Montana Cropland Enrolled Farm Fields Carbon Sequestration Field Sampling, Measurement, Monitoring, and Verification:

Application of Visible-Near Infrared Diffuse Reflectance Spectroscopy (VNIR) and Laser-induced Breakdown Spectroscopy (LIBS)

Final Report

November 2011

U.S. Department of Energy (DOE)
National Energy Technology Laboratory (NETL)
DOE Award Number: DE-FC26-05NT42587

Submitted by: Dr. Lee Spangler

BSCSP Principal Investigator and Director
Montana State University
P.O. Box 173905
Bozeman, MT 59717
spangler@montana.edu
http://www.bigskyco2.org
Phone: (406) 994-4399
Fax: (406) 994-3745

BSCSP Researchers
Ross Bricklemyer
David Brown

Disclaimer

Table of Contents

Executive Summary 6

1. Overview and Organization of Report 8

2. On-the-go VisNIR: Potential and limitations for mapping soil clay and organic carbon 9

    2.1 Introduction 9

    2.2 Materials and Methods 11

    2.3 Results 15

    2.4 Discussion 21

    2.5 Conclusions 23

3. Intact soil core total, inorganic and organic carbon measurement using laser-induced breakdown spectroscopy (LIBS). (Accepted for publication, SSSAJ) 24

    3.1 Introduction 24

    3.2 Material and Methods 27

    3.3 Core Interrogation 28

    3.4 Results 34

    3.5 Discussion 42

    3.6 Conclusions 45

4. Full spectrum laser-induced breakdown spectroscopy (LIBS) for intact soil core total, inorganic and organic carbon measurement. (unpublished data) 47

    4.1 Material and Methods 47

    4.2 Results and Discussion 48

    4.3 Conclusions 51

5. Analysis and Conclusions 53

    5.1 Introduction 53

    5.2 Material and Methods 53

    5.3 Results and Discussion 53

    5.4 Conclusion 55

6. References 57

Figures

Figure 1.Geographical location of the study area with eight (8) selected farm fields. 12

Figure 2. Example of on-the-go north-south oriented transects showing data collection points (open circles) and 100 reference soil sample locations (black crosses). Blank areas within transects were caused by sensor skips and temporary WAAS GPS signal loss. 13

Figure 3: Whole-field out validation of soil organic carbon predicted using partial least squares regression modeling of lab-based (a) and on-the-go VisNIR (b) 18

Figure 4: Hybrid model validation of SOC predicted using PLSR modeling of lab-based (a) & on-the-go VisNIR (b) 18

Figure 5: Whole-field out validation of soil clay content predicted using partial least squares regression modeling of lab-based (a) & on-the-go VisNIR (b) 19

Figure 6: Hybrid model validation of soil clay content predicted using PLSR modeling of lab-based (a) & on-the-go VisNIR (b) 19

Figure 7: Predictive accuracy response to the addition of local samples in hybrid partial least squares regression calibration models for soil clay content & soil organic carbon (SOC) .. 20

Figure 8: A LIBS spectrum (200-300 nm) of a representative soil core. LIBS spectra are typically spectrally rich containing many emission lines for each element in the sample. Some of the emission features associated with the major elements present in the sample are identified. 25

Figure 9: Geographical location of the study area with 3 selected farm fields (A), & randomly selected calibration (triangles) & validation (circles) core locations at the LYD (B), HOR (C), & MAT (D) sites.. 28

Figure 10: LIBS sampling depths on a representative intact soil core & LIBS interrogation point configuration within a depth sample (inset)... 29

Figure 11: Independent validation of predicted soil total carbon (TC), inorganic carbon (IC), & organic carbon (SOC) using LIBS and partial least squares 2 regression models. Nine interrogation spectra were averaged for PLS2 analysis.. 37

Figure 12: Partial least squares regression (PLS2) coefficient values for soil total, inorganic, & organic carbon. The magnitude of the coefficients indicates the relative importance of each emission line. Dashed vertical lines indicate important elemental emission lines for predicting the various forms of soil carbon. All wavelengths were used for predictions.... 38

Figure 13. A simple linear regression of calibration set IC on TC (lab reference measurements) applied to validation set TC: a) measured by total combustion; b) predicted by LIBS PLS2; and c) predicted by LIBS PLS1. 40

Figure 14. Predictive accuracy response, as indicated by the standard error of prediction (SEP), to spectrally averaging multiple LIBS interrogation points for soil total carbon (TC), inorganic carbon (IC), and organic carbon (SOC) determination....................................... 42

Tables

Table 1: Summary Statistics for standard lab mesaured SOC & soil clay content in eight Montana wheatfields 16

Table 2: PLSR model statistics 17

Table 3: Laboratory measured soil total, inorganic & organic carbon summary statistics for 78 intact soil cores from 3 wheat fields in north central MT, 2007 31

Table 4: Soil total carbon (TC), inorganic carbon (IC), and organic carbon (SOC) calibration cross-validation statistics for laser-induced breakdown spectroscopy (LIBS) partial least squares 1 (PLS1) & 2 (PLS2) regression models. 35

Table 5: Soil total carbon (TC), inorganic carblon (IC), and organic carbon (SOC) independent validation statistics for laser-induced breakdown spectroscopy (LIBS) partial least squares 1 (PLS1) & 2 (PLS2) regression models. 36

Table 6: Pearson's correlation coefficients (r) for laboratory measured soil carbon, laser-induced breakdown spectroscopy (LIBS) partial least squares 2 (PLS2) regression coefficients, & PLS2 carbon predictions. 41

Table 7: Prediction statistics from a simple linear regression of calibration set inorganic carbon (IC) on total carbon (TC) using lab reference measurements applied to validation set TC: a) measured by total combustion; b) predicted by laser-induced breakdown spectroscopy (LIBS) partial least squares 2 (PLS2); and c) predicted by LIBS PLS1. 41

Table 8: Laboratory-measured soil total, inorganic, and organic carbon summary statistics for 80 intact soil cores from 7 wheat fields in north central MT & the Cook Agronomy farm near Pullman, WA, 2008. 49

Table 9: Laboratory-measured soil total, inorganic and organic carbon summary statistics for 60 intact soil cores from 6 wheat fields in north central MT, 2009. 50

Table 10: Soil total carbon (TC), inorganic carbon (IC), and organic carbon (SOC) whole-field out independent cross validation statistics for 2008 & 2009 laser-induced breakdown spectroscopy (LIBS) partial least squares 2 (PLS2) regression models. 51

Table 11: Soil total carbon (TC), inorganic carbon (IC), and organic carbon (SOC) independent validation statistics for 2008 and 2009 Visible-near infrared spectroscopy (VisNIR) partial least squares regression models. 54

Table 12: Partial Least Squares model statistics for combined VisNIR-LIBs for predicting total carbon (TC), inorganic carbon (IC), & soil organic carbon (SOC) in intact soil cores. 55

Executive Summary

There is growing need for rapid, accurate, and inexpensive methods to measure, and verify soil organic carbon (SOC) change for national greenhouse gas accounting and the development of a soil carbon trading market. Laboratory based soil characterization typically requires significant soil processing, which is time and resource intensive. This severely limits application for large-region soil characterization. Thus, development of rapid and accurate methods for characterizing soils are needed to map soil properties for precision agriculture applications, improve regional and global soil carbon (C) stock and flux estimates and efficiently map sub-surface metal contamination, among others. The greatest gains for efficient soil characterization will come from collecting soil data *in situ*, thus minimizing soil sample transportation, processing, and lab-based measurement costs. Visible and near-infrared diffuse reflectance spectroscopy (VisNIR) and laser-induced breakdown spectroscopy (LIBS) are two complementary, yet fundamentally different spectroscopic techniques that have the potential to meet this need. These sensors have the potential to be mounted on a soil penetrometer and deployed for rapid soil profile characterization at field and landscape scales. Details of sensor interaction, efficient data management, and appropriate statistical analysis techniques for model calibrations are first needed.

In situ or on-the-go VisNIR spectroscopy has been proposed as a rapid and inexpensive tool for intensively mapping soil texture and organic carbon (SOC). While lab-based VisNIR has been established as a viable technique for estimating various soil properties, few experiments have compared the predictive accuracy of on-the-go and lab-based VisNIR. Eight north central Montana wheat fields were intensively interrogated using on-the-go and lab-based VisNIR. Lab-based spectral data consistently provided more accurate predictions than on-the-go data. However, neither *in situ* nor lab-based spectroscopy yielded even semi-quantitative SOC predictions. There was little SOC variability to explain across the eight fields, and on-the-go VisNIR was not able to capture the subtle SOC variability in these Montana soils. With more variation in soil clay content compared to SOC, both lab and on-the-go VisNIR showed better explanatory power. There are several potential explanations for poor on-the-go predictive accuracy: soil heterogeneity, field moisture, consistent sample presentation, and a difference between the spatial support of on-the-go measurements and soil samples collected for laboratory analyses. Though the current configuration of a commercially available on-the-go VisNIR system allows for rapid field scanning, on-the-go soil processing (i.e. drying, crushing, and sieving) could improve soil carbon predictions.

Laser-induced breakdown spectroscopy (LIBS) is an emerging elemental analysis technology with the potential to provide rapid, accurate and precise analysis of soil constituents, such as carbon, *in situ* across landscapes. The research team evaluated the accuracy of LIBS for measuring soil profile carbon in field-moist, intact soil cores simulating conditions that might be encountered by a probe-mounted LIBS instrument measuring soil profile carbon *in situ*. Over the course of three experiments, more than120 intact soil cores from eight north central Montana wheat fields and the Washington State University (WSU) Cook Agronomy Farm near Pullman, WA were interrogated with LIBS for rapid total carbon (TC), inorganic carbon (IC), and SOC determination. Partial least squares regression models were derived and independently validated at field- and regional scales. Researchers obtained the best LIBS validation predictions for IC followed by TC and SOC. Laser-induced breakdown spectroscopy is fundamentally an elemental analysis technique, yet LIBS PLS2 models appeared to discriminate IC from TC. Regression coefficients from initial models suggested a reliance upon stoichiometric relationships between carbon (247.8 nm) and other elements related to total and inorganic carbon in the soil matrix [Ca (210.2 nm, 211.3 nm, and 220.9 nm), Mg (279.55-280.4 nm, 285.26 nm), and Si (251.6 nm, 288.1 nm)]. Expanding the LIBS spectral range to capture emissions from a broader range of elements related to soil organic matter was explored using two spectrometer systems to improve SOC predictions. Results for increasing the spectral range of LIBS to the full 200-800 nm found modest gains in prediction accuracy for IC, but no gains for predicting TC or SOC. Poor SOC predictions are likely a function of 1) the lack of a consistent/definable molecular composition of SOC, 2) relatively little variation in SOC across field sites, and 3) inorganic carbon constituting the primary form of soil carbon, particularly for Montana soils. Exploration into alternative data reduction and statistical modeling techniques continues in an effort to increase prediction accuracy, model parsimony, and computational efficiency. As research matures for this emerging spectroscopic method, new field-deployable equipment should be developed to exploit the unique ability of LIBS to rapidly characterize soil elemental composition.

VisNIR and LIBS are fundamentally different, yet complementary spectroscopic techniques. The LIBS technique is an elemental analysis method that can quickly determine elemental composition of heterogeneous material, whereas VisNIR is based on the fundamentals of light energy absorption by molecular bond vibrations. The research team attempted to take advantage of these fundamental differences by combining spectral information obtained by VisNIR and LIBS sensors. It was hypothesized that this may provide more accurate, robust, and spatially transferable soil carbon determination than individual sensors currently permit. The 120 cores interrogated with LIBS, as discussed previously, were concurrently scanned with VisNIR. Preliminary exploration of combining VisNIR and LIBS spectra was completed using data in 2007 (78 cores) and PLS regression. Combining LIBS and VisNIR data slightly improved TC predictions, but did not improve IC or SOC predictions compared to individual sensor results. The results suggest that PLS regression is sensitive to large predictor datasets. It is suspected that the 2046 LIBS wavelength predictors dominated the PLS analysis and overwhelmed information found in the 216 VisNIR wavelength predictors. Reduction of LIBS spectra poses a difficult challenge to overcome. Other approaches may be able to use the data at the current resolution to build better predictive spectral models for TC, IC, and SOC, regardless of the number of predictor variables; however alternative multivariate statistical and data mining approaches continue to be tested.

VisNIR and LIBS spectroscopy have the potential to fill the growing need for rapid, accurate, and inexpensive methods to measure, and verify soil organic carbon change. These fundamentally different techniques performed individually mostly as expected; however the combined LIBS-VisNIR data array did not consistently improve predictive accuracies over individual sensors. A field unit with a LIBS-VisNIR array mounted in a soil penetrometer, when fully operational, would allow for rapid soil profile characterization and mapping at field and landscape scales.

1. Overview and Organization of Report

This report summarizes research to study and compare two different spectroscopic techniques for conducting soil characterization: visible and near-infrared diffuse reflectance spectroscopy (VisNIR) and laser-induced breakdown spectroscopy (LIBS).

Section 2 presents a study of VisNIR, and Section 3 presents a study of LIBS. Section 4 presents the results of an updated LIBS study with modified instrumentation. In Section 5, the results of the two techniques are compared and analyzed. Section 5 also discusses research to use the two techniques in combination in order to enhance results.

2. On-the-go VisNIR: Potential and limitations for mapping soil clay and organic carbon

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

Abstract

In situ or on-the-go visible and near infrared (VisNIR) diffuse reflectance spectroscopy has been proposed as a rapid and inexpensive tool for intensively mapping soil texture and organic carbon (SOC). While lab-based VisNIR has been established as a viable technique for estimating various soil properties, few experiments have compared the predictive accuracy of on-the-go and lab-based VisNIR. In this study, eight north central Montana wheat fields were intensively interrogated using on-the-go and lab-based VisNIR. The on-the-go VisNIR system employed a spectrophotometer (350-2224 nm, 8-nm spectral resolution) built into an agricultural shank mounted on a toolbar and pulled behind a tractor. Regional (whole-field out cross-validation) and hybrid (regional model including randomly chosen 'local' calibration samples) spectral models were calibrated using partial least squares regression. Lab-based spectral data consistently provided more accurate predictions than on-the-go data. However, neither *in situ* nor lab-based spectroscopy yielded even semi-quantitative SOC predictions. For hybrid models with 9 local samples included in the calibrations, standard error of prediction (SEP) values were 2.6 and 3.4 g kg^{-1} for lab and on-the-go VisNIR respectively, with σ_{SOC} = 3.2 g kg^{-1}. With an SOC coefficient of variation (CV) = 26.7%, even with a relatively low SEP values, there was little SOC variability to explain. For clay content, hybrid +7 calibrations yielded lab SEP = 53.1 g kg^{-1} and residual product differential (RPD) = 1.8 with on-the-go SEP = 69.4 g kg^{-1} and RPD = 1.4. With more variability (σ_{Clay} = 91.4 g kg^{-1} and CV = 49.6%), both lab and on-the-go VisNIR show better explanatory power. There are a number of potential explanations for degraded on-the-go predictive accuracy: soil heterogeneity, field moisture, consistent sample presentation, and a difference between the spatial support of on-the-go measurements and soil samples collected for laboratory analyses. In terms of predictive accuracy, our results are largely consistent with those previously published by Christy (2008), but on-the-go VisNIR was not able to capture the subtle SOC variability in Montana soils. Though the current configuration of the Veris on-the-go VisNIR system allows for rapid field scanning, on-the-go soil processing (i.e. drying, crushing, and sieving) could improve predictions.

2.1 *Introduction*

Soil properties, including soil organic carbon (SOC) and soil texture, vary spatially across landscapes (McBratney and Pringle, 1997). To cost-effectively capture soil variability, on-the-go visible and near infrared (VisNIR) spectroscopy has been proposed as a rapid and inexpensive method of intensively measuring and mapping SOC, soil texture (i.e. clay content), and other soil properties (Adamchuk et al., 2004; Christy, 2008; Gehl and Rice, 2007). While lab-based VisNIR has been established as a viable technique for the estimation of multiple soil properties (Brown et al., 2006; Morgan et al., 2009; Reeves and McCarty, 2001; Shepherd and Walsh, 2002; Viscarra Rossel et al., 2006; Waiser et al., 2007), and there are a few published studies of in situ VisNIR (Ben-Dor et al., 2008; Morgan et al., 2009; Viscarra Rossel et al., 2009; Waiser et al., 2007), few experiments have evaluated the potential and limitations of VisNIR on-the-go for

SOC or soil clay content (Christy, 2008; Shibusawa et al., 1999; Shonk et al., 1991; Sudduth and Hummel, 1993a; Sudduth and Hummel, 1993b).

Researchers have evaluated lab-based estimation of SOC and soil clay content using VisNIR spectroscopy applied to air-dried, crushed and sieved soil samples for a diverse range of soil materials and calibration-validation designs (Viscarra Rossel et al., 2006). Soil organic carbon content has been estimated using VisNIR spectra with root mean square error (RMSE) ranging from 0.9 g kg^{-1} to 12.7 g kg^{-1}, depending upon soil diversity and validation rigor (Ben-Dor and Banin, 1995; Brown et al., 2006; Islam et al., 2003; McCarty et al., 2002; Morgan et al., 2009; Reeves et al., 1999; Shepherd and Walsh, 2002; Vasques et al., 2008). Similarly, soil clay content has been predicted on prepared soils with RSME = 11 to 95 g kg^{-1} (Ben-Dor and Banin, 1995; Brown et al., 2006; Islam et al., 2003; Janik et al., 1998; Shepherd and Walsh, 2002; Waiser et al., 2007). If these results are screened to only include independently validated calibrations (Brown et al., 2006), the literature suggests that lab-based VisNIR can provide semi-quantitative estimation (RPD = 1.5-2.0; with RPD>2.0 considered quantitative) of SOC and clay content.

On-the-go and static *in situ* VisNIR methods introduce unique challenges for accurate determination of soil properties compared to controlled conditions in the laboratory. Natural soil heterogeneity, macro-aggregation, and field moisture content have been identified as variables that can reduce the predictive accuracy of VisNIR methods (Christy, 2008; Morgan et al., 2009; Sudduth and Hummel, 1993a; Waiser et al., 2007). Sensors moving through the soil can also cause inconsistent soil presentation, smearing, and spectral data that are averaged over some distance traveled, dependent on acquisition time and velocity, which can also degrade accurate VisNIR predictions (Morgan et al., 2009; Sudduth and Hummel, 1993a; Waiser et al., 2007). Both *in situ* and lab-based VisNIR accuracy statistics are affected by soil diversity and validation rigor (Brown et al., 2005; Morgan et al., 2009; Waiser et al., 2007).

Static VisNIR is accomplished by holding a spectrometer fore-optic stationary on a soil face for interrogation, potentially via a penetrometer mounting. This method has been tested on exposed soil faces in pit walls (Ben-Dor et al., 2008; Viscarra Rossel et al., 2009), exterior walls of extracted soil core holes (Ben-Dor et al., 2008), and the exterior walls of intact cores (Morgan et al., 2009; Waiser et al., 2007). Waiser et al. (2007) predicted clay content in field-moist intact soil cores using partial least squares regression (PLSR) models, which were validated using 30% of cores randomly held out (RMSD = 61 g kg^{-1}) and whole-field out cross validation (RMSD = average of 89 g kg^{-1}, 64 - 143 g kg^{-1} across 6 fields). For this study, smearing of the soil surface increased the 30% validation RMSD to 74 g kg^{-1}. As part of the same field campaign, Morgan et al. (2009) predicted SOC in field-moist intact soil cores using partial least squares regression (PLSR) models validated using 30% of cores randomly held out (RMSD = 5.4 g kg^{-1}) and whole-field out cross validation, with RMSD increasing an average of 1.3 g kg^{-1}, -0.4 to 4.5 g kg^{-1} across 6 fields. Smearing had only a minor effect on SOC prediction statistics. Rossel et al. (2009), using a large regional spectral dataset for calibration (1287 laboratory- and 74 field-collected spectra), predicted clay content with RMSE = 7.9% clay. Ben-Dor et al. (2008) predicted SOC[1] with RMSE = 1.2 g kg^{-1}; however, samples used for model validation were not

---

[1] SOC was estimated assuming 58% organic carbon in SOM (Nelson and Sommers, 1982).

completely independent of those used for calibration. Ben-Dor et al. (2008)concluded that the method has good potential for *in situ* soil characterization and requires additional study and validation using independent samples.

On-the-go VisNIR typically involves a spectrophotometer either enclosed within or connected to (via fiber optics) an implement that is inserted into the soil and pulled behind a tractor. Several surface soil properties including surface SOM and SOC have been mapped using on-the-go VisNIR with varying degrees of success. Shonk et al. (1991) used a single wavelength diode (red, 660 nm) to predict soil organic matter (SOM) in real-time with good success ($R^2 > 0.83$, SOM = 1-6%). Sudduth and Hummel (1993a; 1993b) developed an on-the-go NIR spectrometer (1650-2650 nm) system that predicted surface SOC with standard error of prediction (SEP) of 2.3 g kg^{-1} SOC in laboratory tests; however, SEP increased to 5.3 g kg^{-1} in independent field tests. Shibusawa et al. (1999) tested a prototype on-the-go VisNIR (400-1700 nm) instrument; measured SOM was highly correlated to certain NIR wavelengths, however independent validation was not executed. More recently, Christy (2008) reported good predicted accuracy (RMSE = 3.0 g kg^{-1}) for mapping SOC[1] in Kansas (SD = 5.1 g kg^{-1}, SOC = 3.0-26.3 g kg^{-1}) using a shank-mounted on-the-go VisNIR sensor (950-1650 nm) and whole-field cross-validation (n = 8 fields). To the best of our knowledge, there are no studies published in the utility of on-the-go VisNIR for soil clay content estimation.

The primary objective of this study was to compare the predictive accuracy for estimating SOC and soil clay content measured *in situ* with an on-the-go VisNIR sensor versus a lab-based sensor interrogating air-dry, sieved soil samples. Researchers also tested regional (i.e. whole-field out) calibration models versus hybrid regional models (regional models including a few “local” calibration samples), and quantified the change in prediction accuracy of hybrid models with increasing numbers of local samples in the calibration.

2.2 ***Materials and Methods***

2.2.1 Study Area

The study area for this research was in the north central region of Montana, USA (Figure 1). This region is characterized by soils formed in glacial till on gently rolling topography in a frigid, ustic, continental climate. Soils in the study area were not highly weathered and typically were calcareous within 0.5 m of the surface. Aridic intergrades of frigid, ustic, Mollisols, Entisols, and Inceptisols predominated. Cropping systems in the study area generally consisted of reduced tillage, small grain-fallow rotations with a significant acreage managed by direct-seeding (i.e. no-till). All eight sampling sites had a general cropping history of cultivation beginning in the 1920’s progressing to wheat-fallow rotations with multiple tillage operations per year and finally conversion to a direct-seeded wheat-fallow rotation between 2004 and 2005.

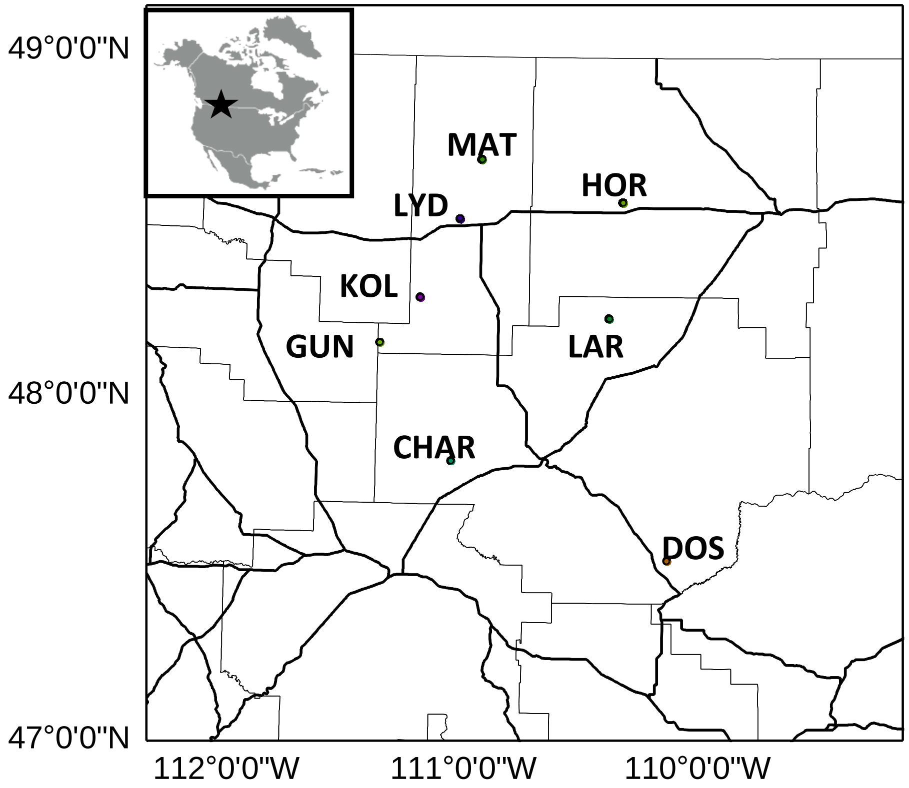

Figure 1.Geographical location of the study area with eight (8) selected farm fields.

2.2.2 Digital and Field Data

The sampling design and data collection for this project was multi-leveled and multi-faceted. Eight farm fields (259 hectares [ha]; 640 acres [ac]) were selected at random from a larger population of farm fields enrolled in a carbon offset marketing pilot. The eight selected fields were each divided into 16.2 ha (40 ac) sub-fields and one sub-field was randomly chosen for VisNIR scanning and soil sampling in the fall of 2006.

The research team intensively scanned each sub-field using a new commercially available on-the-go VisNIR sensor (Veris Technologies Inc., Salina, KS, USA). Fields were scanned approximately a week after rains to allow enough dry down time to permit field access. The Veris system employed a spectrophotometer (350-2224 nanometers [nm], 8-nm spectral resolution) built into an agricultural shank mounted on a toolbar and pulled behind a tractor (Christy, 2008). Spectral measurements were made through a sapphire window mounted on the bottom of the shank with fiber optic cables transferring the diffuse reflected light from the soil to the spectrometer and into a laptop computer for storage. The shank was lowered to approximately 10 cm into the soil and pulled at approximately 5 km/h along 15-m spaced, north-south transects (Figure 2). Approximately 20 spectra were acquired per second and a spectral average was calculated every 2.5-3 s, representing approximately 4 m of travel. Spectral averages were stored on the laptop computer in conjunction with Wide Area Augmentation System (WAAS) real-time corrected GPS data for each point.

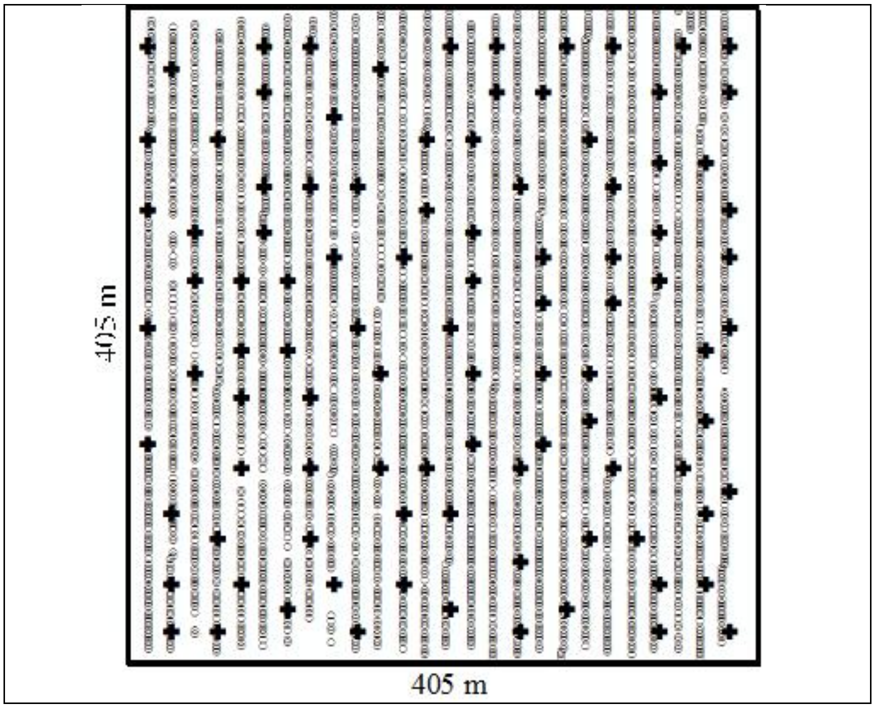

Figure 2. Example of on-the-go north-south oriented transects showing data collection points (open circles) and 100 reference soil sample locations (black crosses). Blank areas within transects were caused by sensor skips and temporary WAAS GPS signal loss.

Soil sampling was designed to capture representative variability within each sub-field. A 15 m grid, following the same transects used for VisNIR scanning, was superimposed over each sub-field and 100 georeferenced grid intersections were chosen randomly with a 30-m spatial inhibition function added to ensure representative coverage of the sub-field (Figure 2). Three surface soil samples (0-10 cm) were taken using a slide hammer and 5 cm diameter coring tube in a 1 m triangular configuration straddling the VisNIR transect at each of the 100 selected intersection points and homogenized for further analysis in the lab. Samples were dried, crushed to pass a 2 mm sieve, and scanned using a lab-based VisNIR spectrometer produced by Analytical Spectral Devices, Inc. (ASD Inc., Boulder CO, USA; FieldSpec Pro, 350-2500 nm, 3 and 10 nm spectral resolutions for the 512 element silicon photo-diode array (350-1025 nm) and InGaS (1025-2500 nm) detectors, respectively). Soil total carbon (TC) was measured by dry combustion (LECO TruSpec, Leco Corp., St. Joseph, MI), soil inorganic carbon (IC) was determined by modified pressure calcimeter method (Sherrod et al., 2002), and soil organic carbon (SOC) was determined by the difference (SOC = TC – IC) (Bricklemyer et al., 2007). Soil clay content was determined by pipette method (Gee and Bauder, 1986). For SOC and clay content, the standard error of laboratory measurement (SEL) was estimated using replicate laboratory measurements as described by Workman (2001).

On-the-go spectral data used for model calibration and validation were extracted using GPS coordinates. Using ArcMap (ESRI, Redlands, CA, USA), on-the-go measurement locations were matched with the nearest georeferenced soil sampling location and the corresponding VisNIR spectral data were extracted for post-processing, partial least squares regression (PLSR) modeling, and accuracy assessment. All spectra (lab-based and on-the-go) were checked for errors and smoothed using cubic smoothing splines with 1st derivatives extracted in 10 nm increments directly from spline fits, following methods outlined in Brown et al. (2006).

2.2.3 Partial Least Squares Modeling

Regional and "hybrid" Partial Least Squares Regression (PLSR) calibration models were derived for both SOC and soil clay content. Regional models were constructed employing a whole-field out cross-validation approach, whereby each field was held out in turn for model validation, with data from the other seven fields used for model calibration. For example, data for field #1 was held out from the calibration set, and data from fields 2-8 were used to calibrate PLSR models to independently predict SOC and clay content for field #1. Then, data from field #2 was held out, and new PLSR models were derived using data from fields 1 and 3-8 to predict SOC and clay content for field #2. The systematic removal of holding out data from whole fields, calibrating new PLSR models using data from the remaining fields, and predicting SOC and clay content for the held out field continued until each of the eight fields had been predicted. For hybrid models, researchers included 1, 2, 3, 4, 5, 7, and 9 randomly selected "local" samples from the held-out validation field in the calibration set, with the same 8-fold cross-validation regional modeling approach. To test for potential loss in prediction accuracy due to differences in the spectral ranges of the two sensors (lab-based = 350-2500 nm, on-the-go = 350-2200 nm), regional and hybrid PLSR calibration models were also derived using lab-acquired spectra reduced to the same spectral range as the on-the-go spectrometer.

The quality of PLSR model fit was evaluated using performance statistics derived from the regression of PLSR predicted vs. laboratory measurements (Brown et al., 2005; Gauch et al., 2003):

$$MSD = \sum_{n} \left(Y_{pred} - Y_{lab}\right)^2 \Big/ N \qquad (1)$$

$$Bias = \sum_{n} \left(Y_{pred} - Y_{lab}\right) \Big/ N \qquad (2)$$

$$\mathrm{SB} = \mathrm{Bias}^2 \qquad (3)$$

$$\mathrm{NU} = (1-b)^2 \times \mathrm{var}(Y_{lab}) \qquad (4)$$

$$\mathrm{LC} = (1-r^2) \times \mathrm{var}(Y_{pred}) \qquad (5)$$

$$SEP = \sqrt{\sum_{n} \left(Y_{pred} - Y_{lab}\right)^2 \Big/ (N-1)} \qquad (6)$$

$$SEP_{adj} = \sqrt{SEP^2 - SEL^2} \qquad (7)$$

$$RPD = {}^{\sigma_{Y-valid}}\!\big/_{SEP_{adj}} \qquad (8)$$

$$RPL = {}^{\left(SEP_{adj} - SEL\right)}\!\big/_{SEL} \qquad (9)$$

where b and r^2 are the slope and square correlation, respectively, from the least-squares regression of Y_{pred} on Y_{lab}, SEP = standard error of prediction, RPD = residual product differential. Lack of accuracy is described by portioning mean squared deviation (MSD) into three components: squared bias (SB), non-unity regression line (NU) and lack of correlation (LC). Researchers derived the RPL statistic (Ratio of Prediction error to Lab reference error) as a simple method to scale SEP relative to the precision error of the standard laboratory reference method (SEL).

2.3 *Results*

Summary statistics for standard laboratory measured soil organic carbon (SOC) and soil clay are presented in Table 1. Values for SOC did not exceed 27.2 g kg^{-1} and clay content ranged between 55 and 483 g kg^{-1} across eight farm fields. Soil organic carbon varied little across all sites (σ = 3.2 g kg^{-1}, CV = 26.7%) compared to soil clay (σ = 91.4 g kg-1, CV = 49.6%). Standard error of lab measurements (SEL_{SOC} = 0.95 g kg^{-1}, SEL_{clay} = 19.7 g kg^{-1}) accounted for 22% and 29% of clay and SOC variation, respectively.

Table 1: Summary Statistics

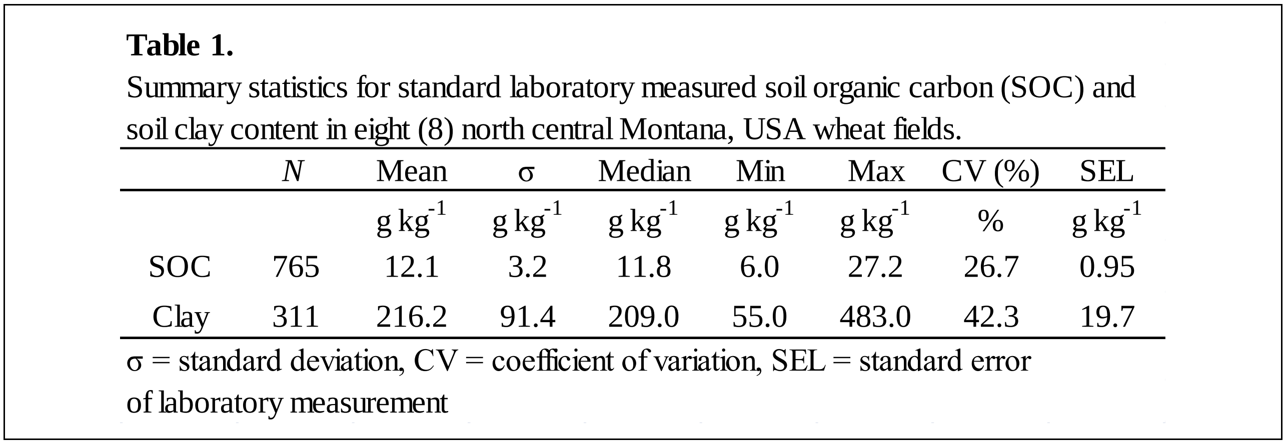

Table 1.
Summary statistics for standard laboratory measured soil organic carbon (SOC) and soil clay content in eight (8) north central Montana, USA wheat fields.

	N	Mean	σ	Median	Min	Max	CV (%)	SEL
		g kg^{-1}	g kg^{-1}	g kg^{-1}	g kg^{-1}	g kg^{-1}	%	g kg^{-1}
SOC	765	12.1	3.2	11.8	6.0	27.2	26.7	0.95
Clay	311	216.2	91.4	209.0	55.0	483.0	42.3	19.7

σ = standard deviation, CV = coefficient of variation, SEL = standard error of laboratory measurement

The research team did not achieve even semi-quantitative levels of validation accuracy for regional SOC modeling (

Table 2: PLSR model statistics

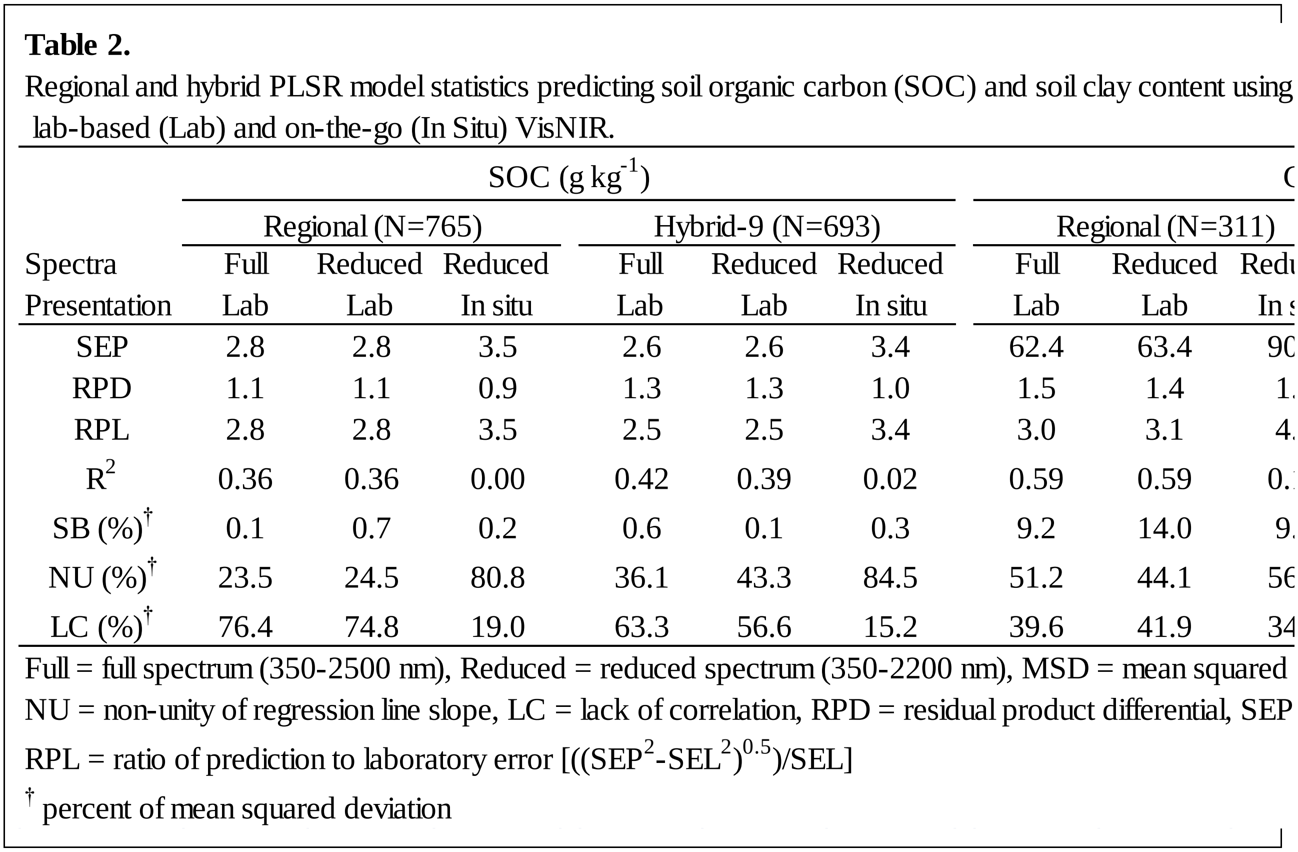

Table 2.
Regional and hybrid PLSR model statistics predicting soil organic carbon (SOC) and soil clay content using lab-based (Lab) and on-the-go (In Situ) VisNIR.

	SOC (g kg^{-1})						C		
	Regional (N=765)			Hybrid-9 (N=693)			Regional (N=311)		
Spectra Presentation	Full Lab	Reduced Lab	Reduced In situ	Full Lab	Reduced Lab	Reduced In situ	Full Lab	Reduced Lab	Redu In s
SEP	2.8	2.8	3.5	2.6	2.6	3.4	62.4	63.4	90
RPD	1.1	1.1	0.9	1.3	1.3	1.0	1.5	1.4	1.
RPL	2.8	2.8	3.5	2.5	2.5	3.4	3.0	3.1	4.
R^2	0.36	0.36	0.00	0.42	0.39	0.02	0.59	0.59	0.
SB (%)†	0.1	0.7	0.2	0.6	0.1	0.3	9.2	14.0	9.
NU (%)†	23.5	24.5	80.8	36.1	43.3	84.5	51.2	44.1	56
LC (%)†	76.4	74.8	19.0	63.3	56.6	15.2	39.6	41.9	34

Full = full spectrum (350-2500 nm), Reduced = reduced spectrum (350-2200 nm), MSD = mean squared
NU = non-unity of regression line slope, LC = lack of correlation, RPD = residual product differential, SEP
RPL = ratio of prediction to laboratory error $[((SEP^2-SEL^2)^{0.5})/SEL]$

† percent of mean squared deviation

). The best regional SOC calibration model (RPD = 1.1, RPL = 1.8) was derived using lab-based spectroscopy applied to prepared samples (Table 2, Figs. 3-6), with *in situ* data yielding less accurate predictions (RPD = 0.9, RPL = 2.3). Though RPL values are actually low, standard error of prediction (SEP) values (2.8 and 3.5 g kg^{-1} for lab-based and *in situ* data, respectively) were similar to σ_{soc} (3.2 g kg^{-1}). The majority of mean squared deviation (MSD) for SOC predictions was attributed to lack of correlation (LC) using prepared sample data (LC = 76.4% of MSD) and non-unity of regression line slope (NU) using *in situ* data (NU = 80.8% of MSD).

Table 2: PLSR model statistics

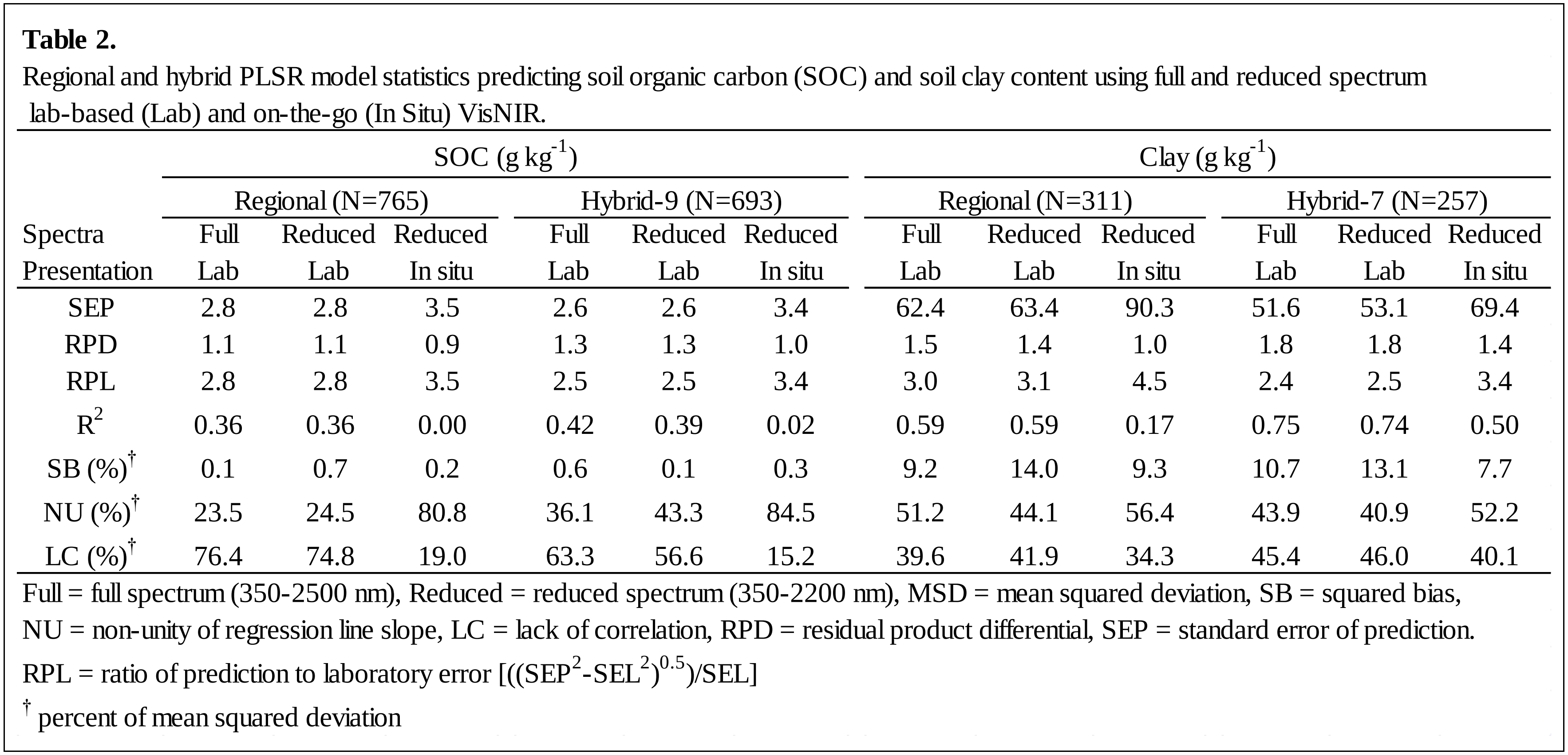

Table 2.

Regional and hybrid PLSR model statistics predicting soil organic carbon (SOC) and soil clay content using full and reduced spectrum lab-based (Lab) and on-the-go (In Situ) VisNIR.

	SOC (g kg^{-1})						Clay (g kg^{-1})					
	Regional (N=765)			Hybrid-9 (N=693)			Regional (N=311)			Hybrid-7 (N=257)		
Spectra Presentation	Full Lab	Reduced Lab	Reduced In situ	Full Lab	Reduced Lab	Reduced In situ	Full Lab	Reduced Lab	Reduced In situ	Full Lab	Reduced Lab	Reduced In situ
SEP	2.8	2.8	3.5	2.6	2.6	3.4	62.4	63.4	90.3	51.6	53.1	69.4
RPD	1.1	1.1	0.9	1.3	1.3	1.0	1.5	1.4	1.0	1.8	1.8	1.4
RPL	2.8	2.8	3.5	2.5	2.5	3.4	3.0	3.1	4.5	2.4	2.5	3.4
R^2	0.36	0.36	0.00	0.42	0.39	0.02	0.59	0.59	0.17	0.75	0.74	0.50
SB (%)†	0.1	0.7	0.2	0.6	0.1	0.3	9.2	14.0	9.3	10.7	13.1	7.7
NU (%)†	23.5	24.5	80.8	36.1	43.3	84.5	51.2	44.1	56.4	43.9	40.9	52.2
LC (%)†	76.4	74.8	19.0	63.3	56.6	15.2	39.6	41.9	34.3	45.4	46.0	40.1

Full = full spectrum (350-2500 nm), Reduced = reduced spectrum (350-2200 nm), MSD = mean squared deviation, SB = squared bias, NU = non-unity of regression line slope, LC = lack of correlation, RPD = residual product differential, SEP = standard error of prediction.

RPL = ratio of prediction to laboratory error [$((SEP^2-SEL^2)^{0.5})/SEL$]

† percent of mean squared deviation

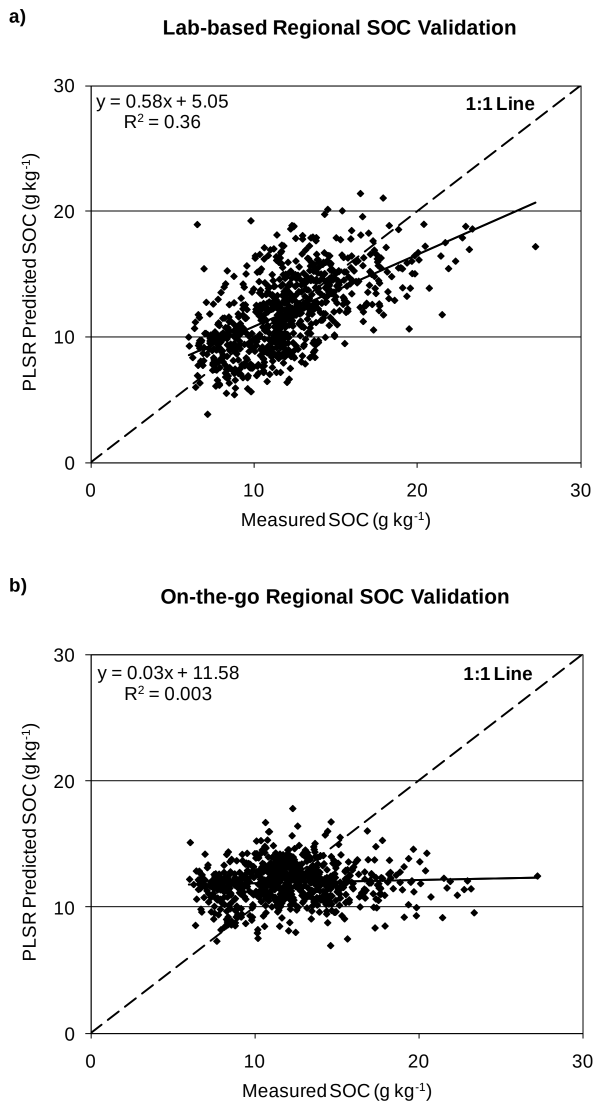

Figure 3: Whole-field out validation of soil organic carbon predicted using partial least squares regression modeling of lab-based (a) and on-the-go VisNIR (b)

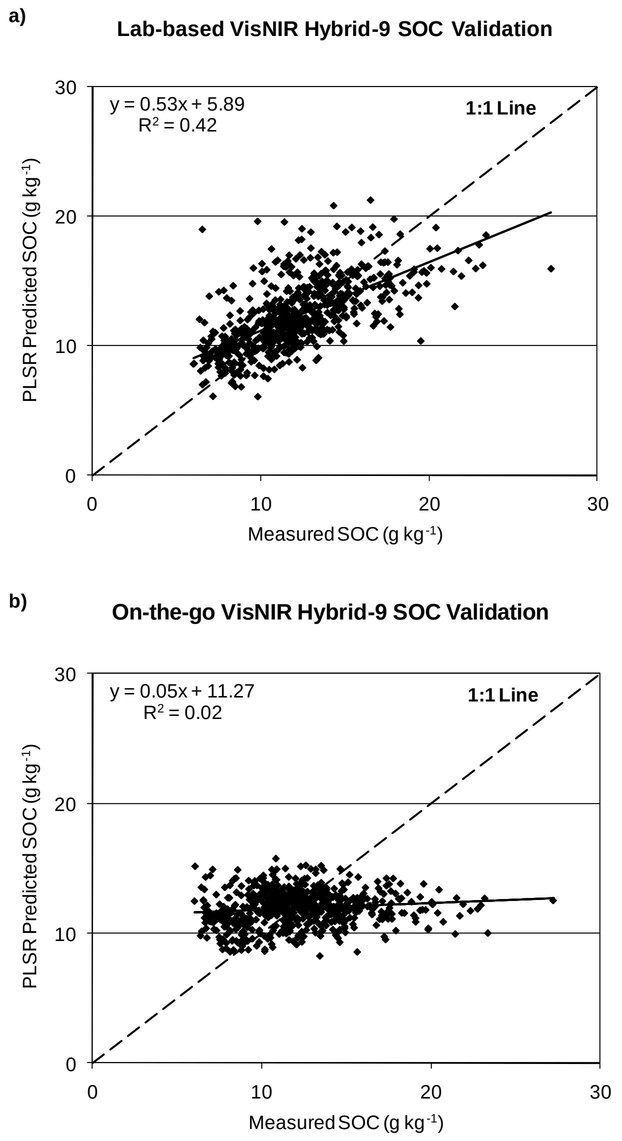

Figure 4: Hybrid model validation of SOC predicted using PLSR modeling of lab-based (a) & on-the-go VisNIR (b)

Regional calibrations for soil clay content were more accurate using lab-based interrogation of prepared samples. The regional clay model derived from prepared sample spectra achieved semi-quantitative accuracy (RPD = 1.5, RPL = 3.0). Calibrations using the *in situ* spectra, again yielded limited prediction accuracy (RPD = 1.0, RPL = 4.5). The lab-based regional clay model SEP (62.4 g kg^{-1}) was substantially lower than both the *in situ* regional model SEP (90.3 g kg^{-1})

and σ_{clay} (91.4 g kg^{-1}) (Table 2, Figures. 3-6). The majority of the MSD for both prepared sample and *in situ* derived regional clay models was attributed to NU (51.2% and 56.4% of MSD, respectively).

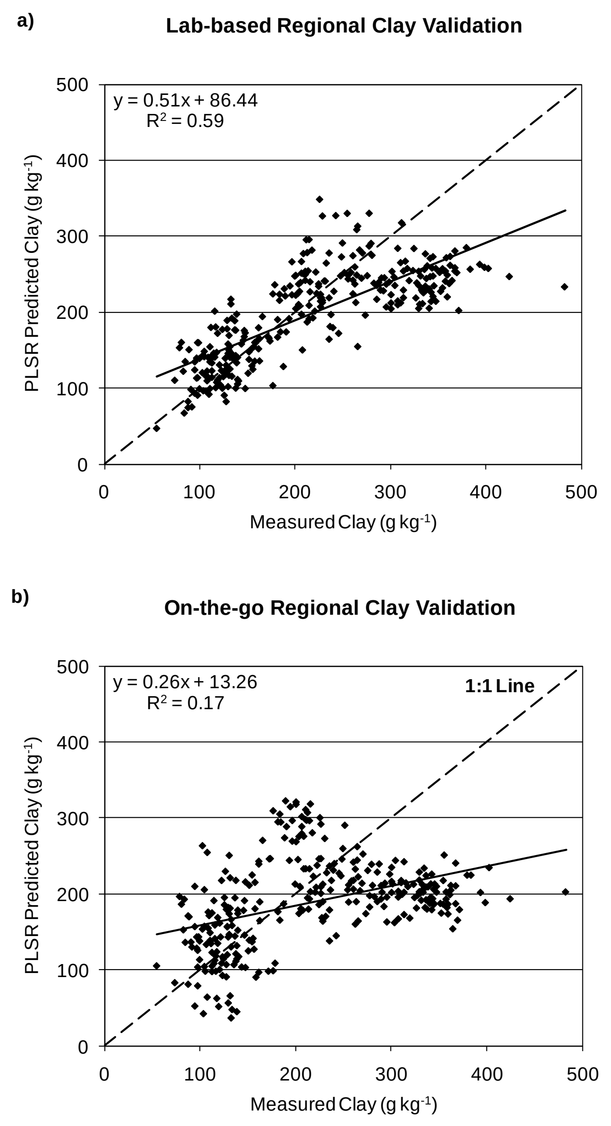

Figure 5: Whole-field out validation of soil clay content predicted using partial least squares regression modeling of lab-based (a) & on-the-go VisNIR (b)

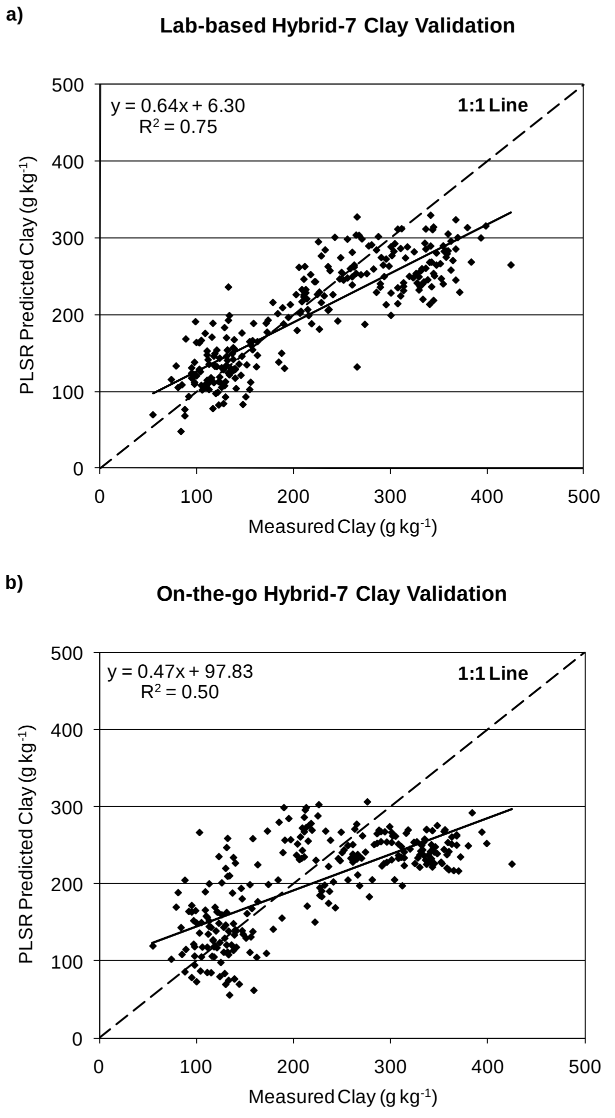

Figure 6: Hybrid model validation of soil clay content predicted using PLSR modeling of lab-based (a) & on-the-go VisNIR (b)

Increasing the number of "local" samples added to regional calibrations (i.e. hybrid models) yielded some improvement in predictive accuracy. Adding 1 to 9 local samples to regional clay models increased RPD values from 1.4 to 1.8 for lab calibrations and 1.0 to 1.4 for *in situ* calibrations (Figure 7). There were, however, only slight accuracy improvements for SOC calibrations when local samples were added to regional models (Figure 7, Table 2).

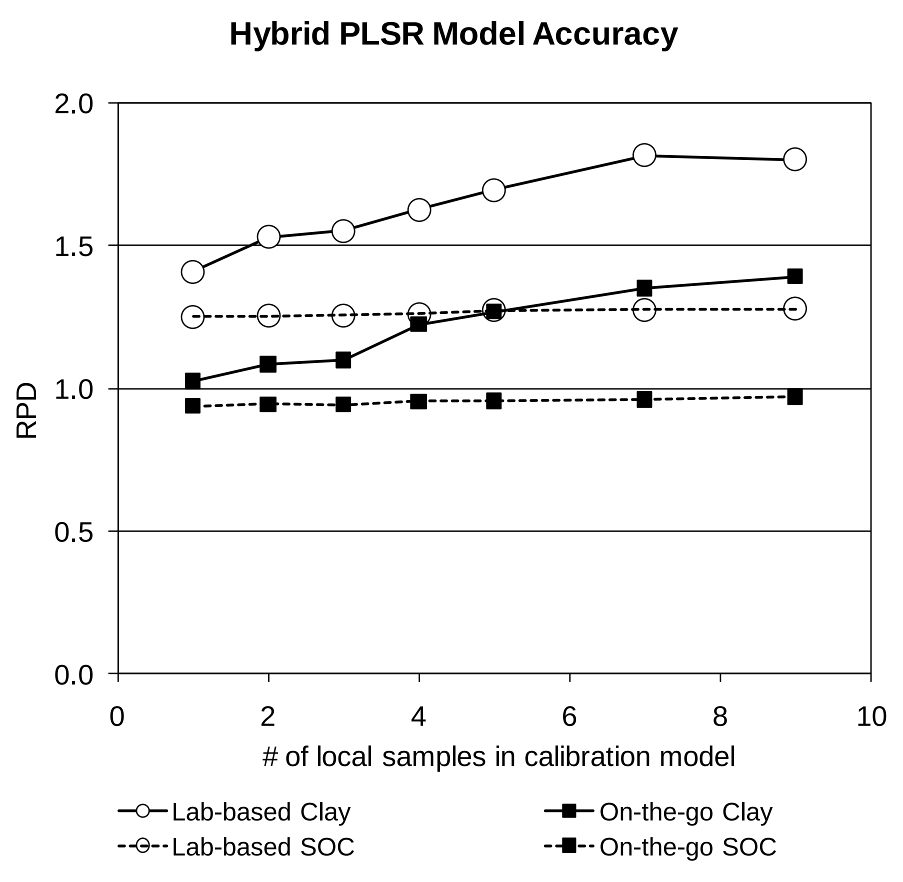

Figure 7: Predictive accuracy response to the addition of local samples in hybrid partial least squares regression calibration models for soil clay content & soil organic carbon (SOC)

Differences in the spectral range of the sensors had little effect on prediction accuracy. Reducing the spectral range of the lab-acquired data to match the spectral range of the on-the-go sensor only reduced R^2 from 0.42 to 0.39 for the hybrid-9 SOC model, and SEP and RPD were unchanged (Table 2). Regional clay models were only slightly degraded with SEP increasing 1.0 g kg^{-1} and RPD decreasing 0.1 (Table 2). Similarly, the hybrid clay model SEP increased 1.5 g kg^{-1} (Table 2).

2.4 *Discussion*

Both lab-based and on-the-go SOC predictive accuracy results, expressed as SEP, were similar to studies reviewed in the introduction, although not highly correlated with laboratory measurements. Both lab-based and on-the-go regional and hybrid SOC model RPD values (0.9 – 1.3) were lower than on-the-go RPD (1.7) reported by Christy (2008). But, measured SOC standard deviation (σ_{soc} = 3.2 g kg^{-1}) in this study, was substantially less than σ_{soc} reported by Christy (5.1 g kg^{-1}, 2008)[2]. Interestingly, SEP values were similar between this and the Christy (2008) study (3.5 and 3.0 g kg^{-1}, respectively). Clay content was more variable than SOC (CV= 42.3% vs. 26.7%, respectively), and clay models had higher RPD values. It is important to remember that given constant prediction error (SEP), increasing target parameter variability (σ_Y) will result in improved r^2 and RPD statistics. So the poor performance of on-the-go VisNIR in this study *could* be due to the low variability in SOC and to a lesser extent clay content. The fact that even the lab-based VisNIR calibrations were very poor for SOC and only semi-quantitative for clay content lends credence to this argument.

Previous studies have highlighted the necessity of a wider spectral range to accurately predict SOC. For example, Brown et al. (2006) reported that reflectance between 2000 – 2500 nm was important for SOC and clay content prediction. Similarly, Mouazen et al (2006) suggested that collecting diffuse VisNIR reflectance between 1700 and 2500 nm improved prediction accuracy for wet, *in situ* soil nitrogen (N), carbon (C), sodium (Na), and magnesium (Mg) measurements; and Sudduth and Hummel (1993a) reported that 1660 – 2620 nm was the most predictive range for organic carbon. The spectral range of the on-the-go sensor in this study was 350 – 2200 nm, compared to 350 – 2500 nm captured by the laboratory instrument. However, differences between the spectral ranges of the lab-based and on-the-go spectrometers did not contribute to meaningful differences in predictive accuracy. The results from this study suggest that 1) spectral information between 2200 and 2500 nm did not improve lab calibrations for SOC and clay, and 2) the reduced spectral range was not responsible for degraded on-the-go sensor prediction accuracy for this study.

On-the-go VisNIR measurements also have unique concerns related to continuously collecting data while moving through the field. Soil passing the sensor during scanning could cause different wavelengths to be captured at different physical locations (Christy, 2008; Sudduth and Hummel, 1993a). This potential problem was not a consideration in this study because the on-the-go instrument employed an array spectrometer that captured the entire spectrum simultaneously by using a grating to separate the reflected light according to wavelength, and then projected the light onto an InGaAs detector with an integration time of 0.042 s. Scanning type spectrometers, such as the lab-based instrument used in this study, measure one wavelength at a time and progress through the entire spectrum with an integration time of 0.1 s to complete a scan. A scanning type spectrometer used on-the-go could degrade accuracy by measuring soil reflectance across different soil scenes as it collects data through the spectrum. The scanning nature of the lab-based instrument was a non-issue in this study because soil samples were stationary when interrogated in the laboratory.

---

[2] σ_{soc} estimated for the Christy (2008) study assuming 58% SOC in soil organic matter (Nelson and Sommers, 1982).

Soil heterogeneity may have caused reduced prediction accuracy for the on-the-go VisNIR sensor compared to the lab-based spectrometer. Precipitation just prior to field sampling operations likely increased spatial and temporal variation of soil moisture across the study area. Oxygen-Hydrogen (O-H) bonds associated with water have two strong absorption features near 1400 and 1900 nm which may have obscured spectral information important to on-the-go SOC and clay predictions. More so, inconsistent soil roughness occurring as soil macro-aggregates were disturbed by pulling a shank through the soil also may have negatively impacted accuracy results. This is consistent with previous studies that have indicated that in situ measurement accuracy can be degraded by heterogeneity in soil moisture, aggregation, and surface roughness (Christy, 2008; Morgan et al., 2009; Shonk et al., 1991; Waiser et al., 2007).

Inconsistent soil presentation and soil smearing are other possible sources of error in this study. Sudduth and Hummel (1993a) reported considerable reflectance value offset as sensor-to-sample height varied from 5 to 25 mm. The quartz-sapphire window of the on-the-go sensor in this study slid along the bottom of a trench opened by the lead edge of the shank. Under ideal conditions, the sensor window would have consistent soil contact, which maintains a constant view height and angle for the reflected VisNIR signal. However, wheat field surfaces are not smooth and pulling a VisNIR sensor behind a tractor over rough and uneven surfaces possibly caused inconsistent soil contact with the bottom of the trench, thus view height and angle variation could have occurred. Additionally, soil smearing, particularly under the moist conditions during field sampling, could have occurred as the shank and sensor window slid along the trench. Simulating a VisNIR sensor being pushed vertically into soil, Waiser et al. (2007) and Morgan et al. (2009) both reported higher RMSD for predicting clay and SOC when field-moist intact soil cores were smeared prior to VisNIR interrogation.

Differences in spatial support (Dungan et al., 2002) for the reference laboratory analyses vs. *in situ* VisNIR interrogations could have impacted the estimated accuracy for *in situ* VisNIR soil calibrations. Data obtained on-the-go was the average of 20 spectra collected at ~ 5 km/h along a ~10-cm deep trench, with each resulting spectrum capturing approximately 4-5 m of travel. For both *in situ* and lab-based VisNIR model calibration and validation, we extracted three 0-10 cm deep, 5 cm diameter soil cores at the vertices of a triangle that measured 1-m along each side and straddled the transect. These samples were composited and homogenized for laboratory VisNIR and reference analysis. Christy (2008), however, collected reference soil samples for model calibration and validation that were 1 cm deep, 3 cm wide, and 1 m long from the bottom of the trench directly interrogated by the on-the-go VisNIR sensor. Support differences between the on-the-go sensor and laboratory analyses could have contributed to the poorer *in situ* vs. lab-based calibrations reported in this study.

Most of the potential problems with on-the-go VisNIR soil spectroscopy are eliminated in the laboratory due to processing steps that remove moisture, remove gravels, break up aggregates, homogenize soil material, and ensure good sample presentation. An instrument that processed soils on-the-go, in the field prior to VisNIR interrogation would likely produce substantially better calibrations. Alternatively, perhaps on-the-go VisNIR should only be applied to fields with substantial variability in soil properties of interest.

Including up to nine ‘local’ samples in hybrid models had differing effects on prediction accuracy for SOC and clay content. Including up to nine ‘local’ samples in hybrid SOC models had no apparent impact on predictive accuracies, likely because there was little SOC variability across the study area (CV = 26.7%). In comparison, clay content had greater variation across the study area (CV = 42.3%) and adding ‘local’ samples to regional clay models increased RPD from 1.5 to 1.8 and 1.0 to 1.4 for lab-based and *in situ* interrogations, respectively. This finding suggests that hybrid models would be more effective for predicting target variables with greater variability.

2.5 *Conclusions*

Lab-based VisNIR spectroscopy provided somewhat more accurate predictions than *in situ* on-the-go VisNIR sensing. In terms of SOC predictive accuracy, our results are largely consistent with those previously published by Christy (2008), but on-the-go VisNIR was not able to capture the subtle SOC variability in Montana soils. Estimating SOC in fields with low SOC variability did not produce usable results for either on-the-go or lab-acquired spectra. Spectra from prepared samples did, however, yield semi-quantitative regional and hybrid calibrations for soil clay. Regional clay models derived from on-the-go VisNIR spectra did not provide useful predictions; however, hybrid on-the-go soil clay models, using up to seven local samples in the calibration approached semi-quantitative predictive levels (RPD = 1.4, RPL = 3.4). This suggests on-the-go VisNIR spectroscopy has potential for mapping soil clay, assuming that local samples are available for recalibration at every field. Results comparing spectral ranges of the two instruments suggest increasing the spectral range of the on-the-go sensor similar to that of a lab-based spectrometer will not improve predictions for this application. Our findings indicate that on-the-go VisNIR might not be effective in mapping fields with relatively low target property variability. Given the better performance with processed soil samples in the laboratory, researchers and equipment designers might consider developing instruments to process soils on-the-go, in the field.

3. Intact soil core total, inorganic and organic carbon measurement using laser-induced breakdown spectroscopy (LIBS). (Accepted for publication, SSSAJ)

Abstract

Laser-induced breakdown spectroscopy (LIBS) is an emerging elemental analysis technology with the potential to provide rapid, accurate and precise analysis of soil constituents, such as carbon, *in situ* across landscapes. Researchers evaluated the accuracy of LIBS measuring soil profile carbon in field-moist, intact soil cores simulating conditions that might be encountered by a probe-mounted LIBS instrument measuring soil profile carbon *in situ*. The team interrogated 78 intact soil cores from three north central Montana wheat fields. Samples from each core were analyzed in the laboratory for total carbon (TC), inorganic carbon (IC), and soil organic carbon (SOC). Partial least squares 2 regression (PLS2) calibration models were derived using 58 cores (227 samples) and independently validated at the field-scale with the remaining 20 cores (79 samples). Researchers obtained the best LIBS validation predictions for IC (r^2 = 0.66, SEP = 5.3 g kg^{-1}, RPD = 1.7) followed by TC (r^2 = 0.63, SEP = 6.0 g kg^{-1}, RPD = 1.6) and SOC (r^2 =0.22, SEP = 3.2 g kg^{-1}, RPD=1.1). Though the standard error of prediction (SEP) for SOC was less than that for TC and IC, low SOC variability resulted in low r^2 and RPD statistics. Laser-induced breakdown spectroscopy is fundamentally an elemental analysis technique, yet LIBS PLS2 models appeared to discriminate IC from TC. Regression coefficients from these models suggested a reliance upon stoichiometric relationships between carbon (247.8 nm) and other elements related to total and inorganic carbon in the soil matrix [Ca (210.2 nm, 211.3 nm, and 220.9 nm), Mg (279.55-280.4 nm, 285.26 nm), and silicon (Si) (251.6 nm, 288.1 nm)]. Expanding the LIBS spectral range to capture emissions from a broader range of elements related to soil organic matter might improve SOC predictions. Results indicate that LIBS spectral data, collected on intact soil cores, can be calibrated to accurately estimate and differentiate between soil total and inorganic carbon concentrations using PLS2 regression analysis. A lack of SOC variability limited our ability to evaluate LIBS SOC prediction capabilities, with σ_{SOC} = 3.47 g kg^{-1} = 2.5 × SEL_{SOC} (standard error of the laboratory reference measurement). Calibration performance statistics from this study were substantially degraded relative to previously published research, a result attributed to the challenges of interrogating intact soil surfaces vs. prepared soil samples. As research matures for this emerging spectroscopic method, new field-deployable equipment should be developed to exploit the unique ability of LIBS to rapidly characterize soil elemental composition.

3.1 *Introduction*

There is a growing need for rapid, accurate, and inexpensive methods to measure and verify soil organic carbon (SOC) sequestration for national greenhouse gas accounting and the development of a soil carbon trading market (Council, 1999; Gehl and Rice, 2007). In particular, techniques for the rapid measurement of SOC *in situ* are required (Christy, 2008; Gehl and Rice, 2007). Laser-induced breakdown spectroscopy (LIBS) is an emerging spectroscopic technique for rapid quantification of soil carbon and other soil constituents (Cremers et al., 2001; Ebinger et al., 2003; Martin et al., 2003; Martin et al., 2004; Martin et al., 2007). Moreover, the LIBS instrument is capable of being mounted in a soil penetrometer (Mosier-Boss et al., 2002) which could be deployed for rapid soil profile characterization and mapping at field and landscape scales.

Laser-induced breakdown spectroscopy (LIBS) is fundamentally an ***elemental analysis*** probe based on atomic emission spectroscopy. Thus LIBS has the potential to complement characterization of soil minerals and organic molecules provided by visible-near infrared diffuse reflectance spectroscopy (Brown et al., 2006; Clark, 1999; Hunt, 1982). Typical LIBS analysis involves directing a focused Neodymium-doped yttrium aluminum garnet (Nd:YAG) laser onto the surface of a target material (Radziemski and Cremers, 1989). The focused laser ablates a small amount of surface material producing expanding plasma, containing electronically excited ions, atoms, and small molecules. As these excited species relax to lower electronic states they emit light at wavelengths indicative of the elemental composition of the ablated sample. Some of the emission is captured by a fiber optic cable, directed into a dispersive spectrometer, and recorded with a charge coupled detector (CCD) (Clegg et al., 2009; Cremers et al., 2001; Ebinger et al., 2003; Martin et al., 2003; Radziemski and Cremers, 1989; Thompson et al., 2006). The resulting spectra show discrete emission lines that represent electronic emissions for most atoms and ions present in the ablated material. LIBS spectra from elementally complex samples, such as soils, are spectrally complex as depicted in Figure 8. The spectrally robust and complex nature of LIBS allows one to detect every element present in the sample probed (above the detection limit) on every laser shot.

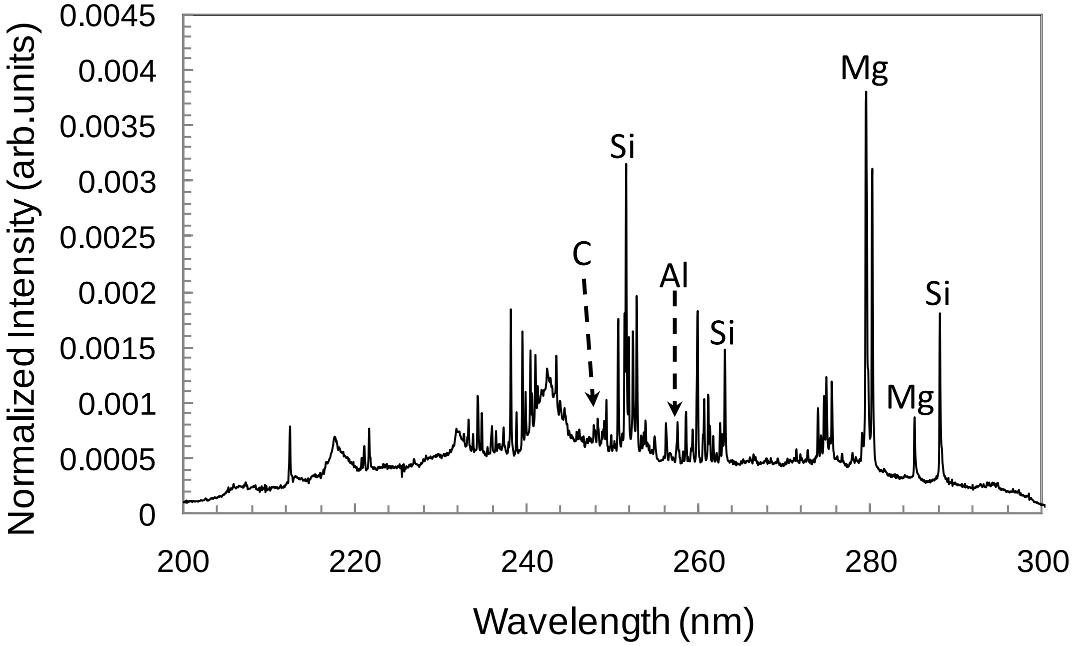

Figure 8: A LIBS spectrum (200-300 nm) of a representative soil core. LIBS spectra are typically spectrally rich containing many emission lines for each element in the sample. Some of the emission features associated with the major elements present in the sample are identified.

Univariate calibrations of LIBS spectra are generally complicated by chemical matrix effects. Chemical matrix effects have been defined as chemical properties of the interrogated sample that impact the relationship between emission line intensity (or area) and the concentration of the element in the sample responsible for producing that line (Cremers and Radziemski, 2006; Eppler et al., 1996; Gornushkin et al., 2002; Häkkänen and Korppi-Tommola, 1998). More specifically, matrix effects are related to the elemental and molecular composition of the sample, plasma composition, within plasma interactions, and laser-sample coupling efficiency. Previously published studies have attempted to compensate for these matrix effects and increase predictive accuracy by employing a variety of approaches (Clegg et al., 2009): (a) peak height or peak area calibration to standards with known composition (Ebinger et al., 2003; Eppler et al., 1996; Martin et al., 2003; Salle et al., 2006); (b) normalization of LIBS spectra to total emission intensity (Clegg et al., 2009; Thompson et al., 2006); (c) normalization of peak height or area to another spectral feature (Cremers et al., 2001; Eppler et al., 1996; Salle et al., 2006); (d) employing a plasma physics model without the use of calibration curves or “calibration-free LIBS” (Salle et al., 2006; Yaroshchyk et al., 2006); (e) spectrally averaging multiple interrogations per sample for calibration and and/or validation (Bousquet et al., 2007; Bousquet et al., 2008; Clegg et al., 2009; Eppler et al., 1996; Martin et al., 2003), and (f) employing chemmometric statistical approaches for predictive model calibration (Bousquet et al., 2007; Clegg et al., 2009; Ferreira et al., 2008; Martin et al., 2005; Martin et al., 2010; Martin et al., 2007; Sirven et al., 2006)

With proper calibration, LIBS can provide a precise and selective method for measuring metal ions such as lead (Pb), beryllium (Be), chromium (Cr), and strontium (Sr) in paint and soils (Sirven et al., 2006; Yamamoto et al., 1996), N, Pb, and Ba in sand (Eppler et al., 1996; Harris et al., 2004), and copper (Cu), zinc (Zn), and arsenic (As) in wood preservatives (Martin et al., 2005). Though there have been relatively few applications of LIBS for soil carbon determination, published calibrations show LIBS spectra to be well correlated with standard dry combustion measurements of total soil carbon with reported r^2 values of 0.93 to 0.99 (Cremers et al., 2001; Ebinger et al., 2003; Martin et al., 2003; Martin et al., 2010; Martin et al., 2007). There is; however, no published literature demonstrating the ability of LIBS to distinguish between total and inorganic soil carbon.

There has been little independent validation of published LIBS calibrations for soil carbon using a large number of soil samples. Cremers et al. (2001) used a subset of 12 Colorado agricultural soil samples from conventionally tilled farms to calibrate a LIBS model ($r^2 = 0.96$) and verified the model with a different subset (N=8) of the same Colorado soils, as well as soils from Los Alamos, NM (N=10) that formed in different parent materials (accuracy = 3 to 14% relative standard deviation). Ebinger et al. (2003) used six randomly chosen soil samples from a dataset of 18 samples collected from three Colorado farm fields to calibrate a model ($r^2 = 0.99$) then used the model to predict the remaining 12 samples ($r^2 = 0.95$). It is not yet standard practice in LIBS spectroscopy to ‘hold-out’ independent samples for validation (Martin et al., 2003; Martin et al., 2010; Martin et al., 2007). While published research shows the potential of LIBS for SOC determination, further work is required with larger sample sets and more rigorous model validation.

The soil samples employed in the studies cited above were pre-treated prior to LIBS interrogation. Pre-treatments included: air-drying, sieving and packing in quartz tubes (Cremers et al., 2001); pelletizing under pressure (Martin et al., 2004; Martin et al., 2010; Martin et al., 2007); and treating with acid to remove carbonates, pelletizing in a tube, and air-drying (Martin et al., 2003). Though LIBS has been proposed as an *in situ* SOC measurement tool (Gehl and Rice, 2007), it remains to be demonstrated that *in situ* results will match those obtained with prepared samples. To our knowledge, the study reported in this paper is the first to measure soil carbon content in field-moist intact cores without soil physical pretreatments.

The objectives of this study were to (1) evaluate the accuracy of field-scale LIBS calibrations for soil profile carbon in field-moist, intact soil cores without soil pretreatments, and (2) determine if TC, IC, and SOC can be differentiated using LIBS calibrations. Though important SOC elements such as H, O and N cannot be detected with the 200-300 nm LIBS spectral range employed in this study, researchers theorized that it might be possible to estimate SOC by subtraction should Ca and Mg emissions support the discrimination of total and inorganic C. The team defined accuracy as agreement between LIBS measurements and standard laboratory-based soil measurements. Field-moist intact cores were used to simulate conditions that might be encountered by a probe-mounted LIBS instrument measuring soil profile carbon *in situ*.

3.2 ***Material and Methods***

3.2.1 Study Area

The "Golden Triangle" region of north central Montana, USA served as our research study area (Figure 9). This region was characterized by soils formed in glacial till on gently rolling topography. Soils were not highly weathered and were typically calcareous within 0.5 m of the surface. Aridic intergrades of frigid, ustic, Mollisols, Entisols, and Inceptisols predominated. Cropping systems in the study area were generally reduced tillage small grain-fallow rotations with a significant acreage managed by direct-seeding or no-till. All three sampling sites had a general cropping history of cultivation beginning in the 1920's progressing to wheat-fallow rotations with multiple tillage operations per year and finally conversion to a direct-seeded wheat-fallow rotation between 2004 and 2005.

3.2.2 Soil Sampling

In 2006, 78 intact cores were obtained from three 16.2 ha sub-fields in north central Montana with locations shown in Figure 9. Soil coring locations were selected based on surface soil (0-10 cm) visible and near-infrared (VisNIR) reflectance acquired for a parallel study focused on that technology (Bricklemyer and Brown, 2010). Intact, 4.45 cm diameter by 50 cm deep soil cores were extracted using a truck-mounted hydraulic soil sampling tube fitted with removable plastic sleeves (Giddings Machine Co., Windsor, CO). The field-moist intact cores were transported to the laboratory and stored under refrigeration prior to interrogation.

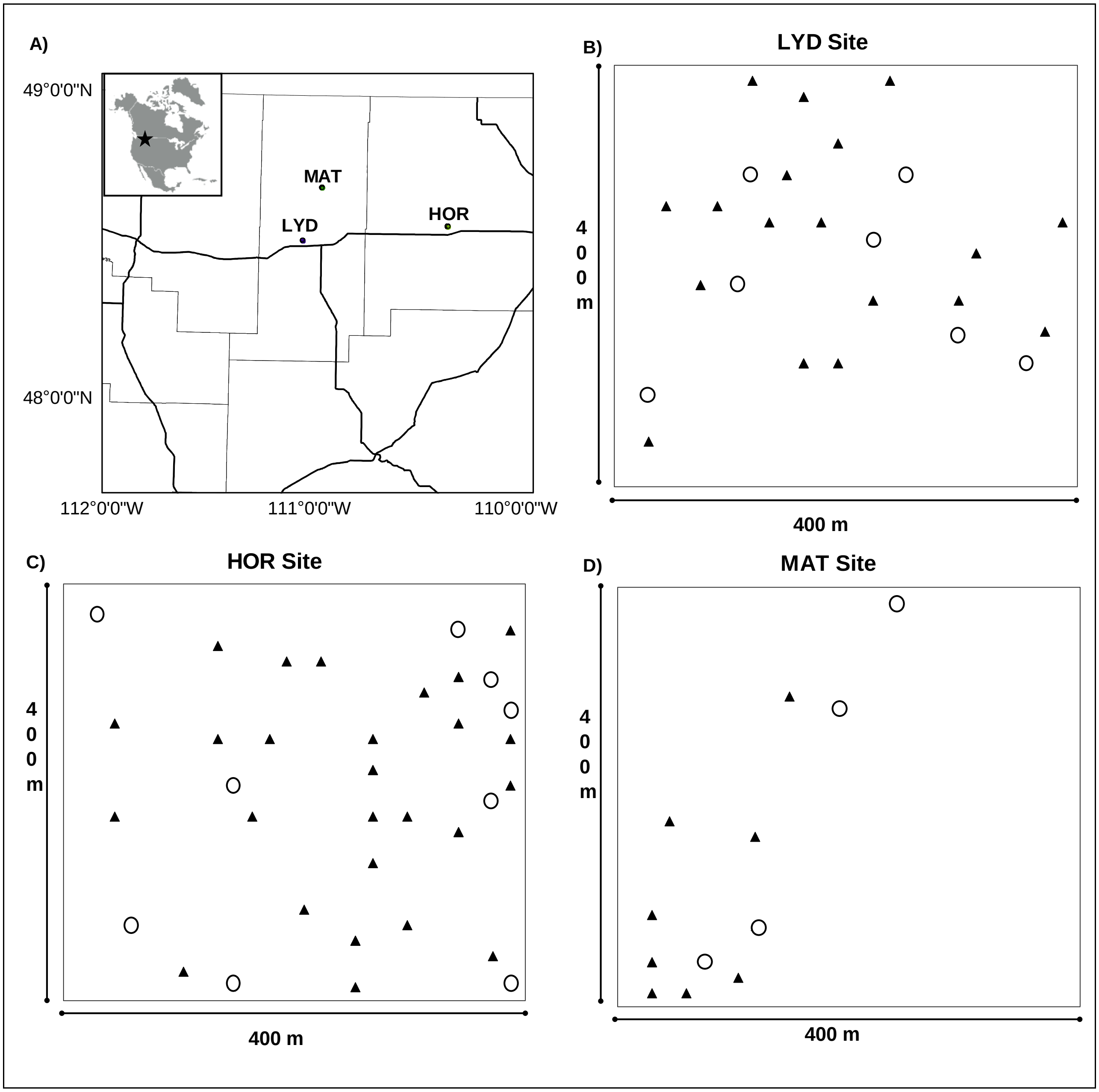

Figure 9: Geographical location of the study area with 3 selected farm fields (A), & randomly selected calibration (triangles) & validation (circles) core locations at the LYD (B), HOR (C), & MAT (D) sites.

3.3 *Core Interrogation*

Researchers interrogated intact soil cores to simulate conditions that might be encountered by a penetrometer-mounted LIBS instrument performing *in situ* soil characterization following the general protocol of Waiser et al. (2007). Each field moist core was interrogated at 8 depths through ~ 3 x 3 cm windows cut in the plastic core sleeve (Figure 10). A prototype Los Alamos

National Laboratory LIBS Core Scanning (LIBS-CS) instrument was used to probe the samples in an argon purged environment at 2.5, 7.5, 12.5, 17.5, 22.5, 27.5, 35, and 45 cm (+/- 1.5 cm) along each intact soil core with 9 interrogation points per depth. Interrogation points were approximately 0.5 cm apart and centered within the 3 x 3 cm window. Each LIBS interrogation point measured approx. 200 μm in diameter and probed a volume of ~ 8×10^{-5} cm^3. The LIBS-CS instrument employed a Big Sky Laser operating at 1064 nm, firing laser pulses at 10 Hz, producing approximately 80mJ/pulse.

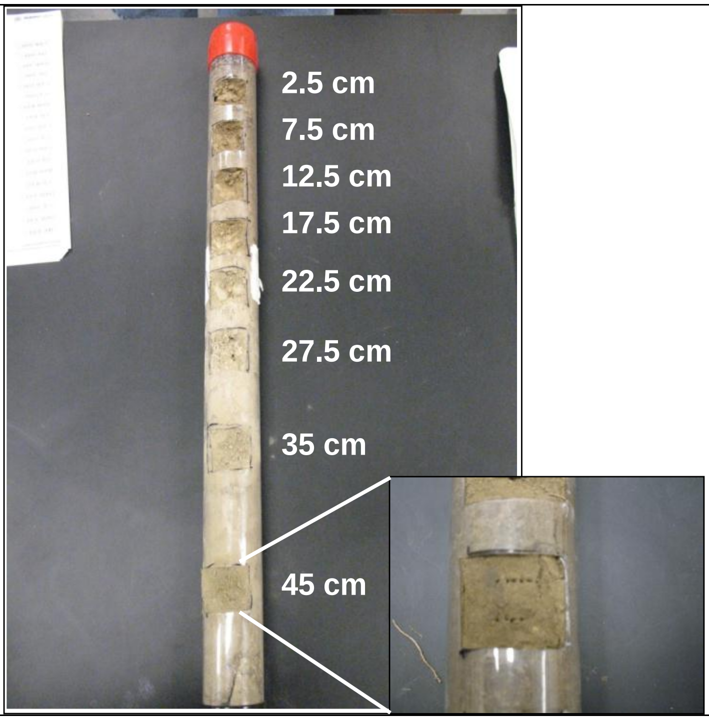

Figure 10: LIBS sampling depths on a representative intact soil core & LIBS interrogation point configuration within a depth sample (inset).

The LIBS spectra were collected with an optical fiber and directed into an Ocean Optics HR2000 spectrometer (200-300 nm, 0.1 nm spectral resolution, 2046 wavelengths). The 200-300 nm spectral range was chosen because, at the time of instrument development, published studies measuring carbon with LIBS typically used univariate calibration using peak height of the carbon emission at 247.8 nm. The spectrometer is readout noise limited and signal-to-noise ratios improve if emissions from multiple laser shots are used (Clegg et al., 2009). For this experiment, the team set the spectrometer to a one-second integration time with 5 averages so that each spectrum (i.e. interrogation point) represented 50 laser shots. An argon purge was used to replace ambient air between the sample and the sensor to reduce plasma interferences from oxygen. A similar argon purge would be operationally feasible for an *in situ* application, a LIBS instrument mounted in a soil penetrometer for example, given the small gas volume needed to purge a 200 μm interrogation point along a soil profile. The LIBS data were normalized to the total integrated intensity by dividing each wavelength value by the sum of all wavelength values for each spectrum prior to spectral model calibration as detailed by Thompson et al. (2006) and Clegg et al. (2009) and then averaged by depth. Normalization reduces the shot-to-shot variability in LIBS data that has been attributed to soil and chemical matrix effects (Mosier-Boss et al., 2002).

3.3.1 Soil sample analysis

Subsamples of soil (~ 4 g) were taken from all interrogation depths for laboratory analysis. To reduce analysis cost, yet capture representative soil variability within cores, researchers used stratified random sampling to choose samples from each core for carbon analysis and LIBS spectral modeling. Sample selection employed VisNIR spectra acquired concurrently with LIBS interrogations (Morgan et al., 2009). VisNIR spectra from the 8 interrogation depths were clustered into 4 spectrally similar groups, per core, using Partitioning Around Mediods (Kaufman and Rousseeuw, 1990). Partitioning Around Mediods (PAM) select a user-defined number of mediods (i.e. center points of clusters) such that within cluster variation is minimized while maximizing differences between clusters. One interrogation depth was randomly chosen from each spectral group (78 cores × 4 depths per core = 312 possible samples) and the corresponding soil sub-sample was analyzed for TC, IC, and SOC, using standard procedures described in Bricklemyer et al. (2005). Total carbon was measured by dry combustion using a LECO TruSpec (LECO Corp., St. Joseph, MI, USA). Inorganic carbon was measured by modified pressure calcimeter method as developed by Sherrod et al. (2002). Soil organic carbon (SOC) was calculated by difference: SOC = TC-IC where TC = total carbon and IC = soil inorganic carbon. Standard carbon measurements were used for LIBS calibration and validation. The final dataset included 306 samples (Table 3) from 78 cores due to incomplete LIBS spectra for six samples.

Table 3: Laboratory measured soil total, inorganic & organic carbon summary statistics for 78 intact soil cores from 3 wheat fields in north central MT, 2007

	Soil Total C[†]	Soil Inorganic C[‡]	Soil Organic C
	----------------------------	g kg $^{-1}$	-----------------------------
Median	15.74	7.95	7.82
Min	1.43	0.00	0.85
Max	56.56	45.27	19.32
Mean	17.09	8.88	8.21
σ	10.18	9.53	3.47
CV	59.6%	107.4%	44.4%
n	306	306	306

σ = standard deviation; CV = coefficient of variation

[†] dry combustion (Leco TruSpec, Leco Corp., St. Joseph, MI, USA)

[‡] modified pressure calcimeter (Sherrod et al., 2002)

3.3.2 Data Examination

Exploratory data analysis was performed on TC, IC, and SOC data prior to spectral modeling. Descriptive statistics, median, mean, and sample standard deviation were calculated and tests of normality were performed to examine the distribution of soil carbon data. The research team calculated the coefficient of variation (CV) to assess overall data variability. Total C, IC, and SOC data were also examined for correlations using Pearson's product moment correlation. Total C, IC, and SOC data were found to be non-normally distributed (data not shown). Inorganic carbon was particularly skewed due to a large number of near-surface samples that did not contain measureable IC concentrations. Using the Box-Cox transformation routine in R (Box and Cox, 1964; Venables and Ripley, 2002), square root transformation was found to be the optimal power transformation to normalize the carbon data. However, since better LIBS modeling results were obtained with non-transformed data, only the latter are reported.

3.3.3 Spectral model calibration, validation, and evaluation

The research team randomly selected data from 58 cores (represented by 227 of the 306 total samples) to calibrate LIBS partial least squares regression models for TC, IC and SOC.

Calibration models were validated using samples from the remaining 20 cores (represented by the remaining 79 samples). With cores from only three fields, researchers were not able to construct a regional calibration with whole-field cross-validation, so results of this study indicate the quality of local within-field calibrations (Brown et al., 2005).

Partial least squares 2 (PLS2) regression models were simultaneously fit to TC, IC, and SOC data. Researchers also fit single variable (PLS1) regression models to TC, IC, and SOC data individually for comparison. PLS2 is fundamentally similar to PLS1 in that variable rotation is employed to account for variability in both the response and predictor space to maximize predictive power (Geladi and Kowalski, 1986; Wold et al., 2001a). However, PLS2 allows several response variables to be modeled simultaneously in order to take advantage of any correlations that may exist between the response variables. Because total C and IC data were highly correlated in this study, PLS2 regression was theoretically the most appropriate TC, IC, and SOC PLS calibration approach (Wold et al., 2001a). All PLS1 and PLS2 analyses were implemented using the Unscrambler® v8.0 (Camo Software Inc. Oslo, Norway).

Chemometric approaches, such as PLS2, can be prone to overfitting (Brown et al., 2005; Dudragne et al., 1998). Therefore, the number of latent variables to be included in each PLS2 model was determined using calibration data only by examining a plot of full leave-one-out cross-validation residual variance against the number of latent variables. A “first minimum” selection criteria, which involved identifying the point at which the calibration model cross-validation residual variance no longer decreased with additional latent variables, was employed to reduce the potential for over-fitting (Brown et al., 2005).

The research team took a multi-pronged approach to explore the extent to which IC predictions were based purely on correlations with TC. Firstly, they computed the pair-wise correlations between validation TC, IC and SOC both lab-measured and LIBS predicted, as well as correlations between TC, IC and SOC PLS2 model coefficients. Secondly, they examined TC, IC, and SOC PLS2 regression coefficients to infer the role of stoichiometry in the construction of latent variables. Finally, they constructed a simple linear regression (SLR) calibration for IC using only lab-measured TC as a predictor, then applied this regression model to predict IC for the validation data set using i) lab-measured TC, ii) TC predicted by LIBS PLS2, and iii) TC predicted by LIBS PLS1.

To quantify the effects of mm-scale soil variability and estimate the minimum number of focused LIBS interrogations required to characterize a 1-2 cm diameter heterogeneous soil material, the team randomly selected and averaged 1, 2, 3, 5 and 7 interrogation points from the 9 acquired at each depth. Previously described spectral processing and validation procedures were repeated; however, PLS1 modeling was used to build models for each of these “reduced” LIBS interrogation scenarios using the R statistical software package “pls” (R, 2008; Wehrens and Mevik, 2007).

The quality of all PLS1 and PLS2 model fits were evaluated using regression of predicted vs. laboratory measurement and the following statistics following Gauch et al. (2003) and Brown et al. (2005):

$$MSD = \sum_{n} \frac{(Y_{pred} - Y_{meas})^2}{N} \qquad (1)$$

$$RMSD = \sqrt{MSD} \tag{2}$$

$$Bias = \sum_n \frac{(Y_{pred} - Y_{meas})}{N} \tag{3}$$

$$SB = Bias^2 \tag{4}$$

$$NU = (1-b)^2 * \mathrm{var}(Y_{meas}) \tag{5}$$

$$LC = (1-r^2) * \mathrm{var}(Y_{pred}) \tag{6}$$

where b and r^2 are the slope and square correlation, respectively from the least-squares regression of Y_{pred} on Y_{meas}. The Mean Squared Deviation (MSD) is partitioned into three independent components describing lack of accuracy due to bias (SB), non-unity regression line (NU), and lack of correlation (LC), with MSD = SB + NU + LC (Brown et al., 2005). The following chemometric modeling statistics were calculated for each model:

$$SEP = \left[\frac{\sum (y-x)}{n-1}\right]^{\frac{1}{2}} \tag{7}$$

$$RPD = \frac{SD_x}{SEP} \tag{8}$$

$$PRL = \frac{SEP_{adj}}{SEL} = \frac{\sqrt{SEP^2 - SEL^2}}{SEL} \tag{9}$$

where SEP = validation standard error of prediction, *y* = predicted value by LIBS, *x* = laboratory measured reference value, RPD = residual product differential, and SD_x = standard deviation of reference measurements (Islam et al., 2003). We introduce a new validation statistic with the acronym PRL (pronounced “pearl”) representing *prediction relative to laboratory error*. The PRL ratio scales adjusted prediction error (SEP_{adj}) relative to the precision of the standard laboratory reference method (SEL). For calibrations approaching laboratory precision, PRL approaches unity, but most predictions errors would be measured as some multiple of the reference method precision (e.g. “two times” SEL).

The standard error of laboratory measurement (SEL) for TC and IC was estimated using replicate laboratory measurements as described by Workman (2001).

$$SEL = \left\{ \frac{\sum_{i=1}^{N} (y_1 - y_2)^2}{2N} \right\}^{1/2} \tag{10}$$

where y_1 = first replicate laboratory measurement, y_2 = second replicate laboratory measurement, and N = number of samples (not number of replicates). For SOC, SEL was calculated using propagation of error estimation with the general equation (Andraos, 1996):

$$\text{SEL: } \Delta f_i \approx df_i \pm \frac{\partial f}{\partial x_i} \Delta x_i \tag{11}$$

where f is a function with $i = 1, 2, \ldots, n$ variables each with an associated uncertainty Δx_i and Δf_i = uncertainty of the function f. For this study, SOC was calculated as an algebraic sum SOC = TC – IC. Solving Eqn. 1 for an algebraic sum, applied to our example, we find:

$$SEL_{SOC} = \pm\left[(SEL_{TC})^2 + (SEL_{IC})^2\right]^{0.5} \tag{12}$$

3.4 ***Results***

3.4.1 Exploratory data analysis

Summary statistics for Lab[3] TC, IC, and SOC are presented in Table 3. For the samples in this study, TC values did not exceed 57 g kg^{-1}, IC values were less than 46 g kg^{-1}, and SOC values never exceeded 20 g kg^{-1}. Concentrations of IC were most variable in the dataset (σ = 10.18, CV = 107%) followed by TC and SOC (σ = 9.53 and 3.47; CV = 59.6% and 44.4%, respectively). The SEL for laboratory measurements were estimated at 0.90, 1.03, and 1.37 g kg^{-1} for TC, IC, and SOC, respectively. Measured SOC was least variable, where σ_{SOC} was just 2.5 times SEL_{SOC}. Variability in SOC was substantially less than IC with highest concentrations of SOC occurring in A horizons and diminished with depth. Variability in IC was a function of pedogenesis where IC was not present or present in low concentrations in A horizons (top 20 cm) and increased sharply in B horizons, the majority of which occurred below 20 cm. The sharp boundary between the A and B horizons is evident in Figure 10.

3.4.2 LIBS chemometric models

Full cross-validated PLS1 and PLS2 calibration model performance statistics are presented in Table 4. Single-response PLS calibrations for TC and IC yielded good explanatory power (RPD=2.1 and 2.3, r^2 = 0.76 and 0.81, respectively); whereas explanatory statistics were notably weaker for the PLS1 SOC calibration (RPD = 1.5, r^2 = 0.55). Conversely, the lowest calibration dataset cross-validation prediction errors, both absolute and relative, were obtained from the PLS1 SOC model with SEP = 2.5 g SOC kg^{-1} soil and PRL = 1.5. The other PLS1 cross-validation prediction errors were 5.1 g TC kg^{-1} and 4.3 g IC kg^{-1} soil, with PRL values of 5.6 and 4.3 for TC and IC, respectively. The cross-validation statistics for PLS2 calibrations were

---

[3] "Lab" refers to laboratory measurements and "LIBS" refers to a LIBS PLS2 calibration, unless otherwise indicated.

uniformly worse than those obtained from PLS1, with SEP equal to 6.2 g TC kg^{-1} soil, 5.4 g IC kg^{-1} soil, and 3.4 g SOC kg^{-1} soil and PRL equal to 6.9, 5.2, and 2.3 for TC, IC, and SOC, respectively.

Table 4: Soil total carbon (TC), inorganic carbon (IC), and organic carbon (SOC) calibration cross-validation statistics for laser-induced breakdown spectroscopy (LIBS) partial least squares 1 (PLS1) & 2 (PLS2) regression models.

Model	*n*	r^2	RPD	SEP (g/kg soil)	PRL	SB† (%)	NU† (%)	LC† (%)
PLS1 LIBS TC	227	0.76	2.1	5.1	5.6	0.0	24.6	75.4
PLS1 LIBS IC	227	0.81	2.3	4.3	4.1	0.5	28.2	71.3
PLS1 LIBS SOC	227	0.55	1.5	2.5	1.5	0.0	45.0	55.0
PLS2 LIBS TC	227	0.65	1.7	6.2	6.9	0.0	26.9	73.1
PLS2 LIBS IC	227	0.69	1.8	5.4	5.2	0.3	31.0	68.8
PLS2 LIBS SOC	227	0.17	1.1	3.4	2.3	0.0	73.5	26.5

RPD = residual product differential, PRL = prediction error relative to lab, SEP = standard error of prediction, SB = squared bias, NU = non-unity, LC = lack of correlation

† percent of mean squared deviation (MSD)

Table 5: Soil total carbon (TC), inorganic carblon (IC), and organic carbon (SOC) independent validation statistics for laser-induced breakdown spectroscopy (LIBS) partial least squares 1 (PLS1) & 2 (PLS2) regression models.

Model	*n*	r^2	RPD	SEP (g/kg soil)	PRL	SB† (%)	NU† (%)	LC† (%)
PLS1 LIBS TC	79	0.68	1.6	5.8	6.3	3.3	5.6	91.1
PLS1 LIBS IC	79	0.60	1.5	5.8	5.6	0.1	19.2	80.7
PLS1 LIBS SOC	79	0.19	1.0	3.4	2.3	3.7	46.5	49.8
PLS2 LIBS TC	79	0.63	1.6	6.0	6.5	2.3	15.4	82.3
PLS2 LIBS IC	79	0.66	1.7	5.3	5.0	0.1	24.5	75.4
PLS2 LIBS SOC	79	0.22	1.1	3.2	2.1	4.0	77.5	18.5

RPD = residual product differential, PRL = prediction error relative to laboratory error, SEP = standard error of prediction, SB = squared bias, NU = non-unity, LC = lack of correlation

† percent of mean squared deviation (MSD)

Independent validation using 20 randomly selected hold-out cores showed that PLS2 models were more stable and robust than PLS1 models for TC, IC and SOC (Table 5). For PLS1 models, validations statistics were substantially degraded relative to calibration statistics, suggesting that the calibrations were overfit. This is a potential issue with chemometric approaches such as PLS (Brown et al., 2005; Dudragne et al., 1998) even when a conservative number of latent variables are chosen. Validation statistics for PLS2 models; however, were consistent with calibration cross-validation statistics. Furthermore, PLS2 validation SEP for SOC and IC were 7 and 10% lower than for PLS1, respectively. Therefore, the remainder of the discussion is focused on the PLS2 calibrations.

For the hold-out validation dataset, the LIBS PLS2 analysis best predicted inorganic carbon (r^2 = 0.66, RPD = 1.7; Table 5, Figure 11) with total carbon explanation nearly as good (r^2 = 0.63, RPD = 1.6; Table 5, Figure 4), and minimal explanatory capability for SOC (r^2=0.22, RPD= 1.1, Table 5, Figure 11). Validation SEP values were 6.0 g TC kg^{-1} soil, 5.3 g IC kg^{-1} soil, and 3.2 g SOC kg^{-1} soil in absolute units, or relative to the lab precision PRL = 6.5, 5.0 and 2.1 for TC, IC and SOC, respectively. For SOC, though validation SEP was only twice SEL_{SOC}, the standard deviation of the target variable (σ_{SOC}) was also very low relative to SEL_{SOC} (1.37 g kg^{-1}). Partitioning MSD into three components, we found the LIBS TC model had low bias (SB=2.3%) and marginal non-unity (NU=15.4%) (Table 5). The LIBS IC model also had low bias (SB=0.1%); however, non-unity was greater (NU=24.5%). Bias remained low for the LIBS SOC model (SB=4.0%), but non-unity greatly increased (NU=77.5%; Table 5, Figure 11). Lack of correlation accounted for the majority of MSD for the LIBS TC and IC models (82.3% and 75.4%, respectively) but only 18.5% of MSD for the SOC model (Table 5).

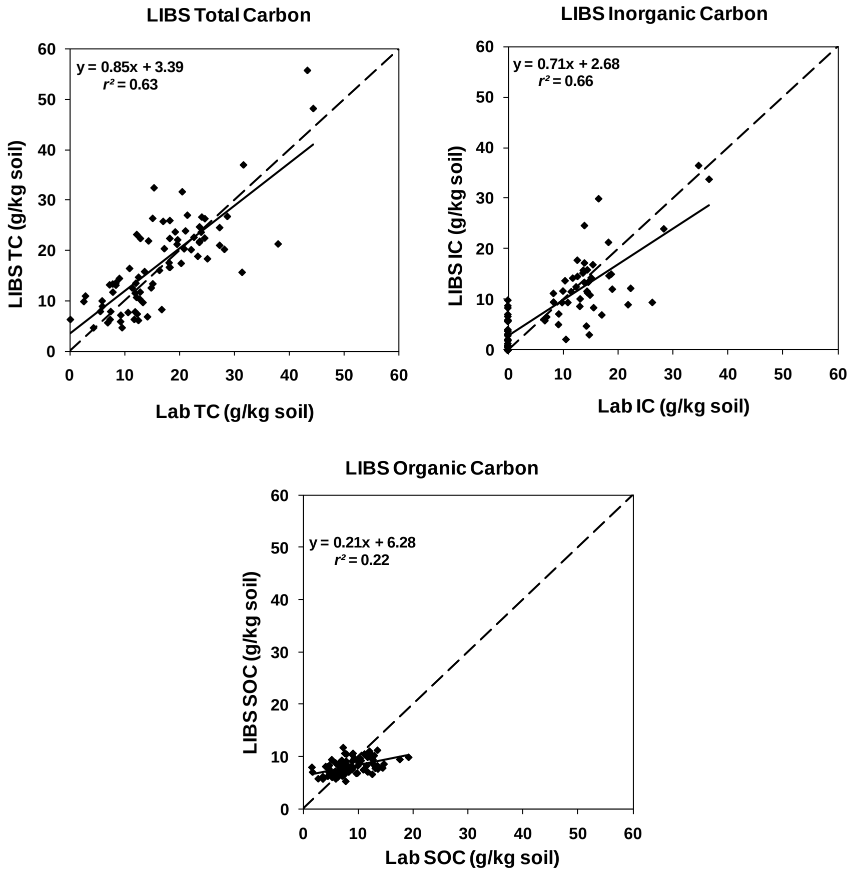

Figure 11: Independent validation of predicted soil total carbon (TC), inorganic carbon (IC), & organic carbon (SOC) using LIBS and partial least squares 2 regression models. Nine interrogation spectra were averaged for PLS2 analysis.

The importance of specific elemental emissions in TC, IC, and SOC PLS2 models (Figure 12) suggests that stoichiometric relationships were used for prediction. The major carbon emission at 247.8 nm was an important predictor for TC, IC, and SOC (relatively); though the magnitude of the regression coefficient varied between models. Researchers expect that the carbon emission line would be an important predictor for the three calibrations. However, Si (silicate mineral framework element), Mg and weak secondary Ca (associated with carbonates) emissions were also useful for predicting IC and TC. The LIBS regression coefficients for SOC were all close to zero, as might be expected with a weak calibration.

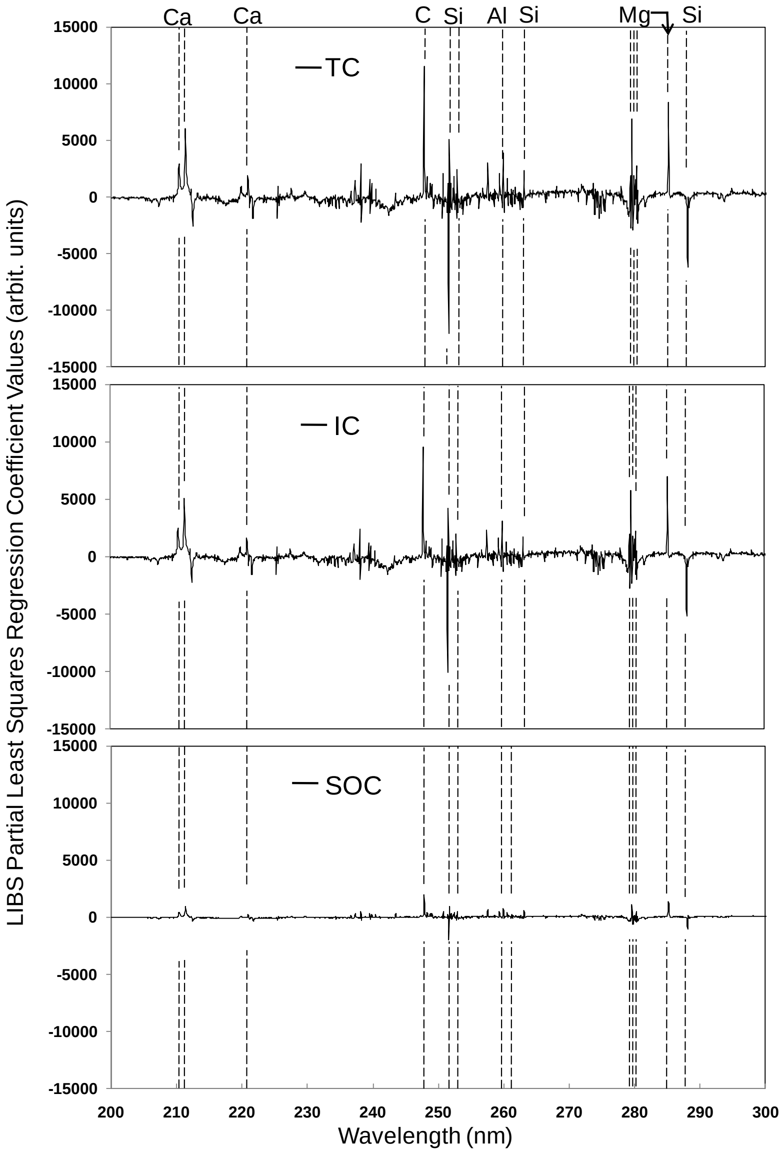

Figure 12: Partial least squares regression (PLS2) coefficient values for soil total, inorganic, & organic carbon. The magnitude of the coefficients indicates the relative importance of each emission line. Dashed vertical lines indicate important elemental emission lines for predicting the various forms of soil carbon. All wavelengths were used for predictions.

Lab TC and IC were strongly correlated for calibration set soils in this study (*r* = 0.972, Table 6), raising questions regarding the ability of LIBS to distinguish TC and IC. Total C and IC LIBS regression coefficients were also highly correlated (*r* = 0.997, Table 6) with the associated TC and IC predictions somewhat less correlated (*r* = 0.976, Table 6). Lab SOC (*r* = 0.302 and -0.076), LIBS SOC regression coefficients (*r* = 0.938 and 0.911), and LIBS SOC predictions (*r* = 0.384 and 0.198) were not as correlated with TC or IC, respectively (Table 6).

Validation results from simple linear regression (SLR) calibrations of IC on TC suggest that PLS2 IC predictions were not based solely on a correlation with TC (Table 7, Figure 13). Using validation set lab TC (total combustion analysis), the research team was able to obtain substantially improved validation IC predictions relative to LIBS PLS2 IC predictions (SEP = 3.4 vs. 5.3 g kg^{-1}). Of course, TC is not usually known at locations where researchers want to predict IC using LIBS. Using validation set TC predicted by LIBS PLS2, the SLR model yields slightly worse predictions for IC than obtained from LIBS PLS2 direct predictions of IC (SEP = 5.4 vs. 5.3 g kg^{-1}). When inorganic C (IC) was completely removed from the estimation of validation set TC (using PLS1), SLR IC validation predictions were substantially degraded relative to PLS2 and PLS1 IC predictions (SEP = 7.7 vs. 5.3 and 5.8 g kg^{-1} for PLS2 and PLS1, respectively).

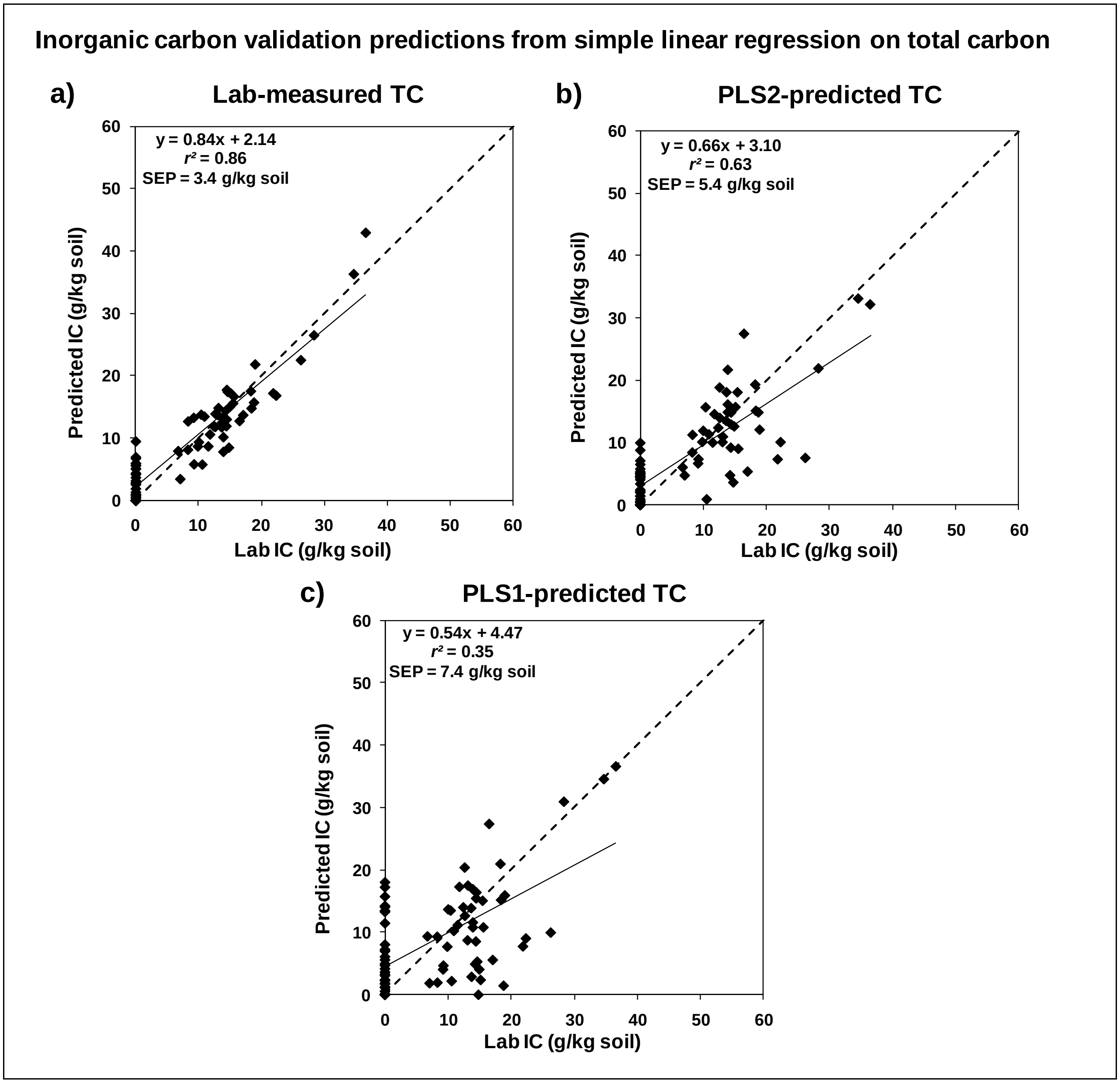

Figure 13. A simple linear regression of calibration set IC on TC (lab reference measurements) applied to validation set TC: a) measured by total combustion; b) predicted by LIBS PLS2; and c) predicted by LIBS PLS1.

Table 6: Pearson's correlation coefficients (r) for laboratory measured soil carbon, laser-induced breakdown spectroscopy (LIBS) partial least squares 2 (PLS2) regression coefficients, & PLS2 carbon predictions.

	TC vs. IC	TC vs. SOC	IC vs. SOC
Carbon Lab Measurements (**y**’s)	0.972	0.302	-0.076
PLS2 Coefficients (**X**’s)	0.997	0.938	0.911
Carbon PLS2 Predictions ($\hat{\mathbf{y}}$’s)	0.976	0.384	0.198

TC = Total soil carbon; IC = inorganic carbon; SOC = soil organic carbon.

Table 7: Prediction statistics from a simple linear regression of calibration set inorganic carbon (IC) on total carbon (TC) using lab reference measurements applied to validation set TC: a) measured by total combustion; b) predicted by laser-induced breakdown spectroscopy (LIBS) partial least squares 2 (PLS2); and c) predicted by LIBS PLS1.

TC Source	*n*	r^2	RPD	SEP (g/kg soil)	PRL	SB (%)	NU (%)	LC (%)
Lab IC	79	0.86	2.6	3.4	3.2	5.3	17.2	77.5
PLS2 LIBS TC	79	0.64	1.6	5.4	5.2	0.1	31.7	68.2
PLS1 LIBS TC	79	0.35	1.2	7.7	7.4	0.4	28.8	70.8

RPD = residual product differential, SEP = standard error of prediction, PRL = prediction relative to laboratory error (SEP/SEL), SB = squared bias reported as percent of mean squared deviation (% MSD), NU = non-unity (% MSD), LC = lack of correlation (% MSD)

Figure 14 shows the effect of spectrally averaging multiple LIBS interrogation points per sample on TC, IC, and SOC predictions. Spectrally averaging up to 9 interrogation points decreased TC SEP by 6.3% and IC SEP by 3.6% relative to using a single interrogation point; however, SEP was unchanged for SOC. Values for SEP ranged from 6.0 – 7.0 g C kg^{-1} soil for TC, 5.3 – 6.2 g C kg^{-1} soil for IC and 3.1 – 3.7 g C kg^{-1} soil for SOC. Only slight accuracy improvement was observed for spectrally averaging more than five LIBS interrogation points per interrogation area (Figure 14).

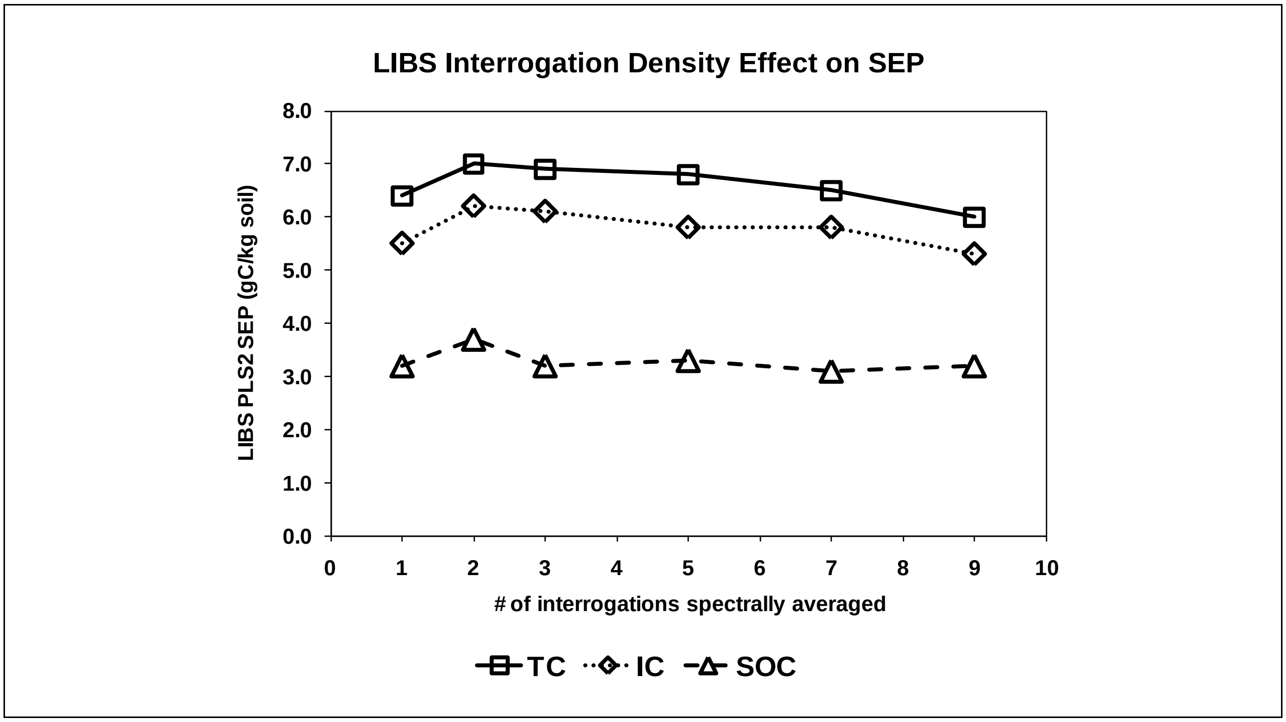

Figure 14. Predictive accuracy response, as indicated by the standard error of prediction (SEP), to spectrally averaging multiple LIBS interrogation points for soil total carbon (TC), inorganic carbon (IC), and organic carbon (SOC) determination.

3.5 *Discussion*

Realistic evaluation of model accuracy and transferability requires testing with truly independent validation samples (Brown et al., 2005; Dardenne et al., 2000), a requirement that has rarely been met for published LIBS soil characterization studies. Of the studies referenced previously, only four report some type of model validation. Cremers et al. (2001) collected soil samples at 0-10 and 10-20 cm depths from three farm fields in Colorado, used 12 soil samples for model calibration, and tested the model using 8 samples from the Colorado farm fields and 10 New Mexico forest soil samples. While the 10 New Mexico forest samples were independent of calibration samples, geographic independence of the Colorado validation samples was unclear. Researchers could not compare validation statistics as it was unclear how model “accuracy of 3-14%” was determined and r^2 was not reported. Ebinger et al. (2003) randomly selected six soil samples from a dataset of 18 samples from three farm fields in Colorado for calibration and then predicted soil carbon for the remaining 12 samples (SEP and RMSE not reported). Martin et al. (2010) reported full cross-validation accuracy for a LIBS PLS1 TC calibration (SEP = 3.1g C kg^{-1}) using soil samples from a single soil series in each of three states–not a geographically independent validation given that the samples were not selected to be geographically independent (Brown et al., 2005; Brown et al., 2006; Dardenne et al., 2000). Lastly, Belkov et al. (2009) calibrated combined- and double-pulse LIBS to “homemade certified standards”, then used those calibrations to predict soil carbon in 15 soil samples truly independent from calibration samples with RMSE = 130 and 200 g C kg^{-1}, respectively.

In this study, the research team was not able to *explain* SOC variance using LIBS ($r^2 = 0.22$, RPD = 1.1) even though the LIBS *prediction error* (SEP) for SOC was less than that for TC and IC (Table 3). Assuming a constant prediction error, r^2 and RPD will both increase with greater variation in the target variable. For the experiment reported in this paper, researchers estimate the RPD performance statistic *assuming* the standard error associated with the LIBS measurement is exactly equal to the standard error of the laboratory reference measurement (SEL):

$$SEP = \sqrt{SEL^2 + SE_{LIBS}^2}$$, where SE_{LIBS} is the LIBS measurement error; therefore

$$RPD = \frac{\sigma_{SOC}}{SEP} = \frac{\sigma_{SOC}}{\sqrt{SEL^2 + SE_{LIBS}^2}} = \frac{(3.47\ \mathrm{g\ kg^{-1}})}{\sqrt{2} \times (1.37\ \mathrm{g\ kg^{-1}})} = 1.8$$, given SE_{LIBS} = SEL.

Even assuming that LIBS interrogation of intact soil material is as precise as standard laboratory analyses of prepared samples (a very optimistic assumption), the best result researchers could obtain would be a semi-quantitative (RPD < 2) explanation of SOC variability due to low SOC variance. If SE_{LIBS} equaled 2-4 times SEL, the RPD statistic would decline to a corresponding range of 1.5 to 1.1. The team can conclude that the LIBS analysis of intact soil has a measurement error greater than or equal to four times the standard error of the laboratory reference measurement (SEL). Beyond this limited conclusion, they were not able to fully evaluate the capability of LIBS to predict SOC due to the low SOC variance of the soils interrogated in this study.

Several other factors could have contributed to *calibration* statistics being lower than previous LIBS soil carbon studies. These factors include (a) greater soil compositional diversity; (b) less variance/range in the target variables; and (c) LIBS measurements taken on *in situ* vs. prepared soils. Soil compositional diversity and range in target variables were likely not significant factors leading to the low calibration results. Previous studies have calibrated LIBS to soils of comparable or greater diversity. For example, Cremers et al. (2001) and Ebinger et al. (2003) calibrated LIBS to samples from three farm fields in Colorado, and Martin et al. (2010) calibrated LIBS using soil samples from Illinois, Michigan, and North Dakota farm fields. Furthermore, the range in measured TC and IC content (55 and 45 gC kg^{-1}, respectively) was similar to soil carbon content ranges reported by others [32 to 57 gC kg^{-1}, (Bel'kov et al., 2008; Belkov et al., 2009; Cremers et al., 2001; Ebinger et al., 2003; Martin et al., 2003; Martin et al., 2010; Martin et al., 2007)]. Researchers therefore suspect that difficulties associated with interrogation of intact cores contributed to relatively low calibration and validation statistics.

Interrogating undisturbed soil core surfaces presents several challenges for acquiring high quality LIBS data compared to prepared samples. Soil samples prepared as pressed “pucks”, as is the standard method for LIBS soils research, provide a target virtually free of moisture with consistent sample density and a smooth surface for interrogation. Sensitivity of LIBS is reduced by sample moisture because a portion of the laser energy is used to vaporize water rather than ablate the targeted sample (Castle et al., 1998). Similarly, low sample density can cause poor ablation efficiency and result in up to 40% variation in LIBS signal, with localized complete signal loss, when analyzing non-pressed soil samples (Bousquet et al., 2008). Uneven sample surfaces can cause the distance from the laser focal point to the sample to vary among and within interrogation areas, with naturally occurring fractures and macropores potentially obstructing collection optics. Finally, prepared samples are free of plant litter and roots that would likely influence soil carbon LIBS calibrations. The results of this experiment suggest that *in situ* LIBS soil C calibrations will be substantially degraded relative to the interrogation of prepared samples.

Soil heterogeneity is another factor that has the potential to reduce the accuracy of LIBS. Soils have sub-cm variation that is evident even with the naked eye, and some LIBS experts expect this short-range soil variation to cause LIBS signal to vary from one location to another, thus reducing the accuracy of the method (Bousquet et al., 2008). This is of particular concern for direct *in situ* application of LIBS undisturbed soil core surfaces. Somewhat surprisingly, the effect of fine-scale (i.e. sub-cm) soil heterogeneity on LIBS prediction was lower than expected. The results from spectrally averaging 1, 2, 3, 5, 7, and 9 interrogation points per depth suggest that short-range soil variation may not be as troublesome of an issue, for no more than 5 interrogations per sample were required for representative *in situ* measurements of small interrogation volumes (~3 cm^3). This finding also implies that interrogation areas were relatively homogenous with respect to carbon content and the soil matrix. Each LIBS interrogation point was the spectral average of 50 laser shots that probed a volume of ~ $8x10^{-5}$ cm^3. Averaging 5 interrogation points equated to ~0.02% of the total sample volume, a very small proportion. Five LIBS interrogations may have been representative of the small sampled volume. However, researchers have yet to determine how representative those small sampled volumes were of the associated core depth increments (e.g. 0-10 cm).

Important PLS2 regression coefficients suggested stoichiometric relationships were used to predict both IC and TC. The strong C emission line at 247.8 nm was an important predictor for TC, IC, and SOC, as would be expected given that C is a component of all three measures of soil carbon. The fact that Mg and Ca emission lines were also important predictors for *both* TC and IC indicates that these PLS2 calibrations were built on carbonate-associated cation content (Sposito, 2008) as well as elemental carbon. For IC this is not surprising, but for TC this highlights the value in using a chemometric approach that extracts information from indirect relationships as opposed to a single-peak calibration. A silicon emission line at 251.6 nm was a significant predictor of TC and IC, despite the fact that Si is not an important component of either organic matter or carbonates and has no stoichiometric relationship with TC or IC. The team attributes the importance of the Si emission to chemical matrix effects associated with the LIBS method. Silicon (Si) is a fundamental backbone element found in silicate soil minerals and likely serves as a “scaling” factor, adjusting predictions relative to the overall amplitude of each spectrum.

While TC and IC were highly correlated in this study (r = 0.972, Table 6), it would be incorrect to infer that LIBS IC predictions were based solely on correlations with elemental or total carbon (TC). While LIBS is fundamentally an elemental technique, as discussed above carbonate-associated stoichiometric relationships clearly played an important role in LIBS PLS2 calibrations for TC and IC. Validation predictions of IC using LIBS PLS2 models were better, though only slightly, than those obtained indirectly from a regression of calibration IC on TC applied to validation LIBS PLS2-predicted TC. The relatively small difference in predictive accuracies between these direct and indirect predictions should not be surprising given that PLS2 calibrations explicitly use response variable correlations. In contrast, the TC PLS1 calibration that did not use the TC-IC correlation yielded TC validation predictions that were substantially less correlated with validation lab-measured IC (r^2 = 0.35 vs. 0.64 for PLS1 and PLS2, respectively). While it is impossible to fully evaluate the potential of LIBS PLS models to distinguish IC and TC given their degree of correlation in this study, the results do suggest that stoichiometric relationships may support such a chemometric separation.

Measuring LIBS emissions beyond the 200-300 nm spectral range recorded with the Ocean Optics HR2000 spectrometer in this study could improve all predictions, particularly for SOC (Clegg et al. 2009, Tucker et al 2010). Previously, researchers focused on univariate calibrations based largely on the C emission at 247.8 nm (Ebinger et al., 2003; Eppler et al., 1996; Martin et al., 2003; Salle et al., 2006) for which the 200 – 300 nm spectral range was adequate. A key finding of this study was that PLS2 models utilize a larger number of peaks, particularly for elements associated with soil C (e.g. Ca and Mg with carbonates) and the soil matrix as a whole (e.g. Si). The spectral range employed in this study did not capture emissions from several other elements associated with organic C such as H (656.27 and 656.29 nm), N (742.3, 744.2, and 746.8nm) and O (777.4nm) and soil inorganic C such as primary Ca emissions (315.89, 317.93, 393.37, 396.85 and 422.60 nm). Clegg et al. (2009) observed that the best calibration models for predicting elemental composition of igneous rocks were generated when the entire LIBS spectrum (200 – 800 nm) was used in PLS2. To date, full-spectrum LIBS PLS calibrations have not been reported for soil carbon calibrations.

3.6 *Conclusions*

To the best of our knowledge, this study represents the first rigorous validation of LIBS calibrations for field-scale characterization of soil carbon using intact soil cores without soil pre-treatment. It also reports the first attempt at differentiating IC from TC with LIBS-derived stoichiometry. Using LIBS with a spectral range of 200-300 nm and employing partial least squares 2 regression (PLS2) modeling, the research team achieved semi-quantitative validation accuracies for total carbon (TC) (r^2 = 0.63, RPD = 1.6, SEP = 6.0 g kg^{-1}, SEL = 0.9 g kg^{-1}) and inorganic carbon (IC) (r^2 = 0.66, RPD=1.7, SEP = 5.3 g kg^{-1}, SEL = 1.03 g kg^{-1}). Soil organic carbon (SOC) predictions were not useful (r^2 =0.22, RPD=1.1) despite a low prediction error (SEP = 3.2 g kg^{-1}). Low SOC variability (σ = 3.47 g kg^{-1}) precluded a rigorous evaluation of LIBS capability to determine SOC. Calibration statistics from this study were substantially degraded relative to previously published research, a result attributed to the challenges of interrogating intact soil surfaces vs. prepared soil samples.

A key finding of this study was that LIBS was able to distinguish IC from TC. Regression coefficients from PLS2 models suggested that calibrations utilized stoichiometric relationships between C and elements related to C in the soil matrix. The primary C emission (247.8 nm) was an important predictor for TC and IC. Additionally, Mg (279.55-280.4 nm, 285.26 nm), Si (251.6 nm, 288.1 nm), and weak secondary Ca emissions (210.2, 211.3, and 220.9 nm), were important predictors for TC and IC. The relatively narrow spectral range (200 – 300 nm) of the LIBS spectrum recorded in this study, however, omitted primary emissions from elements related to soil carbon, including primary O, H, and N. Increasing the spectral range to the full LIBS spectrum (200 – 800 nm) could increase predictive accuracies for *in situ* measurement of both inorganic and organic C.

Additional improvements in LIBS soil characterization might come from two directions. First, the LANL-LIBS-CS instrument design could be improved to compensate for uneven soil core surfaces. Secondly, researchers need to acquire soil cores from a more diverse set of locations to evaluate the potential of developing regional LIBS calibrations. Ultimately, the team would like to see LIBS and VisNIR used in a complementary fashion—simultaneously interrogating soils *in situ* for soil mineralogy and elemental composition.

4. Full spectrum laser-induced breakdown spectroscopy (LIBS) for intact soil core total, inorganic and organic carbon measurement. (unpublished data)

4.1 *Material and Methods*

4.1.1 Instrumentation

Building on the results from previous intact soil core interrogations (Bricklemyer et al., Accepted), two additional experiments using the LANL-LIBS-CS instrument were conducted in 2008 and 2009 that included modification from the original design. In 2008, the LANL-LIBS-CS instrument underwent design modifications to address two of the issues reported to have contributed to poor prediction accuracies in 2007 (Bricklemyer et al., Accepted). The first modification was to increase the focal length from 50 mm to 500 mm. Increasing the focal length of the laser reduces the angle at which the laser energy is redirected. This has the benefit of maintaining more consistent laser energy when the distance from the laser focal point to the sample varies across uneven surfaces. The second modification was to employ an Echelle type spectrometer that captured the full LIBS spectrum (200-800 nm). Results from 2007 suggested that increasing the spectral range to the full LIBS spectrum (200 – 800 nm) could increase predictive accuracies for *in situ* measurement of both inorganic and organic C. The full spectrum would provide spectral emissions for all elements related to inorganic and organic soil C (ex. C, Ca, Mg, O, N, P, K, and H). This second LIBS experiment would test the hypothesis that by providing emissions of all soil elements, stoichiometric relationships suggested previously would increase prediction accuracies for soil IC and OC.

Further testing of the Echelle spectrometer at Los Alamos National Lab, after completion of the 2008 LANL-LIBS-CS experiment, called the quality of the 2008 spectral data into question. Therefore, the experiment was repeated in 2009 where the Echelle spectrometer was replaced by a series of three Ocean Optics spectrometers (identical to the spectrometer used in 2007) that spanned the full LIBS spectrum (225-925 nm).

4.1.2 Study area, soil sampling, core interrogation, soil analysis, spectral model evaluation

See section 3.2 above for details.

4.1.3 Spectral model calibration and validation

Regional LIBS models were constructed for TC, IC, and SOC employing a whole-field out cross-validation approach whereby each field was held out in turn for model validation, with data from the remaining fields used for model calibration. For example, data for field #1 was held out from the calibration set and data from the remaining fields were used to calibrate PLS2 models to independently predict TC, IC, and SOC for field #1. Then, data from field #2 was held out and new PLS2 models were derived using the remaining data. The systematic removal of holding out data from whole fields, calibrating new PLS2 models using data from the remaining fields, and predicting TC, IC, and SOC for the held out field continued until each of the fields had been predicted.

4.2 ***Results and Discussion***

4.2.1 Exploratory data analysis

Summary statistics for Lab[4] TC, IC, and SOC are presented in Tables 8 and 9. For the samples in the 2008 study (Table 8), TC values did not exceed 52 g kg^{-1}, IC values were less than 43 g kg^{-1}, and SOC values never exceeded 32 g kg^{-1}. Concentrations of IC were most variable (σ = 8.25, CV = 150%) followed by TC and SOC (σ = 8.56 and 4.26; CV = 57% and 45%, respectively). The SEL for laboratory measurements were estimated at 0.28, 0.44, and 0.55 g kg^{-1} for TC, IC, and SOC, respectively. Measured SOC was least variable, where σ_{SOC} was just 7.7 times SEL_{SOC}. Variability in SOC was substantially less than IC with highest concentrations occurring in A horizons and diminished with depth. Variation in IC was a function of pedogenesis where IC was not present or present in low concentrations in A horizons (top 20 cm) and increased sharply in B horizons, the majority of which occurred below 20 cm (Figure 10, above).

Similar patterns in the summary statistics were found in 2009 TC, IC, and SOC laboratory measured values (Table 9). Total carbon values did not exceed 45 g kg^{-1}, IC values were less than 32 g kg^{-1}, and SOC values never exceeded 25 g kg^{-1}. Concentrations of IC were again most variable (σ = 7.56 g kg^{-1}, CV = 131%) followed by TC and SOC (σ = 8.18 and 3.39 g kg^{-1}; CV = 55% and 37%, respectively). The SEL for laboratory measurements were estimated at 0.11, 0.60, and 0.60 g kg^{-1} for TC, IC, and SOC, respectively. Measured SOC was least variable, where σ_{SOC} was just 5.7 times SEL_{SOC}. Variation in SOC was substantially less than IC with highest concentrations occurring in A horizons and diminished with depth. Variation in IC was again a function of pedogenesis where IC was not present or present in low concentrations in A horizons (top 20 cm) and increased sharply in B horizons, the majority of which occurred below 20 cm (Figure 10, above).

---

[4] "Lab" refers to laboratory measurements and "LIBS" refers to a LIBS PLS2 calibration, unless otherwise indicated.

Table 8: Laboratory-measured soil total, inorganic, and organic carbon summary statistics for 80 intact soil cores from 7 wheat fields in north central MT & the Cook Agronomy farm near Pullman, WA, 2008.

	Soil Total C[†]	Soil Inorganic C[‡]	Soil Organic C
	---------------------------- g kg $^{-1}$	------------------------------	
Median	12.33	0.40	8.47
Min	4.11	0.00	2.08
Max	51.96	42.79	31.76
Mean	14.91	5.51	9.40
σ	8.56	8.25	4.26
CV	57%	150%	45%
n	316	316	316

σ = standard deviation; CV = coefficient of variation

[†] dry combustion (Leco TruSpec, Leco Corp., St. Joseph, MI, USA)

[‡] modified pressure calcimeter (Sherrod et al., 2002)

Table 9: Laboratory-measured soil total, inorganic and organic carbon summary statistics for 60 intact soil cores from 6 wheat fields in north central MT, 2009.

	Soil Total C[†]	Soil Inorganic C[‡]	Soil Organic C
	------------------------- g kg $^{-1}$	----------------------------	
Median	12.74	0.58	8.78
Min	3.18	0.00	0.22
Max	44.81	31.36	24.56
Mean	14.98	5.77	9.21
σ	8.18	7.56	3.39
CV	55%	131%	37%
n	240	240	240

σ = standard deviation; CV = coefficient of variation

[†] dry combustion (Leco TruSpec, Leco Corp., St. Joseph, MI, USA)

[‡] modified pressure calcimeter (Sherrod et al., 2002)

4.2.2 LIBS chemometric models

For the 2008 whole-field out cross validation, the LIBS PLS2 analysis best predicted inorganic carbon (r^2 = 0.75, RPD = 2.0; Table 10) with total carbon explanation somewhat lower (r^2 = 0.62, RPD = 1.6; Table 10), and virtually no explanatory capability for SOC (r^2=0.03, RPD= 1.0, Table 10). Validation SEP values were 5.3 g TC kg^{-1} soil, 4.2 g IC kg^{-1} soil, and 4.4 g SOC kg^{-1} soil in absolute units, or relative to the lab precision PRL = 19, 9.4 and 8.0 for TC, IC and SOC, respectively. Partitioning MSD into three components, the research team found the LIBS TC model had no bias (SB=0.0%) and marginal non-unity (NU=34.1%) (Table 10). The LIBS IC model also had low bias (SB=1.7%); however, non-unity was lower (NU=19.7%). Bias remained low for the LIBS SOC model (SB=0.3%), but non-unity greatly increased (NU=75.8%; Table 10). Lack of correlation accounted for the majority of MSD for the LIBS TC and IC models (65.8% and 78.6%, respectively) but only 23.9% of MSD for the SOC model (Table 10).

2009 Whole-field out cross validation results followed very similar patterns as 2008 results. Inorganic carbon was again best predicted (r^2 = 0.78, RPD = 2.1; Table 10) with total carbon explanation somewhat lower (r^2 = 0.59, RPD = 1.5; Table 10), and virtually no explanatory capability for SOC (r^2=0.02, RPD= 0.8, Table 10). Validation SEP values were similar to 2008 results with 5.43 g TC kg^{-1} soil, 3.62 g IC kg^{-1} soil, and 3.99 g SOC kg^{-1} soil in absolute units, or relative to the lab precision PRL = 49.3, 5.9 and 6.6 for TC, IC and SOC, respectively. It is important to note that differences in PRL from experiment to experiment must be compared with caution as SEL for TC, IC, and SOC varied by year. Partitioning MSD into three components, we found the LIBS TC model had no bias (SB=0.0%) and marginal non-unity (NU=16.6%, Table 10). The LIBS IC model also had low bias (SB=1.9%); however, non-unity was somewhat reduced (NU=8.8%). Bias was negligible for the LIBS SOC model (SB=0.0%), but non-unity greatly increased (NU=82.4%; Table 10). Lack of correlation accounted for the majority of MSD for the LIBS TC and IC models (83.3% and 89.4%, respectively) but only 17.6% of MSD for the SOC model (Table 10).

Table 10: Soil total carbon (TC), inorganic carbon (IC), and organic carbon (SOC) whole-field out independent cross validation statistics for 2008 & 2009 laser-induced breakdown spectroscopy (LIBS) partial least squares 2 (PLS2) regression models.

Model	*n*	r^2	RPD	SEP	PRL	SB	NU	LC
				g kg^{-1}		--------- % MSD ---------		
2008 TC	316	0.62	1.6	5.33	19.0	0.0	34.1	65.8
2008 IC	316	0.75	2.0	4.15	9.4	1.7	19.7	78.6
2008 SOC	316	0.03	1.0	4.43	8.0	0.3	75.8	23.9
2009 TC	240	0.59	1.5	5.43	49.3	0.0	16.6	83.3
2009 IC	240	0.78	2.1	3.62	5.9	1.9	8.8	89.4
2009 SOC	240	0.02	0.8	3.99	6.6	0.0	82.4	17.6

RPD = residual product differential; SEP = standard error of prediction; PRL = prediction relative to laboratory error (SEP/SEL); SB = squared bias reported as percent of mean squared deviation (% MSD); NU = non-unity (% MSD); LC = lack of correlation (% MSD).

4.3 *Conclusions*

Preliminary results for increasing the spectral range of LIBS to the full 200-800 nm in order to capture elemental emissions of all elements in a given soil sample suggest marginal gains in prediction accuracy for inorganic carbon, but no gains for predicting total and organic carbon. Poor SOC predictions are likely a result of 1) the lack of a consistent/definable molecular composition of SOC, 2) relatively little variation in SOC across field sites, and 3) inorganic carbon being the primary form of soil C, particularly for Montana samples. Continued exploration into alternative data reduction and statistical modeling techniques continues in an effort to increase prediction accuracy, model parsimony, and computational efficiency.

5. Analysis and Conclusions

5.1 *Introduction*

Visible-near infrared diffuse reflectance spectroscopy (VisNIR) and LIBS are fundamentally different, yet complementary spectroscopic techniques. The LIBS technique is an emerging elemental analysis method that can quickly determine elemental composition of heterogeneous material, whereas VisNIR is based on the fundamentals of energy absorption by molecular bond vibrations. Taking advantage of the fundamental spectroscopic differences in these techniques, by combining spectral information obtained by VisNIR and LIBS sensors, may provide more accurate, robust, and spatially transferable soil C determination than individual sensors currently permit.

5.2 *Material and Methods*

VisNIR spectra were collected concurrently with LIBS analysis on intact soil cores (described previously). All VisNIR spectra were checked for errors and smoothed using cubic smoothing splines with 1st derivatives extracted in 10 nm increments directly from spline fits, following methods outlined in Brown et al. (2006). Model calibration and validation strategies were consistent with those described for LIBS in 2007 and 2008.

5.3 *Results and Discussion*

5.3.1 VisNIR chemometric models

Preliminary PLS regression results from 2007 and 2008 experiments found that VisNIR TC prediction accuracy was similar to LIBS; IC prediction accuracy was markedly lower than LIBS; however VisNIR outperformed LIBS for predicting SOC (Table 11). The degraded explanatory power of the 2008 VisNIR SOC model ($r^2 = 0.33$) compared to the 2007 VisNIR model was likely a function of 1) a more rigorous whole-field out cross validation used for 2008 models compared to the random whole-core out cross validation used in 2007, and 2) additional core samples from the WSU Cook Agronomy farm. The Cook farm soils are substantially different from the MT soils in several key characteristics. Cook farm soils are wind deposited, have lower pH, generally lower clay content, higher SOC content, and negligible IC content compared to MT soils.

Table 11: Soil total carbon (TC), inorganic carbon (IC), and organic carbon (SOC) independent validation statistics for 2008 and 2009 Visible-near infrared spectroscopy (VisNIR) partial least squares regression models.

Model	*N*	r^2	RPD	SEP	SB	NU	LC
				g kg^{-1}	------	% MSD	------
2007 VisNIR TC	80	0.61	1.5	6.0	1.7	19.7	78.6
2007 VisNIR IC	80	0.65	1.7	5.4	0.5	29.0	70.5
2007 VisNIR SOC	80	0.51	1.3	2.9	0.3	8.8	90.9
2008 VisNIR TC	316	0.55	1.4	5.9	1.0	20.3	79.2
2008 VisNIR IC	316	0.69	1.8	4.6	1.0	28.2	71.3
2008 VisNIR SOC	316	0.33	1.2	3.5	1.8	42.8	55.4

RPD = residual product differential; SEP = standard error of prediction; SB = squared bias reported as percent of mean squared deviation (% MSD); NU = non-unity (% MSD); LC = lack of correlation (% MSD).

5.3.2 Combined LIBS-VisNIR chemometric models

Preliminary exploration of combining VisNIR and LIBS spectra was completed using 2007 data and PLS regression. Combining 2007 LIBS and VisNIR data slightly improved 2007 TC predictions, but did not improve IC or SOC predictions compared to individual sensor results. The combined LIBS-VisNIR TC model was the best TC predictor with a slightly higher r^2 value than LIBS alone and deviated less from the 1:1 line (See NU, Table 5 and Table 12). The LIBS-VsiNIR IC model (r^2 = 0.68) performed similar to both 2007 VisNIR and LIBS models alone (r^2 = 0.65, and r^2 = 0.66, respectively). Conversely, the LIBS-VisNIR SOC model (r^2 = 0.40) improved predictions compared to LIBS alone (r^2 = 0.22), but did not perform as well as VisNIR alone (r^2 = 0.51).

Table 12: Partial Least Squares model statistics for combined VisNIR-LIBs for predicting total carbon (TC), inorganic carbon (IC), & soil organic carbon (SOC) in intact soil cores.

Model	*N*	r^2	RPD	SEP g kg^{-1}	SB† (%)	NU† (%)	LC† (%)
LIBS-VisNIR TC	80	0.69	1.7	5.5	0.2	5.9	93.9
LIBS-VisNIR IC	80	0.68	1.7	5.1	0.5	27.7	71.8
LIBS-VisNIR SOC	80	0.40	1.0	3.5	0.8	7.1	92.1

RPD = rational product differential, SEP= standard error of prediction, SB = squared bias, NU= non-unity, LC = lack of correlation

† percent of mean squared deviation (MSD)

Researchers suspect that the 2046 LIBS wavelength predictors collected in 2007 are diluting, if not dominating, the PLS analysis and overwhelming the 216 VisNIR wavelength predictors. Although PLS is similar to principal component analysis in that both use statistical rotations of predictor variables, PLS rotates both the *X* and *Y* variables "semi-orthogonally" relative to the response variable to maximize predictive capabilities (Brown et al., 2006; Wold et al., 2001b). One would expect that the combined LIBS-VNIR data would perform at least as well as the best individual sensor results. The results suggest that PLS is sensitive to large predictor datasets. Reduction of LIBS spectra poses a difficult challenge to overcome. High spectral resolution is required for elemental analysis, because elemental emission features are often distinct peaks that occur at very specific wavelengths. For example, the carbon emission peak occurs specifically at 247.8 nm. This precludes a strict re-sampling algorithm (commonly used with VisNIR) that reduces the spectral resolution below a yet undetermined critical resolution required to spectrally resolve elemental features. Other approaches may be able to use the data at the current resolution to build better predictive spectral models for TC, IC, and SOC, regardless of the number of predictor variables. One option may involve using principal components analysis of individual sensor data, then combining those principal components in a PLS regression environment (Hamalainen and Albano, 1992). Alternative multivariate statistical and data mining approaches continue to be tested.

5.4 *Conclusion*

VisNIR and LIBS spectroscopy have the potential to fill the growing need for rapid, accurate, and inexpensive methods to measure, and verify soil organic carbon change. These fundamentally different techniques performed individually mostly as expected; however, the combined LIBS-VisNIR data array did not consistently improve predictive accuracies over individual sensors. Data reduction methods are needed for LIBS spectral data, and other statistical and data mining algorithms need to be tested before integrating these sensors into an operational field unit. A field unit with a LIBS-VisNIR array mounted in a soil penetrometer, when fully operational, would allow for rapid soil profile characterization and mapping at field and landscape scales. Instrumentation modifications to the LIBS sensor will be to extend the focal length of the laser to reduce the effects of uneven core surfaces on spectral intensity and variability. To the best of our knowledge this is the first attempt to combine these two proximal sensing techniques.

6. References

Adamchuk, V.I., J.W. Hummel, M.T. Morgan, and S.K. Upadhyaya. 2004. On-the-go soil sensors for precision agriculture. Comput. Electron. Agric. 44:71-91.

Andraos, J. 1996. On the propagation of statistical errors for a function of several variables. J. Chem. Educ. 73:150-154.

Bel'kov, M.V., V.S. Burakov, V.V. Kiris, V.A. Kononov, S.N. Raikov, V.N. Reshetnikov, and N.V. Tarasenkoa. 2008. Determination of Carbon in Soil by Laser Spectral Analysis. J. Appl. Spectrosc. 75:275-282.

Belkov, M.V., V.S. Burakov, A. De Giacomo, V.V. Kiris, S.N. Raikov, and N.V. Tarasenko. 2009. Comparison of two laser-induced breakdown spectroscopy techniques for total carbon measurement in soils. Spectroc. Acta Pt. B-Atom. Spectr. 64:899-904.

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

Ben-Dor, E., D. Heller, and A. Chudnovsky. 2008. A novel method of classifying soil profiles in the field using optical means. Soil Sci. Soc. Am. J. 72:1113-1123.

Bousquet, B., J.B. Sirven, and L. Canioni. 2007. Towards quantitative laser-induced breakdown spectroscopy analysis of soil samples. Spectroc. Acta Pt. B-Atom. Spectr. 62:1582-1589.

Bousquet, B., G. Travaille, A. Ismael, L. Canioni, K.M.L. Pierres, E. Brasseur, S. Roy, I. le Hecho, M. Larregieu, S. Tellier, M. Potin-Gauder, T. Boriachon, P. Wazen, A. Diard, and S. Belbeze. 2008. Development of a mobile system based on laser-induced breakdown spectroscopy and dedicated to in situ analysis of polluted soils. Spectroc. Acta Pt. B-Atom. Spectr. 63:1085-1090.

Box, G.E.P., and D.R. Cox. 1964. An analysis of transformations (with discussion). J. Royal Stat. Soc. B. 26:211–252.

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

Bricklemyer, R.S., D.J. Brown, J.E. Barefield, and S. Clegg. Accepted. Intact soil core total, inorganic and organic carbon measurement using laser-induced breakdown spectroscopy (LIBS). Soil Sci. Soc. Am. J.

Bricklemyer, R.S., P.R. Miller, K. Paustian, T. Keck, G.A. Nielsen, and J.M. Antle. 2005. Soil organic carbon variability and sampling optimization in Montana dryland wheat fields. J Soil Water Conserv. 60:42-51.

Bricklemyer, R.S., P.R. Miller, P.J. Turk, K. Paustian, T. Keck, and G.A. Nielsen. 2007. Sensitivity of the century model to scale-related soil texture variability. Soil Sci. Soc. Am. J. 71:784-792.

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

Brown, D.J., K.D. Shepherd, M.G. Walsh, M.D. Mays, and T.G. Reinsch. 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma. 132:273-290.

Castle, B.C., K. Talabardon, B.W. Smith, and J.D. Winefordner. 1998. Variables influencing the precision of laser-induced breakdown spectroscopy measurements. Appl. Spectrosc. 52:649-657.

Christy, C.D. 2008. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. Comput. Electron. Agric. 61:10-19.

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

Clegg, S.M., E. Sklute, M.D. Dyar, J.E. Barefield, and R.C. Wiens. 2009. Multivariate analysis of remote laser-induced breakdown spectroscopy spectra using partial least squares, principal component analysis, and related techniques. Spectrochim. Acta Part B. 64:79-88.

Council, N.R. 1999. A U.S. carbon cycle science plan. U.S. Global Change Res. Program, Washington D.C.

Cremers, D.A., and L.J. Radziemski. 2006. Handbook of laser-induced breakdown spectroscopy. First ed. John Wiley & Sons, West Sussex, England.

Cremers, D.A., M.H. Ebinger, D.D. Breshears, P.J. Unkefer, S.A. Kammerdiener, M.J. Ferris, K.M. Catlett, and J.R. Brown. 2001. Measuring total soil carbon with laser-induced breakdown spectroscopy (LIBS). J. of Environ. Qual. 30:2202-2206.

Dardenne, P., G. Sinnaeve, and V. Baeten. 2000. Multivariate calibration and chemometrics for near infrared spectroscopy: which method? J. Near Infrared Spectrosc. 8:229-237.

Dudragne, L., P. Adam, and J. Amouroux. 1998. Time-resolved laser-induced breakdown spectroscopy: Application for qualitative and quantitative detection of fluorine, chlorine, sulfur, and carbon in air. Appl. Spec. 52:1321-1327.

Dungan, J.L., J.N. Perry, M.R.T. Dale, P. Legendre, S. Citron-Pousty, M.J. Fortin, A. Jakomulska, M. Miriti, and M.S. Rosenberg. 2002. A balanced view of scale in spatial statistical analysis. Ecography. 25:626-640.

Ebinger, M.H., M.L. Norfleet, D.D. Breshears, D.A. Cremers, M.J. Ferris, P.J. Unkefer, M.S. Lamb, K.L. Goddard, and C.W. Meyer. 2003. Extending the applicability of laser-induced breakdown spectroscopy for total soil carbon measurement. Soil Sci. Soc. Am. J. 67:1616-1619.

Eppler, A.S., D.A. Cremers, D.D. Hickmott, M.J. Ferris, and A.C. Koskelo. 1996. Matrix effects in the detection of Pb and Ba in soils using laser-induced breakdown spectroscopy. Appl. Spec. 50:1175-1181.

Ferreira, E.C., D.M.B.P. Milori, E.J. Ferreira, R.M. Da Silva, and L. Martin-Neto. 2008. Artificial neural network for Cu quantitative determination in soil using a portable Laser Induced Breakdown Spectroscopy system. Spectroc. Acta Pt. B-Atom. Spectr. 63:1216-1220.

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

Gee, G.W., and J.W. Bauder. 1986. Particle-size Analysis, p. 383-399, *In* A. Klute, ed. Methods of Soil Analysis. Part 1. 2nd ed. Agronomy.

Gehl, R.J., and C.W. Rice. 2007. Emerging technologies for in situ measurement of soil carbon. Clim. Change. 80:43-54.

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

Gornushkin, S.I., I.B. Gornushkin, J.M. Anzano, B.W. Smith, and J.D. Winefordner. 2002. Effective normalization technique for correction of matrix effects in laser-induced breakdown spectroscopy detection of magnesium in powdered samples. Appl. Spec. 56:433-436.

Häkkänen, H.J., and J.E.I. Korppi-Tommola. 1998. Laser-induced plasma emission spectrometric study of pigments and binders in paper coatings: matrix effects. Anal. Chem. 70:4724 - 4729.

Hamalainen, M., and C. Albano. 1992. Principal variations in the chemical composition of peat: Predictive peat scales based on multivariate strategies. J. Chemometrics. 6:41-56.

Harris, R.D., D.A. Cremers, M.H. Ebinger, and B.K. Bluhm. 2004. Determination of nitrogen in sand using laser-induced breakdown spectroscopy. Appl. Spectrosc. 58:770-775.

Hunt, G.R. 1982. Spectroscopic properties of rocks and minerals, p. pp. 295-385, *In* R. S. Carmichael, ed. Handbook of physical properties of rocks. CRC Press, Boca Raton, FL.

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

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

Kaufman, L., and P.J. Rousseeuw. 1990. Finding groups in data: An introduction to cluster analysis John Wiley & Sons, Inc.

Martin, M.Z., S.D. Wullschleger, C.T. Garten, and A.V. Palumbo. 2003. Laser-induced breakdown spectroscopy for the environmental determination of total carbon and nitrogen in soils. Appl. Opt. 42:2072-2077.

Martin, M.Z., N. Labbe, T.G. Rials, and S.D. Wullschleger. 2005. Analysis of preservative-treated wood by multivariate analysis of laser-induced breakdown spectroscopy spectra. Spectrochim. Acta, Part B: Atom. Spectrosc. 60:1179-1185.

Martin, M.Z., S.D. Wullschleger, C.T. Garten, A.V. Palumbo, and J.G. Smith. 2004. Elemental analysis of environmental and biological samples using laser-induced breakdown spectroscopy and pulsed Raman spectroscopy. J. Dispersion Sci. Technol. 25:687-694.

Martin, M.Z., N. Labbe, N. Andre, S.D. Wullschleger, R.D. Harris, and M.H. Ebinger. 2010. Novel Multivariate Analysis for Soil Carbon Measurements Using Laser-Induced Breakdown Spectroscopy. Soil Sci. Soc. Am. J. 74:87-93.

Martin, M.Z., N. Labbe, N. Andre, R. Harris, M. Ebinger, S.D. Wullschleger, and A.A. Vass. 2007. High resolution applications of laser-induced breakdown spectroscopy for environmental and forensic applications. Spectroc. Acta Pt. B-Atom. Spectr. 62:1426-1432.

McBratney, A.B., and M.J. Pringle. 1997. Spatial variability in soil—implications for precision agriculture, p. 3–31, *In* J. Stafford, ed. Precision Agriculture, Vol. 2. BIOS Scientific Publishing, Oxford, UK.

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

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

Mosier-Boss, P.A., S.H. Lieberman, and G.A. Theriault. 2002. Field demonstrations of a direct push FO-LIBS metal sensor. Environ. Sci. Technol. 36:3968-3976.

Mouazen, A.M., J. De Baerdemaeker, and H. Ramon. 2006. Effect of wavelength range on the measurement accuracy of some selected soil constituents using visual-near infrared spectroscopy. J. Near Infrared Spectrosc. 14:189-199.

R, D.C.T. 2008. R: A language and environment for statistical computing. [Online]. Available by R Foundation for Statistical Computing http://www.R-project.org.

Radziemski, L.J., and D.A. Cremers. 1989. Spectrochemical analysis using laser plasma excitation., p. 295-325, *In* L. J. Radziemski and D. A. Cremers, eds. Laser-induced plasmas and applications. Marcel Dekker, New York.

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

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

Salle, B., J.L. Lacour, P. Mauchien, P. Fichet, S. Maurice, and G. Manhes. 2006. Comparative study of different methodologies for quantitative rock analysis by laser-induced breakdown spectroscopy in a simulated Martian atmosphere. Spectroc. Acta Pt. B-Atom. Spectr. 61:301-313.

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

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

Shibusawa, S., M.Z. Li, K. Sakai, A. Sasao, H. Sato, S. Hirako, and A. Otomo. 1999. Spectrometer for real-time underground soil sensing Paper No. 993030. ASAE, St. Joseph, MI, USA.

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

Sirven, J.-B., B. Bousquet, L. Canioni, and L. Sarger. 2006. Laser-induced breakdown spectroscopy of composite samples: Comparison of advanced chemometrics methods. Anal. Chem. 78:1462-1469.

Sposito, G. 2008. The chemistry of soils. 2nd ed. Oxford University Press, New York, NY.

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

Sudduth, K.A., and J.W. Hummel. 1993b. Portable, near-Infrared Spectrophotometer for Rapid Soil Analysis. Trans. ASAE. 36:185-193.

Thompson, J.R., R.C. Wiens, J.E. Barefield, D.T. Vaniman, H.E. Newsom, and S.M. Clegg. 2006. Remote laser-induced breakdown spectroscopy analyses of Dar al Gani 476 and Zagami Martian meteorites. Journal of Geophysical Research-Planets. 111.

Vasques, G.M., S. Grunwald, and J.O. Sickman. 2008. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. Geoderma. 146:14-25.

Venables, W.N., and B.D. Ripley. 2002. Modern Applied Statistics with S. Fourth ed. Springer, New York, NY.

Viscarra Rossel, R.A., S.R. Cattle, A. Ortega, and Y. Fouad. 2009. In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy. Geoderma. 150:253-266.

Viscarra Rossel, R.A., D.J.J. Walvoort, A.B. McBratney, L.J. Janik, and J.O. Skjemstad. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma. 131:59-75.

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

Wehrens, R., and B.H. Mevik. 2007. pls: Partial least squares regression (PLSR) and principal component regression (PCR), pp. http://mevik.net/work/software/pls.html R package version 2.1-0.

Wold, S., M. Sjostrom, and L. Eriksson. 2001a. PLS-regression: a basic tool of chemometrics. Chemometrics Intellig. Lab. Syst. 58:109-130.

Wold, S., H. Trygg, A. Berglund, and H. Antti. 2001b. Some recent developments in PLS modeling. Chemometrics Intellig. Lab. Syst. 58:131-150.

Workman, J.J., Jr. 2001. Handbook of near-infrared analysis, p. 91-128., *In* D. A. Burns and E. W. Ciurczak, eds. Practical Spectroscopy; v. 27., 2nd ed. M. Dekker, New York.

Yamamoto, K.Y., D.A. Cremers, M.J. Ferris, and L.E. Foster. 1996. Detection of metals in the environment using a portable laser-induced breakdown spectroscopy instrument. Appl. Spectrosc. 50:222-233.

Yaroshchyk, P., D. Body, R.J.S. Morrison, and B.L. Chadwick. 2006. A semi-quantitative standard-less analysis method for laser-induced breakdown spectroscopy. Spectrochim. Acta, Part B: Atom. Spectrosc. 61:200-209.