Laboratory-based evaluation of optical performance for a new soil penetrometer visible
and near-infrared (VisNIR) foreoptic

Matteo Poggio¹, David J. Brown¹, Ross S. Bricklemyer²

¹ Crop and Soil Science Department, Washington State University, Pullman, WA
99164, USA

matteo.poggio@wsu.edu

dave.brown@wsu.edu

² Monsanto Company, St. Louis, MO 63167, USA

ross.s.bricklemyer@monsanto.com

***Corresponding author:**

Matteo Poggio

Washington State University

Crop and Soil Science Department

Box 641020

Pullman, WA, 99164

Phone: +1 509 715 9959

Email: matteo.poggio@wsu.edu

Abstract

In situ visible and near infrared (VisNIR) spectroscopy has been gaining interest as a proximal soil sensing technique to rapidly and inexpensively measure soil properties. For this study, our goal was to design a VisNIR soil penetrometer foreoptic that matched the optical performance of a commercial ASD (Analytical Spectral Devices Inc., Boulder, CO, USA) foreoptic and was compatible with any spectrometer having a fibre optic input (e.g. commonly used ASD field spectrometers). Space available for penetrometer optics was constrained by a several practical design requirements: (a) ability to collect *in situ* soil VisNIR reflectance and insertion force simultaneously; (b) thick probe walls strong enough for insertion into dry soil with high clay content; and (c) a narrow enough cross-sectional diameter for insertion using a standard soil coring hydraulic push system. In this paper, we evaluate the optical and chemometric performance of our design relative to a commercial ASD foreoptic under controlled conditions in a laboratory setting, in order to isolate the optical performance of the probe from complications associated with *in situ* use. Both the VisNIR penetrometer and ASD Contact Probe, attached to an ASD AgriSpec spectrometer in the lab, were used to interrogate 389 milled and pressed surface and subsoil samples taken from two US continent-scale transects. The reflectance spectra obtained from these two foreoptics were highly correlated ($r > 0.95$ for most wavelengths) and the 1st derivatives of the spectra were well correlated for all but the detector splice points ($r > 0.9$ for most wavelengths). Partial least squares regression (PLSR) was used to calibrate and validate VisNIR models predicting soil organic carbon content for both foreoptics. Using reflectance data, our VisNIR penetrometer yielded quantitative SOC predictions with standard error of prediction ($SEP = 0.19 \log[g\ kg^{-1}]$) and ratio of prediction to determination ($RPD = 2.34$) highly similar to the ASD contact probe ($SEP = 0.18$ and $RPD = 2.45$). Regression coefficients of PLS calibration models from both foreoptic were comparable and highly correlated ($r = 0.91$). These results suggest that our penetrometer foreoptic provides VisNIR reflectance spectra comparable to a widely-used commercial

foreoptic. In future work, we deploy this penetrometer in the field and evaluate the performance of this foreoptic where soil heterogeneity, smearing and moisture can impact the quality of VisNIR spectra.

Keywords: Proximal soil sensing, diffuse reflectance spectroscopy, VNIR, penetrometer design, *in situ* optical probe.

Abbreviations: VisNIR = visible near-infrared, ASD = Analytical Spectral Device Inc., SOC = soil organic carbon, RMSEP = root mean square error of prediction, SEP = standard error of prediction, LED = light emitting diode, SOM = soil organic matter, EC = electrical conductivity, SWIR = short wave infrared.

1. Introduction

Visible and Near Infrared (VisNIR) diffuse reflectance spectroscopy is an established method to rapidly and inexpensively characterize soils (Ben-Dor and Banin, 1995; O’Rourke and Holden, 2011). With lab-based interrogation, this technique has been used to estimate a number of soil properties related to mineralogy, organic matter and water (Clark et al., 1990; Gholizadeh et al., 2013), such as moisture, clay, total carbon, organic carbon, total N, and CEC (Brown et al., 2006; Demattê et al., 2004; Heinze et al., 2013; Kuang et al., 2012; Shepherd and Walsh, 2002; Viscarra Rossel et al., 2006). Nevertheless, lab-based VisNIR spectroscopy still requires samples to be extracted and transported to laboratory facilities for drying, sieving and scanning. Advancement of *in situ* VisNIR soil spectroscopy would allow more spatially intensive soil interrogation for applications such as constructing maps for precision agriculture (Adamchuk et al., 2004).

Researchers have reported generally good results from the interrogating intact soil cores to simulate *in situ* VisNIR spectroscopy and evaluate the potential for core scanning in the field. Using a commercially available Analytical Spectral Devices (ASD Inc.

Boulder, CO, USA) contact probe and spectrometer, researchers obtained similar accuracies for clay content (Waiser et al., 2007) and SOC (Morgan et al., 2009) using intact core and air-dried ground sample scans. Kusumo et al. (2008) designed a “purpose built contact probe” derived from the ASD contact probe fitting the cross section of a soil core. The use of the sensor on site with extracted soil core provided prediction of total carbon (TC) with root mean square error of prediction (RMSEP) of 1.21% over TC range of 0.26-11.21% and of total nitrogen (TN) with RMSEP = 0.07% over TN range of 0.02-1.01%. Viscarra Rossel et al. (2009) described variation of clay content, soil mineralogy and colors along different profiles using an ASD contact probe. For clay content, the authors obtained slightly more accurate predictions on site (root mean square error [RMSE] = 7.9%) compared to laboratory VisNIR interrogation (RMSE = 8.3%). Regarding the potential transfer of these findings to fully *in situ* spectroscopy, one concern is that the quality of spectra can be diminished when soil smears on the sensor window (Morgan et al., 2009; Waiser et al., 2007).

There are a number of technical challenges specific to on-the-go spectroscopy. First of all, the foreoptic, connections and spectrometer must be robust to dust, vibrations, and variations in humidity and temperature. Secondly, constant distance and angle between soil surface and on-the-go sensor is essential to collect consistent and high quality spectral signature: roughness of soil surface or presence of gravel can vary the distance reducing the quality of delivered spectra (Bricklemyer and Brown, 2010; Schirrmann et al., 2013; Shonk et al., 1991; Sudduth and Hummel, 1993a). Furthermore, on-the-go approach requires the use of array spectrometer capable of analysing all wavelengths simultaneously so that the entire spectrum captures the reflectance of a single soil view (Sudduth and Hummel, 1993a).

A number of VisNIR foreoptics have been developed for on-the-go soil spectroscopy. Shonk et al. (1991) developed an on-the-go sensor with a narrow bandwidth (40 nm) source

light emitting exclusively the red portion of electromagnetic spectrum (red LED, 660 nm peak) to minimize the effects of soil moisture and iron oxides on spectra. Laboratory test showed accurate estimations of soil organic matter [SOM] ($R^2 = 0.87 - 0.98$) over soil samples from 6 different farms with overall large amount of SOM variability (0.38-13.37%) and medium to fine soil texture. The authors reported high accuracy estimation ($R^2 = 0.84 - 0.95$) in a single field test on one of the farms, selected for the wide range of soil organic matter (1.7-6.1%) even though they highlighted the need for more comprehensive test to assess the sensor validity. Sudduth and Hummel, (1993a, 1993b) designed a real-time NIR sensor following the criteria of optimal spectral range (1650 – 2650 nm, single detector spectrophotometer) and robustness to vibration and dust. Although the authors achieved accurate laboratory estimation of SOC (SEP = 0.23%), on field test the sequential scanning approach of the spectrophotometer was unable to capture different wavelengths from unique sample area, providing unsatisfactory predictions of SOC (SEP = 0.53%) on-the-go.

Recently, four different research teams (Bricklemyer and Brown, 2010; Christy, 2008; Huang et al., 2007; Schirrmann et al., 2013) evaluated the performances of an on-the-go NIR sensor manufactured by Veris Technologies Inc. (Salinas, KS, USA). The performance of the Veris on-the-go sensor varied according to study design, soil property range and validation scheme adopted by research teams. Small sample size (single field) and less rigorous validation (leave-one-out cross validation) provided satisfactory results on estimated total carbon as reported by Huang et al. (2007) (RMSE = 0.19% over total carbon range of 0.55-2.89%). A more rigorous cross-validation scheme led to unsatisfactory estimation of soil organic matter, obtained by Schirrmann et al. 2013 (RMSEP = 0.38% with field-out cross validation over SOM range of 0.9-2.9%). Christy (2008) and Bricklemyer and Brown (2010) showed comparable RMSE for SOC estimation (0.40% and 0.35%, respectively) with similar experiment design but small SOC variance (0.6-2.72%) for the field interrogated in the Bricklemyer and Brown (2010) study resulted in unsatisfactory

performances for the Veris sensor. Both Bricklemeyer and Brown (2010) and Schirrmann et al. (2013) suggested that inconsistent contact between sensor and soil might be partially responsible for degradation in VisNIR model performance.

Veris Technologies Inc. designed an on-the-go multi sensor platform for simultaneously collecting spectra and auxiliary measurements (i.e. electrical conductivity [EC] and temperature). Knadel et al. (2011) employed the Veris multi sensor platform to estimate SOC on a 12-ha young moraine field characterized by a mixture of peat and small inclusions of mineral soils (SOC range of 1.44-42.9%). The authors reported a decrease of RMSEP from 7.15% to 5.98% by using EC and spectra together compared to estimation with spectra only, indicating that the issues related to spectral acquisition on-the-go can be partly overcome with other sensory data (i.e. EC).

Foreoptics have also been developed for *in situ* VisNIR profile interrogation. As with on-the-go spectroscopy, *in situ* profile interrogation faces specific challenges: the foreoptic must be robust to soil insertion, the probe must have a relatively small cross-sectional diameter (no larger than the 36-mm diameter or 10 cm² cross-sectional area standard cone penetrometer) to minimize penetration resistance, and the side-viewing optics must be compressed into this small diameter tube. Hummels *et al.* (2004) fitted a bifurcated fiber optic inside a modified soil ASAE standard large cone penetrometer (20.3 mm base diameter tip and 15.88 mm outside diameter - 9.53 inside diameter hollow steel shaft). The bifurcated fiber optic simultaneously delivered the source light and collected the reflection from the soil. The authors tested the sensor in the laboratory on 28-cm diameter by 37-cm deep plastic containers, filled by 3 different soils to estimate soil moisture and cone index. Although they achieved satisfactory results (SEP = 2.38 % for soil moisture with $r^2 = 0.896$), the authors highlighted the need for field testing and suggested that limited signal amplitude due to attenuation in the bifurcated fiber optic was a source of error. Ben-Dor et al. (2008) designed a housing assembly hosting foreoptic sensor and a 45° mirror to direct

light on soil profile, running inside 10 cm diameter drilled hole. The authors tested the sensor on 4 different soil profiles across Israel: with 10 cm vertical step increment, the sensor was capable to distinguish the different soil horizons. Chang et al. (2011) developed micro optical probe suitable for a soil penetrometer housing (25.4 mm inner diameter). The authors achieved the side-viewing configuration by the use of microfabricated hybrid optical waveguides inside chip structures avoiding the use of mirrors: they used highly reflective metal layer on the sharp 90° bending with fiber optics at the terminals of the chip. The study compared spectra acquired with micro optical probe and ASD contact probe showing parallel absorption features pattern although spectra from micro optical probe had lower absolute reflectance and low signal to noise ratio at the edge of spectra (suitable range from 500-2200nm).

Veris Technologies Inc. has commercialized a penetrometer foreoptic for VisNIR profile interrogation *in situ* (Veris P4000™) that also collects data on EC and insertion force. However, this probe can only be deployed with a Veris spectrometer that lacks the spectral resolution and range of the commonly used ASD spectrometers. This 25.4-mm diameter probe contains a light source parallel to a fiber-optic cable that transmits reflected light to the spectrometer, with both facing a single angled concave mirror that directs source light on to the window and redirects reflected light onto the fiber optic cable. To provide room for these optics, the assembly has a square cross section that extends laterally beyond the cone tip diameter. Kweon et al. (2009) tested the Veris profile sensor at 5-8 locations on six fields in Kansas with predominant silty clay and silt loam soils. The authors reported accurate estimations of SOC (RMSE = 0.08-0.22%, varying according the field) and bulk density (RMSE = 0.07-0.13 g/cm³) using spectra and auxiliary data and leave-one-out cross validation, although the authors removed several outliers based on targeted soil properties.

One of the limitations of published research on *in situ* VisNIR soil spectroscopy is that it is difficult to separate and assess potential sources of error due to spectrometer differences, fore optic differences, and soil state (e.g. moisture, aggregation, heterogeneity, smearing). The first objective of this study was to design a penetrometer foreoptic with a performance comparable to commercial ASD foreoptics that can be used with any spectrometer having a fibre optic cable input (including commonly used ASD spectrometers). Specifically, the penetrometer foreoptic needed to simultaneously collect insertion force; be rigid enough for repeated insertion (without bending) into the dry soils; and have a narrow enough cross-section to insert using standard hydraulic soil coring equipment. Our second objective was to evaluate the optical performance of this penetrometer foreoptic relative to a commercial ASD foreoptic, using *exactly* the same spectrometer and soil materials under controlled conditions to quantify any potential performance degradation independent of spectrometer differences and *in situ* soil interrogation issues. Such an instrument would facilitate the use of existing libraries (with corrections where necessary [Ge et al., 2014]) in developing *in situ* calibrations.

2. Materials and Methods

2.1. VisNIR penetrometer foreoptic design

Our penetrometer design consists of three major components: main probe housing, optic assembly and cone tip assembly (Fig. 1). The main probe housing (OD = 31.75 mm) is constructed of stainless steel and it protects the optic assembly during the insertion. The cross section is approximately 11.4 mm larger in diameter than the ASAE standard large cone (Hummels et al., 2004) but has a smaller cross-section (7.9 cm^2) than the ASTM standard of 10 cm^2 (USDA, National Resources Conservation Services. 2012) and a more tapered tip for easier insertion. The probe housing connects to a shaft with an OD of 25.4 mm, with the probe and shaft having 9-mm and 6.35-mm thick walls, respectively, to endure high compression stress without bending while pushing into even dry, high-clay

soils. A 10-mm diameter quartz-sapphire window allows communication between optic assembly and soil profile. The quartz-sapphire window is sealed and centred on a flat area of the main probe housing, machined from housing wall and is 160 mm from the tip of the penetrometer. The window is set back 1.8 mm from the soil profile during insertion.

Figure 1: Penetrometer with three main components: (a) main probe housing and cone tip assembly, (b) current optic assembly and (c) first ASD optic assembly.

The optic assembly, designed to fit inside the main probe housing (ID = 13.45 mm), directs source light (halogen lamp) towards the foreoptic window and conveys diffuse reflected light towards a fiber optic cable (Fig. 1b). Our first optic assembly, purchased from ASD and shown in the Figure 1c, used a bifurcated fiber optic cable to both transmit source light (from an external halogen lamp) down the length of the probe and return reflected light up the probe to the spectrometer. The bifurcated fiber optic cable was presented to a 45° degree flat mirror, diverting source light to the window and reflected light back to the spectrometer fiber optic.

After tests of this optic assembly proved unsatisfactory due to source light attenuation and backscatter (discussed in results), we designed the current optic assembly shown in Figure 1b. The optic apparatus consists of stainless steel cylinder (OD = 11 mm and ID =

6.4 mm) into which the spectrometer fiber optic and halogen lamp are inserted from opposite sides, facing two angled flat mirrors angled as shown in Fig 1b. The mirror facing the fiber optic is angled at 56° relative to the main axis of the probe and optic assembly whereas the mirror facing the lamp is angled at 38° relative to the main axis to reduce the amount of specular reflection (due to window) incident on the fiber optic (Fig. 1b). The configuration of the two mirrors directs incident light through the quartz-sapphire window towards soil profile and diffuse reflected light towards the fiber optic cable. We applied flat black (Krylon®) on the sides of the optic assembly and reflector mirrors to decrease backscattering. A notch inside the main probe housing and two screws secure the optic assembly.

The cone tip assembly is inserted into probe main housing and held in place through a screw. The cone tip assembly is divided into two components: the cone tip and the load cell housing. The cone tip angle is 30° (Fig. 1a). The cone tip and the load cell housing are separated by a rubber O-ring to avoid soil entering inside the load cell housing. Load cell housing (ID = 27.89 mm) hosts Futek mini load cell with a range of 0-3000 lbf for *in situ* insertion force measurements. A precision cable extension potentiometer for linear position measurement (UniMeasure® - HX-PA series) connected to the penetrometer records the depth during insertion.

2.2. Soil samples and laboratory analyses

We evaluated the optical performance of the VisNIR penetrometer under controlled laboratory conditions at Washington State University. The 386 soil samples used in this study were obtained from the United States Geological Service archives in Denver, CO, USA. Samples were archived A- and C-horizon soils originally collected in 2004 along two continental-scale transects as a pilot study for the North American Soil Geochemical

Landscapes Project (Smith et al., 2009, Fig. 2). Samples were air dried, disaggregated, passed through 2-mm sieve, crushed to <150 μm in a ceramic mill, and homogenized by mixing prior to chemical analysis. Total C (TC), inorganic carbon (i.e. carbonate-C; IC), were determined for all samples by a commercial geochemical laboratory (Smith et al., 2009). TC was determined by automated dry combustion analysis (Brown and Curry, 2002) with IC measured by coulometric titration (Brown et al., 2002). Organic C (SOC) was calculated as the difference between TC and IC. Readers are directed to Smith et al. (2009) for project details regarding site selection, field sampling and analysis protocols, and quality assurance/quality control. Approximately 6 g of crushed soil material was placed in a 31 x 8 mm aluminum sample cup (Spex Certiprep, New Jersey, USA) and pressed into a puck under 344.7 MPa for five minutes (Fig. 3a). Pressing soil into pucks allowed us to consistently present soil material flush with different foreoptic windows.

Figure 2: Location of soil sampling sites along the two transect across Canada and USA (Smith et al., 2009)

2.3 Spectra collection and modelling

Each sample puck was scanned in the laboratory using two different foreoptics: (a) the VisNIR penetrometer described previously (Fig. 3c); and (b) an Analytical Spectral Devices (ASD Inc. Boulder, CO, USA) high intensity contact probe (Fig. 3b) designed for contact measurement of solid materials. The ASD contact probe source halogen light-mirror assembly and fiber optic detector are set parallel at a 70° angle to the window, with the fiberoptic recessed slightly behind the light source mirror. The spot size is 10 mm whereas the window has 20 mm diameter. For more details, readers can visit Analytical Spectral Devices website (<http://www.asdi.com/>). Both foreoptics were employed with the same ASD “Agrispec” VisNIR spectrometer having a spectral range of 350-2500 nm and a spectral resolution of 3 nm between 350-1025 nm for VisNIR silicon photo-diode array detector and 10 nm between 1025-2500 nm for SWIR1 (1001-1800 nm) and SWIR2 (1801-2500 nm) InGaS detectors. Each recorded spectrum was the average of 25 scans, acquired internally. For each soil sample, two replicate scans were acquired rotating the puck 90°. After every 50 soil samples, a white Spectralon® and pure kaolinite reference puck were scanned with both foreoptics, to recalibrate the ASD spectrometer and assess the consistency of spectra acquisition, respectively. For the VisNIR penetrometer foreoptic only, a black reference (3 cm x 3 cm x 10 cm cardboard box internally painted with Krylon® flat black) was also scanned every 50 soil samples to control for backscattering.

Figure 3: In laboratory spectra collection with two foreoptic sensors: a) example of pucks, b) ASD contact probe and c) VisNIR penetrometer.

Sample spectra from both foreoptics were corrected for possible offset at splice points between different regions of spectrometer detectors (VisNIR – SWIR1 – SWIR2) using ViewSpec Pro 6.0.15 (ASD Inc. Boulder, CO, USA). Afterwards, reflectance and 1st derivative data from both foreoptics were extracted every 10 nm along the entire spectral range (350-2500 nm) from a weighted cubic smoothing spline fitted to the corrected spectra as developed by (Brown et al., 2006). This preprocessing method removes random noise while retaining absorbance features by smoothing the spectra with a wavelength-dependent signal-to-noise ratio function. Modified from original routine described in Brown et al. (2006), we fitted parameters to increase the smoothing spline tension (“smoothing.spline” function in R Statistical Software – R Core Team, 2013), raising knot gap from 10 to 10.75 nm and smoothing parameter from 0.05 to 0.1. By spacing smoothing splines knots at 10.75 nm—asynchronous with the 10 nm intervals used for reflectance and derivative extraction, the splines better smoothed out random noise. In order to remove as much noise as possible without losing information (Kramer et al., 2004), the spline smoothing parameter

was increased gradually until the kaolinite doublet at 1390 and 1420 (Clark et al., 1990; Crowley and Vergo, 1988) started to merge into a single feature.

We qualitatively compared reflectance and 1st derivative data from the two foreoptic sensors through direct inspection of paired spectra. To quantitatively evaluate the similarity between the two foreoptics, we calculated wavelength-specific Pearson correlation coefficients along the entire range (350-2500 nm) at a 10 nm resolution for both reflectance and the 1st derivative.

We constructed prediction models for SOC with the Partial Least Squares (PLS) regression technique. (Total C and Inorganic C models provided no additional insight and are therefore not reported in this paper.) The PLSR regression routine was implemented using the R statistical software package “pls”, version 2.4-3 (Bjørn-Helge et al., 2013). Whole-location out (surface and subsoil samples) cross-validation was used to assess predictive model accuracy. The performance of PLSR models was evaluated using statistics obtained by comparing PLSR predicted vs. wet-lab measured SOC values as outlined by Gauch et al. (2003) and (Brown et al., 2005):

$$(1) \text{ MSD} = \sum_N \frac{(Y_{pred} - Y_{lab})^2}{N}$$

$$(2) \text{ Bias} = \sum_N \frac{Y_{pred} - Y_{lab}}{N}$$

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

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

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

$$(6) \text{ SEP} = \left[\frac{\sum (Y - \hat{Y} - \text{bias})^2}{N-1} \right]^{0.5}$$

$$(7) \text{ RPD} = \frac{SD_Y}{SEP}$$

where b and r^2 represents the slope and square correlation, respectively, from the least squares regression of Y_{pred} on Y_{lab} . Mean squared deviation (MSD) was used to distinguish the discrepancies between SOC prediction and observed data for both foreoptic sensors. Following (Gauch et al., 2003), MSD was partitioned in bias (SB), non-unity slope of regression line (NU) and lack of correlation (LC). SEP is the validation Standard Error of Prediction, RPD is the residual product differential. Likewise to quantitatively analyze spectral similarity, we calculated wavelength-specific Pearson correlation coefficients for the PLS regression coefficients of the calibration models derived from the two foreoptics.

3. Results

3.1 Design configuration

The narrow space inside the probe housing presented the most difficult challenge in the design of our VisNIR penetrometer foreoptic. While both the first ASD design (Fig. 4a) and the initial alignment of the current (Fig. 4b) optic assembly designs overcame this challenge, backscatter was reduced with our current design (average backscatter = 42%) relative to the ASD design (average backscatter = 51%). The current design with opposing light source and fiber optic preserved the small diameter of the optic assembly needed to fit inside the main housing and still provided a “closely-packed” configuration capable of delivering and collecting a higher intensity signal (Chang et al., 2011). With further experimentation, we found that by moving the center of the two mirrors 1.5 mm up the shaft from the center of the window, backscatter was reduced to an average of 12% (Fig. 4c). Nonetheless, the use of mirrors to redirect light increased backscattering relative to direct light path towards soil and back as in the ASD contact probe (average backscatter < 1%, Fig. 4d).

Figure 4: Backscattering comparison between different designs: (a) 1st design (ASD optic assembly), (b) current design (centered mirrors), (c) current design (up-shifted mirrors) and (d) ASD contact probe.

We found it difficult to consistently present soil material to a foreoptic, particularly with the goal of simulating *in situ* soil contact. Pushing a foreoptic into soil material changes that material for subsequent interrogations, and soil heterogeneity leads to different material presentation with small shifts in probe position or changes in window size. Pressed soil pucks provided sample homogeneity, presentation consistency, and a flat surface for optimum direct soil-window contact, so the foreoptics could be directly compared. This approach could be used to evaluate a variety of foreoptics for soil analysis.

3.2 Direct comparison of spectra

Visual observation of reflectance and first derivative data showed comparable patterns between VisNIR penetrometer and contact probe foreoptics, along visible and near-infrared range (350-2500 nm) (Fig. 5). Our VisNIR penetrometer foreoptic captured important absorption features, like hydroxyl and adsorbed water at 1400 nm and 1900 nm, respectively and structural hydroxyl bond at 2200 nm in clay mineral (Ben-Dor, 2002) with the same relative intensity showed by contact probe and slightly variations in absolute reflectance values as illustrated with two examples in Figure 5. The kaolinite puck scans collected as part of this experiment also show the spectra from the two different foreoptics to be highly similar (Fig. 6).

Figure 5: Comparison of spectra reflectance and first derivative between VisNIR penetrometer and ASD contact probe for two selected soil samples.

The median of VisNIR penetrometer spectra of kaolinite sample had similar reflectance values to the median of contact probe spectra along the entire range 350-2500 nm. The VisNIR penetrometer spectra had a wider interquartile range compared to contact probe with a standard deviation of wavelength-specific reflectance values between 0.4% and 2.9% for VisNIR reflectance and between 0.2% and 1.3% for the contact probe reflectance. First derivative signatures for VisNIR penetrometer and contact probe are highly comparable, slightly diverging at the edges of the spectra (350-470 nm and 2470-2500 nm).

Figure 6: Median and interquartile range of 10 kaolinite scans collected respectively with VisNIR penetrometer and ASD contact probe (a) and zoom of characteristic doublet of kaolinite (b).

In a direct comparison of the two foreoptics, Pearson coefficients for each wave band varied between 0.85 and 0.98 ($n = 386$), with the lowest correlations found in the UV end of the visible range (350-500 nm) (Fig. 7). The Pearson coefficient for first derivative comparison yielded a less consistent correlation, with low r values observed at the edges of the spectra (350 and 2500 nm) and a negative coefficient (-0.22) at 1810 nm, related to the spectral discontinuity at the internal detector splice in the ASD spectrometer (Ge et al., 2011).

Figure 7: Pearson coefficient calculated between VisNIR penetrometer and ASD contact probe for reflectance and first derivative data (total soil samples = 386).

3.3 SOC data and model performance

Summary statistics for SOC measurements (Table 1) show a wide range of values (0.4 g kg⁻¹ to 109.3 g kg⁻¹), a mean of 15.1 g kg⁻¹ and standard deviation of 18 g kg⁻¹. The log-transformed data showed values of mean and median, 0.98 and 0.96, respectively, with an approximately normal distribution with a standard deviation of 0.45 [log(g kg⁻¹)] and coefficient of variation equal to 46%. Since the purpose of this analysis was to compare the performances of our VisNIR penetrometer with ASD contact probe, we did not back transform SOC results prior to computing performance statistics.

Table 1: Summary statistics for standard laboratory measured soil organic carbon (SOC) and logarithmic transformation of soil organic carbon.

	N	Mean	σ	Median	Skewness	Min	Max	CV
<i>SOC</i>	386	15.12	17.98	8.15	2.32	0.40	109.30	118.9
<i>Log (1+SOC)</i>	386	0.98	0.45	0.96	0.17	0.15	2.04	45.7

σ = standard deviation, CV = coefficient of variation [%]

Chemometric model performance statistics for log-SOC were comparable for VisNIR penetrometer and contact probe foreoptics (Table 2). We obtained satisfactory validation accuracy (Table 2 and Fig. 8) with reflectance data from both foreoptics (RPD = 2.34 and = 2.45 for VisNIR penetrometer and ASD contact probe, respectively), as well as with 1st derivative data from both foreoptics (RPD = 2.31 and RPD = 2.42 for VisNIR penetrometer and ASD contact probe, respectively). There was little difference in the standard error of prediction (SEP) regardless of foreoptic used or spectral transformation (reflectance or 1st derivative). A direct comparison of foreoptic cross-validation SOC predictions obtained from reflectance data showed very similar predictions for the contact probe and penetrometer on a sample-by-sample basis ($r^2 = 0.97$, Fig. 9). The VisNIR penetrometer and ASD foreoptics yielded comparable partitionings of mean squared deviation (MSD) (Table 2); with most of the error due to lack of correlation (LC) and minimal bias in predicted SOC with either reflectance data (LC = 90.7% and 89.7% for penetrometer and contact probe, respectively) or first derivative data (LC = 91.2% and 90.3% for penetrometer and contact probe, respectively).

Regression coefficients for the two different PLSR models for reflectance data showed a high degree of correspondence along the entire spectral range (Fig 10). Coefficients deviated more so in the visible range of the spectrum, however wavelength importance was consistent across calibrations with only slight differences in the model coefficient magnitude. Wavelengths at 1400 and 1750 to 2500 nm showed the highest regression coefficient values. These significant wavelengths are associated with specific absorption features typical of organic compounds: absorption at 1400 nm by C-H, N-H, OH⁻ and water molecules. Furthermore, free water and C-O, O-H, C=O, C-H and N-H present strong absorptions at 1760 nm, and between 1900 and 2500 nm (Ben-Dor et al., 1997; Hunt, 1977).

Table 2: Partial least squares statistics predicting soil organic carbon (SOC) using reflectance and 1st derivative data for VisNIR penetrometer and ASD contact probe foreoptics. (Validation SEP for both the contact probe and penetrometer models were only 0.02 units— $\log(\text{g kg}^{-1})$ —greater than calibration SEP values.)

	Penetrometer		Contact Probe	
	R	1 st Der	R	1 st Der
<i>N</i>	386	386	386	386
<i>Latent Variables</i>	15	12	12	11
<i>r</i> ²	0.82	0.82	0.84	0.83
<i>SEP log(g kg⁻¹)</i>	0.19	0.19	0.18	0.19
<i>RPD</i>	2.34	2.31	2.46	2.42
<i>SB (%)</i>	0.1	0.1	0.0	0.0
<i>NU (%)</i>	9.2	8.7	10.2	9.7
<i>LC (%)</i>	90.7	91.2	89.7	90.3

R = Reflectance, 1st Der = First derivative, SEP = standard error of prediction, RPD =

residual product differential, SB = squared bias, NU = non-unity of regression line slope,

LC = lack of correlation. SB, NU and LC are calculated as percent of mean squared

deviation.

Figure 8: Partial least squares log(SOC) model validation.

Figure 9: A comparison of VisNIR-predicted SOC for penetrometer vs. contact probe foreoptics.

Figure 10: PLS calibration model regression coefficients based on reflectance data obtained from the penetrometer and contact probe foreoptics.

4. Discussion

The VisNIR penetrometer and ASD contact probe yielded surprisingly similar spectra—both quantitatively and qualitatively. Though the interquartile ranges of reflectance values in VisNIR penetrometer spectra were wider relative to the contact probe, a low wavelength-specific standard deviation of spectra suggests stability and spectral reproducibility (Ben-Dor et al., 2008). Divergences between the two foreoptics were visible in areas of spectra with a low signal-to-noise ratio such as the edges of spectral range and the transition between the spectrometer detectors. Small fluctuations in backscattering due to source light oscillation as reported by Munday et al. (2004), may have caused the wider interquartile ranges for our VisNIR penetrometer foreoptic spectra, as this foreoptic lacks the electronic stability circuitry found in the ASD contact probe. Collection of external

black reference measurements is recommended to ameliorate potential backscattering effects, even though these effects are limited by the penetrometer design.

The VisNIR penetrometer SOC predictions (SEP, RPD, and R^2) were comparable with the contact probe. Though a rigorous evaluation of model performance is beyond the aim of current paper, the results observed from both foreoptic sensors are consistent with previous studies (Dunn et al., 2002; Islam et al., 2003; Morgan et al., 2009; Sankey et al., 2008; Summers et al., 2011; Vasques et al., 2010, 2008). For the penetrometer calibration, higher regression coefficients were observed around 1400 nm and 1800 to 2500 nm, consistent with previous studies (Sarkhot et al., 2011; Stevens et al., 2013; Vasques et al., 2010). Both the penetrometer and contact probe calibrations tended to underestimate SOC at high values, potentially due to uneven distribution of SOC data towards low values of SOC and inadequate representation of high SOC values. Similarities in validation statistics and regression coefficients suggest that a potentially lower signal-to-noise ratio for the VisNIR penetrometer does not considerably degrade chemometric model performance.

5. Conclusions

We designed a penetrometer foreoptic which incorporates VisNIR technology for *in situ* soil characterization and can be used with any field spectrometer that acquires spectra via a fiber optic cable. Spectra derived from the VisNIR penetrometer were highly correlated with contact probe spectra ($r = 0.86$ to 0.98 , $n = 386$ spectra), and the band-wise first derivatives were well correlated for all wavelengths except at the edges of the ASD Field Spec internal detectors. A comparison of chemometric models for SOC prediction derived from these two foreoptics yielded very similar validation performance statistics (SEP = 0.18 and $0.19 \log[\text{g kg}^{-1}]$) for the contact probe and penetrometer, respectively. In a laboratory setting, spectra derived from the VisNIR penetrometer were comparable in most respects to spectra obtained using a commercial ASD contact probe. We recommend that laboratory evaluation of foreoptic performance be a standard step in the development of *in*

situ VisNIR probes so that various sources of error associated with *in situ* spectroscopy can be segregated and quantified. In future work, we will evaluate the performance of the penetrometer foreoptic in the field for *in situ* soil characterization.

Acknowledgments

Project was funded by USDA-NIFA grant No. 2011-67003-30341 and U.S. DOE-NETL grant No. DE-FC26-05NT42587. Dan Rooney contributed to the initial penetrometer design. Cristine Morgan and Yufeng Ge collaborated on VisNIR penetrometer design revisions. Dave Smith at the US Geological Survey provided soil samples and associated laboratory data.

References

- Adamchuk, V.I., Hummel, J.W., Morgan, M.T., Upadhyaya, S.K., 2004. On-the-go soil sensors for precision agriculture. *Comput. Electron. Agric.* 44, 71–91.
doi:10.1016/j.compag.2004.03.002
- Ben-Dor, E., 2002. Quantitative remote sensing of soil properties, in: *Advances in Agronomy*. Academic Press, pp. 173–243.
- Ben-Dor, E., Banin, A., 1995. Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties. *Soil Sci. Soc. Am. J.* 59, 364.
doi:10.2136/sssaj1995.03615995005900020014x
- Ben-Dor, E., Heller, D., Chudnovsky, A., 2008. A Novel Method of Classifying Soil Profiles in the Field using Optical Means. *Soil Sci. Soc. Am. J.* 72, 1113.
doi:10.2136/sssaj2006.0059
- Ben-Dor, E., Inbar, Y., Chen, Y., 1997. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process. *Remote Sens. Environ.* 61, 1–15. doi:10.1016/S0034-4257(96)00120-4

- Bjørn-Helge M., Wehrens R. and Hovde Liland K. (2013). pls: Partial Least Squares and Principal Component regression. R package version 2.4-3. <http://CRAN.R-project.org/package=pls>
- Bricklemeyer, R.S., Brown, D.J., 2010. On-the-go VisNIR: Potential and limitations for mapping soil clay and organic carbon. *Comput. Electron. Agric.* 70, 209–216. doi:10.1016/j.compag.2009.10.006
- Brown, D.J., Bricklemeyer, R.S., Miller, P.R., 2005. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. *Geoderma* 129, 251–267. doi:10.1016/j.geoderma.2005.01.001
- Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M., Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma* 132, 273–290. doi:10.1016/j.geoderma.2005.04.025
- Brown, Z.A., Curry, K.J., 2002. Total carbon by combustion, in: Taggart, J.E. (Ed.), *Analytical Methods for Chemical Analysis of Geologic and Other Materials*, US Geol. Surv. Open-File Rep. 02-223. US Geological Survey, pp. 1–4 (Chapter R).
- Brown, Z.A., Papp, C., Brandt, Aruscavage, P., 2002. Carbonate carbon by coulometric titration, in: Taggart, J.E. (Ed.), *Nalytical Methods for Chemical Analysis of Geologic and Other Materials*, US Geol. Surv. Open-File Rep. 02-223. US Geological Survey, pp. 1–6 (Chapter S).
- Chang, C.-C., Garcia-Urbe, A., Zou, J., Morgan, C.L.S., 2011. Micro Side-Viewing Optical Probe for VNIR-DRS Soil Measurement. *IEEE Sens. J.* 11, 2527–2532. doi:10.1109/JSEN.2011.2128867
- Christy, C.D., 2008. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. *Comput. Electron. Agric.* 61, 10–19. doi:10.1016/j.compag.2007.02.010

- Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G.A., Vergo, N., 1990. High spectral resolution reflectance spectroscopy of minerals. *J. Geophys. Res. Solid Earth* 95, 12653–12680. doi:10.1029/JB095iB08p12653
- Crowley, J.K., Vergo, N., 1988. NEAR-INFRA-RED REFLECTANCE SPECTRA OF MIXTURES OF KAOLIN-GROUP MINERALS: USE IN CLAY MINERAL STUDIES. *CLAYS CLAY Min. Clays Clay Min.* 36, 310.
- Demattê, J.A.M., Campos, R.C., Alves, M.C., Fiorio, P.R., Nanni, M.R., 2004. Visible–NIR reflectance: a new approach on soil evaluation. *Geoderma* 121, 95–112. doi:10.1016/j.geoderma.2003.09.012
- Dunn, B.W., Batten, G.D., Beecher, H.G., Ciavarella, S., 2002. The potential of near-infrared reflectance spectroscopy for soil analysis — a case study from the Riverine Plain of south-eastern Australia. *Aust. J. Exp. Agric.* 42, 607–614.
- Gauch, H.G., Hwang, J.T.G., Fick, G.W., 2003. Model Evaluation by Comparison of Model-Based Predictions and Measured Values. *Agron. J.* 95, 1442. doi:10.2134/agronj2003.1442
- Ge, Y., Morgan, C.L.S., Ackerson, J.P., 2014. VisNIR spectra of dried ground soils predict properties of soils scanned moist and intact. *Geoderma* 221–222, 61–69. doi:10.1016/j.geoderma.2014.01.011
- Ge, Y., Morgan, C.L.S., Grunwald, S., Brown, D.J., Sarkhot, D.V., 2011. Comparison of soil reflectance spectra and calibration models obtained using multiple spectrometers. *Geoderma* 161, 202–211. doi:10.1016/j.geoderma.2010.12.020
- Gholizadeh, A., Borvka, L., Saberioon, M., Vaát, R., 2013. Visible, Near-Infrared, and Mid-Infrared Spectroscopy Applications for Soil Assessment with Emphasis on Soil Organic Matter Content and Quality: State-of-the-Art and Key Issues. *Appl. Spectrosc.* 67, 1349–1362.
- Heinze, S., Vohland, M., Joergensen, R.G., Ludwig, B., 2013. Usefulness of near-infrared spectroscopy for the prediction of chemical and biological soil properties in different

- long-term experiments. *J. Plant Nutr. Soil Sci.* 176, 520–528.
doi:10.1002/jpln.201200483
- Huang, X., Senthilkumar, S., Kravchenko, A., Thelen, K., Qi, J., 2007. Total carbon mapping in glacial till soils using near-infrared spectroscopy, Landsat imagery and topographical information. *Geoderma* 141, 34–42. doi:10.1016/j.geoderma.2007.04.023
- Hunt, G., 1977. SPECTRAL SIGNATURES OF PARTICULATE MINERALS IN THE VISIBLE AND NEAR INFRARED. *GEOPHYSICS* 42, 501–513.
doi:10.1190/1.1440721
- Islam, K., Singh, B., McBratney, A., 2003. Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. *Soil Res.* 41, 1101–1114.
- Knadel, M., Thomsen, A., Greve, M.H., 2011. Multisensor On-The-Go Mapping of Soil Organic Carbon Content. *Soil Sci. Soc. Am. J.* 75, 1799. doi:10.2136/sssaj2010.0452
- Kramer, R., Workman, J., Reeves, J.B., 2004. Qualitative Analysis. -Infrared Spectrosc. *Agric. agronomymonogra*, 175–206.
- Kuang, B., Mahmood, H.S., Quraishi, M.Z., Hoogmoed, W.B., Mouazen, A.M., van Henten, E.J., 2012. Chapter four - Sensing Soil Properties in the Laboratory, In Situ, and On-Line: A Review, in: Donald L. Sparks (Ed.), *Advances in Agronomy, Advances in Agronomy*. Academic Press, pp. 155–223.
- Kusumo, B.H., Hedley, C.B., Hedley, M.J., Hueni, A., Tuohy, M.P., Arnold, G.C., 2008. The use of diffuse reflectance spectroscopy for in situ carbon and nitrogen analysis of pastoral soils. *Soil Res.* 46, 623–635.
- Kweon, G.Y., Lund, E., Maxton, C., Drummond, P., Jensen, K., 2009. Soil profile measurement of carbon contents using a probe-type VIS-NIR spectrophotometer. *J. Biosyst. Eng.* 34, 382–389.
- Morgan, C.L.S., Waiser, T.H., Brown, D.J., Hallmark, C.T., 2009. Simulated in situ characterization of soil organic and inorganic carbon with visible near-infrared diffuse

- reflectance spectroscopy. *Geoderma* 151, 249–256.
doi:10.1016/j.geoderma.2009.04.010
- Munday, W.L.C., Zapf, C.M., Reeves, J.B., 2004. Testing and Validating Instrument Performance. -*Infrared Spectrosc. Agric. agronomymonogra*, 33–48.
- O' Rourke, S.M., Holden, N.M., 2011. Optical sensing and chemometric analysis of soil organic carbon – a cost effective alternative to conventional laboratory methods? *Soil Use Manag.* 27, 143–155. doi:10.1111/j.1475-2743.2011.00337.x
- Sankey, J.B., Brown, D.J., Bernard, M.L., Lawrence, R.L., 2008. Comparing local vs. global visible and near-infrared (VisNIR) diffuse reflectance spectroscopy (DRS) calibrations for the prediction of soil clay, organic C and inorganic C. *Geoderma* 148, 149–158.
doi:10.1016/j.geoderma.2008.09.019
- Sarkhot, D.V., Grunwald, S., Ge, Y., Morgan, C.L.S., 2011. Comparison and detection of total and available soil carbon fractions using visible/near infrared diffuse reflectance spectroscopy. *Geoderma* 164, 22–32. doi:10.1016/j.geoderma.2011.05.006
- Schirrmann, M., Gebbers, R., Kramer, E., 2013. Performance of Automated Near-Infrared Reflectance Spectrometry for Continuous in Situ Mapping of Soil Fertility at Field Scale. *Vadose Zone J.* 12, 0. doi:10.2136/vzj2012.0199
- Shepherd, K.D., Walsh, M.G., 2002. Development of Reflectance Spectral Libraries for Characterization of Soil Properties. *Soil Sci Soc Am J* 66, 988–998.
- Shonk, J.L., Gaultney, L.D., Schulze, D.G., Van Scoyoc, G.E., 1991. Spectroscopic sensing of soil organic matter content. *Trans. ASAE USA*.
- Smith, D.B., Woodruff, L.G., O'Leary, R.M., Cannon, W.F., Garrett, R.G., Kilburn, J.E., Goldhaber, M.B., 2009. Pilot studies for the North American Soil Geochemical Landscapes Project – Site selection, sampling protocols, analytical methods, and quality control protocols. *Appl. Geochem.* 24, 1357–1368.
doi:10.1016/j.apgeochem.2009.04.008

- Stevens, A., Nocita, M., Tóth, G., Montanarella, L., van Wesemael, B., 2013. Prediction of Soil Organic Carbon at the European Scale by Visible and Near InfraRed Reflectance Spectroscopy. *PLoS ONE* 8, e66409. doi:10.1371/journal.pone.0066409
- Sudduth, K.A., Hummel, J.W., 1993a. Soil organic matter, CEC, and moisture sensing with a portable NIR spectrophotometer. *Trans. ASAE* v. 36(6) p. 1571-1582.
- Sudduth, K.A., Hummel, J.W., 1993b. Portable, near-infrared spectrophotometer for rapid soil analysis. *Trans. ASAE* v. 36(1) p. 185-193.
- Summers, D., Lewis, M., Ostendorf, B., Chittleborough, D., 2011. Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties. *Ecol. Indic.* 11, 123–131. doi:10.1016/j.ecolind.2009.05.001
- Vasques, G.M., Grunwald, S., Harris, W.G., 2010. Spectroscopic models of soil organic carbon in Florida, USA. *J. Environ. Qual.* 39, 923–934. doi:10.2134/jeq2009.0314
- Vasques, G.M., Grunwald, S., Sickman, J.O., 2008. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. *Geoderma* 146, 14–25. doi:10.1016/j.geoderma.2008.04.007
- Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* 131, 59–75. doi:10.1016/j.geoderma.2005.03.007
- Viscarra Rossel, R.A., Cattle, S.R., Ortega, A., Fouad, Y., 2009. In situ measurements of soil colour, mineral composition and clay content by vis–NIR spectroscopy. *Geoderma* 150, 253–266. doi:10.1016/j.geoderma.2009.01.025
- Waiser, T.H., Morgan, C.L.S., Brown, D.J., Hallmark, C.T., 2007. In Situ Characterization of Soil Clay Content with Visible Near-Infrared Diffuse Reflectance Spectroscopy. *Soil Sci. Soc. Am. J.* 71, 389. doi:10.2136/sssaj2006.0211