

Article

# Mobile Proximal Sensing with Visible and Near Infrared Spectroscopy for Digital Soil Mapping

Masakazu Kodaira \*  and Sakae Shibusawa

Institute of Agriculture, Tokyo University of Agriculture and Technology, Fuchu-shi, Tokyo 183-8509, Japan; sshibu@cc.tuat.ac.jp

\* Correspondence: kodaira.phd@gmail.com

Received: 29 February 2020; Accepted: 3 July 2020; Published: 7 July 2020



**Abstract:** The objective of this study was to estimate multiple soil property local regression models, confirm the accuracy of the predicted values using visible near-infrared subsurface diffuse reflectance spectra collected by a mobile proximal soil sensor, and show that digital soil maps predicted by multiple soil property local regression models are able to visualize empirical knowledge of the grower. The parent materials in the experimental fields were light clay, clay loam, and sandy clay loam. The study was conducted in Saitama Prefecture, Japan. To develop local regression models for the 30 chemical and 4 physical properties, a total of 231 samples were collected; to evaluate accuracy of prediction, 65 samples were collected. The local regression models were developed using 2nd derivative pretreatment by the Savitzky–Golay algorithm and partial least squares regression. The local regression models were evaluated using the coefficient of determination ( $R^2$ ), residual prediction deviation (RPD), range error ratio (RER), and the ratio of prediction error to interquartile range (RPIQ). The  $R^2$  accuracy of the 34 local regression models was 0.81 or higher. In the predicted values for 65 unknown samples, the local regression models could ‘distinguish between high and low’ for 3 of the 34 soil properties, but were ‘not useful’ as absolute quantitative values for the other 31 soil properties. However, it was confirmed that the predicted values followed the transition in measured values, and thus that the developed 34 regression models could be used for generating digital soil maps based on relative quantitative values. The grower changed the ridge direction in the field from east–west to north–south just looking at the digital soil maps.

**Keywords:** mobile proximal sensor; local regression model; spectroscopy; partial least squares regression; chemometrics

## 1. Introduction

Proximal soil sensing (PSS), coupled with GNSS (global navigation satellite system) and visible and near-infrared (VNIR) spectroscopy, is a promising approach for detailed characterization of spatial soil heterogeneity. Given that none of the existing on-the-go soil sensors can provide all the soil information required for precision agriculture (including sustainable farming, environmental load reduction, and risk management), the combination of spectroscopy and chemometrics offers a reasonable alternative for characterizing soil composition [1]. In particular, using PSS in agricultural fields can dramatically enhance this technique by enabling innovative approaches based on appropriate field understanding which characterize local soil and environmental conditions in space and time, improving the efficiency of production to minimize environmental side effects and maximize farm incomes by increasing crop quality and/or yield [2].

Conventional PSS can facilitate the measurement and monitoring of the soil’s chemical [3,4], physical [5], and biological [6] attributes to better understand their dynamics and interactions with the environment, while at the same time revealing their broader spatial heterogeneity through

digital soil mapping. In addition, when high accuracy prediction values for soil properties related to environmental risks are obtained by PSS, they can be used to effectively monitor methanogen in paddy fields, as well as soil organic carbon, and offer a sound foundation for the adoption of optimal agronomic practices that also enable carbon sequestration and reduction in greenhouse gas emissions. Thus, sensing by PSS can help us to better articulate the potential of agricultural soil to meet the world's needs with regard to soil regeneration, soil contamination monitoring, food security, global climate change, and environmental sustainability, enabling the design and implementation of innovative management practices and the efficient pursuit of sustainable development goals (SDGs). To effectively pursue environmental conservation and sustainable agriculture, a better understanding of soil, based on high spatial-resolution information, is needed. In precision agriculture, also required is simultaneous detection of soil fertility (for crop production) and soil contamination, and environmental monitoring. Thus, the need for timely and accurate multiple soil property information is greater now than ever before.

VNIR spectroscopy is one means of enabling rapid, non-destructive, straightforward, and comparatively inexpensive determination of the relative quantitative value of multiple soil properties. One advantage of VNIR spectroscopy is that a single spectral dataset enables the simultaneous characterization of various soil properties using regression models. Numerous agricultural spectroscopic methods measure the diffuse reflectance spectra of oven-dried soil in the laboratory. In such cases, however, it is necessary to collect soil samples from the target field, oven-dry and sieve them (usually in a 2-mm sieve), and finally dispose of the remaining soil, and there has been little advancement in this process [7–16]. However, these laboratory measurements using VNIR spectroscopy enable highly accurate predictions of multiple soil properties.

In order to improve the conventional measurement methodology, and to determine the relative quantitative values within a given field and between fields, a number of mobile proximal soil sensors have been developed by researchers which are capable of simultaneously recording subsurface diffuse reflectance spectra and global positioning data for digital soil mapping [17–21]. Current mobile soil sensors may have lower accuracy in prediction values than laboratory measurement values, but still provide benefits due to their intense spatial visualization. Therefore, there is ongoing research and development based on in-situ field measurement using on-the-go PSS, in areas such as variable-rate fertilization, the relationship between yield and environmental load, soil fertility, soil pollution, etc. [22–25]. In most of the previous studies, researchers focused on few soil properties they would like to estimate and created soil maps for variable rate fertilizer using PSS predicted values. Therefore, researchers have found it difficult to provide soil property information required by growers. The grower's requests were that if the spatial variability between and within fields could be understandable by using several soil maps, and if that variability was confirmed by the grower's experience, it could be used for farm management. Furthermore, in actual farm management, crops are cultivated for various purposes (e.g., organic farming, carbon farming, high quality crops, biodiversity, and environmental conservation), making digital soil maps depicting multiple soil properties even more valuable. In our previous studies [26–30], the development of regression models for several soil properties used in soil management was incomplete. Consequently, we desired to provide growers with digital soil maps of all soil properties for soil diagnostics. The aims of this study, then, were to:

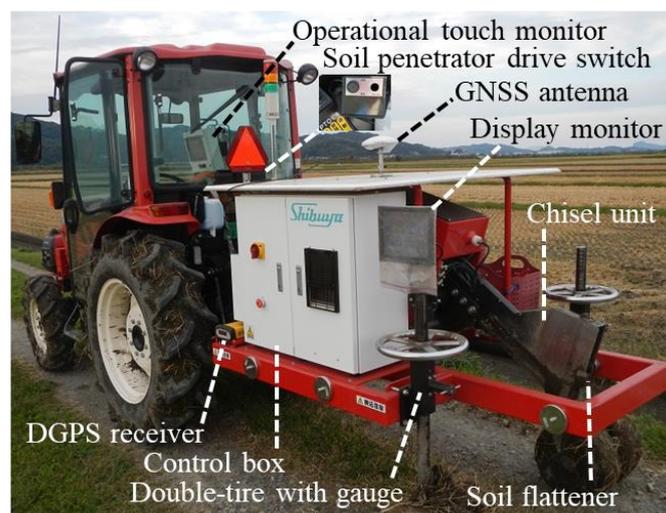
- (i) estimate local regression models for multiple soil properties, using subsurface field-moist soil diffuse reflectance spectra;
- (ii) evaluate the accuracy of predicted values for unknown samples (as both absolute and relative quantitative values);
- (iii) demonstrate the effectiveness of newly-developed digital soil map software for our proximal soil sensor, which can generate digital soil maps immediately after measurement in agricultural fields;
- (iv) describe an instance of decision making by a grower using such digital soil maps.

In this study, a total of 34 local regression models for 30 chemical and 4 physical properties were estimated by 2nd derivative pretreatment and partial least squares regression coupled with full cross-validation.

## 2. Materials and Methods

### 2.1. Mobile Proximal Sensor

A mobile proximal soil sensor (model SAS3000, Shibuya Seiki Co., Ltd., Ehime Pref., Japan) was used to measure visible and near-infrared subsurface diffuse reflectance spectra (VNIR-SDRS) in agricultural fields. The SAS3000 records several types of data simultaneously, to provide high-resolution soil maps of agricultural fields. It is composed of an operational touch monitor, soil penetrator drive switch, display monitor, GNSS antenna (model GA530, Nikon-Trimble Co., Ltd., Tokyo, Japan), DGPS (differential global positioning system) receiver (model SPS351, Nikon-Trimble Co., Ltd., Tokyo, Japan), and control box (Figure 1). The electrical power for the SAS3000 was supplied by a gasoline inverter generator (EU-16i, Honda Motor Co., Ltd., Minato-ku, Tokyo, Japan). An oil control valve in the tractor is connected to the hydraulic power system of the SAS3000, and is used to drive the chisel unit into the subsurface. The core devices of the SAS3000 for measuring the VNIR-SDRS include a tungsten halogen lamp (JCR15V 150W AL, Fuji Electric Lamp Industrial Co., Ltd., Tokyo, Japan), visible range spectrophotometer (C10083CAH, 320 to 1000 nm, 1 nm interval; Hamamatsu Photonics K. K., Shizuoka Pref., Japan), and near infrared range spectrophotometer (C9406GC, 900 to 1700 nm, 7 nm interval, Hamamatsu Photonics K. K., Shizuoka Pref., Japan). The SAS3000 was designed to collect subsurface VNIR-SDRS at depths of 0.05 to 0.30 m (with 0.05 m intervals) from the soil flattener to the soil surface, created by the double-tire of the SAS3000, which is adjusted by a gauge on the double-tire. In this study, the soil flattener depth was set at 0.15 m. The VNIR-SDRS measurement was recorded at 3 s intervals with a tractor speed of 0.28 m s<sup>-1</sup>.

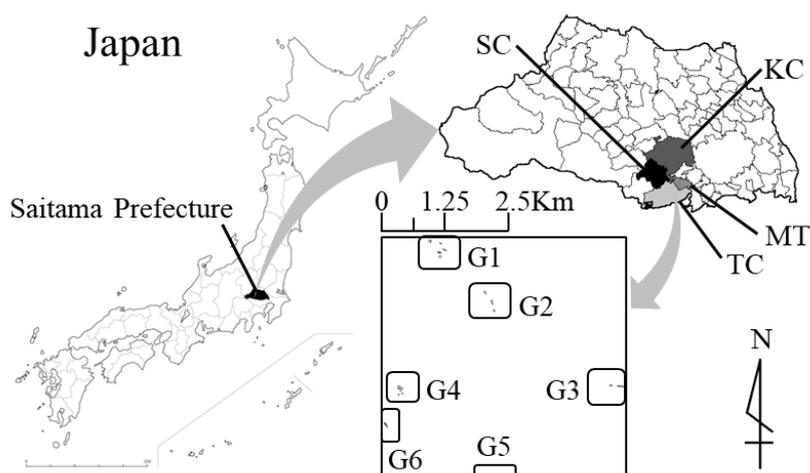


**Figure 1.** The study's mobile proximal soil sensor (SAS3000).

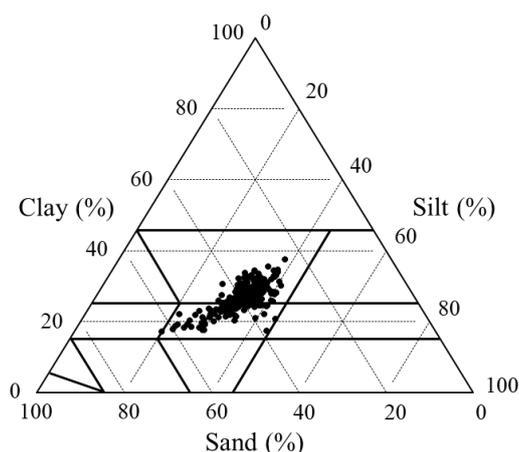
### 2.2. VNIR Spectra, Field-Moist Soil Collection, and Soil Analysis

The experimental sites were in rotational upland fields of taro (*Colocasia esculenta* (L.) Schott; cultivar: Dodare and Hasuba) in Sayama City (SC), Tokorozawa City (TC), Kawagoe City (KC), and Miyoshi Town (MT), Saitama Pref., Japan. To estimate local regression models for multiple soil properties, the experiment was conducted in a total of 24 fields, of 6 growers [G1, 8 fields (total 0.94 ha); G2, 3 fields (0.45 ha); G3, 2 fields (0.41 ha); G4, 7 fields (0.74 ha); G5, 1 field (0.23 ha); G6, 3 fields (0.35 ha); Figure 2]. A total of 231 VNIR-SDRS data and soil samples were collected at the measured VNIR-SDRS data points after harvesting the taro yam plants in combination with tilling, from 2017

to 2018. The parent materials in the experimental sites were distributed on light clay, clay loam, and sandy clay loam as soil texture (Atterberg method) (Figure 3).



**Figure 2.** Location of the experimental sites. SC: Sayama City; KC: Kawagoe City; TC: Tokorozawa City; MT: Miyoshi Town; G1: 8 fields; G2: 3 fields; G3: 2 fields; G4: 7 fields; G5: 1 field; G6: 3 fields.



**Figure 3.** Distribution of soil texture on 231 field-moist soil samples.

The SAS3000 sounds an alarm at the time of each data acquisition, counts the number of data items, and displays the VNIR-SDRS data number on each monitor. Each time a predetermined VNIR-SDRS data number (for example, at 20-data-number intervals in a single sensing line) was displayed on each monitor, a placemark stick was inserted into the soil surface beside the chisel unit, as a soil sampling point. After the SAS3000 had fully passed the placemark stick, we dug up the soil surface smoothed over by the soil flattener and packed the field-moist soil samples in sealable plastic bags, taking two sets of field-moist soil samples from each soil sampling point.

In this study, we focused on 30 chemical and 4 physical properties, including soil moisture content (MC, weight ratio); soil organic matter (SOM, %); pH (1:2.5 water suspension); exchange acidity (y1); carbon-nitrogen ratio (CN); phosphate absorption coefficient (PAC); electrical conductivity (EC,  $\text{mS cm}^{-1}$ ); cation exchange capacity (CEC,  $\text{me } 100 \text{ g}^{-1}$ ); humus rate (HR, %); total carbon (C-t, %); total nitrogen (N-t, %); ammonium nitrogen (N-a,  $\text{mg } 100 \text{ g}^{-1}$ ); hot-water extractable nitrogen (N-h,  $\text{mg } 100 \text{ g}^{-1}$ ); nitrate nitrogen (N-n,  $\text{mg } 100 \text{ g}^{-1}$ ); available phosphorus (P-a,  $\text{mg } 100 \text{ g}^{-1}$ ); exchangeable potassium (K,  $\text{mg } 100 \text{ g}^{-1}$ ); exchangeable calcium (Ca,  $\text{mg } 100 \text{ g}^{-1}$ ); exchangeable sodium (Na,  $\text{mg } 100 \text{ g}^{-1}$ ); exchangeable magnesium (Mg,  $\text{mg } 100 \text{ g}^{-1}$ ); hot-water soluble boron (B-s, ppm); soluble copper (Cu, ppm); easily reducible manganese (Mn, ppm); soluble zinc (Zn, ppm); calcium-magnesium ratio

(CaMg, equivalence ratio; E.R.); magnesium-potassium ratio (MgK, E.R.); base saturation percentage (BSP, %); calcium saturation percentage (CSP, %); exchangeable sodium percentage (ESP, %); available silicate (SiO, mg 100 g<sup>-1</sup>); free iron oxide (Fe, %); dry density (DD, g cm<sup>-3</sup>); sand (S, %); silt (SL, %); and clay (CL, %).

To estimate these local regression models for the SAS3000, two sets of 231 field-moist soil samples and their VNIR-SDRS were collected from the total of 24 fields of the six growers, from 2017 to 2018. In addition, in 2019, two sets of 65 field-moist soil samples and their VNIR-SDRS were collected from a total of 10 fields of 4 of the growers (G2, G3, G4, and G5), to confirm the accuracy of predictions for unknown samples using the respective estimated local regression models. Thus, in total, two sets of 296 field-moist soil samples and their VNIR-SDRS were collected, at the VNIR-SDRS data measurement points, after harvesting the taro plants in combination with tilling, from 2017 to 2019. One set was analyzed at our laboratory in Tokyo University of Agriculture and Technology (TUAT; Tokyo, Japan), and the other at the Agricultural Product Chemical Research Laboratory of the Federation of Tokachi Agricultural Cooperative Association (APCRL; Hokkaido, Japan) and the Oita Laboratory of Sumika Chemical Analysis Service, Ltd. (SCAS; Oita Pref., Japan). The 296 field-moist soil samples sent to TUAT and APCRL, respectively, were transported in a refrigerator car at less than 10 °C. At TUAT, MC was determined using the oven-dry method (110 °C, 24 h) [31]. The samples were then stored in sealable plastic bags at 5 °C until the end of MC measurement. After MC measurement, each oven-dry soil sample was sieved through a 2 mm sieve, and these 2-mm-sieved oven-dry soil samples were used to measure SOM by the ignition loss method, using a muffle furnace (750 °C, 3 h) [31]. MC and SOM measurements were conducted three times, and the average value was adopted as a dataset for the multivariate statistical analysis. At APCRL, the field-moist soil samples were oven-dried (60 °C, 24 h) and 2-mm-sieved, to measure the following 30 soil physical and chemical properties: pH, y1, PAC, EC, CEC, HR, CN, C-t, N-t, N-a, N-h, N-n, P-a, K, Ca, Mg, Na, B-s, Cu, Mn, Zn, CaMg, MgK, BSP, CSP, ESP, DD, S, SL, and CL. These soil properties were assessed using the soil analysis method of the Agricultural Research Division [32]. The 296 2-mm-sieved oven-dry soil samples left by APCRL were sent to SCAS for analysis of SiO and Fe. The SiO solution was extracted using soil nutrient analysis method 14 [33] and the Fe solution was extracted using soil nutrient analysis method 16.3.4 [34]. These sample solutions were assessed with an inductivity coupled plasma optical emission spectrometer SPS5520ICP-OES (Agilent Technologies International Japan, Ltd., Tokyo, Japan).

### 2.3. Absorption Spectral Wavelength of Soil Properties

It has been confirmed that MC, HR, SiO, Fe, and CL have absorption wavelengths at 410 (SiO), 420 to 800 (Fe<sup>3+</sup>), 425 (Fe), 450 (HR), 470 (HR), 570 (HR), 620 (HR), 760 (MC), 930 (SiO), 970 (MC), 1000 (SiO), 1020 to 1070 (Fe<sup>2+</sup>), 1025 (Fe), 1075 (Fe), 1100 to 1600 (HR), 1380 to 1410 (Fe), 1390 (CL), 1400 (CL), 1420 (CL), and 1450 nm (MC) [35–48]. As the spectra measurement range of the SAS3000 was 350 to 1700 nm, it is theoretically possible to estimate regression models for all five soil properties (MC, HR, SiO, Fe and CL) using the device. However, the partial least squares regression (PLSR) spectra analysis range was set at 500 to 1600 nm, meaning that absorbance wavelengths of 350 to 500 nm and 1600 to 1700 nm could not be used in this analysis for regression model estimation. In addition, the absorption wavelengths of the other 29 soil properties (29 SPs) did not fall within this 500 to 1600 nm range. Thus, to estimate the soil property regression models for these 29 SPs, it was necessary to confirm the correlation between their measured values and those of the 5 SPs [49–51]. Specifically, when there was sufficient (0.4 or greater) measured value correlation between the above 5 SPs and the remaining 29 SPs, we were able to estimate regression models.

### 2.4. Spectra Data Processing and Partial Least Squares Regression Setting

To estimate the regression models, two spectrometer (visible and NIR range spectrophotometer) data sets were converted to a VNIR-SDRS data sets, and an absorbance spectra data sets, using the following procedure, based on Kodaira and Shibusawa [27] and using Data Monitor Software (DMS;

Shibuya Seiki Co., Ltd.). To transform the two spectrometer data sets to VNIR-SDRS data sets, the former was connected at 1000 nm, coupled with equal (5-nm-interval) spacing of the wave number resolution, with a wavelength range from 350 to 1700 nm, using the interpolation method. Then, the VNIR-SDRS data sets were converted to absorbance spectra data sets, using white reference spectra data (with a standard reflector: SRS-40-020; Labsphere, Inc., North Sutton, NH, USA) and dark reference spectra data (with light shielding). The white and dark reference spectra data were recorded once in each experimental field before commencing the VNIR-SDRS measurement [27].

Since absorbance spectra data are affected by uncertain noise (caused by factors such as the soil particle size, moisture variability, and light scatter), pretreatment of the spectra data can often improve the regression model accuracy, compared to using untreated absorbance spectra data. In addition, to enhance weak signals and reduce the noise and light scatter that would affect the baseline, smoothing and derivative mathematical pretreatment can be useful, depending on the parameter conditions. For example, the first derivative can eliminate baseline fluctuation, and the second can eliminate multiplicative fluctuation. In addition, the second derivative offers the following beneficial effects:

- (i) the peak waveforms can be separated from the broad absorption wavelengths;
- (ii) the peak wavelength values are enhanced more than with the first derivative;
- (iii) the peak wavelength value corresponds to the original absorbance wavelength (the peak waveform is inverted).

Thus, in this study, the second derivative coupled with smoothing ( $2D_{S-G}$ ) was applied [52].

MLR (multiple regression analysis), PCA (principal component analysis), PCR (principal component regression), NN (neural network) are among the multivariate analysis methods. In this paper, PLSR analysis [53–55] coupled with full (leave-one-out) cross-validation was applied, because:

- (i) even when there are fewer response variables than explanatory variables, it is possible to estimate a regression model;
- (ii) it enables orthogonalization between wavelength variables, to avoid multicollinearity.

$2D_{S-G}$  and PLSR were performed using Unscrambler Ver.9.8 software (U98, CAMO Analytics AS, Oslo, Norway).

In the field measurement using the SAS3000, it is impossible to remove all the gravel and crop residues, roots, creatures, etc. from the soil. Therefore, there are some outliers in absorbance spectra data sets, and it is necessary to exclude outliers from the absorbance spectra data sets. In this study, the selection criteria for identifying a given sample as an outlier was the largest residual variance value, based on the residual sample variance plot in U98 after the PLSR [56,57]. This process of absorbance spectra data outlier exclusion was repeated until the coefficient of determination ( $R^2$ ) of the regression model was 0.82 or higher. In this study, we decided to exclude outliers up to a maximum of 115 samples (roughly half of the 231 total samples). In this study, the following  $2D_{S-G}$  and PLSR parameters were set using U98:

$2D_{S-G}$  setting was,

- (i) the wavelength range used for calculation was 350 to 1700 nm;
- (ii) the 'polynomial order' was second order;
- (iii) the 'number of smoothing points' was 3 to 41 (only odd numbers could be selected), and the guideline for selection was the 'number of smoothing points' with the highest  $R^2$ .

PLSR setting was,

- (i) the wavelength range used for calculation was 500 to 1600 nm, with reference to the report by Kodaira and Shibusawa [27];

- (ii) the 'number of PCs' (principal components; i.e., the number of PLSR factors) was 20; limited to this number, for each regression model, to prevent overfitting;
- (iii) the remaining settings were default settings.

### 2.5. Evaluation of Regression Models and Predicted Values

The regression models and predicted values for the unknown samples were evaluated using the following performance indicators: the  $R^2$  value, the root mean square error (RMSE), residual prediction deviation (RPD), range error ratio (RER), and the ratio of prediction error to interquartile range (RPIQ). Generally, a model that performs well would have large values of  $R^2$ , RPD, RER, and RPIQ, and a small RMSE value. These performance indicators were computed for each soil property according to the following equations:

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

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

$$\text{RPD} = \frac{\text{S.D.}}{\text{RMSE}} \quad (3)$$

where  $y_i$  is the measured value of the  $i$ th sample measured by conventional soil analysis,  $\hat{y}_i$  is the predicted value of the  $i$ th sample,  $\bar{y}_i$  is the average of the predicted values,  $N$  is the number of samples, and S.D. is the standard deviation of the observed soil property in the full cross-validation or prediction dataset;

$$\text{RER} = \frac{\text{Range}}{\text{RMSE}} \quad (4)$$

where the RER is the ratio of the range of the measured values of soil properties in the full cross-validation set or the prediction set to the RMSE;

$$\text{RPIQ} = \frac{\text{IQ}}{\text{RMSE}} \quad (5)$$

where IQ is the interquartile range of the observed soil property in the full cross-validation or prediction dataset (IQ = Q3 – Q1, quantifying the spread in the central 50% of the data (Q2), with Q1 as the median of the first half, and Q3 the median of the second half). RPD and RPIQ (ratio of performance to IQ) are used in the spectroscopic literature to make data more comparable, accounting for the differences in the spread of the data. Their explanatory power in terms of classifying the model performance is the same as  $R^2$ , but they are also useful for comparing datasets with different data spreads; with the additional advantage that RPIQ can be used for skewed data [57,58].

Previous studies defined criteria to assess the relative fit of the regression model using performance indicators, as following:  $R^2 > 0.90$ ,  $0.82 < R^2 \leq 0.90$ ,  $0.66 < R^2 \leq 0.82$ ,  $0.50 < R^2 \leq 0.66$ , and  $R^2 \leq 0.5$ ;  $\text{RPD} > 3.0$ ,  $2.5 < \text{RPD} \leq 3.0$ ,  $2.0 < \text{RPD} \leq 2.5$ ,  $1.5 < \text{RPD} \leq 2.0$ , and  $\text{RPD} \leq 1.5$ ;  $\text{RER} > 20$ ,  $15 < \text{RER} \leq 20$ ;  $10 < \text{RER} \leq 15$ ,  $8 < \text{RER} \leq 10$ , and  $\text{RER} \leq 8$  [59,60]. RPIQ follows the same classification format as RPD. In this study, based on these indicators, the regression model performance was classified into five categories: Category A is excellent, B is good (successful for RER), C indicates approximate quantitative prediction (moderately successful for RER), D indicates capability to distinguish between high and low (moderately useful for RER), and E means not useful (screening for RER).

### 2.6. Digital Soil Mapping Software

Traditionally, ArcGIS Ver.10.0 software (ESRI Japan Co., Chiyoda-ku, Tokyo, Japan) has been used to display digital soil maps, but parameter setting of ArcGIS is complicated and is both time consuming and very expensive [27]. Thus, in this study, the digital soil maps were prepared using

Soil Map Viewer (SMV, Figure 4) for the SAS3000. The generated digital soil maps are displayed with colored dots representing the soil sensing points, meaning that no interpolation method (e.g., Inverse Distance Weighting, Kriging) is required. The features of SMV are:

- (i) classification of predicted values can be specified up to 5 levels;
- (ii) the boundary line of each field is displayed (boundary position data is required);
- (iii) statistics data of the mean, max, min, and coefficient of variation can be confirmed;
- (iv) the absorbance spectra, the absorbance spectra after pretreatment and predicted value at each measurement point can be confirmed;
- (v) a histogram of the selected area is displayed.

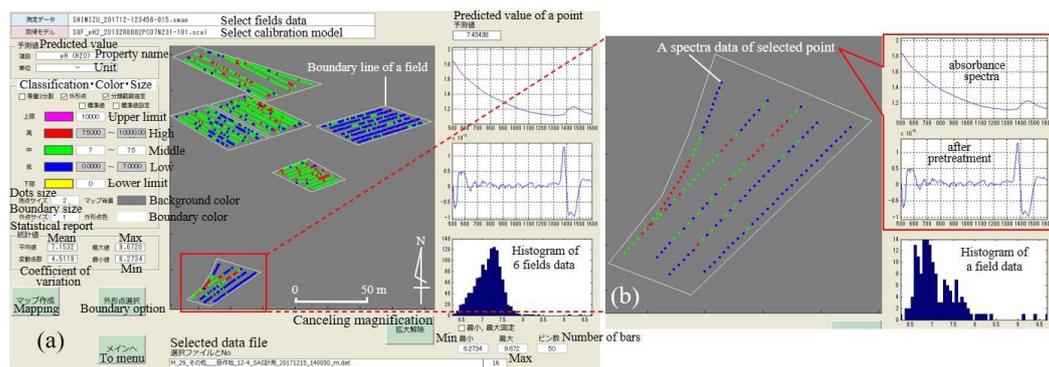


Figure 4. In-situ digital soil mapping system: (a) view of six fields; (b) close-up of one field.

As a result, it is possible to display the 34 soil properties soil map and talk with growers on site immediately after the measurement, and to provide a sound analytical foundation for soil management. In particular, (iv) can support outlier determination as unusual spectra. SMV was jointly developed for the SAS3000 by Shibuya Seiki Co., Ltd. and the authors.

### 2.7. Workflow of Regression Model Estimation and Digital Soil Mapping

Figure 5 shows the workflow of the regression model estimation and digital soil mapping. The regression models are estimated by performing C, D, and E for each soil property in the workflow.

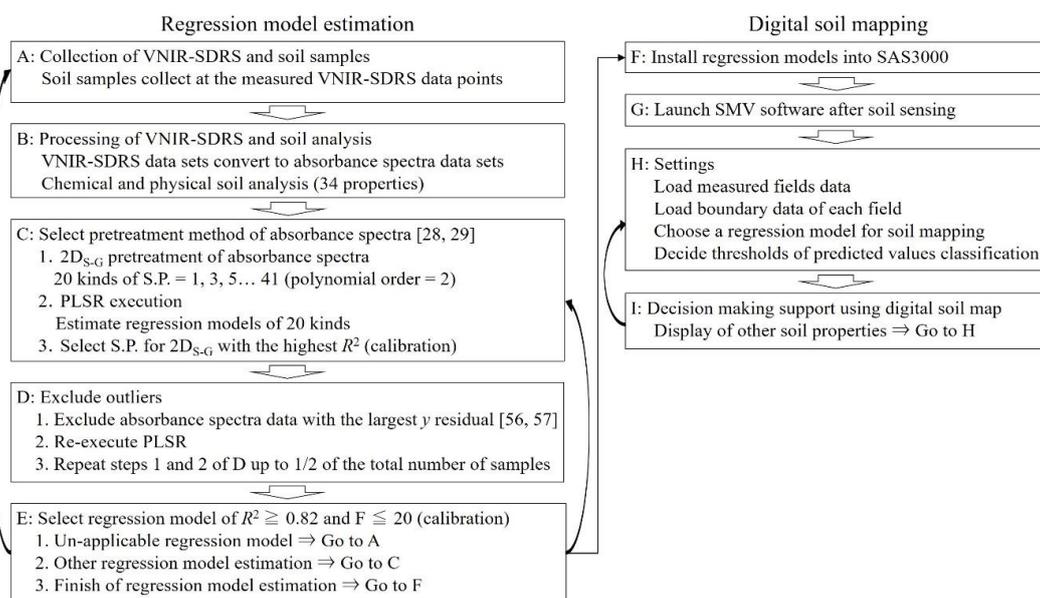


Figure 5. Workflow of regression model estimation and digital soil mapping.

### 3. Results

#### 3.1. Evaluation and Accuracy of Local Regression Models

Table 1 shows the correlation between measured values. As the absorption wavelengths of CL, Fe, HR, MC, and SiO (5 SPs, red character) have been confirmed to fall within the range of 500 to 1600 nm, it is theoretically possible to estimate local regression models for these properties using PLSR. The measured values for BSP, Ca, CaMg, CEC, CSP, C-t, Cu, DD, MgK, N-h, N-t, P-a, PAC, pH, S, SL, SOM and y1 (18 SPs, green character) were confirmed to be in correlation with those of the 5 SPs; while those of B-s, CN, EC, ESP, K, Mg, Na, N-a, N-n and Zn (10 SPs, blue character) were confirmed to be in correlation with those of these 18 SPs. Consequently, it was confirmed that the 10 SPs had an indirect correlation with the 5 SPs, through their correlation with the 18 SPs. There was no confirmed correlation with measured values in the case of Mn. Therefore, with the exception of Mn, it was confirmed that local regression model estimation was possible for both the 10 SPs and the 18 SPs. For Mn, it is statistical estimate.

The accuracy of the respective local regression models estimated using PLSR is shown in Table 2. Without outlier exclusion, the smoothing points (S.P.) of  $2D_{S-G}$  for the absorbance spectra with the highest  $R^2$  were 5 to 33. The average absorbance spectra of the 231 samples, after pretreatment by  $2D_{S-G}$ , are shown in Figure 6. The S.P. of B-s, Fe, MC, N-t, and P-a were 30 or more, meaning that greater noise removal was required than in the case of the other soil properties. On the other hand, the S.P. of CSP and EC were 10 or less, suggesting that it was necessary to utilize small spectral fluctuation. When outliers were excluded until  $R^2$  exceeded 0.82, the number of PLSR factors (F) for the local regression models was 2 to 14, except in the case of Cu (Table 2). The accuracy of the Cu local regression model was  $R^2 = 0.81$ , and F was 15 when the maximum outlier was 115. Although F for B-s was 2, the number of correlations between measured values was 3, so the result was inferred to be 'not underfitting' (Table 1). As F for y1 and DD was 14, there was concern about overfitting; however, the result was inferred to be 'not overfitting', because the F selected in the previous study was 19 [61]. The local regression model  $R^2$  accuracy of full cross-validation was classified as Category C for Cu, Category D for N-n, and Category B for the other local regression models. The accuracy of  $R^2$ , RPD, and RER for full cross-validation was 0.65 to 0.84 for  $R^2$ , 1.70 to 2.50 for RPD, and 7.15 to 14.6 for RER. In the full cross-validation RPD classification, the local regression models for BSP, CL, CEC, Cu, CSP, DD, EC, MgK, N-a, N-h, N-n, pH, SiO, and y1 were classified as Category D. The full cross-validation RER of CaMg, DD, EC, pH, N-a, and y1 was classified as Category E. Therefore, in order to estimate these local regression models, it was confirmed that the range of measured values should be expanded. Scatter plots of the three soil properties with  $R^2 > 0.82$  (B-s, C-t, HR) and the four soil properties with low  $R^2$  (BSP, Mn, N-n, SiO) are shown in Figure 7 as examples of the scatter plots of measured values versus predicted full cross-validation values. The three soil properties with high  $R^2$  accuracy were distributed on the 1:1 line, but the four soil properties with low  $R^2$  accuracy were distributed away from the 1:1 line.

**Table 1.** Correlation coefficients between the measured values and between the regression coefficients of the local regression models.

Property	Correlation coefficient between measured value																
	BSP	B-s	Ca	CaMg	CEC	CL	CN	CSP	C-t	Cu	DD	EC	ESP	Fe	HR	K	MC
BSP	—	0.20	0.94	0.67	0.59	-0.10	0.09	0.98	0.64	-0.70	-0.45	-0.24	0.06	-0.21	0.64	-0.08	-0.05
B-s	0.03	—	0.16	-0.15	0.25	0.13	0.04	0.12	0.22	-0.16	0.00	0.01	0.00	0.16	0.22	0.18	-0.16
Ca	0.31	-0.07	—	0.76	0.77	-0.19	0.11	0.96	0.76	-0.69	-0.53	-0.26	-0.04	-0.22	0.76	-0.15	0.05
CaMg	0.33	-0.02	0.42	—	0.40	-0.12	0.36	0.80	0.76	-0.40	-0.29	-0.20	-0.18	-0.34	0.76	-0.43	-0.01
CEC	0.28	0.07	0.04	0.16	—	-0.31	-0.10	0.58	0.56	-0.70	-0.52	-0.11	-0.10	0.05	0.56	0.14	0.25
CL	0.10	0.30	-0.23	-0.30	-0.04	—	0.30	-0.12	-0.07	0.26	0.65	0.00	0.24	-0.01	-0.07	-0.01	-0.64
CN	0.20	0.08	0.07	0.08	-0.04	0.39	—	0.18	0.36	0.16	0.19	-0.41	-0.23	-0.27	0.36	-0.60	-0.22
CSP	0.18	0.06	0.12	0.02	0.00	0.08	0.13	—	0.72	-0.65	-0.46	-0.27	-0.01	-0.27	0.72	-0.21	-0.03
C-t	0.27	0.24	0.26	0.59	0.10	-0.01	0.39	0.04	—	-0.39	-0.29	-0.15	-0.19	-0.36	1.00	-0.27	-0.06
Cu	-0.26	-0.36	-0.01	-0.06	-0.14	0.16	0.03	-0.01	-0.17	—	0.77	0.03	-0.08	-0.15	-0.39	-0.24	-0.19
DD	0.15	0.14	-0.26	0.06	-0.03	-0.13	0.03	-0.03	0.25	0.05	—	0.23	0.19	0.11	-0.29	0.07	-0.49
EC	0.01	0.08	-0.04	-0.03	-0.04	0.01	0.01	-0.22	0.02	-0.14	0.11	—	0.48	0.10	-0.15	0.44	-0.03
ESP	0.01	0.23	-0.07	-0.11	-0.12	0.22	-0.33	-0.04	-0.10	0.01	0.07	0.05	—	0.07	-0.19	0.32	-0.34
Fe	-0.10	-0.04	-0.02	0.07	0.17	-0.41	-0.76	-0.01	-0.27	-0.07	-0.04	-0.06	-0.11	—	-0.36	0.33	0.26
HR	0.23	0.13	0.35	0.28	0.18	-0.07	0.17	0.06	0.38	-0.08	0.13	0.01	-0.20	-0.07	—	-0.27	-0.06
K	-0.29	0.17	-0.17	-0.18	0.08	-0.26	-0.78	-0.08	-0.32	-0.02	0.04	0.00	0.10	0.78	-0.24	—	-0.12
MC	-0.04	-0.29	0.25	0.09	0.14	-0.41	-0.06	-0.01	-0.04	-0.05	-0.15	-0.08	-0.70	0.26	0.24	0.09	—
Mg	-0.18	-0.06	-0.10	-0.27	0.05	0.18	-0.28	0.00	-0.31	-0.14	-0.20	0.04	-0.21	0.28	-0.47	0.46	0.28
MgK	-0.13	-0.03	-0.03	-0.08	-0.16	0.03	0.28	-0.06	-0.02	0.09	-0.20	-0.04	0.08	-0.37	-0.07	-0.04	-0.10
Mn	-0.37	0.09	-0.43	-0.42	-0.07	0.29	-0.41	-0.02	-0.31	0.10	0.08	-0.02	0.38	0.17	-0.44	0.48	-0.26
Na	-0.03	0.08	0.12	0.01	-0.02	-0.03	-0.49	-0.05	-0.25	0.04	-0.20	-0.01	0.49	0.32	-0.23	0.06	-0.08
N-a	-0.06	0.09	-0.07	-0.03	-0.12	0.03	-0.03	0.01	0.09	-0.09	0.16	-0.03	-0.04	0.08	-0.07	0.06	0.00
N-h	0.24	0.16	0.18	0.26	0.01	0.05	0.12	0.21	0.36	-0.01	0.33	-0.02	-0.02	-0.01	0.42	-0.40	-0.04
N-n	0.21	0.05	0.26	0.17	-0.05	-0.02	0.14	0.01	0.29	-0.02	0.29	0.07	-0.15	-0.01	0.34	-0.14	0.18
N-t	0.16	0.11	0.18	0.26	0.12	-0.10	0.15	0.17	0.29	-0.08	0.07	-0.04	-0.31	0.01	0.76	-0.37	0.35
P-a	-0.42	0.15	0.16	0.14	0.07	0.19	0.37	0.21	0.24	-0.21	-0.01	0.06	0.18	-0.36	0.38	-0.38	-0.15
PAC	0.42	-0.25	0.12	0.04	0.22	-0.58	-0.56	-0.02	-0.18	-0.06	-0.24	0.00	-0.23	0.60	-0.01	0.40	0.50
pH	0.44	-0.08	0.35	0.39	0.22	-0.37	-0.25	0.12	0.12	0.04	-0.07	-0.12	0.01	0.27	0.19	0.07	0.27
S	0.12	-0.46	0.30	0.13	0.04	-0.13	0.28	-0.08	0.13	0.21	0.06	-0.06	-0.34	-0.30	0.23	-0.11	0.49
SiO	0.09	-0.13	-0.19	0.19	0.28	-0.56	-0.49	0.02	0.00	-0.08	-0.13	-0.04	-0.10	0.50	0.06	0.23	0.35
SL	0.14	0.02	0.24	0.09	-0.05	-0.16	-0.55	0.03	-0.17	-0.14	-0.26	0.03	0.66	0.28	-0.17	0.03	-0.25
SOM	0.08	0.07	0.18	0.25	0.14	-0.13	0.18	0.19	0.25	-0.01	0.03	-0.03	-0.43	0.01	0.67	-0.21	0.37
y1	-0.06	0.08	-0.06	-0.08	-0.09	0.18	0.04	0.02	0.02	0.26	0.38	0.01	0.04	0.02	-0.07	-0.16	-0.27
Zn	-0.09	0.15	-0.27	-0.11	-0.14	0.29	0.23	0.12	0.02	0.26	0.15	0.00	0.14	-0.12	-0.37	0.00	-0.58
Nc	3	1	2	3	0	4	7	0	1	0	0	0	4	5	5	6	5

Correlation coefficient between regression coefficients

Table 1. Cont.

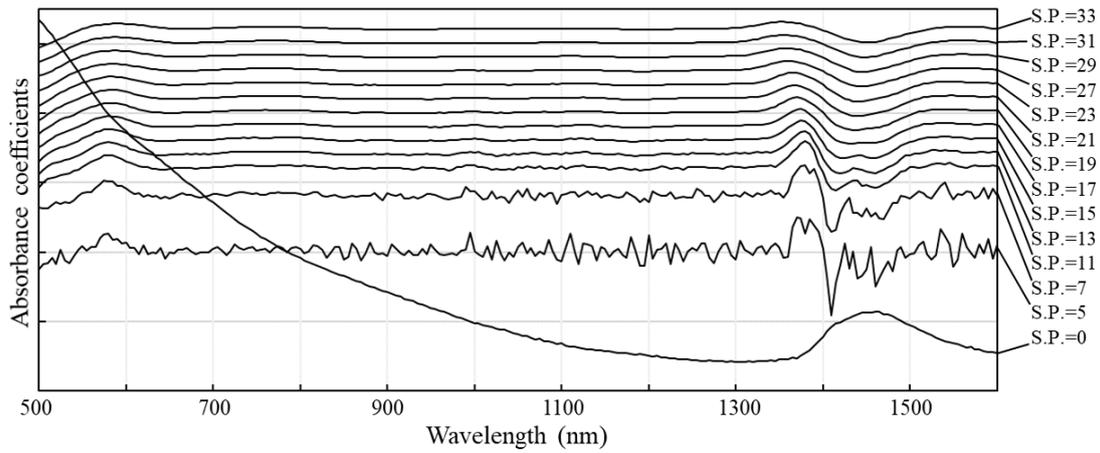
Property	Mg	MgK	Mn	Na	N-a	N-h	N-n	N-t	P-a	PAC	pH	S	SiO	SL	SOM	y1	Zn	Nc
BSP	0.10	0.11	-0.30	0.27	-0.29	0.30	-0.17	0.65	0.39	0.14	0.89	-0.02	0.23	0.13	0.43	-0.66	-0.07	12
B-s	0.47	0.14	-0.14	0.08	-0.15	0.16	-0.06	0.24	0.58	-0.23	0.11	-0.28	-0.20	0.24	0.00	-0.09	0.47	3
Ca	-0.03	0.08	-0.30	0.25	-0.30	0.33	-0.15	0.77	0.34	0.23	0.89	0.05	0.26	0.12	0.56	-0.64	-0.13	12
CaMg	-0.58	-0.16	-0.30	-0.01	-0.19	0.43	-0.20	0.69	0.05	0.10	0.57	0.06	0.12	0.03	0.69	-0.38	-0.33	13
CEC	0.23	0.09	-0.17	0.27	-0.17	0.24	0.06	0.63	0.30	0.38	0.73	0.15	0.36	0.06	0.40	-0.47	0.09	12
CL	0.00	-0.06	0.01	0.10	0.15	0.32	-0.14	-0.17	0.30	-0.70	-0.22	-0.85	-0.55	0.42	-0.16	0.28	0.26	6
CN	-0.33	0.06	-0.16	-0.27	-0.29	0.02	-0.38	0.09	-0.07	-0.30	-0.03	-0.29	-0.27	0.20	0.29	-0.07	-0.31	2
CSP	-0.10	0.06	-0.31	0.21	-0.30	0.33	-0.20	0.71	0.33	0.14	0.86	0.01	0.21	0.11	0.54	-0.64	-0.16	12
C-t	-0.31	-0.09	-0.38	0.03	-0.18	0.54	-0.09	0.96	0.25	0.00	0.50	-0.01	-0.09	0.10	0.76	-0.27	-0.08	10
Cu	-0.27	-0.04	0.08	-0.34	0.07	-0.27	-0.08	-0.46	-0.26	-0.41	-0.82	-0.17	-0.54	0.01	-0.27	0.56	-0.18	11
DD	-0.10	-0.16	0.06	-0.02	0.33	0.33	-0.03	-0.36	0.08	-0.59	-0.50	-0.57	-0.47	0.30	-0.14	0.57	0.39	12
EC	0.04	-0.24	0.14	0.42	0.51	0.34	0.75	-0.04	-0.14	-0.01	-0.33	0.05	-0.07	-0.09	-0.04	0.52	0.16	7
ESP	0.17	-0.06	0.35	0.93	0.44	0.25	0.39	-0.13	0.11	-0.28	-0.02	-0.31	-0.20	0.28	-0.22	0.28	0.17	3
Fe	0.27	-0.03	0.25	0.08	0.10	-0.20	0.17	-0.31	0.00	0.23	-0.06	0.04	0.28	-0.06	-0.22	0.02	0.31	0
HR	-0.31	-0.92	-0.38	0.03	-0.18	0.54	-0.09	0.96	0.25	0.00	0.50	-0.01	-0.09	0.10	0.76	-0.27	-0.08	11
K	0.43	-0.47	0.14	0.36	0.34	0.16	0.52	-0.10	0.27	-0.04	0.01	-0.02	-0.01	0.05	-0.38	0.16	0.53	7
MC	0.05	0.20	-0.01	-0.25	-0.12	-0.39	0.05	0.00	-0.43	0.85	0.12	0.70	0.76	-0.55	0.20	-0.21	-0.30	7
Mg	—	0.43	0.01	0.23	-0.07	-0.23	0.05	-0.22	0.27	0.13	0.19	-0.04	0.19	0.07	-0.40	-0.17	0.37	5
MgK	-0.19	—	0.00	-0.03	-0.30	-0.32	-0.30	-0.11	-0.03	0.22	0.14	0.04	0.22	-0.01	0.00	-0.27	-0.10	3
Mn	0.41	-0.14	—	0.29	0.23	-0.20	0.23	0.01	-0.12	-0.02	-0.24	-0.08	-0.01	0.13	-0.24	0.15	0.06	0
Na	0.06	-0.05	0.32	—	0.37	0.33	0.39	0.11	0.21	-0.13	0.25	-0.24	-0.07	0.30	-0.06	0.09	0.19	2
N-a	0.06	-0.07	-0.01	-0.01	—	0.43	0.34	-0.10	-0.12	-0.15	-0.28	-0.10	-0.16	0.02	-0.04	0.69	0.21	4
N-h	-0.40	-0.29	-0.24	-0.06	0.08	—	0.12	0.57	0.27	-0.29	0.24	-0.36	-0.23	0.28	0.47	0.27	0.29	6
N-n	-0.14	-0.24	-0.30	-0.03	0.02	0.36	—	0.02	-0.08	0.06	-0.22	0.22	-0.02	-0.23	-0.11	0.30	0.15	2
N-t	-0.37	0.00	-0.30	-0.21	-0.01	0.42	0.09	—	0.29	0.09	0.54	0.07	-0.03	0.05	0.73	-0.26	0.01	11
P-a	-0.38	0.20	-0.42	0.02	-0.01	0.19	0.02	0.41	—	-0.44	0.32	-0.42	-0.35	0.40	-0.12	-0.20	0.64	6
PAC	0.40	-0.18	0.08	0.05	-0.02	-0.15	-0.03	0.01	-0.33	—	0.34	0.72	0.86	-0.51	0.26	-0.37	-0.31	8
pH	0.07	-0.27	0.00	0.19	-0.08	0.11	0.13	0.19	0.02	0.45	—	0.07	0.48	0.10	0.37	-0.72	-0.02	12
S	-0.10	0.11	-0.46	-0.31	-0.08	0.08	0.28	0.17	0.02	0.10	0.14	—	0.59	0.84	0.11	-0.17	-0.31	7
SiO	0.23	-0.30	0.13	0.16	-0.12	-0.01	0.12	0.05	-0.27	0.75	0.65	0.04	—	-0.44	0.16	-0.43	-0.25	9
SL	-0.21	-0.10	0.23	0.68	0.01	0.04	-0.15	-0.19	0.09	0.08	0.25	-0.38	0.17	—	-0.02	0.02	0.26	6
SOM	0.03	-0.05	-0.25	-0.31	-0.05	0.32	0.06	0.89	0.33	0.06	0.16	0.06	0.06	-0.26	—	-0.12	-0.24	10
y1	-0.16	0.01	0.05	0.01	0.08	0.20	0.07	-0.05	-0.08	-0.22	-0.16	-0.08	-0.16	-0.11	-0.03	—	0.16	10
Zn	0.00	-0.01	0.40	-0.06	0.02	0.01	-0.04	-0.34	-0.32	-0.29	-0.20	-0.33	-0.20	-0.16	-0.28	0.37	—	3
Nc	5	0	9	3	0	4	0	4	3	9	3	3	5	3	3	0	2	—

N<sub>C</sub>: Number of confirmed correlations (0.40 or higher in grey).

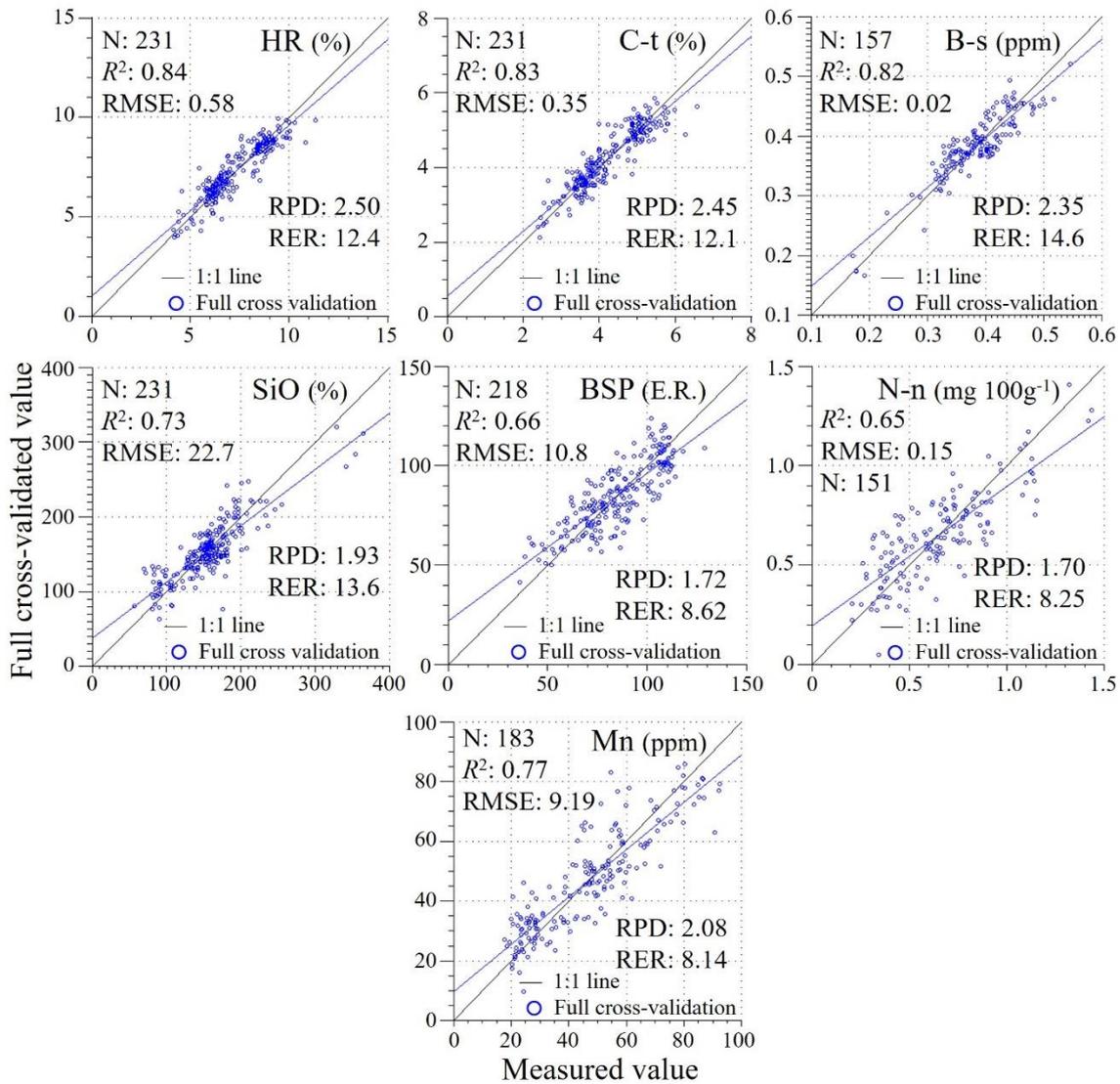
**Table 2.** Accuracy of the 34 regression models and evaluation of the full cross-validated results.

Property	N	S.P.	F	Range		Calibration		Full Cross-Validation									
				Min	Max	RMSE	R <sup>2</sup>	S.D.	Range	RMSE	R <sup>2</sup>	RPD		RER			
BSP	218	13	10	36.0	–	129	7.78	0.82	18.5	92.7	10.8	0.66	C	1.72	D	8.62	D
B-s	157	33	2	0.17	–	0.55	0.03	0.82	0.06	0.37	0.03	0.82	B	2.35	C	14.6	C
Ca	219	19	8	259	–	1486	131	0.82	310	1231	152	0.76	C	2.04	C	8.09	D
CaMg	224	15	7	2.39	–	20.3	2.12	0.82	5.04	17.9	2.47	0.76	C	2.05	C	7.26	E
CEC	205	13	7	29.0	–	50.6	2.02	0.82	4.77	23.2	2.52	0.72	C	1.89	D	9.23	D
CL	184	15	8	18.2	–	37.3	1.34	0.83	3.29	19.2	1.65	0.75	C	1.99	D	11.6	C
CN	178	27	6	10.7	–	13.6	0.22	0.82	0.53	2.98	0.24	0.79	C	2.20	C	12.4	C
CSP	213	7	5	28.4	–	110	8.22	0.82	19.5	81.6	9.93	0.74	C	1.96	D	8.21	D
C-t	231	15	9	2.39	–	6.57	0.26	0.90	0.84	4.18	0.35	0.83	B	2.45	C	12.1	C
Cu	116	15	5	0.24	–	0.73	0.04	0.81	0.09	0.49	0.05	0.71	C	1.86	D	10.6	C
DD	231	19	14	0.66	–	1.04	0.03	0.86	0.09	0.38	0.05	0.69	C	1.80	D	7.35	E
EC	120	5	4	0.09	–	0.18	0.01	0.82	0.02	0.09	0.01	0.67	C	1.73	D	7.70	E
ESP	173	27	6	0.06	–	0.38	0.03	0.82	0.07	0.32	0.03	0.79	C	2.16	C	10.4	C
Fe	143	31	6	4.01	–	5.33	0.12	0.82	0.29	1.33	0.14	0.78	C	2.15	C	9.75	D
HR	231	27	8	4.12	–	11.3	0.51	0.88	1.46	7.20	0.58	0.84	B	2.50	B	12.4	C
K	192	29	7	22.1	–	110	9.08	0.82	21.5	87.8	10.1	0.78	C	2.12	C	8.67	D
MC	166	31	6	35.4	–	50.4	1.27	0.82	2.99	14.9	1.38	0.79	C	2.17	C	10.8	C
Mg	183	29	7	28.9	–	114	8.12	0.83	19.5	84.6	9.20	0.78	C	2.12	C	9.20	D
MgK	155	17	7	1.33	–	5.30	0.37	0.83	0.89	3.97	0.46	0.73	C	1.93	D	8.58	D
Mn	183	21	7	17.6	–	92.5	7.90	0.83	19.1	74.8	9.19	0.77	C	2.08	C	8.14	C
Na	182	17	7	1.17	–	5.02	0.36	0.82	0.86	3.85	0.42	0.76	C	2.04	C	9.08	C
N-a	199	11	8	0.31	–	5.83	0.56	0.82	1.34	5.52	0.77	0.67	C	1.73	D	7.15	E
N-h	218	23	10	2.20	–	12.2	0.82	0.82	1.96	10.0	1.01	0.74	C	1.94	D	9.95	D
N-n	151	21	10	0.21	–	1.43	0.11	0.82	0.25	1.23	0.15	0.65	D	1.70	D	8.25	D
N-t	229	33	8	0.19	–	0.49	0.03	0.83	0.06	0.30	0.03	0.79	C	2.16	C	10.3	C
P-a	151	31	5	0.62	–	54.4	5.17	0.82	12.3	53.7	5.55	0.80	C	2.21	C	9.67	D
PAC	231	13	8	1540	–	2699	79.1	0.86	215	1159	102	0.77	C	2.10	C	11.4	C
pH	212	13	7	5.52	–	7.44	0.20	0.82	0.47	1.92	0.24	0.73	C	1.92	D	7.90	E
S	165	21	5	24.4	–	52.3	1.79	0.82	4.23	28.0	2.00	0.78	C	2.12	C	14.0	C
SiO	231	13	8	56.7	–	364	17.5	0.84	43.8	307	22.7	0.73	C	1.93	D	13.6	C
SL	134	21	6	27.7	–	38.3	0.97	0.82	2.31	10.5	1.12	0.77	C	2.07	C	9.44	D
SOM	219	33	8	13.8	–	21.3	0.70	0.82	1.65	7.42	0.77	0.79	C	2.15	C	9.66	D
y1	217	23	14	0.06	–	1.19	0.10	0.84	0.26	1.13	0.15	0.67	C	1.75	D	7.70	E
Zn	165	29	5	2.24	–	13.2	0.94	0.82	2.24	11.0	1.03	0.79	C	2.18	C	10.7	C

N: Number of samples; S.P.: smoothing points; F: PLSR factor; A, B, C, D, and E: accuracy indicator categories.



**Figure 6.** Average absorbance spectra of 231 samples after 2D<sub>S-G</sub> pretreatment. The baselines of the absorbance coefficients have been shifted and the amplitudes adjusted. S.P.: smoothing points of 2D<sub>S-G</sub>. S.P. = 0: absorbance spectra before 2D<sub>S-G</sub> pretreatment.



**Figure 7.** Full cross-validation scatter plots for seven soil properties.

PAC showed response variable correlation with CL, Cu, DD, MC, P-a, S, SiO, and SL (Table 1). Furthermore, PAC showed correlation between local regression model coefficients for BSP, CL, CN, Fe, K, MC, Mg, pH, and SiO (Table 1). PAC's response variable and local regression coefficient correlation showed a positive correlation with MC and SiO, and a negative correlation with CL. In particular, the correlation with 5 SPs is very important. Given these results, even if the PAC has no absorption wavelength, it is assumed that the PAC local regression model is reliable. The same applies to the other local regression models in Table 1. In contrast, Mn showed no correlation with the measured values of any soil property (Table 1). However, in the local regression model coefficients, a negative correlation was confirmed with Ca, CaMg, CN, HR, P-a, and S; a positive correlation was confirmed with Mg, K, and Zn (Table 1). Thus, only nine coefficient correlations were confirmed, and the Mn local regression model was assumed to be of low reliability.

### 3.2. Evaluation and Accuracy of Predicted Values for Unknown Samples

The accuracy of the predicted values for the 65 unknown samples, using the 34 local regression models, is shown in Table 3. In the case of  $R^2$ , RPD, and RPIQ, the local regression models for 10 soil properties (CaMg, CL, C-t, HR, MC, N-t, PAC, S, SOM and SiO) were classified as Category C or D, with the other 24 local regression models classified as Category E. The RER of HR, N-t, and PAC was Category D, while the others were Category E. The RER of HR, N-t, PAC was Category D; others were Category E. Therefore, in the overall dataset for this experiment, the predicted values for HR, N-t, and PAC, in the 65 unknown samples, using the local regression models, were considered to 'distinguish between high and low', as absolute quantitative values for each sample. The predicted values for properties other than HR, N-t, and PAC were evaluated to be 'not useful' as absolute quantitative values for each sample. To investigate why the accuracy of the predictions decreased, the transition in the measured and predicted values for the 34 soil properties in the 65 unknown samples is shown in Figure 8, where the two blue lines are the upper and lower predictable limits (predictable range), the black line is the measured value, and the red line is the predicted value. Large error is confirmed in the case of 22 soil properties (B-s, BSP, Ca, CEC, CN, CSP, Cu, EC, ESP, Fe, MC, Mg, K, MgK, Na, N-n, P-a, PAC, pH, SL, y1, and Zn), where the unknown samples fell outside the predictable range of the respective local regression models. This result suggests that each of these local regression models must be re-calibrated by PLSR, using a new dataset that includes the unknown samples outside the predictable range. The mean of the measured values (M.M.) and predicted values (M.P.), and error percentage (E.P.), are shown in Table 3. A total of 15 properties (B-s, BSP, CEC, Cu, EC, ESP, MgK, Na, N-a, N-h, N-n, P-a, SiO, y1, and Zn) showed an error of 10% or more, suggesting that bias adjustment is required. On the other hand, it was visually confirmed that the predicted values followed the transition in the measured values. Therefore, it was determined that these predicted values estimated by the local regression models could be used for soil maps displaying relative quantitative values. In particular, in this study, 13 soil properties (CL, CN, C-t, DD, Fe, HR, MC, N-t, PAC, S, SL, SiO, and SOM) could be used as relative quantitative values.

**Table 3.** Evaluation of the predicted values for 34 soil properties estimated using the local regression models.

Property	Prediction (N = 65)														
	S.D.	RMSE	Range	IQ	$R^2$		RPD		RPIQ		RER		M.M.	M.P.	E.P.(%)
BSP	22.9	25.1	105	21.1	0.08	E	0.91	E	0.84	E	4.17	E	78.7	89.6	13.9
B-s	0.15	0.14	0.76	0.14	0.41	E	1.06	E	0.99	E	5.36	E	0.45	0.38	16.1
Ca	318	360	1346	361	0.08	E	0.88	E	1.00	E	3.74	E	922	970	5.28
CaMg	3.15	3.39	11.0	5.48	0.27	E	0.93	E	1.62	D	3.25	E	9.89	9.91	0.24
CEC	4.15	7.31	19.6	5.44	0.00	E	0.57	E	0.74	E	2.67	E	46.8	41.3	11.7
CL	3.97	3.10	15.8	6.49	0.39	E	1.28	E	2.09	C	5.09	E	27.1	27.1	0.01
CN	0.50	0.44	2.46	0.46	0.34	E	1.13	E	1.03	E	5.55	E	12.7	12.6	0.45
CSP	22.0	26.1	99.6	25.8	0.05	E	0.84	E	0.99	E	3.82	E	69.5	65.9	5.11
C-t	0.79	0.51	3.93	0.72	0.65	D	1.54	D	1.41	E	7.68	E	4.69	4.47	4.55
Cu	0.49	0.55	2.05	0.23	0.06	E	0.90	E	0.42	E	3.73	E	0.49	0.32	36.1
DD	0.04	0.05	0.21	0.06	0.26	E	0.83	E	1.14	E	3.98	E	0.75	0.78	3.83
EC	0.04	0.05	0.17	0.05	0.01	E	0.82	E	0.93	E	3.47	E	0.15	0.13	16.0
ESP	0.14	0.21	0.92	0.13	0.00	E	0.65	E	0.60	E	4.34	E	0.31	0.15	50.8
Fe	0.27	0.29	1.72	0.33	0.15	E	0.95	E	1.14	E	6.01	E	4.54	4.49	1.15
HR	1.36	0.69	6.77	1.24	0.77	C	1.98	D	1.80	D	9.83	D	8.08	8.24	2.02
K	19.4	18.9	88.5	23.1	0.12	E	1.02	E	1.22	E	4.67	E	40.6	43.3	6.52
MC	6.43	4.41	29.2	7.65	0.68	C	1.46	E	1.73	D	6.62	E	46.1	45.7	0.70
Mg	23.2	26.1	121	18.7	0.02	E	0.89	E	0.71	E	4.62	E	69.1	69.3	0.34
MgK	2.03	2.31	9.40	2.15	0.03	E	0.88	E	0.93	E	4.07	E	4.56	3.41	25.2
Mn	9.02	13.3	37.9	8.94	0.01	E	0.68	E	0.67	E	2.85	E	38.9	36.2	6.99
Na	2.12	3.25	14.2	1.82	0.03	E	0.65	E	0.56	E	4.37	E	4.54	2.20	51.6
N-a	0.36	1.24	2.18	0.37	0.10	E	0.29	E	0.30	E	1.75	E	0.62	1.38	121
N-h	1.49	1.78	9.39	1.26	0.19	E	0.84	E	0.71	E	5.27	E	5.67	6.47	14.2
N-n	0.95	0.95	4.13	0.60	0.28	E	1.00	E	0.63	E	4.34	E	1.14	0.72	36.9
N-t	0.06	0.04	0.28	0.05	0.67	C	1.64	D	1.39	E	8.04	D	0.37	0.37	0.54
P-a	24.6	35.6	104	30.7	0.05	E	0.69	E	0.86	E	2.94	E	48.2	22.3	53.7
PAC	243	132	1078	314	0.78	C	1.84	D	2.38	C	8.16	D	2160	2100	2.76
pH	0.57	0.74	2.40	0.46	0.00	E	0.77	E	0.63	E	3.25	E	7.01	6.68	4.67
S	5.58	3.83	23.7	9.19	0.53	D	1.46	E	2.40	C	6.19	E	39.9	39.8	0.34
SiO	40.6	35.2	208	23.9	0.54	D	1.15	E	0.68	E	5.90	E	141	162	15.0
SL	2.76	3.22	12.3	3.66	0.00	E	0.86	E	1.14	E	3.82	E	33.0	32.4	1.95
SOM	1.42	1.14	6.79	1.50	0.53	D	1.25	E	1.31	E	5.96	E	18.1	18.7	3.20
y1	0.39	0.41	2.34	0.11	0.01	E	0.96	E	0.28	E	5.76	E	0.26	0.32	21.9
Zn	3.32	3.42	15.2	2.64	0.06	E	0.97	E	0.77	E	4.43	E	8.32	7.48	10.1

M.M.: mean of 65 samples' measured values; M.P.: mean of 65 samples' predicted values; E.P.: error percentage of predicted values to measured values; A, B, C, D, and E: accuracy indicator categories.

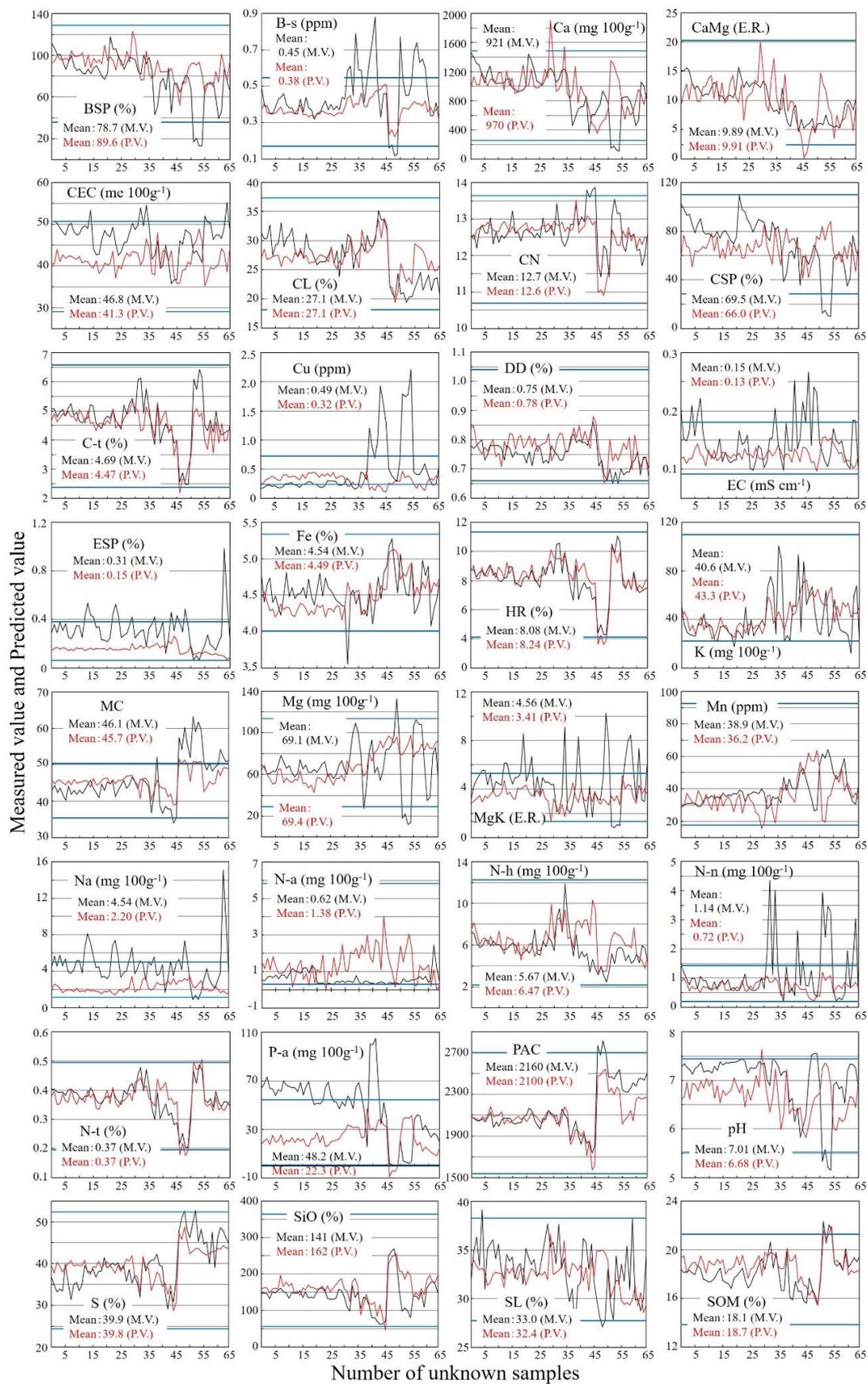
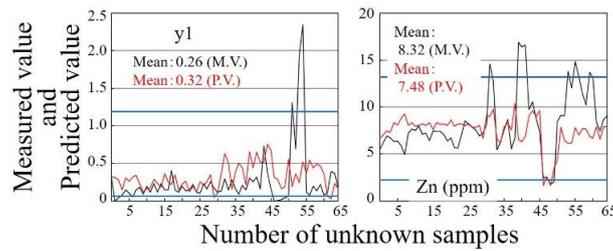


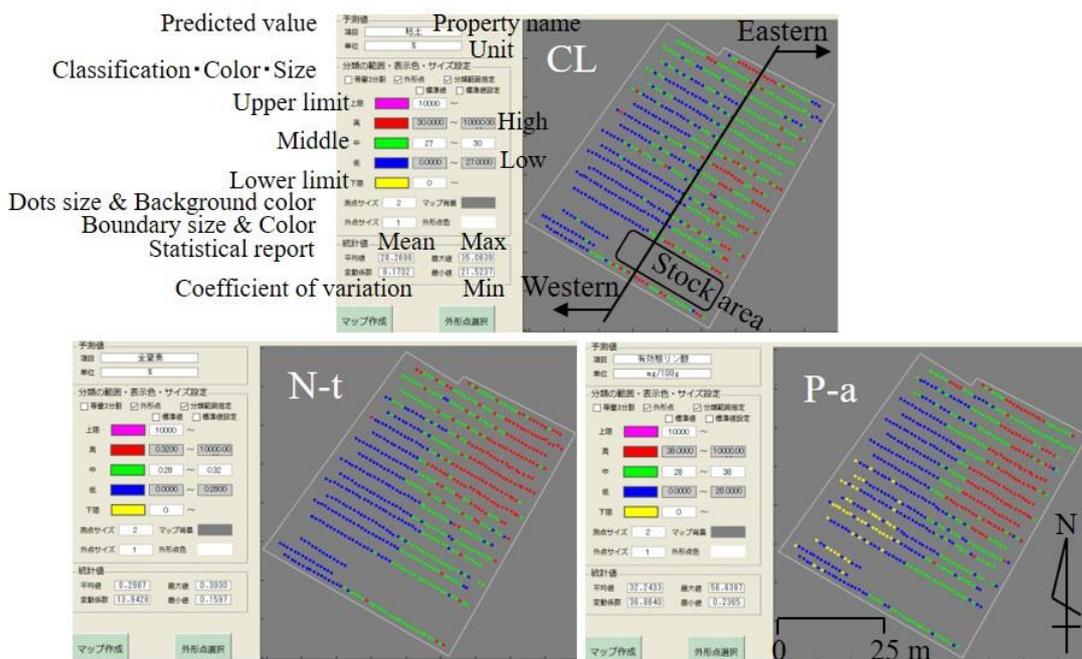
Figure 8. Cont.



**Figure 8.** Transition in the measured and predicted values of the 34 soil properties for 65 unknown samples. The blue lines represent the upper and lower predictable limits. The black line is the measured value and the red line the predicted value. M.V.: measured value; P.V.: predicted value.

### 3.3. In-Situ Digital Soil Mapping

In precision agriculture, it is necessary to understand site-specific variability in agricultural fields, and growers require in-situ digital soil mapping of relative quantitative values. After soil sensing by the SAS3000, several soil property maps were provided at the field, using SMV, to a taro grower, who then focused on the soil variability of CL, N-t, and P-a in his field (Figure 9). The soil variation among CL, N-t, and P-a in the field was divided into eastern and western regions. At that time, the grower noted that the variability was consistent with top soil dressing because, in the past, fresh red soil was brought in only in the eastern region. Additionally, the growth, quality, and yield of taro also differed between the east and west regions. In the grower’s experience, the western region contained a large amount of gravel and small taros. To improve the variability in the yield, and the quality and efficiency of the farm work, one scenario for precision agricultural soil management in this field might involve site-specific fertilization through topdressing and basal dressing, top-soil dressing of the western region, changing the ridge direction the field, etc. In fact, the grower changed the ridge direction in the field from east-west to north-south, resulting in greater farm work efficiency (for example, top dressing, pest control, and harvesting), with no soil improvement costs. Note that, the area without dots in the southeastern part of the field does not indicate a breakdown of the SAS3000, but rather a region in which taro plants were stored underground, and which was not measured.



**Figure 9.** Visualization of soil variability in a field using Soil Map Viewer. CL: clay; N-t: total nitrogen; P-a: available phosphorus.

## 4. Discussion

### 4.1. Regression Model Accuracy and Actual Operation

With the range of prediction for each soil property, the local regression models is shown as 'Range' in Table 2. In the soil diagnosis of upland fields in Japan, the reference value range is 10 to 30 mg 100 g<sup>-1</sup> for P-a (according to the Agricultural Policy Division, [62] standard). In this case, the local regression model enables effective soil diagnosis. However, since the reference value range is 2.0 to 40 ppm for Zn, the Zn local regression model is not usable, which requires a new local regression model for Zn, by adding a dataset with multiple Zn measured data of 13 ppm and more (e.g., 14, 15..., 40 ppm or more). The same logic applies in the case of the other local regression models.

### 4.2. Performance Indicators and Regression Model

Parameters  $R^2$ , RPD, RER, and RPIQ are performance indicators for evaluating the absolute quantitative accuracy of the predicted values, and it is important to note that the criteria for defining 'excellent' and 'not useful' models using such performance indicators are rather arbitrary, and there is no statistical basis for how these classification thresholds are determined [63]. Shree et al. [63] suggest that it is more important for a user to evaluate a model's performance with respect to the specific objectives of the given study, and we concur.

### 4.3. Digital Soil Map for Growers

In this study, CL, N-t, and P-a soil maps were used for grower decision making. The properties of soil maps required vary depending on the crops and varieties, the agricultural equipment available, and the philosophy and cultivation experience of the grower. The six growers suggested that it was important to understand the relative variability of multiple soil properties within a given field and between fields. As a specification from growers, mobile proximal soil sensor has been required that can provide various soil properties maps. Therefore, it is highly important, for sustainable farm management, that a mobile soil sensing system has the ability to provide multiple soil property maps.

### 4.4. Potential of a Mobile Proximal Soil Sensor

The results of this study, using the SAS3000, are particularly promising in suggesting that the method can be used for in-situ digital soil mapping, and rapidly, cheaply, and with less labor compared to conventional laboratory measurement, and the grower's use of the relative quantitative soil maps to modify his crop management suggests the possibility of using the SAS3000 as information technology for precision agriculture. As potential applications of the proposed method, the local regression models, and SMV for digital soil mapping can be used not only for variable-rate fertilization, crop selection, land use, and site-specific soil management, but also for effective monitoring of soil organic carbon (e.g., SOM and C-t), and to provide a sound analytical foundation for agronomic practices that enable carbon sequestration and reduction in greenhouse gas emissions [10,64–66].

## 5. Conclusions

The dataset used in this study comprised soil analysis values and VINR soil reflectance spectra, coupled with GNSS data, from the rotational upland fields of six growers in Saitama Prefecture, Japan. The local regression models for 30 chemical and 4 physical soil properties, which showed an  $R^2$  accuracy greater than 0.81, were estimated using 2nd derivative pretreatment and PLSR. The evaluation of full cross-validation, based on the  $R^2$  and RPD performance indicators, revealed no 'excellent' or 'not useful' local regression model estimations, while six soil properties were classified as 'screening' by the RER indicator. In the performance indicator-based evaluation of the predicted values of 34 soil properties for 65 unknown samples, estimated using the local regression models, 3 of the 34 models could 'distinguish between high and low', but the other 31 local regression models were 'not useful',

as absolute quantitative values. The study demonstrated that local regression model-based predictions for 13 of the 34 soil properties in the 65 unknown samples could provide relative quantitative values for the properties, based on the comparative transition in measured values and predicted values. The prediction error percentage (10% or less) for the 65 unknown samples confirmed that 20 soil properties could be successfully estimated as relative quantitative values, suggesting that the proposed method could generate useful digital soil maps displaying relative quantitative soil property values. Therefore, we concluded that the SAS3000 digital soil mapping equipment provided relative quantitative accuracy.

**Author Contributions:** M.K. and S.S. designed the study and performed the experiments; M.K. analyzed the results and wrote the article; S.S. contributed to the manuscript editing process. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by grants from the Project of the NARO Bio-oriented Technology Research Advancement Institution (project ID16822280 of the special scheme project on regional developing strategy), and partially funded by the Ministry of Agriculture, Forestry and Fisheries, Japan.

**Acknowledgments:** We thank APCRL and SCAS for analyzing measure to the soil properties, and warmly thank all those who assisted us in the study. Finally, we would like to warmly thank the growers who provided the upland fields for the study.

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

## References

1. Adamchuk, V.I.; Hummel, J.W.; Morgan, M.T.; Upadhyaya, S.K. On-the-go soil sensors for precision agriculture. *Comput. Electron. Agric.* **2004**, *44*, 71–91. [[CrossRef](#)]
2. Shibusawa, S. Precision farming Japan model. *Agric. Inf. Res.* **2003**, *12*, 125–133.
3. Debaene, G.; Niedźwiecki, J.; Pecio, A. Visible and near-infrared spectrophotometer for soil analysis: Preliminary results. *Pol. J. Agron.* **2010**, *3*, 3–9.
4. Mouazen, A.M.; Alhwaitmel, S.A.; Kuang, B.; Waive, T. Multiple on-line soil sensors and data fusion approach for delineation of water holding capacity zones for site specific irrigation. *Soil Tillage Res.* **2014**, *143*, 95–105. [[CrossRef](#)]
5. Tümsavas, Z.; Tekin, Y.; Ulusoy, Y.; Mouazen, A.M. Prediction and mapping of soil clay and sand contents using visible and near-infrared spectroscopy. *Biosyst. Eng.* **2019**, *177*, 90–100. [[CrossRef](#)]
6. Bernard, G.B.; Didier, B.; Bodovololona, R.; Oumar, B.; Cécile, V. Near infrared reflectance spectroscopy (NIRS) could be used for characterization of soil nematode community. *Soil Biol. Biochem.* **2011**, *43*, 1649–1659.
7. Matsunaga, T.; Uwasawa, M. Application of near infrared spectrometry to quantitative analysis of soil physical and chemical properties. *Jpn. J. Soil Sci. Plant Nutr.* **1992**, *63*, 712–714.
8. Chang, C.-W.; Laird, D.A.; Mausbach, M.J.; Hurburgh, C.R., Jr. Near-infrared reflectance spectroscopy-principal components regression analysis of soil properties. *Soil Sci. Soc. Am. J.* **2001**, *65*, 480–490. [[CrossRef](#)]
9. Islam, K.; Singh, B.; McBratney, A. Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. *Aust. J. Soil Res.* **2003**, *41*, 1101–1114. [[CrossRef](#)]
10. Stenberg, B. Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. *Geoderma* **2010**, *158*, 15–22. [[CrossRef](#)]
11. 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](#)]
12. Nawar, S.; Buddenbaum, H.; Hill, J.; Kozak, J.; Mouazen, A.M. Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. *Soil Tillage Res.* **2016**, *155*, 510–522. [[CrossRef](#)]
13. Adeline, K.R.M.; Gomez, C.; Gorretta, N.; Roger, J.-M. Predictive ability of soil properties to spectral degradation from laboratory Vis-NIR spectroscopy data. *Geoderma* **2017**, *288*, 143–153. [[CrossRef](#)]
14. 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](#)]
15. Śliwińska, A.; Smolinski, A.; Kucharski, P. Simultaneous analysis of heavy metal concentration in soil samples. *Appl. Sci.* **2019**, *9*, 4705. [[CrossRef](#)]

16. Recena, A.; Fernández-Cabanás, V.M.; Delgado, A. Soil fertility assessment by Vis-NIR spectroscopy: Predicting soil functioning rather than availability indices. *Geoderma* **2019**, *337*, 368–374. [[CrossRef](#)]
17. Sudduth, K.A.; Hummel, J.W. Soil organic matter, CEC, and moisture sensing with a portable NIR spectrophotometer. *Trans. ASAE* **1993**, *36*, 1571–1582. [[CrossRef](#)]
18. Shibusawa, S.; Hirako, S.; OTOMO, A.; Li, M.Z. Real-time underground soil spectrophotometer. *J. Jpn. Soc. Agric. Mach.* **1999**, *61*, 131–133.
19. Mouazen, A.M.; Maleki, M.R.; De Baerdemaeker, J.; Ramon, H. On-line measurement of some selected soil properties using a VIS-NIR sensor. *Soil Tillage Res.* **2007**, *93*, 13–27. [[CrossRef](#)]
20. Christy, C.D. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. *Comput. Electron. Agric.* **2008**, *61*, 10–19. [[CrossRef](#)]
21. Rodionov, A.; Welp, A.; Damerow, L.; Berg, T.; Amelung, W.; Pätzold, S. Towards on-the-go field assessment of soil organic carbon using Vis–NIR diffuse reflectance spectroscopy: Developing and testing a novel tractor-driven measuring chamber. *Soil Tillage Res.* **2015**, *145*, 93–102. [[CrossRef](#)]
22. Maleki, M.R.; Mouazen, A.M.; De Ketelaere, B.; Ramon, H.; De Baerdemaeker, J. On-the-go variable-rate phosphorus fertilisation based on a visible and near-infrared soil sensor. *Biosyst. Eng.* **2008**, *99*, 35–46. [[CrossRef](#)]
23. Umeda, H.; Shibusawa, S.; Okayama, T.; Sakuma, D.Y.; Kaho, T.; Ninomiya, K. Study of the precision farming with soil maps describing environmental load using a real-time soil sensor. *J. Jpn. Soc. Agric. Mach.* **2011**, *73*, 37–44.
24. Tekin, Y.; Kuang, K.; Mouazen, A.M. Potential of on-line visible and near infrared spectroscopy for measurement of pH for deriving variable rate lime recommendations. *Sensors* **2013**, *13*, 10177–10190. [[CrossRef](#)]
25. 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](#)]
26. Kodaira, M.; Shibusawa, S.; Ninomiya, K.; Kato, Y. Farm mapping techniques for effective soil management in large-scale farming. *Agric. Inf. Res.* **2009**, *18*, 110–121. [[CrossRef](#)]
27. Kodaira, M.; Shibusawa, S. Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping. *Geoderma* **2013**, *199*, 64–79. [[CrossRef](#)]
28. Kodaira, M.; Shibusawa, S. Multivariate Regression model estimation and soil mapping of multiple soil properties by using a tractor-mounted soil analyzing system. *J. Jpn. Soc. Agric. Mach. Food Eng.* **2016**, *78*, 401–415.
29. Kodaira, M.; Shibusawa, S. Development of calibration models for SiO<sub>2</sub>, Fe<sub>2</sub>O<sub>3</sub>, and Na<sub>2</sub>O using subsurface diffuse reflectance spectra. *J. Jpn. Soc. Agric. Mach. Food Eng.* **2018**, *80*, 114–122.
30. Kodaira, M.; Shibusawa, S. Estimation of nitrogen available using soil spectra on paddy subsurface. *J. Jpn. Soc. Agric. Mach. Food Eng.* **2020**, *82*, 85–87.
31. Japanese Geotechnical Society. *Soil Test-Basics and Guidance*, 2nd ed.; The Japanese Geotechnical Society; Maruzen: Tokyo, Japan, 2011; pp. 18–66.
32. Agricultural Research Division. *Analytical Methods for Soil and Crop Nutrition Diagnosis 2012*; Local Independent Administrative Agency Hokkaido Research Organization: Hokkaido, Japan, 2012.
33. Imai, H. *Soil Nutrient Analysis*, 15th ed.; Available Silicate; Yokendo: Tokyo, Japan, 2002; pp. 278–279.
34. Imai, H. *Soil Nutrient Analysis*, 15th ed.; Free Iron Oxide; Yokendo: Tokyo, Japan, 2002; pp. 329–330.
35. Palmer, K.F.; Williams, D. Optical properties of water in the near infrared. *J. Opt. Soc. Am.* **1974**, *64*, 1107–1110. [[CrossRef](#)]
36. Soga, N.; Ota, R.; Nasu, H.; Yoshimoto, M. Preparation of and optical properties of amorphous solids containing high concentration of transition metal oxides. *J. Soc. Mater. Sci. Jpn.* **1982**, *31*, 986–991. [[CrossRef](#)]
37. Matsunaga, T.; Uwasawa, M. Near infrared diffuse reflectance spectra of the upland soils in Japan. *Jpn. J. Soil Sci. Plant. Nutr.* **1992**, *63*, 403–410.
38. Matsunaga, T.; Uwasawa, M. Near infrared diffuse reflectance spectra of clay minerals. *Jpn. J. Soil Sci. Plant Nutr.* **1993**, *64*, 329–331.
39. Uchida, K.; Inoue, T. A practical alteration zoning based upon in-situ reflective spectroscopy measurement of rook samples in the vicinity of the Omori mine, Shimane prefecture, Japan. *J. Soc. Resour. Geol.* **1993**, *43*, 65–77.

40. 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](#)]
41. Kumada, K. Absorption spectra of humic acids. *Soil Sci. Plant Nutr.* **1955**, *1*, 29–30. [[CrossRef](#)]
42. Kumada, K.; Sato, O. Chromatographic separation of green humic acid from podzol humus. *Soil Sci. Plant Nutr.* **1962**, *8*, 31–33. [[CrossRef](#)]
43. Mamiya, M. The near-infrared absorption spectra of ammonium ion in the solution. *Jpn. Anal.* **1965**, *14*, 253–258. [[CrossRef](#)]
44. Waiser, T.H.; Morgan, C.L.S.; Brown, D.J.; Tom, H.C. In situ characterization of soil clay content with visible near-infrared diffuse reflectance spectroscopy. *Soil Sci. Soc. Am. J.* **2007**, *71*, 389–396. [[CrossRef](#)]
45. Iwamoto, M.; Kawano, S.; Uozumi, J. *An Introduction of Near-Infrared Spectroscopy*; Saiwai-Shobo: Tokyo, Japan, 2002.
46. Ozaki, Y.; Kawata, S. *Near-Infrared Spectroscopy*; Japan Scientific Societies Press: Tokyo, Japan, 2008.
47. Mohanty, B.; Gupta, A.; Das, B.S. Estimation of weathering indices using spectral reflectance over visible to mid-infrared region. *Geoderma* **2016**, *266*, 111–119. [[CrossRef](#)]
48. Isono, Y.; Kimura, T.; Nakajima, S. Study of Near-Infrared Diffuse Reflectance Spectroscopy to Engineering Geological Material. Japan Society of Engineering Geology Chugoku-Shikoku Branch, Japan. Available online: <http://www.jseg.or.jp/chushikoku/ronnbunn/PDF/PDF23/2311.pdf> (accessed on 12 December 2019).
49. Clark, D.H.; Mayland, H.F.; Lamb, R.C. Mineral analysis of forages with near infrared reflectance spectroscopy. *Agron. J.* **1987**, *79*, 485–490. [[CrossRef](#)]
50. Umeda, T.; Koyama, F.; Takamuku, K. Determination of mineral concentrations of beef cattle manure with near infrared reflectance spectroscopy. *Bull. Fukuoka Agric. Res. Cent.* **2001**, *20*, 105–108.
51. Zornoza, R.; Guerrero, C.; Mataix-Solera, J.; Scow, K.M.; Arcenegui, V.; Mataix-Beneyto, J. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. *Soil Biol. Biochem.* **2008**, *40*, 1923–1930. [[CrossRef](#)] [[PubMed](#)]
52. Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [[CrossRef](#)]
53. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
54. Miyashita, Y.; Sasaki, S. *Chemometrics-Chemical Pattern Recognition and Multivariate Analysis*; Kyoritsu Shuppan: Tokyo, Japan, 1995; pp. 55–72.
55. Ozaki, Y.; Uda, A.; Akai, T. *Multivariate Analysis for Chemists Introduction to Chemometrics*; Kodansha: Tokyo, Japan, 2002; pp. 51–70.
56. Mouazen, A.M.; Kuang, B. On-line visible and near infrared spectroscopy for in-field phosphorous management. *Soil Tillage Res.* **2016**, *155*, 471–477. [[CrossRef](#)]
57. Ma, F.; Du, C.W.; Zhou, J.M.; Shen, Y.Z. Investigation of soil properties using different techniques of mid-infrared spectroscopy. *Eur. J. Soil Sci.* **2019**, *70*, 96–106. [[CrossRef](#)]
58. 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](#)]
59. 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. *Adv. Agron.* **2012**, *114*, 155–223.
60. Malley, D.; Martin, P.; Ben-Dor, E. Application in analysis of soils. In *Near Infrared Spectroscopy in Agriculture*; Roberts, C., Workman, J., Reeves, J., III, Eds.; A Three Society Monograph (ASA, SSSA, CSSA): Madison, WI, USA, 2004; pp. 729–784.
61. Kinoshita, R.; Bianca, N.M.-C.; Harold, M.E.; Hively, W.D.; Bilgili, A.V. Strategies for soil quality assessment using visible and near-infrared reflectance spectroscopy in a Western Kenya chronosequence. *Soil Sci. Soc. Am. J.* **2011**, *76*, 1776–1788. [[CrossRef](#)]
62. Agricultural Policy Division. *Hokkaido Fertilization Guide 2015*; Local Independent Administrative Agency Hokkaido Research Organization: Hokkaido, Japan, 2015.
63. Shree, R.S.D.; Jonathan, S.; Skye, W.; Leonardo, R.-L. Accurate and precise prediction of soil properties from a large mid-infrared spectral library. *Soil Syst.* **2019**, *3*, 11.
64. Lal, R. Soil carbon dynamics in cropland and rangeland. *Environ. Pollut.* **2002**, *116*, 353–362. [[CrossRef](#)]

65. Deng, F.; Minasny, B.; Knadel, M.; McBratney, A.; Heckrath, G.; Greve, M.H. Using Vis-NIR spectroscopy for monitoring temporal changes in soil organic carbon. *Soil Sci.* **2013**, *178*, 389–399. [[CrossRef](#)]
66. Viscarra Rossel, R.A.; Webster, R.; Bui, E.N.; Baldock, J.A. Baseline map of organic carbon in Australian soil to support national carbon accounting and monitoring under climate change. *Glob. Change Biol.* **2014**, *20*, 2953–2970. [[CrossRef](#)] [[PubMed](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).