**ORIGINAL ARTICLE**

European Journal of **Soil Science** **WILEY**

# Performance of field-scale lab vs in situ visible/near- and mid-infrared spectroscopy for estimation of soil properties

Isabel Greenberg[1] | Michael Seidel[2,3] | Michael Vohland[2,3,4] |
Bernard Ludwig[1]

[1]Department of Environmental Chemistry, University of Kassel, Witzenhausen, Germany

[2]Geoinformatics and Remote Sensing, Institute for Geography, Leipzig University, Leipzig, Germany

[3]Remote Sensing Centre for Earth System Research, Leipzig University, Leipzig, Germany

[4]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

**Correspondence**
Isabel Greenberg, Department of Environmental Chemistry, University of Kassel, Nordbahnhofstr. 1a, 37213 Witzenhausen, Germany.
Email: isabel.greenberg@uni-kassel.de

**Abstract**

Comparison of laboratory versus in situ visible/near- (visNIR) and mid-infrared (MIR) spectroscopy for prediction of various soil properties is required to demonstrate trade-offs between accuracy and efficiency. Field measurements were made on an arable field in Germany (silt loam Haplic Luvisol) using visNIR (ASD FieldSpec 3 Hi-Res) and MIR (Agilent Technologies 4300 Handheld FTIR) and material was collected for lab visNIR (Foss XDS Rapid Content Analyzer) and MIR (Bruker-TENSOR 27) measurements on dried and ground soil and determination of total, labile (>63 μm light), stabilized (>63 μm heavy + <63 μm oxidizable) and resistant organic carbon (OC) content, total nitrogen ($N_t$), pH and texture. Partial least squares regression models were calculated for five repeated partitions of the dataset ($n = 238$) into training (75%) and test (25%) sets. Lab spectral models outperformed in situ models for total OC (root mean squared error [RMSE] = 0.24–1.0 g kg$^{-1}$), $N_t$ (RMSE = 0.026–0.088 g kg$^{-1}$), pH (RMSE = 0.12–0.28) and texture (RMSE = 0.53–1.5%). For both lab and field spectra, the accuracy of visNIR models was comparable or slightly better than MIR for sand, silt and clay. Spectral estimations for labile (RMSE = 0.34–0.47 g kg$^{-1}$) and stabilized OC (RMSE = 0.41–0.85 g kg$^{-1}$) were slightly (lab spectra) to substantially (field spectra) inferior to estimations from multiple linear regressions using total OC, $N_t$, clay and pH as predictors. Variable importance in the projection scores elucidated differences in spectral prediction mechanisms by spectrometer and OC fraction, and found mineral spectral signatures were highly important. For this field-scale study with 14% median soil gravimetric water content (GWC), the loss of accuracy from lab to field measurement was lower for visNIR than MIR. Analysis of the driest soils (<9% GWC) found field MIR outperformed field visNIR for OC and $N_t$ estimation and vice versa for the wettest soils (>18%), demonstrating the moisture dependence of performance rankings.

**Highlights**

- Lab vs field visible/near- (visNIRS) and mid-infrared (MIRS) spectroscopy require comparison for prediction of soil C fractions, N, pH and texture.
- Lab MIRS prediction of C, N and pH were superior, while texture estimations were comparable or slightly inferior to lab visNIRS.
- At 14% median soil water content, the loss of accuracy from lab to field measurement was lower for visNIRS than MIRS.
- The ranking of field visNIRS vs MIRS performance for C and N estimation is moisture dependent.

**KEYWORDS**

carbon fractions, handheld, labile carbon, mid infrared, near infrared, partial least squares regression, portable, soil organic carbon, soil spectroscopy, stabilized carbon

# 1 | INTRODUCTION

Past research has proven the potential of visible/near- (visNIR) and mid-infrared (MIR) spectroscopy (visNIRS and MIRS) to collect a wide range of information about temporally and spatially heterogeneous soil properties (Soriano-Disla et al., 2014). While IR spectroscopy may be less accurate than traditional laboratory methods, measurement is rapid, non-destructive and cheaper (after investment in a spectrometer), and requires no harmful chemicals (England & Viscarra Rossel, 2018). Thus, the lower accuracy of spectroscopy is compensated by the possibility of collecting more measurements in space and time with the same resources compared to traditional measurements.

In situ measurement of spectra using field devices saves additional resources by eliminating the need for soil transport and preparation (i.e., sieving, grinding and drying). However, the accuracy of estimations may suffer due to soil moisture, structure and heterogeneity (Stevens et al., 2008). For field visNIRS, increasing soil moisture causes, for example, pronounced absorption features at 1400 and 1900 nm, but also a general decrease in the overall albedo (Ben-Dor, 2002; Stenberg et al., 2010). For field MIRS, the most apparent absorption feature of water appears as a broad band centred around 3400 cm$^{-1}$. However, increasing soil moisture also affects the overall MIR reflectance spectrum nonlinearly and in a more pronounced way than the visNIR reflectance spectrum, and thus tends to mask or overlap the spectral signatures of other soil components to a greater extent (Janik et al., 2016; Reeves et al., 2010; Silvero et al., 2020). The performance of field visNIRS has been comparatively well-examined, but there are only a few published studies using portable MIR devices with in situ measurement (e.g., Hutengs et al., 2019; Ji et al., 2016; Reeves et al., 2010).

In addition to soil properties important for classification (e.g., texture) and plant productivity (e.g., total Nitrogen [$N_t$] content, pH), visNIRS and MIRS have applications in monitoring soil organic carbon (OC) dynamics (Greenberg et al., 2020), with implications for both soil quality and climate change. Soil OC content is an important determinant of soil quality due to its role in nutrient cycling, aggregate stability, water infiltration and erosion prevention (Wiesmeier et al., 2019), and is therefore an indicator of land degradation (Decision 22/COP.11; UNCCD, 2013). Furthermore, as soil is the largest reactive C pool in terrestrial ecosystems (Lal, 2013), sequestration of C in the soil is recognized as a critical climate change mitigation strategy by both international organizations (e.g., the Intergovernmental Panel on Climate Change [Edenhofer, 2014]) and national legislation (e.g., the Australian Government's Carbon Credits Act 2011). A 2018 amendment to the latter specifically states that sensors can be used to monitor soil C sequestration on agricultural land in their carbon credit scheme (Australian Government, 2018). Thus, the possibility to accurately and cost-effectively monitor soil OC contents with IR spectroscopy has important implications for the feasibility of creating financial incentives to improve soil management.

Not only measurement of total OC contents, but also quantification of OC fractions with varying residence times is useful for understanding dynamics (von Lützow et al., 2007) since changes in total OC contents occur slowly and are small in comparison to the bulk storage, making these changes difficult to detect (Necpálová et al., 2014). The soil fractionation method of Zimmermann et al. (2007), which applies both physical (dispersion, sieving, density separation) and chemical (oxidation by NaOCl) methods of separation was compared to 20 other fractionation methods and found to have a superior ability to separate fractions with distinct turnover times, low redundancy of the fractions and high OC

recovery and reproducibility (Poeplau et al., 2018). However, due to the time-consuming nature of fractionation procedures, estimation of OC fractions following model calibration using paired laboratory and spectral measurements is desirable.

Several studies have successfully estimated the OC content of fractions for dried and ground soils with lab visNIR and MIR spectrometers. In some cases, these studies have used partial least squares regression (PLSR) coefficients (Baldock et al., 2013; Madhavan et al., 2017) or loadings of the latent variables (Zimmermann et al., 2007; Knox et al., 2015) to indicate the existence of distinct spectral signatures for labile, intermediate and resistant OC fractions. For example, aliphatic peaks around 3000–2800 $cm^{-1}$ indicate the presence of labile OC compounds, whereas aromatic peaks around 1600–1500 $cm^{-1}$ indicate recalcitrant OC (Nocita et al., 2015). However, the often high correlation between fraction OC contents and other soil properties with known spectral signatures (e.g., total OC, clay) also enables indirect estimation of OC fractions, making the predictive mechanisms unclear in some cases (Ludwig et al., 2016). Models for indirectly estimated properties may be less robust since the relationships between properties differ between soils due to other influencing factors (Soriano-Disla et al., 2014). While some studies have reported successful estimation of OC fractions by lab visNIRS and/or MIRS (Baldock et al., 2013; Linsler et al., 2017; Madhavan et al., 2017), comparison with in situ measurement would reveal whether quantification of OC fractions can be made even more efficient.

The objective of our study was therefore (i) to compare the prediction accuracy of PLSR models calculated with multiple partitions of a field-scale sample into training and test sets for a range of bulk soil properties as well as the OC content of soil fractions using standard methods for laboratory versus in situ visNIR and MIR devices. In addition, (ii) the relative usefulness of spectral estimation of fraction OC contents was investigated by comparing PLSR prediction accuracy to that of multiple linear regressions (MLR) using total OC, $N_t$, clay and pH as predictors. Finally, (iii) variable importance in the projection (VIP) scores for PLSR models were calculated to provide insights into the predictive mechanisms for estimation of fraction OC contents by laboratory versus field visNIRS and MIRS.

# 2 | MATERIALS AND METHODS

## 2.1 | Field spectral measurements and sampling

The soil under investigation was an arable, silt loam Haplic Luvisol (16% clay, 80% silt and 4% sand) in

Lüttewitz (Saxony, Germany) (IUSS Working Group WRB, 2015). The site has an elevation of 290 m, annual average temperature of 8.6°C and precipitation of 572 mm (Koch et al., 2009). Management was consistent with standard agricultural practices, including conventional tillage with a moldboard plow to a depth of 30 cm.

Sampling was conducted over 5 days in September 2016. Wheat stubble remained on the field at the time of sampling. Sample points were laid out in a grid across a 52.5 m × 600 m homogeneously managed field (at the intersections of four columns spaced 17.5 m apart and 60 rows spaced 10 m apart, $n = 238$ due to missing field spectra for two sample points). At each sampling point, a 15 × 15 cm sampling area was cleared of wheat stubble and nine spectral measurements were made with each of the two portable spectrometers in direct contact with the soil surface. The individual spectra from each sampling point were then averaged to a mean spectrum. VisNIR spectra were measured using the ASD FieldSpec 3 Hi-Res (Malvern Panalytical, Analytik Ltd, Cambridge, UK) (350–2500 nm) in combination with a contact probe (measurement window of approximately 300 $mm^2$) and 50 co-added scans for each of the nine subsamples per observational unit. The instrument was calibrated against a Spectralon® reference panel at time intervals of ca. 15 min. MIR spectra were measured using the Agilent Technologies 4300 Handheld FTIR (Agilent Technologies, Santa Clara, California) (4000–650 $cm^{-1}$, spectral resolution set to 8 $cm^{-1}$) with a diffuse reflectance sampling interface (measurement window of approximately 3 $mm^2$) and 64 internal scans for each of the nine subsamples per observational unit. The MIR instrument was calibrated every 10 min using a coarse gold-plated reference cap. Following IR measurements, about 210 g soil was collected from each 15 × 15 cm sampling grid to a depth of 2 cm.

## 2.2 | Laboratory analysis and spectral measurements

For all measurements, soils were dried and sieved to <2 mm before analysis. Total C and N contents were analysed on ball-milled soils by dry combustion (Elementar Vario El, Heraeus, Hanau, Germany). Due to the absence of carbonates in the soil, total C was equivalent to total OC. pH was determined in a 0.01 M $CaCl_2$ solution (2.5 g soil per 6.25 ml). Soil texture was determined with the pipette method according to DIN ISO 11277 (2002).

OC fractions were separated by physical and chemical methods according to Zimmermann et al. (2007) for every

second sample point in the field ($n = 117$, three outliers were removed because the distribution of OC among the fractions was starkly different than for other sample points and there was insufficient material to repeat the fractionation). For this, 15 g of soil was sonicated in 75 ml $H_2O$ (Branson Digital Sonifier, Branson Ultrasonics Corporation, Dietzenbach, Germany) at an energy level of 22 J ml$^{-1}$ to break up macroaggregates. The soil was then wet-sieved using a 63-µm sieve, separating sand-size particles from silt- and clay-sized particles. Deviating from the original method, dissolved OC was not collected from the suspension at this stage due to the small size of this fraction (2% of total OC for Zimmermann et al., 2007) and its heterogeneous chemical structure (von Lützow et al., 2007). Analysis continued on the fraction >63 µm by separating particulate organic matter (>63 µm POM) from sand and stable aggregates (>63 µm S + A) with 1.8 g cm$^{-3}$ sodium polytungstate solution. Analysis continued on a 1 g aliquot of the <63 µm material by shaking it with 50 ml of 6% NaOCl solution (adjusted to pH 8 with HCl) in a 25°C water bath for 18 h to remove oxidizable C. This step was repeated two more times to isolate the <63 µm resistant fraction. The mass and OC content of these four fractions (i.e., >63 µm POM, >63 µm S + A, the total <63 µm fraction, and <63 µm resistant fraction) were measured to determine the allocation of total OC. OC contents of the <63 µm oxidizable fraction were derived from the total and resistant <63 µm fractions. For the purposes of spectral predictions, the four fractions were grouped into the following three fractions based on expected turnover time: labile OC (>63 µm POM fraction), stabilized OC (>63 µm S + A and <63 µm oxidizable fractions) and resistant OC (<63 µm resistant fraction) (Zimmermann et al., 2007; Poeplau et al., 2018).

Prior to lab spectral measurements on two replicates per observational unit, soil was dried and ball-milled to a particle size <0.2 mm using a Retsch MM 400 (Haan, Germany) with 10 zirconium oxide balls at 30 Hz for 5 min. VisNIR spectra in the range of 400–2500 nm (25,000–4000 cm$^{-1}$) were measured using a Foss XDS Rapid Content Analyzer (Silver Spring, MD, USA) at 2 nm resolution with 32 co-added scans on approximately 10 g of soil filled into a cell (5 cm diameter) with a quartz window. We re-calibrated the instrument every ca. 30 min using an internal white reference. Due to instrumental artefacts below 500 nm (Stevens et al., 2013), this region was excluded from the spectral ranges of both the lab and field visNIR spectrometers. Due to a detector change at 1100 nm for the lab spectrometer (Si detector from 400 to 1100 nm, PbS detector from 1100 to 2500 nm), 1092 to 1108 nm was also excluded.

Diffuse reflectance infrared Fourier transform spectra of the soils (approximately 1.5 g) in the range of 7000 to 370 to cm$^{-1}$ (1429–27,027 nm) were recorded with a Bruker-TENSOR 27 MIR spectrometer (Ettlingen, Germany) with an A562 integrating sphere detector and the diffuse-reflectance accessory (Ulbricht-Kugel, Ettlingen, Germany). The instrument was calibrated every hour with a gold reference background. The range from 7000 to 4000 cm$^{-1}$ (longwave NIR) was excluded from the analysis. The region <650 cm$^{-1}$ of the lab MIR spectrometer was also excluded in order to match the spectral ranges for lab and field MIR devices and because this region has limited usefulness due to overlapping mineral and organic absorption bands (Nocita et al., 2015). The spectra were measured with 200 scans at approximately 2 cm$^{-1}$ intervals.

For all spectra, the reflectance values of replicate measurements at each sampling point were averaged and converted to absorbance ($\log_{10}(1/\text{reflectance})$) for calculation of the PLSR models.

## 2.3 | Chemometric approach and performance measures

Analysis was carried out for five partitions of the complete dataset ($n = 238$ soils) into model training (75% of the complete dataset, that is, $n = 178$ for bulk soil properties and $n = 88$ for soil OC fractions) and model testing sets (25% of the complete dataset, i.e., $n = 59$ for bulk soil properties and $n = 29$ for soil OC fractions). Calculation of more model partitions (e.g., 100) would be preferable in order to determine the average and SD of model performance for each method with greater certainty; however, our analysis was limited to five partitions due to the time-consuming nature of modelling nine response variables for four spectrometers. The calibration and validation sets resulting from the five dataset partitions were identical across the four spectrometers tested. To investigate the effect of soil moisture on the estimation accuracy of field visNIR and MIR spectral models, analysis was also carried out with only the visNIR and MIR field spectra of the 30 soils with the highest and 30 soils with the lowest gravimetric water content (GWC).

To evaluate the performance of the models, root mean squared error (RMSE), bias and the ratio of performance to interquartile distance (RPIQ) were calculated. RPIQ was calculated rather than ratio of performance to deviation (RPD) due to the non-normality of most properties of interest. For comparability with performance measures given in other studies, consider that for a normally distributed variable and a large sample size, the interquartile range is 1.34896 × the SD (Ludwig

et al., 2019), and thus RPD = 1.40 is equivalent to RPIQ = 1.89 and $R^2 = 0.5$, where $R^2 = 1 - \frac{\text{Residual Sum of Squares}}{\text{Total Sum of Squares}}$ (Minasny & McBratney, 2013).

Separate PLSR analyses were performed for the visNIR and MIR spectral ranges with the statistical software R (version 3.4.4, R Core Team, 2018). Data pretreatment was carried out with the prospect package (Stevens & Ramirez-Lopez, 2013) and PLSR was performed using the pls package (Mevik et al., 2019). Data pretreatment first involved implementation of an automatic selection of all possible combinations of six regions making up the complete spectra. For visNIRS, these regions were 500–834 nm (region 1), 834–1167 (2), 1167–1500 (3), 1500–1833 (4), 1833–2166 (5), and 2166–2500 (6). For MIRS, these regions were 4000–3682.1 cm$^{-1}$ (region 1), 3682.1–3020.5 (2), 3020.5–2358.9 (3), 2358.9–1693.5 (4), 1693.5–1030 (5), and 1030–650 (6). Each of the 63 possible region combinations was then tested in conjunction with the following 13 data pretreatments: (i) use of the full spectra without manipulation, (ii–iv) calculation of moving averages (over 5, 11, 17 or 23 data points), and (vi–xiii) application of the Savitzky–Golay algorithm for the reduction of noise applied with the polynomial degree (PD) set to 2, the order of the derivative (DER) ranging from 1 to 2 (with PD-DER: 2–1 or 2–2), and a window smoothing size of 5, 11, 17 or 23. To determine the optimal number of latent variables, model training included a fivefold cross-validation, which has been shown to produce less biased and less variable error estimates compared to other resampling strategies (Beleites et al., 2005). The maximum number of latent variables was set to 15. In order to create a more robust, parsimonious PLSR model, the optimal number was determined by considering minimization of Akaike Information Criterion (AIC) (Viscarra Rossel & Behrens, 2010), calculated as:

$$\text{AIC}_{\text{PLSR}} = n \times \log_e(\text{RMSE}) + 2v,$$

where $n$ is the sample size and $v$ is the number of latent variables. The model with the optimal data pretreatment (i.e., resulting in the highest RPIQ) was then identified and tested using the remaining 25% of the complete dataset. This process was repeated for all five partitions of the complete dataset into model training and model testing sets.

To determine the relative importance of the wavelength predictors in the PLSR models for lab versus field estimation of total, labile, stabilized and resistant OC contents, variable importance in the projection (VIP) was calculated for wavelength $j$ as:

$$\text{VIP}_j = \sqrt{p \sum_{a=1}^{A} \left[ \text{SS}_a \left( \frac{w_{aj}}{\|w_a\|} \right)^2 \right] \bigg/ \sum_{a=1}^{A} (\text{SS}_a)},$$

where $p$ is the total number of predictors (i.e., wavelengths), $\text{SS}_a$ is the sum of squares explained by the $a$th component (i.e. latent variable), and $\left( \frac{w_{aj}}{\|w_a\|} \right)^2$ is the normalized importance of the $a$th component of the $j$th wavelength (Mehmood et al., 2012).

This analysis was carried out using the VIP() function of the plsVarSel package in R (Liland et al., 2020). VIP scores for a given spectrometer and target variable were calculated for each optimal model created from the five dataset partitions and then averaged for each wavelength. As the average of squared VIP scores equals 1 (Chong & Jun, 2005), graphical comparison of the wavelengths with VIP scores greater than 1 was used to elucidate differences in the prediction mechanisms by spectrometer and OC fraction.

## 2.4 | Descriptive statistics and multiple linear regressions

The statistical software R (version 3.4.4, R Core Team, 2018) was used for all analyses. Descriptive statistics were calculated and the Shapiro–Wilk test was performed to determine normality of the response variables.

MLRs were performed to compare the estimation accuracy of soil fraction OC content made by visNIRS- and MIRS-PLSR to indirect estimates based on soil properties with relationships with labile, stabilized and resistant OC content. For this, the complete dataset was separated into the five partitions of training and test sets identical to those used for PLSR analysis. We applied MLR with step-wise backward elimination starting from the full model using total OC, $N_t$, pH, clay content and all two-, three- and four-way interactions to predict the labile, stabilized and resistant OC content. The most appropriate model was determined using AIC calculated as:

$$\text{AIC}_{\text{MLR}} = n \times \log_e \left( \frac{\text{SSE}}{n} \right) + 2p,$$

where $n$ is the sample size, SSE is the sum of squares of the error, and $p$ is the number of predictors. The step command in the stats package was used to find the model that minimized $\text{AIC}_{\text{MLR}}$ compared to all potential models (Ripley, 2018). The residuals of the MLRs were checked for normality and homogeneity of variance. This MLR equation was then applied to estimate OC contents in the test set and the aforementioned performance measures were calculated.

To provide insights on the effects of soil moisture and texture on the accuracy of in situ spectral predictions, the

**TABLE 1** Descriptive statistics of the complete dataset for soil total organic carbon (OC) content as well as labile, stabilized and resistant OC fractions, total nitrogen ($N_t$) content, pH and texture

| Property | Sample size | Minimum | Maximum | Median | Mean | SD |
|---|---|---|---|---|---|---|
| Total OC (g kg$^{-1}$) | 238 | 7.7 | 18 | 11 | 12 | 1.9 |
| Labile OC (g kg$^{-1}$) | 117 | 0.61 | 3.1 | 1.6 | 1.7 | 0.54 |
| Stabilized OC (g kg$^{-1}$) | 117 | 6.0 | 12 | 8.4 | 8.3 | 1.2 |
| Resistant OC (g kg$^{-1}$) | 117 | 0.046 | 2.0 | 1.1 | 1.0 | 0.32 |
| $N_t$ (g kg$^{-1}$) | 238 | 0.93 | 1.9 | 1.3 | 1.3 | 0.17 |
| pH | 238 | 4.6 | 6.7 | 5.4 | 5.4 | 0.38 |
| Clay (%) | 238 | 9.9 | 22 | 15 | 16 | 2.2 |
| Silt (%) | 238 | 73 | 87 | 80 | 80 | 2.1 |
| Sand (%) | 238 | 2.5 | 7.0 | 3.8 | 4.1 | 0.93 |

residuals and absolute value of the residuals of OC and $N_t$ content estimates were correlated with GWC, clay content and sand content.

# 3 | RESULTS

## 3.1 | Spectra and descriptive statistics of soils

Given that this was a field-scale study, the range in each soil property was relatively low (Table 1). The range in total OC and $N_t$ content for the complete sample ($n=238$) was 7.7–18 g kg$^{-1}$ and 0.93–1.9 g kg$^{-1}$, respectively. Based on median OC contents of the $n=117$ subsample of fractionated soils, 14% of total OC was contained in the labile fraction, 73% in the stabilized fraction, and 9% in the resistant fraction. pH ranged from 4.6 to 6.7 and clay, silt, and sand content ranged 9.9%–22%, 73%–87% and 2.5%–7.0%, respectively. The distributions of soil properties were primarily slightly right-skewed and leptokurtic, with the exception of the normally-distributed stabilized OC fraction and left-skewed distributions of resistant OC and silt content. GWC for the collected soil ranged from 3.4% to 23%, with a median of 14%. Field spectra had lower and more variable reflectance in both the visNIR and MIR regions compared to lab spectra (Figure 1a). Strong absorption by water is apparent at 1400 and 1900 nm in the field visNIR spectra. Substantial loss of detail can be observed in the MIR field versus lab spectra.

## 3.2 | Average model performance and predictive mechanisms for soil OC content

For estimation of total, labile and stabilized OC contents, lab spectrometers decidedly outperformed field spectrometers (Table 2, Figures 2 and 3). Note that Figure 3 shows measured versus estimated values for the test set resulting in the median model performance; thus, although training and test sets were identical for the five partitions across the four spectrometers as well as the MLR analysis, the plotted test sets may differ. Average RMSE for the test sets for lab MIR and lab visNIR estimates of total OC content were 0.24 and 0.29 g kg$^{-1}$, respectively, and average RPIQ were 9.9 and 8.0, respectively, while field MIR and field visNIR had average RMSE of 1.0 and 0.83 g kg$^{-1}$ and RPIQ of 2.3 and 2.9, respectively. Labile OC content estimates based on lab spectra ranged from average RMSE of 0.34–0.37 g kg$^{-1}$ and RPIQ of 1.9–2.1, while performance of field devices range from average RMSE of 0.44–0.47 g kg$^{-1}$ and RPIQ of 1.5–1.6. For all spectrometers, estimates of stabilized OC content had higher RMSE than labile OC, but also higher RPIQ due to the wider range of contents for stabilized OC (6.0–12 g kg$^{-1}$) than labile OC (0.61–3.1 g kg$^{-1}$). For stabilized OC, performance of lab devices ranged from average RMSE of 0.41–0.45 g kg$^{-1}$ and RPIQ = 3.0–3.3, while field visNIR (average RMSE = 0.68 g kg$^{-1}$ and RPIQ = 2.0) outperformed field MIR (average RMSE = 0.85 g kg$^{-1}$, RPIQ = 1.6).

For the MLRs, total OC, $N_t$, clay and pH, as well as several two- and three-way interactions between these soil properties were useful for predicting fraction OC contents according to the stepwise simplification of the maximal models using AIC (Table 2). MLR estimates for labile and stabilized OC contents performed slightly better than the best spectral models. Finally, all models produced poor estimates for resistant OC content (average RMSE = 0.29–0.33 g kg$^{-1}$ and RPIQ = 1.1–1.3). The narrow range in resistant OC content (0.046–2.0 g kg$^{-1}$) is responsible for the lower RPIQ of this fraction compared to labile and stabilized fractions, since RMSEs for resistant OC were comparable and lower than that of labile
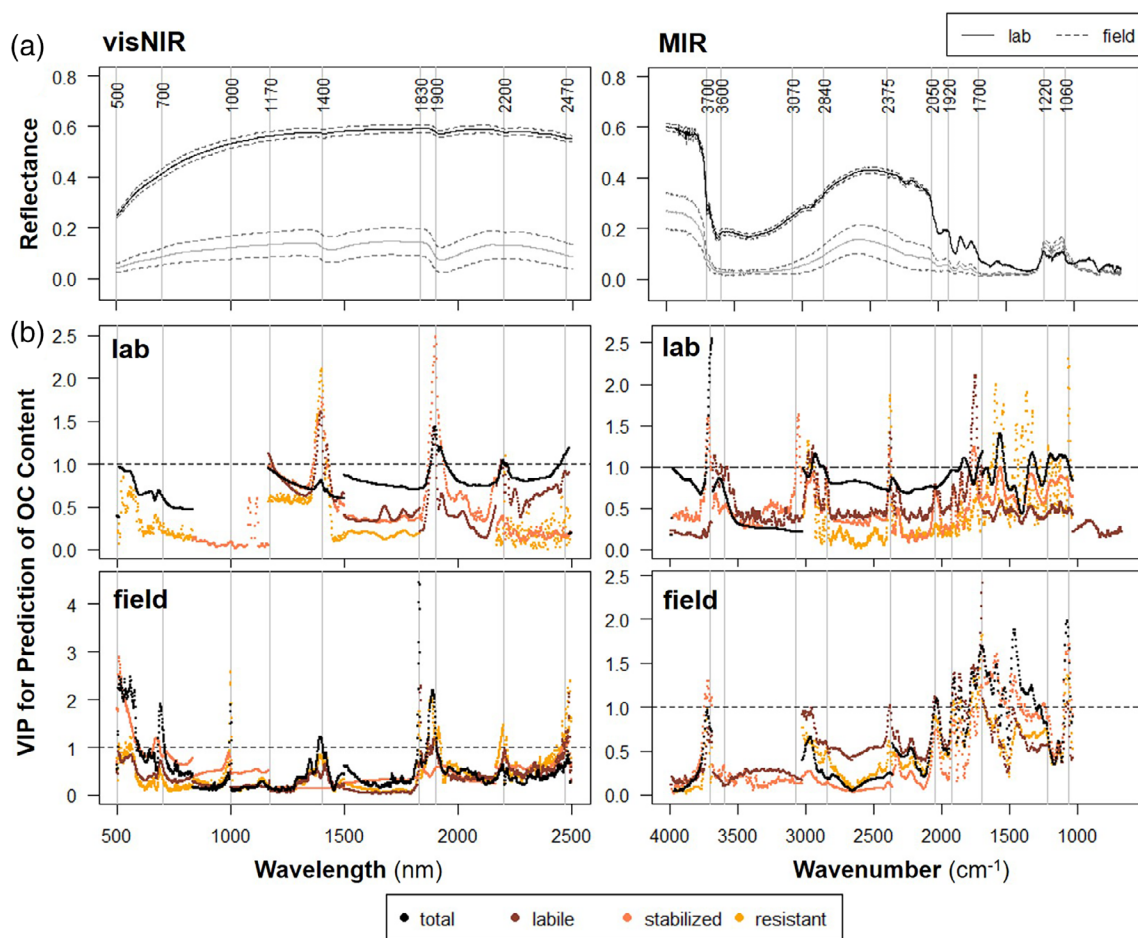
**FIGURE 1** (a) Average lab (black lines) and field (grey lines) visible/near- (visNIR) and mid-infrared (MIR) reflectance spectra of $n = 238$ soils. Dashed lines show $\pm 1$ SDs of the average. (b) Variable importance in the projection (VIP) scores for partial least squares regression models predicting total and fraction OC contents by lab and field visible/near- (visNIR) and mid-infrared (MIR) spectroscopy. The VIP scores shown are the averages of scores for the optimal models created in model training from five partitions of the complete dataset. Vertical grey lines demarcate bands and regions of importance. The dotted horizontal line at VIP = 1 is an importance threshold for ease of interpretation

and stabilized fractions, respectively. The average bias for total and fraction OC contents (Figure 2) showed no consistent trends by spectrometer and the bias for MLR estimates was comparable.

VIP analysis for visNIRS found absorbance from 500 to 700 nm was important for field visNIR prediction of total and stabilized OC (Figure 1b). Wavelengths around 1000 nm aided field visNIR prediction of total and resistant OC, 1170 nm helped with prediction of labile OC by lab visNIR, and 1400 nm was very important for prediction of all OC fractions by lab visNIR and somewhat important for total OC prediction by field visNIR. Absorbance around 1830 nm aided field visNIR prediction of total, labile and resistant OC and 1900 nm was useful for prediction of total, labile and stabilized OC for lab visNIR and total, labile and resistant OC for field visNIR. Absorption around 2200 nm aided prediction of total, stabilized and resistant OC for lab visNIR and

resistant OC for field visNIR and 2470 nm helped with prediction of total OC for lab visNIR and all OC fractions for field visNIR. For MIRS, 3700–3600 $cm^{-1}$ was important for predicting total, labile and stabilized OC for lab MIR and primarily stabilized OC for field MIR. 3070–2840 $cm^{-1}$ aided prediction of all fraction and total OC by lab MIR. The peak around 2375 $cm^{-1}$ was useful for prediction of all OC fractions by lab MIR and 2050 $cm^{-1}$ aided prediction of total and labile by field MIR. 1920–1700 $cm^{-1}$ was especially important for prediction of labile OC for both spectrometers, but also useful for total and all fraction OC prediction for field MIR and total and stabilized OC for lab MIR. 1700–1220 $cm^{-1}$ aided prediction of total and resistant OC for lab MIR and total and stabilized OC for field MIR. 1220–1060 $cm^{-1}$ helped with total and resistant OC prediction for lab MIR, and 1060 $cm^{-1}$ aided total and all fraction OC prediction for field MIR.

**TABLE 2** Parameterization and performance of field and lab visible/near- (visNIR) and mid-infrared (MIR) spectroscopy partial least squares regression (PLSR) models for the prediction of soil properties. Multiple linear regression (MLR) models were additionally calculated for the prediction of fraction organic carbon (OC) contents.

| Property | Predictor | Parameterization | | Testing | |
| | | Selected regions[a]/average data points or selected predictors[a] | Average factors | Average RMSE | Average RPIQ |
| --- | --- | --- | --- | --- | --- |
| Total OC (g kg$^{-1}$) | Lab MIR | 2-3-4-5-6/1255 | 12 | 0.235 | 9.94 |
| | Field MIR | 2-3-4-6/459 | 7 | 1.02 | 2.28 |
| | Lab visNIR | 1-3-4-5-6/694 | 15 | 0.294 | 8.01 |
| | Field visNIR | 1-2-3-4-5-6/1390 | 14 | 0.830 | 2.87 |
| Labile OC (g kg$^{-1}$) | Lab MIR | 1-2-3-4-5-6/963 | 7 | 0.374 | 1.90 |
| | Field MIR | 2-3-4-5-6/497 | 5 | 0.471 | 1.51 |
| | Lab visNIR | 3-4-5-6/430 | 8 | 0.335 | 2.10 |
| | Field visNIR | 1-2-3-4-5-6/862 | 10 | 0.437 | 1.60 |
| | MLR | OC, $N_t$, clay, pH, OC:$N_t$, OC:clay, $N_t$:clay, OC:$N_t$:clay | 5 | 0.307 | 2.28 |
| Stabilized OC (g kg$^{-1}$) | Lab MIR | 2-3-4-5-6/980 | 6 | 0.407 | 3.27 |
| | Field MIR | 2-3-4-5-6/403 | 7 | 0.845 | 1.60 |
| | Lab visNIR | 2-3-4-5-6/465 | 7 | 0.448 | 3.02 |
| | Field visNIR | 1-2-3-4-5-6/1126 | 13 | 0.675 | 1.98 |
| | MLR | OC, $N_t$, clay, pH, OC:clay, OC:pH, $N_t$:pH | 4 | 0.399 | 3.33 |
| Resistant OC (g kg$^{-1}$) | Lab MIR | 2-3-4/687 | 2 | 0.314 | 1.17 |
| | Field MIR | 2-3-4-6/372 | 4 | 0.325 | 1.11 |
| | Lab visNIR | 1-2-3-4-5/298 | 2 | 0.291 | 1.25 |
| | Field visNIR | 1-2-3-4-5-6/1128 | 4 | 0.313 | 1.16 |
| | MLR | OC, $N_t$, clay, pH, OC:$N_t$, OC:clay, $N_t$:clay, OC:pH, $N_t$:pH, clay:pH, OC:$N_t$:pH, OC:clay:pH, $N_t$:clay:pH | 10 | 0.304 | 1.22 |
| Total N (g kg$^{-1}$) | Lab MIR | 1-2-3-4-5-6/1570 | 10 | 0.0256 | 7.44 |
| | Field MIR | 2-3-4-6/444 | 6 | 0.0883 | 2.17 |
| | Lab visNIR | 1-2-3-4-5-6/660 | 14 | 0.0333 | 5.76 |
| | Field visNIR | 1-2-3-4-5-6/1160 | 13 | 0.0723 | 2.71 |
| pH | Lab MIR | 2-3-4-5-6/1184 | 14 | 0.119 | 3.85 |
| | Field MIR | 1-2-3-4-5-6/661 | 11 | 0.284 | 1.62 |
| | Lab visNIR | 1-2-3-4-5-6/879 | 14 | 0.143 | 3.19 |
| | Field visNIR | 1-3-4/857 | 14 | 0.231 | 2.01 |
| Clay (%) | Lab MIR | 1-3-4-5-6/686 | 4 | 1.26 | 2.56 |
| | Field MIR | 1-2-3-4-5-6/462 | 6 | 1.39 | 2.37 |
| | Lab visNIR | 1-3-4-5-6/397 | 9 | 1.25 | 2.59 |
| | Field visNIR | 1-2-4-5-6/791 | 9 | 1.27 | 2.46 |
| Silt (%) | Lab MIR | 1-2-3-4-5-6/1036 | 5 | 1.40 | 2.21 |
| | Field MIR | 1-2-3-4-5-6/528 | 5 | 1.54 | 2.00 |
| | Lab visNIR | 1-2-3-4-5-6/525 | 8 | 1.33 | 2.37 |
| | Field visNIR | 1-2-3-4-6/932 | 10 | 1.45 | 2.07 |
| Sand (%) | Lab MIR | 1-2-3-4-5-6/1199 | 7 | 0.663 | 1.96 |
| | Field MIR | 1-2-3-6/247 | 6 | 0.776 | 1.67 |
| | Lab visNIR | 1-2-3-4-5-6/427 | 8 | 0.532 | 2.43 |
| | Field visNIR | 1-2-3-4-5/1065 | 13 | 0.612 | 2.12 |

*Note*: Average root mean squared error (RMSE) and ratio of prediction to interquartile range (RPIQ) of estimates are given for test sets created from five partitions of the dataset.

[a]Selected in at least one of the models calculated from five partitions of the dataset; visNIR: 500–834 nm (region 1), 834–1167 (2), 1167–1500 (3), 1500–1833 (4), 1833–2166 (5), 2166–2500 (6); MIR: 4000–3682.1 cm$^{-1}$ (1), 3682.1–3020.5 (2), 3020.5–2358.9 (3), 2358.9–1693.5 (4), 1693.5–1030 (5), 1030–650 (6).
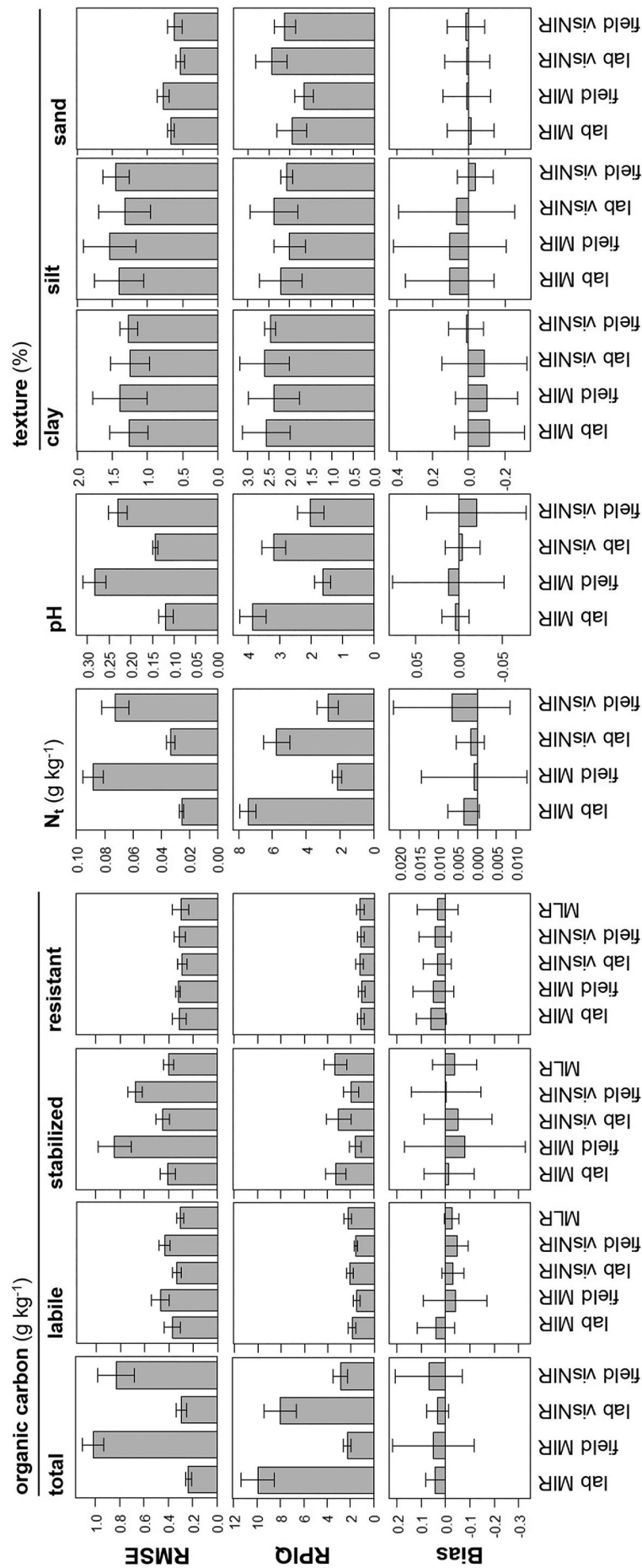
**FIGURE 2** Average root mean squared error (RMSE), ratio of prediction to interquartile range (RPIQ), and bias of lab and field visible/near- (visNIR) and mid-infrared (MIR) spectroscopy partial least squares regression estimates for the five test sets. Estimates from multiple linear regression (MLR) are additionally given for fraction organic carbon contents. Error bars show SDs
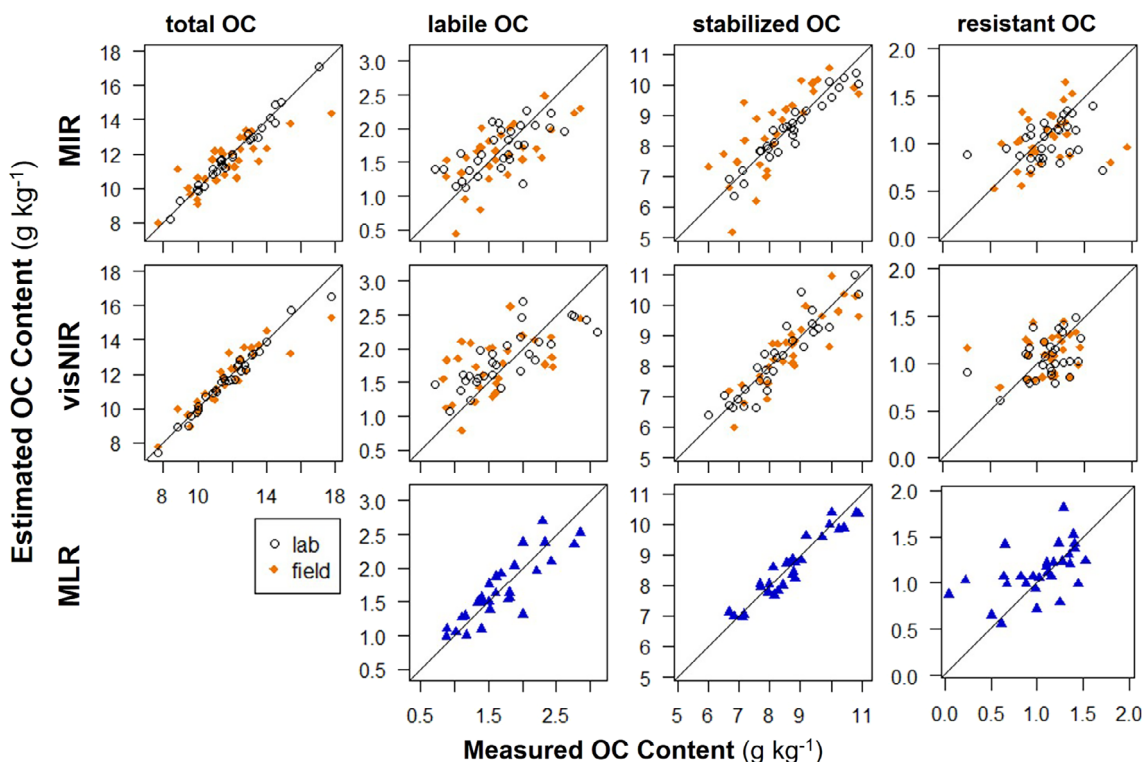
**FIGURE 3** Measured versus estimated soil total, labile, stabilized, and resistant organic carbon (OC) fractions. Shown are the lab and field visible/near- (visNIR) and mid-infrared (MIR) spectroscopy partial least squares regression and multiple linear regression (MLR) estimates for the test set with the median performance (based on ratio of performance to interquartile distance)
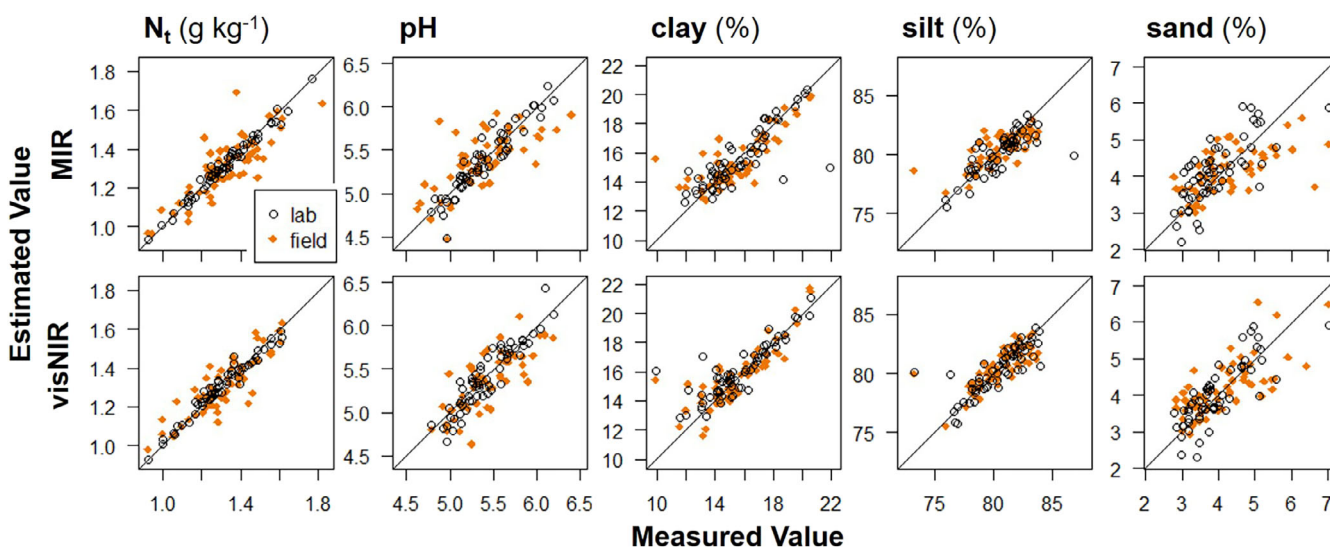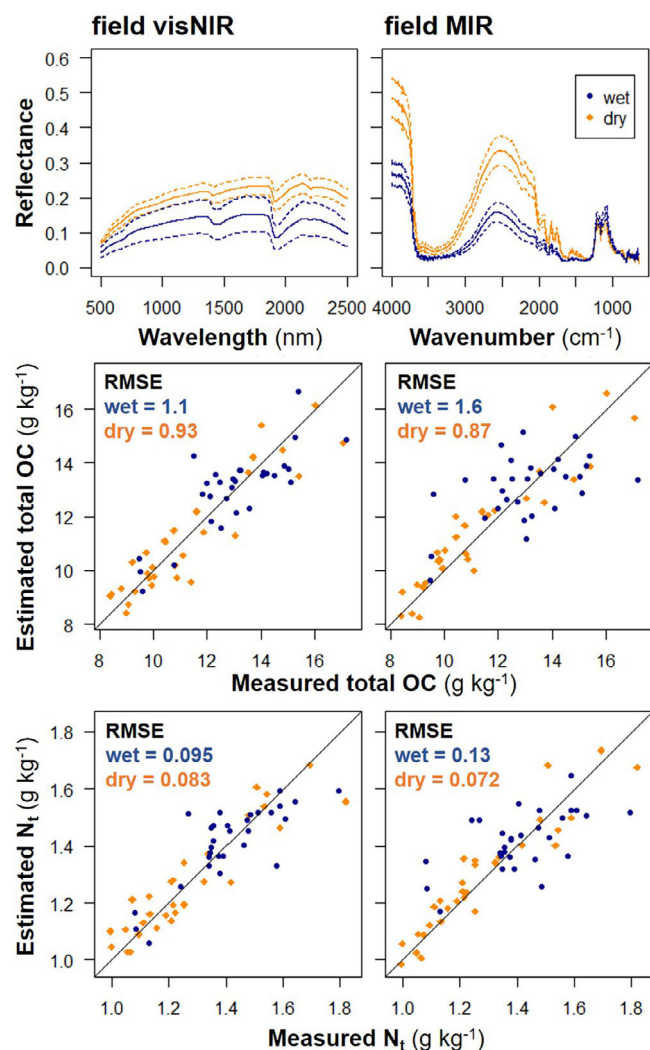


**FIGURE 4** Measured versus estimated total nitrogen content ($N_t$), pH and soil texture. Shown are the field and lab visible/near- (visNIR) and mid-infrared (MIR) spectroscopy partial least squares regression estimates for the test set with the median performance (based on ratio of performance to interquartile distance)

## 3.3 | Average model performance for general soil properties

For $N_t$ and pH, lab devices decidedly outperformed field devices (Table 2, Figures 2 and 4). Note that Figure 4 shows

measured versus estimated values for the test set resulting in the median model performance; thus, although training and test sets were identical for the five partitions across the four spectrometers, the plotted test sets may differ. Average RMSE for lab MIR and visNIR estimates of $N_t$ content were 0.026

**FIGURE 5** Average ± 1 SD field visible/near- (visNIR) and mid-infrared (MIR) reflectance spectra of the 30 wettest and 30 driest soils as well as cross-validation estimated versus measured total organic carbon (OC) and nitrogen ($N_t$) content. RMSE = root mean squared error

and 0.033 g kg$^{-1}$, respectively, and average RPIQ were 7.4 and 5.8, respectively, while field MIR and visNIR had average RMSE of 0.088 and 0.072 g kg$^{-1}$ and RPIQ of 2.2 and 2.7, respectively. Average RMSE for lab MIR and visNIR estimates of pH were 0.12 and 0.14, respectively, and average RPIQ were 3.9 and 3.2, respectively, while field MIR and visNIR had average RMSE of 0.28 and 0.23 and RPIQ of 1.6 and 2.0, respectively. Average bias of $N_t$ content, pH, and also total OC content was highest for the field visNIR device.

For soil texture, either lab or field visNIRS models slightly outperformed lab MIRS models on average, and field MIR estimates were comparable or worse. For clay, the performance of all models was comparable

(average RMSE = 1.3%–1.4%, RPIQ = 2.4–2.6). Silt estimates were slightly better for both lab devices (average RMSE = 1.3%–1.4%, RPIQ = 2.2–2.4) than field devices (average RMSE = 1.5%, RPIQ = 2.0–2.1). Finally, for sand, both visNIR devices were slightly better (average RMSE = 0.53% and 0.61% and RPIQ = 2.4 and 2.1 for lab and field, respectively) than both MIR devices (average RMSE = 0.66% and 0.78% and RPIQ = 2.0 and 1.7 for lab and field, respectively). The narrow range in sand content (2.5%–7.0%) is responsible for the lower RPIQ of sand compared to clay and silt estimates, since RMSEs for sand were consistently lower than that of clay and silt. Field visNIR had the lowest average bias for clay and silt contents compared to other spectrometers.

## 3.4 | Effects of soil moisture and texture on in situ spectral models

The relationship of GWC, clay content and sand content to the residuals and absolute value of the residuals of OC and $N_t$ content estimates were studied using Spearman rank correlations due to non-normality of all properties. No significant correlations were found between clay or sand content and the absolute value of residuals of OC prediction by field visNIRS or MIRS. However, weak positive Spearman rank correlations were found between GWC and the absolute value of residuals of OC estimation by field MIRS ($\rho = 0.21$, $p < 0.01$) and by field visNIRS ($\rho = 0.12$, $p = 0.04$). A separate analysis of the 30 wettest soils (GWC >18.8%) revealed the spectra had lower reflectance and a loss of detail, particularly for MIR, compared to the spectra of the 30 driest soils (GWC <9.0%) (Figure 5). The accuracy of cross-validation estimates for both field devices was higher for dry soils compared to wet soils for OC and $N_t$ contents; however, the improvement in accuracy was much greater for MIR (RMSE of the wettest versus driest soils, respectively, was 1.6 and 0.87 g kg$^{-1}$ for OC content and 0.13 versus 0.072 g kg$^{-1}$ for $N_t$ content) than visNIR (RMSE of the wettest versus driest soils, respectively, was 1.1 and 0.93 g kg$^{-1}$ for OC content and 0.095 and 0.083 g kg$^{-1}$ for $N_t$ content) (Figure 5). Thus, field MIR had more accurate OC and $N_t$ content predictions than field visNIR under drier conditions (GWC <9.0%), and vice versa under wetter conditions (GWC >18.8%). However, attempts to improve field MIR models by excluding regions with O—H vibrations from the spectra (3596–3200 cm$^{-1}$ and/or 1650–1560 cm$^{-1}$; Clark, 1999; Tinti et al., 2015) for the complete dataset or for a subset of the 30 wettest soils found model performance was similar or worse without these regions (data not shown).

# 4 | DISCUSSION

## 4.1 | Average model performance and predictive mechanisms for soil OC content

The performance of models predicting fraction OC contents in our study at the field scale (Table 2, Figures 2 and 3) can be compared to studies using a sample collected from diverse sites. The lab visNIRS and MIRS RPIQs found in this study for labile and stabilized OC content were low compared to other studies, but RMSEs were also lower or comparable (Knox et al., 2015; Madhavan et al., 2017; Zimmermann et al., 2007). This suggests that a relatively homogeneous sample produces a well-calibrated model with a low error rate, but also a small interquartile range of soil properties, resulting in a likewise low RPIQ compared to a sample including soils collected from many sites and management types. Thus the variability of the sample has a large effect on performance measures and their interpretation depends on the context in which the model is applied.

Ludwig et al. (2016) and Linsler et al. (2017) likewise found the accuracy that could be achieved for estimates of labile and stabilized fractions made indirectly with MLR or PLSR using related soil properties (i.e., total OC, $N_t$, pH and texture) were approximately equivalent or better to that which could be achieved with lab visNIR and MIR (Table 2, Figures 2 and 3). However, in contrast to our findings, they found that lab MIRS was able to achieve satisfactory estimates of resistant OC, defined as OC remaining after $Na_2S_2O_8$ oxidation, which were also superior to MLR predictions using related soil properties. This could be due to the greater efficiency of $Na_2S_2O_8$ compared to NaOCl for the isolation of stable OC (Helfrich et al., 2007), and thus separation of a more chemically differentiated fraction.

These contrasting model accuracies depending on the fractionation method applied raises the issue of the underlying spectral prediction mechanisms. Spectroscopy relies on the absorbance of radiation at distinct wavelengths due to vibrations of molecular bonds (Tinti et al., 2015). However, soil OC turnover time is determined not only by molecular structure, but also the physical disconnection between decomposers and organic matter, for example, resulting from aggregate formation (Schmidt et al., 2011). Furthermore, Poeplau et al. (2018) found that the turnover time of light OC is highly related to particle size. The fractionation procedure implemented here considers this complexity by utilizing both physical and chemical methods of separation (Zimmermann et al., 2007). While soil physical properties have been estimated by visNIRS and MIRS with some success, this is due to correlations with other spectrally active soil components, such as soil minerals, OC and carbonates (Soriano-Disla et al., 2014).

The VIP analysis found absorbance from 500 to 700 nm was important for field visNIR prediction of total and stabilized OC, which could be attributed to interactions between visible light and various chromophores, such as iron oxides and organic matter (Ben-Dor, 2002; Stenberg et al., 2010). The greater importance of the visible region for field than lab visNIRS could indicate that the presence of soil moisture was helpful in distinguishing OC visually. For the NIR region, attribution of important wavelengths to specific compounds is challenging due to the presence of overlapping overtones and/or combinations of fundamental vibrations occurring in the MIR range (Soriano-Disla et al., 2014). Wavelengths around 1000 nm, possibly related to amines (Stenberg et al., 2010), aided field visNIR prediction of total and resistant OC. The region around 1170 nm, possibly related to alkyl groups (Stenberg et al., 2010), helped with prediction of labile OC by lab visNIR. Absorbance around 1400 nm, potentially related to water, phenols and kaolinite (Soriano-Disla et al., 2014; Stenberg et al., 2010), was very important for prediction of all OC fractions by lab visNIR and somewhat important for total OC prediction by field visNIR. This may indicate that soil moisture limited the usefulness of this peak for field models. Absorbance at 1830 nm, possibly related to methyl groups (Stenberg et al., 2010), aided field visNIR prediction of total, labile and resistant OC. The peak around 1900 nm, possibly related to carboxylic acids and water, including water in soil pores, absorbed to surfaces, or held within the lattice of clay minerals (Stenberg et al., 2010), helped with prediction of total, labile and stabilized OC for lab visNIR and total, labile and resistant OC for field visNIR. Absorption around 2200 nm, related to alumino-silicates and organic matter (Soriano-Disla et al., 2014), aided prediction of total, stabilized and resistant OC for lab visNIR and resistant OC for field visNIR. Finally, the peak around 2470 nm, perhaps related to methyl groups (Stenberg et al., 2010), was important for prediction of total OC for lab visNIR and all OC fractions for field visNIR.

For MIR, VIP analysis found that lab and field models shared many important wavelengths, but bands from 3700 to 2375 $cm^{-1}$ were of greater importance for lab MIR than field MIR OC predictions. 3700–3600 $cm^{-1}$, related to kaolinite, smectite and illite (Soriano-Disla et al., 2014), was important for predicting total, labile and stabilized OC for lab MIR and primarily stabilized OC for field MIR. 3070–2840 $cm^{-1}$, affected by aliphatic CH (Janik et al., 2007), helped with prediction of all fraction and total OC by lab MIR. The peak around 2050 $cm^{-1}$, related to quartz (Tinti et al., 2015), aided prediction of

total and labile OC by field MIR. 1920–1700 $cm^{-1}$, also containing absorption peaks for quartz and carboxylic acid (Soriano-Disla et al., 2014; Tinti et al., 2015), was especially important for prediction of labile OC for both spectrometers, but also useful for total and all fraction OC prediction for field MIR and total and stabilized OC for lab MIR. 1700–1220 $cm^{-1}$, which contains absorption due to amides, aromatic groups, carboxylate anions, $C—NO_2$, $SO_2O—$, and $P—O—alkyl$ (Soriano-Disla et al., 2014; Janik et al., 2007; Bruker Optics Inc., 2009), helped with prediction of total and resistant OC for lab MIR and total and stabilized OC for field MIR. This region also contains water absorption from 1642 to 1569 $cm^{-1}$ (Tinti et al., 2015), but was nevertheless useful for field MIRS. The usefulness of 1220–1060 $cm^{-1}$ for total and resistant OC prediction for lab MIR, and 1060 $cm^{-1}$ for total and all fraction OC prediction for field MIR can be attributed to signatures of quartz, alumino-silicates, carbohydrate-COH stretching and P—O—Aryl groups (Bruker Optics Inc., 2009; Janik et al., 2007; Soriano-Disla et al., 2014; Tinti et al., 2015). There was generally large overlap in the important peaks for prediction of total and all fraction OC contents by field MIR, which indicates an inability to distinguish the fractions, resulting in poor model performance. For lab MIR, the important wavelengths for each OC fraction were more unique, indicating better differentiation and resulting in superior performance of prediction models. Finally, VIP analysis for both visNIRS and MIRS confirmed that prediction of fraction OC contents with differentiated turnover times was due to a combination of direct and indirect spectral estimation mechanisms: both absorption peaks of specific OC compounds and soil minerals (alumino-silicate clays for both spectral regions and quartz for MIR) were highly important for lab and field predictions of total and fraction OC contents (Figure 1b).

## 4.2 | Average model performance for general soil properties

It was likewise found by Viscarra Rossel et al. (2006) that lab MIR outperformed lab NIR for field-scale estimation of total OC and pH (Table 2). They additionally found lab MIRS was superior to NIR for estimation of clay, silt and sand, whereas the present study found lab or field visNIR was comparable to or slightly outperformed lab MIR for texture estimations. However, Viscarra Rossel et al. (2006) also found that only texture estimations were improved by a combined PLSR analysis of visible, near- and mid-infrared spectral ranges compared to use of MIR alone, supporting our findings that visNIR is comparatively more useful for texture than OC and pH estimation.

Comparing lab versus field spectral performance, OC, $N_t$ and pH predictions by lab spectrometers were far superior, and the best and worse estimations were made by lab and field MIR, respectively. Hutengs et al. (2019) found MIR OC estimations were always more accurate than visNIR estimations when measured under the same conditions (i.e., on sieved, dried and ground soil or in situ with median soil GWC of 7%). When a subset of the driest soils was isolated in the current study (<9% GWC), we likewise found field MIR outperformed field visNIR for OC and $N_t$ estimation (Figure 5). Though the effect was much more dramatic for field MIR, OC and $N_t$ estimation accuracy of both visNIR and MIR was worse for a subset of the wettest soils. The influence of soil moisture on visNIRS was likewise explored by Marakkala Manage et al. (2018), who found that soil texture could be estimated well by visNIR at various moisture levels, whereas OC estimation accuracy decreased with increasing soil moisture. This matches our findings that the loss of accuracy between in situ measurements on field moist soil and lab visNIR estimations on dried soil was much greater for OC content than soil texture.

In addition to soil moisture, the accuracy of models created from spectra measured in the lab versus in situ could be attributed to differences between the benchtop versus handheld spectrometers as well as the effects of soil structure and heterogeneity. The former effect was explored by Hutengs et al. (2018), who found performance of the Agilent 4300 Handheld FTIR field device was as good or better than the Bruker-TENSOR 27 lab device for a range of soil properties when measurements were made on sieved, dried and ground soil. The effects of sample homogeneity, structure and moisture were explored by Hutengs et al. (2019) by comparing OC estimation accuracy for in situ measurement versus measurement on sieved and dried soil, versus measurement on sieved, dried and ground soil using the same field spectrometers used in the present study. Thus, the effects of standard laboratory soil preparation methods were isolated. For visNIR, sieving and drying soils with median water content of 7% decreased total OC RMSE dramatically, while additionally grinding the soils had much less effect. In contrast, both sieving/drying and grinding soils substantially decreased RMSE for MIR. The greater benefit of soil homogenization for MIR could be because the measurement window of the Agilent 4300 sampling interface is two orders of magnitude smaller than that of the visNIR contact probe (ASD FieldSpec). Thus, the poorer performance of the MIR field device compared to the visNIR field device in the present study can be attributed to soil moisture and heterogeneity,

exacerbated by the small measurement window of the Agilent device. The achieved results must also be considered in the context that this is a field-scale study of a conventionally-tilled site, thus the ratio of variance within sub-replicates (replicate measurements within a $15 \times 15$ cm sampling point) to variance between sampling points was relatively high compared to larger-scale studies including a diverse range of soils.

## 5 | CONCLUSIONS

Models created from lab spectra outperformed models from in situ spectra for total OC, $N_t$, and pH, whereas the accuracy of both visNIR devices was comparable or slightly better than both MIR devices for sand, silt and clay. Lab spectral estimations for labile and stabilized OC were slightly inferior to estimations from multiple linear regressions using total OC, $N_t$, clay and pH as predictors, while field spectral estimations were substantially worse. VIP analysis found both spectral signatures of specific OC compounds and soil minerals were key predictors for fraction OC contents, and the importance of various peaks and regions differed between lab and field measurement.

The loss of model accuracy from lab to field measurement was lower for visNIRS than MIRS; however, these results must be considered in the context of the soil moisture at the time of sampling. Analysis of a subset of the driest and wettest soils demonstrated that the comparative benefit of field visNIR versus field MIR for OC and $N_t$ prediction was highly moisture dependent. The ratio of variance within to between soils in a dataset is also expected to impact the relative performance of field visNIR versus MIR due to the higher moisture sensitivity of field MIR and the smaller measurement window of the Agilent device. Thus, the suitability of a particular spectral range (visNIR or MIR) for field measurements might mainly depend on the soil moisture content and variability of the study site. Careful preparation of the soil surface prior to taking in situ spectral measurement is strongly recommended (e.g., removing crop residues, compressing aggregates at the surface, and allowing the soil to sun dry). A systematic rewetting experiment would be useful to determine performance thresholds for the two spectral ranges. Finally, future research could further investigate whether combining both spectral ranges for field measurements would improve model robustness.

## CONFLICT OF INTEREST
The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

## AUTHOR CONTRIBUTIONS
**Isabel Greenberg:** Conceptualization (supporting); data curation (lead); formal analysis (lead); investigation (lead); methodology (equal); software (supporting); validation (lead); visualization (lead); writing – original draft (lead). **Michael Seidel:** Conceptualization (supporting); formal analysis (supporting); methodology (supporting); validation (supporting); writing – review and editing (equal). **Michael Vohland:** Conceptualization (supporting); funding acquisition (equal); resources (supporting); supervision (supporting); writing – review and editing (equal). **Bernard Ludwig:** Conceptualization (lead); data curation (supporting); funding acquisition (equal); methodology (lead); project administration (lead); resources (lead); software (lead); supervision (lead); validation (supporting); writing – review and editing (equal).

## ORCID
*Isabel Greenberg* https://orcid.org/0000-0002-4762-8474
*Michael Vohland* https://orcid.org/0000-0002-6048-1163
*Bernard Ludwig* https://orcid.org/0000-0001-8900-6190

## REFERENCES
Australian Government. (2018). *Carbon credits (carbon farming initiative—measurement of soil carbon sequestration in agricultural systems) methodology determination* 2018.

Baldock, J. A., Hawke, B., Sanderman, J., & Macdonald, L. M. (2013). Predicting contents of carbon and its component fractions in Australian soils from diffuse reflectance mid-infrared spectra. *Soil Research*, *51*, 577. https://doi.org/10.1071/SR13077

Beleites, C., Baumgartner, R., Bowman, C., Somorjai, R., Steiner, G., Salzer, R., & Sowa, M. G. (2005). Variance reduction in estimating classification error using sparse datasets. *Chemometrics and Intelligent Laboratory Systems*, *79*, 91–100. https://doi.org/10.1016/j.chemolab.2005.04.008

Ben-Dor, E. (2002). Quantitative remote sensing of soil properties. *Advances in Agronomy*, *75*, 173–243. https://doi.org/10.1016/S0065-2113(02)75005-0

Bruker Optics Inc. (2009). *Guide for infrared spectroscopy.*

Chong, I.-G., & Jun, C.-H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, *78*, 103–112. https://doi.org/10.1016/j.chemolab.2004.12.011

Clark, R. N. (1999). Spectroscopy of rocks and minerals, and principles of spectroscopy. In A. N. Rencz (Ed.), *Manual of remote sensing: Remote sensing for the earth sciences* (pp. 3–58). John Wiley and Sons.

DIN ISO 11277. (2002). *Bodenbeschaffenheit—Bestimmung der Partikelgrößenverteilung in Mineralböden: Verfahren mittels Siebung und Sedimentation. ISO 11277: 1998/ Cor.1:2002.* Beuth Verlag.

Edenhofer, O. (Ed.). (2014). *Climate change 2014: Mitigation of climate change working group III contribution to the fifth assessment report of the intergovernmental panel on climate change.* Cambridge University Press.

England, J. R., & Viscarra Rossel, R. A. (2018). Proximal sensing for soil carbon accounting. *The Soil*, *4*, 101–122. https://doi.org/10.5194/soil-4-101-2018

Greenberg, I., Linsler, D., Vohland, M., & Ludwig, B. (2020). Robustness of visible/near and mid-infrared spectroscopic models to changes in the quantity and quality of crop residues in soil. *Soil Science Society of America Journal.*, *84*, 963–977. https://doi.org/10.1002/saj2.20067

Helfrich, M., Flessa, H., Mikutta, R., Dreves, A., & Ludwig, B. (2007). Comparison of chemical fractionation methods for isolating stable soil organic carbon pools. *European Journal of Soil Science*, *58*, 1316–1329. https://doi.org/10.1111/j.1365-2389.2007.00926.x

Hutengs, C., Ludwig, B., Jung, A., Eisele, A., & Vohland, M. (2018). Comparison of portable and bench-top spectrometers for mid-infrared diffuse reflectance measurements of soils. *Sensors*, *18*, 993. https://doi.org/10.3390/s18040993

Hutengs, C., Seidel, M., Oertel, F., Ludwig, B., & Vohland, M. (2019). In situ and laboratory soil spectroscopy with portable visible-to-near-infrared and mid-infrared instruments for the assessment of organic carbon in soils. *Geoderma*, *355*, 113900. https://doi.org/10.1016/j.geoderma.2019.113900

IUSS Working Group WRB. (2015). *World reference base for soil resources 2014, update 2015: International soil classification system for naming soils and creating legends for soil maps.* FAO.

Janik, L. J., Skjemstad, J. O., Shepherd, K. D., & Spouncer, L. R. (2007). The prediction of soil carbon fractions using mid-infrared-partial least square analysis. *Soil Research*, *45*, 73. https://doi.org/10.1071/SR06083

Janik, L. J., Soriano-Disla, J. M., Forrester, S. T., & McLaughlin, M. J. (2016). Moisture effects on diffuse reflection infrared spectra of contrasting minerals and soils: A mechanistic interpretation. *Vibrational Spectroscopy*, *86*, 244–252. https://doi.org/10.1016/j.vibspec.2016.07.005

Ji, W., Adamchuk, V. I., Biswas, A., Dhawale, N. M., Sudarsan, B., Zhang, Y., Viscarra Rossel, R. A., & Shi, Z. (2016). Assessment of soil properties in situ using a prototype portable MIR spectrometer in two agricultural fields. *Biosystems Engineering*, *152*, 14–27. https://doi.org/10.1016/j.biosystemseng.2016.06.005

Knox, N. M., Grunwald, S., McDowell, M. L., Bruland, G. L., Myers, D. B., & Harris, W. G. (2015). Modelling soil carbon fractions with visible near-infrared (VNIR) and mid-infrared (MIR) spectroscopy. *Geoderma*, *239-240*, 229–239. https://doi.org/10.1016/j.geoderma.2014.10.019

Koch, H.-J., Dieckmann, J., Büchse, A., & Märländer, B. (2009). Yield decrease in sugar beet caused by reduced tillage and direct drilling. *European Journal of Agronomy*, *30*, 101–109. https://doi.org/10.1016/j.eja.2008.08.001

Lal, R. (2013). Soil carbon management and climate change. *Carbon Management*, *4*, 439–462. https://doi.org/10.4155/CMT.13.31

Liland, K.H., Mehmood, T., & Sæbø, S. (2020). *plsVarSel: Variable selection in partial least squares (R package version 0.9.6).* https://CRAN.R-project.org/package=plsVarSel.

Linsler, D., Sawallisch, A., Höper, H., Schmidt, H., Vohland, M., & Ludwig, B. (2017). Near-infrared spectroscopy for determination of soil organic C, microbial biomass C and C and N fractions in a heterogeneous sample of German arable surface soils. *Archives of Agronomy and Soil Science*, *63*, 1499–1509. https://doi.org/10.1080/03650340.2017.1292030

Ludwig, B., Linsler, D., Höper, H., Schmidt, H., Piepho, H.-P., & Vohland, M. (2016). Pitfalls in the use of middle-infrared spectroscopy: Representativeness and ranking criteria for the estimation of soil properties. *Geoderma*, *268*, 165–175. https://doi.org/10.1016/j.geoderma.2016.01.010

Ludwig, B., Murugan, R., Parama, V. R. R., & Vohland, M. (2019). Accuracy of estimating soil properties with mid-infrared spectroscopy: Implications of different chemometric approaches and software packages related to calibration sample size. *Soil Science Society of America Journal*, *83*, 1542–1552. https://doi.org/10.2136/sssaj2018.11.0413

Madhavan, D. B., Baldock, J. A., Read, Z. J., Murphy, S. C., Cunningham, S. C., Perring, M. P., Herrmann, T., Lewis, T., Cavagnaro, T. R., England, J. R., Paul, K. I., Weston, C. J., & Baker, T. G. (2017). Rapid prediction of particulate, humus and resistant fractions of soil organic carbon in reforested lands using infrared spectroscopy. *Journal of Environmental Management*, *193*, 290–299. https://doi.org/10.1016/j.jenvman.2017.02.013

Marakkala Manage, L. P., Greve, M. H., Knadel, M., Moldrup, P., de Jonge, L. W., & Katuwal, S. (2018). Visible-near-infrared spectroscopy prediction of soil characteristics as affected by soil-water content. *Soil Science Society of America Journal*, *82*, 1333–1346. https://doi.org/10.2136/sssaj2018.01.0052

Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, *118*, 62–69. https://doi.org/10.1016/j.chemolab.2012.07.010

Mevik, B.-H., Wehrens, R., & Liland, K. H. (2019). *pls: Partial least squares and principal component regression* (R package version 2.7–1). https://CRAN.R-project.org/package=pls

Minasny, B., & McBratney, A. (2013). Why you don't need to use RPD. *Pedometron*, *33*, 14–15.

Necpálová, M., Anex, R. P., Kravchenko, A. N., Abendroth, L. J., Del Grosso, S. J., Dick, W. A., Helmers, M. J., Herzmann, D., Lauer, J. G., Nafziger, E. D., & Sawyer, J. E. (2014). What does it take to detect a change in soil carbon stock? A regional comparison of minimum detectable difference and experiment duration in the north Central United States. *Journal of Soil and Water Conservation*, *69*, 517–531. https://doi.org/10.2489/jswc.69.6.517

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Dor, E. B., Brown, D. J., Clairotte, M., Csorba, A., & Dardenne, P. (2015). Soil spectroscopy: An alternative to wet chemistry for soil monitoring.

*Advances in Agronomy*, *132*, 139–159. https://doi.org/10.1016/bs.agron.2015.02.002

Poeplau, C., Don, A., Six, J., Kaiser, M., Benbi, D., Chenu, C., Cotrufo, M. F., Derrien, D., Gioacchini, P., Grand, S., Gregorich, E., Griepentrog, M., Gunina, A., Haddix, M., Kuzyakov, Y., Kühnel, A., Macdonald, L. M., Soong, J., Trigalet, S., ... Nieder, R. (2018). Isolating organic carbon fractions with varying turnover rates in temperate agricultural soils – A comprehensive method comparison. *Soil Biology and Biochemistry*, *125*, 10–26. https://doi.org/10.1016/j.soilbio.2018.06.025

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing https://www.r-project.org/

Reeves, J. B., McCarty, G. W., & Hively, W. D. (2010). Mid- versus near-infrared spectroscopy for on-site analysis of soil. In R. A. V. Rossel, A. B. McBratney, & B. Minasny (Eds.), *Proximal soil sensing* (pp. 133–142). Springer Netherlands.

Ripley, B.D. (2018). *step {stats}: Choose a model by AIC in a Stepwise Algorithm (R package version 3.4.4)*. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/step.html.

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S., & Trumbore, S. E. (2011). Persistence of soil organic matter as an ecosystem property. *Nature*, *478*, 49–56. https://doi.org/10.1038/nature10386

Silvero, N. E. Q., Di Raimo, L. A. L., Pereira, G. S., de Magalhães, L. P., Terra, F. D. S., Dassan, M. A. A., Salazar, D. F., & Demattê, J. A. (2020). Effects of water, organic matter, and iron forms in mid-IR spectra of soils: Assessments from laboratory to satellite-simulated data. *Geoderma*, *375*, 114480. https://doi.org/10.1016/j.geoderma.2020.114480

Soriano-Disla, J. M., Janik, L. J., Viscarra Rossel, R. A., Macdonald, L. M., & McLaughlin, M. J. (2014). The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. *Applied Spectroscopy Reviews*, *49*, 139–186. https://doi.org/10.1080/05704928.2013.811081

Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M., & Wetterlind, J. (2010). Visible and near infrared spectroscopy in soil science. *Advances in Agronomy*, *107*, 163–215. https://doi.org/10.1016/S0065-2113(10)07005-7

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. https://doi.org/10.1371/journal.pone.0066409

Stevens, A., & Ramirez-Lopez, L. (2013). *An introduction to the prospectr package (R package version 0.1.3)*. https://cran.rproject.org/web/packages/prospectr/index.html.

Stevens, A., van Wesemael, B., Bartholomeus, H., Rosillon, D., Tychon, B., & Ben-Dor, E. (2008). Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. *Geoderma*, *144*, 395–404. https://doi.org/10.1016/j.geoderma.2007.12.009

Tinti, A., Tugnoli, V., Bonora, S., & Francioso, O. (2015). Recent applications of vibrational mid-infrared (IR) spectroscopy for studying soil components: A review. *Journal of Central European Agriculture*, *16*, 1–22. https://doi.org/10.5513/JCEA01/16.1.1535

United Nations Convention to Combat Desertification (UNCCD). (2013). *Decision 22/COP.11, report* (WWW document). https://www.unccd.int/sites/default/files/sessions/documents/2019-08/22COP11_0.pdf.

Viscarra Rossel, R. A., & Behrens, T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, *158*, 46–54. https://doi.org/10.1016/j.geoderma.2009.12.025

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. https://doi.org/10.1016/j.geoderma.2005.03.007

von Lützow, M., Kögel-Knabner, I., Ekschmitt, K., Flessa, H., Guggenberger, G., Matzner, E., & Marschner, B. (2007). SOM fractionation methods: Relevance to functional pools and to stabilization mechanisms. *Soil Biology and Biochemistry*, *39*, 2183–2207. https://doi.org/10.1016/j.soilbio.2007.03.007

Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., van Wesemael, B., Rabot, E., Ließ, M., Garcia-Franco, N., Wollschläger, U., Vogel, H. J., & Kögel-Knabner, I. (2019). Soil organic carbon storage as a key function of soils - a review of drivers and indicators at various scales. *Geoderma*, *333*, 149–162. https://doi.org/10.1016/j.geoderma.2018.07.026

Zimmermann, M., Leifeld, J., & Fuhrer, J. (2007). Quantifying soil organic carbon fractions by infrared-spectroscopy. *Soil Biology and Biochemistry*, *39*, 224–231. https://doi.org/10.1016/j.soilbio.2006.07.010