

Identification and Evolution of Soil Organic Carbon Density Caused by Coastal Rapid Siltation Based on Imaging Spectroscopy

Aosheng Wang¹, Guanghui Zheng¹, Danping Wei, Caixia Jiao, Xianli Xie, Rong Zeng, Chengyi Zhao, Giri Kattel², Ying Xiong, and Dian Zhou²

Abstract—This article aimed to investigate the feasibility of using imaging spectroscopy (IS) to predict soil organic carbon density (SOCD) either directly or indirectly through soil organic carbon (SOC). Three methods, partial least square regression, support vector machine (SVM), and random forest, were utilized to calibrate the models and map the SOCD. The results showed that direct prediction was better than indirect prediction and the best model SVM had high R^2 of 0.94 and 0.93 for calibration and validation, respectively. The measured SOCD was mainly concentrated in the surface soil layer from 0 to 40 cm, and the deeper layer tended to be gentle. The continuous depth variation trend of SOCD in the topsoil (0–40 cm) was relatively close to the measured values, while in the deeper layers (below 40 cm), it was much higher than the measured values. This article also found that the best fitting function of measured SOC stocks over time varied from linear to power and then to logarithmic with increasing depth, indicating less efficient accumulation of SOC in deep soil compared to topsoil, resulting in an overall decrease in SOC accumulation rate across soil depth. The best temporal functions for the predicted values differed from the measured values at each depth, but the changing trends of the three functions were basically consistent. It suggests that IS technology has the potential to quantitatively reveal the process of coastal soil evolution and offers a new insight for the rapid monitoring of soil property changes.

Index Terms—Coastline change, imaging spectroscopy, organic carbon density, soil chronosequence, soil evolution, soil profile.

Manuscript received 29 March 2023; revised 6 April 2023; accepted 17 April 2023. Date of publication 21 April 2023; date of current version 11 May 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 41877004 and Grant 42130405, in part by the Innovative and Entrepreneurial Talent Program of Jiangsu Province under Grant R2020SC04, and in part by the Longshan Professorship and the talent under Grant #1511582101011 at Nanjing University of Information Science and Technology. (Corresponding author: Guanghui Zheng.)

Aosheng Wang, Guanghui Zheng, Danping Wei, Caixia Jiao, Rong Zeng, Chengyi Zhao, and Giri Kattel are with the School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: 20201210021@nuist.edu.cn; zgh@nuist.edu.cn; wdpnuist@163.com; jcx@nuist.edu.cn; rzeng@nuist.edu.cn; zhaocy@nuist.edu.cn; gkattel@unimelb.edu.au).

Xianli Xie is with the Institute of Soil Science, Chinese Academy of Science, Nanjing 210008, China (e-mail: xlxie@issas.ac.cn).

Ying Xiong is with the Reading Academy, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: fw804619@student.reading.ac.uk).

Dian Zhou is with the Changwang School of Honors, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: 202083330052@nuist.edu.cn).

This article has supplementary downloadable material available at <https://doi.org/10.1109/JSTARS.2023.3269078>, provided by the authors.

Digital Object Identifier 10.1109/JSTARS.2023.3269078

I. INTRODUCTION

SOIL organic carbon (SOC) is a crucial indicator of soil fertility and has a direct impact on crop and vegetation yields in terrestrial landscapes worldwide [1]. Increasing SOC stocks can offset some of the effects of fossil fuel carbon emissions and contribute to mitigating the greenhouse gas effect [2]. Therefore, SOC is often regarded as one of the most essential components to address the issues related to world food security and global climate change [3], [4], [5]. Understanding the dynamic changes in the regional SOC is conducive to regulating farming practices and land management strategies in many countries including China [6].

Until recently, studies on SOC have mainly focused on the surface layer (0–30 cm), with limited understanding of the deep soil profiles [7]. However, more than 50% of the organic carbon is stored at the depth between 0.3 m and 1 m [5], hence, soil sampling from an appropriate depth is essential for accurately calculating the rate, and potential of the soil profile in SOC sequestration [8], [9]. In addition, despite their high accuracy, the determination of SOC density (SOCD) relies mainly on conventional methods, which require considerable manpower and resources, and have the disadvantages such as longer time sampling with, higher sampling cost, as well as pollution [10]. The development of remote sensing technology has addressed these limitations, enabling real-time, rapid quantitative monitoring of soil properties. This capability holds greater importance for quickly obtaining soil information under the ongoing global climate change scenario [11], [12], [13].

Reflectance spectroscopy has been widely used to estimate SOC and other properties worldwide [14], [15], [16], [17], [18], [19], [20]. However, interestingly the reflectance spectroscopy has been rarely used to estimate the SOCD directly; and it cannot map the SOCD for a certain soil profile. Hence, imaging spectroscopy, as a new technique with high spectral and spatial resolution, has become an effective method for mapping soil properties mostly for profiles with fine spatial resolutions [21], [22], [23]. As a promising soil analysis tool [24], imaging spectroscopy can also effectively compensate for the disadvantages arose the fixed-depth sampling and obtain continuous depth functions for soil properties [25]. Lately, imaging spectroscopy has been successfully applied to a soil property analysis [21], [26] and a profile mapping system [23], [27], [28]. For instance,

Sorenson et al. [29] and Xu et al. [30] mapped SOC and carbon components for soil profile, and extracted their vertical distribution patterns. Their mapping of SOC suggests that imaging spectroscopy helps capture both vertical and lateral variations of soil properties more quickly and efficiently. Today, the advancement of mapping technology has become an essential component of the digital soil morphometrics, and widely believed that it would help estimate, and map SOCD profiles including the SOC stock time- and cost-more efficiently [31].

Apart from mapping, carbon contents and properties in soil spectra are modelled using various calibration methods, which include linear models, such as multiple linear regression, principal components regression, and partial least squares regression (PLSR) [30], [32], [33], [34], [35]. In addition, nonlinear models, such as support vector machine (SVM), random forests (RF), and artificial neural networks, are often used to predict soil properties [32], [36], [37], [38]. Despite having the development of robust modelling, many scholars have explored biases in the accuracy of model predictions in SOC. For example, Hobbey et al. [39] used laboratory imaging spectroscopy and various machine learning methods to predict the distribution of arable land SOC, and found that the RF method relatively produce good performance than other methods. Xu et al. [32] compared the accuracy of PLSR and various machine learning methods for predicting SOC and carbon components and found that SVM generated the best prediction result. However, the optimal model suitable for predicting soil properties varies under different soil formation conditions and requires the development of a specific prediction model for each soil property for the selected study area [13], [30], [32], [40].

The coastal soil in the north of Jiangsu Province of China developed from the sediment of the Yellow and Yangtze Rivers over the past 1000 years and has typical soil formation from the parent materials. Many scholars have studied the centennial or millennial scale dynamics of soil organic carbon in this region using traditional methods [9], [41], [42], yet there is limited research evaluating the potential of imaging spectroscopy technology for investigating soil evolution processes. The objectives of this article are as follows:

- 1) establish prediction models of optimal SOC content and density by using SPXY sample set partitioning methods combined with PLSR, SVM, and RF model;
- 2) directly and indirectly predict the profile SOCD by using imaging spectroscopy;
- 3) compare vertical distribution of SOCD and evolution of the SOCs between predicted and measured values at the millennium scale coastline change in the region.

II. MATERIALS AND METHODS

A. Field Sample Collection and Processing

The study area was in the coastal plain of Dongtai, Jiangsu Province, which has a subtropical monsoonal climate (mean annual temperature 14.6 °C; mean annual precipitation 1042.3 mm). The soil, east of the Fangong Dike, is saline soil, which developed from the sediments transported by the Yellow

and Yangtze Rivers over the past 1000 years. Historical changes of the coastline are shown in Fig. 1(a) [9], [43]. Ten soil pits were nearly vertical to the historical coastline from east to west. The age of soil formation in each profile can be estimated according to the history of the coastline. The detailed information on the sampling sites can be found in Supplementary Table I. The intact profiles were sampled and stored in rectangular wooden boxes (100 cm in length, 20 cm in width, and 5 cm in thickness) [see Fig. 1(b)]. In each soil pit, 10 layer samples were collected for the SOC determination at fixed depth interval: 0–5 cm, 5–10 cm, 10–15 cm, 15–20 cm, 20–30 cm, 30–40 cm, 40–50 cm, 50–60 cm, 60–80 cm, and 80–100 cm [see Fig. 1(c)]. A total of 100 soil samples were obtained from 10 sampling sites. The samples for bulk density (BD) determination were collected from the center of each depth interval using a steel cylinder (100 cm³ volume).

B. Image Spectroscopy Measurement and Preprocessing

1) *Image Spectroscopy Measurement*: The image spectroscopy for the profiles were recorded using an INFINITY V10E imaging spectrometer (Lumenera, Canada) in a dark room. The wavelength ranged between 389 and 1045 nm, and 256 bands were collected. The collection steps were performed in the laboratory: four halogen lamps were used as the light source, the air-dried soil profiles were placed on a mobile platform, and a wooden ruler of 1 m was placed on the side of the wooden box [see Fig. 2(b)]. The profile images were acquired after black and white corrections.

2) *Identification and Removal of Invalid Pixels*: The images were geometrically corrected to remove wooden borders and to retain the original images with a fixed width (400 columns) [see Fig. 2(c)]. Soil profiles were air-dried because the soil water content may affect the model performance [24], [44], [45]. However, the loss of water in the profile can lead to cracks [see Fig. 3(a)], which will cause errors. The spectroscopy curves of soil and crack are obviously distinct, as shown in Fig. 3(b). The soil reflectance increases significantly between 389–600 nm, and then rises with slight fluctuations. However, the reflectance for crack or shadow pixels is generally low. Therefore, the pixels with reflectance below 0.15 at 600 nm were defined as invalid pixels. This reflectance threshold was fine-tuned according to the interprofile variability to classify the invalid pixels.

3) *Computation of Average Spectral Curves for the Layer Samples*: The average spectral curves were extracted by averaging the reflectance of valid pixels in the corresponding depth interval (0–5 cm, 5–10 cm, 10–15 cm, 15–20 cm, 20–30 cm, 30–40 cm, 40–50 cm, 50–60 cm, 60–80 cm, and 80–100 cm). These averaged curves were used as the original spectral data for corresponding layer samples [see Fig. 2(d)].

4) *Pre-Processing of Spectral Curves*: In the process of spectral data collection, certain noise may be generated because of the laboratory environment and the instrument. Then, spectral curve of each pixel in profile image was smoothed and the average spectral curves were smoothed using the Savitzky–Golay method (three orders of 25 points) to remove noise. The bands

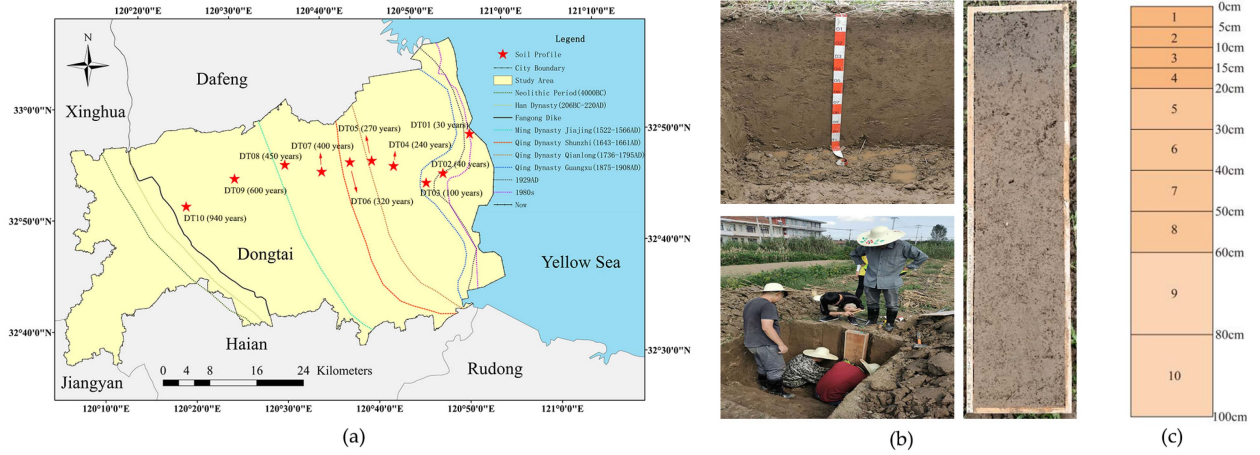


Fig. 1. (a) Sampling location. (b) Process. (c) Depth.

TABLE I
BASIC STATISTICS FOR SOIL PROPERTIES IN COASTAL JIANGSU (CHINA)

Soil Property	Maximum	Minimum	Mean	Standard	CV(%)	Kurtosis	Skewness
SOC (g kg^{-1})	17.03	1.00	4.41	3.41	77.21	1.46	1.10
BD (g cm^{-3})	1.62	1.09	1.40	0.11	7.75	0.03	-0.53
SOCD _{icm} (g m^{-2})	1.86	0.13	0.60	0.42	70.85	-0.49	0.63
SOCs (g m^{-2})	59.47	23.88	44.09	12.16	27.57	-0.64	-0.31

at both ends were removed. We retained the data for 230 bands between the range of 410–1000 nm.

5) *Division for Calibration and Validation Sets*: The SPXY method was used to select the calibration and validation sets based on the independent variable (spectral information) while considering the Euclidean distance from the dependent variable (soil properties) [46]. The distance equations are as follows:

$$d_x p, q = \sqrt{\sum_{j=1}^N [x_p(j) - x_q(j)]^2}; p, q \in [1, N] \quad (1)$$

$$d_y p, q = \sqrt{(y_p - y_q)^2} = |y_p - y_q| \quad (2)$$

$$d_{xy}(p, q) = \frac{d_x p, q}{\max_{p, q \in [1, N]} d_x p, q} + \frac{d_y p, q}{\max_{p, q \in [1, N]} d_y p, q} \quad (3)$$

where x_p, x_q represent two different samples; N represents the number of features of the sample. The distance calculated using (6) was used to select calibration and validation samples. In this article, the ratio of the calibration and validation sets was 7:3 [see Fig. 2(e)].

C. Chemical Analysis

After air-drying and ground, the soil samples were sieved by a 0.15 mm mesh sieve. SOC content was determined using the potassium dichromate oxidation-external heating method. The samples for bulk density were weighed after drying in an oven for 48 h at 105 °C. The following formula was used to calculate

the BD:

$$\text{BD} = \frac{m_t - m_s}{v} \quad (4)$$

where m_t and m_s are the mass of the steel cylinder with the drying soil sample and mass of the empty steel cylinder (g), respectively, and v is the volume of the steel cylinder (100 cm^3).

The soil organic carbon density per centimeter (SOCD_{icm}, g m^{-2}) of layer i refers to the SOC storage of 1 cm interval per unit area. The calculation formula is shown in

$$\text{SOCD}_{icm} = \text{SOC}_i \times \text{BD}_i \times A \times (1 - C_i) / 10. \quad (5)$$

SOCD _{i} for soil layer i was calculated using

$$\text{SOCD}_i = \text{SOC}_i \times \text{BD}_i \times D_i \times (1 - C_i) / 10. \quad (6)$$

The SOC stock (SOC _{s_i}) from surface to certain depth was calculated using

$$\text{SOC}_{s_i} = \sum_{i=1}^n \text{SOCD}_i \quad (7)$$

where i is the number of soil layer; SOC_i , BD_i , D_i , and C_i represent the SOC content (g kg^{-1}), soil BD (g cm^{-3}), soil depth (cm), and the volume (%) occupied by gravel larger than 2 mm, respectively; A represents 1 cm; C_i is negligible because there are no coarse particles in this article area; and 10 is the unit conversion factor.

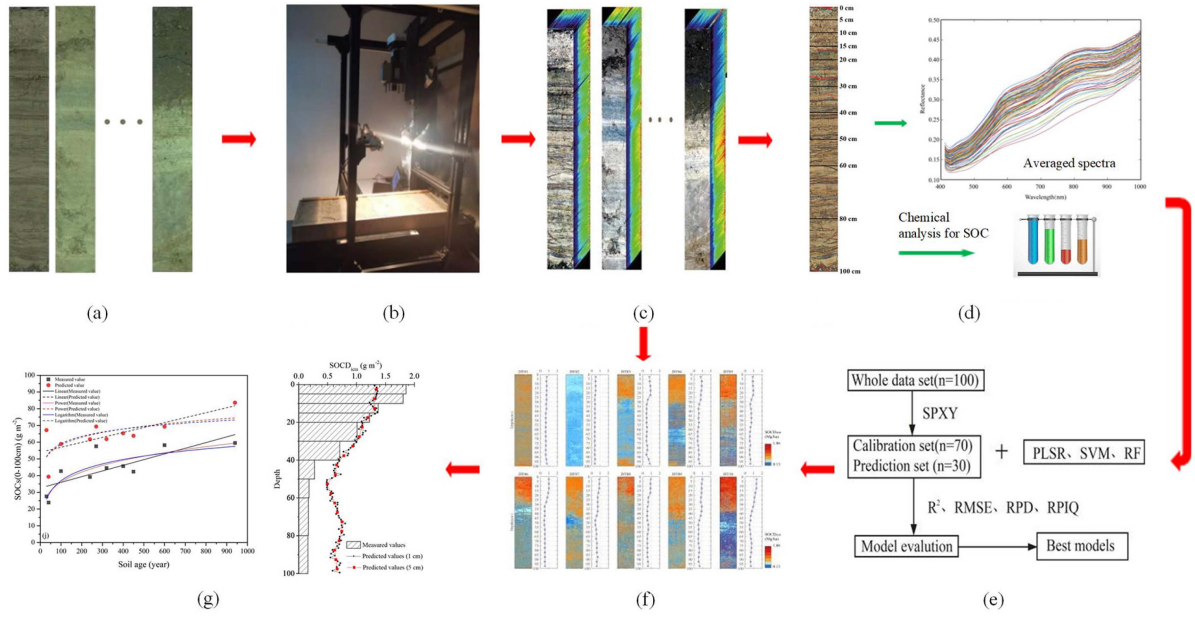


Fig. 2. Flowchart of mapping and analysis for soil profiles via IS. (a) Soil profile. (b) IS scanning system. (c) IS data cubes. (d) Averaged spectra and SOC determination. (e) Vertical distribution for SOCD and chronosequence for SOCs. (f) SOC/SOCD maps. (g) Model development.

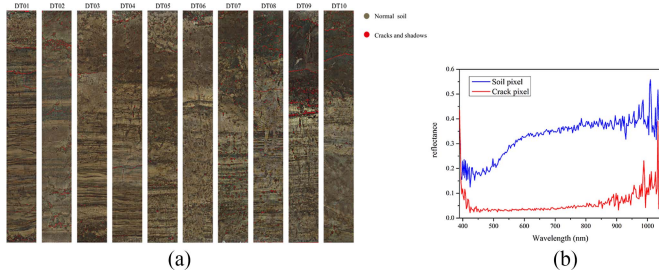


Fig. 3. (a) Identification and removal of invalid pixels. (b) Spectroscopy of normal soil and cracks pixel.

D. Model Development

1) *Indirect and Direct Prediction of SOCD*: Indirect prediction of SOCD involves predicting the organic carbon content using imaging spectroscopy, and then combining it with soil bulk density data to calculate the SOCD. Direct prediction of SOCD refers to directly establishing a regression model between imaging spectroscopy and SOCD. D.B. Predictive models and model evaluation.

The calibrations were built with PLSR, SVM, and RF [see Fig. 2(e)]. PLSR was proposed by Wold et al. [47]. This method combines typical correlation analysis, principal component analysis, and multiple linear regression analysis, effectively solving the multicollinearity problem among independent variables. Its discriminant equation is as follows:

$$R_{i+1}^2 - R_i^2 \leq P \quad (8)$$

where i represents the current number of principal components; P is a preset value usually set to 0.05; If the difference between

R_{i+1}^2 and R_i^2 is less than P when adding a new principal component to the regression model, only the first i principal components are selected to build the model.

SVM can generalize unknown samples for applications such as identification classification, regression modeling, and time series prediction [48]. In this article, the grid search method was used to search the model parameters c and g for the Gaussian radial basis function, and the 10-fold cross-validation was used to determine the best model parameters to minimize the root-mean-squared error (RMSE).

RF can improve prediction accuracy without significantly increasing the amount of computation and is one of the best data mining algorithms available [39]. In this article, the mtry parameter was set to its default value, and the best value of the parameter ntree was selected through 10-fold cross-validation and grid search method.

In this article, three metrics were used to compare the model accuracy: determination coefficient (R^2), RMSE, and RPIQ (RPIQ is the ratio of IQ to RMSEp, IQ is the difference between the third quartile (Q3), and the first quartile (Q1) of the sample observations). The higher R^2 and RPIQ and lower RMSE, the better model is [49], [50]. RPIQ is a more objective index than RPD for the model evaluation of non-normally distributed soil data [49], [50]. The subscript C represents the modeling results of the calibration set. The subscript P represents the independent validation results.

The best model was used to map SOC or SOCD for profiles [see Fig. 2(f)]. Then, the predicted values for 1 cm, 5 cm, and corresponding intervals were extracted by averaging the corresponding pixel values and used to validate the results and analyze the vertical distribution and chrono-sequence [see Fig. 2(g)].

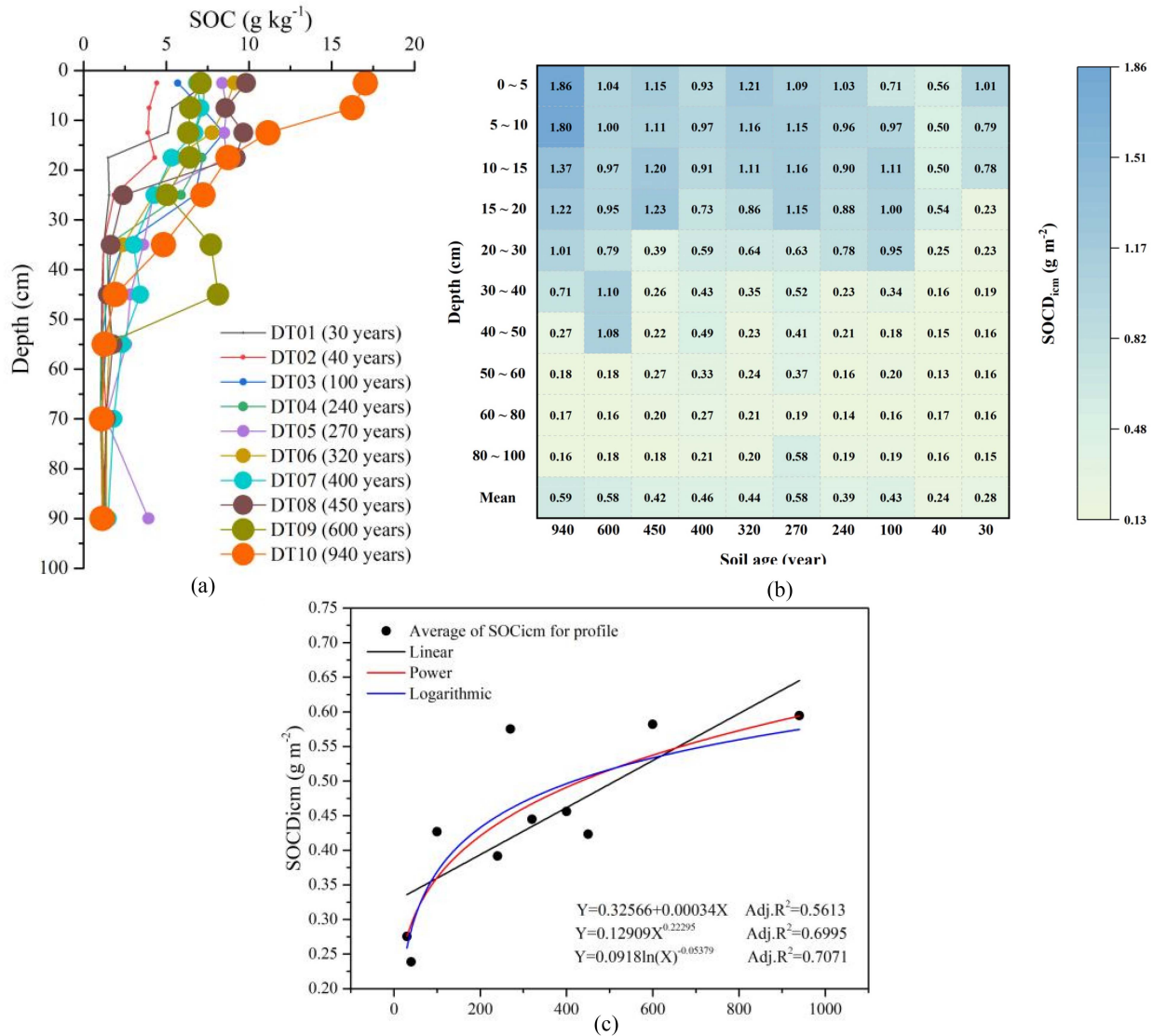


Fig. 4. (a) Distribution of SOC, (b) SOCD_{icm} in soil profiles and the average (c) SOCD_{icm} changing with time.

III. RESULTS

A. Characteristics of Soil Properties

1) *Identification of SOC Variation:* The soil properties are shown in Table I. The SOC content of coastal soil in northern Jiangsu was recorded generally low. The mean value, was 4.41 g kg⁻¹ with a large difference in range (1–17.03 g kg⁻¹). The coefficient of variation (CV) was as high as 77.21%. The BD, (1.09–1.62 g cm⁻³) had the standard deviation and CV were 0.11 and 7.75%, respectively. The soil organic carbon density (SOCD_{icm}) ranged from 0.13 to 1.86 g m⁻² with a CV of 70.85%. Among 10 soil profiles studied, the highest and the lowest SOCs were 59.47 g m⁻² and 23.88 g m⁻². Kurtosis can describe the steepness of the data distribution pattern, with three representing a normal distribution and 1.8 representing a uniform distribution. Skewness measures the degree of skewness of the data distribution, with zero representing a normal distribution. None of the soil properties satisfied the normal distribution.

2) *Variation of SOC in Soil Profile:* The SOC content in the coastal Jiangsu was found to be varying with depth, but in certainty and in a regular basis [see Fig. 4(a)]. Except for DT09, the SOC content in the profiles decreased with depth, and tended to be gentle in the deep layer. The anomaly in DT09 may be related to the buried peat layer. In addition, the vertical distribution of SOC in the profiles with different soil formation ages differed significantly. The profiles with a short soil formation age exhibited low SOC accumulation and a relatively uniform vertical distribution.

With an increase in soil age, the accumulation of SOC gradually extended to subsurface layers. The SOC accumulation in profile DT10 was more significant than that in the other profiles.

The vertical distribution of SOCD_{icm} showed the same trend as SOC [see Fig. 4(b)]. The SOCD_{icm} decreased along with depth. In general, SOCD_{icm} in older profiles were higher than those in younger profiles for the upper 40 cm. For the entire profile, the average SOCD_{icm} ranged from 0.27 g m⁻² (DT02)

TABLE II
STATISTICS OF ORGANIC CARBON CONTENT OF SOILS DIVIDED BY SPXY METHOD

Soil Property	Number of Samples	Maximum	Minimum	Mean	Standard	CV(%)	Kurtosis	Skewness
Total	100	1.00	17.03	4.41	3.41	77.21	1.46	1.10
Calibration	70	1.00	17.03	4.84	3.55	73.23	1.60	1.10
Prediction	30	1.09	9.20	3.40	2.85	83.92	-0.95	0.87

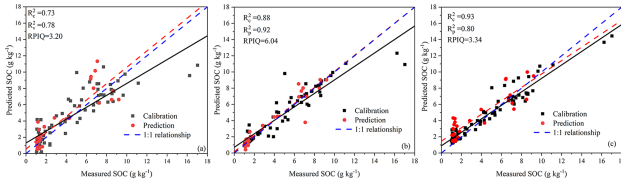


Fig. 5. Modeling results for SOC content [(a) PLSR; (b) SVM; (c) RF].

to 0.60 g m^{-2} (DT10) and generally increased with soil age. The power and logarithmic functions showed comparable results, better than linear function, which indicated the decreasing rate of SOC sequestration in this chronosequence [see Fig. 4(c)].

B. Prediction of SOC Density by Spectroscopy

1) *Indirect Prediction of SOC Density by Spectroscopy:* Imaging spectroscopy was commonly used for estimating SOC content with high estimation accuracy, and SOC content can be converted to SOCD data by BD [23], [27], [29], [30]. Therefore, we first established a SOC estimation model based on averaged spectra and mapped the SOC content for profiles using imaging spectroscopy, and then obtained the SOCD map combined with the corresponding BD for layer samples.

a) *Statistical analysis of SOC content in sample sets:* Table II showed the SOC content statistics of the sample set divided by the SPXY method. The mean and standard deviation of the SOC content in the model set were 4.84 g kg^{-1} and 3.55 g kg^{-1} , respectively. The kurtosis and skewness were 1.60 and 1.10, which had only had a little difference from that of the total sample set. The content of the SOC in the prediction set varied from 1.09 g kg^{-1} to 9.20 g kg^{-1} , and the mean and standard deviation decreased to 3.40 g kg^{-1} and 2.85 g kg^{-1} , respectively. The spatial heterogeneity increased, and the CV was 83.92%. The kurtosis was -0.95, indicating that the data distribution was relatively flat compared with the normal distribution.

b) *Modeling results for SOC content:* Three methods (PLSR, SVM, and RF) were used to establish SOC content prediction model, and the prediction accuracy was compared (see Fig. 5). The results demonstrated that the R^2_c of RF was the highest at 0.93 but that R^2_p was low (0.80). Comparatively, the R^2_c of SVM and PLSR were second to RF, 0.88 and 0.73, respectively. But the R^2_p of SVM was higher than that of RF, at 0.92. The RPIQ of SVM was higher than that of the PLSR and RF, indicating that SVM was more suitable to map SOC content for profiles.

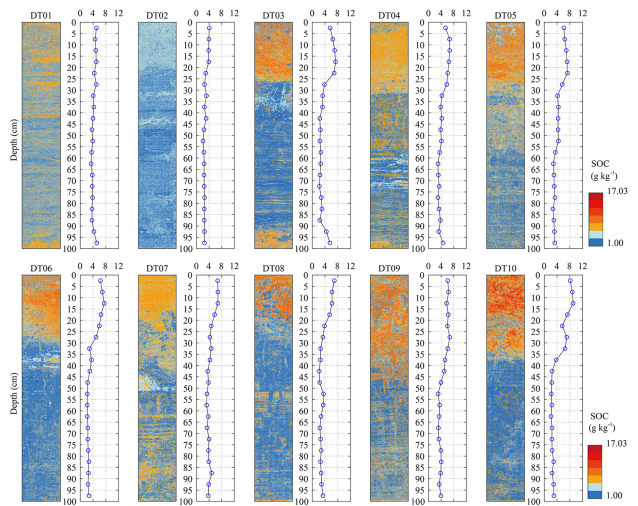


Fig. 6. Estimation results of organic carbon content in profiles with different soil formation ages.

c) *Mapping results of SOC content in profiles:* The SVM model was used to map the SOC content for the profiles, and the results were shown in Fig. 6. There were significant differences in the vertical distribution of SOC in profiles of different age. The SOC contents in DT01 and DT02 were relatively low with mean value 4 g kg^{-1} . The low carbon content could be due to the shorter time of soil formation, especially in DT02 profile. In DT03, DT04, DT05, DT06, DT07, and DT08, the SOC content in the surface layer (30 cm) ranged from $4\text{--}8 \text{ g kg}^{-1}$. The SOC content in DT09 at 0–40 cm layer ranged from $4\text{--}8 \text{ g kg}^{-1}$ while in DT10 at 0–15 cm layer was measured above 8 g kg^{-1} and at 15–40 cm layer, ranged from $4\text{--}8 \text{ g kg}^{-1}$. Interestingly, the SOC content in the 40–100 cm layer was less than 4 g kg^{-1} in almost all profiles. The SOC contents in DT01, DT03, and DT04 increased slightly at 90–100 cm depth, because of the organic residues buried during deposition. The mapping results showed similar vertical distribution and temporal trend of the SOC content as the measured values presented in Fig. 4(a). It is clear that the SOC was found to be more concentrated in the surface layer, decreased along with depth and eventually stabilized for ageing.

d) *Indirect mapping of SOC_{icm} for profiles:* The SOC_{icm} distribution was calculated by combining the SOC map with the corresponding BD (see Fig. 7). As shown in Figs. 6 and 7, the vertical distribution of the SOC content and density in the profiles were roughly similar. The SOC_{icm} in the younger

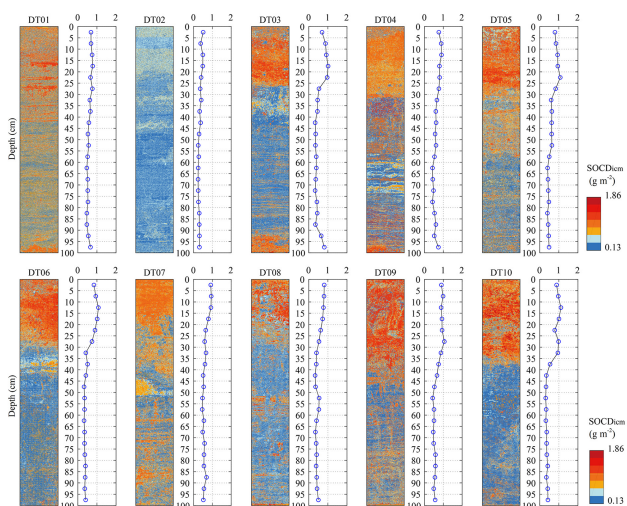


Fig. 7. Indirect estimation results of SOC density for profiles with different soil ages.

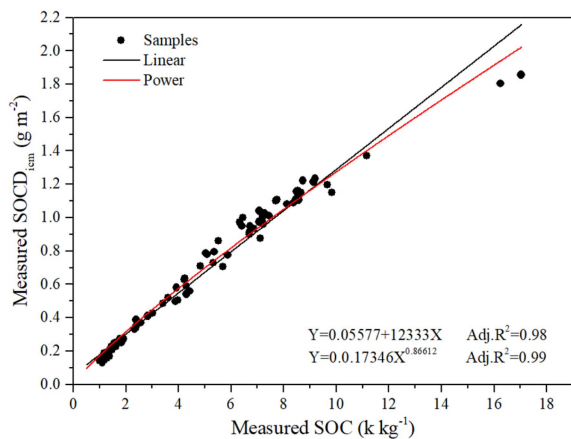


Fig. 8. Correlation between SOC content and $SOCD_{icm}$.

profiles, DT01 and DT02, was low and relatively consistent with, less than 0.50 g m^{-2} . The $SOCD_{icm}$ content at 0–25 cm depth of the older (formed during 100 to 500 years ago) profiles such as DT03, DT04, DT05, DT06, DT07, and DT08, the was relatively higher than the younger profiles. The DT09 and DT10 soil profiles of 600 and 940 years old, where the $SOCD_{icm}$ content was measured highest at 0–35 cm depth. Except for DT01 and DT02, the vertical distribution of $SOCD_{icm}$ varied significantly with depth, and most of the $SOCD_{icm}$ at 35–100 cm depth was less than 0.5 g m^{-2} .

2) *Direct Prediction of Soil Organic Carbon Density by Spectroscopy:*

a) *Correlation between SOC content and SOC density:* Fig. 8 shows the correlation between SOC content and $SOCD_{icm}$. The fitting results for the linear and power functions were R^2 values of 0.98 and 0.99, respectively. It indicated that there was a strong correlation between SOC content and $SOCD_{icm}$. In theory, $SOCD_{icm}$ could be predicted directly using imaging spectral data.

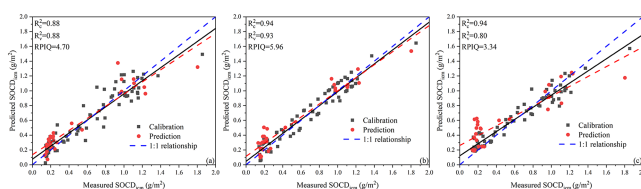


Fig. 9. Modeling results of $SOCD_{icm}$ (a) PLSR; (b) SVM; (c) RF.

b) *Statistical analysis of $SOCD_{icm}$ in sample sets:* The same sample sets as before were used for calibration and prediction sets (see Table III). The mean and standard deviation of the modeling set were 0.64 g m^{-2} and 0.40 g m^{-2} , respectively, and the kurtosis and skewness were -0.58 and 0.42 , which had a small difference with the total set. The CV was 62.55%, which was smaller than the total set. The $SOCD_{icm}$ of the prediction set varied from 0.13 to 1.80 g m^{-2} , with an average value of 0.50 g m^{-2} and a CV as high as 93.27%. The kurtosis and skewness were 0.49 and 1.23, indicating that the data was skewed to the right compared to the mean value.

c) *Modeling results of SOC density:* The smoothed spectral data and $SOCD_{icm}$ were used to build PLSR, SVM, and RF models and the prediction performance was compared (see Fig. 9). The modeling and prediction performance of SVM was higher than those of PLSR and RF, up to 0.94 and 0.93, respectively. The RPIQ for SVM was 5.96, which was much higher than those of the other two models. The SVM got the best model for this article, which was suitable for mapping the $SOCD_{icm}$ for the profiles.

d) *Mapping results of $SOCD_{icm}$ in profiles:* The SVM model was used to map the $SOCD_{icm}$ for the profiles directly (see Fig. 10). The $SOCD_{icm}$ of the invalid pixel was expressed as the average of the nearby values. The results (see Fig. 10) were similar with the maps predicted indirectly by SOC and BD (see Fig. 7). The $SOCD_{icm}$ was low and there was no significant difference in the vertical distribution in younger profiles (DT01 and DT02). In other profiles, the $SOCD_{icm}$ of the surface and subsurface soil were significantly different. These $SOCD_{icm}$ maps also showed that the vertical distribution of $SOCD_{icm}$ in profiles with different ages differed significantly. And the amount and accumulation depth of SOC increased with age.

3) *Evaluation of Mapping Results:* To evaluate the mapping accuracy for SOC and $SOCD_{icm}$, the maps for SOC and $SOCD_{icm}$ were averaged according to the sampling depths. Then, the averaged mapping values for the 100 layer samples were compared with the measured values. The results showed that the R^2 and RMSE of the predicted SOC content reached 0.77 and 2.07 g kg^{-1} , and the RPIQ was 2.79, indicating that the model produced a good mapping result for SOC [see Fig. 11(a)]. The R^2 for indirect prediction of $SOCD_{icm}$ was 0.68, RMSE was 0.27 g m^{-2} , and RPIQ was 2.85, respectively, indicating that the indirect mapping of $SOCD_{icm}$ by the model was generally effective [see Fig. 11(b)]. The mapping result for direct prediction of $SOCD_{icm}$ ($R^2 = 0.78$, RPIQ = 2.83) was slightly better compared to the indirect prediction [see Fig. 11(c)]. These three scatter plots also showed a phenomenon: the high value tend

TABLE III
STATISTICS OF ORGANIC CARBON CONTENT OF SOILS DIVIDED BY SPXY METHOD

Soil Property	Number of Samples	Max	Min	Mean	Std	CV(%)	Kurtosis	Skewness
Total	100	0.13	1.86	0.60	0.42	70.85	-0.49	0.63
Calibration	70	0.14	1.86	0.64	0.40	62.55	-0.58	0.42
Prediction	30	0.13	1.80	0.50	0.46	93.27	0.49	1.23

TABLE IV
PREDICTION OF SOIL ORGANIC CARBON/MATTER CONTENT BASED ON LABORATORY IMAGING SPECTROSCOPY

Spectral range /nm	Sample number	Multivariate method	R_{cv}^2	R_p^2	RPD	Literature Source
450–1671	400	PLSR	0.75	0.75	--	O'Rourke et al. [35]
400–990	66	PLSR	--	0.81	--	Steffens and Buddenbaum [23]
400–990	66	SVM	--	0.97	--	Steffens and Buddenbaum [23]
450–1671	366	PLSR	0.76	0.75	1.98	Askari et al. [22]
400–2500	18	PLSR	0.73	0.74	--	Ogen et al. [33]
400–990	33	PLSR	--	0.85	2.60	Hobley et al. [39]
400–990	33	RF	--	0.91	3.30	Hobley et al. [39]
400–1000	96	PLSR	0.71	0.76	2.02	Tahmasbian et al. [26]
400–1010	306	PLSR	0.82	0.88	2.85	Xu et al. [30]
400–1010	306	SVM	0.91	0.97	5.50	Xu et al. [30]

Notes: R_{cv}^2 represents the cross-validation results; R_p^2 represents the independent validation results; RPD is the relative percent deviation.

to be underestimated and low value tend to be overestimated. As a result, the vertical variation for SOC and $\text{SOCD}_{i_{cm}}$ in the profiles was underestimated.

The SOC_i is the accumulated SOCD_i from surface to certain depth (i th layer), representing the amount of SOC stock within certain depth in the profile. The R^2 was 0.84 and the RPIQ was 2.45, indicating that the map for $\text{SOCD}_{i_{cm}}$ can be used to estimate SOC stock for the profiles with better accuracy [see Fig. 11(d)]. The SOC_i were underestimated within 40 cm and then the values were overestimated beyond 40 cm due to the phenomenon of overestimation for low values and the underestimation for high values. That is, the underestimation for high $\text{SOCD}_{i_{cm}}$ led to the underestimation within 40 cm soil. The overestimation for low $\text{SOCD}_{i_{cm}}$ and the greater thickness of subsoil led to the overestimation for SOC_i for deeper soil and the whole profile.

C. Temporal Function of SOC's Accumulation

According to the relationship between soil formation age and SOC's at different depths, the evolution of SOCD in this region can be inferred on a millennium scale (7). The scatter diagrams between the measured and predicted SOC's at different depths and soil ages are shown in Fig. 12. The R^2 values for the linear, logarithmic, and power functions were used to determine the best fitting results for SOC's and soil age. In general, the measured SOC's were greater than the predicted values for surface soil within 20 cm, and the predicted SOC's were greater than the measured values for deeper soil (>50 cm).

The temporal function of measured SOC's varies continuously with the sampling depth. The accumulation of surface SOC's continued at a stable rate and was roughly linear on a millennial

time scale. With increasing soil depth, the fitted function of SOC's with time was closed to a power function and then changed to a logarithmic function, indicating that the deep soil could not reach the organic carbon stock accumulation rate of the surface soil, thus slowing the rate of SOC's accumulation in the whole soil profile. For the predicted values of SOC's, the linear function was the best fit for all soil depths, which was a result of the overestimation of low values and underestimation of high values of SOCD from the spectral data.

IV. DISCUSSION

A. Calibration Methods and Their Performance

Many scholars have used imaging spectroscopy to predict SOC/SOM (see Table IV). PLSR was used in earlier studies while SVM or RF has been used in recent studies. Our research indicates that the models established by SVM and RF methods are superior to PLSR, which is consistent with previous research results (see Table IV). This is mainly because machine-learning techniques can effectively handle the nonlinear relationship between spectral features and soil properties [34]. However, although these nonlinear models have higher predictive accuracy, there is a "black box" problem where the model structure is complex and the transparency is generally low [51], making it difficult to reveal the functional relationship between spectra and soil properties. It is also difficult to understand the contribution of each band to property prediction in model construction. Linear models have lower predictive accuracy but simpler structures and typically have lower computational costs [13], [30], which are advantageous for identifying response characteristics between spectra and properties from large amounts of data. Therefore, the predictive performance of different models is

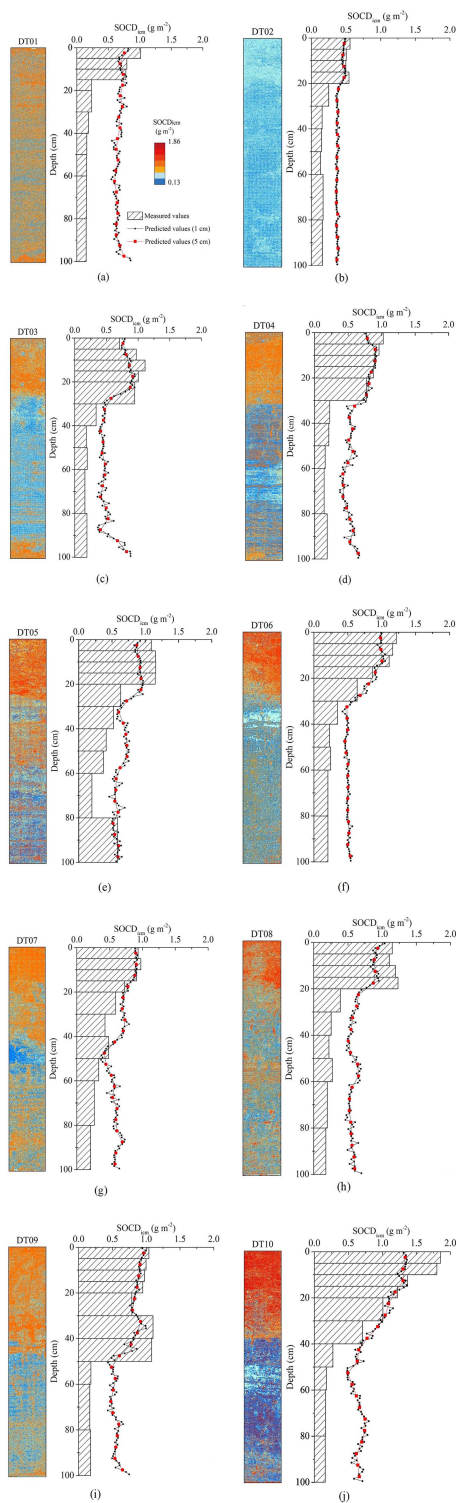


Fig. 10. Estimation results of organic carbon density in profiles with different soil formation ages. (a)-(j) Represent the mapping results of SOCD for DT01 to DT10 based on the best model.

related to their underlying mechanisms, and each method has advantages and disadvantages that must be considered based on research purposes and data availability [32].

Previous studies showed that the imaging spectroscopy could effectively map the SOC content for profiles [22], [27], [29],

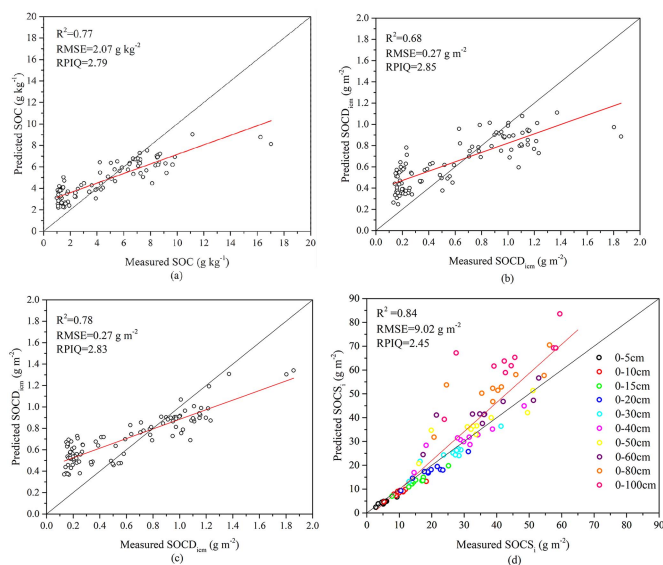


Fig. 11. Scatter plot of mapped and measured SOC, SOCD_{icm} , and SOCS_i : (a) SOC; (b) indirect mapping for SOCD_{icm} ; (c) direct mapping for SOCD_{icm} ; (d) SOCS_i .

[30], [33], [35], [39]. However, SOCD_{icm} needs to be calculated based on the BD [52], [53], [54], which is complicated and laborious. Therefore, a question arises whether SOCD_{icm} can be predicted by spectral data directly with improved performance. In this article, we compared the consistency between direct and indirect predictions of SOCD_{icm} with the measured values, and the results showed that direct prediction accuracy was better than that of indirect prediction [see Fig. 11(b) and (c)]. Three modeling methods achieved good performance in predicting SOCD_{icm} (see Fig. 9), with SVM method obtaining the best model, with R^2 values of 0.94 and 0.93 for the calibration and prediction sets, respectively. This means that it is possible to predict SOCD_{icm} using spectral data directly and accurately, thus filling the gap in BD data in SOC inventory calculations, which is of great significance for estimating global soil carbon stocks. The 1:1 relationship indicated that all three models exhibited both over- and underestimation of SOCD_{icm} predictions (see Fig. 9). The reason for overestimation of high values may have been due to the limited number of soil samples with high SOCD_{icm} concentrations, which were only found in the top 0–10 cm layer of the DT10 profile with contents exceeding 1.5 g m^{-2} [see Fig. 10(j)]. Underestimation of low values may have been related to the weakening of the spectral prediction ability of soil organic matter with decreasing content [55]. Zheng et al. [13] used imaging spectroscopy combined with multivariate regression to predict the content of five soil fertility properties in Cixi area, and the results showed that all six models exhibited over- and underestimation of the five soil properties, which was consistent with this article. Since the prediction of soil properties by spectroscopy is influenced by multiple factors such as spectral bands, soil characteristics, and preprocessing procedures [13], [30], [56], the underlying mechanisms of these prediction errors were not fully understood, and further research was needed for more accurate prediction.

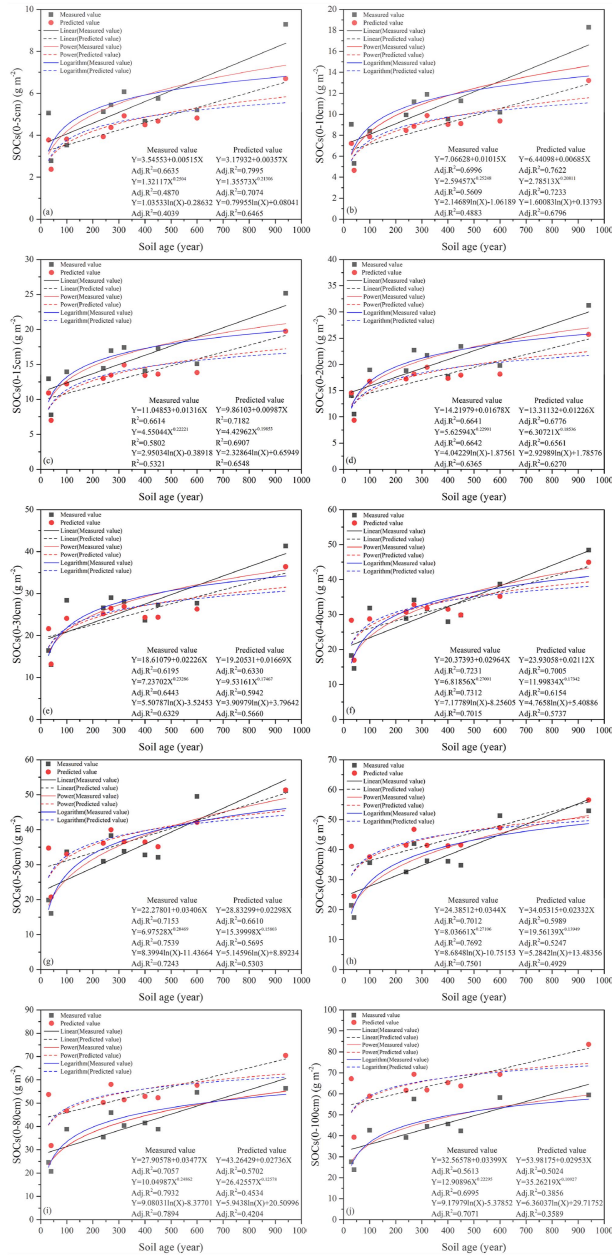


Fig. 12. Temporal function of soil organic carbon stock for different depths.

In addition, the variable soil parent material and carbon source can also affect the model's predictive performance [57], [58], [59]. The soil samples selected in this article developed in the typical parent material area of the Yellow River and Yangtze River accumulated under the interaction of land and sea in the past 1000 years, which has reference significance for studying soil chronosequence in other coastal areas. However, soils developed on different parent materials from other regions may have different applicable models and results [60], [61]. For instance, Liu et al. [60] investigated the predictive performance of three regression models for soil organic carbon content developed on different parent materials (trachyte and basalt) and found that the best model and spectral response characteristics differed.

Therefore, the model we used was applicable to this region and was not considered globally applicable. It is, however, necessary to collect soil samples from different regions for SOC/SOM and type evaluations and to extend the spectral prediction model to a larger range.

B. Variation of SOCD in Soil Profile

Many scholars have already conducted research on the vertical variation of SOCD, and their findings indicated that the surface soil had a higher SOCD compared to the lower layers [9], [62], [63], [64]. In this article, we also observed a similar distribution pattern of SOCD [see Fig. 4(a)]. The reason may be that the decomposition of plant leaves and roots formed organic matter that accumulated rapidly in the surface layer, leading to a significant increase in the surface SOCD [5], [65]. Imaging spectroscopy is an effective new method for mapping soil properties, which allows us to observe the vertical distribution characteristics of soil properties in more detail compared to traditional reflectance spectroscopy [24], [30]. To evaluate and validate the prediction results, we extracted the continuous depth variation curve of SOCD from the soil profile mapping results, and displayed the vertical distribution of measured values on the right side of the profile map (see Fig. 10). The continuous depth variation trend of SOCD in the topsoil (0–40 cm) was relatively close to the measured values, while in the deeper layers (below 40 cm), it was much higher than the measured values. The main reason for this was that the model had poor accuracy in estimating the SOCD for lower contents (less than 0.5 g m^{-2}). The mapping results of soil organic matter content in the soil profile by Zheng et al. [13] showed that the best model overestimated the organic matter content in soil samples with values lower than 5 g kg^{-1} to varying degrees, consistent with this article. These results indicated that when mapping for SOC/SOCD, the mapping results in the low-content areas may have been unreliable. Although many studies have conducted high-resolution mapping of various soil properties in both vertical and horizontal directions, all the validation results of their established models have exhibited overestimation of high values and underestimation of low values [13], [30], [32], [40]. Further research is needed to determine the accuracy in estimating the content of attribute in high and low value regions. In addition, the generalized ability of the model may also affect the mapping effect. Steffens and Buddenbaum [23] pointed out that a high-precision regression model does not always represent the best mapping results. In their study, the accuracy of PLSR was the lowest, but its generalized potential for reasonable mapping of profile properties was the highest, while SVM had the highest accuracy but the poorest robustness and lower generalization potential.

This article filled the gap in the application of imaging spectroscopy technology in soil chronosequence research by mapping SOCD in soil profiles at different ages. The results suggest that imaging spectroscopy technology could be used to directly map the topsoil (0–40 cm) SOCD of older profiles (DT03–DT10), but the predicted accuracy of younger profiles (DT01, DT02) and lower SOCD areas (below 40 cm) needed to be improved. Nonetheless, imaging spectroscopy technology

provide a fast and non-destructive method for observing the spatial structure of soil properties in profiles [30]. Compared to the limitations of traditional fixed-depth sampling intervals, this technology could better serve the study of profile horizon division, dynamic changes in soil properties, and achieve better soil survey tasks [13], [23], [30].

C. Temporal Variation of Organic Carbon Density

Soil chronosequence is of greater significance to explore the direction and rate of overall soil evolution, providing essential information for soil genetic theory [66], [67], [68]. In this article, power functions and logarithmic functions were found to be better fit the changes in 1 cm average SOCD over time compared to linear functions [see Fig. 4(c)], indicating that the soil carbon sequestration rate gradually slowed down over time for the entire profile. This was mainly because the deep soil could not accumulate organic carbon as efficiently as the topsoil, leading to a decrease in the accumulation rate of organic carbon in the entire depth of the soil [see Fig. 4(b)]. The best-fit function for the topsoil (0–15 cm) was linear, but with increasing soil depth, it became power functions and logarithmic functions (see Fig. 12), which also supported this view. However, our results were different from that of other soil evolution studies in which the accumulation rate of SOM with time at different depths all showed a linear growth trend [13]. The main reason may be that soil with different parent materials had different carbon sequestration capacity [69].

For the predicted values of SOC_s, the linear function was the best fit for the 0–15 cm topsoil, which was consistent with the measured results (see Fig. 12). However, the best fitting function for predicted values at other depths was different from the measured results, with the best fitting function for predicted values being linear and for measured values being power function. Traditional sampling methods were commonly used in previous studies to analyze the direction and rate of soil evolution [69], [70], [71], [72], making it difficult to directly compare the results of the best fitting functions obtained from imaging spectroscopy with those of other studies. Nevertheless, the fitting trends of SOC_s over time showed minimal differences among the three functions for both measured and predicted values. This suggests that imaging spectroscopy technology holds great potential for quantitatively revealing the process of soil evolution. Despite the current limitations in accuracy, this article provided a new insight for the rapid monitoring of soil property changes under global climate change scenario. Future research should focus on exploring more intelligent prediction models [73], [74], [75], improve the generalized ability of the models, and reducing the extent of overestimation and underestimation of high and low values, respectively. This will enable more accurate soil property mapping to support strategic planning for regional and even national-level sustainable development goals related to soil conservation and management.

D. Limitations and Perspectives of This Article

The results of this article indicated that utilizing imaging spectroscopy data for direct or indirect estimation of SOCD

distribution in soil profiles was a feasible option for studying soil evolution processes. However, there were still some limitations to this approach. First, the soil samples used for chemical analysis were grounded, which may have reduced the influence of environmental factors. However, the surface of intact soil profiles was much rougher and may decrease the accuracy of estimating SOC/SOCD. Therefore, it is necessary to develop algorithms to eliminate the light scattering effect caused by soil surface roughness. Second, the developed model for estimating SOCD and the fitted soil chronosequence function in this article are only applicable to the studied region. Under different environmental conditions, the accumulation rate and vertical distribution of SOC may present varying results. Therefore, it is necessary to compare the results obtained from applying imaging spectroscopy technology to study soil evolution to the other region.

In the future research, it is desired to collect soil profile samples under differential parent material conditions to construct soil chronosequence and use more advanced imaging spectrometers to detect more wavelengths to enrich soil information. Additionally, it may be possible to extend the developed model in the article to larger spatiotemporal scales. These can provide a theoretical basis for the soil evolution and pedogenesis.

V. CONCLUSION

This article utilized PLSR, SVM, and RF methods to predict the SOCD directly or indirectly in the entire depth (0–100 cm) of soil profiles at different ages. The results showed that the accuracy of direct prediction was better than that of indirect prediction, and the SVM method produced the best model, with R^2 values of 0.94 and 0.93 for the calibration set and validation set, respectively. The variation of measured SOCD in soil profile showed that SOC mainly accumulates in the topsoil (0–40 cm), with little accumulation in the deeper soil (below 40 cm). The mapping of SOCD achieved good results in the 0–40 cm layer of older profiles (DT03-DT10), but the prediction accuracy in younger profiles (DT01, DT02) and deep soil layers needed to be further improved. Power functions and logarithmic functions were found to be better fit the changes in 1 cm average SOCD over time compared to linear functions, indicating that the soil carbon sequestration rate in this area gradually slowed down over time for the entire profile. For the predicted values of SOC_s, the best temporal functions were different from that of the measured values at most depths, but the fitting trends of changes with time by three functions were basically consistent. It suggests that the IS technology had the potential to quantitatively reveal the process of coastal soil evolution. Future research should focus on exploring more intelligent prediction models, and improve the existing generalization in the ability and prediction accuracy of the models.

REFERENCES

- [1] E. G. Jobbágy and R. B. Jackson, "The vertical distribution of soil organic carbon and its relation to climate and vegetation," *Ecological Appl.*, vol. 10, no. 2, pp. 423–436, Apr. 2000.

- [2] J. M. F. Johnson, A. J. Franzluebbers, S. L. Weyers, and D. C. Reicosky, "Agricultural opportunities to mitigate greenhouse gas emissions," *Environ. Pollut.*, vol. 150, no. 1, pp. 107–124, Nov. 2007.
- [3] R. Lal, "Soil carbon sequestration impacts on global climate change and food security," *Science*, vol. 304, no. 5677, pp. 1623–1627, 2004.
- [4] K. Paustian, J. Lehmann, S. Ogle, D. Reay, G. P. Robertson, and P. Smith, "Climate-smart soils," *Nature*, vol. 532, no. 7597, pp. 49–57, 2016.
- [5] B. Minasny et al., "Soil carbon 4 per mille," *Geoderma*, vol. 292, pp. 59–86, 2017.
- [6] E. Liu et al., "Long-term effects of no-tillage management practice on soil organic carbon and its fractions in the northern China," *Geoderma*, vol. 213, pp. 379–384, 2014.
- [7] N. H. Batjes, "Total carbon and nitrogen in the soils of the world," *Eur. J. Soil Sci.*, vol. 65, no. 1, pp. 10–21, 2014.
- [8] K. R. Olson and M. Al-Kaisi, "The importance of soil sampling depth for accurate account of soil organic carbon sequestration, storage, retention and loss," *Catena*, vol. 125, pp. 33–37, 2015.
- [9] C. Jiao, G. Zheng, X. Xie, X. Cui, G. Shang, and C. Zhao, "Rate of soil organic carbon sequestration in a millennium coastal soil chronosequence in northern Jiangsu, China," *Catena*, vol. 193, 2020, Art. no. 104627.
- [10] R. V. Rossel, D. Brus, C. Lobsey, Z. Shi, and G. McLachlan, "Baseline estimates of soil organic carbon by proximal sensing: Comparing design-based, model-assisted and model-based inference," *Geoderma*, vol. 265, pp. 152–163, 2016.
- [11] R. R. Poppiel, A. F. S. Paiva, and J. A. M. Demattê, "Bridging the gap between soil spectroscopy and traditional laboratory: Insights for routine implementation," *Geoderma*, vol. 425, 2022, Art. no. 116029.
- [12] S. Chen et al., "Digital mapping of GlobalSoilMap soil properties at a broad scale: A review," *Geoderma*, vol. 409, 2022, Art. no. 115567.
- [13] G. Zheng, A. Wang, C. Zhao, M. Xu, C. Jiao, and R. Zeng, "Evolution of paddy soil fertility in a millennium chronosequence based on imaging spectroscopy," *Geoderma*, vol. 429, Jan. 2023, Art. no. 116258.
- [14] R. A. V. Rossel et al., "Diffuse reflectance spectroscopy for estimating soil properties: A technology for the 21st century," *Eur. J. Soil Sci.*, vol. 73, no. 4, 2022, Art. no. e13271.
- [15] F. Nyarko, F. M. Tack, and A. M. Mouazen, "Potential of visible and near infrared spectroscopy coupled with machine learning for predicting soil metal concentrations at the regional scale," *Sci. Total Environ.*, vol. 841, 2022, Art. no. 156582.
- [16] A. F. S. Paiva, R. R. Poppiel, N. A. Rosin, L. T. Greschuk, J. T. F. Rosas, and J. A. Demattê, "The Brazilian Program of soil analysis via spectroscopy (ProBASE): Combining spectroscopy and wet laboratories to understand new technologies," *Geoderma*, vol. 421, 2022, Art. no. 115905.
- [17] Y. Liu, X. Xie, M. Wang, Q. Zhao, and X. Pan, "Removing the effects of iron oxides from vis-NIR spectra for soil organic matter prediction," *Soil Sci. Soc. Amer. J.*, vol. 82, no. 1, pp. 87–95, 2018.
- [18] X. Jia, S. Chen, Y. Yang, L. Zhou, W. Yu, and Z. Shi, "Organic carbon prediction in soil cores using VNIR and MIR techniques in an alpine landscape," *Sci. Rep.*, vol. 7, no. 1, 2017, Art. no. 2144.
- [19] Q. Jiang, Q. Li, X. Wang, Y. Wu, X. Yang, and F. Liu, "Estimation of soil organic carbon and total nitrogen in different soil layers using VNIR spectroscopy: Effects of spiking on model applicability," *Geoderma*, vol. 293, pp. 54–63, 2017.
- [20] Z. Shi, W. Ji, R. V. Rossel, S. Chen, and Y. Zhou, "Prediction of soil organic matter using a spatially constrained local partial least squares regression and the Chinese vis-NIR spectral library," *Eur. J. Soil Sci.*, vol. 66, no. 4, pp. 679–687, 2015.
- [21] S. Adar, Y. Shkolnisky, and E. Ben-Dor, "Change detection of soils under small-scale laboratory conditions using imaging spectroscopy sensors," *Geoderma*, vol. 216, pp. 19–29, 2014.
- [22] M. S. Askari, S. M. O'Rourke, and N. M. Holden, "A comparison of point and imaging visible-near infrared spectroscopy for determining soil organic carbon," *J. Near Infrared Spectrosc.*, vol. 26, no. 2, pp. 133–146, 2018.
- [23] M. Steffens and H. Buddenbaum, "Laboratory imaging spectroscopy of a stagnic Luvisol profile—High resolution soil characterisation, classification and mapping of elemental concentrations," *Geoderma*, vol. 195, pp. 122–132, 2013.
- [24] E. Ben-Dor et al., "Using imaging spectroscopy to study soil properties," *Remote Sens. Environ.*, vol. 113, pp. S38–S55, 2009.
- [25] A. K. Nayak et al., "Current and emerging methodologies for estimating carbon sequestration in agricultural soils: A review," *Sci. Total Environ.*, vol. 665, pp. 890–912, May 2019.
- [26] I. Tahmasbian et al., "Laboratory-based hyperspectral image analysis for predicting soil carbon, nitrogen and their isotopic compositions," *Geoderma*, vol. 330, pp. 254–263, 2018.
- [27] P. Sorenson, S. Quideau, and B. Rivard, "High resolution measurement of soil organic carbon and total nitrogen with laboratory imaging spectroscopy," *Geoderma*, vol. 315, pp. 170–177, 2018.
- [28] I. Burud, C. Moni, A. Flo, C. Futsaether, M. Steffens, and D. P. Rasse, "Qualitative and quantitative mapping of biochar in a soil profile using hyperspectral imaging," *Soil Tillage Res.*, vol. 155, pp. 523–531, 2016.
- [29] P. Sorenson, S. Quideau, B. Rivard, and M. Dyck, "Distribution mapping of soil profile carbon and nitrogen with laboratory imaging spectroscopy," *Geoderma*, vol. 359, 2020, Art. no. 113982.
- [30] S. Xu, M. Wang, and X. Shi, "Hyperspectral imaging for high-resolution mapping of soil carbon fractions in intact paddy soil profiles with multivariate techniques and variable selection," *Geoderma*, vol. 370, 2020, Art. no. 114358.
- [31] E. J. Jones and A. B. McBratney, "What is digital soil morphometrics and where might it be going?," *Digital Soil Morphometrics*. Berlin, Germany: Springer, 2016, pp. 1–15.
- [32] S. Xu, M. Wang, X. Shi, Q. Yu, and Z. Zhang, "Integrating hyperspectral imaging with machine learning techniques for the high-resolution mapping of soil nitrogen fractions in soil profiles," *Sci. Total Environ.*, vol. 754, 2021, Art. no. 142135.
- [33] Y. Ogen, C. Neumann, S. Chabrilat, N. Goldshleger, and E. B. Dor, "Evaluating the detection limit of organic matter using point and imaging spectroscopy," *Geoderma*, vol. 321, pp. 100–109, 2018.
- [34] R. V. Rossel and T. Behrens, "Using data mining to model and interpret soil diffuse reflectance spectra," *Geoderma*, vol. 158, no. 1/2, pp. 46–54, 2010.
- [35] S. M. O'Rourke, I. Argentati, and N. M. Holden, "The effect of region of interest size on model calibration for soil organic carbon prediction from hyperspectral images of prepared soils," *J. Near Infrared Spectrosc.*, vol. 19, no. 3, pp. 161–170, 2011.
- [36] L. Wei, Y. Zhang, Z. Yuan, Z. Wang, F. Yin, and L. Cao, "Development of visible/near-infrared hyperspectral imaging for the prediction of total arsenic concentration in soil," *Appl. Sci.*, vol. 10, no. 8, 2020, Art. no. 2941.
- [37] S. Nawar and A. Mouazen, "On-line vis-NIR spectroscopy prediction of soil organic carbon using machine learning," *Soil Tillage Res.*, vol. 190, pp. 120–127, 2019.
- [38] N. K. Wijewardane, Y. Ge, S. Wills, and T. Loecke, "Prediction of soil carbon in the conterminous United States: Visible and near infrared reflectance spectroscopy analysis of the rapid carbon assessment project," *Soil Sci. Soc. Amer. J.*, vol. 80, no. 4, pp. 973–982, 2016.
- [39] E. Hobbey, M. Steffens, S. L. Bauke, and I. Kögel-Knabner, "Hotspots of soil organic carbon storage revealed by laboratory hyperspectral imaging," *Sci. Rep.*, vol. 8, no. 1, Sep. 2018, Art. no. 13900.
- [40] X. Meng, Y. Bao, Y. Wang, X. Zhang, and H. Liu, "An advanced soil organic carbon content prediction model via fused temporal-spatial-spectral (TSS) information based on machine learning and deep learning algorithms," *Remote Sens. Environ.*, vol. 280, Oct. 2022, Art. no. 113166.
- [41] J. Li et al., "Effect of reclamation on soil organic carbon pools in coastal areas of eastern China," *Front. Earth Sci.*, vol. 12, no. 2, pp. 339–348, Jun. 2018.
- [42] Y. Xu et al., "Effects of agricultural reclamation on soil physicochemical properties in the mid-eastern coastal area of China," *Land*, vol. 10, no. 2, pp. 142–153, 2021.
- [43] R. Zhang, "Land-forming history of the Huanghe River delta and coastal plain of north Jiangsu (in Chinese)," *Acta Geographica Sinica Beijing*, vol. 39, no. 2, pp. 173–184, 1984.
- [44] V. Allory et al., "Quantification of soil organic carbon stock in urban soils using visible and near infrared reflectance spectroscopy (VNIRS) in situ or in laboratory conditions," *Sci. Total Environ.*, vol. 686, pp. 764–773, 2019.
- [45] B. Minasny et al., "Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon," *Geoderma*, vol. 167, pp. 118–124, 2011.
- [46] R. K. H. Galvão, M. C. U. Araujo, G. E. José, M. J. C. Pontes, E. C. Silva, and T. C. B. Saldanha, "A method for calibration and validation subset partitioning," *Talanta*, vol. 67, no. 4, pp. 736–740, Oct. 2005.
- [47] S. Wold, H. Martens, and H. Wold, "The multivariate calibration problem in chemistry solved by the PLS method," in *Proc. Matrix Pencils, Conf. Held Pite Havsbud*, 2006, pp. 286–293.
- [48] B. D. Yonas, S. Velickov, D. Solomatine, and B. A. Michael, "Model induction with support vector machines: Introduction and applications," *J. Comput. Civil Eng.*, vol. 15, no. 3, pp. 208–216, Jul. 2001.

- [49] V. Bellon-Maurel, E. Fernandez-Ahumada, B. Palagos, J.-M. Roger, and A. McBratney, "Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy," *Trends Anal. Chem.*, vol. 29, no. 9, pp. 1073–1081, 2010.
- [50] P. Roudier, C. Hedley, and C. Ross, "Prediction of volumetric soil organic carbon from field-moist intact soil cores," *Eur. J. Soil Sci.*, vol. 66, no. 4, pp. 651–660, 2015.
- [51] M. García, S. Saatchi, S. Ustin, and H. Balzter, "Modelling forest canopy height by integrating airborne LiDAR samples with satellite Radar and multispectral imagery," *Int. J. Appl. Earth Observ. Geo-Inf.*, vol. 66, pp. 159–173, 2018.
- [52] G. Pan, L. Li, L. Wu, and X. Zhang, "Storage and sequestration potential of topsoil organic carbon in China's paddy soils," *Glob. Change Biol.*, vol. 10, no. 1, pp. 79–92, Jan. 2004.
- [53] H. Zhang et al., "Distribution and determinants of organic carbon and available nutrients in tropical paddy soils revealed by high-resolution sampling," *Agriculture, Eco-Syst. Environ.*, vol. 320, Oct. 2021, Art. no. 107580.
- [54] Z. Zhuo et al., "Soil organic carbon storage, distribution, and influencing factors at different depths in the dryland farming regions of Northeast and North China," *Catena*, vol. 210, Mar. 2022, Art. no. 105934.
- [55] P. Krishnan, J. D. Alexander, B. J. Butler, and J. W. Hummel, "Reflectance technique for predicting soil organic matter," *Soil Sci. Soc. Amer. J.*, vol. 44, no. 6, pp. 1282–1285, Nov. 1980.
- [56] D. J. Brown, K. D. Shepherd, M. G. Walsh, M. D. Mays, and T. G. Reinsch, "Global soil characterization with VNIR diffuse reflectance spectroscopy," *Geoderma*, vol. 132, no. 3, pp. 273–290, Jun. 2006.
- [57] R. A. V. Rossel and W. S. Hicks, "Soil organic carbon and its fractions estimated by visible–near infrared transfer functions," *Eur. J. Soil Sci.*, vol. 66, no. 3, pp. 438–450, May 2015.
- [58] T. L. Henderson, M. F. Baumgardner, D. P. Franzmeier, D. E. Stott, and D. C. Coster, "High dimensional reflectance analysis of soil organic matter," *Soil Sci. Soc. Amer. J.*, vol. 56, no. 3, pp. 865–872, May 1992.
- [59] C. L. S. Morgan, T. H. Waiser, D. J. Brown, and C. T. Hallmark, "Simulated in situ characterization of soil organic and inorganic carbon with visible near-infrared diffuse reflectance spectroscopy," *Geoderma*, vol. 151, no. 3, pp. 249–256, Jul. 2009.
- [60] J. Liu, J. Han, Y. Zhang, H. Wang, H. Kong, and L. Shi, "Prediction of soil organic carbon with different parent materials development using visible–near infrared spectroscopy," *Spectrochimica Acta A, Mol. Biomol. Spectrosc.*, vol. 204, pp. 33–39, Nov. 2018.
- [61] S. Priori, N. Bianconi, and E. A. C. Costantini, "Can γ -radiometrics predict soil textural data and stoniness in different parent materials? A comparison of two machine-learning methods," *Geoderma*, vol. 226/227, pp. 354–364, Aug. 2014.
- [62] S. Liu et al., "Estimation of plot-level soil carbon stocks in China's forests using intensive soil sampling," *Geoderma*, vol. 348, pp. 107–114, Aug. 2019.
- [63] Y.-J. Ma, T. Xie, and X.-Y. Li, "Spatial variation of soil organic carbon in the Qinghai Lake watershed, northeast Qinghai-Tibet Plateau," *Catena*, vol. 213, Jun. 2022, Art. no. 106187.
- [64] Y. Sun, J. Ma, and C. Li, "Content and densities of soil organic carbon in urban soil in different function districts of Kaifeng," *J. Geographical Sci.*, vol. 20, no. 1, pp. 148–156, Feb. 2010.
- [65] J.-F. Soussana et al., "Matching policy and science: Rationale for the '4 per 1000-soils for food security and climate' initiative," *Soil Tillage Res.*, vol. 188, pp. 3–15, 2019.
- [66] R. J. Huggett, "Soil chronosequences, soil development, and soil evolution: A critical review," *Catena*, vol. 32, no. 3, pp. 155–172, Jun. 1998.
- [67] P. R. Stevens and T. W. Walker, "The chronosequence concept and soil formation," *Quart. Rev. Biol.*, vol. 45, no. 4, pp. 333–350, Dec. 1970.
- [68] D. H. Yaalon, "Conceptual models in pedogenesis: Can soil-forming functions be solved?," *Geoderma*, vol. 14, no. 3, pp. 189–205, Oct. 1975.
- [69] L.-M. Huang, A. Thompson, G.-L. Zhang, L.-M. Chen, G.-Z. Han, and Z.-T. Gong, "The use of chronosequences in studies of paddy soil evolution: A review," *Geoderma*, vol. 237/238, pp. 199–210, Jan. 2015.
- [70] L.-M. Huang, A. Thompson, and G.-L. Zhang, "Long-term paddy cultivation significantly alters topsoil phosphorus transformation and degrades phosphorus sorption capacity," *Soil Tillage Res.*, vol. 142, pp. 32–41, Sep. 2014.
- [71] Y.-J. Luo, L.-M. Huang, and D.-G. Yuan, "Characteristics and controls of inorganic and organic phosphorus transformation during long-term paddy soil evolution," *Soil Tillage Res.*, vol. 224, Oct. 2022, Art. no. 105524.
- [72] L. Wissing, A. Kölbl, V. Vogelsang, J.-R. Fu, Z.-H. Cao, and I. Kögel-Knabner, "Organic carbon accumulation in a 2000-year chronosequence of paddy soil evolution," *Catena*, vol. 87, no. 3, pp. 376–385, Dec. 2011.
- [73] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2016, *arXiv:1409.0473*.
- [74] G. Ruan et al., "Improving wheat yield prediction integrating proximal sensing and weather data with machine learning," *Comput. Electron. Agriculture*, vol. 195, Apr. 2022, Art. no. 106852.
- [75] A. Vaswani et al., "Attention is all you need," *Adv. Neural Inf. Process. Syst.*, vol. 30, pp. 5998–6008, 2017.



Aosheng Wang received the bachelor's degree from the Nanjing University of Information Science and Technology, Nanjing, China, in 2019. He is currently working toward the master's degree from Nanjing University of Information Science and Technology, Nanjing, China, both in geography.

His research interests include mainly in soil remote sensing and soil evolution.



Guanghui Zheng received the Ph.D. degree in geography from the Nanjing University, Nanjing, China, in 2010.

He is currently an Associate Professor with the School of Geographic Sciences, Nanjing University of Information Science and Technology, Nanjing, China. His main research interests include soil remote sensing and soil evolution.



Danping Wei received the bachelor's degree from the Huaiyin Normal University, Huai'an, China, in 2019 and the master's degree from the Nanjing University of Information Science and Technology, Nanjing, China, in 2022, both in geography.

Her main research interests include soil remote sensing and soil evolution.



Caixia Jiao received the Ph.D. degree in geography from the Nanjing Normal University, Nanjing, China, in 2006.

She is currently an Associate Professor with the School of Geographic Sciences, Nanjing University of Information Science and Technology, Nanjing, China. Her main research interest includes soil remote sensing.



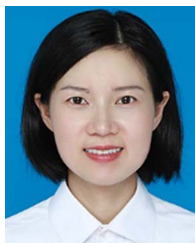
Xianli Xie received the Ph.D. degree in geography from the Nanjing Normal University, Nanjing, China, in 2004.

She is currently an Associate Researcher with the Nanjing Institute of Soil Science, Chinese Academy of Sciences, Beijing, China. Her main research interest includes soil remote sensing.



Giri Kattel received the Ph.D. degree in geography from the University College London, London, U.K, in 2004.

He is currently a Professor with the School of Geographic Sciences, Nanjing University of Information Science and Technology, Nanjing, China. His main research interest includes global environmental change.



Rong Zeng received the Ph.D. degree in resource environment and remote sensing information from the Nanjing Institute of Soil Science, Chinese Academy of Sciences, Beijing, China, in 2018.

She is currently a Lecturer with the School of Geographic Sciences, Nanjing University of Information Science and Technology, Nanjing, China. Her main research interests include soil hyperspectral remote sensing and digital soil mapping.



Ying Xiong is currently working toward the bachelor's degree in maths and applied mathematics with the Reading Academy, Nanjing University of Information Science and Technology, Nanjing, China.

Her research interests include global environmental change and soil remote sensing.



Chengyi Zhao received the Ph.D. degree in soil science from the China Agricultural University, Beijing, China, in 2002.

He is currently a Professor with the School of Geographic Sciences, Nanjing University of Information Science and Technology, Nanjing, China. His main research interests include intelligent remote sensing and artificial intelligence technology research and development, geographical Big Data and cloud platform technology application, and surface process simulation.



Dian Zhou is currently working toward the bachelor's degree in human geography and urban and rural planning from Nanjing University of Information Science and Technology, Nanjing, China.

Her research interests include coastal area's human-natural relationship, remote sensing, and human geography.