

Article

Using Mid-Infrared Spectroscopy to Optimize Throughput and Costs of Soil Organic Carbon and Nitrogen Estimates: An Assessment in Grassland Soils

Paulina B. Ramírez ^{1,*}, Samantha Mosier ², Francisco Calderón ¹  and M. Francesca Cotrufo ²

¹ Columbia Basin Agricultural Research Center, Oregon State University, Pendleton, OR 97801, USA

² Department of Soil and Crop Sciences & Natural Resources Ecology Laboratory, Colorado State University, Fort Collins, CO 80523, USA

* Correspondence: paulina.ramirezcisterna@oregonstate.edu

Abstract: Low-cost techniques, such as mid-infrared (MIR) spectroscopy, are increasingly necessary to detect soil organic carbon (SOC) and nitrogen (N) changes in rangelands following improved grazing management. Specifically, Adaptive Multi-Paddock (AMP) grazing is being implemented to restore grassland ecosystems and sequester SOC often for commercialization in C markets. To determine how the accuracy of SOC and N predictions using MIR spectroscopy is affected by the number of calibration samples and by different predictive models, we analyzed 1000 samples from grassland soils. We tested the effect of calibration sample size from 100 to 1000 samples, as well as the predictive ability of the partial least squares (PLS), random forest (RF) and support vector machine (SVM) algorithms on SOC and N predictions. The samples were obtained from five different farm pairs corresponding to AMP and Conventional Grazing (CG), covering a 0–50 cm soil depth profile along a latitudinal gradient in the Southeast USA. Overall, the sample size had only a moderate influence on these predictions. The predictive accuracy of all three models was less affected by variation in sample size when >400 samples were used. The predictive ability of non-linear models SVM and RF was similar to classical PLS. Additionally, all three models performed better for the deeper soil samples, i.e., from below the A horizon to the –50 cm depth. For topsoil samples, the particulate organic matter (POM) content also influenced the model accuracy. The selection of representative calibration samples efficiently reduces analysis costs without affecting the quality of results. Our study is an effort to improve the efficiency of SOC and N monitoring techniques.

Keywords: machine learning; multi-paddock grazing; particulate organic matter; rangeland monitoring; sample size



Citation: Ramírez, P.B.; Mosier, S.; Calderón, F.; Cotrufo, M.F. Using Mid-Infrared Spectroscopy to Optimize Throughput and Costs of Soil Organic Carbon and Nitrogen Estimates: An Assessment in Grassland Soils. *Environments* **2022**, *9*, 149. <https://doi.org/10.3390/environments9120149>

Academic Editors: Sutie Xu, Virginia L. Jin, Navreet Kaur Mahal and Jing Hu

Received: 26 October 2022

Accepted: 22 November 2022

Published: 25 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The ability to rapidly screen thousands of soil samples with a cost-effective and reproducible methodology is an extremely attractive prospect for improving soil carbon and health. Therefore, high-throughput technologies are needed as a low-cost alternative capable of accelerating screening and monitoring soil more effectively. Mid-infrared spectroscopy (MIR) as a tool for soil analysis is a promising alternative for scaling up conventional laboratory assays in vast areas with a high carbon sequestration potential in order to mitigate climate change impacts [1,2].

Grassland soils have a significant potential for soil organic carbon (SOC) sequestration at scale [1]. They also stock high amounts of nitrogen (N) per unit of SOC [3]. N availability may limit SOC sequestration in these soils depending on environmental conditions [4,5]. Over the years, several methods have been proposed to estimate a wide range of soil properties in grasslands and rangelands. Remote sensing, for example, has been widely applied to monitor temporal and spatial patterns [6,7]. Unfortunately, the use of remote technology to predict soil properties still has some limitations, resulting in inaccurate values

with substantial uncertainties when estimating SOC and N stocks [8,9]. This is mostly due to visible near-infrared (VNIR) and shortwave (SWIR) spectral ranges, which have a low spatial resolution and can be affected by atmospheric distortion leading to an extremely low signal-to-noise ratio [10]. On the other hand, MIR has demonstrated the potential to achieve high accuracy for C and N predictions even in soils formed with different parent materials [1,2]. However, the precision and performance of MIR at the project scale have not been tested to quantify SOC and N content under different land management practices.

Given the huge demand for these analyses to improve soil health and promote ecosystem service markets, rigorous testing of MIR approaches is needed to optimize throughput and the cost of quantifying SOC and N in response to changes in management practices. For example, this could help compare the SOC sequestration potential by the adoption of regenerative grazing management, such as Adaptive Multi-Paddock (AMP), with conventional grazing (CG) [11,12]. Despite the well-known potential of MIR spectroscopy to obtain highly accurate soil C and N information, research on sample size and adequate calibration set size has not received much attention. It has been challenging to get clear and consistent guidelines for the optimal calibration set size, given that predictive accuracy is also partly determined by sampling design (i.e., spatial scales, sampling depth, land use) as well as the choice of model algorithms [13,14].

Consequently, over the last few decades, chemometrics/machine learning tools have been increasingly applied to spectral data in order to maximize the models' predictive accuracy for estimating soil parameters. The predictive power of emerging machine learning algorithms can offer substantial gains in accuracy relative to conventional chemometric methods. Partial least squares (PLS) is a linear and commonly chemometric technique used to estimate different soil properties. However, PLS is sometimes less accurate when using non-linear data [15,16]. Therefore, greater attention to non-linear modeling techniques is being paid, as such methods offer greater flexibility over linear methods, owing to their ability to capture more complicated relationships between specific spectral reflectance signatures and soil properties [17].

Additionally, a strategy for selecting an optimal calibration sample size is critical in order to build models with a satisfactory predictive ability. This is mainly because larger datasets may not always reduce model uncertainty. While many samples are often required to obtain robust calibrations in order to detect changes in SOC or N stock following improvements in management [18], analyzing a large number of samples for C and N on an elemental analyzer slows throughput and increases costs. Furthermore, measurements can be limited due to low helium supply [19]. Therefore, this study aims to optimize calibration sample sizes by evaluating how SOC and N estimation accuracy is affected by the number of calibration samples using different predictive models. The specific objectives of this study are to: (1) identify the optimal conditions for building SOC and N calibrations (e.g., sample size, sampling locations) in grazed pasture soils, (2) evaluate different machine learning algorithms for SOC and N predictions, and (3) compare the cost-effectiveness of using an optimal calibration dataset without affecting the model's accuracy.

2. Materials and Methods

2.1. Study Sites and Soil Sampling

The study sites and sampling details have been previously presented by [12]. Briefly, sites represented a latitudinal gradient from Adolphus, Kentucky, through Woodville, Mississippi (Figure 1). The samples were collected in different soil series as defined in USDA Soil Taxonomy including Emory silt loam, Hartsell fine sandy loam, Loring silt loam, Trimble gravelly silt loam, and Cumberland gravelly/non-gravelly, ranging from loam to silty clay loam. The selection of the most representative five pairs of neighboring AMP and CG farms was based on the farms that most closely represented our definition of AMP grazing with a neighbor practicing CG, which is the most common and representative grazing. At each farm, 42 soil cores were sampled following the VM0021 "Soil Carbon Quantification Method" [20], for a total of 420 cores. At each farm pair, soil cores were distributed in

two representative catenae on a common soil type across paired farms, with three sampling zones (e.g., upper, middle, and lower slope) per catena (Supplemental Figure S3), and seven cores per sampling zone. Soil cores (1 m deep) were collected with a Giddings hydraulic probe mounted on an ATV using 5 cm in diameter sleeves, and further separated by depth in the laboratory. For this study we used the A horizon (0 to approx. 20 cm; topsoil) and the increment depths below A to 30 and 30–50 cm.

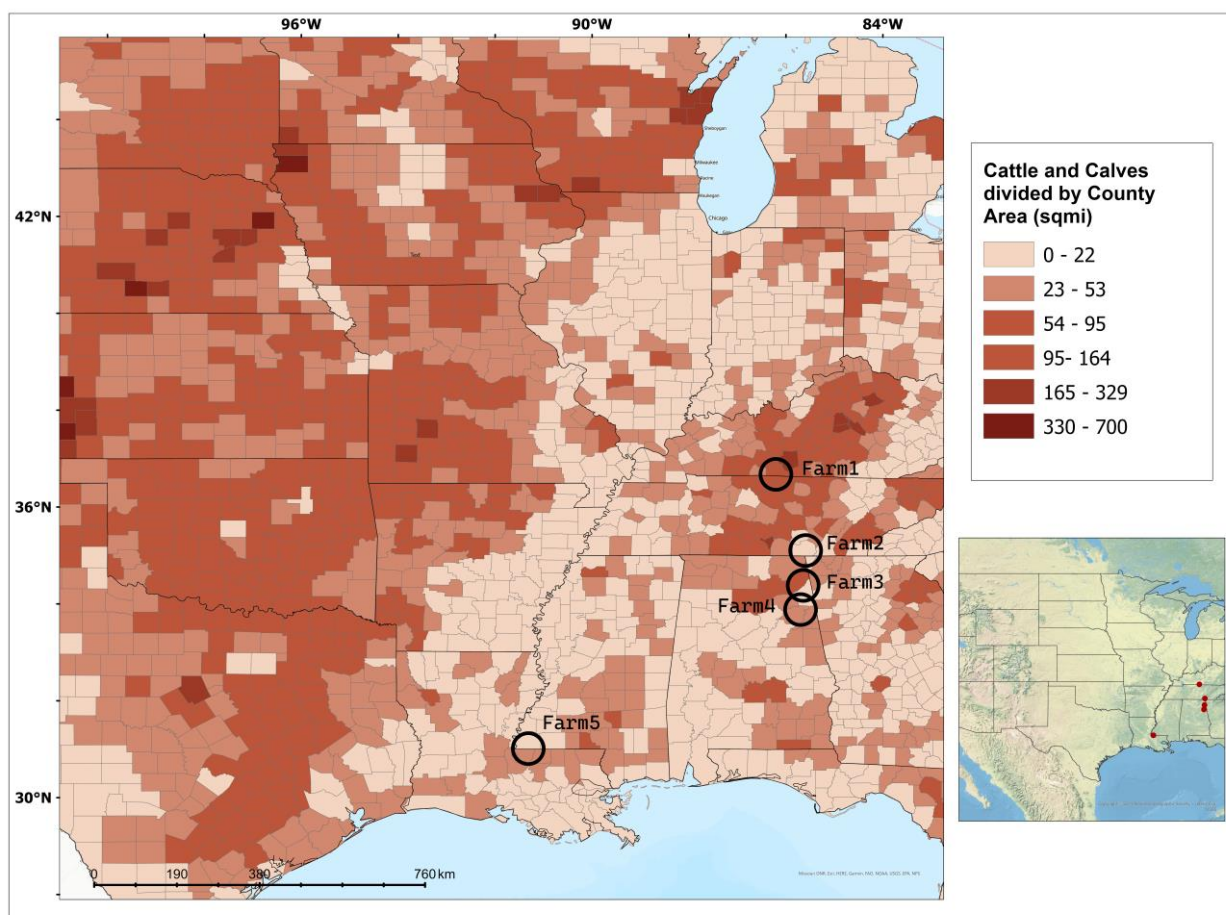


Figure 1. Study area showing the five pairs of adaptive and conventional grazing farms. Study sites are located in a latitudinal gradient from Adolphus, Kentucky, through Woodville, Mississippi. The map was elaborated in ArcGIS Pro using existing data sources USDA—NASS 2017, Census data.

2.2. Soil Analysis

Soil C and N concentrations were initially determined by dry combustion on a Costech ECS 4010 elemental analyzer (Costech Analytical Technologies, Valencia, CA, USA) on 2 mm sieved, finely ground and oven-dried soil samples as described in detail in [11]. In the few soils positive to the fizz test, soil inorganic C concentrations were quantified using an acid pressure transducer connected to a voltage meter [21] and subtracted from the total C concentration to determine SOC concentrations.

Soil organic matter physical fractionations were separated by size and density [22] on the 2 mm sieved samples composited by sampling zone and depth layer, to obtain a light particulate organic matter (POM), heavy POM, and a mineral associated organic matter (MAOM) as described in [12]. All fractions were analyzed for %C and %N on an elemental analyzer as described above for the bulk soils. For the purposes of this study, the light and heavy POM C values were used.

2.3. MIR Measurements

For spectral analysis, air-dried soil samples (<2 mm) were ground using a mortar and a pestle and subsequently analyzed using a Digilab FTS 7000 spectrometer (Varian, Inc., Palo Alto, CA, USA) with a Pike AutoDIFF diffuse reflectance autosampler (Pike Technologies, Madison, WI, USA) for spectral analysis. The MIR (4000–400 cm^{-1}) pseudo absorbance was obtained using a KBr background and deuterated triglycine sulfate detector. Each spectrum was made of 64 co-added scans and 4 cm^{-1} resolutions.

2.4. Sample Selection

A Kennard–Stone (KS) algorithm was used to split the whole dataset (1612 spectra) into a training set (1000 spectra), and a validation dataset of 612 spectra (Figure 2). The KS was performed to ensure two subsets that follow the statistical distribution of the original dataset [23]. From the training subset, ten sample sizes of an increasing number of calibration samples, ranging between 100 and 1000, were randomly obtained to avoid optimistically biased performance estimates. Five replicates of each sample size were generated using the R package dplyr [24].

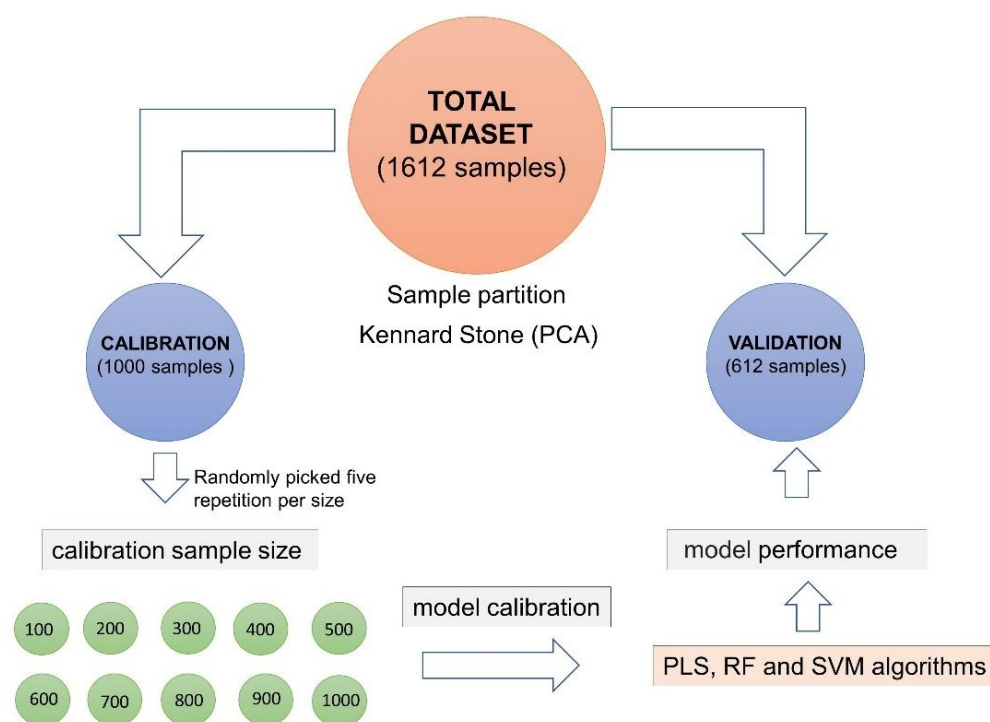


Figure 2. Flowchart of training and validation of the calibration models.

2.5. Model Calibration

Prior to applying the machine learning algorithms, the spectra were preprocessed using the Savitzky–Golay smoothing filter method. The partial least squares (PLS), random forest (RF), and support vector machine (SVM) models were trained comparatively to develop calibration models for predicting C and N content. The machine learning analyses were implemented using the Caret package in the R software [25], using the default method for optimizing hyperparameters. These algorithms were selected because they differ in their linear and non-linear functional capabilities. PLS is a linear regression model that can work efficiently with spectral data at a lower computational capacity [26]. While SVM and RF have non-linear fitting capabilities, they differ in computational demand [27,28].

2.6. Model Validation

To objectively reflect the models' predictive performance, the validation was carried out using an external dataset (610 samples), which was not involved in the model training process. Therefore, we assessed the model's accuracy as a function of the number of calibration samples while keeping the validation dataset constant. This type of validation is more rigorous in order to avoid over-optimistic calibrations that may be unable to cope with unknown samples [29]. Different metrics were computed to quantify the overall model performance; root mean square error (RMSE), coefficient of determination R-square (R^2), mean absolute error (MAE), and Nash–Sutcliffe efficiency coefficient (NSE). As regards these statistics, the best model should have the highest R^2 and NSE, and the lowest RMSE and MAE. The metrics RMSE and MAE are scale-dependent metrics with the same unit of measurement as the dependent variable, whereas NSE and R metrics are dimensionless metrics. Compared with MAE, the RMSE give a relatively high weight to large errors because the errors are squared before averaging. The sensitivity of the RMSE to outliers is the most common concern with the use of this metric; however, RMSE tends to become larger than MAE as the sample size increases [30,31]. The R^2 value close to 1 indicates that the predicted values may fit the measured data, whereas NSE shows how well the predicted data fit to the 1:1 line. When NSE = 1, it indicates a perfect match of the model to the measured data [32,33]. The metrics were calculated as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (1)$$

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (2)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (4)$$

where y_i and $\hat{y}_i$ correspond to the measured and predicted values, respectively; $\bar{y}_i$ is the mean value of the measured value and n is the number of measured data.

ANOVA was conducted to test statistical differences between medians of sample sizes. Subsequently, Fisher's Least Significant Differences test (LSD) was applied when the Kruskal–Wallis results had statistical significance. After testing the effect of sample size, we selected the smallest size that was able to provide more accurate estimates for SOC in order to evaluate the impact of sampling depths and farm soils on C estimates. Model residuals (residual = measured – predicted) were standardized by dividing them using standard deviation [34]. The positive and negative values of the standardized residuals also indicate whether the expected values might be over- or underestimated by the model. The effect of particulate organic matter (POM) on the model accuracy was evaluated by plotting the standardized residual corresponding to the C model against the heavy sand-sized and free light POM content.

3. Results

3.1. Descriptive Analysis of Soil Data

This dataset covers a broad range of SOC and N concentrations and climatic conditions (Table 1). In AMP grazing sites, SOC ranged from 1.28 ± 0.86 to 1.87 ± 1.36 and from 0.13 ± 0.09 to 0.21 ± 0.14 for N in the 0–20-cm depth increment. SOC and N concentrations were higher in farm pair 5 relative to the other in 0–20 cm depth increments. For these soil profiles (farm pair 5), the climate is slightly warmer and wetter with MAT of 19 °C and MAP of 1649 mm. For CG soils, lower values of SOC ranging 1.03 ± 0.63 and 1.4 ± 0.87

were found in the topsoil (0–20-cm) compared with AMP grazing. In the deepest soil depth increment (30–50 cm), SOC ranged from 0.26 ± 0.42 to 0.13 ± 0.07 , while the N concentrations differed slightly among sites.

Table 1. Soil and climate characteristics in the neighboring adaptive multi-paddock (AMP) and conventional (CG) grazed pairs included in this study, data are from [11]. Mean and standard deviation of soil organic carbon (SOC), nitrogen (N), and bulk density (BD) are shown for each depth increment. Climate variables correspond to mean annual precipitation (MAP) and mean annual temperature (MAT).

	Location	MAT (°C)	MAP (mm)	Grazing Practice	Year	n	Depth (cm)	SOC (%)	N (%)	BD (g/cm ³)
Farm 1	Adolphus, Kentucky	13.8	1316	AMP	13	40	0–20	1.38 ± 0.83	0.16 ± 0.09	1.11 ± 0.20
						17	20–30	0.29 ± 0.08	0.04 ± 0.01	1.22 ± 0.11
						20	30–50	0.24 ± 0.07	0.05 ± 0.01	1.12 ± 0.15
				CG	6	91	0–20	1.40 ± 0.87	0.16 ± 0.09	1.14 ± 0.14
						43	20–30	0.29 ± 0.09	0.04 ± 0.01	1.10 ± 0.26
						29	30–50	0.18 ± 0.04	0.04 ± 0.01	1.15 ± 0.12
Farm 2	Sequatchie, Tennessee	14.7	1432	AMP	12	96	0–20	1.51 ± 0.96	0.17 ± 0.09	1.34 ± 0.12
						37	20–30	0.47 ± 0.33	0.07 ± 0.02	1.45 ± 0.10
						44	30–50	0.29 ± 0.14	0.05 ± 0.01	1.53 ± 0.10
				CG	-	91	0–20	1.28 ± 0.77	0.15 ± 0.08	1.41 ± 0.16
						44	20–30	0.30 ± 0.14	0.05 ± 0.02	1.55 ± 0.10
						44	30–50	0.26 ± 0.42	0.05 ± 0.04	1.61 ± 0.10
Farm 3	Fort Payne, Alabama	15.1	1417	AMP	29	91	0–20	1.28 ± 0.86	0.13 ± 0.09	1.45 ± 0.19
						44	20–30	0.26 ± 0.08	0.03 ± 0.01	1.62 ± 0.1
						43	30–50	0.14 ± 0.04	0.02 ± 0	1.67 ± 0.05
				CG	17	78	0–20	1.03 ± 0.63	0.11 ± 0.06	1.52 ± 0.11
						39	20–30	0.23 ± 0.08	0.03 ± 0.01	1.67 ± 0.12
						33	30–50	0.13 ± 0.07	0.02 ± 0	1.74 ± 0.12
Farm 4	Piedmont, Alabama	15.7	1352	AMP	24	92	0–20	1.35 ± 0.97	0.15 ± 0.10	1.4 ± 0.15
						44	20–30	0.31 ± 0.16	0.04 ± 0.01	1.40 ± 0.10
						36	30–50	0.18 ± 0.11	0.03 ± 0.01	1.47 ± 0.12
				CG	-	88	0–20	1.28 ± 0.75	0.12 ± 0.07	1.31 ± 0.21
						40	20–30	0.34 ± 0.15	0.04 ± 0.01	1.45 ± 0.16
						31	30–50	0.25 ± 0.26	0.03 ± 0.02	1.50 ± 0.19
Farm 5	Woodville, Mississippi	19	1649	AMP	10	90	0–20	1.87 ± 1.36	0.21 ± 0.14	1.24 ± 0.19
						45	20–30	0.3 ± 0.13	0.05 ± 0.01	1.44 ± 0.08
						45	30–50	0.15 ± 0.05	0.03 ± 0.01	1.41 ± 0.34
				CG	38	89	0–20	1.46 ± 1.00	0.16 ± 0.10	1.22 ± 0.22
						45	20–30	0.27 ± 0.08	0.05 ± 0.01	1.42 ± 0.09
						43	30–50	0.15 ± 0.03	0.04 ± 0.01	1.47 ± 0.22

This high soil variability along the latitudinal gradient (Figure 1) encompassed soil-specific spectra characteristics in each farm pair (Supplementary Figure S1). The PCA score plots of the MIR spectra generated a clear separation of the five farm pairs into clusters, indicating high heterogeneity in the soil composition of the spectral database (i.e., organic matter, iron oxides and clay), while there were spectral similarities between CG and AMP soils in each pair. Pair 3 appears to show a greater separation than the other 4 farm pairs.

The calibration and validation datasets of the measured SOC and N concentration followed a right-skewed distribution (Supplementary Figure S2). In fact, most of the

SOC values are in the 0.01–1% range, while the highest values are at 5%. Similarly, most values ranged from 0.01 to 0.1% for N concentration, with the highest values at 0.5%. As expected, the sample distributions of the validation and calibration datasets split by the Kennard-Stone algorithm were the same.

3.2. Model Comparison and Influence of Training Sample Size

The performance of the PLS, SVM, and RF models for predicting SOC (Figure 3) and N (Figure 4) values improved rapidly when the training dataset was increased to 400 samples. The standard deviation of all the different sample sizes was higher in a small-sized set (100–200 samples), both for SOC and N. We can observe a rapid decrease in RMSE and MAE in the 100 to 400 sample range. RMSE and MAE have similar values for all datasets, with RMSE sometimes being slightly larger. The prediction accuracy did not improve with over 400 samples for any model. A similar pattern of results was observed when the three different models were used to predict N (Figure 4). In addition, all models underestimated SOC when concentrations were above 2.0% (Figure 5). The same behavior was observed for N concentrations above 0.20% (Figure 6). The effect of sample size differed among models; for example, according to RMSE, R-square, and Nash–Sutcliffe criteria, the PLS model performed better than the RF and SVM models when sample size increased (Figures 3 and 4). The RF model yielded more accurate C and N predictions with sample sizes in the range of 400–600. However, when the sample size used for calibration was higher ($n = 1000$), the prediction was more scattered, tending to strongly overestimate both SOC and N (Figures 5 and 6).

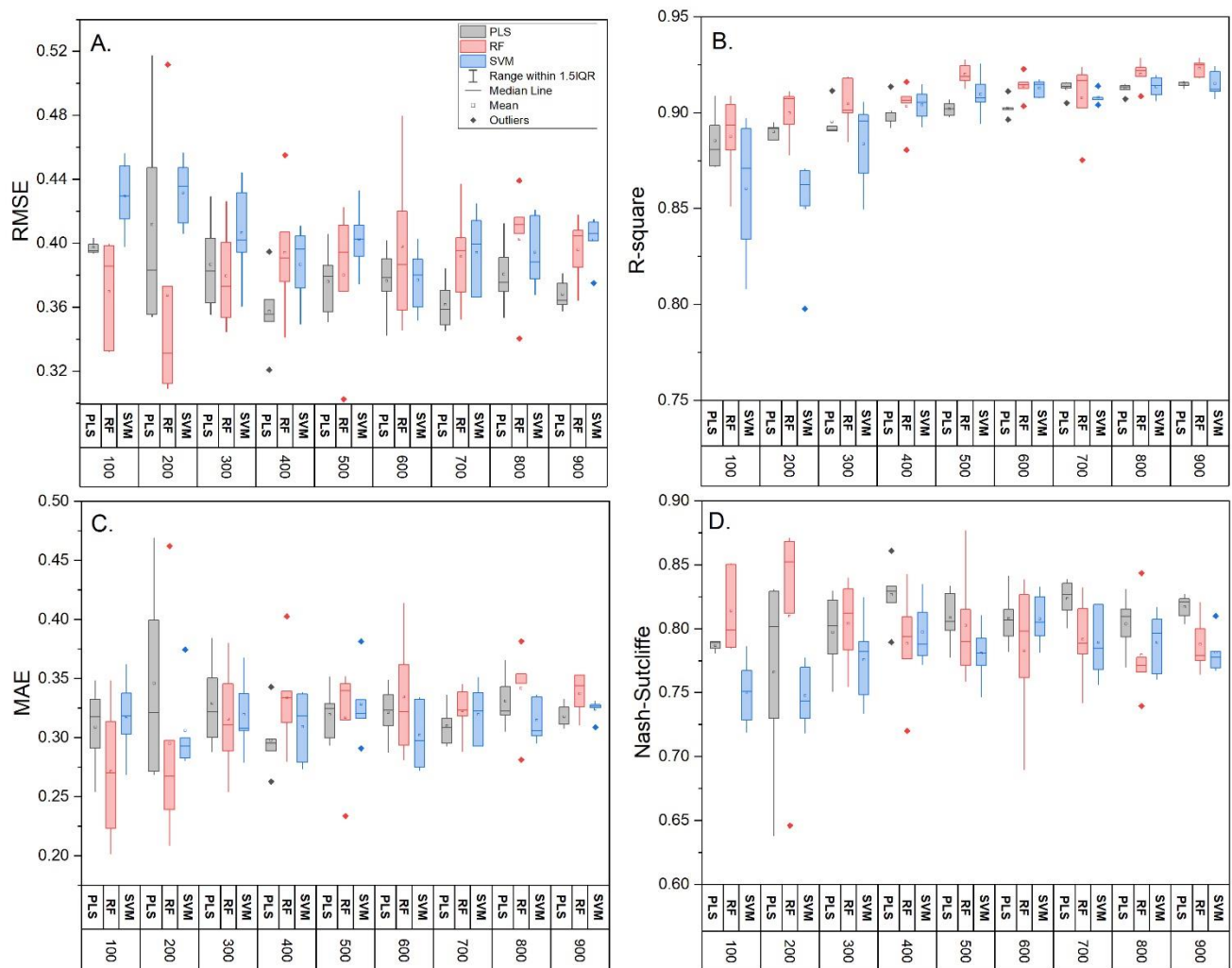


Figure 3. Box-plots graph showing the ability performance of partial least squares (PLS), random forest (RF) and support vector machine (SVM) models for predicting soil organic carbon (SOC) content using various calibration sampling size (100–900 samples). Each boxplot represents the results for the five random repetitions for each sample sizes. The performance metrics include (A) coefficient of determination root mean square error (RMSE), (B) coefficient of determination (R-square), (C) mean absolute error (MAE), and (D) Nash–Sutcliffe.

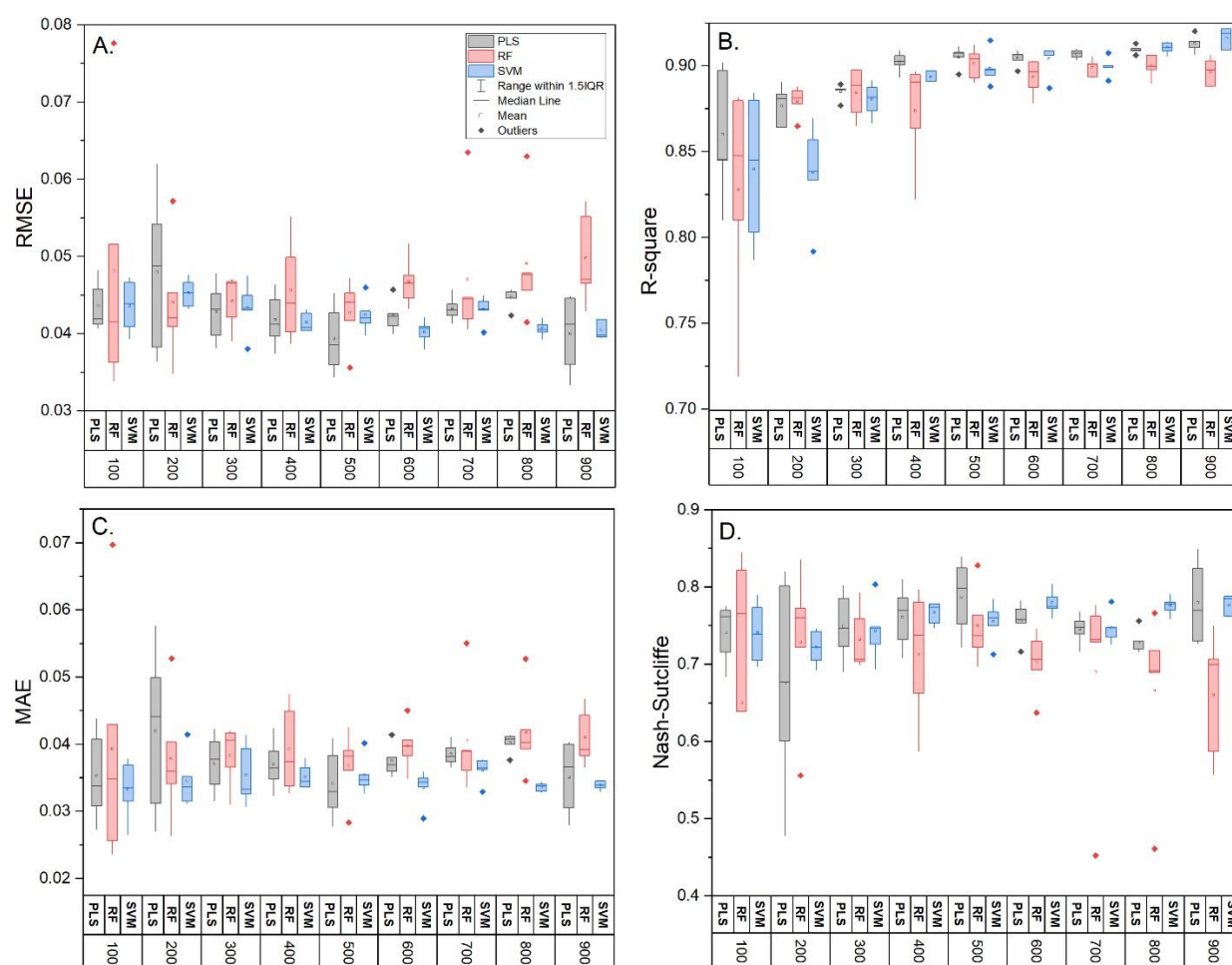


Figure 4. Box-plots graph showing the ability performance of partial least squares (PLS), random forest (RF) and support vector machine (SVM) models in for predicting soil total nitrogen content using various calibration sampling size (100–900 samples). Each boxplot represents the results for the five random repetitions for each sample sizes. The performance metrics include (A) coefficient of determination root mean square error (RMSE), (B) coefficient of determination (R-square), (C) mean absolute error (MAE), and (D) Nash–Sutcliffe.

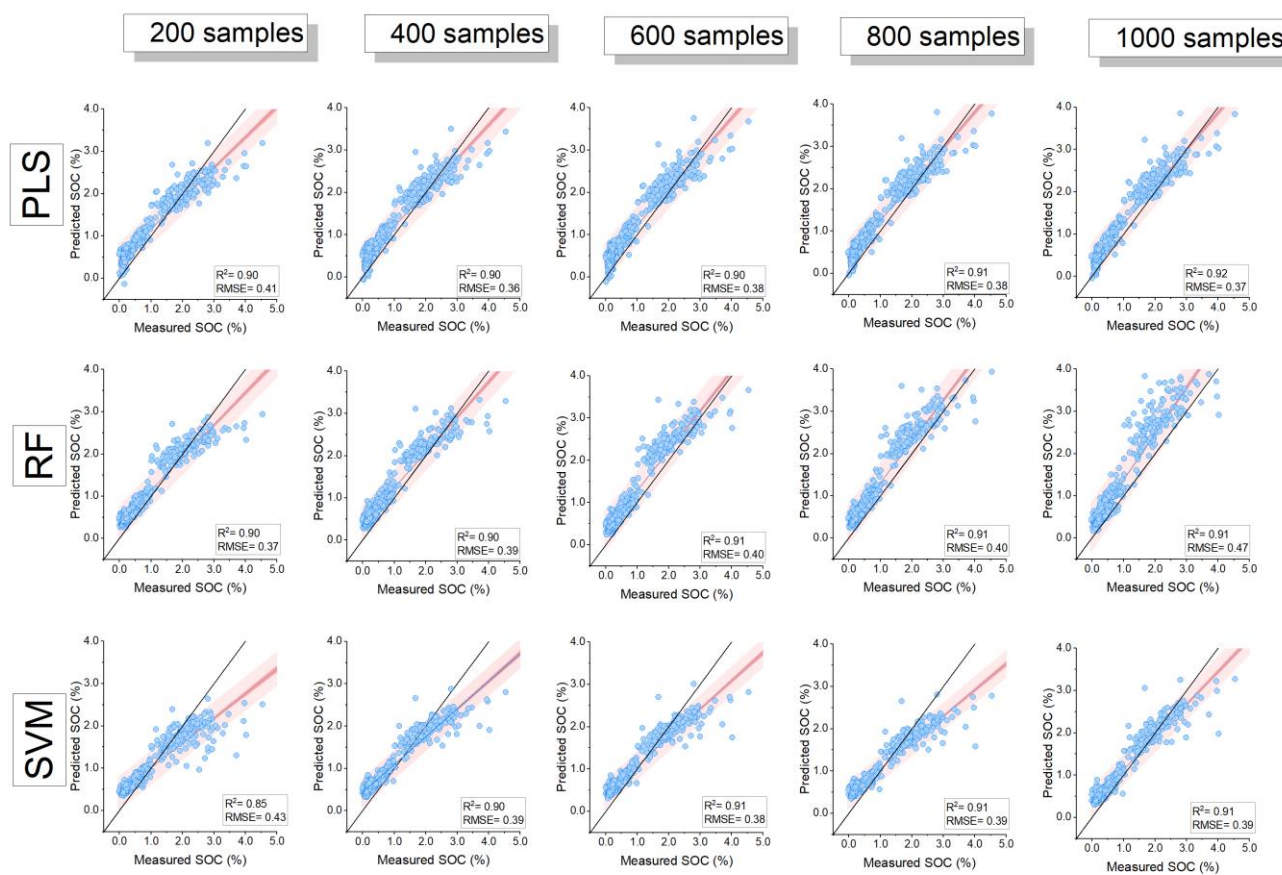


Figure 5. Comparison between the soil organic carbon (SOC) values measured by elemental analyses and the MID predicted values using the three different training algorithms, partial least squares (PLS), random forest (RF) and support vector machine (SVM). The validation included and external test set comprised of 612 observations. For each sample size, the graphs represent an average of the five calibrations randomly extracted from the total dataset ($n = 1000$ samples).

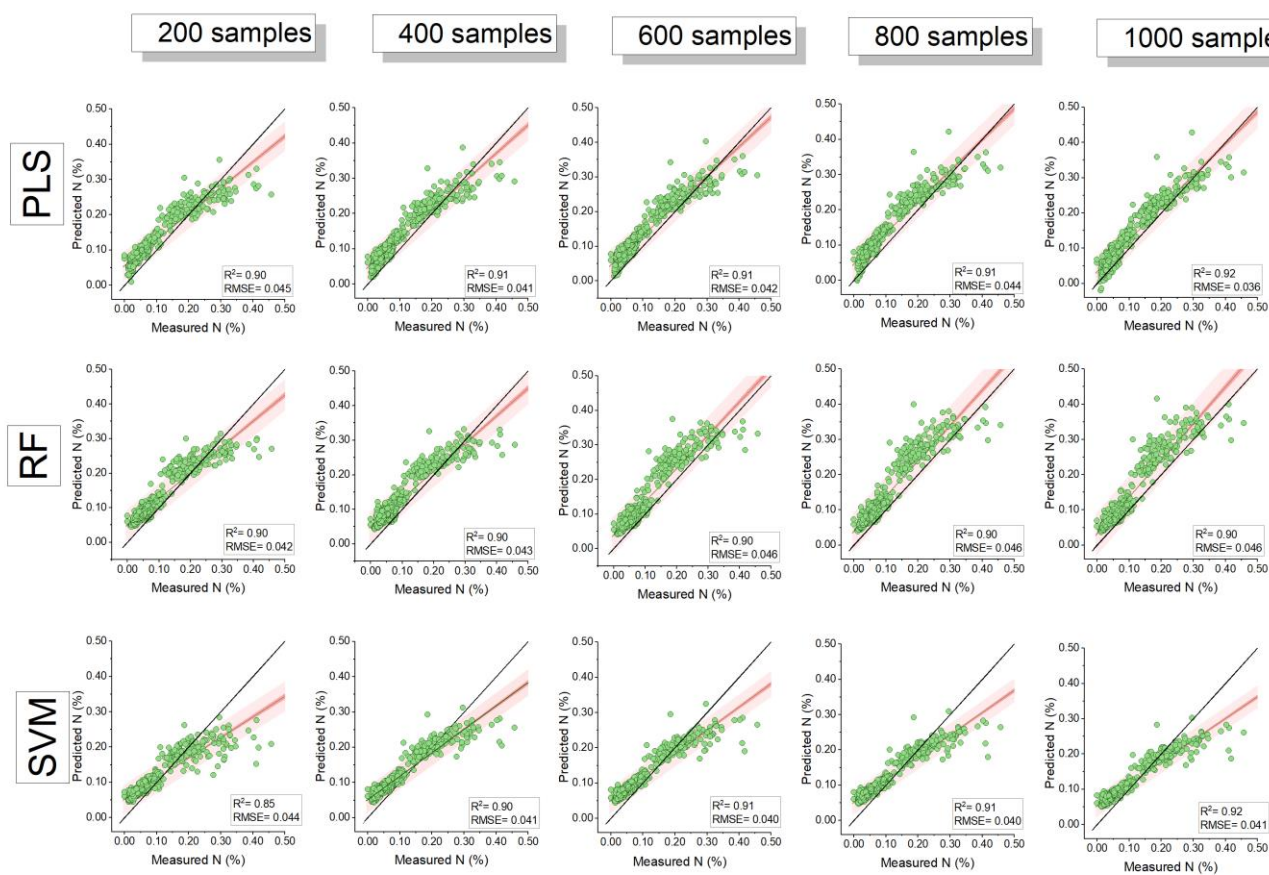


Figure 6. Comparison between the soil nitrogen (N) values measured by elemental analyses and the MID predicted values using the three different training algorithms, partial least squares (PLS), random forest (RF) and support vector machine (SVM). The validation included and external test set comprised of 612 observations. For each sample size, the graphs represent an average of the five calibrations randomly extracted from the total dataset ($n = 1000$ samples).

Since the performance of different modeling approaches plateaued at approximately 400 samples, we tested the models calibrated using 400 samples as our “optimum” for analyzing residuals. To test the statistical significance as the sample size increased from 100 to 400, the metrics values of the three statistical approaches were evaluated with pairwise comparison (p -value < 0.05) (Table 2). PLS, RF, and SVM models performed similarly when the sample size was approximately 400 samples for SOC and N. However, the predictive ability of PLS and SVM algorithms improved significantly, increasing the number of calibration samples from 100 to 400. With 400 samples, the mean values of RMSE were 0.35 ± 0.02 , 0.39 ± 0.04 and 0.39 ± 0.02 for SOC, as well as 0.042 ± 0.004 , 0.046 ± 0.007 , 0.041 ± 0.001 for N, using PLS, RF, and SVM, respectively. The values of the R-square of the three algorithms were 0.90, representing a good fit between the predicted and measured values. Nash–Sutcliffe coefficients exceeded 0.8 for PLS at 400 samples, tending to be less likely to over-underestimate predicted values of SOC and N.

Table 2. Mean and standard deviation of performance metric in terms of average root mean square error (RMSE), coefficient of determination (R-square), mean annual error (MAE) and Nash–Sutcliffe coefficient as a function of sample size using partial least squares (PLS), random forest (RF) and support vector machine (SVM) models based on five replicates randomly selected of each sample size. For each method, five calibrations were built for each sample size. The entire dataset included 1000 samples.

Sample Size	Model	Soil Organic Carbon (%)				Total Soil Nitrogen (%)			
		RMSE	R-Square	MAE	Nash–Sutcliffe	RMSE	R-Square	MAE	Nash–Sutcliffe
100 (n = 5)	PLS	0.40 ± 0 ^a	0.89 ± 0.01 ^a	0.31 ± 0.03 ^a	0.79 ± 0 ^a	0.044 ± 0.003 ^a	0.86 ± 0.03 ^a	0.035 ± 0.007 ^a	0.74 ± 0.04 ^a
	RF	0.37 ± 0.03 ^a	0.89 ± 0.02 ^a	0.27 ± 0.05 ^a	0.81 ± 0.03 ^a	0.048 ± 0.018 ^a	0.83 ± 0.06 ^a	0.039 ± 0.019 ^a	0.65 ± 0.25 ^a
	SVM	0.43 ± 0.02 ^a	0.86 ± 0.03 ^a	0.32 ± 0.03 ^a	0.75 ± 0.02 ^a	0.044 ± 0.003 ^a	0.84 ± 0.04 ^a	0.033 ± 0.005 ^a	0.74 ± 0.04 ^a
400 (n = 5)	PLS	0.36 ± 0.02 ^b	0.90 ± 0.01 ^a	0.30 ± 0.03 ^a	0.83 ± 0.02 ^b	0.042 ± 0.004 ^a	0.90 ± 0.01 ^b	0.037 ± 0.004 ^a	0.76 ± 0.04 ^a
	RF	0.39 ± 0.04 ^a	0.90 ± 0.01 ^a	0.33 ± 0.04 ^a	0.79 ± 0.04 ^a	0.046 ± 0.007 ^a	0.87 ± 0.03 ^a	0.039 ± 0.007 ^a	0.71 ± 0.08 ^a
	SVM	0.39 ± 0.02 ^b	0.90 ± 0.01 ^b	0.31 ± 0.03 ^a	0.80 ± 0.02 ^b	0.041 ± 0.001 ^a	0.89 ± 0.00 ^b	0.035 ± 0.002 ^a	0.77 ± 0.01 ^a
1000 (n = 1)	PLS	0.37	0.92	0.32	0.82	0.036	0.92	0.030	0.83
	RF	0.47	0.90	0.37	0.70	0.046	0.90	0.038	0.71
	SVM	0.39	0.91	0.31	0.79	0.041	0.92	0.034	0.77

Different letters within each column indicate significant differences ($p < 0.05$) among model performance metric as a function of sample size in each model.

3.3. Influence of Sampling Depth and Farm Site on Model Accuracy

The plateau of the performance metrics was shown in Figures 3 and 4, where the calibrations exhibit only slight changes for sample sizes larger than 400. Each residual obtained by model calibrations using 400 samples was plotted to assess model fit at different depths (Figure 7). Since the training algorithms retained comparable predictive abilities for SOC and N, the following results were based only on C values. The models satisfactorily predicted SOC at 10–50 cm depth. However, high positive residuals were observed in the upper A horizon (0–10 cm), which had higher and more variable SOC concentrations (Table 1). This suggests that all models underestimated C content in topsoils with high SOC. In the upper A horizon, the SVM algorithm outperformed other algorithms. However, SVM could not improve the accuracy of the predicted values for SOC concentrations between 0.5–1.0%, especially those at lower (10–50 cm) depths. It is noteworthy that grazing practices and farm sites exhibited a moderate effect on SOC prediction results (Supplementary Figure S3).

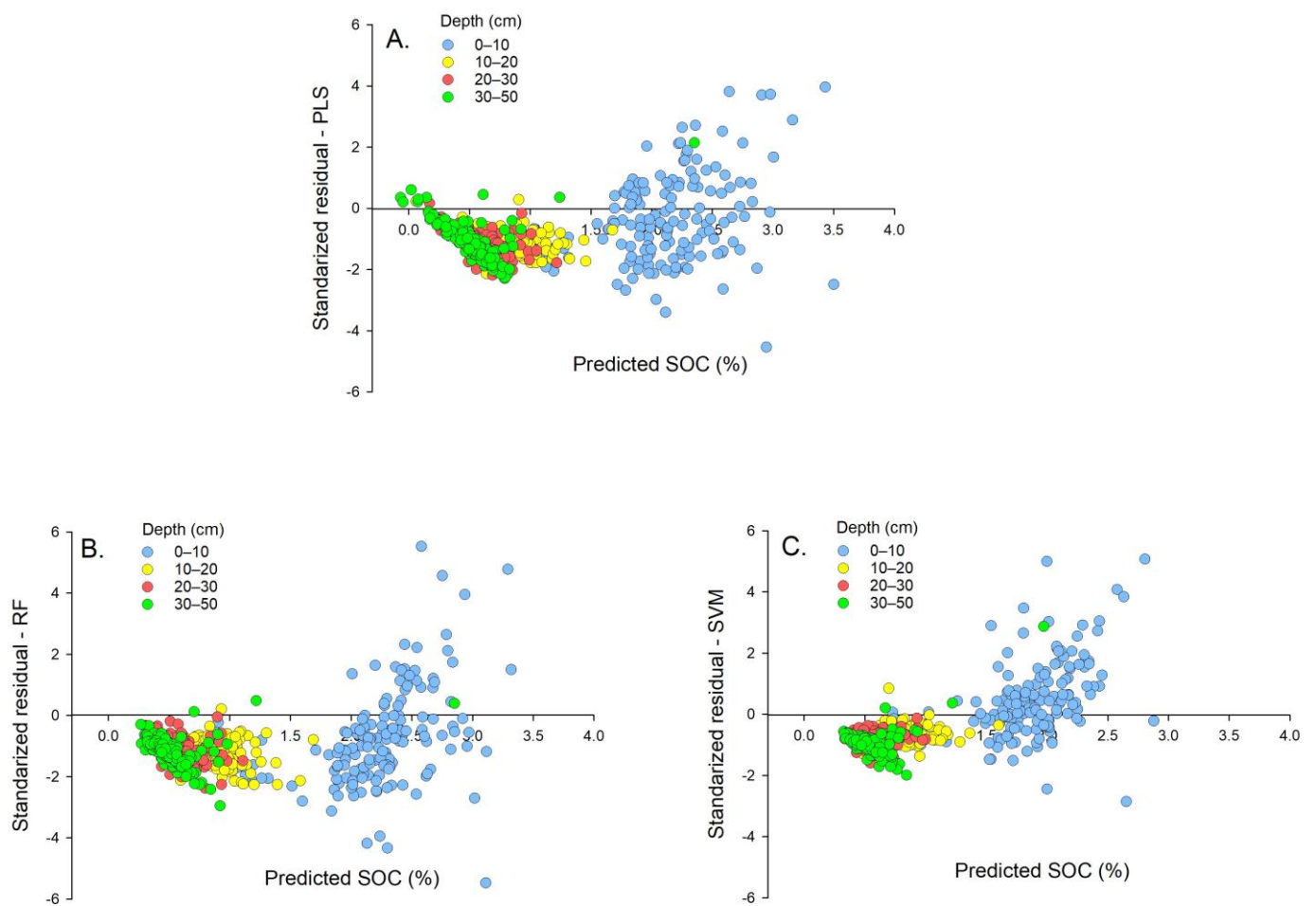


Figure 7. Effect of sampling depth on soil organic carbon (SOC) prediction using (A) partial least squares (PLS), (B) random forest (RF), and (C) support vector machine (SVM) models. The training dataset contains 400 samples. All standardized residuals (residual = measured – predicted) were then plotted against their predicted C content. If the calibration model is appropriate, the residuals should be distributed randomly around the $y = 0$ line.

Because of the tendency of all models to underestimate SOC content in the upper A horizon, model residuals for SOC content were analyzed against heavy sand-sized OM (heavy POM) and free light POM (Figure 8), since POM is typically a SOC fraction that increases with high SOC values [3]. The results indicate that SOC prediction accuracy decreases as POM increases, both in heavy and light fractions, despite having somewhat different dynamics (Figure 8). These relationships were characterized by linear ($R^2 = 0.35$) and exponential growth ($R^2 = 0.21$) behavior in heavy POM and free light POM, respectively.

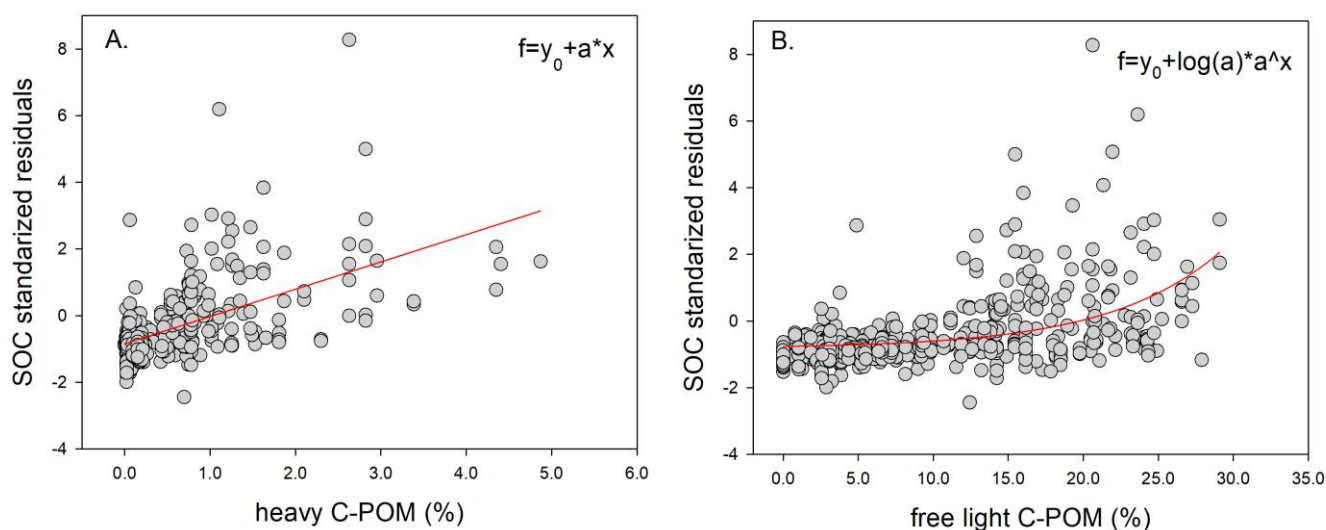


Figure 8. Effect of C concentration stored in particulate organic matter (POM) on under- or over-estimation of SOC by the models. Plots represent standardized residual (residual = measured – predicted) relative to heavy sand-sized OM (heavy POM) (A) and free light POM (B). Linear and exponential growth (modified simple exponent, 2 parameters) regressions were used to fit patterns for heavy POM and free light POM, respectively.

3.4. Cost Analysis

To evaluate the cost and time advantage of using MIR to estimate SOC and N in projects with large quantities of samples (e.g., over 400), MIR spectroscopy (PerkinElmer, Spectrum 3 FT-IR spectrometer), and dry combustion (Costech ECS CHNSO elemental analyzer), techniques were compared in terms of equipment, maintenance, consumables, and technician costs. The cost analysis assumes both instruments require the same sample preparation (i.e., sieving, oven-dried finely ground). Table 3 shows that adopting MIR technology does appear to be cost and time effective. The use of spectroscopy could increase throughput time from 4 to 12 samples per hour, whereas decrease cost by 2.5 times per sample. Although both methodologies have a similar instrument cost, the cost per sample was lower for MIR. In the dry combustion method, consumables associated with carrier gas (helium), purge gas (oxygen), and other supplies increase the cost. The labor cost per sample using dry combustion and MIR methods is USD 4.38/sample and USD 1.46/sample, respectively (labor costs were estimated at the rate of USD 17.5/h of labor). For this number of samples ($n = 400$), the total dry combustion cost is nearly USD 7000, while it is only USD 2000 for spectroscopy. Likewise, annual maintenance costs associated with using MIR are USD 3000, whereas the costs for the elemental analyzer are USD 1700. Using an optimal dataset to perform robust calibration from a large data pool might save approximately USD 4200 and 180 h (7.5 days) of technician time.

Table 3. Cost comparison analysis (USD) for measuring 400 and 1000 samples using FT-IR spectrometer and combustion analyzer.

	One-Time Cost (USD)	Yearly Cost (USD)	Data Acquisition Cost (USD)				Cost CN Analysis (USD)	
			Sample Preparation		Technician		n = 400	n = 1000
Method	Instrument ^(a)	Maintenance	Equipment	Lab. Supplies (<2 mm) ^(d)	Cost (Labor/h) (USD)	Time (Sam- ple/h)		
FT-IR Spectrometer	50,000	3500	9000 ^(c)	0.17	17.5	12.0 ^(e)	768.0	1920.0
Dry combustion analyzer ^(a)	80,000	1700 ^(b)	N/A	1.70	17.5	4.0 ^(f)	2780.0	6950.0

^(a) PerkinElmer Spectrum™ 3 FT-IR spectrometer + Autodiff II to automate multiple samples analysis; CHNSO Costech analyzer. ^(b) Maintenance costs are factored into the \$1.70 cost per sample in a total of 1000 samples. ^(c) Mixer mill with metal grinding balls. ^(d) Consumables and other supplies. ^(e) Time (hours) used to place the samples in the metal cups + analysis time. ^(f) Time (hours) used to place the samples in tin foil cups + analysis time.

4. Discussion

We investigated the effect of calibration set size on SOC and soil N predictions in AMP and CG grazing systems across a latitudinal gradient in the Southeastern U.S. Additionally, the sample size was evaluated using three different models of increasing complexity (PLS > SVM > RF). Our study demonstrates that MIR spectroscopy could be used in large-scale grassland SOC and soil health projects as an alternative for obtaining reliable SOC and soil N estimates with higher throughput and lower costs than elemental analyses using dry combustion. We started with a highly variable set of grassland soils in terms of their MIR spectra, SOC, and N values as well as grazing management in order to include representative conditions for soil analyses in large-scale regional soil grassland projects.

In general, we observed a moderate effect of sample size on the ability to predict SOC and N in grassland soils. As sample size increased, calibration only slightly improved the model's predictive ability, which plateaued around 400 samples. In fact, our results also indicated that there would not be any advantage in increasing the calibration sample size beyond 400 samples. Thus, for example, large-scale grassland projects with over 400 soil samples could analyze 400 samples for C and N using dry combustion in an elemental analyzer, scan all the samples using MIR spectroscopy, and estimate the remaining unknown concentrations with the model predictions. Furthermore, we observed that a larger sample size often resulted in worse predictions. This occurred when the RF model was trained using 800–1000 samples. By contrast, when the sample size was too small (<200 samples) to train the models, the validations were biased toward producing over-estimates.

Our results indicate that all the models underestimated SOC and N content over 2.0% and 0.20%, respectively. Contrary to our expectations, these trends were not improved by applying the non-linear RF and SVM models. Because our dataset was strongly skewed toward low SOC and N concentrations, our suggestion for optimal calibration sample size values does not apply to other datasets with significantly different distributions. Therefore, generalizations could be highly inaccurate for other datasets since we must also take into consideration the distribution of calibration and validation datasets, as well as the range we want to predict more accurately. Similarly, in previous studies [35,36], accuracy has been shown to improve with a higher number of training samples. However, few studies have investigated the impact of calibration sample size on SOC and N.

The PLS model had the smallest median RMSE and a lower underestimation of SOC content in the entire range of calibration sample sizes when compared with RF and SVM. Hence, we did not find any advantage of applying non-linear algorithms such as RF and SVM to our dataset. Therefore, PLS was particularly useful for achieving higher accuracy with less computational power. These results are consistent with previous studies [35],

where a smaller sample size is required in order to yield useful predictions using PLS instead of more complex algorithms. In addition, RF had a notably lower performance, even though we expected it to substantially improve its predictive ability when the sample size was increased. It has been suggested that the RF model is able to deal with many predictors and few samples, resulting in low overfitting [37]. Therefore, the RF training algorithms in high-dimensional datasets such as ours (800–1000 samples) might need more trees and a higher degree of tree depth, which could be a serious problem as the observed variables increase. In contrast, the predictive ability of SVM was distinctly lower when smaller sample sizes were used (100–300 samples). These observations are similar to those reported in other studies [17,38], where PLS outperformed SVM.

We have also investigated other factors influencing the tendency to overestimate measured values when calibrations ($n = 400$) were used. This tendency is not surprising, considering that the relatively poorly predicted C content above 2% at the 0–10 cm depth was influenced mainly by the POM content above this threshold. There was, however, no such effect on accuracy when farm pair location and grazing practices were evaluated. The full potential of MIR to predict soil properties is not always satisfactorily achieved if some of the unknowns' compositional or analytical profile is substantially different from that of the samples in the calibration set [39]. The light component of POM—which mainly consists of partially decomposed plant material with different chemical characteristics—might affect the performance of these models, resulting in less reliable estimates at <10 cm depth [40]. Therefore, the topsoil samples collected play a critical role, while the selection of algorithm and sample size might be less important. In this case, when analyzing shallow samples with high POM concentrations the use of elemental analysis by dry combustion may be more suitable than spectroscopy for C estimates. We also recommend additional work to train MIR to specifically quantify SOC content in POM [41,42].

A rough cost estimate shows that measuring SOC and N using MIR is significantly less expensive than dry combustion techniques. As other groups have shown for these approaches [43,44], MIR spectroscopy offers a cost-effective alternative to conventional methods for C and N estimates. For the MIR method, the highest SOC and N concentrations tended to be under-estimated; however, we believe it will not adversely affect the reliable assessment of these parameters, while it would clearly reduce the costs. In our case, there is a tolerable margin of error for predictions, which depends on the data and application. For instance, for the purposes of restoring soil, the error in relatively organic-poor samples is significantly diminished when compared to the undisturbed ones.

5. Conclusions

We have defined an optimal calibration set size for SOC and soil N predictions for grazing systems using models of increasing complexity coupled with MIR spectroscopy. Overall, we demonstrated a high predictive performance of PLS, RF and SVM models using datasets with about 400 samples. We found that large sample sets did not necessarily improve the accuracy of the three training algorithms. Moreover, an optimal dataset for one algorithm is not always optimal for other machine learning models. Consequently, there is no cut-off criteria for choosing the ideal sample size, since prediction accuracy depends significantly on each dataset's sample distribution, training algorithm, and sampling depth. The spectroscopic method was highly accurate for SOC and N estimates; however all models overestimated the highest values of SOC and N content. In addition, the non-linear models were not able to improve upon the classical PLS performance. Our results also suggest that POM will likely introduce uncertainty into the models. Lastly, MIR spectroscopy provides a cost-effective alternative to conventional SOC and N combustion analysis. This study offers insight into MIR methodology as an efficient, scalable, and cost-effective tool that provides reliable C and N estimates for grazing soils.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/environments9120149/s1>, Figure S1: Principal component analysis

(PCA) score plot showing farm pairs (1 to 5) with high spatial spectral heterogeneity between soils. Axis 1 accounts 42% of variance in the data set and Axis 2 accounts for 25%; Figure S2: Frequency histograms for calibration (1000 samples) and validation dataset (612 samples) dataset for A) soil organic carbon (SOC) and (B) total soil nitrogen (N). The predictions were achieved by keeping the same frequency distribution between the calibration and validation dataset using a Kennard-stone sampling algorithm; Figure S3: Effect of sampling grazing practice and farm-pair soil on soil organic carbon (SOC) prediction using partial least squared (PLS). The training dataset contains 400 samples. All standardized residuals (residual = measured – predicted) were then plotted against their predicted C content. If the calibration model is appropriate, the residuals should be distributed randomly around the $y = 0$ line.

Author Contributions: Conceptualization, M.F.C. and S.M.; soil sampling methodology, M.F.C., S.M., F.C., P.B.R.; software, P.B.R.; validation, F.C., P.B.R.; MID IR methodology and formal analysis, P.B.R.; investigation, S.M. and M.F.C.; resources, M.F.C. and F.C.; data curation, S.M.; writing—original draft preparation, P.B.R.; writing—review & editing, M.F.C. and P.B.R.; visualization, P.B.R.; supervision, M.F.C. and F.C.; project administration, M.F.C.; funding A-acquisition, M.F.C. All authors have read and agreed to the published version of the manuscript.

Funding: The soil data were obtained from a project supported by the Foundation for Food and Agricultural Research grant number 514752. Carbon center (CBARC) # 207411120004003S.

Data Availability Statement: Not applicable.

Conflicts of Interest: P.B.R. and F.C. declare no conflict of interest. M.F.C. and S.M. are cofounders of Cquester Analytics LLC, a service facility which offers both EA and MID analyses of soils.

References

1. Sanderman, J.; Savage, K.; Dangal, S.R.S. Mid-Infrared Spectroscopy for Prediction of Soil Health Indicators in the United States. *Soil Sci. Soc. Am. J.* **2020**, *84*, 251–261. [\[CrossRef\]](#)
2. Seybold, C.A.; Ferguson, R.; Wysocki, D.; Bailey, S.; Anderson, J.; Nester, B.; Schoeneberger, P.; Wills, S.; Libohova, Z.; Hoover, D.; et al. Application of Mid-Infrared Spectroscopy in Soil Survey. *Soil Sci. Soc. Am. J.* **2019**, *83*, 1746–1759. [\[CrossRef\]](#)
3. Cotrufo, M.F.; Ranalli, M.G.; Haddix, M.L.; Six, J.; Lugato, E. Soil Carbon Storage Informed by Particulate and Mineral-Associated Organic Matter. *Nat. Geosci.* **2019**, *12*, 989–994. [\[CrossRef\]](#)
4. Keller, A.B.; Borer, E.T.; Collins, S.L.; DeLancey, L.C.; Fay, P.A.; Hofmockel, K.S.; Leakey, A.D.B.; Mayes, M.A.; Seabloom, E.W.; Walter, C.A.; et al. Soil Carbon Stocks in Temperate Grasslands Differ Strongly across Sites but Are Insensitive to Decade-Long Fertilization. *Glob. Chang. Biol.* **2022**, *28*, 1659–1677. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Rocci, K.S.; Barker, K.S.; Seabloom, E.W.; Borer, E.T.; Hobbie, S.E.; Bakker, J.D.; MacDougall, A.S.; McCulley, R.L.; Moore, J.L.; Raynaud, X.; et al. Impacts of Nutrient Addition on Soil Carbon and Nitrogen Stoichiometry and Stability in Globally-Distributed Grasslands. *Biogeochemistry* **2022**, *159*, 353–370. [\[CrossRef\]](#)
6. Reinermann, S.; Asam, S.; Kuenzer, C. Remote Sensing of Grassland Production and Management—A Review. *Remote Sens.* **2020**, *12*, 1949. [\[CrossRef\]](#)
7. Jones, M.O.; Naugle, D.E.; Twidwell, D.; Uden, D.R.; Maestas, J.D.; Allred, B.W. Beyond Inventories: Emergence of a New Era in Rangeland Monitoring. *Rangel. Ecol. Manag.* **2020**, *73*, 577–583. [\[CrossRef\]](#)
8. Asner, G.P.; Wessman, C.A.; Bateson, C.A.; Privette, J.L. Impact of Tissue, Canopy, and Landscape Factors on the Hyperspectral Reflectance Variability of Arid Ecosystems. *Remote Sens. Environ.* **2000**, *74*, 69–84. [\[CrossRef\]](#)
9. Angelopoulou, T.; Tziolas, N.; Balafoutis, A.; Zalidis, G.; Bochtis, D. Remote Sensing Techniques for Soil Organic Carbon Estimation: A Review. *Remote Sens.* **2019**, *11*, 676. [\[CrossRef\]](#)
10. Chabrillat, S.; Ben-Dor, E.; Cierniewski, J.; Gomez, C.; Schmid, T.; van Wesemael, B. Imaging Spectroscopy for Soil Mapping and Monitoring. *Surv. Geophys.* **2019**, *40*, 361–399. [\[CrossRef\]](#)
11. Stanley, P.L.; Rowntree, J.E.; Beede, D.K.; DeLonge, M.S.; Hamm, M.W. Impacts of Soil Carbon Sequestration on Life Cycle Greenhouse Gas Emissions in Midwestern USA Beef Finishing Systems. *Agric. Syst.* **2018**, *162*, 249–258. [\[CrossRef\]](#)
12. Mosier, S.; Apfelbaum, S.; Byck, P.; Calderon, F.; Teague, R.; Thompson, R.; Cotrufo, M.F. Adaptive Multi-Paddock Grazing Enhances Soil Carbon and Nitrogen Stocks and Stabilization through Mineral Association in Southeastern U.S. Grazing Lands. *J. Environ. Manag.* **2021**, *288*, 112409. [\[CrossRef\]](#)
13. Ng, W.; Minasny, B.; Malone, B.; Filippi, P. In Search of an Optimum Sampling Algorithm for Prediction of Soil Properties from Infrared Spectra. *PeerJ* **2018**, *2018*, e5722. [\[CrossRef\]](#)
14. Zhang, Y.; Saurette, D.D.; Easher, T.H.; Ji, W.; Adamchuck, V.I.; Biswas, A. Comparison of Sampling Designs for Calibrating Digital Soil Maps at Multiple Depths. *Pedosphere* **2022**, *32*, 588–601. [\[CrossRef\]](#)

15. Araújo, S.R.; Wetterlind, J.; Demattê, J.A.M.; Stenberg, B. Improving the Prediction Performance of a Large Tropical Vis-NIR Spectroscopic Soil Library from Brazil by Clustering into Smaller Subsets or Use of Data Mining Calibration Techniques. *Eur. J. Soil Sci.* **2014**, *65*, 718–729. [CrossRef]
16. Wold, S.; Sjöström, M.; Eriksson, L. PLS-Regression: A Basic Tool of Chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [CrossRef]
17. Deiss, L.; Margenot, A.J.; Culman, S.W.; Demyan, M.S. Tuning Support Vector Machines Regression Models Improves Prediction Accuracy of Soil Properties in MIR Spectroscopy. *Geoderma* **2020**, *365*, 114227. [CrossRef]
18. Lucà, F.; Conforti, M.; Castrignanò, A.; Matteucci, G.; Buttafuoco, G. Effect of Calibration Set Size on Prediction at Local Scale of Soil Carbon by Vis-NIR Spectroscopy. *Geoderma* **2017**, *288*, 175–183. [CrossRef]
19. Anderson, S.T. Economics, Helium, and the U.S. Federal Helium Reserve: Summary and Outlook. *Nat. Resour. Res.* **2018**, *27*, 455–477. [CrossRef]
20. VM0021 Soil Carbon Quantification Methodology, v1.0-Verra. Available online: <https://verra.org/methodology/vm0021-soil-carbon-quantification-methodology-v1-0/> (accessed on 18 October 2022).
21. Sherrod, L.A.; Dunn, G.; Peterson, G.A.; Kolberg, R.L. Inorganic Carbon Analysis by Modified Pressure-Calimeter Method. *Soil Sci. Soc. Am. J.* **2002**, *66*, 299–305. [CrossRef]
22. Leuthold, S.J.; Haddix, M.L.; Lavallee, J.; Cotrufo, M.F. Physical Fractionation Techniques. *Ref. Modul. Earth Syst. Environ. Sci.* **2022**. [CrossRef]
23. Kennard, R.W.; Stone, L.A. Computer Aided Design of Experiments. *Technometrics* **1969**, *11*, 137–148. [CrossRef]
24. Wickham, H.; François, R.; Henry, L.; Müller, K. Dplyr: A Grammar of Data Manipulation. 2022. Available online: <https://dplyr.tidyverse.org> (accessed on 25 October 2022).
25. Kuhn, M. Building Predictive Models in R Using the Caret Package. *J. Stat. Softw.* **2008**, *28*, 1–26. [CrossRef]
26. Pirouz, D.M. An Overview of Partial Least Squares. *SSRN Electron. J.* 2006. [CrossRef]
27. Vapnik, V. The Support Vector Method of Function Estimation. In *Nonlinear Modeling*; Springer: Boston, MA, USA, 1998; pp. 55–85. [CrossRef]
28. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
29. Ballabio, D.; Todeschini, R.; Todeschini, R.O.; Bahmani, A.; Consonni, V. Evaluation of model predictive ability by external validation techniques. *J. Chemom.* **2010**, *24*, 194–201. [CrossRef]
30. Chai, T.; Draxler, R.R. Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? -Arguments against Avoiding RMSE in the Literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [CrossRef]
31. Willmott, C.J.; Matsuura, K. Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance. *Clim. Res.* **2005**, *30*, 79–82. [CrossRef]
32. Jierula, A.; Wang, S.; Oh, T.M.; Wang, P. Study on Accuracy Metrics for Evaluating the Predictions of Damage Locations in Deep Piles Using Artificial Neural Networks with Acoustic Emission Data. *Appl. Sci.* **2021**, *11*, 2314. [CrossRef]
33. Ahmed, M.; Ahmad, S.; Ali Raza, M.; Kumar, U.; Ansar, M.; Abbas Shah, G.; Parsons, D.; Hoogenboom, G.; Palosuo, T.; Seidel, S.; et al. Models Calibration and Evaluation. In *Systems Modeling*; Springer: Singapore, 2020; pp. 151–178. [CrossRef]
34. Anscombe, F.J.; Tukey, J.W. The Examination and Analysis of Residuals. *Technometrics* **1963**, *5*, 141–160. [CrossRef]
35. Ng, W.; Minasny, B.; de Sousa Mendes, W.; Melo Demattê, J.A. The Influence of Training Sample Size on the Accuracy of Deep Learning Models for the Prediction of Soil Properties with Near-Infrared Spectroscopy Data. *SOIL* **2020**, *6*, 565–578. [CrossRef]
36. Ramirez-Lopez, L.; Schmidt, K.; Behrens, T.; van Wesemael, B.; Demattê, J.A.M.; Scholten, T. Sampling Optimal Calibration Sets in Soil Infrared Spectroscopy. *Geoderma* **2014**, *226*, 140–150. [CrossRef]
37. Rossel, R.A.V.; Behrens, T. Using Data Mining to Model and Interpret Soil Diffuse Reflectance Spectra. *Geoderma* **2010**, *158*, 46–54. [CrossRef]
38. Tange, R.I.; Rasmussen, M.A.; Taira, E.; Bro, R. Benchmarking Support Vector Regression against Partial Least Squares Regression and Artificial Neural Network: Effect of Sample Size on Model Performance. *J. Near Infrared Spectrosc.* **2017**, *25*, 381–390. [CrossRef]
39. Janik, L.J.; Merry, R.H.; Forrester, S.T.; Lanyon, D.M.; Rawson, A. Rapid Prediction of Soil Water Retention Using Mid Infrared Spectroscopy. *Soil Sci. Soc. Am. J.* **2007**, *71*, 507–514. [CrossRef]
40. Ramírez, P.B.; Calderón, F.J.; Haddix, M.; Lugato, E.; Cotrufo, M.F. Using Diffuse Reflectance Spectroscopy as a High Throughput Method for Quantifying Soil C and N and Their Distribution in Particulate and Mineral-Associated Organic Matter Fractions. *Front. Environ. Sci.* **2021**, *9*, 153. [CrossRef]
41. Baldock, J.A.; Beare, M.H.; Curtin, D.; Hawke, B. Stocks, Composition and Vulnerability to Loss of Soil Organic Carbon Predicted Using Mid-Infrared Spectroscopy. *Soil Res.* **2018**, *56*, 468–480. [CrossRef]
42. Janik, L.J.; Skjemstad, J.O.; Shepherd, K.D.; Spouncer, L.R. The Prediction of Soil Carbon Fractions Using Mid-Infrared-Partial Least Square Analysis. *Aust. J. Soil Res.* **2007**, *45*, 73–81. [CrossRef]
43. Li, S.; Viscarra Rossel, R.A.; Webster, R.; Raphael Viscarra Rossel, C.A. The Cost-Effectiveness of Reflectance Spectroscopy for Estimating Soil Organic Carbon. *Eur. J. Soil Sci.* **2021**, *73*, e13202. [CrossRef]
44. Bellon-Maurel, V.; McBratney, A. Near-Infrared (NIR) and Mid-Infrared (MIR) Spectroscopic Techniques for Assessing the Amount of Carbon Stock in Soils-Critical Review and Research Perspectives. *Soil Biol. Biochem.* **2011**, *43*, 1398–1410. [CrossRef]