

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/374120210>

# Diffuse reflectance mid-infrared spectroscopy is viable without fine milling

Article in *Soil Security* · September 2023

DOI: 10.1016/j.soisec.2023.100104

CITATIONS

0

READS

16

9 authors, including:



**Jonathan Sanderman**

Woodwell Climate Research Center

148 PUBLICATIONS 9,914 CITATIONS

[SEE PROFILE](#)



**José Lucas Safanelli**

Woodwell Climate Research Center

53 PUBLICATIONS 1,027 CITATIONS

[SEE PROFILE](#)

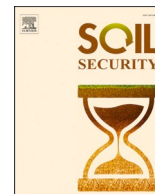
Some of the authors of this publication are also working on these related projects:



Geotechnologies for mapping agricultural soils of São Paulo State [View project](#)



Qualidade da água subterrânea na bacia do rio Marombas, Mesorregião Serrana do Estado de Santa Catarina [View project](#)



# Diffuse reflectance mid-infrared spectroscopy is viable without fine milling

Jonathan Sanderman<sup>a,\*</sup>, Colleen Smith<sup>a</sup>, José Lucas Safanelli<sup>a</sup>, Cristine L.S. Morgan<sup>b</sup>, Jason Ackerman<sup>b</sup>, Nathaniel Looker<sup>b</sup>, Cara Mathers<sup>c</sup>, Rebecca Keating<sup>c</sup>, Ashok A. Kumar<sup>c</sup>

<sup>a</sup> Woodwell Climate Research Center, Falmouth MA 02540 USA

<sup>b</sup> Soil Health Institute, Morrisville, NC 27560 USA

<sup>c</sup> Indigo Ag, Boston MA 02129 USA

## ARTICLE INFO

### Keywords:

Soil spectroscopy  
Laboratory analysis  
Chemometrics  
Soil carbon  
Soil texture

## ABSTRACT

While diffuse reflectance Fourier transform mid-infrared spectroscopy (mid-DRIFTS) has been established as a viable low-cost surrogate for traditional soil analyses, the assumed need for fine milling of soil samples prior to analysis is constraining the commercial appeal of this technology. Here, we reevaluate this assumption using a set of 2380 soil samples collected across North American agricultural soils. Cross-validation indicated that the best preprocessing (standard normal variate) and model form (memory-based learning) resulted in very good and nearly identical predictions for the <2 mm preparation and fine-milled preparation of these soils for total organic carbon (TOC), clay, sand, pH and bulk density (BD). Application of larger models built from the USDA NRCS mid-DRIFTS library also resulted in minimal performance differences between the two sample preps. Lower predictive performance of the existing library was attributed to less-than-perfect spectral representativeness of the library. Regardless of model form, there was very little variability between replicates of the <2 mm prep, suggesting that the lack of fine milling did not lead to more heterogeneous subsamples. Additionally, there was no relationship between residual error and soil texture, implying these results should be robust across most soil types. Overall, in agreement with other recent findings, these results suggest that routine scanning of standard <2 mm preparation does not degrade predictive performance of mid-DRIFTS-based inference systems. With good standard operating procedures including quality control and traditional analysis on a small percent of samples, mid-DRIFTS can become a routine tool in commercial soil laboratories.

## 1. Introduction

Traditional soil analytical methods are labor-intensive and as result often costly. As the need for soil information rapidly expands, lower-cost higher-throughput methods are being sought. Diffuse reflectance spectroscopy, in particular diffuse reflectance Fourier transform mid-infrared spectroscopy (mid-DRIFTS) is now a widely accepted complement to traditional analytical methods (Nocita et al., 2015; Shepherd et al., 2022). When mid-DRIFTS is coupled with a high quality spectral library, such as the USDA NRCS National Soil Survey Center - Kellogg Soil Survey Laboratory (NSSC-KSSL) mid-DRIFTS spectral library, accurate and unbiased estimates of at least 20 soil properties can be made on hundreds of samples in a single day (Dangal et al., 2019; Sanderman et al., 2020; Ng et al., 2022), compared to at least a week of laboratory work for the traditional analytical methods.

Results from mid-DRIFTS are highly encouraging even for

applications such as detecting subtle management-induced changes in soil carbon (Sanderman et al., 2021); however, it is generally thought that samples need to be dried, sieved, and finely milled prior to spectral acquisition. The need for fine milling greatly reduces the commercial viability of mid-DRIFTS. The existing literature is decidedly mixed with respect to the need for fine milling with mid-DRIFTS. Working with a diverse set of 860 Irish soil samples, Bachion de Santana and Daly (2022) found no difference in predictive performance between ball-milled and <2 mm samples for a range of soil properties. However, Wijewardane et al. (2021), working with a set of 500 US soils, found that fine milling significantly improved the predictive performance for all eight of the different soil properties tested. In both of these studies, a Bruker FTIR with a HTS-XT accessory was utilized. In a study of approximately 400 samples, Deiss et al. (2020) also found that fine milling improved predictive performance for many but not all soil properties tested, but the improvement in predictive performance was marginal enough for the

\* Corresponding author.

E-mail address: [jsanderman@woodwellclimate.org](mailto:jsanderman@woodwellclimate.org) (J. Sanderman).

<https://doi.org/10.1016/j.soisec.2023.100104>

Received 17 June 2023; Received in revised form 24 August 2023; Accepted 11 September 2023

Available online 22 September 2023

2667-0062/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

authors to conclude that prediction accuracy on <2 mm samples would be adequate for many applications. This last point is important to consider: the objectives of a particular use case dictate the necessary accuracy and precision of analytical data or mid-DRIFTS-based predictions, with higher uncertainties acceptable for applications such as soil resource mapping but not for tracking temporal changes (e.g., organic carbon accumulation) in response to management.

The overall goal of this study was to better understand the limitations of mid-DRIFTS spectral inference without fine milling. Previous studies all found that good models could be built directly from <2 mm sieved samples, but results were conflicting as to whether better accuracy and precision could be achieved with fine milling. One limitation of these prior investigations was the sample size, which ranged from about 400 to 850 samples. Here, we analyzed 2380 samples from across North American agricultural soils to address the following questions: 1) Can unbiased predictions be achieved from scanning samples that have been through standard sample preparation of sieving to 2 mm?; 2) Does fine milling improve the predictions of organic carbon, pH, clay, sand and bulk density?; 3) Is there more variability between replicates without fine milling?; 4) Are there specific soil parameters that control the performance of predictive models on <2 mm soils?; and, finally, 5) Can we use existing predictive models (i.e. USDA NRCS NSSC-KSSL database) built on finely milled samples or should we build models directly

from the spectra collected on the non-finely milled samples?

## 2. Methods and materials

### 2.1. Soil samples

We used soil samples from two separate sampling efforts: the North American Project to Evaluate Soil Health Measurements (NAPESHM; Norris et al., 2020) and a subset of the 2021 samples from Indigo Ag's soil carbon experiment (SCX) fields. The NAPESHM data set consists of 1953 topsoil samples from 124 long-term agricultural research sites in the United States, Canada, and Mexico. Composite samples (0–7.6 cm depth for bulk density, 0–15 cm depth for all other properties and mid-DRIFTS spectra) were collected from multiple treatment replicates per site (Bagnall et al., 2022). The SCX dataset consists of 427 samples that fall within the US geographic range of the NAPESHM data (Fig. 1). Between one and four 1 m soil cores with an internal diameter of 3.5 cm were extracted from 73 SCX fields located in 11 states. Soil cores were divided into 0–15, 15–30, 30–60 and 60–100 cm depth increments.

### 2.2. Laboratory data

We have focused on five soil properties - total organic carbon (TOC,

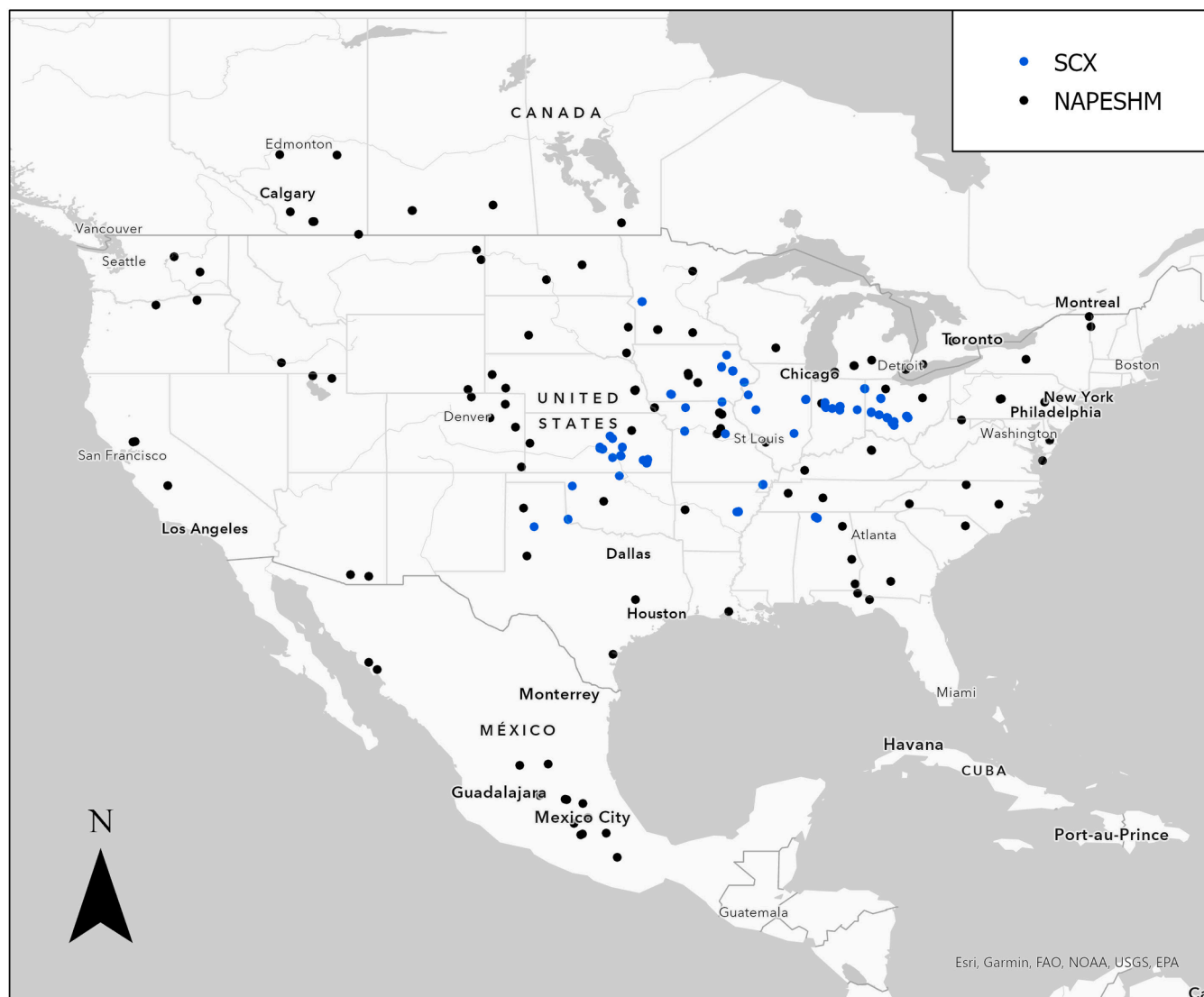


Fig. 1. SCX and NAPESHM site locations.

wt %), pH, bulk density (BD,  $\text{g cm}^{-3}$ ), clay (wt %), and sand (wt %) - which vary in ease of mid-DRIFTS prediction, with TOC typically being extremely well predicted, pH and texture being predicted well, and BD sometimes being predicted with acceptable accuracy.

For the SCX samples, TOC was determined by elemental analysis with carbonate removal, when necessary, by acidification. pH was measured in a 1:1 water solution. Clay and sand were determined by particle size analysis using the hydrometer method after dispersion with sodium hexametaphosphate (Gee and Bauder, 1979). Bulk density was determined as the dry mass of <2 mm fraction per unit volume from core samples (3.5 cm i.d.). Given difficulties in estimating BD from hydraulic cores, BD values < 0.6 and > 1.8  $\text{g cm}^{-3}$  were excluded from further analysis in this paper.

For the NAPESHM samples, TOC was determined as the difference between total carbon via elemental analysis and inorganic carbon via pressure calcimeter. pH was measured in water, but with a 1:2 solution ratio. Bulk density was determined on the <2 mm fraction using a composite of two surface cores (7.6 cm length x 7.6 cm diameter). Clay and sand content were measured using the pipette method and wet sieving, respectively, after dispersion using a solution of sodium hexametaphosphate and sodium carbonate (Gee and Bauder, 1986). Despite small differences in methodologies for TOC, pH, texture, and BD, these two datasets have been combined for further analyses (Table 1).

### 2.3. Spectral acquisition

Mid-DRIFTS spectra were obtained on all samples after drying and sieving to 2 mm (henceforth referred to as <2 mm samples). A third of these samples were scanned in triplicate and the remainder scanned in duplicate. These <2 mm samples were then fine milled using a Retsch Mixer Mill 400 (Retsch GmbH, Germany), with hardened steel 50 ml vials and 1–3 grinding balls (sample dependent) set at a frequency of 25 Hz for 3 min (Baldock et al., 2013), before scanning. Fine-milled samples were only scanned once each. All fine-milled and <2 mm samples were scanned at the Woodwell Climate Research Center (Falmouth, MA) on a Bruker Vertex 70 FTIR (Bruker Optics, Billerica, MA, USA) with a wide-range beamsplitter/detector combination using a Pike AutoDiff (Pike Technologies, Fitchburg, WI, USA) diffuse reflectance accessory. The instrument was purged with ultra-dry CO<sub>2</sub>-free air from a Parker Hannifin FTIR Purge Gas generator (Parker Hannifin Corp., Lancaster, NY, USA). Sixty scans from 6000 to 180  $\text{cm}^{-1}$  were collected at 2  $\text{cm}^{-1}$  resolution and co-added for each sample to minimize noise. Background reflectance was collected on a mirror at the beginning of each set of 60 samples and was subtracted from sample reflectance before converting spectra to pseudo-absorption units. After background correction, replicate spectra were averaged to produce one spectrum per sample except, as noted below, when we designed a test of inter-replicate variability.

### 2.4. Chemometrics

First, outliers in the dataset were identified by building a partial least

squares regression (PLSR) model to predict TOC, and any sample falling more than 3 standard deviations ( $z$  threshold) from the 1:1 line was flagged as a potential outlier and removed from further analysis. This exercise identified 27 samples as outliers in the <2 mm dataset and 24 samples in the fine-milled dataset, with 15 of these samples being common between the <2 mm and fine-milled datasets. The outliers can be due to spurious analytical data or spurious spectral data. After removing these outliers, the initial sample set was trimmed to only include samples that were scanned both as fine-milled and <2 mm sample preparations resulting in a final number of 2380 unique samples being retained for all subsequent analyses.

Second, a series of chemometric models to estimate TOC, pH, clay, sand and BD were built on both the <2 mm and fine-milled spectra using cross-validation to assess model performance. Data were randomly split into 10 folds and blocked based on site in order to ensure that very similar samples were not included in both the calibration and validation sets. Models were calibrated using 9 folds and validated against the remaining fold and this procedure was repeated 10 times for PLSR models but only once for memory-based learning (MBL) models due to the computational cost. For example, 10-fold CV repeated 10 times took about 20 min per model combination for PLSR while a single repeat of an ensemble MBL (details explained below) would take 1–2 h using a 1.8 GHz, 4 core processor with 16 GB of RAM. Computing time of MBL depends on the test set size, training set size, and the range of hyperparameters tested, as dissimilarity calculation and local fitting optimization are run on-the-fly by the MBL algorithm.

Models were built using both PLSR and Memory Based Learning (MBL), a local modeling approach which finds the most appropriate subset of the training data and then builds a PLSR model using only those most similar spectra for each sample to be predicted. PLSR models were optimized based on reduction in the cross-validation root mean square error (RMSE) using up to 20 principal components. MBL models were optimized by trying three different dissimilarity measures (principal component analysis (pca), partial least squares regression (pls) and correlation coefficient (cor)) combined with a dissimilarity distance threshold ranging from 0.5 to 3.0 (by increments of 0.5) with the number of minimum and maximum neighbors ranging from 50 to 200, along with weighted PLSR automatically tuned by nearest neighbor validation. Different spectral preprocessing including baseline offset correction (BOC), Savitzky-Golay smoothing with first derivative (SG-FD), and Savitzky-Golay smoothing with standard normal variate (SNV) were tested along with log transformation of analytical data where appropriate. For each MBL model, all combinations of dissimilarity method, distance threshold and range of neighbors were tested, and the highest performing combination was reported for each combination of training set and soil property. All spectral preprocessing and statistics were performed using R statistical software environment (R Core Team, 2023) using the *pls* (v2.8–1; Liland et al., 2021), *resemble* (v2.2.1; Ramirez-Lopez et al., 2020), *prospectr* (v0.2.6; Stevens and Ramirez-Lopez, 2020), *spatialsample* (v0.2.1; Silge and Mahoney, 2022), and *tidyverse* (v1.3.2; (Wickham et al., 2019)) packages.

**Table 1**  
Summary statistics for each dataset and for the combination of datasets.

	NAPESHM					SCX					Both Labs Together				
	TOC	BD	Ph	CLAY	SAND	TOC	BD	PH	CLAY	SAND	TOC	BD	PH	CLAY	SAND
Minimum	0.20	0.61	3.95	0.13	3.00	0.10	0.65	4.50	2.24	1.35	0.10	0.61	3.95	0.13	1.35
1st Q <sup>a</sup>	1.09	1.09	5.53	13.25	18.00	0.46	1.08	6.10	22.64	8.46	0.98	1.09	5.61	1445	15.60
Median	1.46	1.21	6.28	21.41	30.60	0.80	1.19	6.610	29.07	20.21	1.36	1.21	6.38	22.87	28.20
Mean	1.63	1.20	6.42	22.26	34.78	0.97	1.18	6.73	27.80	24.80	1.51	1.20	6.47	23.23	32.99
3rd Q <sup>b</sup>	2.04	1.33	7.33	28.04	46.80	1.31	1.28	7.50	34.05	30.24	1.94	1.33	7.36	30.24	45.23
Max	6.01	1.70	9.03	61.94	90.50	3.44	1.78	8.60	51.35	94.69	6.01	1.78	9.03	61.94	94.69
n <sup>c</sup>	1953	1942	1953	1953	1953	427	351	427	427	427	2380	2293	2380	2380	2380

<sup>a</sup> first quartile.

<sup>b</sup> third quartile.

<sup>c</sup> number of samples with given soil property.

Further analysis was then conducted to understand the spectral diversity among replicates and whether or not TOC model performance was impacted when a different sample preparation was used for model building versus prediction. Out of the subset of <2 mm samples that were scanned in triplicate, 200 individual samples were selected randomly for further analysis. Two PLSR models were built using the <2 mm averaged data and the finely milled data, excluding the selected 200 samples from both datasets. Data were randomly split into 10 folds using site as a grouping factor. These models were calibrated using 9 folds and validated against the remaining fold. The 200-sample subsets of both sample preparations and as individual versus averaged spectra for the <2 mm preparation were then run as a test set for both models. Prediction performance was then compared between the two test sets.

Finally, predictions for all soil properties (TOC, pH, clay, sand and BD) were made using models built with the USDA NRCS NSSC-KSSL spectral library (hereafter denominated KSSL). The KSSL mid-DRIFTS spectral library has been steadily increasing in size and spectral diversity, being one of the largest publicly available mid-DRIFTS soil spectral libraries that represent the major soil types of the United States and other territories (Wijewardane et al., 2018; Ng et al., 2022). In combination with regional datasets, it is being used to develop global estimation services of soil properties (Shepherd et al., 2022). In this study, two different subsets of the KSSL library were used for model building. First, a subset of 15,000 samples was selected using the conditioned Latin hypercube sampling (cLHS) algorithm (Minasny and McBratney, 2006) to obtain a subset of samples that were representative of the entire 80,000+ sample library. This is the same subset that was used in Sanderman et al. (2021). Second, a similar sized subset that was more spectrally comparable to the samples scanned in this study was selected. For this second subset, *prospectr* R package was used to search the 100 closest spectral neighbors of each fine-milled sample ( $n = 2380$ ) with the *search\_neighbor* function. Principal component analysis (PCA), with column centering and scaling, was run before the neighbor search to reduce spectral dimensionality while accounting for 99.99 % of the original cumulative variance. This procedure yielded 18,000 unique samples. This second subset is subsequently called the nearest neighbor (NN) subset.

PLSR models were then built using both subsets by randomly splitting the KSSL data into 10 folds for defining the best number of factors (tested up to 20) that minimized the cross-validation RMSE. The NAPESHM & SCX spectra that were shared across both the fine-milled and <2 mm scans ( $n = 2380$ ) were then run through the models as test sets. MBL models were also built using the KSSL subsets as the calibration set used for neighbors search and the NAPESHM & SCX datasets as the test sets, again for <2 mm average and fine-milled sets using the same optimization options as for the cross validation. In addition to SNV pretreatment, the spectra acquired at Woodwell were also transformed using Spectral Space Transformation (Du et al., 2011), a calibration transfer technique where the spectra on a secondary instrument are mathematically transformed to “look” like they were acquired on the primary instrument to minimize differences of the scanning instrument and environment. The method is based on the transformation of a secondary instrument onto the spectral space of the primary one by estimating a transformation matrix with singular value decomposition, the same algorithm employed in PCA. A set of standard samples measured in both laboratories ( $n = 493$ ) was used to construct the transformation matrix. The KSSL Vertex was set as the primary instrument while the Woodwell Vertex was treated as secondary. The number of orthogonal features used by SST was set to retain 99.99 % of the original cumulative variance. Detailed information about SST can be found in Pittaki-Chrysodonta et al. (2021).

Model performance was reported using multiple goodness of fit tests including root mean square error (RMSE), bias, coefficient of determination ( $R^2$ ), Lin’s concordance correlation coefficient (CCC), and ratio of performance to interquartile range (RPIQ =  $(Q3-Q1)/RMSE$ ).

### 3. Results and discussion

#### 3.1. Can good unbiased predictions be achieved from scanning samples that have been through standard sample preparation of sieving to 2 mm? Does fine milling improve the predictions of key soil properties?

Before exploring predictive model results, it is informative to examine the similarity or dissimilarity between spectra acquired with and without fine milling. There were only very minor differences in spectral shape between the fine-milled and <2 mm samples with a very small offset between the average spectra below  $1200\text{ cm}^{-1}$  (Fig. 2A and S3) but little differentiation between the two sets of samples in principal component space (Fig. 2B).

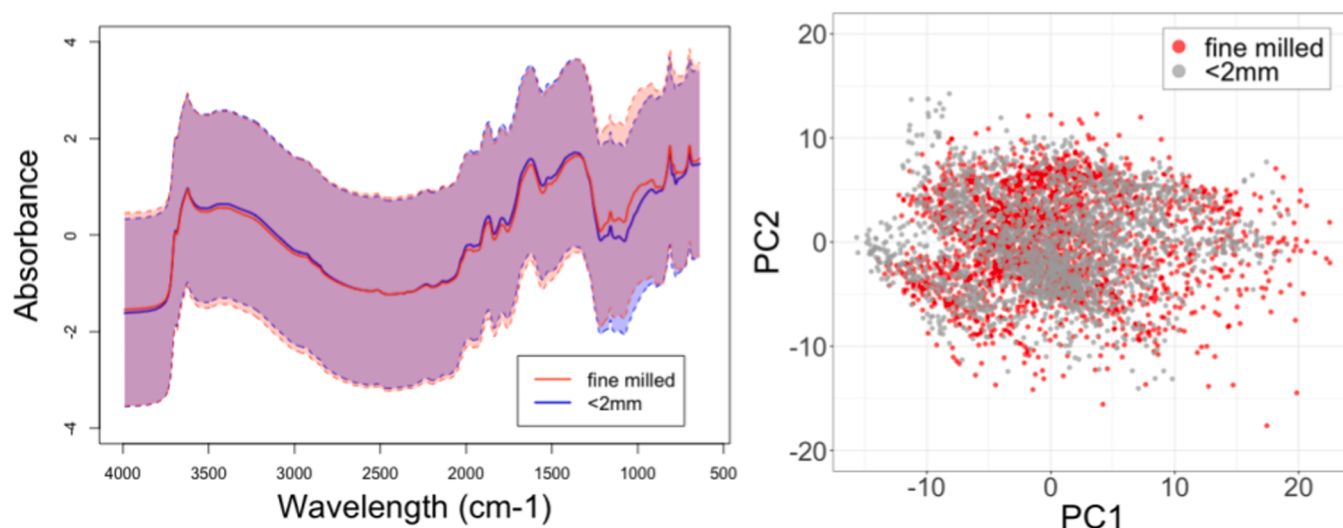
To determine if standard sample preparation (i.e. the <2 mm or fine earth fraction) without fine milling can produce unbiased and accurate predictions, predictive models were built from the assembled spectral libraries using a cross validation scheme where all samples from entire sites were withheld to assess model performance. Pretreatment options all had fairly similar performance with SNV slightly outperforming BOC and SG-FD pretreatments (Table S1). Because of this finding, only the results from SNV pretreatment are highlighted onwards. Additionally, there were only small differences in predictive performance between PLSR and MBL models with MBL producing marginally better results for all properties except TOC (Tables S2 & S3). Overall, as seen in Fig. 3, there was very little difference in the performance between the models built using the fine-milled and <2 mm spectra. Focusing on the <2 mm spectra using MBL and SNV pretreatment, RMSEs were  $0.13\text{ g cm}^{-3}$ , 0.46, 0.34 %, 4.8 % and 8.4 % for BD, pH, TOC, clay and sand, respectively, with negligible bias (Table S3). Utilizing Ng et al. (2022)’s accuracy categorization scheme, clay, pH, sand and TOC when predicted from the <2 mm spectra would be considered highly accurate models and BD would be considered to have medium accuracy. Performance for these five soil properties were generally similar to prior studies (Soriano-Disla et al., 2014), even studies using much larger datasets (Sanderman et al., 2020; Ng et al., 2022). Bulk density is a soil property that is hard to predict because there are no direct spectral features for BD and as such model performance tends to vary from medium to unacceptable across studies with the results from this study falling in the middle.

A second test of the impact of sample preparation is to apply models built on scans of the finely milled samples to the averaged <2 mm scans and vice versa. A set of 200 samples (the same set used in Section 3.2) was withheld from both sample preparation datasets, then PLSR models were built to estimate TOC on the remaining samples. Each of these two models were then applied to spectra obtained on both sample preparations. These PLSR models performed very similarly to the full subset regardless of sample preparation (Table 2). When the model was trained on the finely milled samples, the Lin’s CCC of the held-out test sets were 0.86 and 0.85 for the averaged <2 mm and finely milled samples, respectively. When the model was trained on the averaged <2 mm samples, the Lin’s CCC of the held-out test sets were 0.87 and 0.86 for the averaged <2 mm and finely milled samples, respectively. Table 2 suggests that there was a very minor drop in performance (i.e., RMSE increased by 0.01 – 0.03 %) when the different sample preparation was predicted primarily because a small amount of bias was found. Despite these small differences, the upshot of this analysis is that sample preparation doesn’t impact model performance.

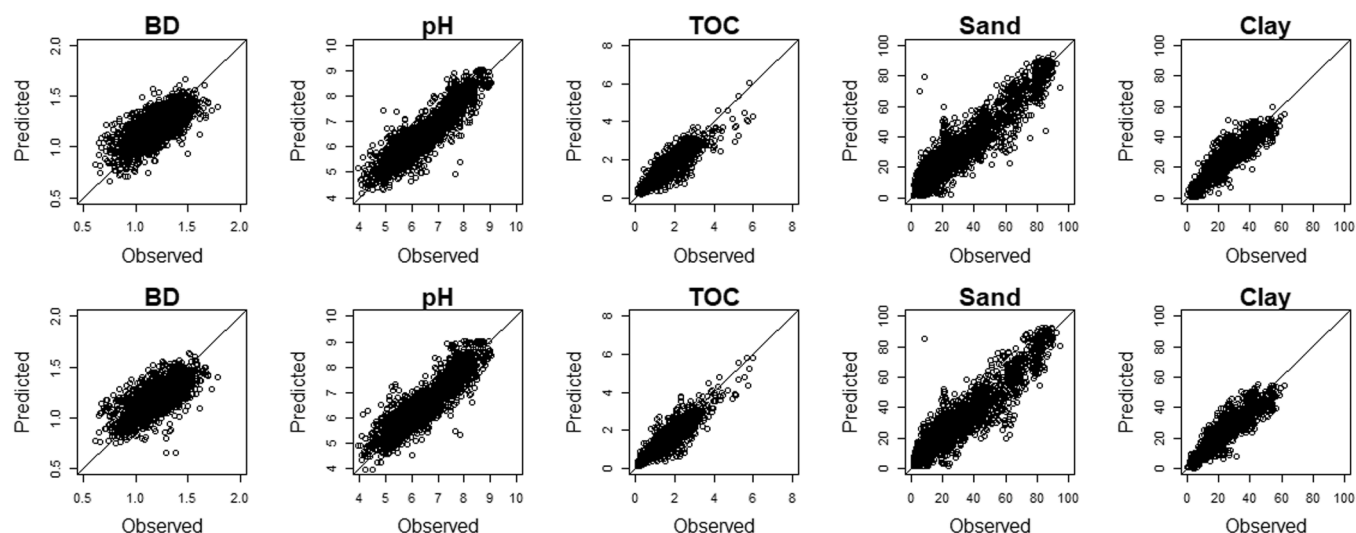
#### 3.2. Is there more variability between replicates without fine milling?

A subset of 200 <2 mm samples were scanned in triplicate to better understand if there was additional heterogeneity introduced by skipping the fine milling step (Fig. 4). As described above, these samples were withheld from the library and a new PLSR model was built with the remaining <2 mm average spectra. The average error between predicted TOC value from each replicate and the predicted TOC value from the





**Fig. 2.** (A) Average spectra with shading showing 5–95th percentile of data for the <2 mm and fine-milled samples. (B) Principal Component Analysis of the same samples. Figure S3 gives examples of 4 contrasting spectra.



**Fig. 3.** Plots of observed versus predicted data for best MBL models built from finely milled spectra (top panels) and <2 mm spectra (bottom panels) for BD (g cm<sup>-3</sup>), pH and TOC (%), sand (%) and clay (%). Model performance reported in Table S3.

**Table 2**

Partial least squares regression models for total organic carbon (TOC) built on each sample preparation and applied to 200 held-out spectra of both sample preparations.

Training set	Test set	RMSE	Bias	R <sup>2</sup>	CCC	RPIQ
fine milled	fine milled	0.30	0.00	0.86	0.92	3.52
fine milled	avg <2 mm	0.31	0.02	0.85	0.92	3.35
avg <2 mm	avg <2 mm	0.29	0.00	0.87	0.93	3.63
avg <2 mm	fine milled	0.32	-0.06	0.85	0.90	3.20

average of the replicate scans was 0.08 %. The RMSE of the predictions from the averaged spectra compared to observed laboratory data was 0.31 %, while the RMSE of the prediction from each replicate relative to the prediction from the averaged spectra was only 0.13 % suggesting that the replicate-to-replicate variance is significantly smaller than the prediction uncertainty. This finding suggests that replication may not be necessary to overcome the additional heterogeneity in the <2 mm samples compared to fine-milled samples.

### 3.3. Are there specific soil parameters that control the performance of predictive models?

We hypothesized that if there were significant differences in spectral model performance between the fine-milled and <2 mm spectra, these differences would be greatest with high sand content because the lower wavenumbers (higher wavelengths) would be most impacted by particle size. We did not find any difference in performance between the two sample preparations, but this question was still worth exploring to see if texture controls residuals of TOC predictive performance. The <2 mm TOC model error was found to be unrelated to sand content (Fig. 5) across a sand content range of 1.1 - 95.1 % (interquartile range of 15.3 - 45.4 %). Predictive performance, measured as relative absolute model error ( $100 \times \text{abs}(\text{obs} - \text{pred})/\text{obs}$ ), did not vary across sand content quartiles averaging 19.8 % relative absolute error. The RPIQ actually improves with increasing sand quartile from 2.2 to 3.3 (data not shown), mostly because there is more variance in TOC at high sand levels.

While soil texture does not appear to control TOC model performance, we did notice a systematic difference in model performance between the two sample sets. It is well known that the quality of the

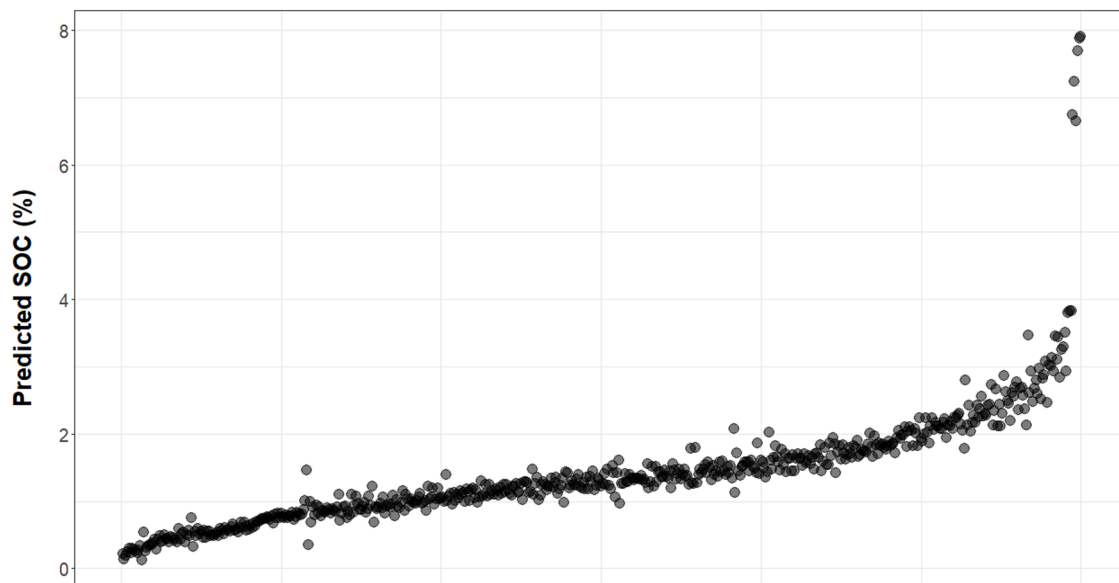


Fig. 4. Predictions of TOC from individual replicates for 200 samples. Samples are indexed from lowest to highest observed TOC on x-axis.

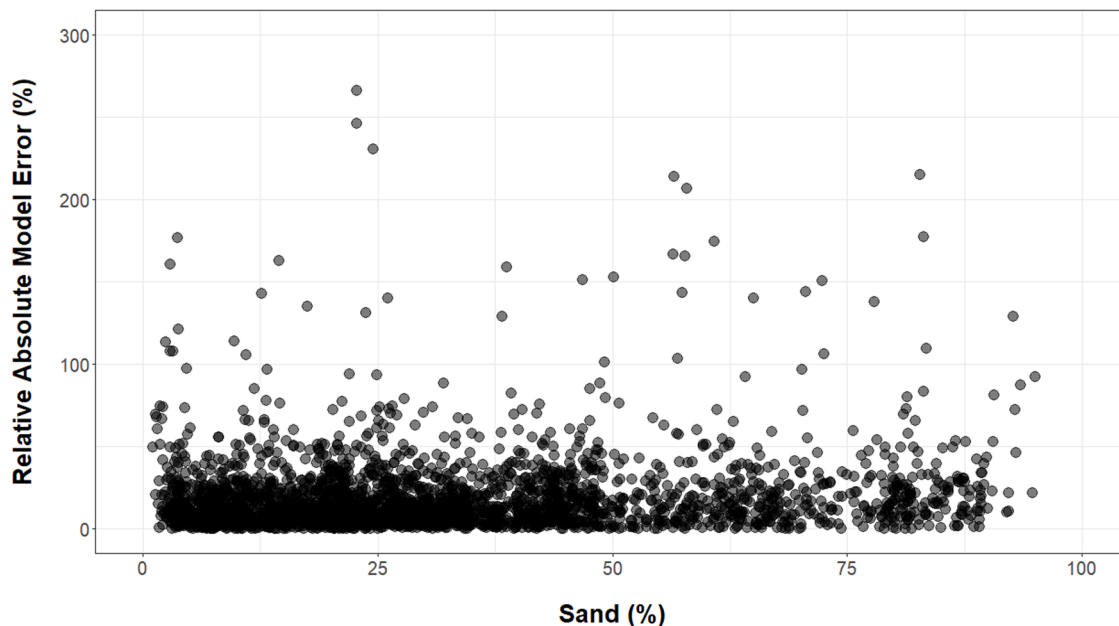


Fig. 5. Relative model error was unrelated to sand content of the <2 mm sample.

analytical data will have a large impact upon model performance (Shepherd et al., 2022), both when building models with cross-validation but especially when using an existing spectral library that was built on analytical data collected in a different lab. Dividing the cross-validation <2 mm averaged soil spectra PLSR model performance for TOC (Table S2) into the two datasets, Lin's CCC was 0.92 and 0.86 for NAPESHM and SCX, respectively. When breaking down the prediction performance of the KSSL PLSR models on the <2 mm spectra reported in Table 3, Lin's CCC = 0.85 and  $R^2 = 0.81$  for the NAPESHM samples and these same statistics were lower at 0.81 and 0.68, respectively, for the SCX samples. Similar differences were seen for all soil properties. The outlier screening may also suggest a difference in data quality. For the NAPESHM dataset, the outlier rate was just under 1 % while in the SCX dataset, the outlier rate was 3.4 %.

#### 3.4. Can existing predictive models (i.e. USDA NRCS NSSC-KSSL database) built on finely milled samples be used on <2 mm preparation samples?

To test this question, we first subset the full KSSL library into a generally representative subset using cLHS ( $n \sim 15,000$ ) and a subset of the most spectrally similar samples ( $n = 18,000$ ) before model building. Two model types (PLSR and MBL) were applied using two pretreatment methods (SNV and SST). Generally, the best model performance regardless of sample preparation or KSSL subset was found using the MBL model with SNV standardization (Table 3). Calibration transfer using SST rarely improved model results. Similar to the findings with cross-validation (Section 3.1), there was little difference in predictive performance between the two sample preparation methods with the exception of TOC and sand where the best fine-milled sample performance was about 7 % higher than for the <2 mm samples (Table 3). The

**Table 3**

Performance, measured by Lin's Concordance Correlation Coefficient (CCC), of KSSL models for bulk density (BD, g cm<sup>-3</sup>), clay (wt%), pH, sand (wt%) and total organic carbon (TOC, wt%) applied to the fine-milled and <2 mm samples in this study. Other performance metrics including R<sup>2</sup>, RMSE, bias, and RPIQ are given in Tables S4–S7.

Preparation	Subset <sup>a</sup>	Model <sup>b</sup>	BD	Clay	pH	Sand	TOC
fine milled	NN	MBL_SNV	0.34	0.74	0.79	0.81	0.85
fine milled	cLHS	MBL_SNV	0.38	0.78	0.77	0.81	0.78
fine milled	NN	MBL_SST	0.21	0.70	0.49	0.56	0.66
fine milled	cLHS	MBL_SST	0.16	0.60	0.46	0.49	0.67
fine milled	NN	PLSR_SNV	0.49	0.81	0.76	0.69	0.57
fine milled	cLHS	PLSR_SNV	0.46	0.73	0.79	0.72	0.60
fine milled	NN	PLSR_SST	0.35	0.64	0.69	0.71	0.68
fine milled	cLHS	PLSR_SST	0.41	0.64	0.69	0.78	0.69
<2mm	NN	MBL_SNV	0.24	0.73	0.80	0.76	0.69
<2mm	cLHS	MBL_SNV	0.32	0.75	0.77	0.74	0.79
<2mm	NN	MBL_SST	0.28	0.66	0.48	0.57	0.68
<2mm	cLHS	MBL_SST	0.25	0.62	0.47	0.55	0.67
<2mm	NN	PLSR_SNV	0.51	0.78	0.69	0.71	0.43
<2mm	cLHS	PLSR_SNV	0.50	0.67	0.74	0.73	0.52
<2mm	NN	PLSR_SST	0.48	0.62	0.64	0.69	0.73
<2mm	cLHS	PLSR_SST	0.52	0.58	0.67	0.77	0.72

<sup>a</sup> NN refers to the nearest-neighbor KSSL subset and cLHS refers to the subset of the KSSL chosen using conditioned Latin hypercube sampling.

<sup>b</sup> Two models (MBL = memory-based learning, PLSR = partial least squares regression) and two preprocessing methods (SNV = standard normal variate, SST = spectral space transformation) were tested in combination.

finding of a Lin's CCC of 0.85 for TOC was similar to the findings in Sanderman et al. (2021), who implemented a nearly identical modeling approach using the cLHS subset of the KSSL on finely milled samples from various US long-term agricultural trials scanned on the same instrument used in this current study, where Lin's CCC was 0.92 for TOC across all 1377 samples in that study.

While the results on the <2 mm averaged spectra were fairly good for TOC, pH, clay and sand, performance for BD was not acceptable (Lin's CCC varied from 0.25 to 0.51 depending on model choice for BD). Additionally, these results were in all cases inferior to the cross-validation performance (Tables S2 & S3) where Lin's CCC was 5, 15, 12, 14 and 40 % higher for TOC, pH, clay, sand and BD, respectively.

Here we further explore some of the reasons for the lower performance of the KSSL models. When applying a model built from a pre-existing library there are several sources of variability that need to be considered. First, the spectra were acquired on different instruments and in the case of this study with different sample preparations as well. Second, the geographic or spectral distribution of the library might not match the sample set. Third, the analytical data were also acquired in different laboratories. The third point was discussed in more detail in section 3.3; here, we discuss points one and two in more detail.

The representation of the test samples in the training space is a factor of both the instrument the spectra were acquired on and the geographic/spectral representativeness of the library relative to the test samples. If there is a systematic offset between the Woodwell instrument and the KSSL instrument as Dangal and Sanderman (Dangal and Sanderman, 2020) suggested, then the projection of the test set into the spectral space of the training library will be biased and the nearest neighbors in the library will not necessarily be representative of that test sample. In Fig. 6, we have projected the <2 mm SNV transformed spectra before and after SST calibration transfer onto the scores of the PLSR model for TOC built on the cLHS subset of the KSSL library (the same projections done on the nearest neighbor subset are shown in Fig. S1). The projection of the fine-milled samples appears almost identical in both cases (data not shown). Before SST, the SNV-transformed <2 mm spectra mostly fall within the PLS scores for factors 1–4; however, after SST, many of the samples are sitting on the edge or outside of the KSSL spectral space (Fig. 6). Assuming that SST did properly correct systematic differences in spectra, which our prior research

(Pittaki-Chrysodonta et al., 2021; Sanderman et al., 2023) suggests, then the library isn't actually capturing the variability in our test data. The poorer performance of the KSSL models before SST transfer is likely due to spectral misalignment with the training set. After SST, the PLSR models improved significantly, with RMSE dropping by 50 % (Table S5), despite the fact that a proportion of samples now fall outside of the calibration space. However, the MBL models did not improve. Prediction performance of large PLSR models is less sensitive to the location of the spectra within the calibration space, so with proper alignment, the PLSR models improved greatly. The MBL models did not improve likely because there weren't enough KSSL spectra close to the test samples that were falling outside of the calibration space.

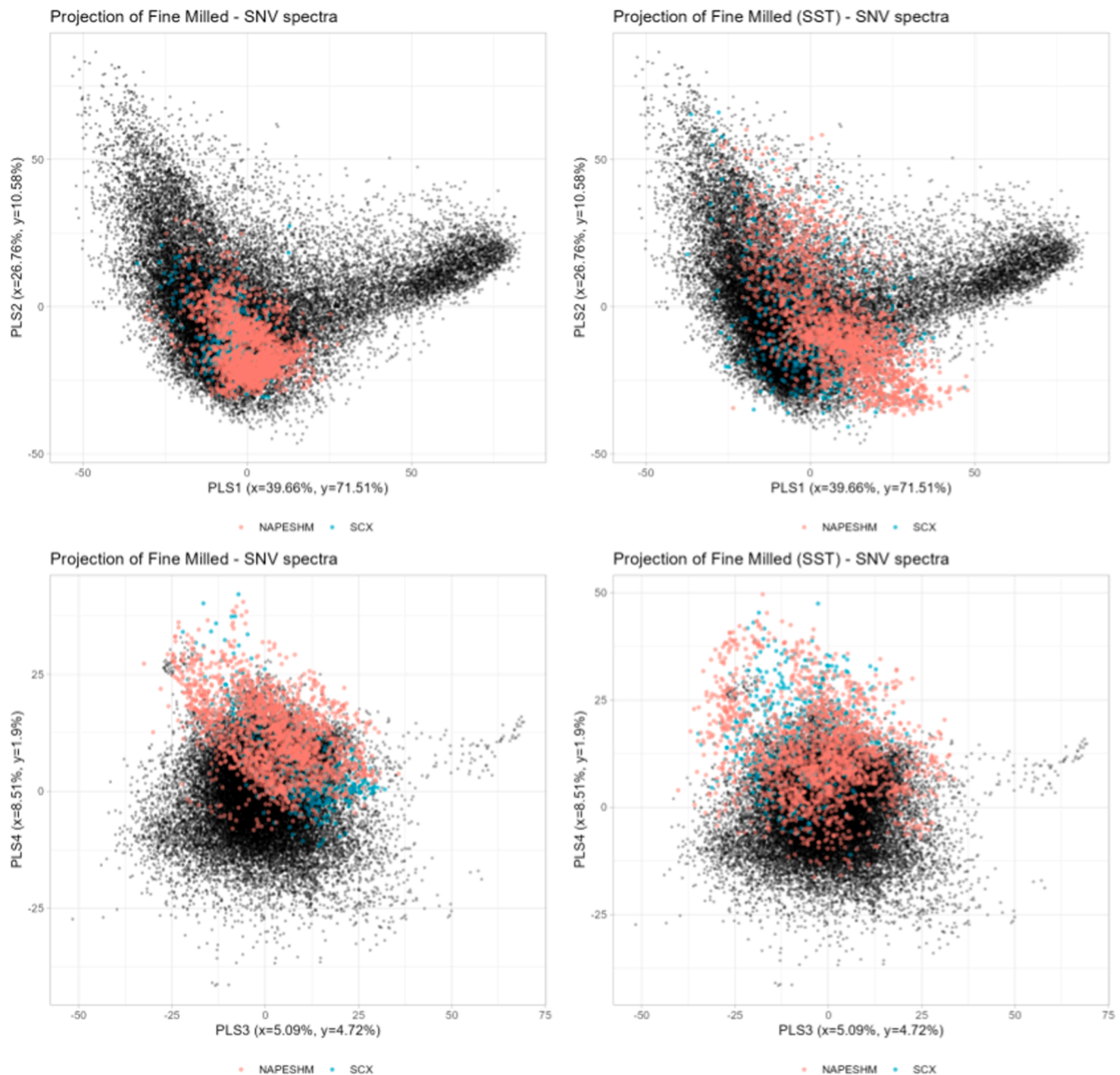
An examination of the geographic representativeness of these samples supports the notion, seen in Fig. 6, that the KSSL library does not capture the soil variability of the samples in this study. Out of the 2380 samples, only 1147 and 1074 overlapped at the county level with the cLHS and NN subsets of the full KSSL library, respectively. The NAPESHM dataset contains 558 samples from Canada and Mexico, but many samples from typical agricultural regions of the US are also underrepresented (Fig. S2).

### 3.5. Potential implications of adoption

Given the moderate to high variance of most soil properties, soil sampling protocols often require large sample sizes for measurement confidence and to assess field-scale soil heterogeneity (Pennock et al., 2007). Additionally, to account for seasonal variation and ensure consistency (Franzuebbers et al., 1995; Alletto and Coquet, 2009), soil samples are also typically collected in the same season each year, leading to a high volume of samples requiring processing concurrently. As ecosystem services and soil carbon markets mature, and the demand for consistent soil testing grows, commercial labs could reach functional capacity and become strained. The body of evidence suggests that mid-DRIFTS spectroscopy is a cost-effective and practical complement to traditional methods of soil analysis for numerous soil properties (Baldock et al., 2013; Gomez et al., 2020; Sanderman et al., 2021; Summerauer et al., 2021; Ng et al., 2022), and multiple commercial labs are increasingly exploring or already adopting this technology. Critically, the additional uncertainty introduced by mid-DRIFTS estimation of many soil properties is minor compared to the field-scale variance in these same properties (Viscarra Rossel et al., 2016; Li et al., 2022). Since mid-DRIFTS estimates are typically unbiased, especially when some local samples are included in the training set (Seidel et al., 2019), the cost savings that mid-DRIFTS spectroscopy affords can allow for more samples to be analyzed and thus a more precise estimate of a field can be achieved. Eliminating fine milling from spectroscopy protocols, which we found in this study, in agreement with other recent work (Bachion de Santana and Daly, 2022), did not measurably impact accuracy or variance of predictions for TOC, pH, clay and sand, can further expedite soil analysis, reduce overall costs, and fit better into the typical workflow of a commercial laboratory where fine milling is often seen as a time prohibitive step. Just focusing on TOC as an example, a typical elemental analyzer can analyze about 80 unknown samples in a day, after a full day of fine milling, but a FTIR spectrometer with an auto-sampler attachment can analyze at least 400 samples for the same amount of technician time. If carbonates are present, this comparative advantage is increased further since a second measurement of carbonates or pretreatment to remove carbonates would be required for the traditional analyses.

Ultimately, the quality of predictions will largely depend on the spectral library used. The libraries need to be representative of the range of soil properties that will be encountered, and care needs to be taken to ensure that newly scanned samples are compatible with the library. Additionally, it is important that laboratories follow strict quality control and quality assurance procedures for acquisition of spectra (e.g., ICRAF, 2021; Soil Survey Staff, 2022) and for determination of the





**Fig. 6.** Fine-milled SNV-transformed spectra before (left) and after SST calibration transfer (right) projected onto the scores of the PLSR model for TOC built on the cLHS subset of the KSSL library. Percent of spectral variance (x) and TOC variance (y) explained for each factor shown in parentheses on the axes.

accompanying wet chemical data. As part of good practice for soil spectroscopy, we recommend routinely running a small percentage of soils through traditional analyses which would serve three important purposes: 1) it would allow for rigorous evaluation of the spectral-based estimates; 2) model predictions can be improved through spiking with local samples from that project; and 3) in the longer-term, a more complete library will be built.

#### 4. Conclusions

The results from this investigation using 2380 soil samples from across North American agricultural soils support the conclusions of [Bachion de Santana and Daly \(2022\)](#) and [Deiss et al. \(2020\)](#) that fine milling of soil samples prior to spectral analysis is not needed for soil properties that are typically well predicted. Model performance for TOC,

pH, clay and sand were quite good when spectra were collected on the <2 mm soil samples. There was little improvement in predictive results for these properties when models were built on finely milled samples. Models trained on one sample preparation and applied to the other performed almost equally as well. Additionally, there was generally only small variability between replicate aliquots of the same <2 mm sample. These findings support the notion that a laboratory can build their own spectral library based upon scans of <2 mm samples and be able to produce accurate and unbiased estimates of TOC, pH, clay and sand content. However, BD was more difficult to predict and is likely not a good candidate for analysis via mid-DRIFTS spectroscopy.

This research also demonstrated that <2 mm spectra could be predicted using an existing spectral library built on finely milled samples. In contrast with the findings of [Wijewardane et al. \(2021\)](#), who found lower predictive performance for the <2 mm preparation, the <2 mm

soil spectra in this investigation performed as well or better than the fine-milled soil spectra for four key soil properties (TOC, pH, clay and sand). The overall predictive performance using the KSSL models was slightly lower than when models were built from the samples used in this study alone; however, given that there are multiple reasons for the performance of the existing KSSL library-based models to suffer when applied to new soils scanned on a different instrument using a different preparation method, these results are still encouraging. Given the inconsistent findings across studies, the performance of any spectroscopy-based inference system using an existing library needs to be thoroughly tested before routine application to new samples, regardless of sample preparation method.

### Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Please note that Indigo Ag, Inc. provided funding for this study, employs several co-authors, and is an organization working to support sustainable agriculture and soil organic carbon accounting.

### Data availability

Data will be made available on request.

### Acknowledgements

We kindly acknowledge the contribution of Matthew Jacques in the Woodwell Climate laboratory; Lauren Matosziuk for insightful discussions that informed the hypotheses to test. The NAPESHM (North American Project to Evaluate Soil Health Measurements) project is a broader effort titled, "Assessing and Expanding Soil Health for Production, Economic, and Environmental Benefits". The NAPESHM project was funded by the Foundation for Food and Agricultural Research (grant ID 523926), General Mills, and The Samuel Roberts Noble Foundation. The Soil Health Institute acknowledges its great appreciation for the collaborative efforts of partnering scientists for NAPESHM and the scientists and technicians that maintained the long-term research sites sampled in this project. We are also thankful to Cornell Soil Health Laboratory and the Soil Water and Environmental Lab at The Ohio State University for performing analysis on samples. Indigo Ag, Inc. provided funding for this study, employs several co-authors, and is an organization working to support sustainable agriculture and soil organic carbon accounting. JLS was supported by USDA NIFA award #2020-67021-32467.

### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.soisec.2023.100104](https://doi.org/10.1016/j.soisec.2023.100104).

### References

- Alletto, L., Coquet, Y., 2009. Temporal and spatial variability of soil bulk density and near-saturated hydraulic conductivity under two contrasted tillage management systems. *Geoderma* 152. <https://doi.org/10.1016/j.geoderma.2009.05.023>.
- Bachion de Santana, F., Daly, K., 2022. A comparative study of MIR and NIR spectral models using ball-milled and sieved soil for the prediction of a range soil physical and chemical parameters. *Mol. Biomol. Spectroscopy* 279. <https://doi.org/10.1016/j.saa.2022.121441>.
- Bagnall, D.K., Morgan, C.L.S., Bean, G.M., Liptzin, D., Cappellazzi, S.B., Cope, M., Greub, K.L.H., Rieke, E.L., Norris, C.E., Tracy, P.W., Aberle, E., Ashworth, A., Tavarez, O.B., Bary, A.I., Baumhardt, R.L., Gracia, A.B., Brainard, D.C., Brennan, J.R., Reyes, D.B., Bruhjell, D., Carlyle, C.N., Crawford, J.J.W., Creech, C.F., Culman, S.W., Deen, B., Dell, C.J., Derner, J.D., Ducey, T.F., Duiker, S.W., Dyck, M.F., Ellert, B.H., Entz, M.H., Solorio, A.E., Fonte, S.J., Fonteyne, S., Fortuna, A.M., Foster, J.L., Fultz, L.M., Gamble, A.V., Geddes, C.M., Griffin-LaHue, D., Grove, J.H., Hamilton, S.K., Hao, X., Hayden, Z.D., Honsdorf, N., Howe, J.A., Ippolito, J.A., Johnson, G.A., Kautz, M.A., Kitchen, N.R., Kumar, S., Kurtz, K.S.M., Larney, F.J., Lewis, K.L., Liebman, M., Ramirez, A.L., Machado, S., Maharjan, B., Gamiño, M.A.M., May, W.E., McClaran, M.P., McDaniel, M.D., Millar, N., Mitchell, J.P., Moore, A.D., Moore, P.A., Gutiérrez, M.M., Nelson, K.A., Omondi, E.C., Osborne, S.L., Alcalá, L.O., Owens, P., Pena-Yewtukhiw, E.M., Poffenbarger, H.J., Lira, B.P., Reeve, J.R., Reinbott, T.M., Reiter, M.S., Ritchey, E.L., Roozeboom, K.L., Rui, Y., Sadeghpour, A., Sainju, U.M., Sanford, G.R., Schillinger, W.F., Schindelbeck, R.R., Schipanski, M.E., Schlegel, A.J., Scow, K.M., Sherrod, L.A., Shober, A.L., Sidhu, S.S., Moya, E.S., St. Luce, M., Strock, J.S., Suyker, A.E., Sykes, V.R., Tao, H., Campos, A.T., Van Eerd, L.L., van Es, H.M., Verhulst, N., Vyn, T.J., Wang, Y., Watts, D.B., Wright, D.L., Zhang, T., Honeycutt, C.W., 2022. Selecting soil hydraulic properties as indicators of soil health: measurement response to management and site characteristics. *Soil Sci. Soc. Am. J.* 86 <https://doi.org/10.1002/saj2.20428>.
- 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*. <https://doi.org/10.1071/SR13077>.
- Dangal, S., Sanderman, J., Wills, S., Ramirez-Lopez, L., Dangal, S.R.S., Sanderman, J., Wills, S., Ramirez-Lopez, L., 2019. Accurate and precise prediction of soil properties from a large mid-infrared spectral library. *Soil Sys.* 3, 11. <https://doi.org/10.3390/soilsystems3010011>.
- Dangal, S.R.S., Sanderman, J., 2020. Is standardization necessary for sharing of a large mid-infrared soil spectral library? *Sensors (Switzerland)* 20. <https://doi.org/10.3390/s20236729>.
- Deiss, L., Culman, S.W., Demyan, M.S., 2020. Grinding and spectra replication often improves mid-DRIFTS predictions of soil properties. *Soil Sci. Soc. Am. J.* 84 <https://doi.org/10.1002/saj2.20021>.
- Du, W., Chen, Z.P., Zhong, L.J., Wang, S.X., Yu, R.Q., Nordon, A., Littlejohn, D., Holden, M., 2011. Maintaining the predictive abilities of multivariate calibration models by spectral space transformation. *Anal. Chim. Acta* 690, 64–70. <https://doi.org/10.1016/j.aca.2011.02.014>.
- Franzluebbers, A.J., Hons, F.M., Zuberer, D.A., 1995. Tillage and crop effects on seasonal dynamics of soil CO<sub>2</sub> evolution, water content, temperature, and bulk density. *Appl. Soil Ecol.* 2 [https://doi.org/10.1016/0929-1393\(94\)00044-8](https://doi.org/10.1016/0929-1393(94)00044-8).
- Gee, G.W., Bauder, J.W., 1986. Particle size analysis. In: Klute, A. (Ed.), *Method of Soil Analysis, Part 1: Physical and Mineralogical Methods*. Soil Science Society of America, Madison, Wisconsin USA.
- Gee, G.W., Bauder, J.W., 1979. Particle size analysis by hydrometer: a simplified method for routine textural analysis and a sensitivity test of measurement parameters. *Soil Sci. Soc. Am. J.* 43 <https://doi.org/10.2136/sssaj1979.03615995004300050038x>.
- Gomez, C., Chevallier, T., Moulin, P., Boufferra, I., Hmaidi, K., Arrouays, D., Jolivet, C., Barthès, B.G., 2020. Prediction of soil organic and inorganic carbon concentrations in Tunisian samples by mid-infrared reflectance spectroscopy using a French national library. *Geoderma* 375, 114469. <https://doi.org/10.1016/j.geoderma.2020.114469>.
- ICRAF, 2021. SOP 002: sample analysis using Bruker ALPHA Fourier transform mid-infrared spectrometer. [https://www.worldagroforestry.org/sites/agroforestry/files/SOP%20for%20sample%20analysis%20on%20Bruker%20Alpha%20Spectrometer\\_0.pdf](https://www.worldagroforestry.org/sites/agroforestry/files/SOP%20for%20sample%20analysis%20on%20Bruker%20Alpha%20Spectrometer_0.pdf). Accessed 21 Aug 2023.
- Li, S., Viscarra Rossel, R.A., Webster, R., 2022. The cost-effectiveness of reflectance spectroscopy for estimating soil organic carbon. *Eur. J. Soil Sci.* 73 <https://doi.org/10.1111/ejss.13202>.
- Liland, K.H., Mevik, B.-H., & Wehrens R. 2021. Pls: partial least squares and principal component regression. R package version 2.8-1. <https://CRAN.R-project.org/package=pls>.
- Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Comp. Geosci.* 32 <https://doi.org/10.1016/j.cageo.2005.12.009>.
- Ng, W., Minasny, B., Jeon, S.H., McBratney, A., 2022. Mid-infrared spectroscopy for accurate measurement of an extensive set of soil properties for assessing soil functions. *Soil Security* 6. <https://doi.org/10.1016/j.soisec.2022.100043>.
- Nocita, M., Stevens, A., van Wesemael, B., Brown, D.J., Shepherd, K.D., Towett, E., Vargas, R., Montanarella, L., 2015. Soil spectroscopy: an opportunity to be seized. *Glob. Chang. Biol.* 21, 10–11. <https://doi.org/10.1111/gcb.12632>.
- Norris, C.E., Bean, G.M., Cappellazzi, S.B., Cope, M., Greub, K.L.H., Liptzin, D., Rieke, E.L., Tracy, P.W., Morgan, C.L.S., Honeycutt, C.W., 2020. Introducing the North American project to evaluate soil health measurements. *Agron. J.* <https://doi.org/10.1002/agj2.20234>.
- Pennock, D., Yates, T., Braidek, J., 2007. Soil sampling designs, in: *Soil Sampling and Methods of Analysis: Second Edition*. doi:10.1201/9781420005271-7.
- Pittaki-Chrysodonta, Z., Hartemink, A.E., Sanderman, J., Ge, Y., Huang, J., 2021. Evaluating three calibration transfer methods for predictions of soil properties using mid-infrared spectroscopy. *Soil Sci. Soc. Am. J.* 85 <https://doi.org/10.1002/saj2.20225>.
- R Core Team, 2023. R: A Language and Environment For Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Ramirez-Lopez, L., and Stevens, A., and Viscarra Rossel, R., and Lobsey, C., and Wadoux, A., and Breure, T. 2020. resemble: regression and similarity evaluation for memory-based learning in spectral chemometrics. R package Vignette R package version 2.2.1. <http://l-ramirez-lopez.github.io/resemble/>.
- Sanderman, J., Baldock, J.A., Dangal, S.R.S., Ludwig, S., Potter, S., Rivard, C., Savage, K., 2021a. Soil organic carbon fractions in the Great Plains of the United States: an application of mid-infrared spectroscopy. *Biogeochemistry*. <https://doi.org/10.1007/s10533-021-00755-1>.
- Sanderman, J., Gholizadeh, A., Pittaki-Chrysodonta, Z., Huang, J., Safanelli, J.L., Ferguson, R., 2023. Transferability of a large mid-infrared soil spectral library

- between two Fourier-transform infrared spectrometers. *Soil Sci. Soc. Am. J.* <https://doi.org/10.1002/saj2.20513>.
- Sanderman, J., Savage, K., Danga, S.R.S., 2020. Mid-infrared spectroscopy for prediction of soil health indicators in the United States. *Soil Sci. Soc. Am. J.* 84, 251–261. <https://doi.org/10.1002/SAJ2.20009>.
- Sanderman, J., Savage, K., Danga, S.R.S., Duran, G., Rivard, C., Cavigelli, M.A., Gollany, H.T., Jin, V.L., Liebig, M.A., Omondi, E.C., Rui, Y., Stewart, C., 2021b. Can agricultural management induced changes in soil organic carbon be detected using mid-infrared spectroscopy? *Remote Sens.* 13. <https://doi.org/10.3390/rs13122265>.
- Seidel, M., Hutengs, C., Ludwig, B., Thiele-Bruhn, S., Vohland, M., 2019. Strategies for the efficient estimation of soil organic carbon at the field scale with vis-NIR spectroscopy: spectral libraries and spiking vs. local calibrations. *Geoderma* 354. <https://doi.org/10.1016/j.geoderma.2019.07.014>.
- Shepherd, K.D., Ferguson, R., Hoover, D., van Egmond, F., Sanderman, J., Ge, Y., 2022. A global soil spectral calibration library and estimation service. *Soil Security* 7. <https://doi.org/10.1016/j.soisec.2022.100061>.
- Silge, J., & Mahoney, M. 2022. spatialsample: spatial resampling infrastructure. R package version 0.2.1. <https://CRAN.R-project.org/package=spatialsample>.
- Soil Survey Staff, 2022. Soil Survey Investigations Report No. 42, Version 6.0. U.S. Department of Agriculture, Natural Resources Conservation Service.
- Soriano-Disla, J.M., Janik, L.J., Viscarra Rossel, R., 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. *Appl. Spectrosc. Rev.* 49, 139–186. <https://doi.org/10.1080/05704928.2013.811081>.
- Stevens, A., & Ramirez-Lopez, L. 2020. An introduction to the prospector package. R package Vignette R package version 0.2.6. <https://cran.r-project.org/web/packages/prospectr/vignettes/prospectr.html>.
- Summerrauer, L., Baumann, P., Ramirez-Lopez, L., Barthel, M., Bauters, M., Bukombe, B., Reichenbach, M., Boeckx, P., Kearsley, E., Van Oost, K., Vanlauwe, B., Chiragaga, D., Heri-Kazi, A.B., Moonen, P., Sila, A., Shepherd, K., Bazirake Mujinya, B., Van Ranst, E., Baert, G., Doetterl, S., Six, J., 2021. The central African soil spectral library: a new soil infrared repository and a geographical prediction analysis. *Soil* 7. <https://doi.org/10.5194/soil-7-693-2021>.
- Viscarra Rossel, R.A., Brus, D.J., Lobsey, C., Shi, Z., McLachlan, G., 2016. Baseline estimates of soil organic carbon by proximal sensing: comparing design-based, model-assisted and model-based inference. *Geoderma* 265. <https://doi.org/10.1016/j.geoderma.2015.11.016>.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the Tidyverse. *J. Open Source Software* 4. <https://doi.org/10.21105/joss.01686>.
- Wijewardane, N.K., Ge, Y., Sanderman, J., Ferguson, R., 2021. Fine grinding is needed to maintain the high accuracy of mid-infrared diffuse reflectance spectroscopy for soil property estimation. *Soil Sci. Soc. Am. J.* 85 <https://doi.org/10.1002/saj2.20194>.
- Wijewardane, N.K., Ge, Y., Wills, S., Libohova, Z., 2018. Predicting physical and chemical properties of US soils with a mid-infrared reflectance spectral library. *Soil Sci. Soc. Am. J.* 82, 722. <https://doi.org/10.2136/sssaj2017.10.0361>.