



## Article

# Evaluating the Precision and Accuracy of Proximal Soil vis–NIR Sensors for Estimating Soil Organic Matter and Texture

Nandkishor M. Dhawale <sup>1,\*</sup>, Viacheslav I. Adamchuk <sup>1,\*</sup> , Shiv O. Prasher <sup>1</sup> and Raphael A. Viscarra Rossel <sup>2</sup>

<sup>1</sup> Department of Bioresource Engineering, McGill University, Sainte-Anne-de-Bellevue, QC H9X 3V9, Canada; shiv.prasher@mcgill.ca

<sup>2</sup> Soil and Landscape Science, School of Molecular and Life Sciences, Faculty of Science and Engineering, Curtin University, G.P.O. Box U1987, Perth, WA 6845, Australia; R.Viscarra-Rossel@curtin.edu.au

\* Correspondence: nandkishor.dhawale@mail.mcgill.ca (N.M.D.); viacheslav.adamchuk@mcgill.ca (V.I.A.)

**Abstract:** Measuring soil texture and soil organic matter (SOM) is essential given the way they affect the availability of crop nutrients and water during the growing season. Among the different proximal soil sensing (PSS) technologies, diffuse reflectance spectroscopy (DRS) has been deployed to conduct rapid soil measurements in situ. This technique is indirect and, therefore, requires site- and data-specific calibration. The quality of soil spectra is affected by the level of soil preparation and can be accessed through the repeatability (precision) and predictability (accuracy) of unbiased measurements and their combinations. The aim of this research was twofold: First, to develop a novel method to improve data processing, focusing on the reproducibility of individual soil reflectance spectral elements of the visible and near-infrared (vis–NIR) kind, obtained using a commercial portable soil profiling tool, and their direct link with a selected set of soil attributes. Second, to assess both the precision and accuracy of the vis–NIR hyperspectral soil reflectance measurements and their derivatives, while predicting the percentages of sand, clay and SOM content, in situ as well as in laboratory conditions. Nineteen locations in three agricultural fields were identified to represent an extensive range of soils, varying from sand to clay loam. All measurements were repeated three times and a ratio spread over error (RSE) was used as the main indicator of the ability of each spectral parameter to distinguish among field locations with different soil attributes. Both simple linear regression (SLR) and partial least squares regression (PLSR) models were used to define the predictability of % SOM, % sand, and % clay. The results indicated that when using a SLR, the standard error of prediction (SEP) for sand was about 10–12%, with no significant difference between in situ and ex situ measurements. The percentage of clay, on the other hand, had 3–4% SEP and 1–2% measurement precision (MP), indicating both the reproducibility of the spectra and the ability of a SLR to accurately predict clay. The SEP for SOM was only a quarter lower than the standard deviation of laboratory measurements, indicating that SLR is not an appropriate model for this soil property for the given set of soils. In addition, the MPs of around 2–4% indicated relatively strong spectra reproducibility, which indicated the need for more expanded models. This was apparent since the SEP of PLSR was always 2–3 times smaller than that of SLR. However, the relatively small number of test locations limited the ability to develop widely applicable calibration models. The most important finding in this study is that the majority of vis–NIR spectral measurements were sufficiently reproducible to be considered for distinguishing among diverse soil samples, while certain parts of the spectra indicate the capability to achieve this at  $\alpha = 0.05$ . Therefore, the innovative methodology of evaluating both the precision and accuracy of DRS measurements will help future developers evaluate the robustness and applicability of any PSS instrument.

**Keywords:** proximal soil sensing; diffuse reflectance spectroscopy; visible and near-infrared spectra; measurement precision and accuracy



**Citation:** Dhawale, N.M.; Adamchuk, V.I.; Prasher, S.O.; Viscarra Rossel, R.A. Evaluating the Precision and Accuracy of Proximal Soil vis–NIR Sensors for Estimating Soil Organic Matter and Texture. *Soil Syst.* **2021**, *5*, 48. <https://doi.org/10.3390/soilsystems5030048>

Academic Editor: Heike Knicker

Received: 16 June 2021

Accepted: 16 August 2021

Published: 20 August 2021

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



**Copyright:** © 2021 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

Soil texture and soil organic matter (SOM) affect plant nutrient availability. Measuring these properties using conventional laboratory methods is laborious and time consuming [1,2]. Wherever possible, the use of proximal soil sensing (PSS) is expected to replace conventional laboratory methods to help reduce time- and labor-intense operations during soil sampling and analysis [3,4]. Among different PSS technologies, diffuse reflectance spectroscopy (DRS) has been deployed to conduct rapid soil measurements in laboratories and in situ. DRS is an indirect analytical technique [5] that was proposed as an alternative to laboratory tests that would allow for relatively inexpensive and rapid measurements of soil properties.

A soil spectroscopic system is based on the sample's absorption of electromagnetic radiation at wavelengths in the ranges of 100–400 nm (UV), 400–700 nm (vis), 700–2500 nm (NIR and SWIR regions), and 2500–25,000 nm (mid-IR) [6]. Soriano-Disla et al. [7] reviewed a number of studies that simultaneously determined different soil attributes using this method. For example, the authors of [8] simultaneously determined soil organic carbon (SOC), moisture, and total nitrogen (TN) using NIR spectroscopy. The authors of [9] simultaneously determined pH, clay, silt, sand, SOC content and cation exchange capacity using mid-IR.

Similarly, numerous studies support the use of diffuse reflectance spectroscopy to simultaneously assess various soil properties in the laboratory as well as directly in the field [10–34]. The Veris<sup>®</sup> P4000 (Veris Technologies, Inc., Salina, KS, USA) is an example of a ruggedized hyperspectral instrument developed to obtain vis–NIR spectra at multiple depths within the soil profile. The data obtained from similar instrument was used previously to predict soil phosphorus, soil texture and SOM content [35,36]. In addition, the authors of [37] showed the potential use of similar probing instruments to produce three-dimensional digital soil maps of agricultural fields, by integrating multiple proximal sensor data obtained from different sensing methods.

Since the use of spectroscopic technique to predict the quantities of soil properties is indirect, it is important to develop calibration models using data-specific pedo-transformation models against measurements obtained according to standardized laboratory protocols [3]. Depending on the linear and nonlinear data types, a variety of simple to complex calibration models were studied by various researchers through the performance of multivariate data analysis, e.g., principal component regression (PCR) [16], partial least squares regression (PLSR) [38], artificial neural networks (ANN) [39,40], and support vector machines (SVM) [41], to name a few.

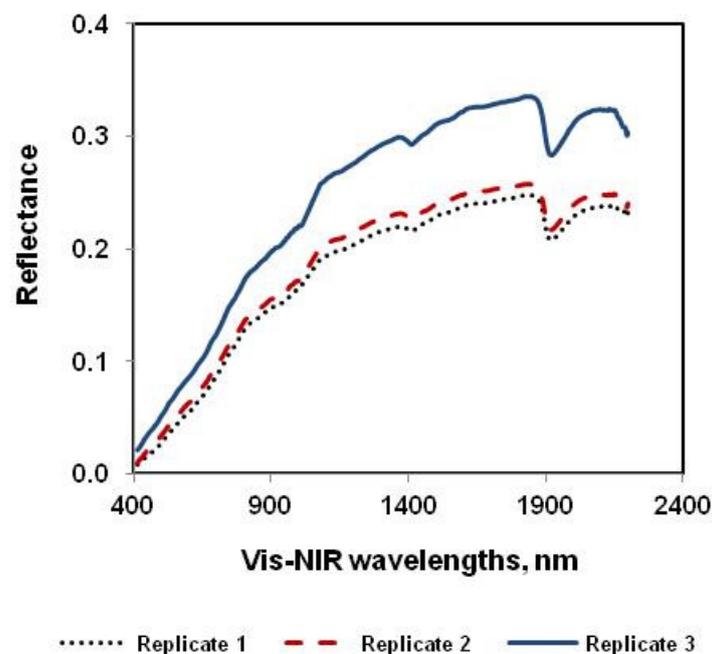
However, a good calibration model can be developed only when a good set of data is obtained from the vis–NIR instrument. This is especially the case when soil is made up of heterogeneous material; the quality of soil spectra can be affected by the level of soil preparation. Generally spectral measurements need to be accurate and precise representations of the targeted material (soil in this case). However, there are a variety of factors that affect the quality of spectral measurements obtained using vis–NIR instruments in both the laboratory and in the field; for example, viewing geometry, orientation of measurement, spectral averaging and calibration of the spectral data [42–44]. In addition to this environment, excess water content in soil, and excess soil aggregates and debris can also influence the data collection procedure [38]. The sampling process is also known to influence measurement uncertainty [45].

Therefore, it is clear that the data acquisition and processing may influence the performance of the calibration models [46–49].

To help improve data acquisition procedures in DRS, it is important to understand the random fluctuation of soil spectra and their derivatives from one scan to the next. As with any measurement, the quality of soil spectra obtained using DRS while predicting any soil attribute can be evaluated using precision and accuracy. Precision (reproducibility) and accuracy (predictability) are considered to be very important characteristics of a given measurement system [50,51]. Where accuracy is not a quantity and precision has a much

more complex meaning, both can be described as the difference of the instrument reading to the standard, and the degree of random variations in the instrument's output while measuring a constant or similar objects (the same soil in the case of the research discussed in this paper) under specified conditions [52]. Other way around inaccuracy may also be defined as combining both bias and imprecision. Where bias is another quantitative term describing the difference between the average of measurements made on the same object (same soil in the case of this research) and its true value [53].

As is illustrated in Figure 1, repeated measurements of soil spectra for a given soil sample can significantly vary from one measurement to another. These differences can be associated with (1) random noise within the data acquisition process that includes inconsistencies of the soil detector interface and (2) soil micro-variability that can occur at very small separation distances. In practice, multiple measurements and the averaging of several consecutive soil spectra were used to diminish the effect of random noise [54]. However, it should be noted that certain parts of the soil spectra, as well as particular parameters constructed using a combination of spectral data values, can be relatively repeatable from spectrum to spectrum. It was also observed that some of these soil spectra parameters do not vary substantially when testing the same soil, but vary considerably when using soil samples with different attributes; therefore, they might be the best candidates to be used for the site-specific calibration process.



**Figure 1.** Differences in ex situ soil spectra when measured on the same soil sample thrice.

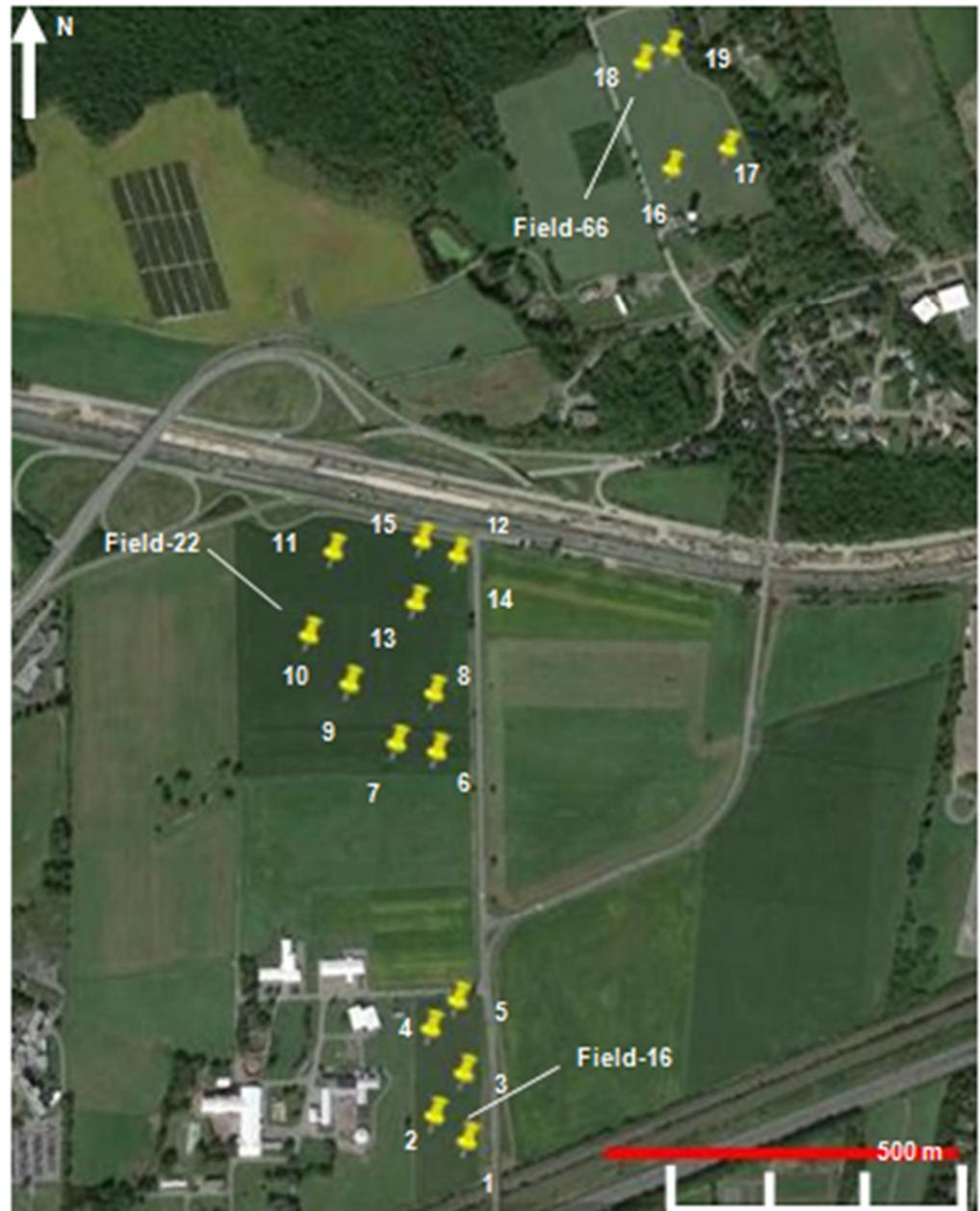
The main objective of this research was twofold: First, to develop a novel method to help improve data processing, focusing on the reproducibility of individual soil reflectance measurements and the transformations of the visible and near-infrared (vis–NIR) kind, obtained using a commercial portable soil profiling tool, and to determine their direct link with a selected set of soil attributes. Second, to assess both the precision and accuracy of the vis–NIR hyperspectral soil reflectance measurements and their transformations, while predicting the percentages of sand, clay and SOM content, in situ as well as in laboratory conditions.

## 2. Materials and Methods

### 2.1. Data Collection

Three fields at McGill University Macdonald research farm (45°24'42.2" N 73°56'23.1" W; Sainte-Anne-de-Bellevue, QC, Canada) were used for this project. In total, nineteen loca-

tions were selected in these fields (Figure 2) to represent diverse soil conditions ranging from sand to clay loam. The fields had 2.5, 4.5 and 12 ha areas and crops were grown according to a rotation involving soybean, corn (grain and silage) and alfalfa. All data were collected in 2012 when an alfalfa/grass mix was grown in each of the three fields.



**Figure 2.** Google earth view of McGill University Macdonald research farm site, illustrating sampling points.

In each location, composite soil samples were obtained from the top (0–20 cm) layer of soil. A 4 cm diameter, stainless steel auger was used to take five samples from within a 0.5 m radius. These soil cores were mixed, air dried, and sieved through a 2 mm mesh. Then, each sample was divided into two subsamples: one to be used for conventional laboratory analysis and the other for *ex situ vis-NIR* spectral measurements.

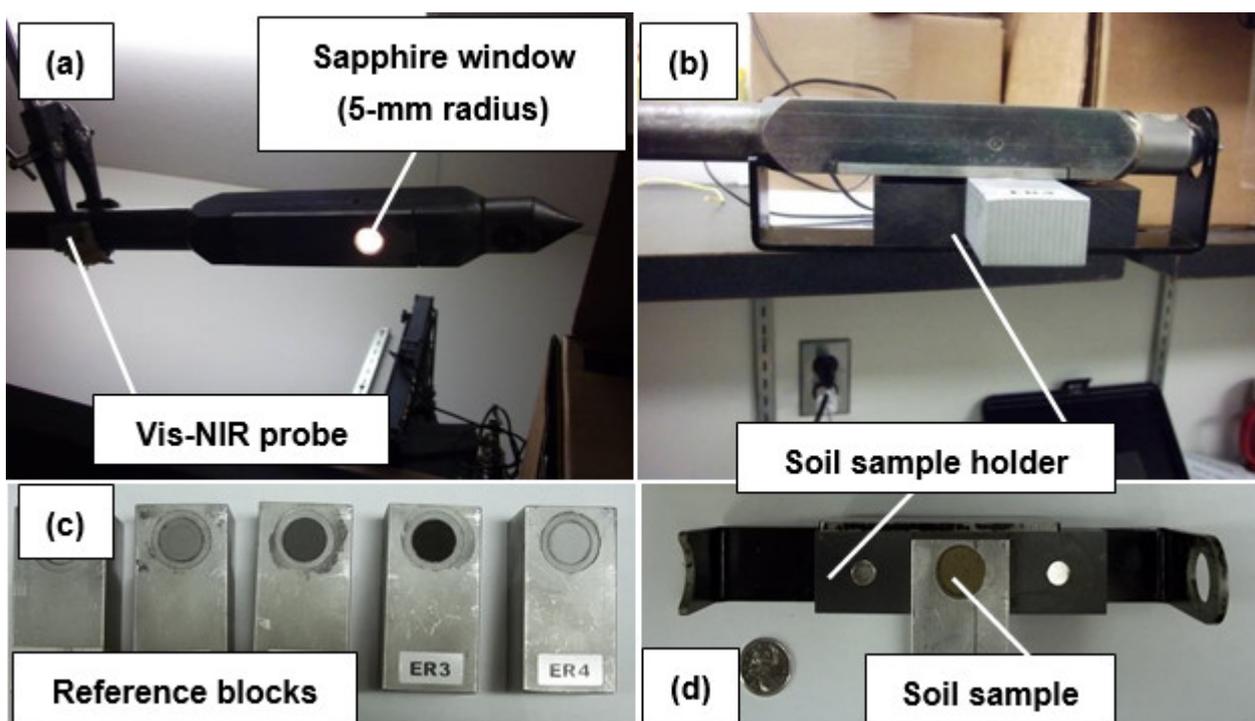
Three categories have been established for soil particles: sand, silt, and clay [55]. These three groups are called soil separates. The three groups are divided by their particle size. Clay particles are the smallest, while sand particles are the largest. Particle size

analysis (fractions of sand, silt, and clay) as well as SOM content were evaluated for each laboratory sample, as summarized in Table 1. The particle size analysis was conducted using hydrometers [1] and SOM was determined using the loss-on-ignition technique [2].

**Table 1.** Statistical results of the percentages of sand, clay, and SOM content in 19 soils.

| Statistics | % Sand | % Clay | % SOM |
|------------|--------|--------|-------|
| Min        | 26     | 2      | 3.9   |
| Median     | 56     | 9      | 5.8   |
| Max        | 93     | 33     | 25.6  |
| Mean       | 56     | 13     | 7.2   |
| SD         | 18     | 9      | 4.8   |

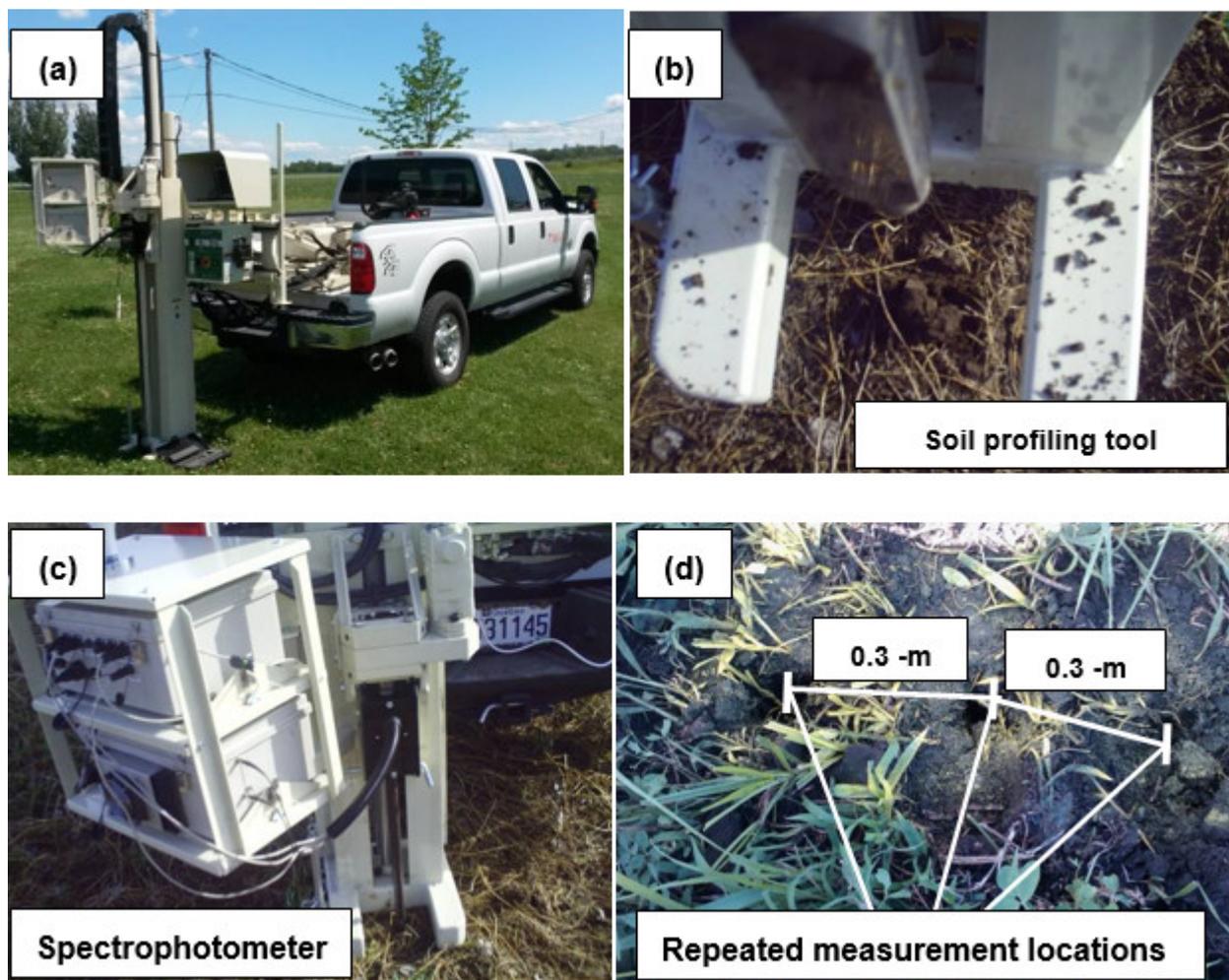
The core of the Veris P4000 instrument is a combined dual type spectrophotometer operating in the visible and near-infrared parts of the electromagnetic spectra (Figure 3). One of the two spectrophotometers was USB2000 (Ocean Optics, Dunedin, FL, USA) operating between 342 and 1023 nm with a spectral resolution of 6 nm. The other spectrophotometer was C9914GB (Hamamatsu Photonics, K.K., Tokyo, Japan), which collected data between 1070 and 2220 nm with a spectral resolution of 4 nm. The instrument included its own light source and was capable of maintaining a constant distance between measured soil surfaces and detectors by means of a sapphire contact probe with fibre-optic cables.



**Figure 3.** Illustration of the vis–NIR soil profiling tool during laboratory measurements (sapphire window on vis–NIR probe (a), soil sample holder (b), reference blocks (c) and soil sample (d)).

All ex situ measurements were conducted as triplicates in a different randomised order for each replicate. A specially designed sample holder (Figure 3b) was filled with <1 g of the soil sample and placed near the optical window of the soil profiling tool. As was recommended in the user manual provided by the manufacturer, at the beginning of each spectral measurement session, the instrument was calibrated by measuring the dark current followed by the white reference measurements using the specially provided reference blocks, as shown in Figure 3c. The instrument was re-calibrated after every

20 samples. Soil spectra were interpolated to about 5 nm of the spectral resolution, yielding a total of 380 data points (wavelengths) per spectrum. To minimize the instrument noise, each spectrum was the average of 25–30 scans ( $\sim 6 \text{ scans s}^{-1}$ ). The in situ measurements were collected using the recommended equipment setup (Figure 4) for topsoil profiles down to 20 cm, while penetrating soil at a speed approximately  $2 \text{ cm s}^{-1}$ . Three measurements were conducted consecutively along a straight line that was less than 0.5 m long. The instrument was re-calibrated in a similar manner as discussed for the ex situ measurements. The soil spectra data collection procedures for in situ and ex situ measurements were the same. However, in this case, the average of 50–60 scans represented soil spectra collected at different depths from 0 to 20 cm.

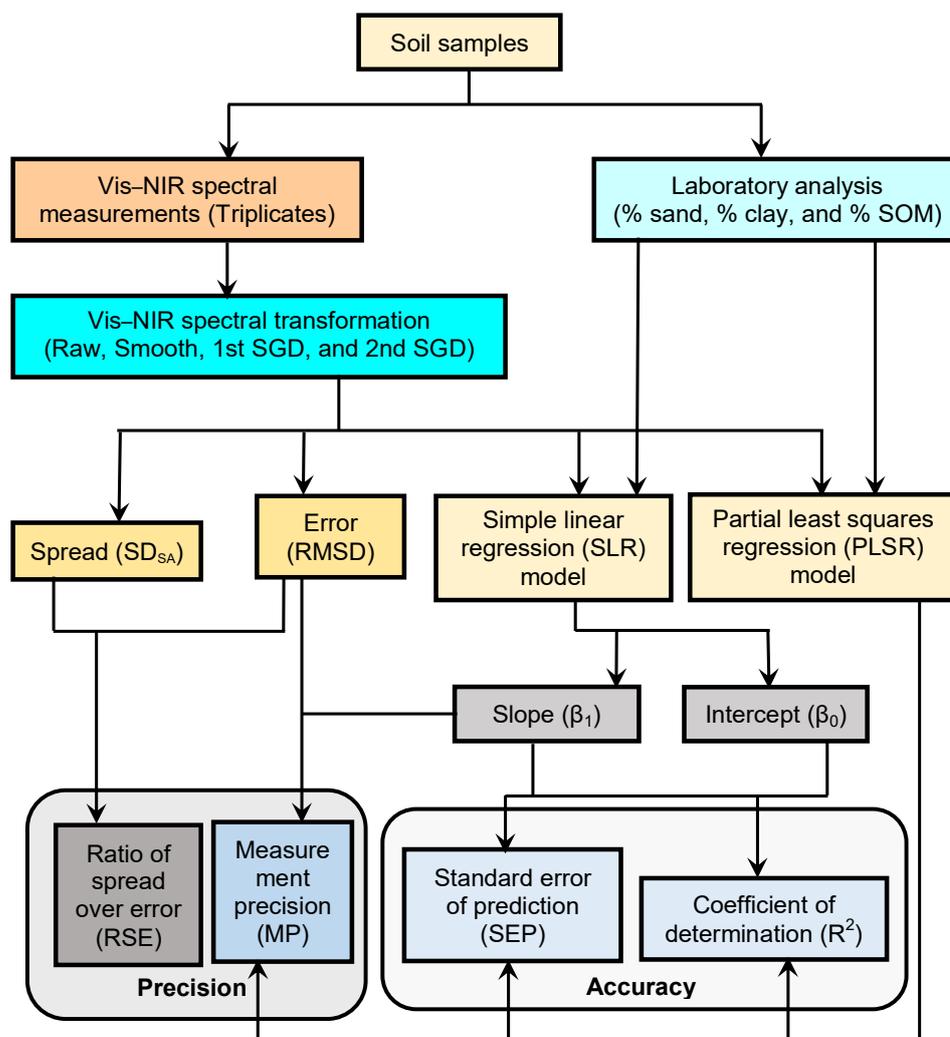


**Figure 4.** Illustration of the vis–NIR soil profiling tool during in situ measurements (pickup truck mounted instrument (a), soil profiling tool (b), both spectrometers inside the enclosures (c) and repeated measurement locations (d)).

## 2.2. Data Processing

As discussed previously (in Section 2.1) and as shown in Figure 5, the DRS measurements were collected in triplicate on representative soil samples using the soil profiling tool (ex situ and in situ), whereas the representative soil samples were analyzed in laboratories to obtain the standard measurements of the soil attributes of interest. Next, the DRS measurements were processed to create three more data sets (four in total: Raw, Smooth, 1st SGD, and 2nd SGD). Later, to understand the random fluctuation of soil spectra within the same samples and between different samples, the spread ( $SD_{SA}$ ) and the root mean squared deviation (RMSD) were calculated on the DRS measurements for all four data sets.

Simultaneously, the best SLR and PLSR models were fit on all four DRS measurement data sets against the known standard measurements of the soil attributes of interest. Finally, while focusing on the reproducibility of soil reflectance and its direct link with soil attributes, precision- and accuracy-related components such as the ratio of spread over error (RSE), measurement precision (MP), standard error of prediction (SEP), and coefficient of determination ( $R^2$ ), respectively, were calculated. These are explained in greater detail later in the text.



**Figure 5.** Illustration of the novel data processing methodology.

All raw spectral data were processed using MATLAB 2012a (The MathWorks, Inc. Natick, MA, USA) and ParLeS 3.1 software (University of Sydney, Sydney, Australia), as described by [56]. All spectra exhibited a step discontinuity from 1023 to 1070 nm, caused by the transition from one detector to another. After removing the relatively noisy parts of the spectra at the edges of the detection ranges for each spectrophotometer (342–409 nm, 1014–1075 nm, and 2206–2220 nm), all resultant spectra consisted of a total of 363 measurements at different wavelengths. The spectral data were then corrected for offset and processed using a multiplicative scattering correction (MSC) algorithm [57] and mean centering (MC). In addition to these “raw” spectra, the following spectra treatments were pursued: (1) 3-point Savitzky–Golay smoothing, (2) 11-point first order Savitzky–Golay derivative (1st SGD), and (3) 11-point second order Savitzky–Golay derivative (2nd SGD) [58].

Among the different statistics considered for assessing the precision of spectral data, the RSE was used:

$$RSE = \frac{SD_{SA}}{RMSD} \quad (1)$$

where  $SD_{SA}$  is the standard deviation of nineteen sample mean values;  $RMSD$  is the root mean squared deviation calculated based on three replicated measurements for each of the nineteen samples.

$$SD_{SA} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} \quad (2)$$

where  $n = 19$  is the number of soil samples;  $X_i$  is an average measured or calculated spectra related value for the  $i$ th sample;  $\bar{X}$  is the average of all  $X_i$  values.

$$RMSD = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (x_{ij} - X_i)^2}{n \cdot (m-1)}} \quad (3)$$

where  $m = 3$  is the number of replicated measurements;  $x_{ij}$  is a  $j$ th replicate of the measured or calculated spectra-related value for the  $i$ th sample

$$X_i = \frac{\sum_{j=1}^m x_{ij}}{m} \quad (4)$$

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (5)$$

Similar to the frequently used ratio of prediction over deviation (RPD) [59], high RSE means a relatively strong ability of a given measurement to distinguish different soil samples. The use of RSE was reported earlier in [60,61] and is directly related to ANOVA F statistics used to compare the means of repeated measurements. Based on the degrees of freedom involved, the difference among the soil samples (means of three measurements) can be detected at  $\alpha = 0.05$  when RSE is greater than  $0.79 \left(\sqrt{\frac{1}{m} F_{stat}}\right)$ . This analysis evaluates measurement precision with the underlying hypothesis that a particular parameter that does not change when measuring the same soil sample, and for which the change is at its maximum when measuring different samples, should be considered reliable.

The percentages of sand, clay, and SOM were predicted by fitting SLR models on each individual measured or treated spectral value versus these properties. A coefficient of determination ( $R^2$ ) was the main indicator of the ability of a single spectral value to explain the variability in a particular soil property. However, the standard error of prediction (SEP) was used as a measure of the accuracy of soil property estimates obtained using each SLR model:

$$\hat{y}_{ij} = \beta_0 + \beta_1 x_{ij} \quad (6)$$

where  $x$  is the measured value of a given soil property for the  $i$ th sample;  $y$  is the predicted value of a given soil property for the  $i$ th sample and the  $j$ th replicate;  $\beta_0$  (intercept) and  $\beta_1$  (slope) are coefficients of SLR.

$$SEP = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (y_i - \hat{y}_{ij})^2}{n \cdot m}} \quad (7)$$

where  $y$  is the measured value of a given soil property for the  $i$ th sample.

While SEP indicated the total error associated with each individual measurement and is primarily linked to the accuracy of prediction, RMSD is linked to measurement precision. However, since RMSD is expressed in the units of spectra measurements and related calculated parameters, measurement precision (MP) can be expressed in physical units as follows:

$$MP = \beta_1 \cdot RMSD \quad (8)$$

When comparing different spectral wavelengths and transformation techniques, it is important to identify measurements that have the maximum RSE and the minimum RMSD, MP, and SEP. The RMSD is an indicator of measurement reproducibility. However, without considering the spread of values across different samples, it is impossible to conclude if the given values are strong values to distinguish different soil samples from each other. Therefore, the RSE is involved in electing candidates to differentiate between the samples' levels of disregard of the prediction property. Neither RMSD nor RSE depend on the model used to predict a given soil property.

Because of the SLR approach [62] used to test one-input soil property prediction functions, RMSD can be expressed, in terms of the percentages of sand, clay, or SOM, as MP. The MP estimate is then evaluated together with SEP, which is the ultimate indicator of the predictability of soil properties. Unlike RMSD, the MP as well as the SEP can be compared across different spectra transformation procedures as both are expressed in physical units. From a sensor development point of view, a small RMSD and MP indicates a stable soil–detector interface. A high RSE means that the sensor can be applied to a particular set of soils. Finally, a small SEP (high  $R^2$  for a given set of samples) indicates the sensor's ability to predict the soil property of interest. The SEP is always greater than the MP, and the greater this difference, the less uncertain the linear relationship between the measured value and the property. Small differences between the SEP and MP indicate the applicability of the prediction model when reliable measurement estimates are obtained. In other words, small differences between the SEP and MP indicate the potential for improved predictability by averaging multiple unbiased measurements, but larger differences signify the limitations of the model and that alternative prediction methods, such as PLSR, should be involved.

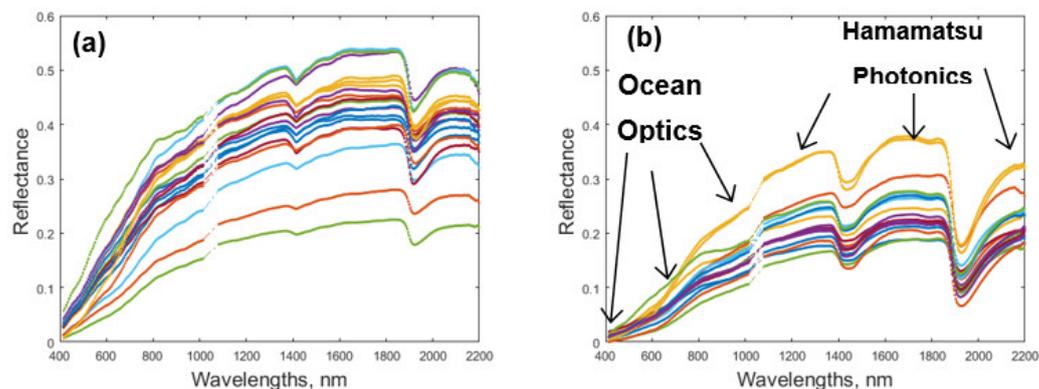
In soil spectroscopy, PLSR is one of the most widely used techniques to aggregate measurements obtained at multiple wavelengths in a single prediction model [32–35]. The PLSR is a bilinear regression technique that extracts a small number of latent factors, which are a combination of the independent variables, and uses these factors as a regression producer for the dependent variables [63,64]. The PLSR analysis is normally evaluated using the leave-one-out cross validation technique, and the RMSD,  $R^2$  and Akaike Information Criterion (AIC), as described in [65], are the most common model performance indicators.

In this study, the orthogonalised PLSR-1 algorithm described in [66] was applied to (1) raw spectra, (2) smoothed spectra, (3) 1st SGD spectra, (4) 2nd SGD spectra, and (5) all the values were combined to develop calibration models using ParLeS software [56]. The number of factors to use in each model was selected using leave-one-out cross validation. However, due to the limited number of soil samples used ( $n = 19$ ), the selected models were not re-validated on a different set of soil samples; however, this was generally undertaken in other reported studies. The developed models were used to estimate the performance indicators that were comparable to MP and SEP to define the superiority of PLSR over SLR models.

### 3. Results and Discussions

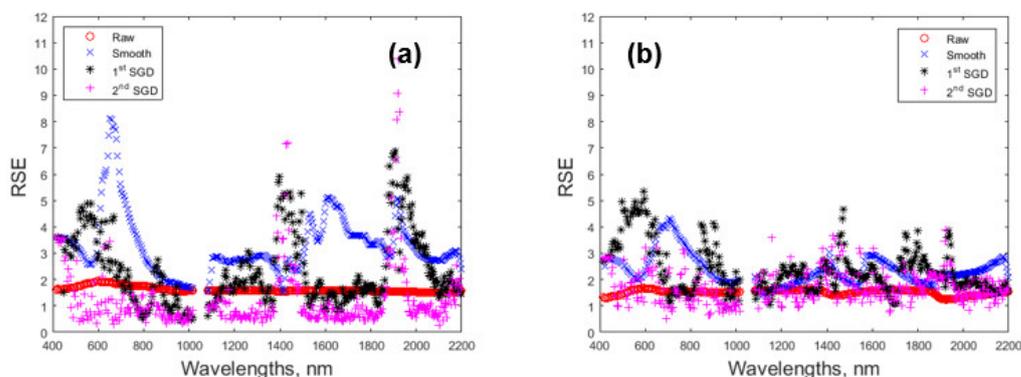
Statistics derived from laboratory results based on the percentages of sand, clay and SOM contents are listed in Table 1. These results demonstrate the variability observed over relatively small fields in terms of their texture and SOM content. It is interesting to note that the standard deviation of % sand in this dataset is two times higher than that of % clay and four times higher than that of % SOM.

Average reflectance spectra of nineteen soil samples, collected both ex situ and in situ, are shown in Figure 6. In general, the laboratory spectra exhibited higher reflectance when compared with those from the field, which might be explained by dry versus wet soil. All spectra showed water absorption inversions near 1400 and 1900 nm [24,38]. Less noticeable inversions were also observed in many other parts of the spectra (1100, 1500, 1600, 1700, and 2100 nm) which could be associated with the primary or secondary effects of various minerals, carbonates and SOM [24,38]. Several noticeable peaks were also observed at certain parts of the spectra, especially in the visible region, that could be associated with soil color.



**Figure 6.** Examples of vis-NIR soil spectra (ex situ (a) and in situ (b)).

Figure 7 illustrates the RSE values produced for the raw and all transformed spectra and Table 2 summarizes the relevant statistical results. It is noticeable that transforming the raw soil spectra into a 3-point Savitzky–Golay smooth spectrum, and the first and second order Savitzky–Golay derivative spectra, increased the reproducibility in many parts of the spectra. It is clear that the second SGDs had the highest relative reproducibility across certain parts of the ex situ spectrum, while RSE for the first SGDs had the highest relative reproducibility across certain parts of the in situ spectrum. One might expect that ex situ measurements were more reproducible than those conducted in natural soil conditions within a 0.5 m distance from each other. However, in both cases, first SGD yielded data with comparable relative reproducibility in certain parts of the spectrum. In large part, each transformed spectra had some portions with the ability to separate the means of the three replicated measurements between different soil samples ( $RSE > 0.79$ ).

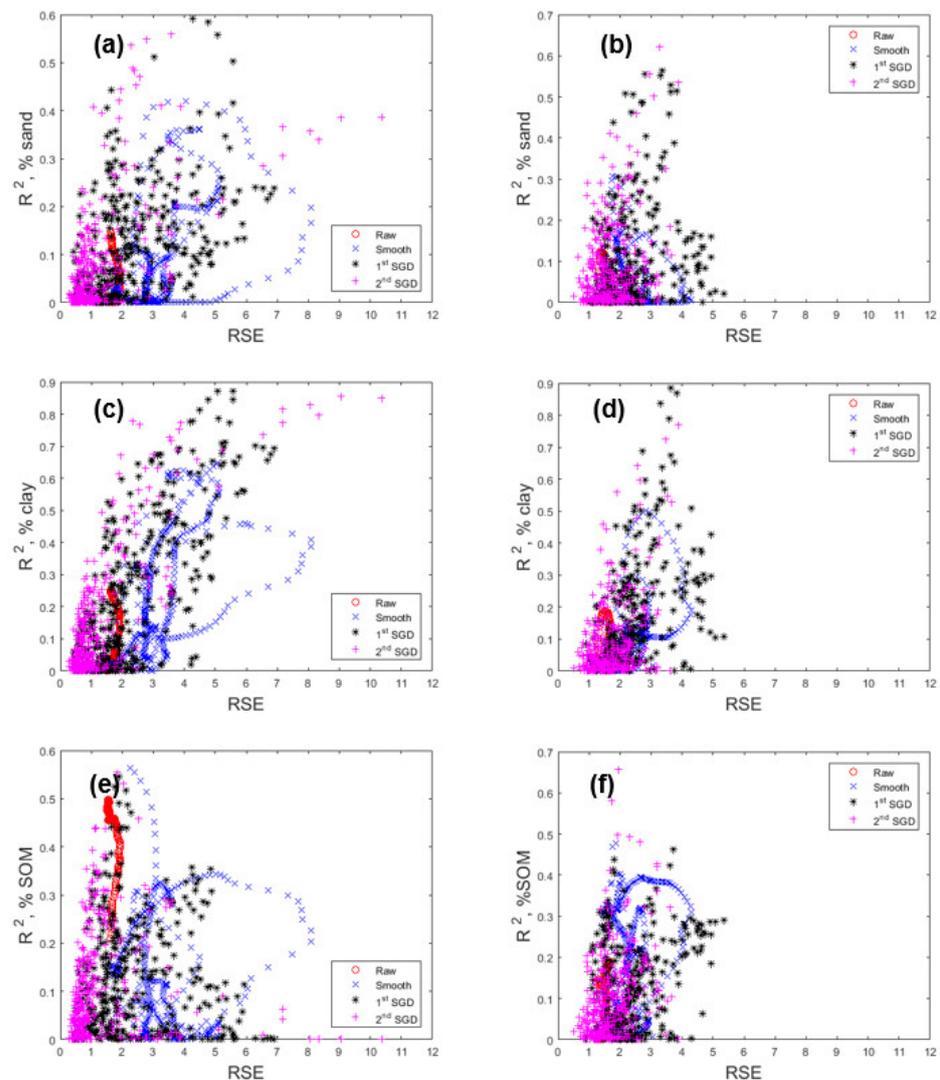


**Figure 7.** RSE estimates for soil spectra and their derivatives (ex situ (a) in situ (b)).

**Table 2.** Ratio of spread over error (RSE) comparison.

| Test    | Transformed Factors | Min  | Median | Mean | Max   |
|---------|---------------------|------|--------|------|-------|
| Ex situ | Raw                 | 1.69 | 1.76   | 1.81 | 2.18  |
|         | Smooth              | 1.62 | 3.04   | 3.37 | 8.71  |
|         | 1st SGD             | 0.47 | 1.91   | 2.49 | 6.93  |
|         | 2nd SGD             | 0.32 | 0.86   | 1.32 | 10.56 |
| In situ | Raw                 | 1.24 | 1.53   | 1.51 | 1.67  |
|         | Smooth              | 1.38 | 2.23   | 2.35 | 4.30  |
|         | 1st SGD             | 0.89 | 2.14   | 2.33 | 5.34  |
|         | 2nd SGD             | 0.52 | 1.59   | 1.70 | 3.88  |

Figure 8 illustrates the coefficients of determination versus the RSE's, while predicting the percentages of sand, clay, and SOM contents using SLR. Table 3 summarizes the wavelengths and transformation methods indicating the highest level of predictability for the soil properties of interest. Selection of these wavelengths was predominantly guided by the coefficient of determination as the weighting between  $R^2$  and RSE is arbitrary.



**Figure 8.** % sand, % clay, and % SOM correlations versus the calculated RSE's for spectra and its derivatives (ex situ (a,c,e) and in situ (b,d,f)).

**Table 3.** vis–NIR wavelengths (nm) indicating the highest correlation ( $R^2$ ) and ratio of spread over error (RSE) and with % sand, % clay and SOM content.

| Soil Property | Test    | Raw      | Smooth   | 1st SGD <sup>a</sup>   | 2nd SGD <sup>b</sup>        |
|---------------|---------|----------|----------|------------------------|-----------------------------|
| % sand        | Ex situ | -        | -        | 1932, 1936, 1940       | 1413                        |
|               | In situ | -        | -        | 1447, 1452, 1940, 1944 | 1457, 1462, 1466, 1923      |
| % clay        | Ex situ | -        | -        | 1936, 1940, 1944, 1948 | 1433, 1915, 1919, 1923      |
|               | In situ | -        | -        | 1931, 1936, 1940, 1944 | 1919, 1923                  |
| % SOM         | Ex situ | 733, 786 | 712, 744 | 2114, 2117, 2121, 2148 | -                           |
|               | In situ | -        | 749, 770 | -                      | 602, 1738, 1742, 2159, 2182 |

<sup>a</sup> First order Savitzky–Golay derivative. <sup>b</sup> Second order Savitzky–Golay derivative.

Alternatively, Figure 9 illustrates SEP versus MP while predicting the soil properties. All values were calculated using both the SLR and PLSR calibration models. To facilitate the comparison, the standard deviation of soil properties (laboratory measurements) was shown as a benchmark for SEP and the isolines represented positions in the chart with the constant sum of MP and SEP. To assure that both error terms are small, the most promising wavelengths and spectra transformation methods were selected as well as those placed between two isolines, indicating the lowest sum of MP and SEP. Table 4 summarizes these wavelengths and spectra transformation methods. Tables 5 and 6 list key model performance indicators for each variable of prediction listed in Table 4 along with all PLSR models. Figure 10 points out selected wavelength using the average of all raw spectra.

It was observed that wavelengths near 1900 nm (short water absorbance band) transformed using either the first or second order derivative were found to be better correlated with the percentages of sand and clay content as compared to other parts of the spectra. Smooth and first SGD spectra corresponding to different visible wavelengths were superior when predicting SOM content.

From Tables 5 and 6, it is obvious that PLSR results had the lowest combination of MP and SEP with no overwhelming difference among the types of spectra transformations involved. In most cases, PLSR that was based on second SGD spectra indicated the best performance and the prediction soil properties did not improve when using all four methods of spectral transformation. This was an expected observation as PLSR uses several spectral factors and SLR just one; therefore, it is not capable of relating the difference in sensor response to a given soil property.

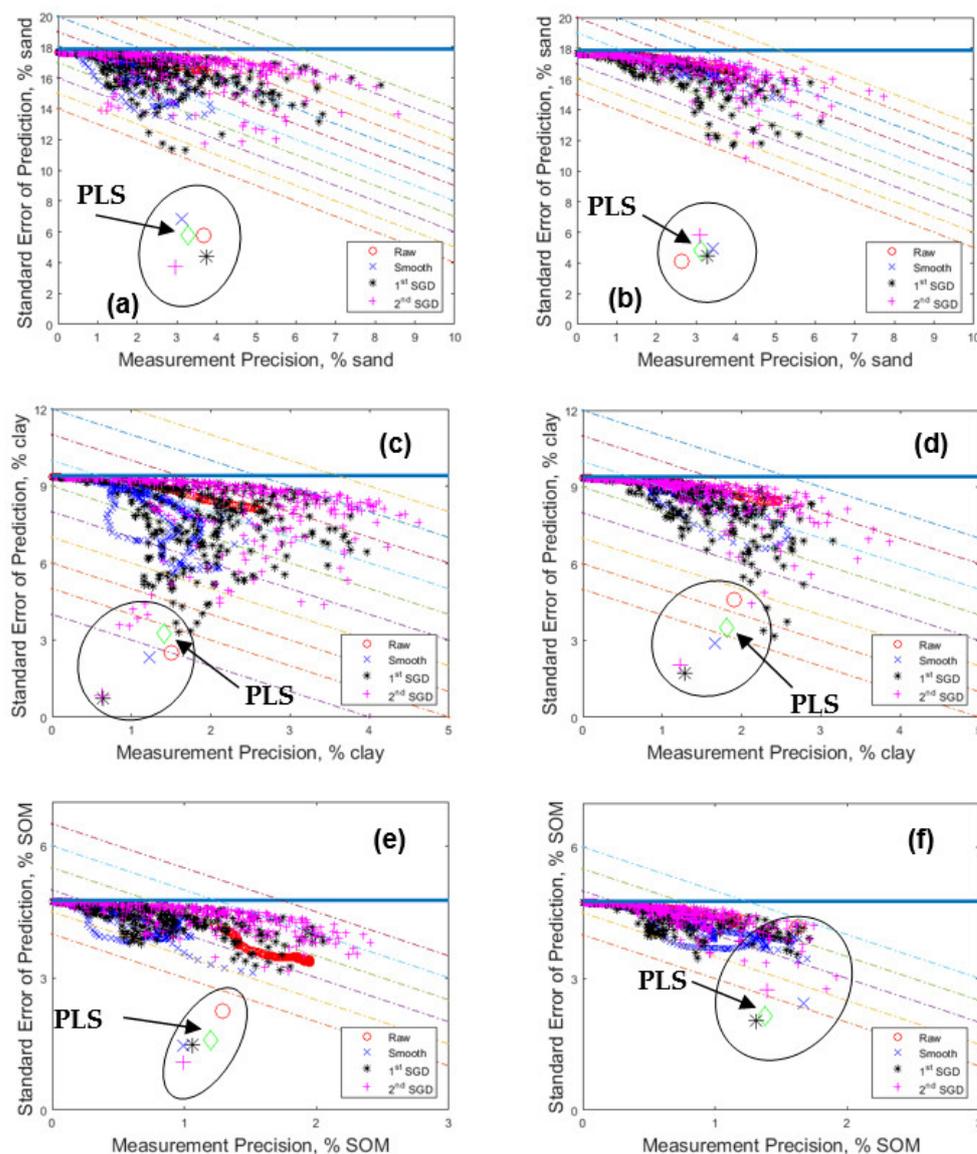
As illustrated in Figure 10 (and also summarized in Tables 5 and 6), when SLR models were applied to ex situ measurements, the best predictor wavelengths (i.e., single variable predictor while considering the sum of both precision (MP) and accuracy (SEP)) were found to be 1936 nm and 680 nm to correspond well with % sand and % SOM, respectively. However, the measurements needed to be transformed into the first SGD, whereas 1924 nm corresponded well with % clay only after transforming the measurements into the second SGD, which was similar to the results obtained when SLR models were applied to in situ measurements.

Then, the best predictors' wavelengths (i.e., single variable predictors while considering the sum of precision (MP) and accuracy (SEP)) were found to be both 1940 nm and 1747 nm, which corresponded well with % clay and % SOM, respectively. However, the measurements needed to be transformed into the first SGD, whereas 1462 nm corresponded well with % sand only after transforming the measurements into the second SGD.

In general, the results indicate that when using a single predicting factor, the SEP for sand was about 10–12% with no significant difference between in situ and ex situ measurements. The MPs of around 3% indicate relatively strong spectra reproducibility, which indicates the need for a more expanded model. This is apparent since the SEP of PLRS was two times smaller, indicating the ability to predict sand within a margin of 4%.

The percentage of clay, on the other hand, had 3–4% SEP and 1–2% MP, indicating both the reproducibility of the spectra and the ability of a single spectra parameter to accurately predict clay. In this case, PLSR also reduced the errors by two times, but only using first, or second, order derivative spectra.

Finally, the SEP for SOM was only a quarter lower than the standard deviation obtained in laboratory measurements [67], indicating that SLR is not an appropriate model for this soil property for the given set of soils. PLSR reduced SEP to about 1% SOM, which is reasonable considering the extensive spread of soil texture in this soil set. Similar to sand and clay, minor increases in error estimates corresponding to in situ versus ex situ measurements were much smaller than expected considering that different physical locations within a 0.5 m distance were used each time.



**Figure 9.** Precision versus accuracy comparison for the vis-NIR soil spectra, their derivatives and PLSR when predicting % sand, % clay, and % SOM using ex situ (a,c,e) and in situ (b,d,f) measurements. (Solid line represents the standard deviation of sand content (laboratory measurements); isolines represent the positions in the chart with the constant sum of MP and SEP; diamond symbol represents PLSR results on all spectra combined).

Altogether, these results are comparable with other reports pertaining to the use of DRS for the measurement of soil texture and SOM, as previously summarized in [6,16,19,20,24]. The relatively small number of locations means that the results for PLSR are rather optimistic and, once the multivariate regression approach is applied, the expected results should be between the present PLSR and the best SLR performance indicators. However, the most important finding in this study is that the majority of spectral measurements are sufficiently reproducible to be considered for distinguishing among diverse soil samples, while certain parts of the spectra indicate the capability to achieve this at  $\alpha = 0.05$ .

**Table 4.** vis–NIR wavelengths (nm) indicating the lowest critical error's when predicting % sand, % clay and SOM content.

| Soil Property | Test    | Raw | Smooth                      | 1st SGD <sup>a</sup>                               | 2nd SGD <sup>b</sup>     |
|---------------|---------|-----|-----------------------------|--|--------------------------|
| % sand        | Ex situ | -   | -                           | 1931 <sup>c</sup> , 1936 <sup>c</sup> , 1940, 1944 | 1413 <sup>c</sup>        |
|               | In situ | -   | -                           | 1452, 1457, 1931,1936, 1940, 1944                  | 1462 <sup>c</sup> , 1923 |
| % clay        | Ex situ | -   | -                           | 1940, 1944, 1948                                   | 1915, 1919, 1923         |
|               | In situ | -   | -                           | 1931, 1936, 1940 <sup>c</sup> , 1944 <sup>c</sup>  | 1923                     |
| % SOM         | Ex situ |     | 674 <sup>c</sup> , 679, 685 | 551, 556 <sup>c</sup> , 562, 568, 585              | -                        |
|               | In situ | -   | 421, 722, 728               | 421, 585, 590, 1747 <sup>c</sup> , 1751            | 1605, 1543               |

<sup>a</sup> First order Savitzky–Golay derivative. <sup>b</sup> Second order Savitzky–Golay derivative. <sup>c</sup> vis–NIR avelengths that are common between in situ and ex situ measurements.

**Table 5.** The vis–NIR soil spectral bands which were capable of exhibiting the highest possible degree in terms of the precision and accuracy components, when predicting % sand, % clay and SOM content during ex situ measurements.

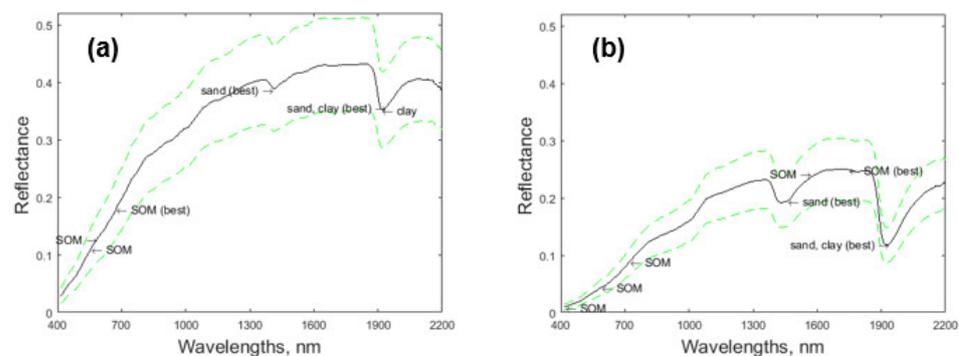
| Soil Property | Spectra              | Model             | Wavelengths, nm | RSE <sup>a</sup> | R <sup>2</sup> | MP <sup>b</sup> | SEP <sup>c</sup> |
|---------------|----------------------|-------------------|-----------------|------------------|----------------|-----------------|------------------|
| % sand        | 2nd SGD <sup>d</sup> | PLSR <sup>f</sup> | -               | -                | -              | 2.96            | 3.78             |
|               | 1st SGD <sup>e</sup> | SLR <sup>g</sup>  | 1931            | 4.26             | 0.59           | 3.22            | 11.31            |
|               | 1st SGD              | SLR               | 1936            | 4.77             | 0.58           | 2.87            | 11.40            |
|               | 2nd SGD              | SLR               | 1413            | 3.58             | 0.56           | 3.70            | 11.74            |
|               | 1st SGD              | SLR               | 1940            | 5.06             | 0.56           | 2.65            | 11.75            |
|               | 1st SGD              | SLR               | 1944            | 5.57             | 0.50           | 2.29            | 12.46            |
| % clay        | 1st SGD              | PLSR              | -               | -                | -              | 0.63            | 0.73             |
|               | 1st SGD              | SLR               | 1944            | 5.57             | 0.87           | 1.59            | 3.33             |
|               | 1st SGD              | SLR               | 1940            | 5.06             | 0.87           | 1.74            | 3.35             |
|               | 2nd SGD              | SLR               | 1920            | 9.06             | 0.86           | 0.98            | 3.55             |
|               | 2nd SGD              | SLR               | 1924            | 10.38            | 0.85           | 0.85            | 3.61             |
|               | 2nd SGD              | SLR               | 1915            | 8.07             | 0.83           | 1.08            | 3.86             |
| % SOM         | 2nd SGD              | PLSR              | -               | -                | -              | 0.99            | 1.09             |
|               | 1st SGD              | SLR               | 557             | 4.24             | 0.36           | 0.67            | 3.78             |
|               | 1st SGD              | SLR               | 551             | 4.86             | 0.35           | 0.59            | 3.79             |
|               | 1st SGD              | SLR               | 585             | 4.26             | 0.34           | 0.65            | 3.84             |
|               | 1st SGD              | SLR               | 563             | 4.46             | 0.33           | 0.62            | 3.86             |
|               | Smooth               | SLR               | 685             | 6.70             | 0.31           | 0.40            | 3.92             |
|               | 1st SGD              | SLR               | 568             | 4.89             | 0.31           | 0.54            | 3.93             |
|               | 1st SGD              | SLR               | 680             | 7.35             | 0.30           | 0.36            | 3.95             |
|               | Smooth               | SLR               | 674             | 7.67             | 0.29           | 0.34            | 3.99             |

<sup>a</sup> Ratio of Spread over error. <sup>b</sup> Measurement precision. <sup>c</sup> Standard error of prediction. <sup>d</sup> Second order Savitzky–Golay derivative. <sup>e</sup> First order Savitzky–Golay derivative. <sup>f</sup> Partial least squares regression (PLSR) calibration model. <sup>g</sup> Simple linear regression (PLSR) model.

**Table 6.** Summary of vis–NIR factors demonstrating reproducibility, precision and accuracy during in situ measurements.

| Soil Property | Spectra              | Model             | Wavelengths, nm | RSE <sup>a</sup> | R <sup>2</sup> | MP <sup>b</sup> | SEP <sup>c</sup> |
|---------------|----------------------|-------------------|-----------------|------------------|----------------|-----------------|------------------|
| % sand        | Raw                  | PLSR <sup>f</sup> | -               | -                | -              | 2.64            | 4.09             |
|               | 2nd SGD <sup>d</sup> | SLR <sup>g</sup>  | 1462            | 1.90             | 0.49           | 4.25            | 10.86            |
|               | 1st SGD <sup>e</sup> | SLR               | 1452            | 3.75             | 0.40           | 3.93            | 11.67            |
|               | 1st SGD              | SLR               | 1944            | 2.34             | 0.50           | 3.96            | 11.83            |
|               | 2nd SGD              | SLR               | 1924            | 2.17             | 0.39           | 3.35            | 12.06            |
|               | 1st SGD              | SLR               | 1940            | 2.74             | 0.56           | 3.54            | 12.13            |
|               | 1st SGD              | SLR               | 1936            | 3.30             | 0.58           | 3.31            | 12.33            |
|               | 1st SGD              | SLR               | 1457            | 4.31             | 0.36           | 3.50            | 12.35            |
| % clay        | 1st SGD              | PLSR              | -               | -                | -              | 1.28            | 1.73             |
|               | 1st SGD              | SLR               | 1940            | 2.74             | 0.87           | 2.41            | 3.17             |
|               | 1st SGD              | SLR               | 1936            | 3.30             | 0.85           | 2.28            | 3.38             |
|               | 1st SGD              | SLR               | 1944            | 2.34             | 0.87           | 2.58            | 3.75             |
|               | 1st SGD              | SLR               | 1932            | 3.64             | 0.78           | 2.32            | 4.28             |
|               | 2nd SGD              | SLR               | 1924            | 2.17             | 0.85           | 2.12            | 4.46             |
|               | % SOM                | 1st SGD           | PLSR            | -                | -              | -               | 1.31             |
| 1st SGD       |                      | SLR               | 1747            | 3.07             | 0.00           | 0.86            | 3.46             |
| 2nd SGD       |                      | SLR               | 1605            | 1.67             | 0.01           | 0.96            | 3.57             |
| 1st SGD       |                      | SLR               | 1752            | 2.67             | 0.00           | 0.85            | 3.59             |
| 1st SGD       |                      | SLR               | 1544            | 1.39             | 0.00           | 0.96            | 3.59             |
| Smooth        |                      | SLR               | 422             | 2.77             | 0.28           | 0.74            | 3.76             |
| Smooth        |                      | SLR               | 728             | 3.87             | 0.33           | 0.72            | 3.78             |
| Smooth        |                      | SLR               | 723             | 3.97             | 0.34           | 0.69            | 3.82             |
| 1st SGD       |                      | SLR               | 591             | 4.60             | 0.00           | 0.48            | 3.98             |
| 1st SGD       |                      | SLR               | 422             | 2.80             | 0.00           | 0.51            | 3.99             |
| 1st SGD       |                      | SLR               | 585             | 5.07             | 0.00           | 0.52            | 3.99             |

<sup>a</sup> Ratio of Spread over error. <sup>b</sup> Measurement precision. <sup>c</sup> Standard error of prediction. <sup>d</sup> Second order Savitzky–Golay derivative. <sup>e</sup> First order Savitzky–Golay derivative. <sup>f</sup> Partial least squares regression (PLSR) calibration model. <sup>g</sup> Simple linear regression (PLSR) model.

**Figure 10.** The identified regions from the vis–NIR soil spectra capable of exhibiting the highest possible degree in terms of the precision and accuracy component, when predicting % sand, % clay, and SOM in ex situ (a) and in situ (b) measurement environments.

Since SEP was typically at least two times greater than MP, it is clear that a simplistic SLR modeling method does not account for a substantial part of variability among different soil samples; this might be explained by the interactions among different soil properties as well as the fact that different soil properties cause different degrees of change in specific spectral parameters. This would imply that a suitable prediction model should involve multiple such parameters, yet their number should be relatively small to assure model applicability outside a given set of soil samples. This approach is appropriate when analyzing other soil properties and PSS systems geared to investigate the spatial heterogeneity of specific soil attributes. However, the goal and the purpose of this paper was to use the simplest method that focused on the reproducibility of soil reflectance spectral elements and to assess direct relationships with soil attributes rather than model development [38–41,68]. The latter is the subject of follow-up work.

#### 4. Conclusions

Based on the results of this study, it can be concluded that vis-NIR spectral measurements obtained using the commercially available portable equipment are compatible in terms of precision and accuracy when they are used under laboratory conditions or in the field, and when soil water content varies substantially. Data processing by performing smoothing, and first and second order derivative transformation, of the spectra has increased measurement reproducibility, which enhanced the system's ability to separate different soil samples. This was clearly noticeable after applying the innovative RSE calculation method on each individual spectra and its smooth or derivative transformations. Naturally, this indicated the ability to predict soil properties using a simple SLR technique that defined the difference between soil samples. From the soil attributes tested using the SLR method, clay content was shown to be predictable using spectral data derivatives around the short water absorption band. The prediction error was higher for sand and unacceptably high for SOM. Use of the popular but complex PSLR technique reduced prediction errors by between a quarter and one third, and the results were found to be comparable with previous studies pertaining to the use of DRS for measuring soil texture and SOM. However, a relatively small number of soil samples require caution when inferring results in practice. The most important finding in this study is that the majority of vis-NIR spectral measurements are sufficiently reproducible to be considered for distinguishing among diverse soil samples, while certain parts of the spectra indicate the capability of achieving this at  $\alpha = 0.05$ . The methodology for evaluating both the precision and accuracy components of spectral measurement errors was found to be useful when assessing the performance of a given proximal soil sensing instrument.

**Author Contributions:** Key author, N.M.D.; writing—original draft preparation, N.M.D. and V.I.A.; writing—review and editing, N.M.D., V.I.A. and R.A.V.R.; conceptualization, V.I.A.; methodology, V.I.A.; funding acquisition, V.I.A.; study site expert, S.O.P. and V.I.A.; resources, V.I.A. and S.O.P.; soil scientist, R.A.V.R.; writing data analytic code, N.M.D.; data collection, N.M.D.; data preprocessing, N.M.D.; data processing, N.M.D.; supervision, V.I.A. and S.O.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported in part by funds provided through National Science and Engineering Research Council of Canada (NSERC) Discovery Grant, number RGPIN/402298-2012, as well as the Agriculture and Agri Food Canada (AAFC) Agricultural Greenhouse Gases Program (AGGP), grant number 1585-16-3-4-52 using infrastructure supported by the Canada Foundation for Innovation (CFI) Leader Opportunity Fund (LOF).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request.

**Acknowledgments:** The authors are grateful to Frederic Rene-Laforest, Luan Pan, Ahmad Suhaizi Mat Su, Marine Louargant, and Ian Chan for their assistance with data collection.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Gee, G.W.; Bauder, J.W. Particle-size Analysis. In *Methods of Soil Analysis, Part 1—Physical and Mineralogical Methods*; American Society of Agronomy: Madison, WI, USA, 1986; pp. 383–411.
2. Schulte, E.E.; Hopkins, B.G. Estimation of organic matter by weight loss-on-ignition. In *Soil Organic Matter: Analysis and Interpretation*; Special Publication No. 46; Magdoff, F.R., Tabatabai, M.A., Hanlon, E.A., Jr., Eds.; Soil Science Society of America: Madison, WI, USA, 1996; pp. 21–31.
3. Viscarra Rossel, R.V.; Adamchuk, V.I.; Sudduth, K.A.; McKenzie, N.; Lobsey, C. Proximal Soil Sensing: An Effective Approach for Soil Measurements in Space and Time. In *Advances in Agronomy*; Academic Press: Cambridge, MA, USA, 2011; Volume 113, pp. 243–291.
4. Adamchuk, V.; Ji, W.; Viscarra Rossel, R.V.; Gebbers, R.; Tremblay, N.; Shannon, D.; Clay, D.; Kitchen, N. Proximal Soil and Plant Sensing. In *ASA, CSSA, and SSSA Books*; Wiley: Hoboken, NJ, USA, 2018; pp. 119–140.

5. Coûteaux, M.-M.; Berg, B.; Rovira, P. Near infrared reflectance spectroscopy for determination of organic matter fractions including microbial biomass in coniferous forest soils. *Soil Biol. Biochem.* **2003**, *35*, 1587–1600. [[CrossRef](#)]
6. Viscarra, R.R.A.; Walvoort, D.J.J.; McBratney, A.B.; Janik, L.J.; Skjemstad, J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59–75. [[CrossRef](#)]
7. Disla, J.S.; Janik, L.J.; Viscarra Rossel, R.V.; Macdonald, L.; McLaughlin, M.J. The Performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties. *Appl. Spectrosc. Rev.* **2014**, *49*, 139–186. [[CrossRef](#)]
8. Dalal, R.; Henry, R. Simultaneous Determination of Moisture, Organic Carbon, and Total Nitrogen by Near Infrared Reflectance Spectrophotometry. *Soil Sci. Soc. Am. J.* **1986**, *50*, 120–123. [[CrossRef](#)]
9. McBratney, A.; Minasny, B.; Viscarra Rossel, R.V. Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. *Geoderma* **2006**, *136*, 272–278. [[CrossRef](#)]
10. Krishnan, P.; Alexander, J.D.; Butler, B.J.; Hummel, J.W. Reflectance Technique for Predicting Soil Organic Matter. *Soil Sci. Soc. Am. J.* **1980**, *44*, 1282–1285. [[CrossRef](#)]
11. Morra, M.J.; Hall, M.H.; Freeborn, L.L. Carbon and nitrogen analysis of soil fractions using near-infrared reflectance spectroscopy. *Soil Sci. Soc. Am. J.* **1991**, *55*, 288–291. [[CrossRef](#)]
12. Henderson, T.L.; Baumgardner, M.F.; Franzmeier, D.P.; Stott, D.; Coster, D.C. High Dimensional Reflectance Analysis of Soil Organic Matter. *Soil Sci. Soc. Am. J.* **1992**, *56*, 865–872. [[CrossRef](#)]
13. Sudduth, K.; Hummel, J.W. Portable, Near-infrared Spectrophotometer for Rapid Soil Analysis. *Trans. ASAE* **1993**, *36*, 185–193. [[CrossRef](#)]
14. Sudduth, K.; Hummel, J.W. Soil Organic Matter, CEC, and Moisture Sensing with a Portable NIR Spectrophotometer. *Trans. ASAE* **1993**, *36*, 1571–1582. [[CrossRef](#)]
15. Ben-Dor, E.; Banin, A. Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties. *Soil Sci. Soc. Am. J.* **1995**, *59*, 364–372. [[CrossRef](#)]
16. Chang, C.W.; Laird, D.A.; Mausbach, M.J.; Hurburgh, C.R. Near infrared reflectance spectroscopy-principal components regression analysis of soil properties. *Soil Sci. Soc. Am. J.* **2001**, *65*, 480–490. [[CrossRef](#)]
17. Chang, C.-W.; Laird, D.A. Near-infrared reflectance spectroscopic analysis of soil C and N. *Soil Sci.* **2002**, *167*, 110–116. [[CrossRef](#)]
18. Martin, P.D.; Malley, D.F.; Manning, G.; Fuller, L. Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. *Can. J. Soil Sci.* **2002**, *82*, 413–422. [[CrossRef](#)]
19. Shepherd, K.; Walsh, M. Development of reflectance spectral libraries for characterization of soil properties. *Soil Sci. Soc. Am. J.* **2002**, *66*, 988–998. [[CrossRef](#)]
20. Cozzolino, D.; Moron, A. The potential of near-infrared reflectance spectroscopy to analyze soil chemical and physical characteristics. *J. Agric. Sci.* **2003**, *140*, 65–71. [[CrossRef](#)]
21. McBratney, A.; Santos, M.M.; Minasny, B. On digital soil mapping. *Geoderma* **2003**, *117*, 3–52. [[CrossRef](#)]
22. Malley, D.F.; Martin, P.D.; Ben-Dor, E. Application in Analysis of Soils. In *Agronomy Monographs*; Wiley: Hoboken, NJ, USA, 2015; pp. 729–784.
23. Christy, C. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. *Comput. Electron. Agric.* **2008**, *61*, 10–19. [[CrossRef](#)]
24. Viscarra Rossel, R.V. The Soil Spectroscopy Group and the development of a global soil spectral library. *NIR News* **2009**, *20*, 14–15. [[CrossRef](#)]
25. Stenberg, B. Effects of soil sample pre-treatment's and standardized rewetting as interacted with sand classes on vis-NIR predictions of clay and soil organic carbon. *Geoderma* **2010**, *158*, 15–22. [[CrossRef](#)]
26. Hodge, A.M.; Sudduth, K.A. *Comparison of Two Spectrometers for Profile Soil Carbon Sensing*; Paper No. 121338240; ASABE: St. Joseph, MI, USA, 2012.
27. St. Luce, M.; Ziadi, N.; Nyiraneza, J.; Tremblay, G.; Zebarthc, B.; Whalen, J.; Laterrière, M. Near infrared reflectance spectroscopy prediction of soil nitrogen supply in humid temperate regions of Canada. *Soil Sci. Soc. Am. J.* **2012**, *76*, 1454–1461. [[CrossRef](#)]
28. Kodaira, M.; Shibusawa, K. Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping. *Geoderma* **2013**, *199*, 64–79. [[CrossRef](#)]
29. Piikki, K.; Wetterlind, J.; Söderström, M.; Stenberg, B. Three-dimensional digital soil mapping of agricultural fields by integration of multiple proximal sensor data obtained from different sensing methods. *Precis. Agric.* **2014**, *16*, 29–45. [[CrossRef](#)]
30. Dhawale, N.; Adamchuk, V.I.; Prasher, S.O.; Viscarra Rossel, R.V.; Ismail, A.A.; Kaur, J. Proximal soil sensing of soil texture and organic matter with a prototype portable mid-infrared spectrometer. *Eur. J. Soil Sci.* **2015**, *66*, 661–669. [[CrossRef](#)]
31. Forrester, S.T.; Janik, L.J.; Disla, J.S.; Mason, S.; Burkitt, L.; Moody, P.; Gourley, C.J.P.; McLaughlin, M.J. Use of handheld mid-infrared spectroscopy and partial least-squares regression for the prediction of the phosphorus buffering index in Australian soils. *Soil Res.* **2015**, *53*, 67–80. [[CrossRef](#)]
32. Ji, W.; Adamchuk, V.I.; Biswas, A.; Dhawale, N.; Sudarsan, B.; Zhang, Y.; Viscarra Rossel, R.V.; Shi, Z. Assessment of soil properties in situ using a prototype portable MIR spectrometer in two agricultural fields. *Biosyst. Eng.* **2016**, *152*, 14–27. [[CrossRef](#)]
33. Chen, S.; Li, S.; Ma, W.; Ji, W.; Xu, D.; Shi, Z.; Zhang, G. Rapid determination of soil classes in soil profiles using vis-NIR spectroscopy and multiple objectives mixed support vector classification. *Eur. J. Soil Sci.* **2019**, *70*, 42–53. [[CrossRef](#)]

34. Knadel, M.; de Jonge, L.W.; Tuller, M.; Rehman, H.U.; Jensen, P.W.; Moldrup, P.; Greve, M.H.; Arthur, E. Combining visible near-infrared spectroscopy and water vapor sorption for soil specific surface area estimation. *Vadose Zone J.* **2020**, *19*, 20007. [[CrossRef](#)]
35. Dhawale, N.M.; Adamchuk, V.; Viscarra, R.; Prasher, S.; Whalen, J.K.; Ismail, A. *Predicting Extractable Soil Phosphorus Using Visible/Near-Infrared Hyperspectral. Soil Reflectance Measurements*; Paper No. CSBE13-047; The Canadian Society for Bioengineering: Renfrew, ON, Canada, 2013.
36. Wetterlind, J.; Piikki, K.; Stenberg, B.; Söderström, M. Exploring the predictability of soil texture and organic matter content with a commercial integrated soil profiling tool. *Eur. J. Soil Sci.* **2015**, *66*, 631–638. [[CrossRef](#)]
37. Piikki, K.; Söderström, M. Digital soil mapping of arable land in Sweden—Validation of performance at multiple scales. *Geoderma* **2019**, *352*, 342–350. [[CrossRef](#)]
38. Ji, W.; Viscarra Rossel, R.A.; Shi, Z. Accounting for the effects of water and the environment on proximally sensed vis–NIR soil spectra and their calibrations. *Eur. J. Soil Sci.* **2015**, *66*, 555–565. [[CrossRef](#)]
39. Goldshleger, N.; Chudnovsky, A.; Ben-Dor, E. Using Reflectance Spectroscopy and Artificial Neural Network to Assess Water Infiltration Rate into the Soil Profile. *Appl. Environ. Soil Sci.* **2012**, *2012*, 1–9. [[CrossRef](#)]
40. Shen, Z.; Viscarra Rossel, R.A. Automated spectroscopic modelling with optimised convolutional neural networks. *Sci. Rep.* **2021**, *11*, 1–12. [[CrossRef](#)]
41. Morellos, A.; Pantazi, X.-E.; Moshou, D.; Alexandridis, T.; Whetton, R.; Tziotziou, G.; Wiebesohn, J.; Bill, R.; Mouazen, A. Machine Learning based Prediction of Soil Total Nitrogen, Organic Carbon and Moisture Content by Using vis–NIR Spectroscopy. *Biosyst. Eng.* **2016**, *152*, 104–116. [[CrossRef](#)]
42. Milton, E.J. Principles of field spectrometry. *Int. J. Remote. Sens.* **1987**, *8*, 1807–1827. [[CrossRef](#)]
43. Deering, D.W. Field measurements of bidirectional reflectance. In *Theory and Applications of Optical Remote Sensing*; Arsar, G., Ed.; Wiley: New York, NY, USA, 1989; pp. 14–61.
44. Rollin, E.M.; Milton, E.J.; Emery, D.R. Reference panel anisotropy and diffuse radiation—Some implications for field spectroscopy. *Int. J. Remote Sens.* **2000**, *21*, 2799–2810. [[CrossRef](#)]
45. Leite, V.D.J.; de Oliveira, E.C.; Aucélio, R.Q. Impact of the sampling process on the measurement uncertainty, a case study: Physicochemical parameters in diesel. *Accredit. Qual. Assur.* **2021**, *26*, 1–9. [[CrossRef](#)]
46. Chodak, M.; Ludwig, B.; Khanna, P.; Beese, F. Use of near infrared spectroscopy to determine biological and chemical characteristics of organic layers under spruce and beech stands. *J. Plant Nutr. Soil Sci.* **2002**, *165*, 27–33. [[CrossRef](#)]
47. Udelhoven, T.; Emmerling, C.; Jarmer, T. Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression: A feasibility study. *Plant Soil* **2003**, *251*, 319–329. [[CrossRef](#)]
48. Madari, B.E.; Reeves, J.B., III; Machado, P.L.O.A.; Guimaraes, C.M.; Torres, E.; McCarty, G.W. Mid and near-infrared spectroscopic assessment of soil compositional parameters and structural indices in two Ferralsols. *Geoderma* **2006**, *136*, 245–249. [[CrossRef](#)]
49. Nduwamungu, C.; Ziadi, N.; Parent, L.-É.; Tremblay, G.F.; Thuriès, L. Opportunities for, and limitations of, near infrared reflectance spectroscopy applications in soil analysis: A review. *Can. J. Soil Sci.* **2009**, *89*, 531–541. [[CrossRef](#)]
50. Holman, J.P. *Experimental Methods for Engineers*, 7th ed.; McGraw Hill: New York, NY, USA, 2001.
51. Milton, E.J.; Rollin, E.M.; Emery, D.R. Advances in field spectroscopy. In *Advances in Environmental Remote Sensing*; Danson, F.M., Plummer, S.E., Eds.; Wiley: Oxford, UK, 1995.
52. De Bièvre, P. The 2012 International Vocabulary of Metrology: "VIM". *Accredit. Qual. Assur.* **2012**, *17*, 231–232. [[CrossRef](#)]
53. Viscarra Rossel, R.A.; Webster, R. Predicting soil properties from the Australian soil visible–near infrared spectroscopic database. *Eur. J. Soil Sci.* **2012**, *63*, 848–860. [[CrossRef](#)]
54. Drongelen, W.V. Signal Averaging. In *Signal Processing for Neuroscientists*; Drongelen, W.V., Ed.; Academic Press: Cambridge, MA, USA, 2007; pp. 55–70.
55. *Soil Survey Division Staff. Soil Survey Manual*; USDA-NRCS; US Government Print Off: Washington, DC, USA, 1993; pp. 63–65.
56. Viscarra Rossel, R.A. ParLeS: Software for chemometric analysis of spectroscopic data. *Chemom. Intell. Lab. Syst.* **2008**, *90*, 72–83. [[CrossRef](#)]
57. Geladi, P.; MacDougall, D.; Martens, H. Linearization and Scatter-Correction for Near-Infrared Reflectance Spectra of Meat. *Appl. Spectrosc.* **1985**, *39*, 491–500. [[CrossRef](#)]
58. Savitzky, A.; Golay, M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [[CrossRef](#)]
59. Bellon-Maurel, V.; Fernandez-Ahumada, E.; Palagos, B.; Roger, J.-M.; McBratney, A. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *TrAC Trends Anal. Chem.* **2010**, *29*, 1073–1081. [[CrossRef](#)]
60. Adamchuk, V.I.; Morgan, M.T.; Brouder, S.M. Development of an On-the-go Soil pH Mapping Method: Analysis of Measurement Variability. *Appl. Eng. Agric.* **2006**, *22*, 223–344. [[CrossRef](#)]
61. Dhawale, N.M.; Adamchuk, V.I.; Prasher, S.O.; Viscarra, R.R.A.; Ismail, A.A. Analysis of the repeatability of soil spectral data obtained using different measurement techniques. In *Proceedings of the 3rd Global Workshop on Proximal Soil Sensing*, Potsdam, Germany, 26–29 May 2013; Gebbers, R., Ed.; ATB Leibniz-Institut für Agrartechnik Potsdam-Bornim: Potsdam, Germany, 2013; pp. 161–165.

- 
62. Li, X.; Ren, J.; Zhao, K.; Liang, Z. Correlation between Spectral Characteristics and Physicochemical Parameters of Soda-Saline Soils in Different States. *Remote Sens.* **2019**, *11*, 388. [[CrossRef](#)]
  63. Geladi, P.; Kowalski, B.R. Partial least-squares regression: A tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17. [[CrossRef](#)]
  64. De Jong, S.; Kiers, H. Principal covariates regression. *J. Chemom. Intell. Lab. Syst.* **1992**, *14*, 155–164. [[CrossRef](#)]
  65. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control.* **1974**, *19*, 716–723. [[CrossRef](#)]
  66. Martens, H.; Næs, T. *Multivariate Calibration*; Springer: Berlin/Heidelberg, Germany, 1984; pp. 147–156.
  67. Miller, R.O. A western evaluation of soil testing laboratory performance. *Better Crops* **2006**, *90*, 26–29.
  68. Vayssade, J.-A.; Paoli, J.-N.; Gée, C.; Jones, G. DeepIndices: Remote Sensing Indices Based on Approximation of Functions through Deep-Learning, Application to Uncalibrated Vegetation Images. *Remote Sens.* **2021**, *13*, 2261. [[CrossRef](#)]