeneous within even a few centimeters while the depth resolution of penetrometer sensors is potentially less than 1 cm. Soil property estimates were improved by carful localization of EC_m and VNIR diffuse reflectance. The ability of soil sensors to estimate soil

properties may be diminished if parity is not achieved between the measured soil property and location of sensor measurement. Further sensor development should consider this problem.

As seen in previous research in claypan soils and around the world, VNIR diffuse reflectance was successful at estimating clay and OC, and based on this research will likely be successful *in-situ* for claypan landscapes. The precision and accuracy of the estimates derived here for clay and OC are particularly useful given the depth resolution obtainable by the penetrometer method, and the fact that many spatial locations can be visited. These are important results because soil property estimates with greater depth and spatial resolution are needed as inputs for high resolution land use and soil management decision making.

REFERENCES

Abdu, H., Robinson, D.A., and Jones, S.B., 2007. Comparing bulk soil electrical conductivity determination using the DUALEM-1S and EM38-DD electromagnetic induction instruments. Soil Science Society of America Jounal, 71(1): 189.

Adamchuk, V.I., Hummel, J.W., Morgan, M.T., and Upadhyaya, S.K., 2004. On-the-go soil sensors for precision agriculture. Computers and Electronics in Agriculture, 44(1): 71.

Ahuja, L.R., Ma, L. and Timlin, D.J., 2006. Trans-disciplinary soil physics research critical to synthesis and modeling of agricultural systems. Soil Science Society of America Journal, 70(2): 311.

Andrade-Sanchez, P., Upadhyaya, S.K., and Jenkins, B.M., 2007. Development, construction, and field evaluation of a soil compaction profile sensor. Transactions of the ASABE, 50(3): 719.

Atekwana, E.A., Sauck, W.A., and Werkema, D.D., 2000. Investigations of geoelectrical signatures at a hydrocarbon contaminated site. Journal of Applied Geophysics, 44(2-3): 167.

Ayers, P.D. and Perumpral, J.V., 1982. Moisture and density effect on cone index. Transactions of the ASAE, 24: 1169.

Beck, F.P., Clark, P.J., and Puls, R.W., 2000. Location and characterization of subsurface anomalies using a soil conductivity probe. Ground Water Monitoring and Remediation, 20(2): 55.

Ben-Dor, E. and Banin, A., 1995. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. Soil Science Society of America Journal, 59(2): 364.

Benson, A.K., Payne, K.L., and Stubben, M.A., 1997. Mapping groundwater contamination using DC resistivity and VLF geophysical methods--a case study. Geophysics, 62(1): 80.

Bramley, R.G.V. and Janik, L.J., 2005. Precision agriculture demands a new approach to soil and plant sampling and analysis: examples from Australia. Communications in Soil Science and Plant Analysis, 36(1): 9.

Brown, D.J., Shepherd, K.D., Walsh, M.G., Mays, M. D., and Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma, 132(3-4): 273.

Brus, D.J., Knotters, M., van Dooremolen, W.A., van Kernebeek, P., and van Seeters, R.J.M., 1992. The use of electromagnetic measurements of apparent soil electrical conductivity to predict the boulder clay depth. Geoderma, 55(1-2): 79.

Butler, J.J., Healey, J.M., McCall, G.W., Garnett, E.J., and Loheide, S.P., 2002. Hydraulic tests with direct-push equipment. Ground Water, 40(1): 25.

Chang, C.-W., Laird, D.A., Mausbach, M.J., and Hurburgh, C.R., 2001. Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. Soil Science Society of America Journal, 65(2): 480.

Chong, I.-G. and Jun, C.-H., 2005. Performance of some variable selection methods when multicollinearity is present. Chemometrics and Intelligent Laboratory Systems, 78(1-2): 103.

Christy, C.D., 2008. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. Computers and Electronics in Agriculture, 61(1): 10.

Chung, S.O., Sudduth, K.A., Plouffe, C., and Kitchen, N.R., 2008. Soil bin and field tests of an on-the-go soil strength profile sensor. Transactions of the ASABE, 51(1): 5.

Cleveland, W.S. and Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting. Journal of the American Statistical Association, 83(403): 596.

Corwin, D.L. and Lesch, S.M., 2005. Characterizing soil spatial variability with apparent soil electrical conductivity I. Survey protocols. Computers and Electronics in Agriculture, 46(1-3): 103.

Dobos, E., Carre, F., Hengl, T., Reuter, H.I., and Toth, G., 2006. Digital Soil Mapping as a support to production of functional maps. Office for Official Publications of the European Communities, Luxemburg, pp. 68.

Doolittle, J.A., Sudduth, K.A., Kitchen, N.R., and Indorante, S.J., 1994. Estimating depths to claypans using electromagnetic induction methods. Journal of Soil and Water Conservation, 49(6): 572.

Drummond, P., Christy, C., and Lund, E., 2000. Using an automated penetrometer and soil EC probe to characterize the rooting zone. In: P. Robert (Editor), 5th International Conference on Precision Agriculture. American Society of Agronomy, Minneapolis, MN, USA.

Evans, S.D., Lindstrom, M.J., Voorhees, W.B., Moncrief, J.F., and Nelson, G.A., 1996. Effect of subsoiling and subsequent tillage on soil bulk density, soil moisture, and corn yield. Soil and Tillage Research, 38(1-2): 35.

Filzmoser, P., Maronna, R. and Werner, M., 2008. Outlier identification in high dimensions. Computational Statistics and Data Analysis, 52(3): 1694.

Follmer, L.R., 1983. Sangamon and Wisconsinan pedogenesis in the Midwestern United States. In: S.C. Porter (Editor), The Late Pleistocene. Late Quaternary Environments of the United States. University of Minnesota Press, Minneapolis, pp. 138.

Gessler, P.E., Chadwick, O.A., Chamran, F., Althouse, L., and Holmes, K., 2000. Modeling soil-landscape and ecosystem properties using terrain attributes. Soil Science Society of America Journal, 64(6): 2046.

Grunwald, S., Lowery, B., Rooney, D.J., and McSweeney, K., 2001. Profile cone penetrometer data used to distinguish between soil materials. Soil and Tillage Research, 62(1-2): 27.

Hummel, J.W., Ahmad, I.S., Newman, S.C., Sudduth, K.A., and Drummond, S.T., 2004. Simultaneous soil moisture and cone index measurement. Transactions of the ASAE, 47(3): 607.

Jaynes, D.B. and Colvin, T.S., 1997. Spatiotemporal variability of corn and soybean yield. Agronomy Journal, 89(1): 30.

Jiang, P., Anderson, S.H., Kitchen, N.R., Sudduth, K.A., and Sadler, E.J., 2007. Estimating plant-available water capacity for claypan landscapes using apparent electrical conductivity. Soil Science Society of America Journal, 71(6): 1902.

Kaspar, T.C. et al., 2003. Relationship between six years of corn yields and terrain attributes. Precision Agriculture, 4(1): 87.

Kitchen, N.R., Sudduth, K.A., and Drummond, S.T., 1999. Soil electrical conductivity as a crop productivity measure for claypan soils. Journal of Production Agriculture, 12(4): 607.

Lerch, R.N. et al., 2005. Development of a conservation-oriented precision agriculture system: water and soil quality assessment. Journal of Soil and Water Conservation, 60(6): 411.

Lund, E.D., Christy, C.D., and Drummond, P.E., 1999. Practical applications of soil electrical conductivity mapping. Precision agriculture, 99: 11â€“15.

Lund, E.D., Christy, C.D., and Drummond, P.E., 2000. Using yield and soil electrical conductivity (EC) maps to derive crop production performance information. Proc. 5th International Conference on Precision Agriculture.

McBratney, A.B., 1992. On variation, uncertainty and informatics in environmental soil-management. Australian Journal of Soil Research, 30(6): 913.

McNeill, J.D., 1980. Electromagnetic terrain conductivity measurement at low induction numbers, Geonics Ltd., Missisauga, Ontario.

McSweeney, K. et al., 1994. Towards a new framework for modeling the soil-landscape continuum. Soil Science Society of America Special Publication, 33: 127.

Mevik, B.H. and Wehrens, R., 2007. The pls package: principal component and partial least squares regression in R. Journal of Statistical Software, 18(2).

Mouazen, A.M., Dumont, K., Maertens, K., and Ramon, H., 2003. Two-dimensional prediction of spatial variation in topsoil compaction of a sandy loam field-based on measured horizontal force of compaction sensor, cutting depth and moisture content. Soil and Tillage Research, 74(1): 91.

Myers, D.B., Kitchen, N.R., Sudduth, K.A., Sharp, R.E., and Miles, R.J., 2007. Soybean root distribution related to claypan soil properties and apparent soil electrical conductivity. Crop Science, 47(4): 1498.

Perumpral, J.V., 1987. Cone penetrometer applications-A review. Transactions of the ASAE, 30(4): 939.

Post, W.M. and Kwon, K.C., 2000. Soil carbon sequestration and land-use change: processes and potential. Global Change Biology, 6(3): 317.

Raper, R.L., Washington, B.H., and Jarrell, J.D., 1999. A tractor-mounted multiple-probe soil cone penetrometer. Applied Engineering in Agriculture, 15(4): 287.

Reedy, R.C. and Scanlon, B.R., 2003. Soil water content monitoring using electromagnetic induction. Journal of Geotechnical and Geoenvironmental Engineering, 129(11): 1028.

Richard, G. et al., 2008. Continuous monitoring of agricultural soil physical conditions using proximal sensors for soil tillage management, First Global Workshop on High Resolution Digital Soil Sensing and Mapping. University of Sydney, Sydney, Australia, Unpaginated.

Richards, S.J., 1941. A soil penetrometer. Soil Science Society of America Proceedings, 6: 164.

Rooney, D.J. and Lowery, B., 2000. A profile cone penetrometer for mapping soil horizons. Soil Science Society of America Journal, 64(6): 2136.

Ruhe, R.V., 1969. Quaternary Landscapes in Iowa. University of Iowa Press, Ames, IA. pp. 255.

Schulmeister, M.K. et al., 2003. Direct-push electrical conductivity logging for high-resolution hydrostratigraphic characterization. Ground Water Monitoring and Remediation, 23(3): 52.

Shepherd, K.D. and Walsh, M.G., 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Science Society of America Journal, 66(3): 988.

Sudduth, K.A. and Hummel, J.W., 1991. Evaluation of reflectance methods for soil organic matter sensing. Transactions of the ASAE, 34(4): 1900

Sudduth, K.A. et al., 2005. Relating apparent electrical conductivity to soil properties across the north-central USA. Computers and Electronics in Agriculture, 46(1-3): 263.

Sun, Y., Ma, D., Schulze Lammers, P., Schmittmann, O., and Rose, M., 2006. On-the-go measurement of soil water content and mechanical resistance by a combined horizontal penetrometer. Soil and Tillage Research, 86(2): 209.

Taylor, H.M. and Gardner, H.R., 1963. Penetration of cotton seedling taproots as influenced by bulk density, moisture content, and strength of soil. Soil Science, 96(3): 153.

Team, R.D.C., 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Thompson, P.J., Jansen, I.J., and Hooks, C.L., 1987. Penetrometer resistance and bulk density as parameters for predicting root system performance in mine soils. Soil Science Society of America Journal, 51(5): 1288.

Udawatta, R.P., Krstansky, J.J., Henderson, G.S., and Garrett, H.E., 2002. Agroforestry practices, runoff, and nutrient loss: a paired watershed comparison. Journal of Environmental Quality, 31(4): 1214.

Vaz, C.M.P., Bassoi, L.H., and Hopmans, J.W., 2001. Contribution of water content and bulk density to field soil penetration resistance as measured by a combined cone penetrometer-TDR probe. Soil and Tillage Research, 60(1-2): 35.

Viscara-Rossel, R.A., Jeon, Y.S., Odeh, I.O.A., and McBratney, A.B., 2008. Using a legacy soil sample to develop a mid-IR spectral library. Australian Journal of Soil Research, 46(1): 1.

Viscarra Rossel, R.A. and McBratney, A.B., 1998. Laboratory evaluation of a proximal sensing technique for simultaneous measurement of soil clay and water content. Geoderma, 85(1): 19.

Waiser, T.H., Morgan, C.L.S., Brown, D.J., and Hallmark, C.T., 2007. *In situ* characterization of soil clay content with visible near-infrared diffuse reflectance spectroscopy. Soil Science Society of America Journal, 71(2): 389.

Weller, U., Zipprich, M., Sommer, M., Castell, W.Z., and Wehrhan, M., 2007. Mapping clay content across boundaries at the landscape scale with electromagnetic induction. Soil Science Society of America Journal, 71(6): 1740.

Wilson, J.P., 1996. GIS based land surface/subsurface modeling: new potential for new models, Proceedings Third International NCGIA Conference on Integrating GIS and Environmental Modelling. University of New Mexico, Santa Fe, New Mexico, USA., Unpaginated CD.

Wold, S., Sjostrom, M., and Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems, 58(2): 109.

VITA

Brent Myers was born in Bloomfield, IA on October 6th, 1970. He was raised in nearby Memphis, MO where his Mother and Father, David and Sarah Myers owned a farrow to finish confinement hog operation and a grain farm. Much of his youth was spent at the farm exploring the glacially formed landscape and the nature it supported. These experiences and observation of natural resources and food production led quite naturally to an interest in plants and soils. Brent completed a B.S. in Plant Science in 2001 and followed on for an M.S. in Soil Science, finishing in 2005. These studies made clear the role that soil property variation had in natural and agro-ecosystems, and also made clear the need for high resolution soil property data. Brent formulated this Ph.D. program with this in mind.

Brent married Melissa Myers on Aug. 8, 1998 and had a son David on March 8, 2004. They are expecting their next son Luke before this dissertation returns from the printer. Brent took a post-doctoral position at the University of Florida, and he and his family will be moving to Gainesville, FL shortly after completing this work.