



Article

Evaluation of a Micro-Electro Mechanical Systems Spectral Sensor for Soil Properties Estimation

Konstantinos Karyotis ^{1,*}, Theodora Angelopoulou ^{2,3} , Nikolaos Tziolas ² , Evgenia Palaiologou ², Nikiforos Samarinas ⁴ and George Zalidis ^{2,5}

- ¹ School of Science and Technology, International Hellenic University, 14th km Thessaloniki-N, Moudania, 57001 Thermi, Greece
 - ² Laboratory of Remote Sensing, Spectroscopy, and GIS, Department of Agriculture, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece; agtheodor@agro.auth.gr (T.A.); ntziolas@auth.gr (N.T.); epalaiolo@agro.auth.gr (E.P.); zalidis@agro.auth.gr (G.Z.)
 - ³ Centre of Research and Technology—Hellas (CERTH), Institute for Bio-Economy and Agri-Technology (iBO), Thessaloniki, 57001 Thermi, Greece
 - ⁴ Faculty of Engineering, School of Rural and Surveying Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece; smnikiforos@topo.auth.gr
 - ⁵ Interbalkan Environment Center, 18 Loutron Str., 57200 Lagadas, Greece
- * Correspondence: kkaryotis@ihu.edu.gr

Abstract: Soil properties estimation with the use of reflectance spectroscopy has met major advances over the last decades. Their non-destructive nature and their high accuracy capacity enabled a breakthrough in the efficiency of performing soil analysis against conventional laboratory techniques. As the need for rapid, low cost, and accurate soil properties' estimations increases, micro electro mechanical systems (MEMS) have been introduced and are becoming applicable for informed decision making in various domains. This work presents the assessment of a MEMS sensor (1750–2150 nm) in estimating clay and soil organic carbon (SOC) contents. The sensor was first tested under various experimental setups (different working distances and light intensities) through its similarity assessment (Spectral Angle Mapper) to the measurements of a spectroradiometer of the full 350–2500 nm range that was used as reference. MEMS performance was evaluated over spectra measured from 102 samples in laboratory conditions. Models' calibrations were performed using random forest (RF) and partial least squares regression (PLSR). The results provide insights that MEMS could be employed for soil properties estimation, since the RF model demonstrated solid performance over both clay ($R^2 = 0.85$) and SOC ($R^2 = 0.80$). These findings pave the way for supporting daily agriculture applications and land related policies through the exploration of a wider set of soil properties.

Keywords: clay; soil organic carbon; MEMS; soil spectroscopy; NIR; random forest; machine learning; SWIR



Citation: Karyotis, K.; Angelopoulou, T.; Tziolas, N.; Palaiologou, E.; Samarinas, N.; Zalidis, G. Evaluation of a Micro-Electro Mechanical Systems Spectral Sensor for Soil Properties Estimation. *Land* **2021**, *10*, 63. <https://doi.org/10.3390/land10010063>

Received: 17 December 2020
Accepted: 10 January 2021
Published: 13 January 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

The world is facing a new era in agriculture where technological advances are largely utilized to manage agricultural systems by assisting in the estimation of temporal and spatial changes in soil and crop production [1]. These advances that range from robotics and drones to computer vision software and sensor development have completely transformed modern agriculture. Precision agriculture or smart farming is highly connected with these technologies and along with Internet of Things (IoT) services are becoming the future of farming applications [2]. Walter et al. [3] highlighted that agriculture undergoes a fourth revolution due to the increasing use of information and communication technology that could enable the continuous monitoring of the farm through a reliable data flow.

However, not all types of data could be gathered effortlessly. Due to the spatial variation of soil attributes, soil analysis is a very important task for agricultural monitoring of soil fertility, yet it is time consuming and costly, while it is a destructive technique that uses

chemical resources and hence has environmental impact [4]. For that reason, scientists have moved towards finding reliable alternative solutions to this matter. Bowers and Hanks [5] did one of the first studies that assessed soil's reflectance and the factors that can influence it. Each part of the electromagnetic spectrum in the VNIR–SWIR region (400–2500 nm) can provide valuable information related to physicochemical properties of soils. Therefore, many studies have evaluated the potential of soil reflectance spectroscopy as an alternative to soil chemical analysis with increasingly accurate estimations [6] regarding several soil properties, i.e., total nitrogen [7,8], total phosphorus [9], and pH [10]. Two of the main soil properties that have been estimated with the use of soil reflectance spectroscopy are clay [10–12] and soil organic carbon (SOC) content [13–16]. Clay content plays an important role in cultivation practices, water, and nutrient retention among others and therefore to the general crop growth [17], while SOC is an important component of soil, affecting its vital functions, and also enhances resilience to climate change [18]. Both of these properties are considered significant soil quality indicators [19,20]. Clay minerals have demonstrated absorptions in the SWIR region (1400–3000 nm), which are mainly caused by vibrational transitions of molecules [21]. More specifically, it has been reported that the spectral regions around 1900 and 2200 nm could provide valuable information, as they are closely related to soil minerals [22,23]. It has been observed that SOC can also influence various regions of the electromagnetic spectrum. Specifically in the visible region due the change of soil color, and at regions between 1700–1900 nm and from 2100–2400 nm [24]. In view of the fact that soil reflectance spectroscopy in laboratory conditions and specifically in the VNIR–SWIR region has been largely evaluated, providing good predictions, large soil spectral libraries (SSL) are being created in many regions in order to cover a wider range of soil attributes in a qualitative and quantitative manner [25–27].

There are also many efforts for in situ applications, due to the need for fast and accurate information that will eventually be utilized by farmers [28–32]. Soil spectroscopy is more rapid than chemical laboratory soil analysis. However, VNIR–SWIR spectroradiometers are often too expensive and power demanding, which hinders their wide-scale deployment for user-friendly soil monitoring. Nowadays, there is a rapid development of wireless sensor networks that has triggered the design of low-cost and small sensor devices with IoT empowered as a feasible tool for informed decision-making in the domain of agriculture [33]. Specifically, the micro electro mechanical systems (MEMS) were first presented in the mid '80s [34]. MEMS are small portable devices, based on semiconductor microfabrication technologies. Over the recent years, MEMS have been inaugurated to the spectroscopy field through the release of numerous miniaturized spectral radiometers, covering various domains of the electromagnetic spectrum. These sensors have been utilized in the food industry to determine different allergens in powdered foods [35], for estimating the moisture content of natural fibers [36], for estimating milk ingredients [37], for the determination of rosmarinic acid content of dried and powdered *Rosmarini folium* [38], and for the detection of microplastics in soil [39]. To that end, the technological evolution of miniaturized NIR sensors that are already being used in different sectors could be evaluated for soil analysis purposes. The interest in their use for soil analysis lies in the fact that specific wavelengths hold information for various soil properties as highlighted above. Although these sensors have lower spectral and radiometric resolution, they could cover critical spectral ranges for soil properties estimation. However, to the authors' knowledge, there are only few studies using miniaturized sensors for soil attributes estimation [40,41], while preliminary experiments have been performed by commercial sensor providers, such as Spectral Engines.

The aim of this work was to evaluate the performance of the Spectral Engines S2.2 MEMS handheld sensor (1750–2150 nm) for SOC and clay content estimation at laboratory conditions. Firstly, the most suitable MEMS sensor configuration for reliable measurements was evaluated by testing different set ups, adjusting the height and the light intensity of the sensor in a subset dataset (six soil samples). These measurements were compared to reference measurements with a spectroradiometer with spectral range from 350–2500 nm.

In the next step, after selecting the optimal sensor configuration, measurements in the full dataset were performed. Finally, the results regarding both the most accurate measurement methodology, and the modelling evaluation for SOC and clay estimation using partial least squares regression (PLSR) and the random forests (RF) models are provided and discussed.

2. Materials and Methods

2.1. Sample Dataset

In this study, a set of 102 topsoil samples (0–30 cm) were utilized, which will be denoted as full dataset Figure 1. The full dataset was comprised of soil samples from two different regions. Specifically, the 45 topsoil samples were collected using a random soil sampling strategy that was adapted during the field campaign to specifically include bare soil samples from the broader rural area around lake Zazari, located in Western Macedonia, Greece. The remaining 57 samples were selected from the Republic of North Macedonia (MKD), part of GEO-CRADLE's initiative open SSL. For more details (e.g., sampling strategies and soil types), the reader is referred to Tziolas et al. [42]. The physical samples were sent to the Laboratory of Remote Sensing and Geographic Information Systems of the Aristotle University of Thessaloniki, Thessaloniki, Greece). The GEO-CRADLE SSL is in compliance with GEOSS data principles and Open Database License standards, and it is publicly available through the GEO-CRADLE regional data hub.

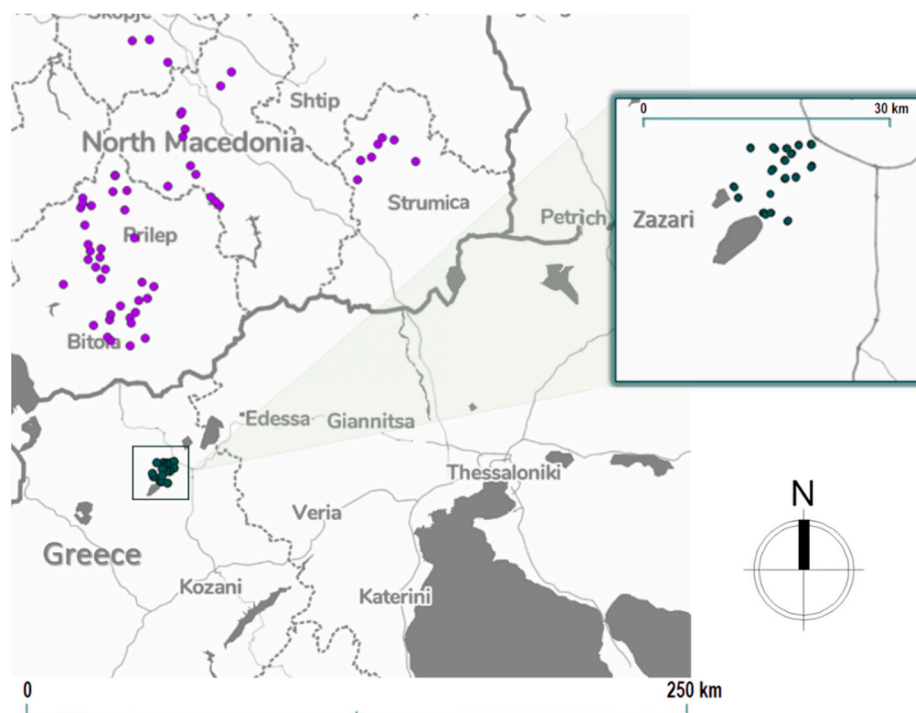


Figure 1. Spatial distribution of soil samples along the MKD and Zazari region (Greece).

The initial MKD part of GEO-CRADLE is a regional VNIR–SWIR SSL, comprised of the spectral signatures and their key soil properties of 124 topsoil samples across the MKD region. The conclusive 57 soil samples were selected from the full dataset, applying the Kennard Stone algorithm [43] over the predictor space, in order to achieve adequate variability of the target soil properties (clay and SOC) in the dataset. A sample dataset constituting of six indicative soil samples (hereinafter referred to as subset dataset, Figure 2) representing different soil classes was selected from the complete dataset and was utilized for the assessment the most suitable MEMS sensor configuration for the subsequent spectral measurements.

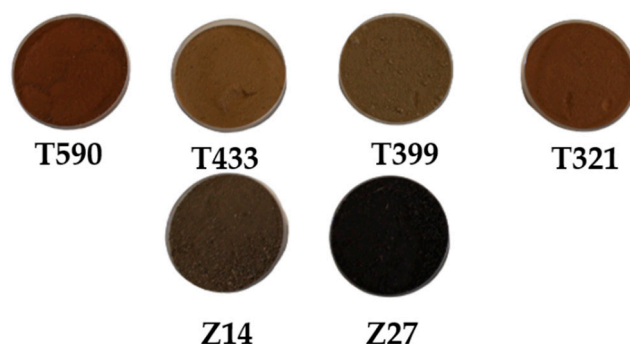


Figure 2. Subset of soil samples. Prefix Z denotes samples from Zazari region, while the T denotes samples from the region of the Republic of North Macedonia.

2.2. Sample Preparation and Chemical Analysis

The soil samples were carefully cleaned from possible vegetation, roots, stones, or other non-soil particles and air-dried. Each sample was crushed, passed through a 2 mm sieve, and then separated into two equal sub-samples. The first sub-sample was used for spectroscopic analysis, while the second was used for chemical analysis. The chemical analyses, in the case of the samples from the area of Zazari, were carried out in the laboratory of the Inter-Balkan Environment Center, while in the case of the samples from the area of the Republic of North Macedonia, they were taken from the open GEO-CRADLE database. The Vougioukos method [44] was used to determine the physicochemical parameters and specifically to determine the granulometric composition of the soil. The total organic carbon was determined via the Walkley–Black method [45].

2.3. Spectral Measurements

The spectral measurements were performed with two different sensors. The first one was the Spectral Evolution PSR+ 3500 instrument (Spectral Evolution Inc., Lawrence, MA, USA) covering the electromagnetic spectrum from 350 to 2500 nm. The spectral resolution of PSR+ 3500 is 2.8 nm @ 700 nm; 8 nm @ 1500 nm; 6 nm @ 2100 nm. PSR+ 3500 derived spectral signatures are graphical representations of one-dimensional functions, mapping the electromagnetic spectrum of VNIR–SWIR to the sensed reflectance. The second spectral device was a MEMS based S2.2 sensor (Spectral Engines OY, Helsinki, Finland) that employs a patented Fabry–Pérot interferometer and covers the spectral range from 1750 to 2150 nm, with wavelength interval of 5 nm. Two tungsten vacuum lamps compose the built-in illumination source. The device is compact with minimum dimensions ($25 \times 25 \times 17.5 \text{ mm}^3$), weighing 15 g, and is equipped with a 5 V re-chargeable battery supply, meeting portability characteristics and able to operate in situ and withstand ambient temperatures between 10 and 50 °C.

Spectral measurements are based on the principles of energy matter interactions. The values of the relative reflectance R_o of each sample are calculated as follows:

$$R_o = \frac{I_o - I_d}{I_w - I_d} \quad (1)$$

where I_o is the reflectance of the sample, I_d the reflectance of the dark measurement (which is sensed through using each device with the sensor's aperture closed), and I_w is the reflectance value of the respective white-reference material used for each sensor.

2.3.1. Reference Spectral Measurements

All measurements of the full dataset were performed inside a totally opaque dark box to prevent interferences with external light conditions and decrease temperature deviations. The steps of the followed protocol were performed by the following order: (i) measurement of a white reference panel (SpectralonTM) with over 98% reflectance in the

VNIR-SWIR region and 95% in the SWIR for irradiance to reflectance conversion, (ii) three measurements of two sand dunes Wylie Bay and Lucky Bay that were used as internal soil standards according to Ben Dor et al. [46], and (iii) three consecutive measurements from each soil sample that was rotated 90° each time and averaged for the formation of the spectral signature. This procedure was repeated after five soil sample measurements. The sensors were cross-calibrated through Wylie Bay and Lucky Bay fine earth samples that were used as diffusion gold standards [47], through the transformation:

$$C_{\lambda} = 1 - \left(\frac{(\rho_{S,\lambda} - \rho_{SBM,\lambda})}{\rho_{S,\lambda}} \right) \quad (2)$$

$$R_{c,\lambda} = R_{o,\lambda} \times R_{\lambda} \quad (3)$$

According to the above set of formulas, a correction coefficient is calculated for each sensed spectral band λ , $\rho_{S,\lambda}$ is the reflectance of standardization media as measured from each of both setups, with λ denoting the spectral band, $\rho_{SBM,\lambda}$ is the reference standard as defined from Commonwealth Scientific and Industrial Research Organization (CSIRO), and the standardized reflectance of the each sample will be the element-wise product of the vector containing the calculated correction factor R_{λ} with the sensed reflectance for each wavelength $R_{o,\lambda}$.

2.3.2. MEMS Sensor Configuration and Experimental Setup

Considering the lack of an established protocol for measuring soil samples with the MEMS sensor, ten different scenarios regarding the sensor's configuration were evaluated. Each scenario was conceptualized to test different instrumentation set-ups in order to select the optimal configuration. The adjusting parameters that were tested include the sensor's distance from the sensed surface and the device's light intensity. The device's aperture angle is 7° while the sensor's distance from its protective glass is 12 mm, resulting in a circular sensing area of 1.43 mm diameter that does not cover the minimum soil particle size limit of 2 mm of coarse sand [48]. By that means, robust and representative measurements could be questionable when it comes to non-homogeneous materials such as soil samples. Under that consideration, a mounting bracket that would raise the device from the sensing area and functionally expand the sensor's field of view to a circular area over 2 mm diameter was tested. Three cylinders at heights (H) of 15, 20, and 35 mm with interior surface of maximum reflectance were evaluated. The tested tubes increased the sensor's sensing area's diameter, resulting to diameters of 2.87, 3.3, and 3.9 mm, respectively. For each tube, the measurement was performed over three scenarios connected to Light Intensity (LI), which was set to 80%, 90%, and 100%, respectively. The experimental set up of the MEMS sensor along with the differences of the sensed area alterations according to height changes is shown in Figure 3.

The default sensor configuration was also assessed, i.e., direct contact of the soil sample with the sensor's lens and 100% LI. The above-mentioned measurements were performed at the subset dataset. The procedure of MEMS measurements included the exact same steps that were followed for PSR+ 3500 spectral measurements. The sensed spectral signatures were compared to the PSR+ 3500 reference spectral signatures in terms of the Spectral Angle Mapper similarity [49], and the optimal scoring methodology was selected for the modeling part. The Spectral Angle Mapper angle is calculated as:

$$a = \cos^{-1} \left(\frac{\sum_{\lambda=1}^n t_{\lambda} \rho_{\lambda}}{\sqrt{\sum_{i=1}^n t_{\lambda}^2} \sqrt{\sum_{i=1}^n \rho_{\lambda}^2}} \right) \quad (4)$$

where n the total count of spectral bands, t_{λ} the sensed reflectance at λ nm, and ρ_{λ} the corresponding reflectance of PSR+ 3500 measurements.

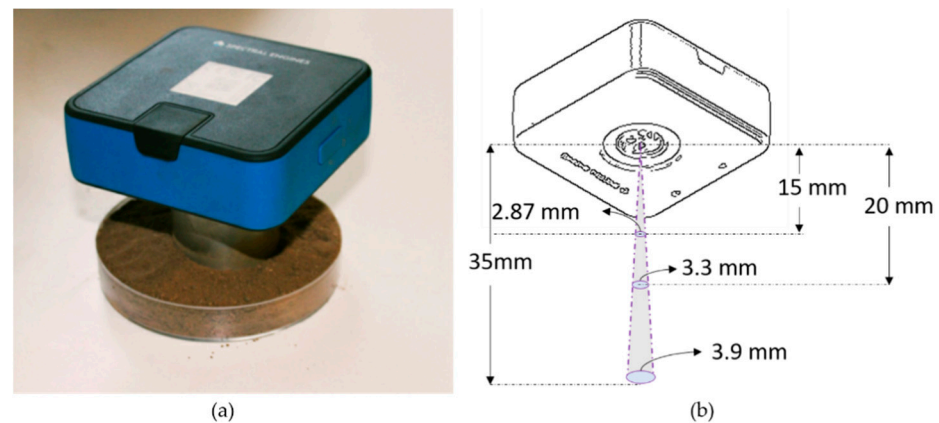


Figure 3. (a) MEMS measurement with the application of the 35 mm fitted tube. (b) MEMS' field of view corresponding diameter alterations for various heights.

2.3.3. MEMS Sensor Spectral Measurements

In order to perform the spectral measurements to the full dataset with the MEMS sensor, the best evaluated set up from the previous step was utilized. The followed procedure was according to the steps described in the previous section, i.e., white reference measurement, internal soil standards measurements, and finally soil sample measurements.

2.4. Spectral Data Preprocessing

A common practice in soil spectroscopy is to test different pre-processing techniques, which may help to enhance the absorption bands or perform scatter correction. In that regard, reflectance (R) measurements were transformed to absorbance values through log transformation ($\log_{10}(1/R)$). Then, the Savitzky–Golay algorithm [50] was performed to apply local polynomial regression on the first derivative of absorbance spectra. Through the first derivative, any possible baseline signal was removed [51], and a third degree polynomial was fitted through Partial Least Squares to spectral neighborhoods with 101 nm width for full PSR+ 3500 measurements and 21 nm for MEMS with the help of the Savitzky–Golay algorithm. The conclusive part of spectral preprocessing includes the range scaling, which was conducted through the application of z -transformation with the standard normal variates (SNV) technique [52]. Range scaling aims to reduce multiplicative effect of light scattering through centering and scaling, resulting in an average value of 0 and standard deviation of 1 for each spectral band, respectively. Reflectance measurements along with the preprocessing transformations are shown in Figure 4. PSR+ 3500 sensor presents higher sensitivity in reflectance sensing according to Figure 4c compared to MEMS, since it presents smoother changes for the first derivative curve. Due to this sensor characteristic, along with the lower spectral resolution, the reflectance's variability of each sensed wavelength is lower for the MEMS sensor.

2.5. Model Calibration and Validation

Two modeling iterations were conducted for both devices and each of the target chemical properties, one by fitting RF algorithm and other by PLSR method.

RF has been proven as an effective prediction algorithm in various domains [53,54]. RF is a classifier constituted from an ensemble of tree-based classifiers, $RF = \{T(x, \Theta_1), T(x, \Theta_2), \dots, T(x, \Theta_k)\}$, where k is the amount of trees and Θ_k independent and identically distributed vectors that act as the class recommendation of each tree for input x [55]. The RF model is tuned from three parameters, which are the number of variables that are used for the development of each tree and ranges from 1 to the total number of variables; the number of trees developed (if it is greater than 500 there is not statistical significant difference on the results) [56]; and the minimum node size, which will determine the minimum size of each intermediate node and the resulting leaves. The value selection of the minimum node size alters the role of RF in terms of producing coarser outputs for size

alterations; thus, for minimum node size equal to 1, the RF acts as a classifier, while for larger values, it performs regression results [57].

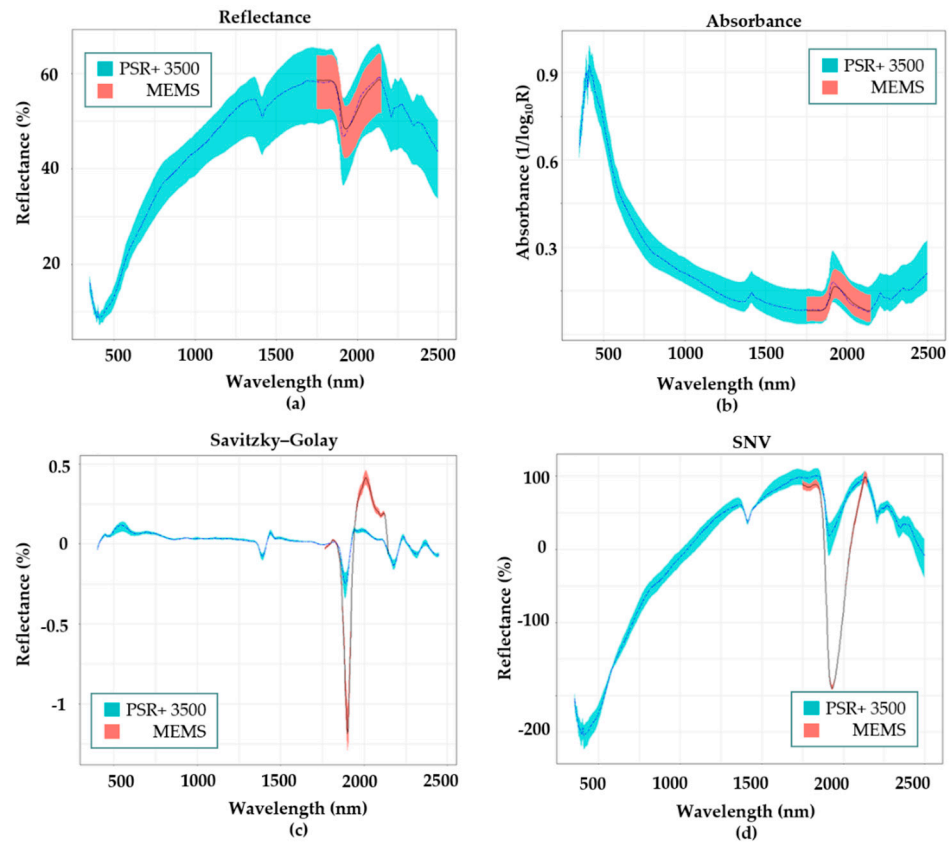


Figure 4. (a) Median reflectance measurements of complete dataset with PSR+ 3500 (with blue) and MEMS (red) ranging from first to third quartile. (b) Absorbance transformation. (c) First derivative transformation combined to third degree polynomial fitting. (d) Standard Normal Variates normalization.

According to Wold et al. [58], PLSR is one of the most used approaches in estimating chemical properties values through multivariable linear modeling due to its simplicity without performance decrease tradeoff. PLSR is capable of analyzing strongly correlated multivariate data by constructing factors of latent variables, aiming to explain the covariance between the predictor variables and their responses.

The accuracy assessment of the iterated models was performed through the calculation of standard modeling metrics between the measured values y and the estimated values \hat{y} for the metrics of coefficient of determination (R^2), root mean squared error (RMSE), and ratio of performance to InterQuartile distance (RPIQ) that are given by:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (5)$$

$$\text{RMSE} = \sqrt{\frac{\sum_i (y_i - \hat{y}_i)^2}{N}} \quad (6)$$

$$\text{RPIQ} = \frac{\text{Interquartile distance}}{\text{RMSE}} \quad (7)$$

Due to the underlying data size, leave-one-out cross validation (LOOCV) was applied for calibration and validation dataset splitting, resulting in increased model's accuracy in terms of the abovementioned metrics and error unbiasedness. Data preparation and pre-

processing was performed under R base package, while the modelling part was completed with Caret [59].

3. Results and Discussion

3.1. Physicochemical Analysis of Soil Samples

The results from the physicochemical analyses show an average clay content of 34.81%, while the average SOC content is 1.33%, which is typical for the Southern Balkan region soils that have low SOC content Table 1.

Table 1. Descriptive statistics of clay and SOC content (n = number of samples).

Region	n	Clay%				SOC%			
		Min	Max	Mean	Std	Min	Max	Mean	Std
Zazari	57	5.00	57.90	27.56	16.41	0.52	5.00	1.77	1.27
MKD	45	16.80	89.70	40.53	15.27	0.37	1.85	0.96	0.30
Total	102	5.00	89.70	34.81	16.99	0.37	5.00	1.33	0.96

3.2. MEMS Sensor Optimal Configuration

The results of the spectral analysis of the six samples constituting the subset dataset displayed distinguishable trends and characteristics for each sample. A visual inspection of the different set ups regarding the heights and light intensities was initially performed. As shown in Figure 5, the different sensor set ups resulted in differences at the samples spectral signatures. More specifically, it is presented how different LI (100%, 90%, and 80%) affect the spectral signatures at a certain height, i.e., 15, 20, and 35 mm. It was observed that the greater differences from the reference measured were at minimum LI, and hence, that is a very important factor during the measurements. Regarding the comparison of the default MEMS sensor set up and the PSR+ 3500 measurement, it was observed that the reflectance of the latter was significantly lower (Figure 5d). The signal-to-noise ratio representing the electromagnetic radiation reflected from soil samples is positively correlated to the LI, while it is negatively correlated to the distance between the sensor and the target. This comes in accordance to Beer–Lambert’s Law, stating that the radiation that passes through an area per time, decreases exponentially due to its conversion to Joule heating on its way to the absorbing medium. This change can be described from the formula:

$$L_i(x) = L_i(0) \times e^{-Ax} \quad (8)$$

where L_i stands for the light intensity, x is the length of the medium between the light source and the target, and A is the medium’s absorbance coefficient.

To quantitatively assess the spectral similarity of the spectra produced from the MEMS sensor against the PSR+ 3500 reference measurements, the Spectral Angle Mapper was utilized. The results with the setup of 20 mm height and 100% light intensity showed the highest similarity to PSR+ 3500 measurements, compared to the rest scenarios, in terms of minimum Spectral Angle Mapper distance, as shown in Figure 6. The default setup (proximal) that implies 100% LI and zero distance between the sensor and the sample, which is suggested from the manufacturer, presented the least optimal sensor configuration, with an average spectral distance of 0.052 rad. These findings demonstrate that field of view plays an important role on the sensor’s accuracy and verified the initial hypothesis indicating that in order to cover the minimum soil particle size limit, the sensor’s field of view must be increased. This was achieved through the increase of distance between the soil sample and the device. The other nine evaluated scenarios presented more accurate measurements in comparison to the default. More specifically, the order from the less accurate scenario to the most accurate is shown in Table 2. To that end, the complete dataset was measured in respect of the insights derived from the preliminary spectral analysis, setting the light intensity to 100% and using the cylinder of 20 mm height.

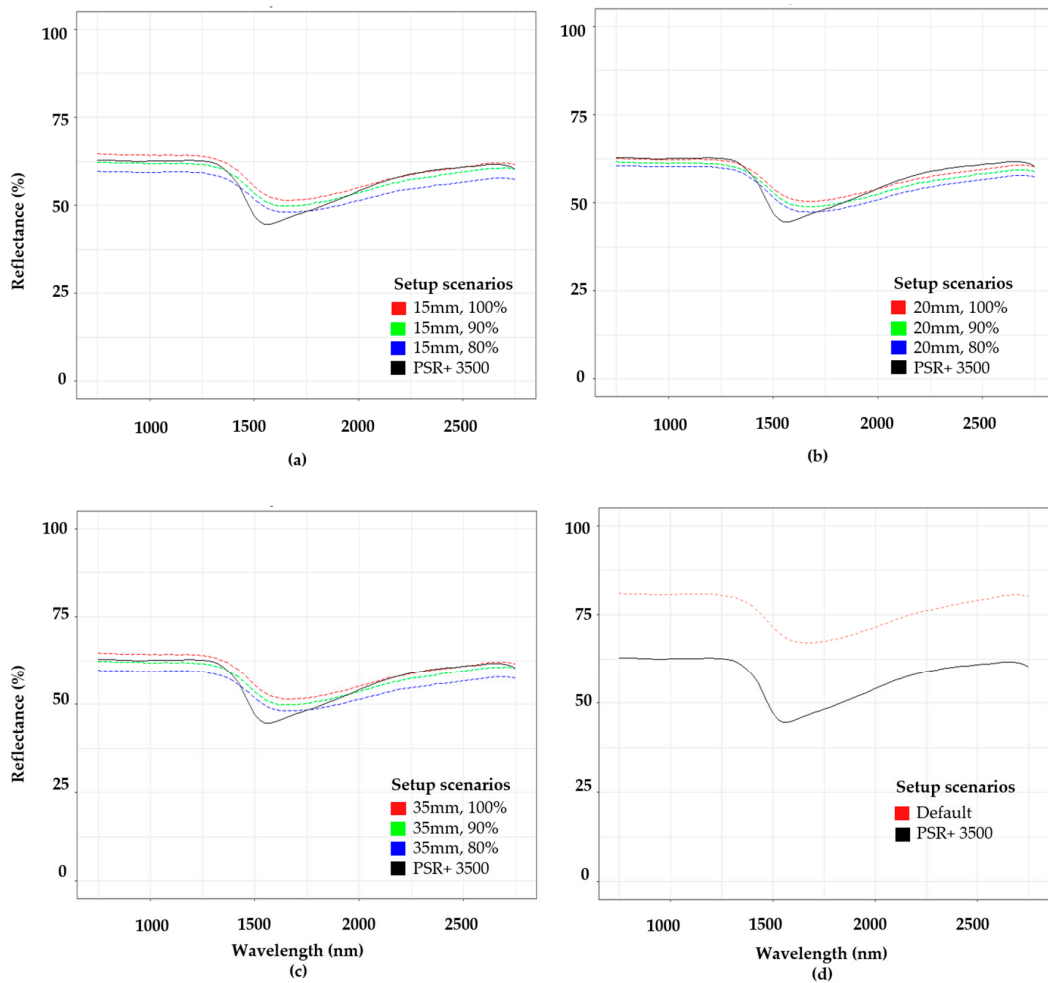


Figure 5. Median reflectance of subset dataset soil samples over ten different scenarios. A black solid line in each figure denotes the PSR+ 3500 median reflectance values. (a) Reflectance percentage for 15 mm height and three different light intensities. (b) Reflectance values with 20 mm distance from soil sample to MEMS sensor device and the three tested light intensities. (c) Reflectance values for 35 mm height from soil sample and the three tested light intensities. (d) Default setup (proximal sensed samples with 100% light intensity).

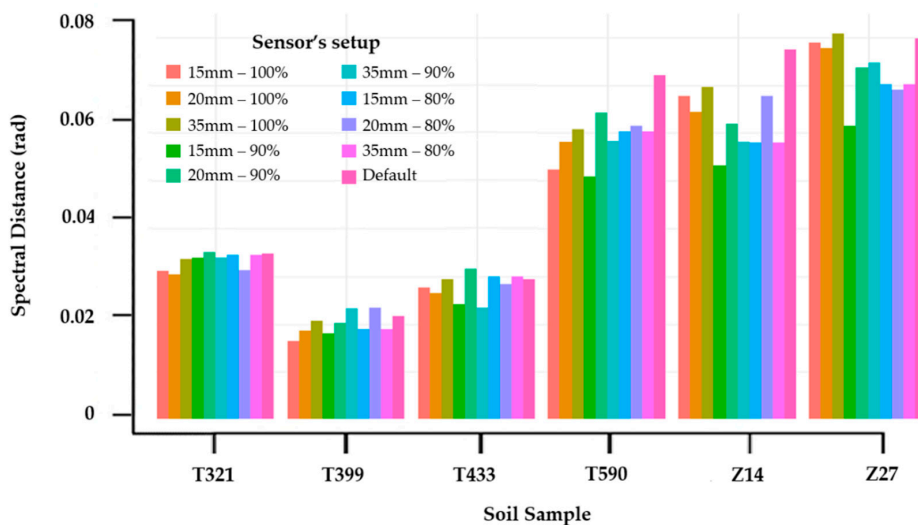


Figure 6. Angular distance calculated from Spectral Angle Mapper for each sample of the subset dataset for ten measuring scenarios. The used reference values are the spectral signatures obtained from PSR+ 3500, restricted to the 1750–2150 nm range.

Table 2. Classification of evaluated scenarios according to their performance, evaluated with Spectral Angle Mapper (H: height and LI: light intensity).

Scenario	Spectral Angle Mapper (rad)
Default set up	0.052
H: 15 mm, LI: 90%	0.049
H: 20 mm, LI: 80%	0.048
H: 35 mm, LI: 80%	0.047
H: 15 mm, LI: 80%	0.046
H: 15 mm, LI: 100%	0.046
H: 20 mm, LI: 90%	0.045
H: 35 mm, LI: 100%	0.045
H: 35 mm, LI: 90%	0.045
H: 20 mm, LI: 100%	0.040

3.3. Reference Spectral Measurements Compared to The MEMS Sensor Spectral Measurements

The calibration models were generated from the use of the RF classifier and the PLSR statistical method, while the preprocessing methods included the transformation from absorbance to reflectance, the Savitzky–Golay first derivative, and the SNV transformation, which were all applied in both calibration models. The results of the full dataset with LOOCV presented higher metric values with the use of the PSR+ 3500 sensor and the RF algorithm with $R^2 = 0.93$ and $RMSE = 6.97\%$ for clay content and $R^2 = 0.91$ and $RMSE = 0.32\%$ for SOC content (Table 3). The PLSR showed less accuracy, supporting the hypothesis that it is more efficient in predicting linear relationships. However, in the case of soil analysis, the relationship between soil properties and spectral data are usually presented as nonlinear [60], and this can explain the low predictive performance of PLSR. In contrast, RF could describe complex linear and nonlinear relationships and interactions while it reduces the probability of model overfitting. High accuracy for clay content prediction was also reported by Waiser et al. [28] and [61], with R^2 ranging from 0.85–0.88 values. However, the results are in contrast with the studies of [13,62], who reported better prediction models with the use of PLSR for clay content rather than SOC. In a study using 1534 legacy soil samples from Brazil, Ref. [63] noted that PLSR and RF models had similar predictive accuracies, however inferior to this study. This could be attributed to the local character of the present study and the smaller number of soil samples. Regarding SOC predictions, the results are similar to [14,56], who reported better results for the RF approach against the PLSR. However, Ref. [64] found that PLSR outperformed the use of RF, which highlights that differences between studies regarding the size of the dataset, and the range of the values of the modelled properties could significantly alter the accuracy of prediction. The respective scatter plots are shown in Figure 7.

Table 3. Evaluation metrics for the fitted models for Clay and SOC content.

Sensor	Method	Clay %			SOC %		
		R^2	RMSE	RPIQ	R^2	RMSE	RPIQ
PSR+ 3500	RF	0.93	6.97	2.03	0.91	0.32	1.31
MEMS	RF	0.85	8.17	1.84	0.80	0.46	1.47
PSR+ 3500	PLSR	0.70	8.17	2.11	0.88	0.29	2.65
MEMS	PLSR	0.31	14.12	0.95	0.38	0.94	1.06

The variable importance analysis plot represents the most important wavelengths selected from the RF calibration models (Figure 8). It could be observed that in most cases, the respective wavelengths of the PSR+ 3500 correspond with the MEMS sensor. Clay minerals in general present distinctive absorptions in the NIR–SWIR region. Regarding the PSR+ 3500 sensors, significant wavelengths above 2200 indicate the presence of clay minerals such as Kaolinite, which is in agreement with Lee et al. [23], that have also

demonstrated the following specific wavelengths as important for clay estimation (1904, 2177, 2201, 2213, 2321, 2492 nm). The narrow absorption bands near the 1900 nm region result from the presence of hydroxyls and water, vermiculite [63], and montmorillonitic clay [64]. Regarding SOC, RF variable importance analysis of PSR+ 3500 full spectrum indicated that the regions of the most significant wavelengths were found mostly centered in the visible region. This probably occurs due to the effect that SOC have in the soil's color as a result of absorptions from chromophores, as the literature review reveals [5,65,66]. More specifically, it has been observed that the yellow and red colors of the soil are due to the presence of hematite and goethite [67]. SOC in general is closely related to NIR wavelengths indicating C–O, C=O, and N–H compounds, while the 1852, 1930, 2033, 2060, and 2208 nm were indicated fundamental regions by Viscarra Rossel et al. [64]. Shi et al. [68] also reported the significance of the 1800–2450 nm region that corresponds with the part of the spectral region of the used MEMS sensor. The above suggest that the use of a MEMS spectral sensor is property related, and therefore there is a need for careful selection of its spectral range.

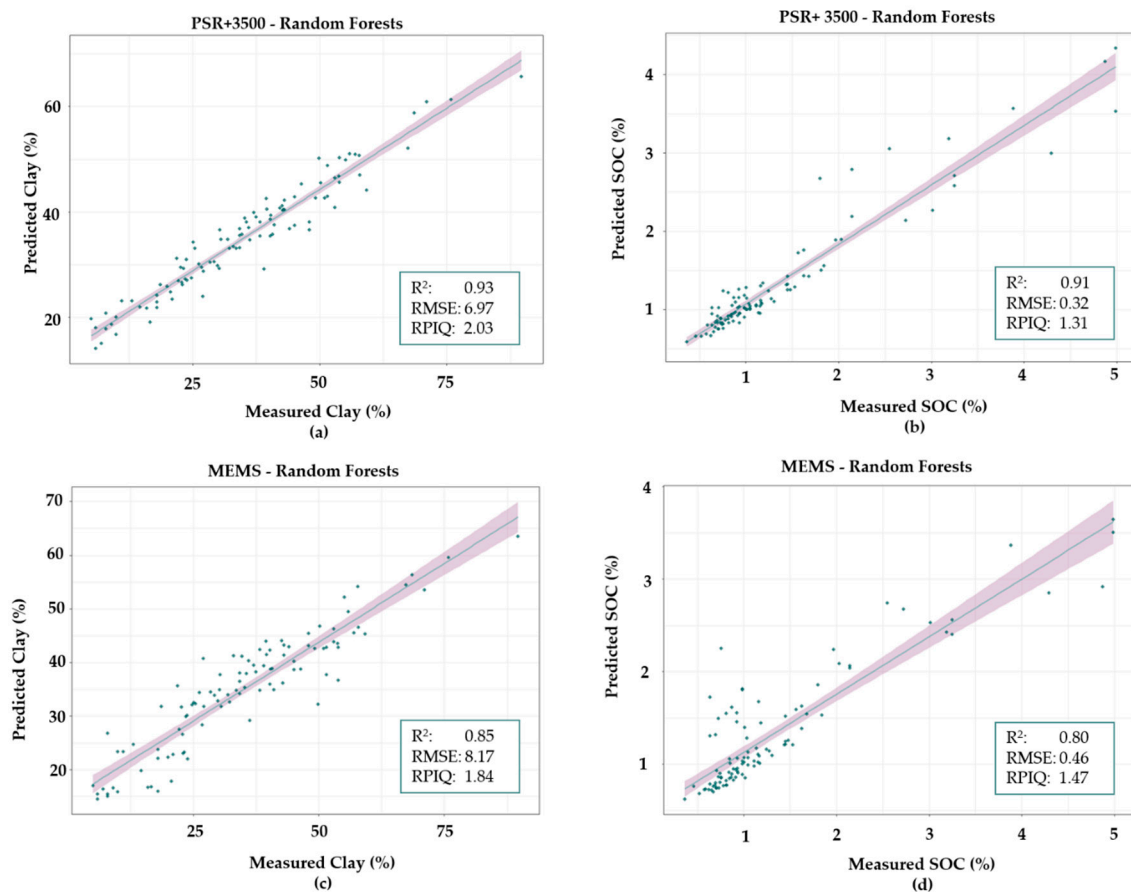


Figure 7. Predicted versus measured soil attributes scatterplots for RF fitted models using different sensors. PSR+ 3500 (a,b), and MEMS (c,d).

Comparing the results between both sensors, it can be observed that their accuracies are very close. The RF model gave an $R^2 = 0.85$ and RMSE = 8.17% for clay content, while for SOC content, the results were slightly lower with $R^2 = 0.80$ and RMSE = 0.46%. This could be attributed to the selection of a suitable spectral region of MEMS sensor for specifically estimating clay and SOC. However, the PLSR results of the MEMS sensor were significantly lower compared to all the models. The results for the models based on PSR+ 3500 full spectrum are in complete accordance with [69], where a single dimensional multi-channel convolutional neural network was applied to VNIR–SWIR spectra for the estimation of

clay and SOC content, with the result of R^2 of 0.86 in both cases. Viscarra Rossel et al. [70] reported similar prediction accuracies for both clay and SOC, using only the NIR region of the electromagnetic spectrum that supports the choice of the respective MEMS sensor for the specific soil attributes. Moreover, the findings of the current study demonstrate the potential of miniaturized spectrometers with reduced wavelength ranges compared to other alternatives, as demonstrated by Tang et al. [40], with similar or even better predictive performance. This can be explained by the choice of RF for the development of regression spectroscopic models.

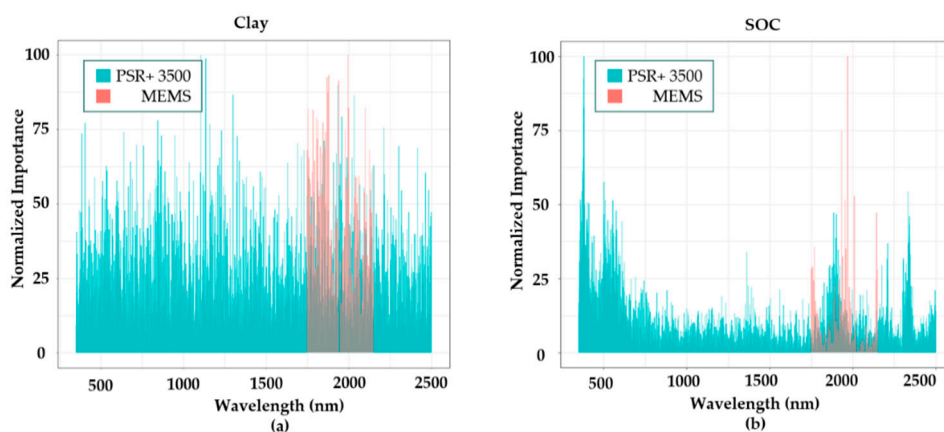


Figure 8. Normalized variable importance scores for each wavelength based on RF modeling. (a) Clay; (b) SOC.

Despite that, the lower accuracy of the MEMS sensors compared to the PSR+ 3500 is presumable due to the smaller spectral range and spectral resolution and hence missing specific wavelengths that could enhance model's predictive performance. As it is shown from various studies mentioned above, both the visible spectral region and wavelengths above 2200 nm can hold additional valuable information that could result in better prediction models.

The evaluated device meets all portability characteristics; thus, it can be the basis for a set of in situ soil spectroscopy applications and services. It is noteworthy that a remarkable success of collecting data in this way can reflect widespread interest for daily farming applications (e.g., fertilization by farmers and/or agri-consultants) and for checking of farming compliance to various international policies and treaties. However, there are limitations that need to be surpassed in order to acquire representative spectral measurements that are not taken under laboratory conditions, like the absence of developed standards and protocols for measurements and sample preparation. In this work, diffuse reflectance analysis was performed according to the standardization protocol proposed by Ben Dor et al. [46] that only concerns reflectance measurements in the laboratory, indicating that the inter-calibration of MEMS and PSR+ 3500 was achieved with the help of standardization through Wylie Bay and Lucky Bay fine earth samples and their standard measurements provided from CSIRO. These standardization media, due to their nature, present characteristics that may be impractical for in situ use (i.e., hard to maintain pure). To that end, there is a need for further assessment of different materials that could be alternatives of the internal soil standards.

4. Conclusions

The elaborated work provides insights that MEMS spectral sensors can support the global effort of assuring crop quality through monitoring soil and accurately estimating key properties playing the role of soil quality indices. The spectral signatures acquired with the MEMS sensor presented high similarity with corresponding PSR+ 3500 in the common spectral range, which encourages further research of soil properties that

are sensitive to vibrations due to electromagnetic radiation with wavelengths inside the 1750–2150 nm region.

Within the framework of this work, ten different measurement setups were evaluated including the default setup, over a sub-sample of six instances. The measurements were compared to PSR+ 3500 measurements, and Spectral Angle Mapper indicated that the MEMS sensor needs to operate with maximum light intensity and have 20 mm distance from soil sample. This is an aspect of MEMS that needs to be further assessed, and especially the capability to obtain in situ measurements of reflectance of rough soil under high scatter effect due to the existence of aggregates with size larger than 2 mm.

The modeling part indicated that RF outperformed PLSR in both clay and SOC predictions, resulting in satisfactory values of accuracy metrics for the case of RF. On the other hand, PLSR modeling over MEMS signatures needs further research, since the device's coarse spectral resolution and lower spectral sensitivity presumably provided less information compared to PSR+ 3500 related to the studied properties. The fitted models revealed correlations between clay and reflectance at the regions of 1990 and 1880 nm and for SOC over the region of 1750, 2000, and 2100 nm that come in accordance with the absorption features of clay and SOC as reported in the literature.

Nonetheless, the need for the analysis of a broader set of soil samples with high spatial variability covering different soil classes is highlighted, under real field conditions. In addition, the limited spectral range of the sensor constitutes that a sensor could mainly be used for specific soil properties estimation. Moreover, this novel approach needs to be tested in the field under real conditions (e.g., roughness), along with the full chain of interconnected systems, to pave the way for the development of tailored downstream services and associated tools aimed at the various stakeholders.

Author Contributions: Conceptualization: K.K., T.A., and G.Z.; methodology: T.A., K.K., E.P., and N.S.; software: K.K.; validation: K.K. and N.T.; formal analysis: T.A.; investigation: T.A., K.K., and E.P.; data curation: K.K.; writing—original draft preparation: K.K. and T.A.; writing—review and editing: N.T. and G.Z.; visualization: K.K.; supervision: G.Z.; and project administration: G.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly funded by the “DIONE” project under grant agreement No. 870378, supported by European Union's Horizon 2020 research and innovation program.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: <http://datahub.geocradle.eu/dataset/regional-soil-spectral-library>. Additional data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy restrictions.

Acknowledgments: The GEO-CRADLE topsoil soil spectral library used in this work was made available by the GEO-CRADLE initiative through the Regional Data Hub (<http://datahub.geocradle.eu/dataset/regional-soil-spectral-library>).

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

References

1. Salam, A.; Shah, S. Internet of things in smart agriculture: Enabling technologies. In Proceedings of the IEEE 5th World Forum on Internet of Things, WF-IoT 2019-Conference Proceedings, Limerick, Ireland, 15–18 April 2019; pp. 692–695.
2. Muangprathub, J.; Boonnam, N.; Kajornkasirat, S.; Lekbangpong, N.; Wanichsombat, A.; Nillaor, P. IoT and agriculture data analysis for smart farm. *Comput. Electron. Agric.* **2018**, *156*, 467–474. [[CrossRef](#)]
3. Walter, A.; Finger, R.; Huber, R.; Buchmann, N. Smart farming is key to developing sustainable agriculture. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6148–6150. [[CrossRef](#)] [[PubMed](#)]
4. Viscarra Rossel, R.A.; Adamchuk, V.I.; Sudduth, K.A.; McKenzie, N.J.; Lobsey, C. *Proximal Soil Sensing: An Effective Approach for Soil Measurements in Space and Time*; Academic Press: Cambridge, MA, USA, 2011; Volume 113, pp. 243–291.
5. Bowers, S.A.; Hanks, R.J. Reflection of Radiant Energy from Soils. *Soil Sci.* **1965**, *100*, 130–138. [[CrossRef](#)]
6. Nocita, M.; Stevens, A.; van Wesemael, B.; Aitkenhead, M.; Bachmann, M.; Barthès, B.; Ben Dor, E.; Brown, D.J.; Clairotte, M.; Csorba, A.; et al. Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring. *Adv. Agron.* **2015**, *132*, 139–159. [[CrossRef](#)]

7. Cozzolino, D.; Cynkar, W.U.; Damberg, R.G.; Shah, N.; Smith, P. In Situ Measurement of Soil Chemical Composition by Near-Infrared Spectroscopy: A Tool Toward Sustainable Vineyard Management. *Commun. Soil Sci. Plant Anal.* **2013**, *44*, 1610–1619. [[CrossRef](#)]
8. Wetterlind, J.; Stenberg, B.; Söderström, M. Increased sample point density in farm soil mapping by local calibration of visible and near infrared prediction models. *Geoderma* **2010**, *156*, 152–160. [[CrossRef](#)]
9. Heinze, S.; Vohland, M.; Joergensen, R.G.; Ludwig, B. Usefulness of near-infrared spectroscopy for the prediction of chemical and biological soil properties in different long-term experiments. *J. Plant Nutr. Soil Sci.* **2013**, *176*, 520–528. [[CrossRef](#)]
10. Rossel, R.A.V.; Behrens, T. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma* **2010**, *158*, 46–54. [[CrossRef](#)]
11. Piekarczyk, J.; Kazmierowski, C.; Krolewicz, S.; Cierniewski, J. Effects of Soil Surface Roughness on Soil Reflectance Measured in Laboratory and Outdoor Conditions. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 827–834. [[CrossRef](#)]
12. Jaconi, A.; Vos, C.; Don, A. Near infrared spectroscopy as an easy and precise method to estimate soil texture. *Geoderma* **2019**, *337*, 906–913. [[CrossRef](#)]
13. Moura-Bueno, J.M.; Dalmolin, R.S.D.; ten Caten, A.; Dotto, A.C.; Dematté, J.A.M. Stratification of a local VIS-NIR-SWIR spectral library by homogeneity criteria yields more accurate soil organic carbon predictions. *Geoderma* **2019**, *337*, 565–581. [[CrossRef](#)]
14. Gholizadeh, A.; Saberioon, M.; Carmon, N.; Boruvka, L.; Ben-Dor, E. Examining the Performance of PARACUDA-II Data-Mining Engine versus Selected Techniques to Model Soil Carbon from Reflectance Spectra. *Remote Sens.* **2018**, *10*, 1172. [[CrossRef](#)]
15. Padarian, J.; Minasny, B.; McBratney, A.B. Using deep learning to predict soil properties from regional spectral data. *Geoderma Reg.* **2019**, *16*, e00198. [[CrossRef](#)]
16. Angelopoulou, T.; Balafoutis, A.; Zalidis, G.; Bochtis, D. From laboratory to proximal sensing spectroscopy for soil organic carbon estimation—A review. *Sustainability* **2020**, *12*, 443. [[CrossRef](#)]
17. Churchman, G.J. Game Changer in Soil Science. Functional role of clay minerals in soil. *J. Plant Nutr. Soil Sci.* **2018**, *181*, 99–103. [[CrossRef](#)]
18. FAO. *Soil Organic Carbon the Hidden Potential*; FAO: Rome, Italy, 2017; ISBN 9789251096819.
19. Nabiollahi, K.; Golmohamadi, F.; Taghizadeh-Mehrjardi, R.; Kerry, R.; Davari, M. Assessing the effects of slope gradient and land use change on soil quality degradation through digital mapping of soil quality indices and soil loss rate. *Geoderma* **2018**, *318*, 16–28. [[CrossRef](#)]
20. Zeraatpisheh, M.; Bakhshandeh, E.; Hosseini, M.; Alavi, S.M. Assessing the effects of deforestation and intensive agriculture on the soil quality through digital soil mapping. *Geoderma* **2020**, *363*, 114139. [[CrossRef](#)]
21. Fang, Q.; Hong, H.; Zhao, L.; Kukulich, S.; Yin, K.; Wang, C. Visible and Near-Infrared Reflectance Spectroscopy for Investigating Soil Mineralogy: A Review. *J. Spectrosc.* **2018**, 1–14. [[CrossRef](#)]
22. Stenberg, B.; Viscarra Rossel, R.A.; Mouazen, A.M.; Wetterlind, J. Visible and near infrared spectroscopy in soil science. In *Advances in Agronomy*; Academic Press: Cambridge, MA, USA, 2010; Volume 107, pp. 163–215. [[CrossRef](#)]
23. Lee, K.S.; Lee, D.H.; Sudduth, K.A.; Chung, S.O.; Kitchen, N.R.; Drummond, S.T. Wavelength identification and diffuse reflectance estimation for surface and profile soil properties. *Trans. ASABE* **2009**, *52*, 683–695. [[CrossRef](#)]
24. 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](#)]
25. Viscarra Rossel, R.A.; Behrens, T.; Ben-Dor, E.; Brown, D.J.; Dematté, J.A.M.; Shepherd, K.D.; Shi, Z.; Stenberg, B.; Stevens, A.; Adamchuk, V.; et al. A global spectral library to characterize the world's soil. *Earth Sci. Rev.* **2016**, *155*, 198–230. [[CrossRef](#)]
26. Brown, D.J. Using a global VNIR soil-spectral library for local soil characterization and landscape modeling in a 2nd-order Uganda watershed. *Geoderma* **2007**, *140*, 444–453. [[CrossRef](#)]
27. Stevens, A.; Nocita, M.; Tóth, G.; Montanarella, L.; van Wesemael, B. Prediction of Soil Organic Carbon at the European Scale by Visible and Near InfraRed Reflectance Spectroscopy. *PLoS ONE* **2013**, *8*, e66409. [[CrossRef](#)] [[PubMed](#)]
28. Waiser, T.H.; Morgan, C.L.S.; Brown, D.J.; Hallmark, C.T. In Situ Characterization of Soil Clay Content with Visible Near-Infrared Diffuse Reflectance Spectroscopy. *Soil Sci. Soc. Am. J.* **2007**, *71*, 389. [[CrossRef](#)]
29. Kuang, B.; Mahmood, H.S.; Quraishi, M.Z.; Hoogmoed, W.B.; Mouazen, A.M.; van Henten, E.J. Sensing soil properties in the laboratory, in situ, and on-line. A review. In *Advances in Agronomy*; Academic Press: Cambridge, MA, USA, 2012.
30. Viscarra Rossel, R.A.; Cattle, S.R.; Ortega, A.; Fouad, Y. In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy. *Geoderma* **2009**, *150*, 253–266. [[CrossRef](#)]
31. Roudier, P.; Hedley, C.B.; Ross, C.W. Prediction of volumetric soil organic carbon from field-moist intact soil cores. *Eur. J. Soil Sci.* **2015**, *66*, 651–660. [[CrossRef](#)]
32. Knadel, M.; Thomsen, A.; Schelde, K.; Greve, M.H. Soil organic carbon and particle sizes mapping using vis-NIR, EC and temperature mobile sensor platform. *Comput. Electron. Agric.* **2015**, *114*, 134–144. [[CrossRef](#)]
33. Vincent, D.R.; Deepa, N.; Elavarasan, D.; Srinivasan, K.; Chauhdary, S.H.; Iwendu, C. Sensors Driven AI-Based Agriculture Recommendation Model for Assessing Land Suitability. *Sensors* **2019**, *19*, 3667. [[CrossRef](#)]
34. Sharma, N.; Pant, B.D.; Mathur, J. MEMS devices used in agriculture—a review. *J. Biosens. Bioelectron.* **2019**, *10*, 267. [[CrossRef](#)]
35. Rady, A.; Fischer, J.; Reeves, S.; Logan, B.; James Watson, N. The Effect of Light Intensity, Sensor Height, and Spectral Pre-Processing Methods When Using NIR Spectroscopy to Identify Different Allergen-Containing Powdered Foods. *Sensors* **2019**, *20*, 230. [[CrossRef](#)]

36. Salzmann, M.; Schledjewski, R. Controlling Moisture Content of Natural Fibres in RTM-Process. In *Key Engineering Materials*; Trans Tech Publisher: Zurich, Switzerland, 2019; Volume 809, pp. 610–614.
37. Uusitalo, S.; Aernouts, B.; Sumen, J.; Hietala, E.; Utriainen, M.; Frondelius, L.; Kajava, S.; Pastell, M. Comparison of milk analysis performance between NIR laboratory analyser and miniaturised NIR MEMS sensors. In *ICAR Conference Technical Series No. 24: New Traits and Adding New Value to the Recording and ID Services in the Animal Production, Prague, Czech Republic, 17–21 June 2019*; ICAR: Rome, Italy, 2019; pp. 111–115.
38. Kirchler, C.G.; Pezzei, C.K.; Beć, K.B.; Mayr, S.; Ishigaki, M.; Ozaki, Y.; Huck, C.W. Critical evaluation of spectral information of benchtop vs. portable near-infrared spectrometers: Quantum chemistry and two-dimensional correlation spectroscopy for a better understanding of PLS regression models of the rosmarinic acid content in Rosmarin. *Analyst* **2017**, *142*, 455–464. [[CrossRef](#)] [[PubMed](#)]
39. Luco Colilles, A. *Evaluation of Two Compact Near-Infrared Spectrometers for the Detection of Microplastics in Soil Samples, Report*; Universitat de Barcelona: Barcelona, Spain, 2019.
40. Tang, Y.; Jones, E.; Minasny, B. Evaluating low-cost portable near infrared sensors for rapid analysis of soils from South Eastern Australia. *Geoderma Reg.* **2020**, *20*, e00240. [[CrossRef](#)]
41. Ng, W.; Husnain; Anggria, L.; Siregar, A.F.; Hartatik, W.; Sulaeman, Y.; Jones, E.; Minasny, B. Developing a soil spectral library using a low-cost NIR spectrometer for precision fertilization in Indonesia. *Geoderma Reg.* **2020**, *22*, e00319. [[CrossRef](#)]
42. Tziolas, N.; Tsakiridis, N.; Ben-Dor, E.; Theocharis, J.; Zalidis, G. A memory-based learning approach utilizing combined spectral sources and geographical proximity for improved VIS-NIR-SWIR soil properties estimation. *Geoderma* **2019**, *340*, 11–24. [[CrossRef](#)]
43. Kennard, R.W.; Stone, L.A. Computer aided design of experiments. *Technometrics* **1969**, *11*, 137–148. [[CrossRef](#)]
44. Bouyoucos, G.J. Hydrometer method improved for making particle size analyses of soils 1. *Agron. J.* **1962**, *54*, 464–465. [[CrossRef](#)]
45. Allison, L.E.; Black, C. Methods of soil analysis. Part 2. Chemical and microbiological properties. *Agron. Monogr.* **1965**, *9*, 1367–1378.
46. Ben Dor, E.; Ong, C.; Lau, I.C. Reflectance measurements of soils in the laboratory: Standards and protocols. *Geoderma* **2015**, *245–246*, 112–124. [[CrossRef](#)]
47. Pimstein, A.; Notesco, G.; Ben-Dor, E. Performance of Three Identical Spectrometers in Retrieving Soil Reflectance under Laboratory Conditions. *Soil Sci. Soc. Am. J.* **2011**, *75*, 746–759. [[CrossRef](#)]
48. Roderick, G.L. *A History of Particle-Size Limits*, 1st ed.; Iowa State University: Ames, Iowa, 1962.
49. Kruse, F.A.; Lefkoff, A.B.; Boardman, J.W.; Heidebrecht, K.B.; Shapiro, A.T.; Barloon, P.J.; Goetz, A.F.H. The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ.* **1993**, *75*, 746–759. [[CrossRef](#)]
50. Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *26*, 1627–1639. [[CrossRef](#)]
51. Gholizadeh, A.; Boruvka, L.; Saberioon, M.M.; Kozák, J.; Vašát, R.; Nemecek, K. Comparing different data preprocessing methods for monitoring soil heavy metals based on soil spectral features. *Soil Water Res.* **2015**, *10*, 218–227. [[CrossRef](#)]
52. Barnes, R.J.; Dhanoa, M.S.; Lister, S.J. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* **1989**, *43*, 772–777. [[CrossRef](#)]
53. Silva, E.B.; Giasson, É.; Dotto, A.C.; Caten, A.T.; Demattê, J.A.M.; Bacic, I.L.Z.; da Veiga, M. A regional legacy soil dataset for prediction of sand and clay content with VIS-NIR-SWIR, in southern Brazil. *Rev. Bras. Cienc. Do Solo* **2019**, *43*. [[CrossRef](#)]
54. Lamichhane, S.; Kumar, L.; Wilson, B. Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. *Geoderma* **2019**, *352*, 395–413. [[CrossRef](#)]
55. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
56. de Santana, F.B.; de Souza, A.M.; Poppi, R.J. Visible and near infrared spectroscopy coupled to random forest to quantify some soil quality parameters. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2018**, *191*, 454–462. [[CrossRef](#)]
57. Svetnik, V.; Liaw, A.; Tong, C.; Christopher Culberson, J.; Sheridan, R.P.; Feuston, B.P. Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1947–1958. [[CrossRef](#)]
58. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
59. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **2008**, *28*, 1–26. [[CrossRef](#)]
60. Vohland, M.; Besold, J.; Hill, J.; Fründ, H.-C. Comparing different multivariate calibration methods for the determination of soil organic carbon pools with visible to near infrared spectroscopy. *Geoderma* **2011**, *166*, 198–205. [[CrossRef](#)]
61. Terra, F.S.; Demattê, J.A.M.; Viscarra Rossel, R.A. Spectral libraries for quantitative analyses of tropical Brazilian soils: Comparing vis-NIR and mid-IR reflectance data. *Geoderma* **2015**, *255–256*, 81–93. [[CrossRef](#)]
62. Knadel, M.; Stenberg, B.; Deng, F.; Thomsen, A.; Greve, M.H. Comparing predictive abilities of three visible-near infrared spectrophotometers for soil organic carbon and clay determination. *J. Near Infrared Spectrosc.* **2013**, *21*, 67–80. [[CrossRef](#)]
63. Viscarra Rossel, R.A.; Chappell, A.; De Caritat, P.; McKenzie, N.J. On the soil information content of visible-near infrared reflectance spectra. *Eur. J. Soil Sci.* **2011**, *62*, 442–453. [[CrossRef](#)]
64. Krishnan, P.; Alexander, J.D.; Butler, B.J.; Hummel, J.W. Reflectance Technique for Predicting Soil Organic Matter¹. *Soil Sci. Soc. Am. J.* **1980**, *44*, 1282. [[CrossRef](#)]

65. Dalal, R.C.; Henry, R.J. Simultaneous Determination of Moisture, Organic Carbon, and Total Nitrogen by Near Infrared Reflectance Spectrophotometry1. *Soil Sci. Soc. Am. J.* **1986**, *50*, 120. [[CrossRef](#)]
66. Rodionov, A.; Pätzold, S.; Welp, G.; Pude, R.; Amelung, W. Proximal field Vis-NIR spectroscopy of soil organic carbon: A solution to clear obstacles related to vegetation and straw cover. *Soil Tillage Res.* **2016**, *163*, 89–98. [[CrossRef](#)]
67. Irons, J.R.; Weismiller, R.A.; Peterson, G.W. Soil Reflectance. *Remote Sens. Earth Sci.* **1989**, *3*, 66–106.
68. Shi, T.; Chen, Y.; Liu, H.; Wang, J.; Wu, G. Soil organic carbon content estimation with laboratory-based visible-near-infrared reflectance spectroscopy: Feature selection. *Appl. Spectrosc.* **2014**, *698*, 831–837. [[CrossRef](#)]
69. Tsakiridis, N.L.; Keramaris, K.D.; Theocharis, J.B.; Zalidis, G.C. Simultaneous prediction of soil properties from VNIR-SWIR spectra using a localized multi-channel 1-D convolutional neural network. *Geoderma* **2020**, *367*, 114208. [[CrossRef](#)]
70. Viscarra Rossel, 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](#)]