

Application of Mid-Infrared Spectroscopy in Soil Survey

Cathy A. Seybold*

Rich Ferguson

Doug Wysocki

Scarlett Bailey

USDA–NRCS

National Soil Survey Center

100 Centennial Mall N.

Rm. 152

Lincoln, NE 68508

Joe Anderson

Brian Nester

USDA–NRCS

760 S. Broadway

Salina, KS 67401

Phil Schoeneberger

Skye Wills

Zamir Libohova

Dave Hoover

USDA–NRCS

National Soil Survey Center

100 Centennial Mall N.

Rm. 152

Lincoln, NE 68508

Pam Thomas

USDA–NRCS

Room 4836 South Building

14th & Independence Ave, SW

Washington, DC 20250

Development of more efficient (rapid) and cost-effective methodologies are needed in soil survey to meet the demand for quantitative data in digital soil mapping and updates. The objective of this study was to pilot the application of mid-infrared (MIR)-diffuse reflectance spectroscopy (DRS) coupled with partial least squares regression (PLSR) in a soil survey field office where soil samples are processed, MIR spectra are acquired, and predictions are obtained using calibration models developed and validated from the Kellogg Soil Survey Laboratory spectral library. Mid-infrared models were built for total C, organic C, CaCO_3 equivalent, total clay, cation exchange capacity, 1500 kPa water, and pH in water and CaCl_2 for Mollisols of the central United States. Validation results (from the MIR library) using Lin's concordance correlation (r_c) of measured versus predicted values showed that most properties predicted very well ($r_c = 0.967\text{--}0.996$), whereas models for total clay in B horizons and 1500 kPa water in B horizons predicted fairly well ($r_c = 0.844\text{--}0.955$). Models for pH predicted the least well ($r_c = 0.750\text{--}0.921$). The MIR-DRS coupled with PLSR was successful in predicting soil properties for completely independent samples that were collected, processed, and MIR scanned in a soil survey field office. Predicted results using r_c ranged from 0.697 to 0.992, with pH in water having the lowest r_c and CaCO_3 having the highest r_c . All properties except pH had an acceptable level of accuracy for use in soil survey and a marginal level of acceptable accuracy for total clay. Direct calibration transfer was feasible.

Abbreviations: CEC, cation exchange capacity; DRS, diffuse reflectance spectroscopy; KSSL, Kellogg Soil Survey Laboratory; ME, mean error; MIR, mid-infrared; NIR, near-infrared; PLSR, partial least squares regression; RH, relative humidity; RMSE_e , root mean square error of estimation; RMSE_p , root mean square error of prediction; RPD, residual prediction deviation.

Soil is one of the basic natural resources, and its inventory and assessment are critical (Soil Survey Staff, 2011). The ability to inventory soil properties in soil survey has been hampered by the inability to rapidly measure soil properties. Soil analysis using traditional laboratory technology (e.g., wet chemistry) is expensive and time consuming. It can take months to get characterization samples through the laboratory. Also, it is impractical to measure soil properties everywhere in a soil survey area. Therefore, soil properties are often estimated, or pedotransfer functions are used. In addition, soil survey in the United States has made advances in its use of digital soil mapping, which exploits the quantitative relationships between soil observations and digital environmental data and products. Digital soil mapping will be used to some degree to complete most of the unmapped areas of the United States and for updating existing soil surveys. The goal is to have full inventory coverage by 2026. As soil survey moves toward greater use of digital soil mapping, the demand for measured soil property data will increase tremendously. More efficient (rapid) and cost-effective methodologies are required to meet this demand.

Infrared spectroscopy, both near-infrared (NIR) and mid-infrared (MIR), allows for rapid acquisition of soil information and prediction of various soil proper-

Core Ideas

- Mid-infrared diffuse reflectance spectroscopy was piloted for use in soil survey.
- Calibrations were built for Mollisols of the central United States for eight soil properties.
- A method was developed for getting MIR predictions in soil survey field offices.
- Successful soil property predictions were obtained using independent samples.
- Direct calibration transfer was viable.

Soil Sci. Soc. Am. J. 83:1746–1759

doi:10.2136/sssaj2019.06.0205

Received 28 June 2019.

Accepted 3 Sept. 2019.

*Corresponding author (cathy.seybold@usda.gov).

© Soil Science Society of America. This is an open access article distributed under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

ties (Brown et al., 2006; Janik et al., 1998; Minasny et al., 2009; Shepherd and Walsh, 2002; Tinti et al., 2015; Viscarra-Rossel et al., 2006; Zovko et al., 2018). Infrared spectroscopy offers high analysis throughput and is a relatively inexpensive method for estimating soil properties. It has the advantage that various physical, chemical, and biological soil properties can be simultaneously assessed from a single spectrum. Where more spatially dense analyses are required, such as in digital soil mapping, infrared spectroscopy can meet the demand for inputs (Madejova and Komadel, 2001). Mid-infrared spectroscopy has been shown to give better predictions of soil properties compared with NIR spectroscopy (McCarty et al., 2002; Minasny and McBratney, 2008; Pirie et al., 2005; Reeves et al., 2001). Mid-infrared spectroscopy detects the fundamental vibrations of minerals and organic matter, which have strong absorptions, whereas NIR spectroscopy detects their overtones and combinations of overtones that are much weaker and greatly overlap (Shepherd and Walsh, 2007).

Mid-infrared (MIR)-diffuse reflectance spectroscopy (DRS) has been shown to be an ideal technique for quantitative analysis in soils (Jia et al., 2017). The technique reduces sample preparation to a minimum and has undergone substantial development over the last couple of decades such that reliable calibrations for diverse soils can be achieved with no sample dilution (Janik et al., 1998; Siebielec et al., 2004). Partial least squares regression (PLSR) is the most common chemometric method used in the prediction of soil properties from diffuse infrared reflectance spectra (Minasny and McBratney, 2008) and can quantify various soil properties with a high degree of accuracy and precision (Soriano-Disla et al., 2014). Mid-infrared DRS coupled with PLSR can successfully predict several soil properties, such as organic C, cation exchange capacity (CEC), total clay, calcium carbonate, soil pH (D'Acqui et al., 2010; Reeves and Smith, 2009; Viscarra Rossel et al., 2006), and matric suction water contents (Janik et al., 2007a). Other properties successfully predicted include soil organic C fractions (Knox et al., 2015), plasticity characteristics of clay soils (Kasprzhitskii et al., 2018), Atterberg limits (Waruru et al., 2015), soil metal contents (Siebielec et al., 2004), and the characteristics of soil organic matter (Cox et al., 2000; Madari et al., 2006), peat (Holmgren and Norden, 1988), and particulate organic matter (Bornemann et al., 2010). Because of these successful applications, the use of MIR-DRS and PLSR for predicting soil properties in soil survey has great potential.

A key limitation in deploying MIR spectroscopy in production has been the transfer of models between geographic areas, libraries, and instruments and the inconsistency of reference data used in developing calibrations. The Kellogg Soil Survey Laboratory (KSSL) in Lincoln, NE, has a growing national MIR spectral library that represents soil samples from over 17,000 pedons from across the United States. The reference data associated with the spectra are consistent throughout the library, where methods have changed little over time. All spectra in the library were acquired on one instrument. Local calibrations can be developed for any geographic region in the United States using one spectral library with quality reference data.

Calibration transfer is where a calibration is moved from one instrument to another and statistical accuracy and precision are retained (Workman and Mark, 2017). In general, calibrations obtained using one instrument are usually unsuitable for use in a second instrument due to dissimilarities between spectral point spacing and range, abscissa shift, and ordinate response (Soriano-Disla et al., 2014). Transfer of calibrations developed using the KSSL spectral library to another instrument would need to have a similar optical bench and method for direct calibration transfer to be successful.

The main purpose of this study was to pilot the application of MIR-DRS coupled with PLSR in soil survey. The objective was to develop a procedure for use in soil survey field offices where soil samples are processed, MIR spectra are acquired (using a compact single-sample spectrometer), and predictions obtained using calibration models developed from the KSSL spectral library. To test the procedure, an MIR pilot project was set up in a soil survey field office where MIR predictions were compared with conventional laboratory measurements. The feasibility of direct calibration transfer of models built from spectra obtained from the KSSL production spectrometer to spectra obtained from a compact single-sample field office spectrometer was tested. The soil properties investigated in this pilot study were total C, organic C, CaCO_3 equivalent, total clay, CEC, 1500 kPa water, and pH in water and CaCl_2 , which are soil properties typically characterized in soil survey. The information is used in soil interpretations, conservation planning, and soil classification and mapping.

MATERIALS AND METHODS

Soil Survey Database and Mid-Infrared Library

The National Cooperative Soil Survey Soil Characterization Database is housed at the KSSL in Lincoln, NE. The database contains more than 38,000 pedons, with measured chemical and physical properties representing geographically diverse soils from across the conterminous United States, Hawaii, and Alaska. Many soil samples from these characterization studies were archived as air-dried (35–37°C), <2-mm sieved soil and are stored in 0.5-L wax-lined cardboard containers. Since 2011, all samples entering the laboratory are MIR scanned. Archive samples were also being scanned in geographic areas of special interest. The ever-growing KSSL MIR library is made up of spectra from over 80,000 samples.

Spectra in the KSSL MIR library were scanned from soil samples that were air-dried (35–37°C), passed through a 2-mm sieve, and ground to <180 μm . For mineral soils, ~15 g of the <2-mm soil was ground using a Planetary Ball Mill with silicon nitride bowls and balls (Pulverisette 5, Fritsch). For organic soils, ~50 g of the <2-mm soil was ground using a Wiley mill (Pulverisette 16, Fritsch). The fine ground sample was stored in a 20-mL glass scintillation vial. To acquire spectra of a sample, about 40 mm³ of the fine ground sample was loaded into a 6-mm-diameter well of a 96-well spot plate in quadruplicate. The sample in each well was lightly pressed using a 6-mm-diameter, flat-bottomed press rod. A micro-vacuum system was used to remove excess soil around the wells. Each well (with sample) was then scanned in sequential manner on the VERTEX 70/HTS-XT Fourier transform infrared

spectrometer (Bruker Optics Inc.) to acquire spectra from 7500 to 600 cm^{-1} in the diffuse reflectance mode and optimized for the MIR range. The MIR spectra range below 600 cm^{-1} contains wavebands with low signal-to-noise ratios and thus was not acquired (Waruru et al., 2015). The spectrometer was equipped with a mercury cadmium telluride detector cooled by liquid nitrogen. The spectral resolution was 4 cm^{-1} , and 32 scans per spectrum were collected. A reference (background) spectrum was collected before each sample using an empty well with an anodized aluminum bottom. No purge gas was used in the optical bench. Spectra were expressed in absorbance units [$\log(1/\text{reflectance})$]. For modeling, only the MIR range from 4000 to 600 cm^{-1} was used.

Measured (or “reference”) data from the characterization database that were used included: total clay (<0.002 mm) (pipette method, 3A1a1a), total C (dry combustion, 4H2a1–3a1), CaCO_3 equivalent (electronic manometer method, 4E1a1a1a1–2), organic C content (determined by difference between the total C and $\text{CaCO}_3\text{-C}$, 4E1a1a1a1–2), CEC (1.0 N NH_4OAc at pH7, 4B1a1a1a1a–b1), water retention at –1500 kPa (<2-mm sieved, 3C2a1a–b), pH (in 1:1 water, 4C1a2a1a–b1, and 0.01 M CaCl_2 in 1:2 soil/solution suspensions, 4C1a2a2a–b1), and effervescence class (1.0 N HCl, 1B1b2d4). All determinations were made on air-dried (30–35°C, 1B1b2), crushed, and sieved (<2 mm, 1B1b2b) soil samples and are reported on an oven-dry basis (air-dry/oven-dry ratio, 1B1b2b4). All methods are described in Soil Survey Staff (2014) by the alphanumeric codes (shown above in parentheses). Information on the soil taxonomic classifications (i.e., soil order) and horizon designations (e.g., Ap, Bt) from the soil profile descriptions were also used and are defined in Soil Survey Staff (2014b).

Data Used and Stratification for Model Development

The National Soil Survey Center is in the Central Plains of the United States where Mollisols are a dominant taxonomic soil order. These soils were chosen to be modeled because of their proximity to the KSSL and to the field office in Salina, KS, where the pilot project will be conducted. Spectra and associated measured data for Mollisols from 15 states in the central United States (Colorado, Iowa, Illinois, Kansas, Minnesota, Missouri, Montana, North Dakota, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, Wisconsin, and Wyoming) were queried from the database (KSSL MIR library) and are referred to hereafter as the “Mollisols dataset.” The majority of the pedons are in Kansas and Nebraska (Fig. 1), which is the anticipated area where most of the samples for the pilot project will be collected. Properties

modeled were total C, organic C, CaCO_3 equivalent, total clay, CEC, 1500 kPa water, and pH in water and CaCl_2 . Preliminary analysis and previous studies (Janik et al., 1998; Minasny et al., 2009; Wijewardane et al., 2018) have shown that these properties can be predicted successfully by MIR-DRS. The Mollisols dataset was further stratified by A and B master horizons for each property (Table 1) to improve prediction accuracy, as shown by Wijewardane et al. (2018).

Model names in Table 1 indicate the stratification of the calibration dataset. In the model name, master horizon is listed first followed by the property being predicted. Next, the model name reflects whether the dataset was further stratified based on clay content or on the presence or absence of effervescence.

To reduce model errors for total clay and 1500 kPa water, further stratification (or subsetting) of the data was used. Janik and Skjemstad (1995) found that subsetting the data significantly improved linearity and precision of several soil properties. In the present study, for total clay in B horizons, the B_clay_over_20 model was developed because preliminary analysis showed an improvement in the RMSE if a separate calibration was built for soils with >20% clay. If the general B_clay model predicted a clay content >20%, then the B_clay_over_20 model was used to predict the clay content. For 1500 kPa water in B horizons, the B_1500_kPa_water_≥35_clay model was developed because preliminary analysis showed an improvement in the overall RMSE when a separate calibration was built for soils >35% clay. When the general B_clay model predicted clay >35%, then the B_1500_kPa_water_≥35_clay model was used to predict the 1500-kPa water content.

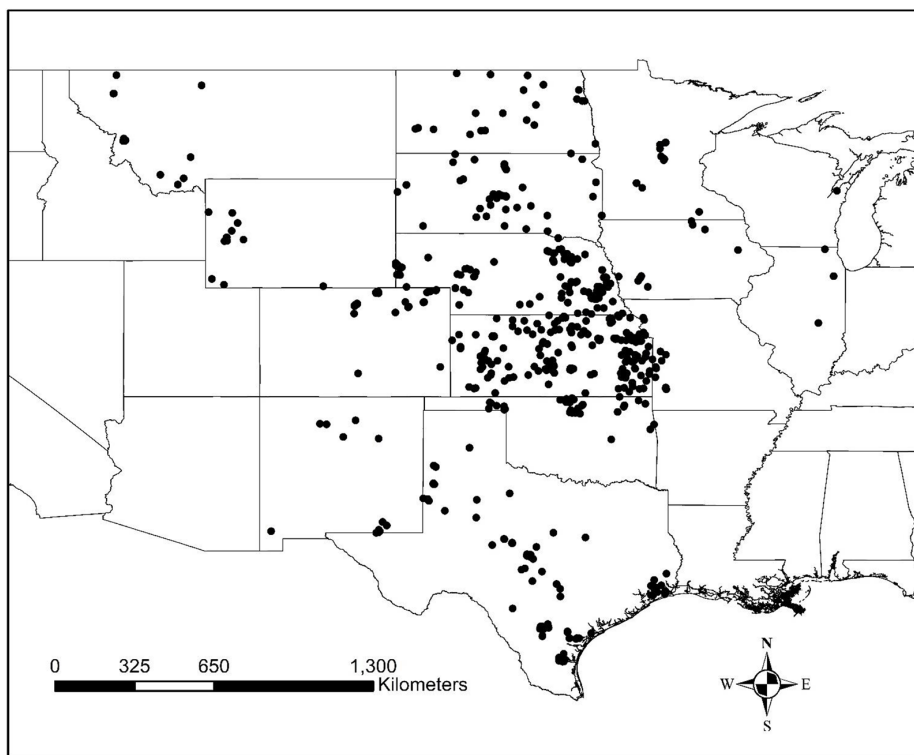


Fig. 1. The central United States showing the location of Mollisol pedons that were used in model calibration and validation from the MIR library.

Table 1. Stratification of data for each mid-infrared (MIR) model calibration/validation dataset obtained from the MIR library and characterization database.

Soil property	MIR model name	Data stratification for model development
Total C, %	A_Total_Carbon	Mollisol + A horizon + total C > 0.1%
	B_Total_Carbon	Mollisol + B horizon + total C > 0.1%
Organic C, %	A_Organic_Carbon	Mollisol + A horizon + OC > 0.1%
	B_Organic_Carbon	Mollisol + B horizon + OC > 0.1%
CEC, cmol _c kg ⁻¹	A_CEC	Mollisol + A horizon
	B_CEC	Mollisol + B horizon
CaCO ₃ equiv., %	A_Carbonates	Mollisol + A horizon + CaCO ₃ equiv. > 0.1%
	B_Carbonates	Mollisol + B horizon + CaCO ₃ equiv. > 0.1%
Total clay, %	A_Clay	Mollisol + A horizon
	B_Clay	Mollisol + B horizon
	B_Clay_Over_20	Mollisol + B horizon + total clay > 20%
1500 kPa water, %	A_1500_kPa_water	Mollisol + A horizon
	B_1500_kPa_water	Mollisol + B horizon
	B_1500_kPa_water_≥ 35_Clay	Mollisol + B horizon + total clay ≥ 35%
pH in water	A_Water_pH_Negative_Effervescence	Mollisol + A horizon + negative effervescence
	A_Water_pH_Positive_Effervescence	Mollisol + A horizon + positive effervescence
	B_Water_pH_Negative_Effervescence	Mollisol + B horizon + negative effervescence
pH in CaCl ₂	B_Water_pH_Positive_Effervescence	Mollisol + B horizon + positive effervescence
	A_CaCl ₂ _pH_Negative_Effervescence	Mollisol + A horizon + negative effervescence
	A_CaCl ₂ _pH_Positive_Effervescence	Mollisol + A horizon + positive effervescence
	B_CaCl ₂ _pH_Negative_Effervescence	Mollisol + B horizon + negative effervescence
	B_CaCl ₂ _pH_Positive_Effervescence	Mollisol + B horizon + positive effervescence
Presence of carbonates score (yes = 1, no = -1)	A_Effervescence_Score	Mollisol + A horizon + CaCO ₃ equiv. > 0.5%
	B_Effervescence_Score	Mollisol + B horizon + measured CaCO ₃ Equiv. > 1%
		Excluded- positive efferv. w/no measured carbonates Excluded- measured carbonates w/no positive efferv.

For the prediction of pH (both in water and CaCl₂), preliminary analysis showed an improvement in the RMSE if the calibration dataset was stratified based on the presence of carbonates, and spectra were stratified accordingly before model development (Table 1). A calibration model was needed to predict the presence of carbonates, and thus an “effervescence score” calibration model was developed for the A and B horizons. Samples with measurable carbonates (CaCO₃ equivalent) were assigned a score of +1, and samples without carbonate were assigned a score of -1. For B horizons, the effervescence class was used to assist in determining measurable carbonates (Table 1). A predicted positive effervescence score for an unknown sample indicates the presence of carbonates, and a negative score indicates its absence. The corresponding pH models were then used to predict pH.

The CaCO₃ equivalent models were developed using a calibration dataset consisting of only soils containing measurable carbonates (Table 1). The effervescence score models were used to predict the presence of carbonates and to determine whether to accept the prediction from the CaCO₃ equivalent models. If a sample had a predicted positive effervescence score, the prediction from the CaCO₃ equivalent model was accepted. If a predicted negative effervescence score was obtained, the sample contained no carbonates.

Model Development

A calibration model was developed for each subset (or strata) of data (Table 1) for a total of 24 models. Calibrations were developed using PLSR in OPUS QUANT2 software (Bruker Optics Inc.) in the mid-infrared region 4000 to 600 cm⁻¹. In

OPUS, the calibration spectra were first mean centered. Then, using principle component analysis, spectrally redundant samples were identified and removed based on 20 weighting factors. Redundancy in 20 weighting factors was required to more conservatively eliminate redundant samples. The remaining spectra were refactorized and divided into spectrally equivalent calibration and test sets (50/50 split) using the Kennard–Stone algorithm. The only exception was that cross-validation (leave-one-out) was used for total clay of A horizons to take advantage of all the calibration reference data and to obtain the smallest RMSE. The PLSR models were optimized by successively trying a combination of 12 preprocessing methods and 40 predefined wavelength ranges. Preprocessing reduces the impact of the non-relevant spectral information and often leads to more robust calibration models. The combination of preprocessing method and wavelength range that produced the smallest RMSE was selected to refine and validate a model. The resulting preprocessing method varied among the soil properties, as was also shown by Minasny and McBratney (2008). Up to 1.5% of the calibration data points were identified as outliers and removed. Models were validated with the test dataset. During model validation, the percentage of outliers removed ranged from ~0 to 1.5%. Outliers were identified in OPUS based on ANOVA of residuals and an F probability > 0.99. The number of PLS loadings was optimized to minimize the RMSE while ensuring that the RMSE of prediction (RMSE_p) for the test set was approximately equal to the RMSE of estimation (RMSE_E) of the calibration set to avoid overfitting the calibration set. Supplemental Table S1 contains

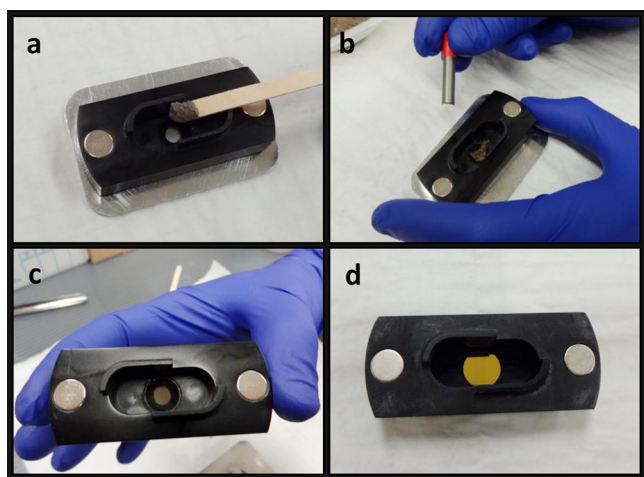


Fig. 2. (a) A spatula tip full of fine ground soil ($<180\ \mu\text{m}$) is being loaded into the 5-mm hole of the ALPHA sample holder that is sitting on an aluminum plate. (b) A flat-bottom press rod (6-mm diameter) is used to lightly press the sample in the hole of the sample holder. (c) The sample holder showing the pressed sample after removal of excess soil. (d) The reference holder showing the gold mirror reference in the center.

the summary statistics of the calibration and validation datasets (from the MIR library) for each model.

Field Office Spectrometer and Calibration Transfer

A compact, single-sample Fourier transform infrared spectrometer (ALPHA, Bruker Optics Inc.) was used to acquire MIR spectra in the field office. Before deployment to the field office, the spectrometer was tested at the KSSL to assess the feasibility of direct calibration transfer from the VERTEX 70 to the ALPHA. The ALPHA was equipped with the front-reflectance module (Bruker Optics Inc.). Spectral resolution was $4\ \text{cm}^{-1}$, and 128 scans per spectrum were acquired from 4050 to $550\ \text{cm}^{-1}$. A reference (background) spectrum was collected before each sample spectrum to help correct for atmospheric and instrument drift. Scanning time was about 2 min each for sample and reference. To assess the feasibility of direct calibration transfer, the ALPHA was tested with 120 soil samples that had also been scanned on the VERTEX 70 spectrometer. The test samples were selected from the Mollisols dataset to reflect the full range in texture, pH, carbonate content, and organic C content (Supplemental Table S2). The selection resulted in 52 A horizons and 68 B horizons.

Fine ground samples were scanned on the ALPHA to mimic the scanning procedure for the VERTEX 70 spectrometer. The sample holder of the ALPHA was placed on a clean, flat, aluminum plate ($5\ \text{cm} \times 7.5\ \text{cm}$) (Fig. 2a). A spatula tip full of fine ground sample (about $60\ \text{mm}^3$) was loaded and lightly pressed, using a 6-mm diameter flat-bottomed press rod, into the 5-mm-diameter hole of the sample holder (Fig. 2a–c). The sample holder was then inverted and lightly tapped on the aluminum plate to remove excess soil around the pressed sample in the hole. Samples were scanned one at a time on the ALPHA. Between samples, the sample holder was wiped cleaned using a clean tissue and cotton swab without the use of water or heat. A reference scan was obtained using a gold mirror reference (Fig. 2d), and

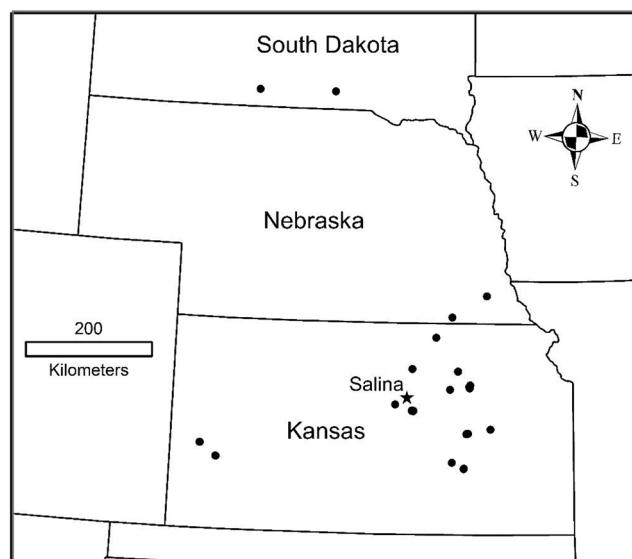


Fig. 3. Location of pedons (or cluster of pedons) collected by the soil survey field office located in Salina, KS, to be mid-infrared scanned on the ALPHA. Because of scale, some of the pedons located near each other could not be separated on the map.

then the sample was scanned. One spectrum was obtained per sample. The goal was to have the most efficient method possible and to keep the scanning time to a minimum.

Spectra from the ALPHA and VERTEX 70 were processed through the models. The predicted results from the four replicates from the VERTEX 70 were averaged. In OPUS, outliers in the predicted results were identified by the Mahalanobis distance limit, which is determined based on the distribution of all calibration spectra. The distribution mean was calculated, and 1 SD was used as the limit.

Field Office Mill Calibration

A compact, two-sample ball mill (Retsch Mixer Mill MM200, Verder Scientific) was used in the field office to grind samples. The Mixer Mill MM200 was optimized to grind samples to the same particle size as the KSSL Planetary Ball Mill by adjusting milling times and frequencies. The fine grinds from the Mixer Mill MM200 and KSSL Planetary Ball Mill were assessed for spectral similarity. The similarity of spectra was compared using OPUS Quick Compare software version 7.2 (Bruker Optics Inc.). Twelve test samples from the KSSL archive, ranging in texture from sand to clay (Supplemental Table S3), were used in the optimization procedure.

A 5-cm^3 scoop of the $<2\text{-mm}$ soil was loaded into a 25-mL milling cup (stainless steel grinding jars, Mixer Mill MM200) with five 10-mm stainless steel grinding balls. The ALPHA spectrometer was used to obtain spectra of the fine ground test samples derived from both mills; then, spectra were compared using the OPUS Quick Compare software feature. After a few iterations, a grinding time of 15 min at 25 Hz on the Mixer Mill MM200 produced grinds that were $\geq 99.3\%$ spectrally similar to the Planetary Ball Mill grinds for all 12 test samples (Supplemental Table S3).

Soil Sample Collection and Processing in Field Office

There were 195 unknown samples from 44 pedons collected by field office staff as part of the pilot study (Fig. 3). These samples served as a completely independent test set (prediction dataset) for assessing model performance because samples were not part of model calibration and validation. Soil samples were collected and described during routine soil survey activities as described in Schoeneberger et al. (2012). The range of soil classifications are presented in Supplemental Table S4.

In the soil survey field office, soil samples were air-dried at 35 to 37°C for 3 d. The air-dried soil was then passed through a 2-mm sieve and dried again at 35 to 37°C for 1 d. The soil was thoroughly mixed using gloved hands to homogenize it and stored in a 0.5-L wax-lined paper container. A 5-cm³ subsample of the <2-mm soil was finely ground on the Mixer Mill MM200 (15 min at 25 Hz) and stored in a 20-mL glass scintillation vial.

Obtaining Predictions in the Field Office

Spectra were obtained using the ALPHA spectrometer set-up in the field office using the procedures described above. Field staff started out by acquiring only one spectrum per sample as instructed. After 30 samples were acquired, field staff experimented with collecting replicate spectra per sample (scanning subsamples of the fine grinds) to potentially improve prediction accuracy relative to measured data obtained on the same samples at the KSSL. Scanning of four subsamples showed the most improved accuracy. Thus, four subsamples were scanned for the remainder of the study.

Two related spectral data sets were processed through the calibration models using OPUS QUANT2 software (Bruker Optics Inc.). The first dataset consisted of one spectrum from each of the 195 samples collected (single-spectra dataset). When four spectra were acquired, only the first spectrum collected was used. The second dataset consisted of 660 spectra collected from 165 samples that were scanned in quadruplicate (four-spectra dataset). Each spectrum was processed through the calibration models, and results were averaged for each sample for the four-spectra dataset. Direct calibration transfer from the VERTEX 70 to the ALPHA was assumed.

After processing the spectra through the QUANT2 software, all predictions were transferred to an Excel spreadsheet for further processing using a macro developed with the Visual Basic Editor. Decisions on which model predictions to use were programmed into the macro, making processing of predictions efficient and free of operator errors. The programming rules consider master horizon designation, the presence of carbonates, predicted clay content, and whether or not a predicted value was a Mahalanobis distance outlier as it prepares the final report.

Conventional Laboratory Analyses

The <2-mm and fine ground samples were sent to the KSSL for conventional and spectral analysis. The suite of properties determined on the <2-mm samples were total clay (<0.002 mm), total C, CaCO₃ equivalent, organic C content, CEC, water retention at 1500 kPa, pH (in water and 0.01 M CaCl₂), effervescence class, air-dry/oven-dry ratio, and MIR scans (after grind-

ing to <180 µm on Planetary Ball Mill). All methods are previously described. The MIR scans were also obtained on the field office fine ground samples. All MIR scans were acquired on the VERTEX 70/HTS-XT spectrometer in quadruplicate as previously described. The measured data were compared with the predicted values from spectra acquired on the ALPHA located in the field office. Predictions acquired from the VERTEX 70 and ALPHA spectrometers were also compared to confirm the validity of assuming direct calibration between the spectrometers.

ALPHA Location Bias

Environmental or field office conditions (e.g., operator variability) were evaluated as a potential source of bias between the two locations of the ALPHA (KSSL and field office). Predictions acquired from the ALPHA when it was located in the KSSL were compared with predictions from the same ALPHA when it was located in the field office. Twenty-five soil samples from the KSSL archive were selected from the 120 samples that were used to test the ALPHA at the KSSL. Subsamples (from the KSSL <2-mm air-dried soil in a 0.5-L container) were ground at the KSSL and spectra acquired on ALPHA (located in KSSL). Another set of 25 subsamples were ground (from KSSL <2 mm air-dried soil in 0.5 L container) in the field office for spectral data collection using the same ALPHA unit (located in field office). Predictions were obtained from one spectrum per sample.

Statistics

Measured versus predicted results were compared using the RMSE, mean error (ME), and R^2 . The RMSE is a measure of the spread of data around the least squares line and is calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{\text{pred}} - y_{\text{obs}})^2}$$

where y_{pred} is the predicted value and y_{obs} is the observed or measured value. The RMSE_E is designated for the calibration and RMSE_p is designated for the validation. The ME is an average measure of the departure from the least squares line and is calculated as:

$$\text{ME} = \frac{1}{N} \sum_{i=1}^N (y_{\text{pred}} - y_{\text{obs}})$$

The R^2 is a measure of how close the data are to the least squares line and summarizes the explanatory power of the model. The ratio of the SD to the RMSE_p or RPD is often used in MIR studies. It has been shown that the residual prediction deviation (RPD) is a function of the R^2 (Minasny and McBratney, 2013); thus, only R^2 is used here. Also, RPD may be inappropriate for use on samples with non-normal distributions (Bellon-Maurel et al., 2010). Higher R^2 values, lower RMSE values (for a given model), and MEs closer to zero indicate more accurate predictions. This assumes the slope of the least squares line (regression line) is not different than one (slope of the 1:1 line).

Lin's concordance correlation coefficient (r_c) was used to determine the degree of agreement between the measured and predicted variables by measuring variation from the 1:1 line (Lin,

1989; Minasny and McBratney, 2008; Tranter et al., 2008). It contains measurements of both accuracy and precision. The r_c values range from 0 to 1, with higher values indicating greater agreement and 1 representing complete agreement between all measured and predicted pairs. Also, r_c was used to measure the agreement between predictions from the two spectrometers. It is calculated for paired datasets as:

$$r_c = \frac{2s_{xy}}{s_x^2 + s_y^2 + (\bar{x} - \bar{y})^2}$$

where

$$s_x^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$$

$$s_y^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2$$

$$s_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

where $\bar{x}$ and $\bar{y}$ are the sample means for populations X and Y , and x_i and y_i are paired i th values from populations X and Y (Lin, 1989).

Partial least squares can give negative predictions of soil properties (Minasny and McBratney, 2008) when the measured values are very close to zero, such as for organic C, total C, and total clay. For predictions from the field office, there were only a

few slightly negative values, which were set to zero before calculating statistics. No outliers were removed unless specified.

RESULTS AND DISCUSSION

Descriptive statistics of the soil properties (Mollisols dataset) are shown in Supplemental Table S1 for both the calibration and validation datasets for each model. Calibration soil samples were diverse, with soils ranging in total clay from 0.4 to 71%. Organic C ranged up to 7%, and soil pH values ranged from 3.8 to 9. Calcium carbonate equivalent ranged from not present to 51%. The samples from the Mollisols dataset provided a wide range in soil properties for modeling.

Mid-Infrared Model Calibration and Validation

Table 2 shows the goodness of model fit (R^2 , RMSE, and ME) for the prediction of soil properties for Mollisols of the central United States for both calibration and validation datasets (from the MIR library). The number of samples includes the four replicates of each sample. For model development, replicate spectra were averaged in the dataset before model development. Individual replicate calibration and test spectra were always individually processed through the models. Validation plots of predicted versus measured data are shown in Supplemental Fig. S1a–v. In validation plots of measured versus predicted properties, the r_c ranged from 0.750 for pH in water (positive effervescence) for B horizons to 0.996

Table 2. Results of each model calibration and validation. Models were developed from spectra obtained from the mid-infrared library and VERTEX 70 spectrometer.

Model	Units	Calibration			Validation				
		N	R^2	RMSE _E †	N	R^2	RMSE _P ‡	ME	r_c
A_Total_Carbon	%	2984	0.97	0.28	2978	0.97	0.27	−0.011	0.984
B_Total_Carbon	%	4240	0.97	0.16	4238	0.97	0.16	0.0059	0.984
A_Organic_Carbon	%	2507	0.97	0.22	2517	0.97	0.23	0.0054	0.985
B_Organic_Carbon	%	3556	0.93	0.12	3560	0.93	0.12	0.0006	0.962
A_CEC	cmol _c kg ^{−1}	2828	0.97	1.6	2845	0.96	1.6	−0.038	0.982
B_CEC	cmol _c kg ^{−1}	4154	0.95	1.8	4154	0.94	1.9	0.030	0.967
A_Carbonates	%	594	0.99	0.77	594	0.99	0.78	0.44	0.996
B_Carbonates	%	1885	0.99	0.32	1860	0.99	1.01	0.021	0.994
A_Clay	%	5925	0.96	2.2	5925	0.95	2.4§	−0.032	0.978
B_Clay	%	4307	0.93	3.3	4323	0.92	3.4	−0.22	0.915
B_Clay_Over_20	%	3811	0.93	2.7	3813	0.91	3.1	−0.030	0.955
A_1500_kPa_water	%	2979	0.91	1.7	2991	0.91	1.7	0.014	0.951
B_1500_kPa_water	%	4226	0.90	1.6	4236	0.89	1.7	0.037	0.941
B_1500_kPa_water_≥ 35_Clay	%	2259	0.75	1.9	2256	0.72	2.1	0.11	0.844
A_Water_pH_Negative_Effervescence		1935	0.86	0.26	1932	0.86	0.27	−0.014	0.921
A_Water_pH_Positive_Effervescence		793	0.86	0.19	784	0.84	0.20	−0.0064	0.899
B_Water_pH_Negative_Effervescence		1406	0.73	0.32	1405	0.69	0.34	−0.017	0.819
B_Water_pH_Positive_Effervescence		2157	0.63	0.24	2507	0.60	0.25	−0.0002	0.750
A_CaCl2_pH_Negative_Effervescence		1935	0.79	0.31	1937	0.76	0.34	0.021	0.863
A_CaCl2_pH_Positive_Effervescence		792	0.75	0.25	804	0.72	0.28	−0.0026	0.835
B_CaCl2_pH_Negative_Effervescence		1407	0.77	0.31	1406	0.72	0.34	0.023	0.840
B_CaCl2_pH_Positive_Effervescence		2040	0.66	0.19	2040	0.63	0.19	0.0046	0.781
A_Effervescence_Score		3026		0.21	3022		0.22	−0.0012	0.856
B_Effervescence_Score		3569		0.23	3560		0.24	0.011	0.970

† Root mean square error (RMSE) of estimation.

‡ RMSE of prediction.

§ RMSE of cross validation.

for CaCO_3 equivalent in A horizons (Table 2). The higher the r_c value, the more accurate the model. In general, models for CaCO_3 equivalent, total C, organic C, CEC, total clay in A horizons, and 1500 kPa water in A horizons predicted very well ($r_c = 0.967$ – 0.996), and models for total clay in B horizons and 1500 kPa water in B horizons predicted fairly well ($r_c = 0.844$ – 0.955). Models for pH predicted the least well ($r_c = 0.750$ – 0.921).

There was slight curvature observed in validation plots of measured versus predicted total C and organic C of the A horizons and CEC of the B horizons (Supplemental Figures S1a, S1c, and S1f). Nonlinear relationships between concentrations and MIR intensities are known to occur in MIR-DRS spectroscopy (Janik and Skjemstad, 1995; Janik et al., 2007a). They have usually been corrected or accounted for by transforming the intensities with a Kubelka–Munk function (Brimmer and Griffiths, 1988) or converting the reflectance spectra to absorbance using the $\log(1/\text{reflectance})$ transformation (Janik et al., 1998; Jia et al., 2017). If curvature still exists, the reference data can be transformed (e.g., $\log_e$, $\sqrt{x}$, $1/x$) before calibration development and then back transformed (Baldock et al., 2013; Janik et al., 2007a, 2007b; Waruru et al., 2015) to linearize the relationship. Janik et al. (2007a) indicated that the curvature can result from a different chemistry or sample composition corresponding to low concentration values compared with high concentration values. Predictions using local linear models developed from subsets of data were effective in removing regression nonlinearities in MIR studies by Bornemann et al. (2008), Janik and Skjemstad (1995), and Janik et al. (1995). The goal of a transformation is to increase the linear relationship between the measured and predicted values. Therefore, any linearization transformation should be realized in the improvement of the RMSE and r_c . Improving calibrations to remove regression curvature would be a goal for future modeling efforts.

Calibration Transfer

Preliminary results obtained using the ALPHA suggested that the calibrations built for the VERTEX 70 could be directly transferred to the ALPHA. The feasibility of direct calibration transfer was tested using 120 samples from the Mollisols dataset. Summary statistics for the test samples are presented in Supplemental Table S2. There is a wide range in properties: total

clay ranged from 0.4 to 77%, organic C ranged up to 6.4%, and CaCO_3 equivalent ranged from not present to 34%.

Measured versus predicted results from the ALPHA and VERTEX 70 spectrometers are shown in Table 3. For all properties, the RMSEs were slightly higher for predictions obtained using the ALPHA than using the VERTEX 70 spectrometer (Table 3). This suggests that the laboratory method using the VERTEX 70 spectrometer outperformed the field method using the ALPHA (as expected), but the loss in model performance was small. Prediction errors for C were within 0.02%, whereas for total clay errors were within 0.7% between the spectrometer methods.

The strength of agreement between the ALPHA predictions versus the VERTEX 70 predictions was tested using Lin's concordance correlation coefficient (Table 3). Predictions of CaCO_3 equivalent had an $r_c > 0.990$, which is classified as "almost perfect" according to McBride (2005) in the strength-of-agreement classification of new laboratory methods (compared with the "gold standard"). Prediction agreement for total C, organic C, total clay, and pH in CaCl_2 had an $r_c > 0.950$ and ≤ 0.990 , which is classified as "substantial," whereas prediction agreement for CEC and 1500 kPa water were classified as "moderate." An almost perfect strength-of-agreement suggests that direct calibration transfer can be used with little loss in accuracy.

Greater prediction errors for a property will contribute to a lower strength-of-agreement between spectrometers. This was shown when the number of replicate spectra (from one to four) acquired on the ALPHA resulted in lower prediction errors and improved agreement (i.e., higher r_c values) for all properties between the spectrometers (Table 4). Total C, organic C, and CaCO_3 equivalent had an $r_c > 0.990$, indicating an "almost perfect" agreement. Total clay, CEC, and 1500 kPa gravimetric water content had a "substantial" agreement, with an $r_c > 0.95$ and < 0.99 . Only pH had a lesser agreement between spectrometers. Properties with low MIR prediction errors, such as CaCO_3 equivalent and C, are advantageous to use in assessing calibration transfer because the confounding errors associated with prediction and direct calibration transfer are reduced.

Calibration transfer is an area of ongoing research. In the absence of a proven protocol for performing calibration transfer, the above results suggest provisional acceptance of direct calibration transfer for all properties. Efforts are underway to find a calibration transfer that will result in lower prediction errors between the spectrometers (J. Sanderman, personal communication, 2019).

Table 3. Comparison of prediction errors and agreement between spectrometers (VERTEX 70 and ALPHA). Test of calibration transfer on 120 samples.

Property	VERTEX 70, measured vs. predicted, RMSE	ALPHA, measured vs. predicted, RMSE	ALPHA predictions vs. VERTEX 70 predictions, r_c
Total C, %	0.27	0.29	0.985
Organic C, %	0.25	0.26	0.988
CEC,† $\text{cmol}_c \text{ kg}^{-1}$	1.9	3.0	0.943
CaCO_3 equivalent, %	0.9	1.0	0.992
Clay, %	2.8	3.5	0.978
1500 kPa water, %	1.58	2.18	0.901
pH (water)	0.28	0.53	0.886
pH (CaCl_2)	0.49	0.54	0.951
Effervescence score	0.3	0.5	0.906

† Cation exchange capacity.

Location and Operator Bias

The ability of field office staff versus laboratory staff to successfully follow the sample drying, processing, scanning, and prediction procedures was assessed by the evaluation of errors obtained in both locations, on the same ALPHA, using a standard set of 25 samples. The r_c of the measured versus ALPHA-predicted properties between locations were nearly the

same (Table 5). Total C, CEC, CaCO₃ equivalent, and total clay predicted slightly better at the KSSL with slightly higher r_c values, whereas organic C, 1500 kPa water, and pH predicted better at the field office. The error appears to be random between locations, suggesting there was very little location or operator bias. With the exception of pH, the r_c of ALPHA predictions from KSSL versus ALPHA predictions from the field office were >0.9. Calcium carbonate equivalent had a near-perfect strength of agreement between the two locations ($r_c = 0.996$). The near-perfect strength of agreement indicates that location and operator bias are contributing very little to the prediction error. As indicated previously, higher prediction errors of the other properties contribute to the confounding effect of location and operator bias.

Differences in environmental conditions and variability between locations were also a concern because the water content of air-dried soils varies based on the relative humidity (RH). The amount and type of clay affects the amount of water adsorbed under a given RH, which only takes a couple of minutes to equilibrate when exposed to a humidity change (Hatch et al., 2012). Water also has a strong spectral response. Total clay, CEC, and 1500 kPa water predictions are based, in part, on the clay mineral signatures in the spectra and thus would be affected by variation in RH and air-dried water contents. The high r_c values (>0.9) for measured versus predicted total clay, CEC, and 1500 kPa between locations indicate that this was not an issue.

Soil Samples Collected for Prediction: Field Office

Soil samples collected by the field office staff (prediction dataset) were predominantly from pedons classified as Mollisols, with one Alfisol, Inceptisol, and Vertisol (Supplemental Table S4). Soil classifications indicate a wide range of properties within limited parent materials and clay mineralogy. There were shallow soils, soils containing natric horizons, paleosols, high carbonate horizons, high clay horizons, and high shrink-swell layers. The dominant taxonomic family mineralogy classes were smectitic and mixed. The soil parent materials were predominantly loess, and a few had been formed from clayey sediments. Some soils had formed in alluvium or colluvium and some from residuum weathered from shales and limestone.

Summary statistics for the 195 samples collected are presented in Table 6. There was a wide range in properties. Total clay ranged from 1 to 74%, and CaCO₃ equivalent ranged from not present to 75%. Several of the argillic ho-

Table 4. Comparison of predictions from one versus four subsamples scanned on ALPHA using Lin's concordance correlation coefficient (r_c). Agreement between VERTEX 70–predicted and ALPHA–predicted soil properties.

Property	ALPHA, measured vs. predicted, r_c		VERTEX 70 predicted vs ALPHA predicted, r_c
	One subsample	Four subsamples	
Total C, %	0.984	0.988	0.993
Organic C, %	0.972	0.981	0.990
CaCO ₃ equivalent, %	0.968	0.992	0.997
Total clay, %	0.897	0.910	0.965
CEC,† cmol _c kg ⁻¹	0.905	0.936	0.978
1500 kPa water, %	0.917	0.943	0.950
pH (water)	0.566	0.697	0.857
pH (CaCl ₂)	0.818	0.847	0.932

† Cation exchange capacity.

Table 5. Comparison of predictions between locations of the ALPHA spectrometer ($n = 25$).

Property	ALPHA, measured vs. predicted, r_c		ALPHA predictions at KSSL vs. ALPHA predictions at field office, r_c
	At KSSL†	At field office	
Total C, %	0.971	0.969	0.986
Organic C, %	0.983	0.988	0.974
CEC,‡ cmol _c kg ⁻¹	0.950	0.948	0.930
CaCO ₃ equivalent, %	0.989§	0.985§	0.996§
Total Clay, %	0.975	0.967	0.963
1500 kPa water, %	0.940	0.942	0.953
pH (water)	0.681	0.821	0.691
pH (CaCl ₂)	0.722	0.800	0.889

† Kellogg Soil Survey Laboratory.

‡ Cation exchange capacity.

§ Based on 15 samples.

rizons (B horizons) contained high clay contents (up to 75%). Organic C ranged from <0.1% in the B horizons up to 5.1% in the A horizons, which is typical of Mollisols.

Soil Property Predictions: Field Office

Predicted versus measured soil properties for 165 samples (where four subsamples were scanned and results averaged) are shown in Fig. 4a–h. The lower the RMSEP and higher the R^2 and r_c values, the greater the prediction accuracies.

Carbon

Organic and total C showed similar prediction accuracies, with an RMSEP of 0.20 and 0.23%, respectively (Fig. 4a and 4b). The RMSEP for organic C is lower than that for total C because the property range of the organic C dataset is smaller and lower,

Table 6. Summary statistics for the soil samples collected by the field office (true test samples—not used in calibration/validation of models).

Property	<i>n</i>	Min.	Max.	Median	Mean	SD
Total C, %	195	<0.1	9.2	1.4	1.7	1.5
Organic C, %	195	<0.1	5.1	0.7	1.1	1.1
CEC,† cmol _c kg ⁻¹	195	2.1	46.7	22.4	22.7	8.7
CaCO ₃ equivalent, %	129	<0.1	74.8	4.0	7.8	12.0
Total clay, %	195	1.1	73.6	33.7	34.5	13.6
1500 kPa water, %	195	5.0	27.7	15.9	15.8	4.4
pH (water)	195	4.4	8.9	7.8	7.4	0.9
pH (CaCl ₂)	195	4.4	8.1	7.3	6.9	1.0

† Cation exchange capacity.

resulting in a smaller standard deviation (Table 6). The RMSE depends on the range of values in the dataset. The R^2 values were high at 0.97 for organic C and 0.98 for total C. The MEs were

small at 0.06% for organic C and 0.02% for total C; these values indicate a slight overall overprediction (Fig. 4a and 4b). The overprediction was greater for organic C and was more pronounced

at the higher organic C contents (>2%). Other researchers have indicated that the presence of carbonates can interfere with or reduce the ability to quantify organic C using PLSR (McCarty et al., 2002; Reeves et al., 2005). In our study, the r_c for soils containing $\geq 2\%$ carbonates was 0.964 ($n = 68$); the r_c for soils without carbonates or with <2% carbonates was 0.985 ($n = 97$). The higher prediction accuracy obtained on soils with <2% carbonates indicates that there could be some interference of carbonates in the prediction of organic C. Total C was also slightly overpredicted at the higher total C contents (Fig. 4b). The four Mahalanobis outliers that have >7% total C are high in inorganic C (53–75% CaCO_3 equivalent), which are outside the range of the calibration dataset. Incorporating these samples into the calibrations could improve predictions (Guerrero et al., 2016). The r_c was 0.981 for organic C and 0.984 for total C, indicating very good (or substantial) agreement between the measured and predicted values. Overall, the predictions for organic C and total C obtained in the field office were considered acceptable for use in soil survey.

Cation Exchange Capacity

The CEC had an RMSE_p of $2.7 \text{ cmol}_c \text{ kg}^{-1}$ and an R^2 of 0.90 (Fig. 4c). The ME is $-1.6 \text{ cmol}_c \text{ kg}^{-1}$, indicating an overall underprediction of CEC. As the CEC increases, the underprediction slightly increases over the entire range (Fig. 4c). A few outliers were underpredicted at the high CEC end. The four Mahalanobis outliers, which contain high carbonate contents, were slightly overpredicted by the CEC calibration models. The CECs at the low end were also slightly underpredicted. Subsetting the data or developing more locally weighted calibrations could be tried in future studies to improve predictions. Overall, the predictions for CEC were considered good, with an r_c of 0.936. Clay minerals (type and amount) and organic matter contribute to the prediction of CEC in pedotransfer functions, all of which have spectral responses. Therefore, a reasonable model performance was not unexpected. The predictions would be sufficiently accurate to separate soils into their family cation-exchange activity classes in soil taxonomy.

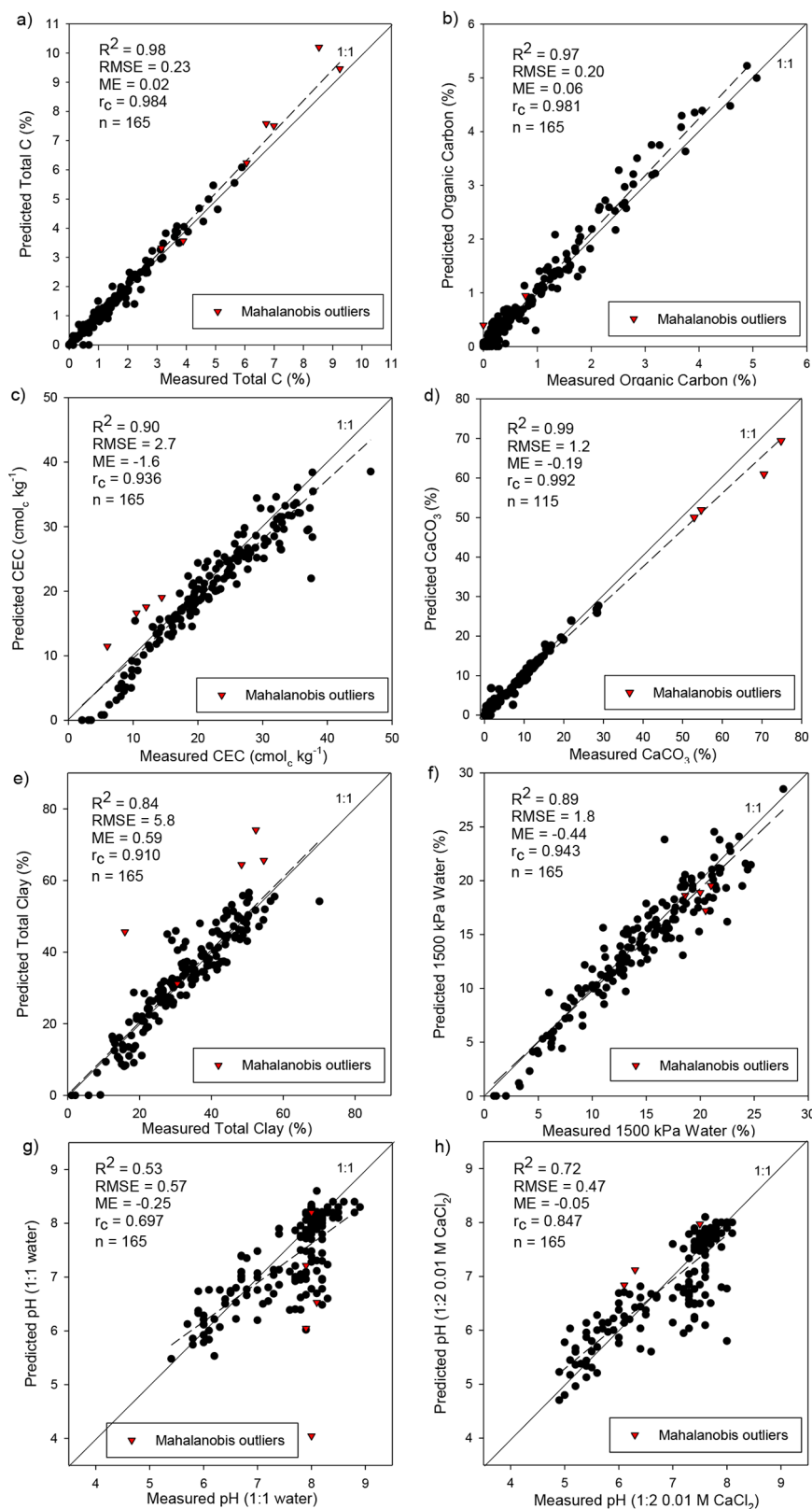


Fig. 4. Mid-infrared prediction of soil properties on samples collected by the field office. Predictions were obtained from spectra acquired on the ALPHA. Plots of measured versus predicted values are shown for (a) total carbon, (b) organic carbon, (c) cation exchange capacity (CEC), (d) CaCO_3 equivalent, (e) total clay, (f) 1500 kPa water, (g) pH in 1:1 water, and (h) pH in 1:2 0.01 M CaCl_2 .

The predictions were considered acceptable for use in soil survey. However, improving calibrations to improve the ME would be a goal of future modeling efforts.

Calcium Carbonate Equivalent

The CaCO_3 equivalent had an RMSE_p of 1.2% and an R^2 of 0.99 (Fig. 4d). The ME was -0.19% , indicating a slight underprediction. At about 15% CaCO_3 equivalent, the underprediction starts to increase as the carbonate content increases (Fig. 4d). The four Mahalanobis outliers in the plot are outside the CaCO_3 equivalent range of the calibration dataset, which means that extrapolation by the model was required to predict the high carbonate contents of these samples. Expanding the calibration dataset to include these and other higher-carbonate-content soils could improve predictions. Because carbonate's major peaks occur in a region relatively free from peaks of other soil components, it has high predictability (Janik and Skjemstad, 1995). Predictions for CaCO_3 equivalent were the best among all the soil properties, with an r_c of 0.992. The predictions of CaCO_3 equivalent were considered excellent and acceptable for use in soil survey.

Total Clay

Total clay had an RMSE_p of 5.8% and an R^2 of 0.84 (Fig. 4e). The ME was small at 0.59%, indicating an overprediction, which spanned along the entire range of the data (Fig. 4e). The high-carbonate samples (53–75% CaCO_3 equivalent) were overpredicted by the calibration models, which resulted in the four Mahalanobis outliers in Fig. 4e. The r_c of soils containing $\geq 5\%$ carbonates was 0.834 ($n = 62$); the r_c of soils containing carbonates or with $< 5\%$ carbonates was 0.955 ($n = 103$). This suggests that total clay in samples containing high carbonates is not predicted as well by the clay calibration models. Several samples with high carbonate contents were overpredicted by the clay calibration models (Fig. 4e). In laboratory measurements of total clay, both carbonate clay and silicate clay are included. The carbonate clay can be more difficult to predict and separate from the silt- and sand-sized carbonates in spectra because the difference is only due to particle size in unground samples. Adding the high-carbonate samples to the calibrations may improve predictions. Subsetting the data based on the presence or absence of carbonate or carbonate clays and developing calibration models for each may be another approach to improve predictions. Total clay was not predicted as well (moderate agreement), with an r_c of 0.910. Predictions of total clay were considered marginal for use in soil survey because a field soil scientist can detect a 3% difference in clay content through hand texturing. Taxonomic separations are based on being able to detect a 3% clay difference (Soil Survey Staff, 1999). Removing the five Mahalanobis outliers in the plot reduced the RMSE_p to 4.9%. The goal of future modeling efforts is to improve prediction accuracies for total clay to around 3%.

1500 kPa Water Contents

The 1500 kPa gravimetric water contents had an RMSE_p of 1.8% and an R^2 of 0.89 (Fig. 4f). The ME is -0.44% , indicating an overall slight underprediction of 1500 kPa water. As the 1500 kPa water increases past 7%, the underprediction slightly increases (Fig. 4f). The crossover point is at a 1500 kPa water content of 7%. There are several outliers scattered around the 1:1 and regression lines (Fig. 4f). Water retention measurements using pressure plates, especially at the more negative water potentials, are prone to measurement errors (Bittelli and Flury, 2009), which is most likely contributing to the prediction error and outliers. Water retention at 1500 kPa (using sieved samples) is highly correlated to the total clay content (Seybold and Harms, 2012) and thus would be linked to the spectral signatures of the clay minerals. Janik et al. (2007a) indicated that their predictions of water retention (volumetric) were based on the link between soil texture, bulk density, and the spectral signatures of some clay minerals, organic matter, and quartz. In our study, the 1500 kPa gravimetric water contents predicted better than that for CEC, with an r_c of 0.943. Predictions of 1500 kPa gravimetric water contents were considered acceptable for use in soil survey.

Soil pH

Soil pH in water and CaCl_2 have RMSE_p values of 0.57 and 0.47, respectively, and R^2 values of 0.53 and 0.72, respectively (Fig. 4g and 4h). The pH in CaCl_2 is predicted with greater accuracy than pH in water. Soil pH measured in a salt suspension (0.01 M CaCl_2) is relatively less variable throughout the season and analytically more stable and accurate (Kissel et al., 2009) and thus has a lower prediction error. The MEs were -0.25 for pH in water and -0.05 for pH in CaCl_2 . Of all the properties, soil pH in water had the poorest predictions, with an r_c of 0.697. Soil pH in CaCl_2 had the second poorest predictions, with an r_c of 0.847. The poor predictions for pH can be attributed to the fact that this property does not have direct spectral responses (Kuang and Mouazen, 2011). Soil pH or reaction classes as defined in the Soil Survey Manual (Soil Survey Division Staff, 2017) have ranges that span 0.4 to 0.7 pH units depending on the class. In the present study, predicted pH values would not be reliable in identifying the reaction class. Predictions of pH were considered inadequate for use in soil survey. Reducing errors would be a goal of future modeling efforts.

Comparison with Other Studies

Prediction errors for the suite of soil properties in our study were compared with prediction errors from other studies that have used MIR-DRS coupled with PLSR. Our prediction errors (RMSE_p) were generally lower for validation of models developed from datasets of large geographical areas (McCarty et al., 2002; Minasny et al., 2009; Pirie et al., 2005; Waruru et al., 2015) and global scale datasets (Bornemann et al., 2008; Terhoeven-Urselmans et al., 2010). Lower prediction errors were obtained by others for validation of models developed from field scale datasets (Bornemann et al., 2010; Viscarra-Rossel et al., 2006), where

the soil variability would be much lower than in our study. Xie et al. (2011) also obtained lower prediction errors for validation of models developed from a one-soil-type and parent material dataset where only soil management practices varied. Greater sample variability or diversity generally leads to higher prediction errors. Reeves and Smith (2009) have shown that extreme sample diversity from continental-scale studies spanning great distances with widely varied parent materials, land uses, and climate can result in poor predictions where suitable calibrations are not possible. In our study area (where sampled), the parent material was primarily loess, and clay mineralogy was predominantly smectite, thus limiting the soil variability. Our calibration models encompassed a much broader range in parent materials and mineralogy. Reducing the soil variability in the calibration dataset to only reflect the variability that will be encountered in the pilot study area would most likely improve predictions (Viscarra-Rossel et al., 2008). Also, we used a large sample size in our calibrations (on average $n = 668$; Supplemental Table S1), which may have created more reliable and representative models compared with models based on smaller datasets for similar-size geographical areas (Ng et al., 2018).

Quadruplicate versus Single Subsample Scanned on ALPHA

Results show that there is improved accuracy when soils are scanned in quadruplicate instead of acquiring just one spectrum per sample. Concordance correlations of predicted versus measured values for all soil properties show greater agreement (higher r_c) when quadruplicate scans were acquired instead of single scans (Table 4). The differences in r_c values are small, ranging from 0.004 units for total C to 0.131 units for pH in water. This suggests that there is within-sample variability of the finely ground sample that can be reduced by acquiring replicate scans and averaging the results. Also, there is variability in the packing density of the pressed soil pellet in the sample holder that could have been reduced by taking replicate scans. Scanning four subsamples increases the total scanning time from 5 to 20 min. Field staff preferred the improved accuracy from scanning samples in quadruplicate over minimizing the time spent scanning.

Predicted property results from one spectrum acquired per sample on the ALPHA versus conventional laboratory measured data for 195 soil samples from 44 pedons are shown in Table 7.

Table 7. Predicted property results from one spectrum acquired per sample on the ALPHA versus conventional laboratory measured data for 195 test samples (collected by field office).

Property	Measured vs. predicted, RMSE	Measured vs. predicted, mean error	Measured vs. predicted, r_c
Total C, %	0.26	0.00	0.985
Organic C, %	0.25	0.05	0.974
CaCO ₃ equivalent, %	2.7	-0.45	0.969
Total clay, %	5.9	0.62	0.915
CEC †, cmol _c kg ⁻¹	3.5	-1.6	0.891
1500 kPa water, %	2.2	-0.26	0.913
pH (water)	0.77	-0.26	0.578
pH (CaCl ₂)	0.56	-0.08	0.818

† Cation exchange capacity.

The r_c values ranged from 0.578 for pH in water to 0.985 for total C. Results indicate that moderate and greater prediction accuracies ($r_c > 0.90$) can be obtained when predictions are from one spectrum acquired per sample on the ALPHA. Depending on the circumstances, more rapid measurements may be desired, and one scan would be preferred. If C is the measurement focus and if there are many samples to run, one spectrum per sample would provide a substantial level of agreement ($r_c = 0.985$). Because carbonates can be highly segregated in the soil, there is a chance that its presence could be missed if only one subsample is scanned. It is recommended that at least duplicate scans be acquired on the ALPHA for CaCO₃ equivalent.

CONCLUSIONS

The objective of this study was to pilot the application MIR-DRS coupled with PLSR in a soil survey field office. Calibration models were built and validated for Mollisols of the central United States using the KSSL MIR spectral library. Models for CaCO₃ equivalent, total C, organic C, CEC, total clay in A horizons, and 1500 kPa in A horizons predicted very well, whereas models for total clay in B horizons and 1500 kPa water in B horizons predicted well. Models for pH did not predict as well. A field office method was developed for processing samples to <180 μm and acquiring MIR spectra using a compact single-sample spectrometer (ALPHA). Quadruplicate spectrum acquired per sample on the ALPHA reduced prediction errors by lessening the effect of sample heterogeneity and differences in packing density of the finely ground sample. Mid-infrared diffuse reflectance spectroscopy coupled with PLSR and principle component analysis was successful in predicting soil properties for completely independent samples that were collected, processed, and MIR scanned in a soil survey field office.

Not all properties predicted had an acceptable level of accuracy for use in soil survey. Total and organic C and CaCO₃ equivalent predicted very well, CEC and 1500 kPa water predicted moderately well, and all had an acceptable level of accuracy. Total clay had a marginal level of acceptability. Soil pH (in water and CaCl₂) did not have an acceptable level of accuracy. However, pH is relatively easy to directly measure, and this is not seen as a significant limitation in the use of MIR spectroscopy. Additional research is needed on lowering prediction errors for total clay and soil pH.

Direct calibration transfer of models developed from spectra acquired on the KSSL spectrometer (VERTEX 70) to spectra acquired on the ALPHA had an “almost perfect” agreement between predictions using Lin’s concordance correlations for those properties that predicted very well. The VERTEX 70 slightly outperformed the ALPHA in a comparison of prediction errors. In conclusion, direct calibration transfer was viable. A key factor to the success of direct calibration transfer was the consistency between the laboratory and field office sample preparation methods, optical benches, and a lack of inter-location bias. The MIR-DRS method described here can provide rapid predictions of soil properties in soil survey

field offices with a level of accuracy comparable to that of the KSSL MIR predictions.

SUPPLEMENTAL MATERIAL

Supplemental material is available with the online version of this article. Supplemental Table 1 contains summary statistics for each MIR model calibration and validation data subset (of the Mollics dataset). Supplemental Table 2 contains summary statistics of calibration transfer test dataset. Supplemental Table 3 contains the test of spectra similarity between fine grinds obtained from the Planetary Ball Mill and Mixer Mill. Supplemental Table 4 contains taxonomic classifications of soils sampled by the field office as part of the prediction dataset. Supplemental Figure 1 contains MIR model validations using spectra from the KSSL MIR library.

REFERENCES

- Baldock, J.A., B. Hawke, J. Sanderman, and L.M. Macdonald. 2013. Predicting contents of carbon and its component fractions in Australian soils from diffuse reflectance mid-infrared spectra. *Soil Res.* 51:577–595. doi:10.1071/SR13077
- Bellon-Maurel, V., E. Fernandez-Ahumada, B. Palagos, J.-M. Roger, and A. McBratney. 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *Trends Analyt. Chem.* 29(9):1073–1081. doi:10.1016/j.trac.2010.05.006
- Bittelli, M., and M. Flury. 2009. Errors in water retention curves determined with pressure plates. *Soil Sci. Soc. Am. J.* 73(5):1453–1460. doi:10.2136/sssaj2008.0082
- Bornemann, L., G. Welp, and W. Amelung. 2010. Particulate organic matter at the field scale: Rapid acquisition using mid-infrared spectroscopy. *Soil Sci. Soc. Am. J.* 74(4):1147–1156. doi:10.2136/sssaj2009.0195
- Bornemann, L., G. Welp, S. Brodowski, A. Rodionov, and W. Amelung. 2008. Rapid assessment of black carbon in soil organic matter using mid-infrared spectroscopy. *Org. Geochem.* 39:1537–1544. doi:10.1016/j.orggeochem.2008.07.012
- Brimmer, P.J., and P.R. Griffiths. 1988. Angular dependence of diffuse reflectance infrared spectra. III. Linearity of Kubelka-Munk plots. *Appl. Spectrosc.* 42:242–247. doi:10.1366/0003702884428293
- 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. doi:10.1016/j.geoderma.2005.04.025
- Cox, R.J., H.L. Peterson, J. Young, C. Cusik, and E.O. Espinoza. 2000. The forensic analysis of soil organic by FTIR. *Forensic Sci. Int.* 108:107–116. doi:10.1016/S0379-0738(99)00203-0
- D'Acqui, L.P., A. Pucci, and L.J. Janik. 2010. Soil properties prediction of western Mediterranean island with similar climatic environments by means of midinfrared diffuse reflectance spectroscopy. *Eur. J. Soil Sci.* 61(6):865–876. doi:10.1111/j.1365-2389.2010.01301.x
- Guerrero, C., J. Wetterlind, B. Stenberg, A.M. Mouazenc, M.A. Gabarrón-Galeote, J.D. Ruiz-Sinogad, R. Zornoza, and R.A. Viscarra-Rossel. 2016. Do we really need large spectral libraries for local scale SOC assessment with NIR spectroscopy? *Soil Tillage Res.* 155:501–509. doi:10.1016/j.still.2015.07.008
- Hatch, C.D., J.S. Wiese, C.C. Crane, K.J. Harris, H.G. Kloss, and F. Baltrusaitis. 2012. Water adsorption on clay minerals as a function of relative humidity: Application of BET and Freundlich adsorption models. *Langmuir* 28:1790–1803. doi:10.1021/la2042873
- Holmgren, A., and B. Norden. 1988. Characterization of peat samples by diffuse reflectance FT-IR spectroscopy. *Appl. Spectrosc.* 42(2):255–262. doi:10.1366/0003702884428284
- Janik, L.J., R.H. Merry, S.T. Forrester, D.M. Lanyon, and A. Rawson. 2007a. Rapid prediction of soil water retention using mid infrared spectroscopy. *Soil Sci. Soc. Am. J.* 71(2):507–514. doi:10.2136/sssaj2005.0391
- 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. doi:10.1071/EA97144
- Janik, L.J., and J.O. Skjemstad. 1995. Characterization and analysis of soils using mid-infrared partial least squares. II. Correlations with some laboratory data. *Aust. J. Soil Res.* 33:637–650. doi:10.1071/SR9950637
- Janik, L.J., J.O. Skjemstad, and M.D. Raven. 1995. Characterization and analysis of soils using mid-infrared partial least squares: I. Correlations with XRF-determined major element composition. *Aust. J. Soil Res.* 33:621–636. doi:10.1071/SR9950621
- Janik, L.J., J.O. Skjemstad, K.D. Shepherd, and L.R. Spouncer. 2007b. The prediction of soil carbon fractions using mid-infrared-partial least square analysis. *Aust. J. Soil Res.* 45:73–81. doi:10.1071/SR06083
- Jia, X., S. Chen, Y. Yang, L. Zhou, W. Yu, and Z. Shi. 2017. Organic carbon prediction in soil cores using VNIR and MIR techniques in an alpine landscape. *Sci. Rep.* 7:2144. doi:10.1038/s41598-017-02061-z
- Kasprzhitskii, A., G. Lazorenko, A. Khater, and V. Yavna. 2018. Mid-infrared spectroscopic assessment of plasticity characteristics of clay soils. *Minerals (Basel)* 8:184. doi:10.3390/min8050184
- Kissel, D.E., L. Sonon, P.F. Vendrell, and R.A. Isaac. 2009. Salt concentration and measurement of soil pH. *Commun. Soil Sci. Plant Anal.* 40:179–187. doi:10.1080/00103620802625377
- Knox, N.M., S. Grunwald, M.L. McDowell, G.L. Bruland, D.B. Myers, and W.G. Harris. 2015. Modelling soil carbon fractions with visible near-infrared (VNIR) and mid-infrared (MIR) spectroscopy. *Geoderma* 239–240:229–239. doi:10.1016/j.geoderma.2014.10.019
- Kuang, B., and A.M. Mouazen. 2011. Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms. *Eur. J. Soil Sci.* 62:629–636. doi:10.1111/j.1365-2389.2011.01358.x
- Lin, L. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45:255–268. doi:10.2307/2532051
- Madari, B.E., J.B. Reeves, III, P.L.O.A. Machado, C.M. Guimaraes, E. Torres, and G.W. McCarty. 2006. Mid- and near-infrared spectroscopic assessment of soil compositional parameters and structural indices in two Ferralsols. *Geoderma* 136:245–259. doi:10.1016/j.geoderma.2006.03.026
- Madejova, J., and P. Komadel. 2001. Baseline studies of the clay minerals society source clays: Infrared methods. *Clays Clay Miner.* 49(5):410–432. doi:10.1346/CCMN.2001.0490508
- McCarty, G.W., J.B. Reeves, III, 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.
- McBride, G.B. 2005. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. NIWA Client Report: HAM2005-062. Report to Ministry of Health. National Institute of Water and Atmospheric Research Ltd. Hamilton, New Zealand. www.medcalc.org/download/pdf/McBride2005.pdf (accessed 1 Nov. 2018).
- Minasny, B., and A.B. McBratney. 2008. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. *Chemom. Intell. Lab. Syst.* 94:72–79. doi:10.1016/j.chemolab.2008.06.003
- Minasny, B., G. Tranter, A.B. McBratney, D.M. Brough, and B.W. Murphy. 2009. Regional transferability of mid-infrared diffuse reflectance spectroscopic prediction for soil chemical properties. *Geoderma* 153:155–162. doi:10.1016/j.geoderma.2009.07.021
- Minasny, B., and A. McBratney. 2013. Why you don't need to use RPD. *Pedometron* 33:13–14.
- Ng, W., B. Minasny, B. Malone, and P. Filippi. 2018. In: search of an optimum sampling algorithm for prediction of soil properties from infrared spectra. *PeerJ* 6:e5722. doi:10.7717/peerj.5722
- Pirie, A., B. Singh, and K. Islam. 2005. Ultra-violet, visible, near-infrared, and mid-infrared diffuse reflectance spectroscopic techniques to predict several soil properties. *Aust. J. Soil Res.* 43:713–721. doi:10.1071/SR04182
- Reeves, J.B., III, B.A. Francis, and S.K. Hamilton. 2005. Specular reflection and diffuse reflectance spectroscopy of soils. *Appl. Spectrosc.* 59(1):39–46. doi:10.1366/0003702052940431
- Reeves, J.B., III, G.W. McCarty, and V.B. Reeves. 2001. Mid infrared diffuse reflectance spectroscopy for the quantitative analysis of agricultural soils. *J. Agric. Food Chem.* 49:766–772. doi:10.1021/jf0011283
- Reeves, J.B., III, and D.B. Smith. 2009. The potential of mid- and near-infrared diffuse reflectance spectroscopy for determining major- and trace-element concentrations in soils from a geochemical survey of North America. *Appl. Geochem.* 24:1472–1481. doi:10.1016/j.apgeochem.2009.04.017
- Schoeneberger, P.J., D.A. Wysocki, and E.C. Benham, and Soil Survey Staff. 2012. Field book for describing and sampling soils. Version 3.0 National Resources Conservation Service. National Soil Survey Center, Lincoln, NE. www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs142p2_052523.pdf (accessed 2 May 2019).

- Seybold, C.A., and D.S. Harms. 2012. Estimation of -1500 kPa water retention in soil survey. *Soil Sci.* 177(4):229–235. doi:10.1097/SS.0b013e31824826af
- 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. doi:10.2136/sssaj2002.9880
- Shepherd, K.D., and M.G. Walsh. 2007. Infrared spectroscopy: Enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. *J. Near Infrared Spectrosc.* 15:1–19. doi:10.1255/jnirs.716
- Siebielec, G., G.W. McCarty, T.I. Stuczynski, and J.B. Reeves, III. 2004. Near- and mid-infrared diffuse reflectance spectroscopy for measuring soil metal content. *J. Environ. Qual.* 33:2056–2069. doi:10.2134/jeq2004.2056
- Soil Survey Division Staff. 2017. Soil survey manual. Agriculture handbook No. 18. USDA, Washington, DC.
- Soil Survey Staff. 1999. Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys. 2nd ed. Agric. Handb. 436. US Gov. Print. Office, Washington, DC.
- Soil Survey Staff. 2011. Soil survey laboratory information manual. Soil Survey Investigations Report No. 45, Version 2.0. USDA–NRCS, Washington, DC.
- Soil Survey Staff. 2014. Kellogg soil survey laboratory methods manual. Soil Survey Investigations Report No. 42. USDA–NRCS, Washington, DC.
- Soil Survey Staff. 2014b. Keys to soil taxonomy, 12th ed. USDA–NRCS, Washington, DC. www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/class/?cid=nrcs142p2_053580 (accessed 2 May 2019).
- Soriano-Disla, J.M., L.J. Janik, R.A. Viscarra Rossel, L.M. Macdonald, and M.J. McLaughlin. 2014. The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. *Appl. Spectrosc. Rev.* 49(2):139–186. doi:10.1080/05704928.2013.811081
- Terhoeven-Urselmans, T., T.G. Vagen, O. Spaargaren, and K.D. Shepherd. 2010. Prediction of soil fertility properties from a globally distributed soil mid-infrared spectral library. *Soil Sci. Soc. Am. J.* 74(5):1792–1799. doi:10.2136/sssaj2009.0218
- Tranter, G., B. Minasny, A.B. McBratney, R.A. Viscarra-Rossel, and B.W. Murphy. 2008. Comparing spectral soil inference systems and mid-infrared spectroscopic predictions of soil moisture retention. *Soil Sci. Soc. Am. J.* 72:1394–1400. doi:10.2136/sssaj2007.0188
- Tinti, A., V. Tugnoli, S. Bonora, and O. Francioso. 2015. Recent applications of vibrational mid-Infrared (IR) spectroscopy for studying soil components: A review. *J. Cent. Eur. Agric.* 16(1):1–22. doi:10.5513/JCEA01/16.1.1535
- Viscarra-Rossel, R.A., D.J.J. Valvoort, 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. doi:10.1016/j.geoderma.2005.03.007
- Viscarra-Rossel, R.A., T.S. Jeon, I.O.A. Odeh, and A.B. McBratney. 2008. Using a legacy soil sample to develop a mid-IR spectral library. *Aust. J. Soil Res.* 46:1–16. doi:10.1071/SR07099
- Workman, J., Jr., and H. Mark. 2017. Calibration transfer chemometrics, part I: Review of the subject. *Spectroscopy (Springf.)* 32(10):18–24.
- Waruru, B.K., K.D. Shepherd, G.M. Ndegwa, A. Sila, and P.T. Kamoni. 2015. Application of mid-infrared spectroscopy for rapid characterization of key soil properties for engineering landuse. *Soils Found.* 55(5):1181–1195 [erratum: 55(6):1520]. doi:10.1016/j.sandf.2015.09.018
- Wijewardane, N.K., Y. Ge, S. Wills, and Z. Libohova. 2018. Predicting physical and chemical properties of US soils with a mid-infrared reflectance spectral library. *Soil Sci. Soc. Am. J.* 82:722–731. doi:10.2136/sssaj2017.10.0361
- Xie, H.T., X.M. Yang, C.F. Drury, J.Y. Yang, and X.D. Zhang. 2011. Predicting soil organic carbon and total nitrogen using mid- and near-infrared spectra for Brookston clay loam soil in Southwestern Ontario, Canada. *Can. J. Soil Sci.* 91:53–63. doi:10.4141/cjss10029
- Zovko, M., D. Romic, C. Colombo, E. Di Iorio, M. Romic, G. Buttafuoco, and A. Castrignano. 2018. A geostatistical Vis-NIR spectroscopy index to assess the incipient soil salinization in the Neretva River valley, Croatia. *Geoderma* 332:60–72. doi:10.1016/j.geoderma.2018.07.005