


Article

Soil Nutrient Estimation and Mapping in Farmland Based on UAV Imaging Spectrometry

Xiaoyu Yang¹, Nisha Bao^{1,*} , Wenwen Li², Shanjun Liu¹, Yanhua Fu³ and Yachun Mao¹

¹ College of Resources and Civil Engineering, Northeastern University, Shenyang 110819, China; yangxiaoyu368@163.com (X.Y.); liusjdr@126.com (S.L.); maoyachun@mail.neu.edu.cn (Y.M.)

² School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ 85287, USA; wenwen@asu.edu

³ JangHo Architecture College, Northeastern University, Shenyang 110169, China; fuyanhua@mail.neu.edu.cn

* Correspondence: baonisha@hotmail.com

Abstract: Soil nutrient is one of the most important properties for improving farmland quality and product. Imaging spectrometry has the potential for rapid acquisition and real-time monitoring of soil characteristics. This study aims to explore the preprocessing and modeling methods of hyperspectral images obtained from an unmanned aerial vehicle (UAV) platform for estimating the soil organic matter (SOM) and soil total nitrogen (STN) in farmland. The results showed that: (1) Multiplicative Scattering Correction (MSC) performed better in reducing image scattering noise than Standard Normal Variate (SNV) transformation or spectral derivatives, and it yielded a result with higher correlation and lower signal-to-noise ratio; (2) The proposed feature selection method combining Successive Projections Algorithm (SPA) and Competitive Adaptive Reweighted Sampling algorithm (CARS), could provide selective preference for hyperspectral bands. Exploiting this method, 24 and 22 feature bands were selected for SOM and STN estimation, respectively; (3) The particle swarm optimization (PSO) algorithm was employed to obtain optimized input weights and bias values of the extreme learning machine (ELM) model for more accurate prediction of SOM and STN. The improved PSO-ELM model based on the selected preference bands achieved higher prediction accuracy (R^2 of 0.73 and RPD of 1.91 for SOM, R^2 of 0.63, and RPD of 1.53 for STN) than support vector machine (SVM), partial least squares regression (PLSR), and the ELM model. This study provides an important guideline for monitoring soil nutrient for precision agriculture with imaging spectrometry.

Keywords: unmanned aerial vehicle; hyperspectral image; extreme learning machine; soil nutrient estimation; feature selection



Citation: Yang, X.; Bao, N.; Li, W.; Liu, S.; Fu, Y.; Mao, Y. Soil Nutrient Estimation and Mapping in Farmland Based on UAV Imaging Spectrometry. *Sensors* **2021**, *21*, 3919. <https://doi.org/10.3390/s21113919>

Academic Editors: Triven Koganti, Ellen Van De Vijver, Maria Knadel, Bo Vangsø Iversen, Barry J. Allred and John Triantifilis

Received: 19 April 2021

Accepted: 4 June 2021

Published: 6 June 2021

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



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Soil organic matter (SOM) and soil total nitrogen (STN) are two important variables that reflect soil quality and soil fertility [1], as they can improve the physical, chemical, and biological properties of soil and provide humic acids and carbon sources for plant growth [2,3]. The geographical distribution of SOM and STN is useful information for crop growers. Therefore, accurate and timely monitoring of their geographical distribution is essential for farmland management in precision agriculture [4].

Due to the presence of functional groups such as C-H, -COOH, -OH, and N-H in soil organic compounds corresponding to the electromagnetic radiation response, there are obvious spectral characteristics in the visible and near-infrared (VIS-NIR) regions. This makes proximal reflectance spectrometry and imaging spectrometry using VIS-NIR beneficial for the quantification of soil properties. Currently, spectrometry in the laboratory has been widely applied to the quantitative inversion of SOM based on the organic matter sensitive bands that exist in the visible range of 550~770 nm and the near-infrared range of 1300~1500 nm [5,6]. A significant correlation was also found between soil spectra and nitrogen content at 1880~1890 nm and 510 nm [7].

However, the spectral characteristic and analysis in the laboratory on collected samples cannot provide a spatially continuous distribution of soil properties in a specific area [8]. Hyperspectral imaging spectrometry has the advantages of capturing rich spectral and spatial information as well as surface information [9]. Hyperspectral remote sensing exploits hyperspectral sensors from satellites, airplanes, and unmanned aerial vehicles (UAVs) to monitor the Earth's surface. The satellite hyperspectral sensors, such as EO-1 Hyperion can be exploited to predict soil organic carbon with a root mean square error (RMSE) of 0.73 at the regional scale [10]. However, EO-1 Hyperion data are only available before 2014, which limits the studies and applications of soil properties estimation in recent years. The GF-5 satellite with a spatial resolution of 30 m and 330 spectral bands was launched in 2018, and the data collected by this satellite have been used in soil estimation and monitoring. Meng et al. developed a regional-scale soil organic carbon prediction model ($R^2 = 0.79$, RPD = 1.46) for random forest using the spectral indices filtered from the hyperspectral data of GF-5 AHSI [11]. Generally, satellite data with moderate spatial resolution are suitable for large regional-scale estimation. As for field or small-scale monitoring, the prediction accuracy of soil properties cannot meet the requirements of precision agriculture management due to mixed pixel issues. The low temporal resolution of satellite-based hyperspectral imagery also limits its use in real-time monitoring of precision agriculture.

Airborne hyperspectral imagery provides data with a high spectral and spatial resolution that may be useful to map culture land which has different types of soils and crops over a large range [12]. Hbirkou et al. [13] predicted SOC over a small-scale bare and fine agricultural soil using the data obtained from aircraft-mounted hyperspectral sensor HyMap and achieved prediction accuracy of RMSEP = 0.76 g kg^{-1} and RPD = 2.08. The aircraft typically maintain an operation height between 1000 to 5000 m above ground to acquire high-resolution images (1 m~5 m resolution). Weather conditions and ground exposure are the primary factors that limit the use of hyperspectral imagers. The aircraft must fly when the sky is clear and there is an adequate exposure of surface soil [12]. Besides, there is often a long planning time for flight missions that are often associated with high operation cost [14].

Compared with aerial systems, the UAVs fly at lower altitudes and are not subject to air traffic control [15], and they have more advantages in practical applications because of their lower cost and higher operational flexibility, which is fundamental for precise agriculture monitoring [16]. UAV-based multispectral imaging is mostly applied to classification and mapping [17], but it fails to meet the needs of quantitative soil estimation due to the limitation of spectral resolution [18]. UAV-based hyperspectral imaging provides an alternative for detecting fine-scale soil properties. Hu et al. [19] demonstrated that the UAV hyperspectral system is effective and crucial for field-scale soil feature monitoring and mapping.

In addition, it is a very challenging research topic for collecting, processing, and analyzing hyperspectral images because of the embedded imagery noise, massive data volume, high dimensionality, and complexity in the data content [17,20,21]. Machine learning techniques, such as support vector machine (SVM), random forest (RF), and extreme learning machine (ELM) have been exploited to analyze hyperspectral features and investigate various soil characteristics. Partial least squares regression (PLSR) is a regression model that has been widely used for soil properties prediction based on hyperspectral features [22]. SVM is a powerful tool to analyze hyperspectral data because of its capability to very efficiently process multiple variables and solve non-linear problems [23]. Honkavaara et al. [24] applied SVM on hyperspectral imagery collected by UAV to predict crop biomass. ELM is a machine learning algorithm that is frequently used to train feed-forward neural networks which have a single hidden layer. The weights in the network is learned through solving a Moore–Penrose (MP) generalized inverse linear problem [25]. Compared with the traditional backpropagation neural network, ELM can provide a higher calculation speed and better generalization performance, as well as avoid issues in the gradient-based

training methods such as difficulties in determining a stopping criteria and a proper learning rate [26]. Ge et al. [27] concluded that the combination of preprocessed spectral indices and ELM algorithm allowed an estimation of soil moisture content with a high accuracy ($R^2_{\text{val}} = 0.907$) using hyperspectral imagery from UAV. Huang et al. [28] demonstrated that ELM tended to achieve better performance in terms of scalability and model generalization and it also yielded a much faster learning speed than support vector machines. However, more hidden neurons may lead ELM to converge slowly [29]. You et al. [30] exploited the use of the particle swarm optimization (PSO) for input weights selection to help ELM with fewer hidden neurons achieve good generalization performance. Compared to other optimization strategies, PSO is easy to implement and has a smaller parameter space [31].

This study focused on improving SOM and STN prediction and mapping accuracy through UAV hyperspectral imagery denoising, preference bands selection, and model optimization. This work aimed to (1) explore the optimal denoising method for UAV hyperspectral imagery by comparing the methods including Multiplicative Scattering Correction (MSC), Standard Normal Variate transformation (SNV), the first derivate (FD), and the second derivate (SD); (2) propose a hyperspectral bands selection method by combining Successive Projections Algorithm (SPA) and Competitive Adaptive Reweighted Sampling algorithm (CARS); (3) establish an accurate SOM prediction model by comparing the performance of traditional methods and improved extreme learning machine algorithm.

2. Data and Methods

2.1. Study Site and Field Sampling

The site ($42^{\circ}23' \text{ N}$ and $122^{\circ}57' \text{ E}$) chosen for this study was located in the north of Shenyang, Liaoning Province, Northeast China (Figure 1). The area belongs to the Songliao Plain, and it is mainly used as continuous agricultural fields as shown in Figure 1b. The area is predominantly flat with an altitude of approximately 60 m and has a continental monsoon climate in the North Temperate Zone. The average annual temperature of this area is 6.7°C , and the average annual precipitation is 600 mm. According to the Genetic Soil Classification of China [32], this area's soil types are mainly brown soil (Luvisols) and meadow soils (Cambisols), which are fertile, especially in terms of organic matter and nitrogen [33]. Crops planted in this area are mainly corn and peanut.

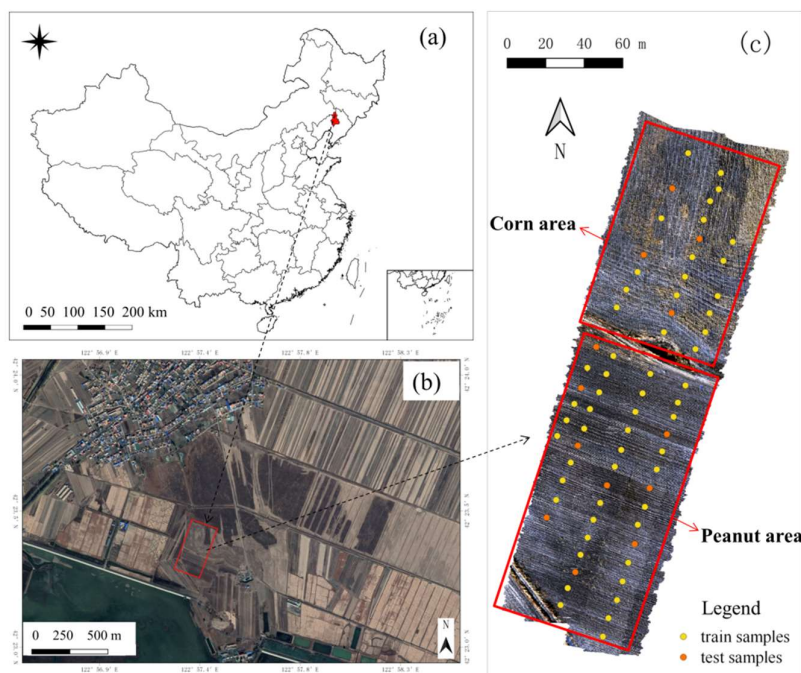


Figure 1. Geographical location of the study site in China (a), its surrounding environment (b), and the soil samples distribution (c).

The soil samples were collected following the grid sampling method, and 68 sampling cells (1 m × 1 m) were collected uniformly at 0~10 cm depth of the study soil in November 2019 after the crops were harvested. From November to April, the farmland is idle with no vegetation cover due to the cold weather, which is a period with good conditions for satellite and airborne image acquisition. A hand-held global positioning system (GPS) was used to record the geographic coordinates at each sampling site [34]. The soil samples were ground, air-dried, and sieved to a size less or equal to 0.25 mm. Finally, the content of SOM and STN was measured through the $K_2Cr_2O_7-H_2SO_4$ oxidation method and the modified Kjeldahl method, respectively [35].

As shown in Figure 2a, SOM content fell within a range of 11.4~30.6 g kg⁻¹, with a mean value of 19.72 g kg⁻¹. The coefficient of variation (CV)—the ratio between standard deviation and the mean value, was 24%. The Shapiro–Wilk (S–W) test confirmed that the SOM content followed a normal distribution. Meanwhile, as shown in Figure 2b, STN content fell within a range of 0.84~2.08 g kg⁻¹, with a mean value of 1.38 g kg⁻¹. The CV of STN content was 22%, and the S–W test confirmed that the STN content followed a normal distribution. The samples were split into 80% for training and 20% for test [36]. Fifty-five train samples and 13 test samples were selected using stratified random sampling based on the content of SOM and STN, then they were fine-tuned according to the statistical and spatial distribution of the samples, which makes the test samples follow a normal distribution and spatially representative as illustrated in Figure 1c. The statistical distribution of the training set and the test set is similar to the entire sample set of SOM and STN.

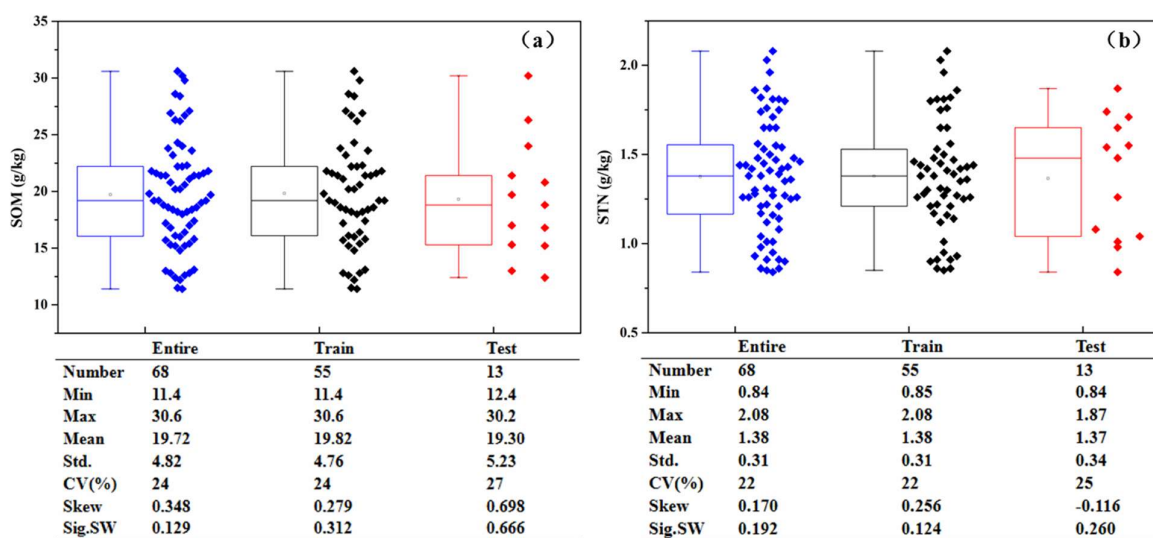


Figure 2. The statistical results of SOM (a) and STN (b) for the entire, training, and test datasets. CV indicates coefficients of variation (%), Skew indicates skewness, and Sig. SW indicates the significance of the Shapiro–Wilk normality test.

2.2. UAV Data Acquisition and Preprocessing

On the day when soil samples were collected, images were also obtained over the study area with Resonon Pika L hyperspectral camera loaded on the DJI M600 Pro UAV platform (Figure 3). The camera captured data in 281 spectral channels from 0.4 to 1.0 μm in the VIS-NIR region with a spectral resolution at 2.1 nm. In order to reduce noise and facilitate data transfer and fast processing, the spectral data were binned to 150 spectral bands each with 4 nm by the interpolation method of spline [37]. In this study, three impact factors were considered for setting up the flight altitude, including regional regulation and limitation of flight altitude, range of study area, and atmosphere influence. Therefore, flight height was set to 100 m, resulting in the spatial resolution of 0.1 m. Image overlap was set at 50% for the side (each parallel flight line). The obtained images were further mosaiced with SpectronPro software (Spectron Pro, Resonon Inc., Bozeman, MT, USA).

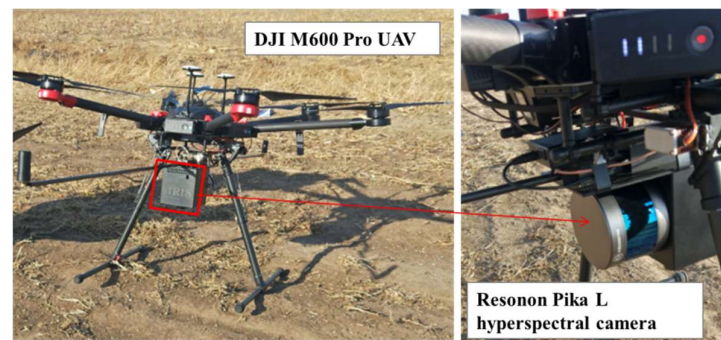


Figure 3. The UAV platform and the imaging hyperspectral sensor.

As shown in Figure 4, a target board was applied for radiation correction and reflectance transformation from the Digital Number (DN) of the original image. Then, a geometric correction was performed on the recorded data with SpectronPro, based on the onboard GPS and inertial measurement unit (IMU). The geometric correction accuracy was improved by base stations set up on the ground based on post-differential GPS processing [38]. It should be noted that the extremely high spatial resolution of UAV images may cause noises such as field monopoly shadows in the quantitative estimation of soil properties [19]. In our case, the spatial distance between two adjacent field monopoly was approximately 1 m, while the spatial resolution of the UAV data was 0.1 m. To eliminate these noises, the sigma filter method was conducted. Furthermore, a series of window sizes from 5 by 5 to 19 by 19 were tested for sigma filter. It was found that the sigma filter with the 11 by 11 window performed best with the highest correlation coefficient (r) value between spectra and SOM, STN. Finally, the noisy or troublesome spectral bands on the edge of spectra were removed. Consequently, 112 bands (403~759 nm and 783~900 nm) were retained, and then the raw spectral reflectance data were obtained.

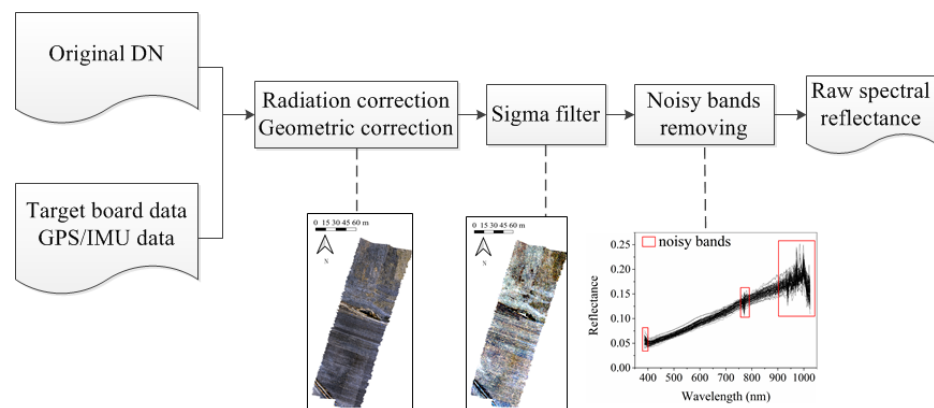


Figure 4. The flowchart of UAV data preprocessing.

2.3. Research Methods

2.3.1. Spectral Denoising Methods

Spectral preprocessing can effectively remove spectral noise. One is scatter correction methods, including Multiplicative Scattering Correction (MSC) and Standard Normal Variate transformation (SNV). The other is spectra derivatives methods, including the first derivate (FD) and the second derivate (SD) [39].

MSC is effective in eliminating the scattering effect and enhancing the spectral absorption information related to the soil properties in the spectral data. The method first establishes an “ideal spectrum” of the samples, i.e., a direct linear relationship between the spectral variation and the content of soil nutrients in the samples. The established spectrum is used as the standard spectrum for all other sample spectra corrections, including baseline

shift and offset corrections. Since the “ideal spectrum” is difficult to obtain, the average spectrum of all spectra is taken as the ideal one in practice. The specific algorithm works as follows: (i) calculate the average spectrum of all sample spectra; (ii) take the average spectrum as the ideal spectrum and perform linear regression to find the linear shift (regression constant b_i) and tilt offset (regression coefficient a_i) of each sample spectrum relative to the standard spectrum in Equation (1); (iii) subtract the linear shift from the original spectrum of each sample and divide it by the tilt offset, as in Equation (2).

$$R_{ik} = b_i + a_i \overline{R}_k \quad (1)$$

where R_{ik} denotes the reflectance of the i -th sample at the k -th band; $\overline{R}_k = \frac{\sum_{i=1}^n R_{ik}}{n}$, $i = 1, 2, \dots, n$, and n is the number of samples.

$$R_{ik_MSC} = \frac{R_{ik} - b_i}{a_i} \quad (2)$$

SNV is applied to reduce the effects of particle size heterogeneity and surface nonspecific scattering. It is assumed that the reflectance of each spectral wavelength point meets specific distribution characteristics, such as normal distribution. Based on the assumption, each sample spectrum can be corrected by subtracting the average value of those spectra from the original sample spectrum and dividing by the standard deviation, which is calculated as follows:

$$R_{ik_SNV} = \frac{R_{ik} - \overline{R}_i}{\sqrt{\frac{\sum_{k=1}^m (R_{ik} - \overline{R}_i)^2}{(m-1)}}} \quad (3)$$

where $\overline{R}_i = \frac{\sum_{k=1}^m R_{ik}}{m}$, $k = 1, 2, \dots, m$, and m is the number of bands.

FD and SD are the most common spectral derivatives methods. They can not only eliminate the baseline shifts and atmospheric scattering in the spectrum but also amplify the subtle changes in the slope of the spectral curve. These two methods are exploited to eliminate other background interference and improve discrimination and sensitivity. The spectral derivatives are expressed as follows:

$$R_{ik_FD} = \frac{R_{i(k+1)} - R_{ik}}{\lambda_{k+1} - \lambda_k} \quad (4)$$

$$R_{ik_SD} = \frac{R_{i(k+2)} - 2R_{i(k+1)} + R_{ik}}{(\lambda_{k+1} - \lambda_k)^2} \quad (5)$$

where λ_k denotes the wavelength of the k -th band.

2.3.2. CARS-SPA Feature Selection

Competitive Adaptive Reweighted Sampling (CARS) is a sampling approach for feature band selection based on Monte Carlo sampling and PLSR [40]. Through supervised feature selection, a subset of bands sensitive to predictor variables can be selected, but there may still be redundant variables with high correlation across the bands. The Successive Projections Algorithm (SPA) is another feature selection algorithm that minimizes the collinearity in the vector space [41]. Based on CARS, SPA is further exploited to eliminate the collinearity between bands, which helps to make the selected feature bands contain the most information and the least internal similarity.

The main steps of the CARS algorithm are as follows: (i) use Monte Carlo sampling approach to select the set of calibration samples and establish the PLSR model; (ii) execute the enforced selection according to the exponentially decreasing function (EDF) based on the regression coefficient generated in each loop; (iii) competitively refine the remaining variables by the adaptive reweighted sampling (ARS); and (iv) evaluate each subset of selected spectral variables with cross validation [42]. The subset of variables with the

lowest root mean square error of cross validation (RMSECV) is considered as the optimal set of feature bands [43].

The SPA algorithm is a forward variable selection algorithm based on vector projection analysis. The variables are selected from those with the largest projection value on the orthogonal subspace. First, we set N as the maximal number of variables to be selected. Then, the algorithm generates K sets of collections of N variables selected from the projected variable space. Finally, the set with the minimal RMSE of the multiple linear regression will be used to determine the optimal initial variable and the number of variables [44].

2.3.3. Predicting Model and Evaluation of the Accuracy

PLSR is a linear modeling technique that projects information from the original independent variable space into a few latent variables, a.k.a., “PLSR components”. This simplifies the interpretation of the correlation between independent and dependent variables because PLSR uses the smallest possible component number. The spectra matrix is used as the independent variable, and the actual values of SOM and STN are used as the dependent variable. PLSR aims to recognize a model with an optimal number of components which yield the lowest RSME and the highest R^2 values. A well fitted PLSR model indicates that the PLSR factors can provide explanation on the variation observed for predictors and responses [45]. In this way, the correlation between the actual values of SOM, STN, and the observed spectra matrix involves the optimal number of components [46]. This regression model was implemented in the R language with the version of 3.2.1 that is equipped with the package “pls” for PLSR modeling.

SVM is a kernel-based statistical machine learning method. As a popular data mining method, SVM has been used to model the VIS-NIR spectra data, and a nonlinear function is obtained by mapping a linear learning machine into a feature space induced by a high-dimensional kernel [47]. In this study, a radial basis function (RBF) kernel was selected, which can solve the non-linear problem using approximate multivariable function. The “e1071 package” with an R interface connected to the library for support vector machines (LIBSVM) was used. The optimization of the kernel-specific SVM parameters including C , e , and a , as well as the selection of the best preprocessing steps was performed by a systematic grid search using the leave-one-out cross-validation strategy applied on the training dataset.

Particle swarm optimization (PSO) is a computational method that simulates the foraging behavior of birds [48]. Exploiting PSO to optimize the selection of the input layer weights and hidden bias values of ELM can decrease the number of hidden layer nodes required by ELM and improve the generalization ability of a trained neural network. The process of ELM exploiting particle swarm optimization (PSO-ELM) is shown in Figure 5. There are four parameters that should be considered in the PSO method, including inertia weight, learning factors, maximum number of iterations, and population size. The parameter of inertia weight (w) has significance for performance of PSO. Shi and Eberhart [49] originally proposed adaptive inertia to assign optimal inertia weight. The strategy of decreasing weight of adaptive inertia was used in this study. The maximum value of w is chosen from the range of 0.8 to 1.2 for increasing the probability of locating the global optimum peak. The minimum value of w is chosen from the range of 0.4 to 0.2 to allow the particles to converge to the located optimum slowly [50]. Therefore, the inertia weights w_{max} and w_{min} were set to 0.8 and 0.4, respectively. The learning rates $c1$ and $c2$ are the acceleration constants with positive values [29], and $c1 > 0$, $c2 \leq 2$ are respectively the individual ($c1$) and global acceleration ($c2$) coefficients. Further, $c1$ and $c2$ are set with the same values, but sometimes different values of $c1$ and $c2$ can help improve the model performance [51]. The appropriate values of $c1$ and $c2$ were obtained by performing a test with a step of 0.1. Finally, the values of $c1$ and $c2$ were respectively set to 2.4 and 1.6 for PSO, which present the best predictive parameters for this model. The maximum number of iterations T was set to 100. The population size was tested from 25 to 200 with a step of 25, and the population size of 100 with the highest prediction accuracy was used

in the PSO. Then, the input weights and hidden biases corresponding to each particle were applied to the ELM model. Meanwhile, 55 training samples were used to train and 13 test samples were used for external validation. The mean square error (MSE) between the predicted and observed values of the test samples was used as the fitness of PSO to calculate the individual extreme value and global extreme value. The position and velocity of the particles were updated by particle fitness iteratively. The individual extremum and global extremum of the particles were updated until the minimum error was obtained or the maximum iteration number was reached. Finally, the input weights and hidden biases that yield the optimal results were used as the input parameters for the ELM model.

The coefficient of determination (R^2) and the mean absolute percentage error (MAPE) were employed to validate the performance of the model. Meanwhile, the residual prediction deviation (RPD, the ratio between the standard deviation and the root mean square error) was used to assess the stability and accuracy of the multivariable models [52]. Existing studies generally assume that the models with $RPD > 2$ are capable of providing accurate prediction of the discussed attributes; the models with RPD between 2 and 1.4 have moderate predictability, while the models with $RPD < 1.4$ have poor predictive power [53].

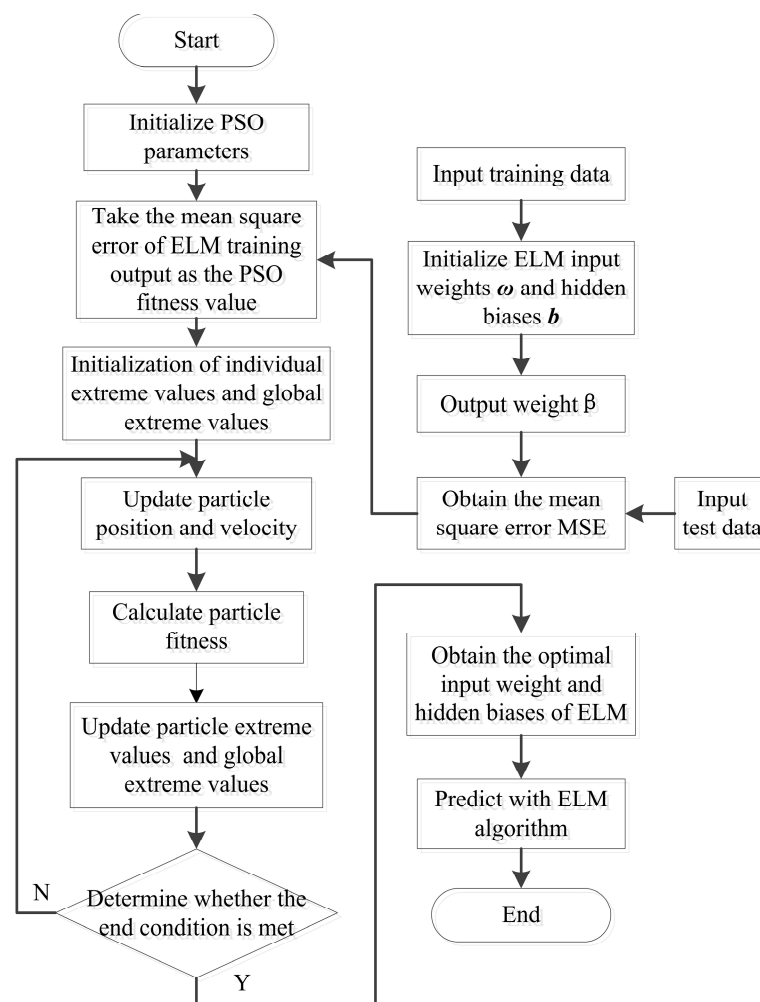


Figure 5. The flowchart of PSO-ELM.

3. Results

3.1. Analysis of Spectral Denoising Effect

The UAV images preprocessed by different methods are illustrated in Figure 6. By visual evaluation, the processed images by MSC and SNV can better present the typical features such as textures and smooth areas than those processed by FD and SD.

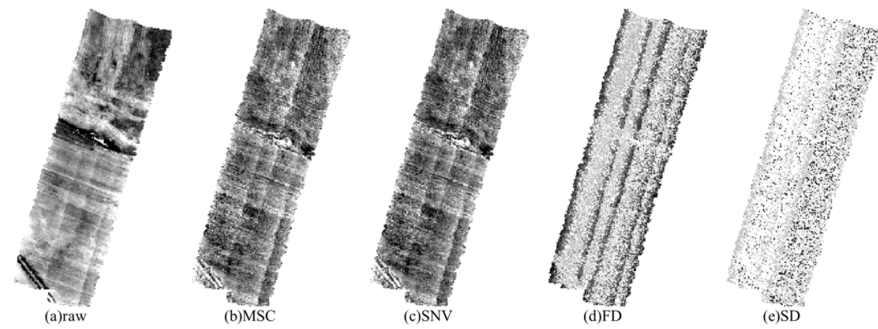


Figure 6. UAV hyperspectral images based on different spectral preprocessing.

As shown in Figure 7, the signal-to-noise ratio (SNR) values of the images processed by MSC and SNV were higher than that of the raw image. Specifically, the MSC technique improved the SNR from 31 to 160 dB at 514 nm, and the SNV technique improved the SNR from 33 to 99 dB at 715 nm. Whereas, the FD and SD techniques generated images with much lower SNR.

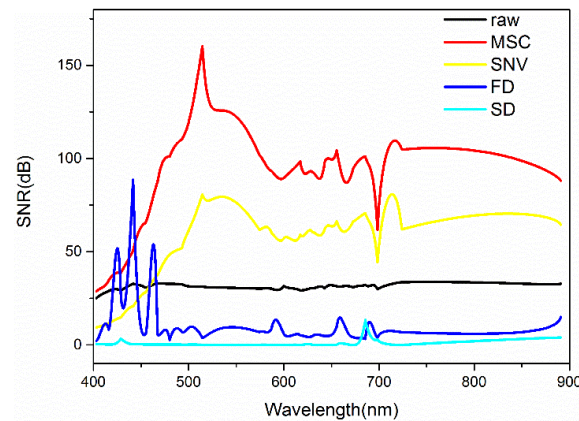


Figure 7. Signal-to-noise ratio of images based on different spectral preprocessing.

The correlation analysis of spectra with SOM and STN is illustrated in Figure 8. It can be seen that all four spectral preprocessing methods improved the correlation between spectra and soil properties. The MSC technique improved the correlation most, with the absolute value of the maximum correlation coefficient between spectral reflectance and soil properties being increased from 0.24 to 0.54 for SOM and from 0.28 to 0.56 for STN. Therefore, MSC can be regarded as the best spectral denoising method, and the subsequent feature selection and modeling were performed based on the spectral data processed by the MSC technique.

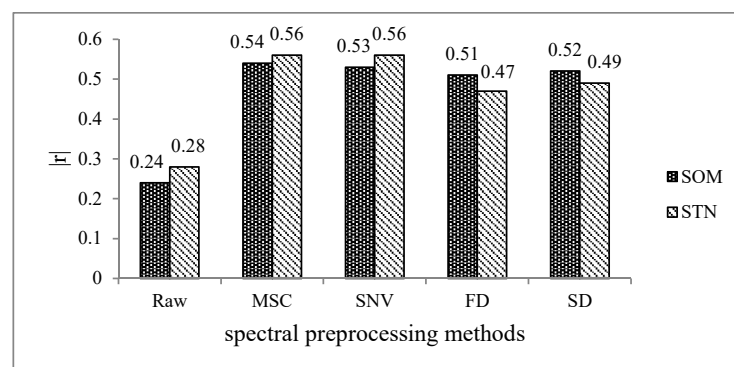


Figure 8. The absolute value of the maximum Pearson's correlation coefficient ($|r|$) between sample spectra and soil properties based on different spectral preprocessing methods.

3.2. Spectral Feature Bands Selection

The feature selection process of 50 CARS runs for the 112 bands of the UAV hyper-spectral image is shown in Figure 9, where the process of screening SOM and STN feature bands are exhibited in Figure 9(a1–a3) and Figure 9(b1–b3), respectively. It can be seen from Figure 9(a1,b1) that in the process of selecting spectral variables, the number of selected band variables gradually decreased with the increase of sampling runs, and the decreasing trend turned from fast to slow. The RMSECV curve in Figure 9(a2,b2) illustrates a trend of first decreasing to the lowest point and then increasing, indicating that the bands unrelated to SOM and STN were eliminated in the process first, and then the bands related to SOM and STN were eliminated. In Figure 9(a3,b3), the location of vertical star markers corresponded to the minimum RMSECV value in the entire variable screening process.

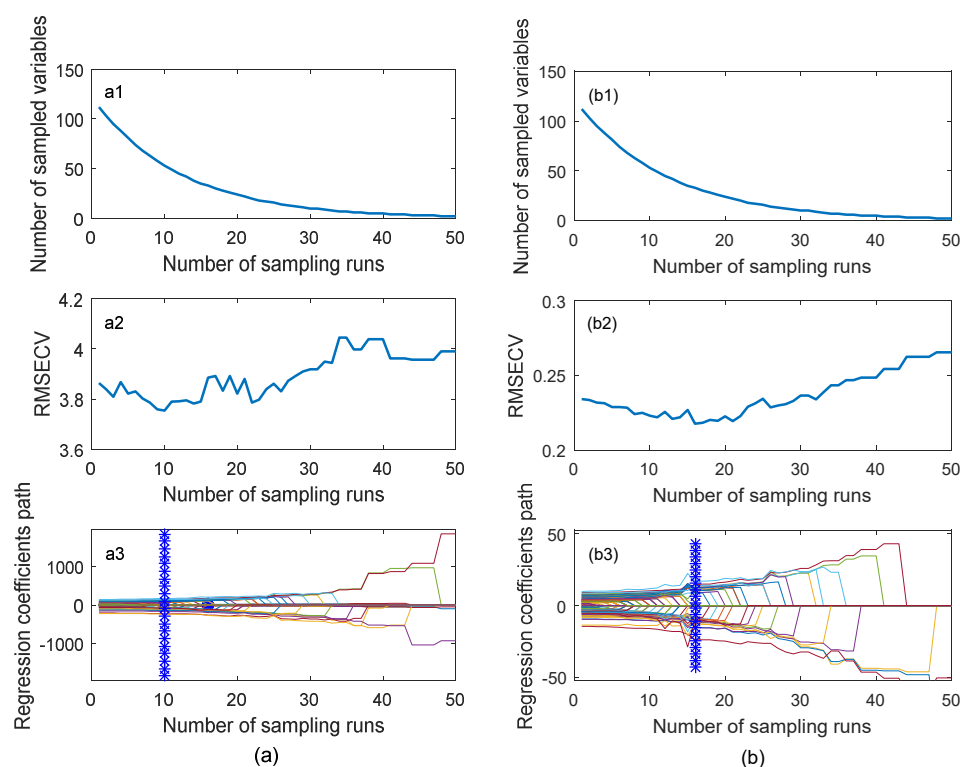


Figure 9. Screening spectral wavelength variables of SOM (a) and STN (b) based on CARS.

The number of SOM variables screened by CARS was 53, and the distribution of their band positions is shown in Figure 10a. It can be seen that the sensitive bands of SOM were located in the interval of 400~900 nm, but the distribution was relatively dense in 400~700 nm. The number of selected STN variables was 33, and the distribution of their band locations was shown in Figure 10b. It can be seen that the sensitive bands of STN were densely distributed in 400~440 nm and 480~540 nm. After the secondary screening of the feature bands through the SPA algorithm, the redundant bands at 400~540 nm were eliminated, and 24 SOM feature bands (Figure 10a) and 22 STN feature bands (Figure 10b) were finally obtained. It can be observed from the partial enlargement in the figures that in the range of 500~550 nm, 506, 532, and 540 nm were reserved by SPA for SOM prediction, and in the range of 480~520 nm, 506, 510, and 519 nm were reserved by SPA for STN prediction.

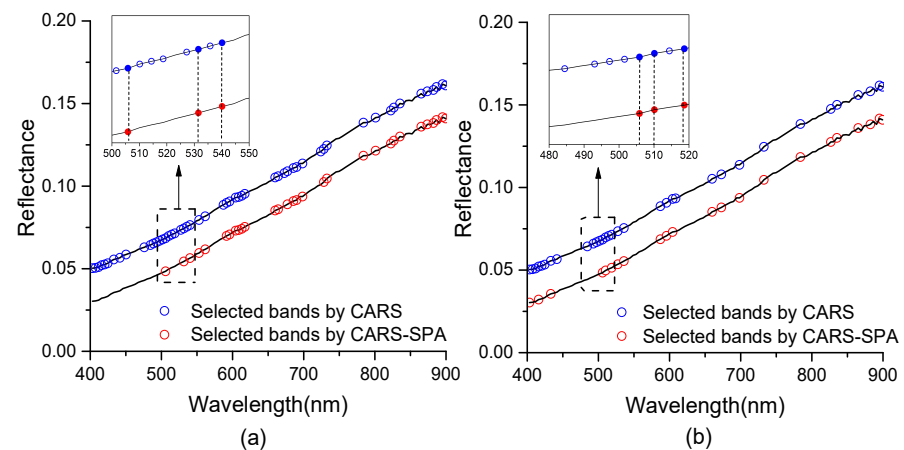


Figure 10. Feature bands of SOM (a) and STN (b) selected by CARS and CARS-SPA.

3.3. Model Accuracy

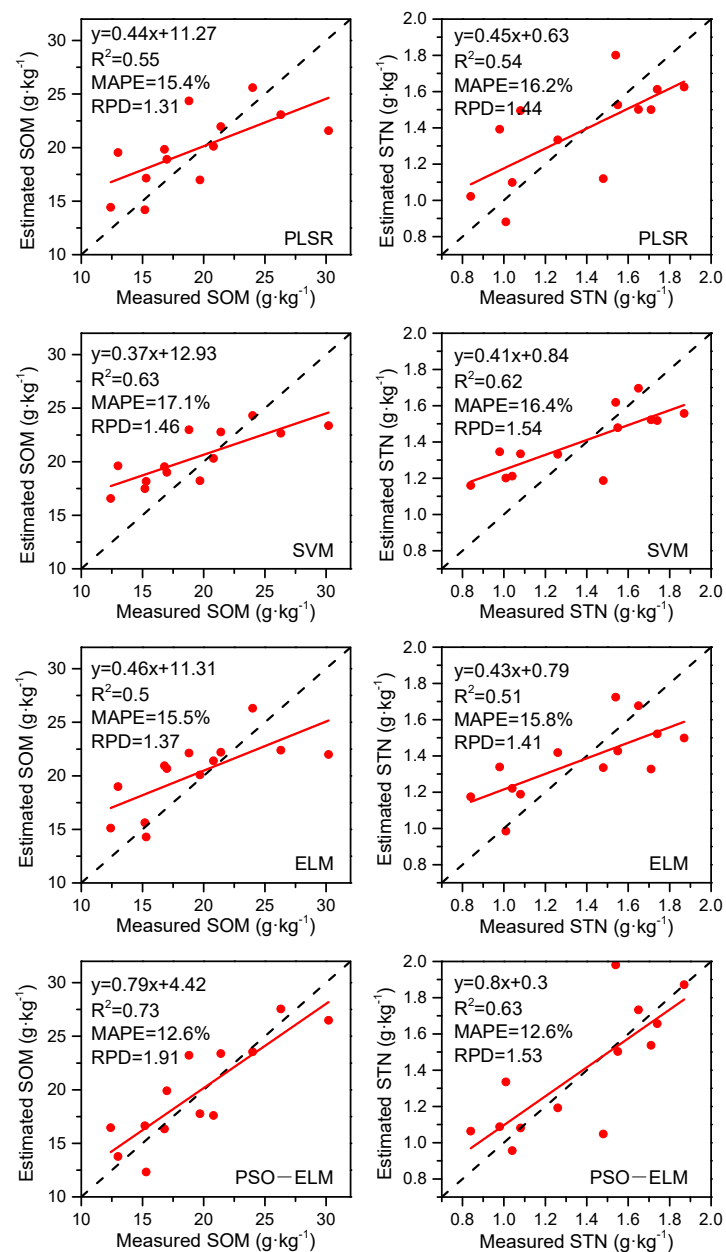
The PLSR, SVM, ELM, and PSO-ELM were conducted to predict SOM and STN based on the hyperspectral feature (Tables 1 and 2). PLSR models based on full spectrum generated prediction accuracy for SOM estimation of $R^2 = 0.55$, MAPE = 15.4%, and RPD = 1.31, for STN estimation of $R^2 = 0.54$, MAPE = 16.2%, and RPD = 1.44. SVM, ELM, and PSO-ELM models with full spectrum as inputs achieved a poor performance with MAPE higher than 19% both for SOM and STN estimation. As PLSR model can select the hyperspectral feature based on “PLSR components”, thus, the CARS and CARS-SPA methods were not conducted in the PLSR model for input feature selection. Although the R^2 values from SVM, ELM, and PSO-ELM were increased by using the CARS method to remove the irrelevant variables in the full spectrum, these models still have poor prediction ability with RPD value < 1.4 . Once the CARS-SPA method was exploited for secondary screening of feature bands, the accuracies of the SOM and STN prediction were further improved. The PSO-ELM model with selected bands by CARS-SPA produced the highest accuracy for SOM prediction with $R^2 = 0.73$, MAPE = 12.6%, and RPD = 1.91 (Figure 11), for STN prediction with $R^2 = 0.63$, MAPE = 12.6%, and RPD = 1.53 (Figure 11). Overall, the spectral variable selection method of CARS-SPA provided better estimates of SOM and STN than the full spectrum and CARS, which was indicated by the greater R^2 and RPD values as well as lower MAPE values.

Table 1. Model accuracies of test samples based on different feature selection methods for predicting SOM.

Modeling Method	Feature Selection Method	R^2	MAPE	RPD
PLSR	Full spectrum	0.55	15.4%	1.31
SVM	Full spectrum	0.18	20.2%	1.14
	CARS	0.35	18.8%	1.26
ELM	CARS-SPA	0.63	17.1%	1.46
	Full spectrum	0.24	24.3%	0.99
PSO-ELM	CARS	0.35	22.5%	1.00
	CARS-SPA	0.50	15.5%	1.37
	Full spectrum	0.26	24.9%	1.05
	CARS	0.55	18.6%	1.42
	CARS-SPA	0.73	12.6%	1.91

Table 2. Model accuracies of test samples based on different feature selection methods for predicting STN.

Modeling Method	Feature Selection Method	R ²	MAPE	RPD
PLSR	Full spectrum	0.54	16.2%	1.44
SVM	Full spectrum	0.34	19.1%	1.26
	CARS	0.57	16.9%	1.46
ELM	CARS-SPA	0.62	16.4%	1.54
	Full spectrum	0.26	22.0%	1.02
	CARS	0.43	16.5%	1.28
PSO-ELM	CARS-SPA	0.51	15.8%	1.41
	Full spectrum	0.27	19.2%	1.05
	CARS	0.58	15.2%	1.41
	CARS-SPA	0.63	12.6%	1.53

**Figure 11.** Scatter plots of the measured and estimated soil properties of test samples based on different models.

3.4. Soil Nutrient Spatial Distribution

The PSO-ELM model established based on the set of feature bands selected by the CARS-SPA algorithm was exploited to invert the SOM and STN pixel by pixel in the study area. The SOM content was between 12 and 30 g kg⁻¹, and the STN content was between 0.8 and 2 g kg⁻¹. The spatial distribution of SOM and STN was shown in Figure 12. It can be seen that the high values of SOM and STN were mainly located in the northeastern corn growing area. According to the nutrient grading standard of agricultural soils, the area percentage of each nutrient grade in corn and peanut areas was counted separately, and the result is shown in the lower right corner of the figure. The analysis indicated that the amount of SOM in the corn area was moderate while that in the peanut area was slightly deficient; the amount of STN in both the corn and peanut areas was moderate.

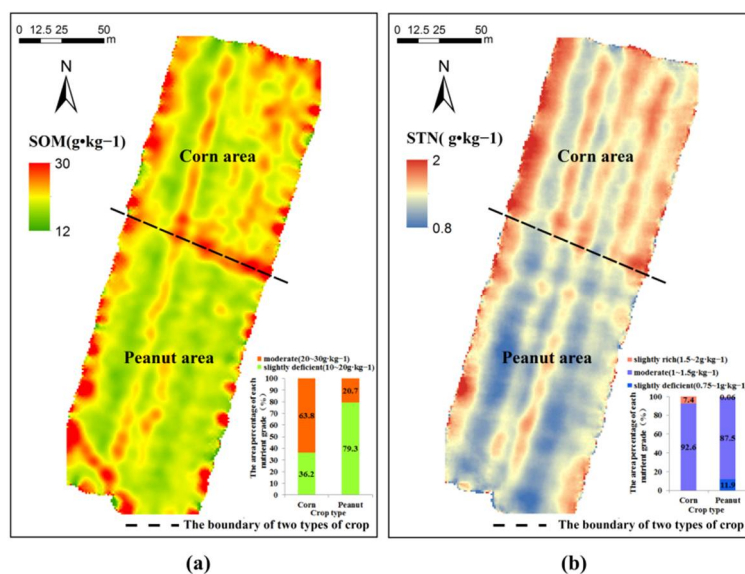


Figure 12. Spatial distribution of SOM (a) and STN (b) using UAV images based on optimal regression models.

4. Discussion

Preprocessing of hyperspectral imagery can facilitate subsequent spectral interpretation and data modeling since unwanted noise can be eliminated while chemical signals of interest can be enhanced [54]. Meanwhile, the development of spectroscopic equipment has brought new challenges for preprocessing spectral data. Preprocessing is exploited to standardize spectra and remove instrumental and physical noise. In this study, the preprocessing based on the MSC, SNV, FD, and SD techniques was considered before spectral feature selection and modeling. It is essential to reduce additive and multiplicative effects caused by light scattering to achieve an accurate quantitative analysis of soil properties on the hyperspectral data obtained from UAV platforms. The SNR of raw imagery without any preprocessing was approximately 30 dB, and a weak correlation was found between the spectral characteristic and soil properties. To achieve optimal prediction performances, it was recommended that powerful preprocessing techniques such as MSC and SNV should be given priority so as to increase the SNR value and the correlation coefficient with soil properties. The MSC technique is based on the Beer–Lambert model, which can reduce the wavenumber independent radiation loss caused by scattering or stray light [55]. The SNV technique has the advantage of handling a constant offset. To remove the baseline offset or handle the baseline polynomial fitting for UAV imagery, MSC is recommended as a better practice than SNV. It was also found that MSC performed better in soil clay estimation on CHRIS images [53]. FD and SD are two common and popular techniques to resolve overlapping analyte signals through conversion of spectra [56]. Since the peak position of spectra is crucial for indicating chemical properties, the first or second derivatives

transformation can reserve or highlight band positions [55]. However, there is no obvious peak position or absorption feature for soil spectra in 400~1000 nm from UAV hyperspectral imagery. This may make the FD or SD method not perform well to reduce or eliminate the scatter effect in the preprocessing of imagery.

To reduce the large dimensionality of the hyperspectral image data and to decrease the computation time, optimal wavelength selection may be more significant than using a full spectrum for the quantitative model [57]. The SVM, ELM, and PSO-ELM models established based on full spectrum obtained the RPD less than 1.4 for SOM and STN, indicating no prediction ability. Additionally, the overlarge number of variables and a small number of samples may cause an overfitting problem in the prediction model [58]. In this study, the input bands that produce high prediction accuracy with minimum collinearity are considered as the optimal wavelengths. CARS based on the backward stepwise method can fast select variables to obtain high prediction accuracy [59]. Exploiting this approach, the effective primary wavelengths were selected from 112 bands of hyperspectral imagery, i.e., 53 bands for SOM and 33 bands for STN, which reduced the MAPE value from 24.9% to 18.6% for SOM and from 19.2% to 15.2% for STN using PSO-ELM. Furthermore, to reduce the collinearity effects from primary wavelengths, the SPA algorithm was exploited for a secondary wavelength selection, which obtained 24 bands for SOM and 22 bands for STN with small collinearity. The experimental results revealed that in contrast to selecting wavelengths only by primary spectral, the coupling of CARS with SPA could reduce the number of bands and improve the prediction accuracy.

Hyperspectral imagery can be considered as multidimensional big data, such as a matrix of 2820×1500 pixels multiplied with 112 bands in our study area covering 0.043 km^2 . A sound and effective learning algorithm is required for soil prediction and mapping on hyperspectral imagery at various scales. The ELM method with random weights of the hidden neurons and inherent steps has benefits in high training speed and easily ensemble [60]. However, ELM usually requires more hidden neurons than traditional algorithms to make good predictions, which might result in slow responses of ELM to new data [26]. To handle this problem, PSO was employed to optimize the input weights and hidden biases, which can improve the generalization performance of ELM with fewer hidden neurons. In this paper, a comparison of the PSO-ELM and ELM regression models was conducted. The greater RPD and R^2 values as well as lower MAPE values indicated that the PSO-ELM methods provided better estimation than the ELM methods for both SOM and STN. It was also found that in the forecasting of CO_2 emissions, the established PSO-ELM model outperformed the ELM [61]. The results of this paper also showed that PSO-ELM performed better than PLSR and SVM with the greater R^2 and lower MAPE values.

5. Conclusions

The experimental results in this study revealed that UAV spectroscopy imaging can be exploited to quantify the SOM and STN in farmland. This study explored the wavelengths selection of UAV hyperspectral imagery and an improved extreme learning machine method for SOM and STN prediction and mapping. The CARS coupled with SPA was a potential strategy for identifying optimal and effective bands for hyperspectral imagery. The following conclusions were made:

1. This work proved that the preprocessing to reduce unwanted noise for UAV imagery is essential to establish optimal models for predicting SOM and STN with VIS-NIR spectroscopy. The MSC technique was highly recommended for preprocessing, which contributed to the image radiance with high SNR value and reflectance spectra with high correlation coefficient.
2. The CARS-SPA approach could select a small number of reasonable wavelengths, which selected 33 and 22 bands that were informative for SOM and STN prediction. Based on these bands, the prediction performance was exhibited by R^2 of 0.73, MAPE

of 12.6%, and RPD of 1.91 for SOM and R^2 of 0.63, MAPE of 12.6%, and RPD of 1.53 for STN with the PSO-ELM model.

3. The PSO was exploited to assign the weights and biases of the ELM model while avoiding randomness. The proposed method of PSO-ELM outperformed the ELM in that it reduced the MAPE by 2.9% and 3.2% and increased R^2 by 0.23 and 0.12 for SOM and STN, respectively. It also performed better than PLSR and SVM with the greater R^2 and lower MAPE values.

Overall, this study provided an alternative approach for soil properties estimation and mapping based on the imaging spectrometry obtained from a UAV platform. In addition, utilizing imagery and field samples on a larger scale will be our next-step research toward developing regional models.

Author Contributions: Conceptualization, literature review, and analysis, N.B.; methodology, data curation, validation, and writing—original draft preparation, X.Y.; writing—review and editing, W.L.; proof-reading, guidance, and regular feedback, S.L.; project administration, Y.F. and Y.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 52074063, U1903216, and 52074064; Fundamental Research Funds for the Central Universities, grant number N2001002.

Acknowledgments: We acknowledge LICA United Technology Limited for acquiring UAV hyperspectral images. We acknowledge Beijing New Olympic Environment Labelling Physical and Chemical Analysis Center for measuring the content of soil nutrient.

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

References

1. Six, J.; Paustian, K. Aggregate-associated soil organic matter as an ecosystem property and a measurement tool. *Soil Biol. Biochem.* **2014**, *68*, A4–A9. [[CrossRef](#)]
2. Jin, X.; Du, J.; Liu, H.; Wang, Z.; Song, K. Remote estimation of soil organic matter content in the Sanjiang Plain, Northeast China: The optimal band algorithm versus the GRA-ANN model. *Agric. Meteorol.* **2016**, *218*, 250–260. [[CrossRef](#)]
3. Pouladi, N.; Møller, A.B.; Tabatabai, S.; Greve, M.H. Mapping soil organic matter contents at field level with Cubist, Random Forest and kriging. *Geoderma* **2019**, *342*, 85–92. [[CrossRef](#)]
4. Jin, X.; Song, K.; Du, J.; Liu, H.; Wen, Z. Comparison of different satellite bands and vegetation indices for estimation of soil organic matter based on simulated spectral configuration. *Agric. Meteorol.* **2017**, *244–245*, 57–71. [[CrossRef](#)]
5. Tian, Y.; Zhang, J.; Yao, X.; Cao, W.; Zhu, Y. Laboratory assessment of three quantitative methods for estimating the organic matter content of soils in China based on visible/near-infrared reflectance spectra. *Geoderma* **2013**, *202*, 161–170. [[CrossRef](#)]
6. Bao, N.; Wu, L.; Ye, B.; Ke, Y.; Wei, Z. Assessing soil organic matter of reclaimed soil from a large surface coal mine using a field spectroradiometer in laboratory. *Geoderma* **2017**, *288*, 47–55. [[CrossRef](#)]
7. Vohland, M.; Ludwig, M.; Thiele-Bruhn, S.; Ludwig, B. Determination of soil properties with visible to near- and mid-infrared spectroscopy: Effects of spectral variable selection. *Geoderma* **2014**, *223–225*, 88–96. [[CrossRef](#)]
8. Hong, Y.; Guo, L.; Chen, S.; Linderman, M.; Mouazen, A.M.; Yu, L.; Chen, Y.; Liu, Y.; Liu, Y.; Cheng, H.; et al. Exploring the potential of airborne hyperspectral image for estimating topsoil organic carbon: Effects of fractional-order derivative and optimal band combination algorithm. *Geoderma* **2020**, *365*, 114228. [[CrossRef](#)]
9. Vaudour, E.; Gilliot, J.M.; Bel, L.; Lefevre, J.; Chehdi, K. Regional prediction of soil organic carbon content over temperate croplands using visible near-infrared airborne hyperspectral imagery and synchronous field spectra. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *49*, 24–38. [[CrossRef](#)]
10. Gomez, C.; Viscarra Rossel, R.A.; McBratney, A.B. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma* **2008**, *146*, 403–411. [[CrossRef](#)]
11. Meng, X.; Bao, Y.; Liu, J.; Liu, H.; Zhang, X.; Zhang, Y.; Wang, P.; Tang, H.; Kong, F. Regional soil organic carbon prediction model based on a discrete wavelet analysis of hyperspectral satellite data. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *89*, 102111. [[CrossRef](#)]
12. Birk, R.; McCord, T. Airborne Hyperspectral Sensor Systems. *IEEE Aerosp. Electron. Syst. Mag.* **1994**, *9*, 26–33. [[CrossRef](#)]
13. Hbirkou, C.; Pätzold, S.; Mahlein, A.-K.; Welp, G. Airborne hyperspectral imaging of spatial soil organic carbon heterogeneity at the field-scale. *Geoderma* **2012**, *175–176*, 21–28. [[CrossRef](#)]
14. Hruska, R.; Mitchell, J.; Anderson, M.; Glenn, N.F. Radiometric and Geometric Analysis of Hyperspectral Imagery Acquired from an Unmanned Aerial Vehicle. *Remote Sens.* **2012**, *4*, 2736–2752. [[CrossRef](#)]
15. Bioucas-Dias, J.M.; Plaza, A. Hyperspectral remote sensing data analysis and future challenges. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 3470. [[CrossRef](#)]

16. Tang, Q.; Zhang, R.; Chen, L.; Xu, G.; Deng, W.; Ding, C.; Xu, M.; Yi, T.; Wen, Y.; Li, L. High-accuracy, high-resolution downwash flow field measurements of an unmanned helicopter for precision agriculture. *Comput. Electron. Agric.* **2020**, *173*, 105390. [[CrossRef](#)]
17. Lu, B.; Dao, P.; Liu, J.; He, Y.; Shang, J. Recent Advances of Hyperspectral Imaging Technology and Applications in Agriculture. *Remote Sens.* **2020**, *12*, 2659. [[CrossRef](#)]
18. Sankey, J.B.; Sankey, T.T.; Li, J.; Ravi, S.; Wang, G.; Caster, J.; Kasprak, A. Quantifying plant-soil-nutrient dynamics in rangelands: Fusion of UAV hyperspectral-LiDAR, UAV multispectral-photogrammetry, and ground-based LiDAR-digital photography in a shrub-encroached desert grassland. *Remote Sens. Environ.* **2021**, *253*, 112223. [[CrossRef](#)]
19. Hu, J.; Peng, J.; Zhou, Y.; Xu, D.; Ruiying, Z.; Jiang, Q.; Fu, T.; Wang, F.; Shi, Z. Quantitative Estimation of Soil Salinity Using UAV-Borne Hyperspectral and Satellite Multispectral Images. *Remote Sens.* **2019**, *11*, 736. [[CrossRef](#)]
20. Zhang, N.; Zhang, X.; Yang, G.; Zhu, C.; Huo, L.; Feng, H. Assessment of defoliation during the *Dendrolimus tabulaeformis* Tsai et Liu disaster outbreak using UAV-based hyperspectral images. *Remote Sens. Environ.* **2018**, *217*, 323–339. [[CrossRef](#)]
21. Abdulridha, J.; Batuman, O.; Ampatzidis, Y. UAV-Based Remote Sensing Technique to Detect Citrus Canker Disease Utilizing Hyperspectral Imaging and Machine Learning. *Remote Sens.* **2019**, *11*, 1373. [[CrossRef](#)]
22. Summers, D.; Lewis, M.; Ostendorf, B.; Chittleborough, D. Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties. *Ecol. Indic.* **2011**, *11*, 123–131. [[CrossRef](#)]
23. Kuang, B.; Tekin, Y.; Mouazen, A.M. Comparison between artificial neural network and partial least squares for on-line visible and near infrared spectroscopy measurement of soil organic carbon, pH and clay content. *Soil Tillage Res.* **2015**, *146*, 243–252. [[CrossRef](#)]
24. Honkavaara, E.; Kaivosoja, J.; Mäkynen, J.; Pellikka, I.; Pesonen, L.; Saari, H.; Salo, H.; Hakala, T.; Markelin, L.; Rosnell, T. Hyperspectral reflectance signatures and point clouds for precision agriculture by light weight UAV imaging system. *Isprs Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *7*, 353–358. [[CrossRef](#)]
25. Li, H.; Jia, S.; Le, Z. Quantitative Analysis of Soil Total Nitrogen Using Hyperspectral Imaging Technology with Extreme Learning Machine. *Sensors* **2019**, *19*, 4355. [[CrossRef](#)] [[PubMed](#)]
26. Figueiredo, E.M.N.; Ludermir, T.B. Investigating the use of alternative topologies on performance of the PSO-ELM. *Neurocomputing* **2014**, *127*, 4–12. [[CrossRef](#)]
27. Ge, X.; Wang, J.; Ding, J.; Cao, X.; Zhang, Z.; Liu, J.; Li, X. Combining UAV-based hyperspectral imagery and machine learning algorithms for soil moisture content monitoring. *Peer J.* **2019**, *7*, e6926. [[CrossRef](#)] [[PubMed](#)]
28. Huang, G.B.; Zhou, H.; Ding, X.; Rui, Z. Extreme Learning Machine for Regression and Multiclass Classification. *Ieee Trans. Syst. Man Cybern. Part. B* **2012**, *42*, 513–529. [[CrossRef](#)]
29. Han, F.; Yao, H.F.; Ling, Q.H. An improved evolutionary extreme learning machine based on particle swarm optimization. *Neurocomputing* **2013**, *116*, 87–93. [[CrossRef](#)]
30. You, X.; Yang, S. In Evolutionary Extreme Learning Machine–Based on Particle Swarm Optimization, Advances in Neural Networks–ISNN 2006. In Proceedings of the Third International Symposium on Neural Networks, Chengdu, China, 28 May–1 June 2006; Springer: Berlin/Heidelberg, Germany, 2006.
31. Langdon, W.B.; Poli, R. Evolving Problems to Learn About Particle Swarm Optimizers and Other Search Algorithms. *Ieee Trans. Evol. Comput.* **2007**, *11*, 561–578. [[CrossRef](#)]
32. Shi, X.Z.; Yu, D.S.; Xu, S.X.; Warner, E.D.; Wang, H.J.; Sun, W.X.; Zhao, Y.C.; Gong, Z.T. Cross-reference for relating Genetic Soil Classification of China with WRB at different scales. *Geoderma* **2010**, *155*, 344–350. [[CrossRef](#)]
33. Peng, M.; Zhao, C.; Ma, H.; Yang, Z.; Yang, K.; Liu, F.; Li, K.; Yang, Z.; Tang, S.; Guo, F.; et al. Heavy metal and Pb isotopic compositions of soil and maize from a major agricultural area in Northeast China: Contamination assessment and source apportionment. *J. Geochem. Explor.* **2020**, *208*, 106403. [[CrossRef](#)]
34. Wang, S.; Adhikari, K.; Zhuang, Q.; Gu, H.; Jin, X. Impacts of urbanization on soil organic carbon stocks in the northeast coastal agricultural areas of China. *Sci. Total Environ.* **2020**, *721*, 137814. [[CrossRef](#)] [[PubMed](#)]
35. Wang, X.; Zhang, F.; Kung, H.-T.; Johnson, V.C. New methods for improving the remote sensing estimation of soil organic matter content (SOMC) in the Ebinur Lake Wetland National Nature Reserve (ELWNNR) in northwest China. *Remote Sens. Environ.* **2018**, *218*, 104–118. [[CrossRef](#)]
36. Dobbin, K.K.; Simon, R.M. Optimally splitting cases for training and testing high dimensional classifiers. *BMC Med. Genom.* **2011**, *4*, 1–8. [[CrossRef](#)]
37. Kanning, M.; Kuhling, I.; Trautz, D.; Jarmer, T. High-Resolution UAV-Based Hyperspectral Imagery for LAI and Chlorophyll Estimations from Wheat for Yield Prediction. *Remote Sens.* **2018**, *10*, 17. [[CrossRef](#)]
38. Su, G.H.; Shu, J.S.; Cui, P.; Lei, G. Research on Location Accuracy of UAV Based on DGPS/INS Technique. *Adv. Mater. Res.* **2011**, *204*, 1525–1528. [[CrossRef](#)]
39. Rinnan, Å.; van den Berg, F.; Engelsen, S.B. Review of the most common pre-processing techniques for near-infrared spectra. *Trac Trends Anal. Chem.* **2009**, *28*, 1201–1222. [[CrossRef](#)]
40. Feng, X.; Zhao, Y.; Zhang, C.; Cheng, P.; He, Y. Discrimination of Transgenic Maize Kernel Using NIR Hyperspectral Imaging and Multivariate Data Analysis. *Sensors* **2017**, *17*, 1894. [[CrossRef](#)]
41. Fan, L.; Zhao, J.; Xu, X.; Liang, D.; Yang, G.; Feng, H.; Yang, H.; Wang, Y.; Chen, G.; Wei, P. Hyperspectral-Based Estimation of Leaf Nitrogen Content in Corn Using Optimal Selection of Multiple Spectral Variables. *Sensors* **2019**, *19*, 2898. [[CrossRef](#)]

42. Xu, S.; Wang, M.; Shi, X. Hyperspectral imaging for high-resolution mapping of soil carbon fractions in intact paddy soil profiles with multivariate techniques and variable selection. *Geoderma* **2020**, *370*, 114358. [[CrossRef](#)]
43. Song, X.; Du, G.; Li, Q.; Tang, G.; Huang, Y. Rapid spectral analysis of agro-products using an optimal strategy: Dynamic backward interval PLS-competitive adaptive reweighted sampling. *Anal. Bioanal. Chem.* **2020**, *412*, 2795–2804. [[CrossRef](#)]
44. Ye, S.; Wang, D.; Min, S. Successive projections algorithm combined with uninformative variable elimination for spectral variable selection. *Chemom. Intell. Lab. Syst.* **2008**, *91*, 194–199. [[CrossRef](#)]
45. 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](#)]
46. Gotelli, N.J.; Associates, S. *A Primer of Ecological Statistics*; Sinauer Associates: Sunderland, MA, USA, 2013.
47. Chih-Chung, C.; Chih-Jen, L. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 1–39. [[CrossRef](#)]
48. Xinxin, L.; Zuojun, L.; Xinzhi, G.; Jie, Z. Bicycling Phase Recognition for Lower Limb Amputees Using Support Vector Machine Optimized by Particle Swarm Optimization. *Sensors* **2020**, *20*, 6533.
49. Shi, Y. In A Modified Particle Swarm Optimizer. In Proceedings of the IEEE Icc Conference, Anchorage, AK, USA, 4–9 May 1998; pp. 69–73.
50. Nickabadi, A.; Ebadzadeh, M.M.; Safabakhsh, R. A novel particle swarm optimization algorithm with adaptive inertia weight. *Appl. Soft Comput.* **2011**, *11*, 3658–3670. [[CrossRef](#)]
51. Pacifico, L.; Ludermir, T.B. Evolutionary extreme learning machine based on particle swarm optimization and clustering strategies. In Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN), Dallas, TX, USA, 4–9 August 2013; pp. 1–6.
52. Chang, C.; Chiang, S.; Smith, J.A.; Ginsberg, I.W. Linear spectral random mixture analysis for hyperspectral imagery. *Ieee Trans. Geosci. Remote Sens.* **2002**, *40*, 375–392. [[CrossRef](#)]
53. Casa, R.; Castaldi, F.; Pascucci, S.; Palombo, A.; Pignatti, S. A comparison of sensor resolution and calibration strategies for soil texture estimation from hyperspectral remote sensing. *Geoderma* **2013**, *197*, 17–26. [[CrossRef](#)]
54. Mishra, P.; Marini, F.; Biancolillo, A.; Roger, J.M. Improved prediction of fuel properties with near-infrared spectroscopy using a complementary sequential fusion of scatter correction techniques. *Talanta* **2020**, *223*, 121693. [[CrossRef](#)] [[PubMed](#)]
55. Kohler, A.; Böcker, U.; Martens, H. Model-Based Pre-Processing in Biospectroscopy. Available online: <https://2009.isiproceedings.org/A5%20Docs/0244.pdf> (accessed on 19 April 2021).
56. Li, L.; Peng, Y.; Li, Y.; Wang, F. A new scattering correction method of different spectroscopic analysis for assessing complex mixtures. *Anal. Chim. Acta* **2019**, *1087*, 20–28. [[CrossRef](#)]
57. Hong, Y.; Chen, Y.; Yu, L.; Liu, Y.; Liu, Y.; Zhang, Y.; Liu, Y.; Cheng, H. Combining Fractional Order Derivative and Spectral Variable Selection for Organic Matter Estimation of Homogeneous Soil Samples by VIS-NIR Spectroscopy. *Remote Sens.* **2018**, *10*, 479. [[CrossRef](#)]
58. Li, H.D.; Liang, Y.Z.; Xu, Q.S.; Cao, D.S. Model population analysis for variable selection. *J. Chemom.* **2010**, *24*, 418–423. [[CrossRef](#)]
59. Zheng, K.; Feng, T.; Zhang, W.; Huang, X.; Zou, X. Variable selection by double competitive adaptive reweighted sampling for calibration transfer of near infrared spectra. *Chemom. Intell. Lab. Syst.* **2019**, *191*, 109–117. [[CrossRef](#)]
60. Tan, C.; Chen, H.; Lin, Z. Brand classification of detergent powder using near-infrared spectroscopy and extreme learning machines. *Microchem. J.* **2021**, *160*, 105691. [[CrossRef](#)]
61. Sun, W.; Wang, C.; Zhang, C. Factor analysis and forecasting of CO₂ emissions in Hebei, using extreme learning machine based on particle swarm optimization. *J. Clean. Prod.* **2017**, *162*, 1095–1101. [[CrossRef](#)]