Computers and Electronics in Agriculture 185 (2021) 106144

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Improving estimation of soil organic matter content by combining Landsat 8 OLI images and environmental data: A case study in the river valley of the southern Qinghai-Tibet Plateau

Qing Yu [a,b], Tianci Yao [a,b], Hongwei Lu [a,*], Wei Feng [a,b], Yuxuan Xue [a,b], Binxiao Liu [a,b]

[a] Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Keywords:
Soil organic matter
Environmental factors
Landsat 8 OLI images
Qinghai-Tibet Plateau
Precision agriculture

ABSTRACT

The Qinghai-Tibet Plateau (QTP) is a typical ecologically fragile area. Once the surface vegetation degenerates, it may not be restored. This requires the development of soil organic matter (SOM) monitoring method without destroying the surface, so as to ensure the sustainable development of plateau agriculture. This work investigated the environmental factors that are significantly related to SOM content in the river valley of the southern QTP. These environmental factors include soil hydrothermal factors (soil moisture content and soil temperature), topographic factors (elevation and slope) and vegetation factor (NDVI). The original band reflectivity (OR) of Landsat 8 OLI images and the band reflectivity after the first-order derivative (FDR) and the second-order derivative (SDR) processing were combined with the above environmental factors to estimate SOM content. The results showed that the accuracy of the model was improved obviously by adding environmental factors. The estimation effect of back propagation neural network (BPNN) model was better than that of geographically weighted regression (GWR) model, partial least squares regression (PLSR) model and multivariable linear regression (MLR) model. GWR model can also meet the estimation requirements, while PLSR and MLR models cannot achieve effectively the estimation of SOM content. FDR-BPNN model considering environmental factors was the best model for estimating SOM content, with R^2 being 0.947, RMSEC being 4.701 g·kg^{-1} and MAEV being 5.485 g·kg^{-1}. Moreover, the model had the lowest uncertainty and the highest stability. This study will provide a good insight for the monitoring of SOM content in the future, and provide basic data support for the implementation of precision agriculture in the QTP.

1. Introduction

The Qinghai-Tibet Plateau (QTP) is one of the most fragile areas all around the world with high terrain and harsh natural conditions. The cultivated land area of the QTP only accounts for 0.5% of the total land area and is mainly distributed in the middle reaches of the Brahmaputra in the south of the QTP (Dai et al., 2014; Li et al., 2019). To protect the fragile environment, chemical fertilizer is strictly controlled (around only 0.1% of average unit usage in China) in the QTP. The growth of crops mainly depends on the supply of nutrients by soil organic matter (SOM). SOM contains a variety of nutrients for crop growth, which can provide various nutrients for crop growth directly after decomposition (Schmidt et al., 2011; Zhao et al., 2016; Zhu et al., 2007). Therefore, non-destructive monitoring of SOM in the QTP is of great theoretical and practical significance for the implementation of regional precision agriculture.

The traditional method for SOM quantification mainly relies on in situ sampling and chemical analysis. It is reliable but costly and time-consuming, as well as incapable of providing detailed spatial distribution (Caddeo et al., 2019; Chacón Iznaga et al., 2014; Gruba et al., 2015). Besides, the earth surface has to be damaged during the sampling process, and thus becomes inapplicable for fragile areas where ecosystem resistance and resilience are weak. Therefore, it is very important to carry out non-destructive surface monitoring for the QTP. How to estimate SOM content effectively, quickly and in large area is thus desired in these fragile areas such as the QTP. In recent years, remote sensing

* Corresponding author.
E-mail address: luhw@igsnrr.ac.cn (H. Lu).

https://doi.org/10.1016/j.compag.2021.106144
Received 16 October 2020; Received in revised form 29 March 2021; Accepted 30 March 2021
Available online 7 April 2021

technology provides a new way for nondestructive monitoring of soil internal components. Many scholars have used the near surface hyperspectral technology to estimate the SOM content, mainly using a portable spectrometer to collect the soil spectrum and establish various estimation models (Allory et al., 2019; Conforti et al., 2015; Dotto et al., 2017; Guo et al., 2017a,b; Nawar and Mouazen, 2018). At present, the commonly used estimation models mainly include linear models such as multiple linear regression (MLR) and partial least square regression (PLSR), as well as machine learning methods such as neural network with high accuracy in spectral inversion research (Caddeo et al., 2019; Dotto et al., 2017; Jin et al., 2016; Li et al., 2015). The accuracy of hyperspectral model for estimating SOM content can meet the requirements of precision agriculture, but the data processing and calculation process are complex and the observation scale is small. There are also a large number of scholars using remote sensing image to estimate SOM content (Jin et al., 2017; Lu et al., 2018; Mirzaee et al., 2016; Pouladi et al., 2019; Wang et al., 2018). This makes up for the limited data of large-scale soil mapping, but its accuracy is not enough to meet the mapping requirements.

The biggest limitation of SOM content estimation by remote sensing technology is the heterogeneous underlying surface. For example, the difference in topography, soil moisture and other characteristics will greatly limit the accuracy of SOM remote sensing inversion (Funes et al., 2019; Grunwald et al., 2017; Ondrasek et al., 2019). Soil is not homogeneous, and its spatial variation is affected by climate, topography, vegetation and other structural factors, as well as fertilization, farming measures, planting system and other random factors. A large number of studies have showed that there is a significant relationship between SOM content and environmental factors (Gruba et al., 2015; Grunwald et al., 2017; Muñoz and Kravchenko, 2011; Ondrasek et al., 2019), which provides a theoretical basis for quantitative prediction of soil properties by environmental attributes.

In this study, environmental factors were taken as co-factors and combined with Landsat 8 OLI images to carry out estimation research on SOM content. The agricultural area in the southern valley of the QTP was taken as the study area. The study area is characterized by high terrain, cold all year round, and fragile ecological environment. The objectives of this study were to: (1) compare the accuracy of estimation of SOM content with and without environmental factors, (2) compare the ability of SOM content estimation of Landsat 8 OLI image before and after derivative processing, (3) compare the advantages and disadvantages of back propagation neural network (BPNN), geographically weighted regression (GWR), PLSR, MLR model in estimating SOM content, and (4) determine the best spatial estimation method of SOM content and map the accurate SOM content spatial distribution in the study area.

2. Materials and methods

2.1. Study area and sample preparation

The study area (28°04′05″–31°03′57″N, 87°22′01″–95°18′50″E), as shown in Fig. 1, is located in the south of the QTP, belonging to the Brahmaputra valley. It has temperate and semi-arid plateau monsoon climate, with an annual average rainfall, sunshine and total radiation of 444.8 mm, about 3,000 h, and 191 kcal/cm^2, respectively. The climate condition is suitable for the growth of crops, and local agriculture develops steadily. As one of the main grain-producing bases in the QTP, currently, this region is cultivated with highland barley, wheat, oats and some other cold resistant crops.

Soil sampling was conducted from September 9 to September 19, 2018. A total of 83 sampling sites were selected considering the principles of randomness and representativeness based on the land-use status, topographic features, digital elevation model (DEM) and so on (Fig. 1). When sampling, soil moisture and temperature were synchronously measured using a TZS-pHW-4G multiparameter tachometer (TOP Instrument, Zhejiang, China), and the latitude, longitude of the sampling sites were also recorded by global positioning system (GPS). Then we used a wooden spatula to collect soil to a depth of 0–20 cm and removed coarse stones and plant debris and roots from the soil. All samples were put into self-sealing bags and brought back to the laboratory for analysis. Soil samples were air dried, and debris such as dead leaves, bricks, tiles and garbage were picked up. After filtering with a 0.25 mm nylon sieve, the samples were further analyzed by the concentrated sulfuric acid–potassium dichromate volumetric heating method. Reagent blanks, duplicate tests, and standard reference materials were applied for analytical quality control.

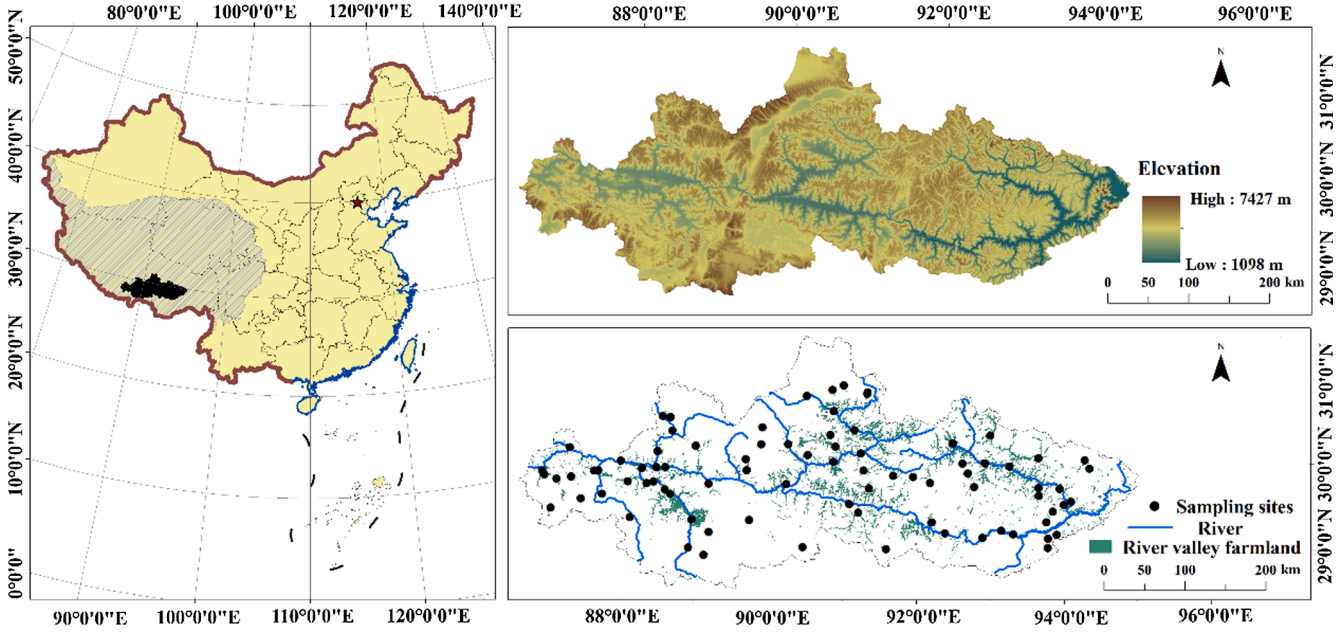


Fig. 1. The study area and spatial distribution of sampling sites.

2.2. Landsat 8 OLI image preprocessing and derivative processing

13 Landsat 8 OLI images were collected from the Geospatial Data Cloud website (http://www.gscloud.cn/) according to the sampling time and cloud amount of this study. In ENVI 5.3 (Exelis Visual Information Solutions, Boulder, Colorado, USA) software, images were preprocessed including radiation calibration, atmospheric correction, image stitching, and clipping. In the radiation calibration process, the useless cirrus wave band (B9) and quality inspection band (BQA) were automatically eliminated, and the two bands of thermal infrared and the panchromatic band were not involved in the calculation because of the different resolutions. Therefore, only 7 bands of images were left after radiometric calibration. In this study, bands 1–7 were selected for analysis. The specific introduction of the original band reflectivity (OR) is shown in Table 1.

After preprocessing, the first-order derivative of band reflectivity (FDR) and the second-order derivative of band reflectivity (SDR) were processed by ENVI 5.3 software, and the formulas are as follows.

$$FDR(\lambda_i) = \frac{R(\lambda_{i+1}) - R(\lambda_{i-1})}{2\Delta\lambda} \quad (1)$$

$$SDR(\lambda_i) = \frac{FDR(\lambda_{i+1}) - FDR(\lambda_{i-1})}{2\Delta\lambda} \quad (2)$$

whereλ_{i-1}, λ_i and λ_{i+1} are wavelengths, $\Delta\lambda$ represents the interval between two adjacent wavelengths,$FDR(\lambda_i)$ is the first derivative reflectivity of wavelength λ_i,$SDR(\lambda_i)$ is the second derivative reflectivity of wavelength λ_i.

2.3. Sources and processing of environmental data

The field surface reflectance is the result of the interaction of various surface factors. It is too ideal to establish SOM estimation model only by reflectance without considering the influence of other factors in the field, which will inevitably lead to the reduction of surveying and mapping accuracy. Therefore, in order to make the prediction results more in line with the actual situation, this study considered the impact of environmental factors, i.e., soil moisture, soil temperature, vegetation factors and terrain factors that have a significant impact on soil spectral reflectance.

2.3.1. Terrain factors

Elevation and slope data were selected to reflect the terrain characteristics of the study area. These two factors affect the hydrological and ecological processes such as surface runoff, plant growth and distribution, and then affect the spatial distribution of soil characteristics (Croft et al., 2012; Guo et al., 2013; Lu et al., 2018; Pouladi et al., 2019). Elevation and slope data were extracted from ASTER GDEM (http://www.gscloud.cn/) through the spatial analysis tool of ArcGIS 10.5 (ESRI, Redlands, California, USA) software. The extracted elevation is shown in Fig. 2(b) and the extracted slope is shown in Fig. 2(c).

2.3.2. Vegetation factor

Normalized difference vegetation index (NDVI) is the most commonly used vegetation index to reflect vegetation coverage, and is the best indicator of plant growth status and vegetation spatial distribution density (Pouladi et al., 2019; Wu et al., 2009). The greater the coverage of surface vegetation, the more abundant surface litter and leaves. The residual roots and leaves are decomposed by microorganisms, thus increasing the content of SOM. NDVI was extracted from Landsat 8 OLI image after preprocessing (Fig. 2(a)), and its formula is:

$$NDVI = (\text{float(Band5)} - \text{float(Band4)})/(\text{float(Band5)} + \text{float(Band4)}) \quad (3)$$

Table 1
Specific information of each band of Landsat 8 OLI image.

Band name	Bandwidth (um)	Spatial resolution (m)
Band 1	0.43–0.45	30
Band 2	0.45–0.51	30
Band 3	0.53–0.59	30
Band 4	0.64–0.67	30
Band 5	0.85–0.88	30
Band 6	1.57–1.65	30
Band 7	2.11–2.29	30

2.3.3. Soil hydrothermal factors

The transformation of SOM is closely related to soil temperature. High temperature can make organic matter decompose quickly, while low temperature can inhibit organic matter decomposition (Guo et al., 2017a,b; Wang et al., 2009; Wu et al., 2009). Soil temperature is of great significance to the humification process, the mineralization process and the nutrient supply of plants, which affects the growth of plants and the formation of soil. In addition, due to the influence of soil moisture content on soil microbial activities, too much or too little water will have a certain impact on SOM content. Based on the observed soil temperature and moisture content data for each sampling site, we used the IDW interpolation tool of ArcGIS software to draw the spatial distribution map of soil temperature (Fig. 2(d)) and soil moisture content (Fig. 2(e)). The spatial distribution maps of soil temperature and soil moisture content were drawn by using IDW interpolation tool of ArcGIS 10.5 software. The resampling tool was used to sample the spatial resolution to 30 m, and the distribution diagrams of soil temperature and soil moisture content after treatment are shown in Figs. 5 and 6.

2.4. Modeling approaches

In addition to the two commonly used regression models, MLR and PLSR, we also used the back propagation neural network (BPNN) model, and the geographically weighted regression (GWR) model. Both models are described below.

BPNN: BPNN is a mechanical learning model. It is a multilayer feedforward neural network trained according to error back propagation algorithm, and it is the most widely used neural network. The basic BP algorithm includes two processes: the forward propagation of signal and the back propagation of error. In forward propagation, the input signal acts on the output node through the hidden layer and generates the output signal through nonlinear transformation. If the actual output does not conform to the expected output, the error back propagation process will be transferred. Error backpropagation is to transmit the output error to the input layer by layer through the hidden layer, and to allocate the error to all cells in each layer, and the error signal obtained from each layer is used as the basis for adjusting the weight of each cell.

BPNN has the ability of arbitrary complex pattern classification and excellent multi-dimensional function mapping, which solves some problems that cannot be solved by simple perceptron. In terms of structure, BP network has input layer, hidden layer and output layer; in essence, BP algorithm takes the square of network error as the objective function and uses gradient descent method to calculate the minimum value of the objective function (Tian et al., 2013).

GWR: GWR is a new spatial analysis method, which detects the non-stationarity of spatial relationships by embedding the spatial structure into the linear regression model (Fotheringham et al., 1998). It is an extension of the general linear regression model, embedding the actual geographic location into the regression parameters:

$$y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \beta_2(u_i, v_i)x_{i2} + \cdots + \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \quad i = 1, 2, 3, \cdots, n, \quad (4)$$

where x_{i1}, x_{i2}, …, x_{ik} represent measured value of the variable k at the sampling site i, (u_i, v_i) represents the longitude and latitude of the sampling site i, ε_i represents the random error at the sampling site i, $\beta_k(u_i, v_i)$ represents the value of the kth regression parameter of sampling

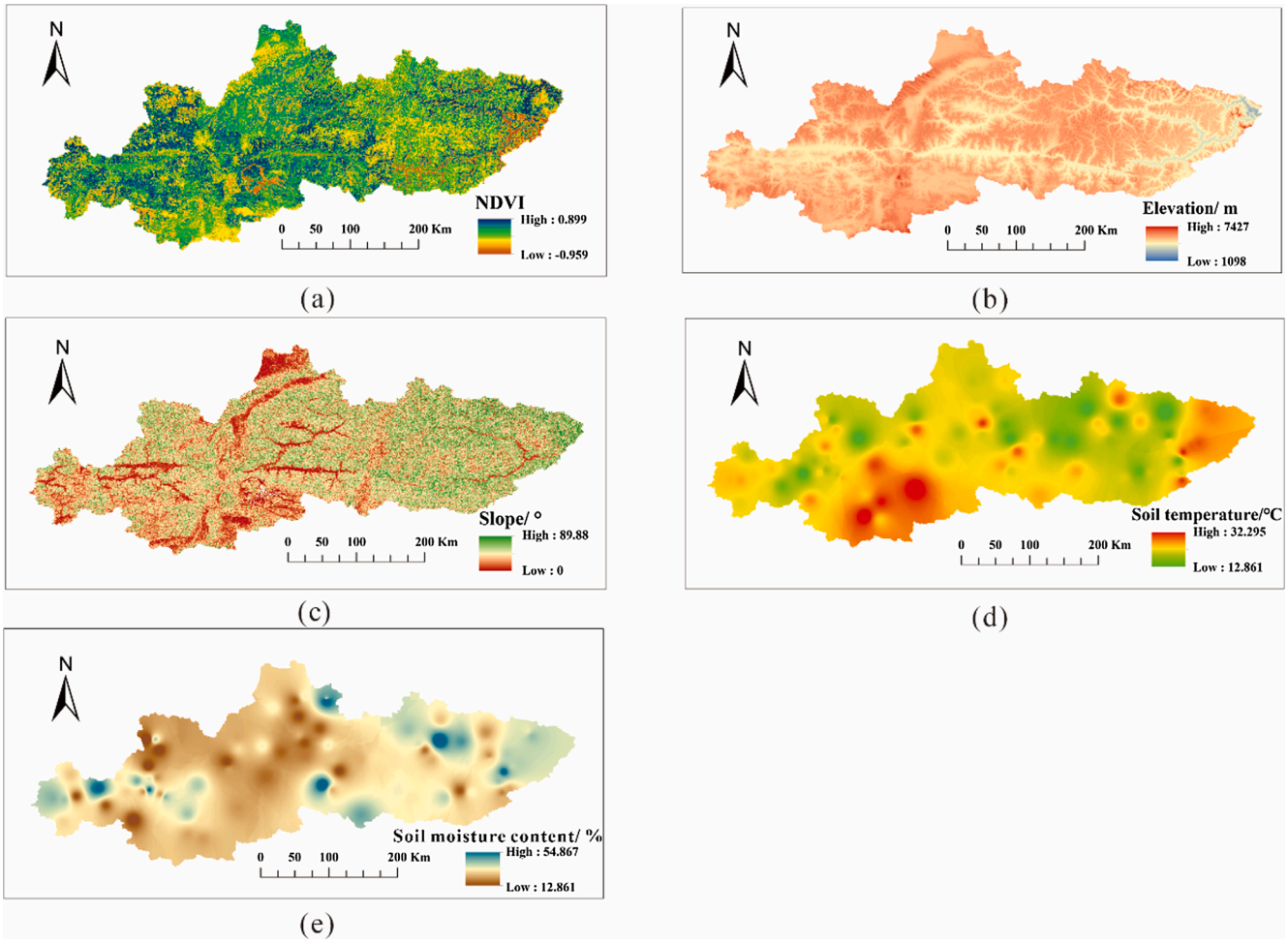


Fig. 2. Spatial distribution of environmental factors in the study area.

site *i*. The basic principle of geographically weighted regression is to use the method that the distribution characteristics of the research area have the quantitative relationship between two or more variables, and consider the local characteristics as the weight when processing the data (Chacón Iznaga et al., 2014). It is characterized by the assumption that the regression coefficient is the location function of the location of the observation point in the linear regression model, and the spatial characteristics of the data are included in the model.

2.5. Flowchart and model performance evaluation

The technical flow chart of this study is shown in Fig. 3, and the data used is shown in Table 2. Kennard-Stone (K-S) algorithm is a kind of clustering algorithm, which has been widely used in the division of data sets, especially suitable for dividing samples for spectral analysis (Nawar and Mouazen, 2018; Vohland et al., 2016). In this study, K-S algorithm was used to categorize 83 samples into calibration dataset and validation dataset according to the Euclidean Distance between samples. The calibration dataset included 55 samples and the validation dataset included 28 samples. Model determination coefficient (R^2), root mean square error of calibration (RMSEC), root mean square error of validation (RMSEV), mean absolute error of calibration (MAEC) and mean absolute error of validation (MAEV) were used to verify the accuracy of the model. The calculation formulas of RMSE and MAE are shown in formula (5) - (6). The larger R^2 is, the smaller the RMSEC is, the smaller the MAEC is, indicating that the modeling effect is better. The smaller the RMSEV is, the smaller the MAEV is, indicating that the model prediction accuracy is higher.

$$RMSE = \sqrt{\sum_{i=1}^{n}(\varepsilon_i - \varepsilon_i^{'})^2/n} \tag{5}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\varepsilon_i - \varepsilon_i^{'}\right| \tag{6}$$

where ε_i is the measured SOM content, $\varepsilon_i^{'}$ is the predicted SOM content, $\overline{\varepsilon i}$ is the mean value of the measured SOM content, and n is the number of sampling sites.

2.6. Model uncertainty analysis

In this study, $d-factor$ was used for uncertainty analysis of the model. The greater the value of $d-factor$, the greater the uncertainty of the model. The smaller the value of $d-factor$, the less the uncertainty of the model, and the more stable the model. The formula of $d-factor$ is as follows:

$$\overline{d_r} = \frac{1}{n}\sum_{i=1}^{n}(Y_{Ui} - Y_{Li}) \tag{7}$$

$$d-factor = \frac{\overline{d_r}}{\sigma_Y} \tag{8}$$

where $\overline{d_r}$ is the average distance between the indicative upper limit Y_{Ui} and the lower limit Y_{Li}. n is the number of samples, and σ_Y is the standard deviation of SOM measured content.

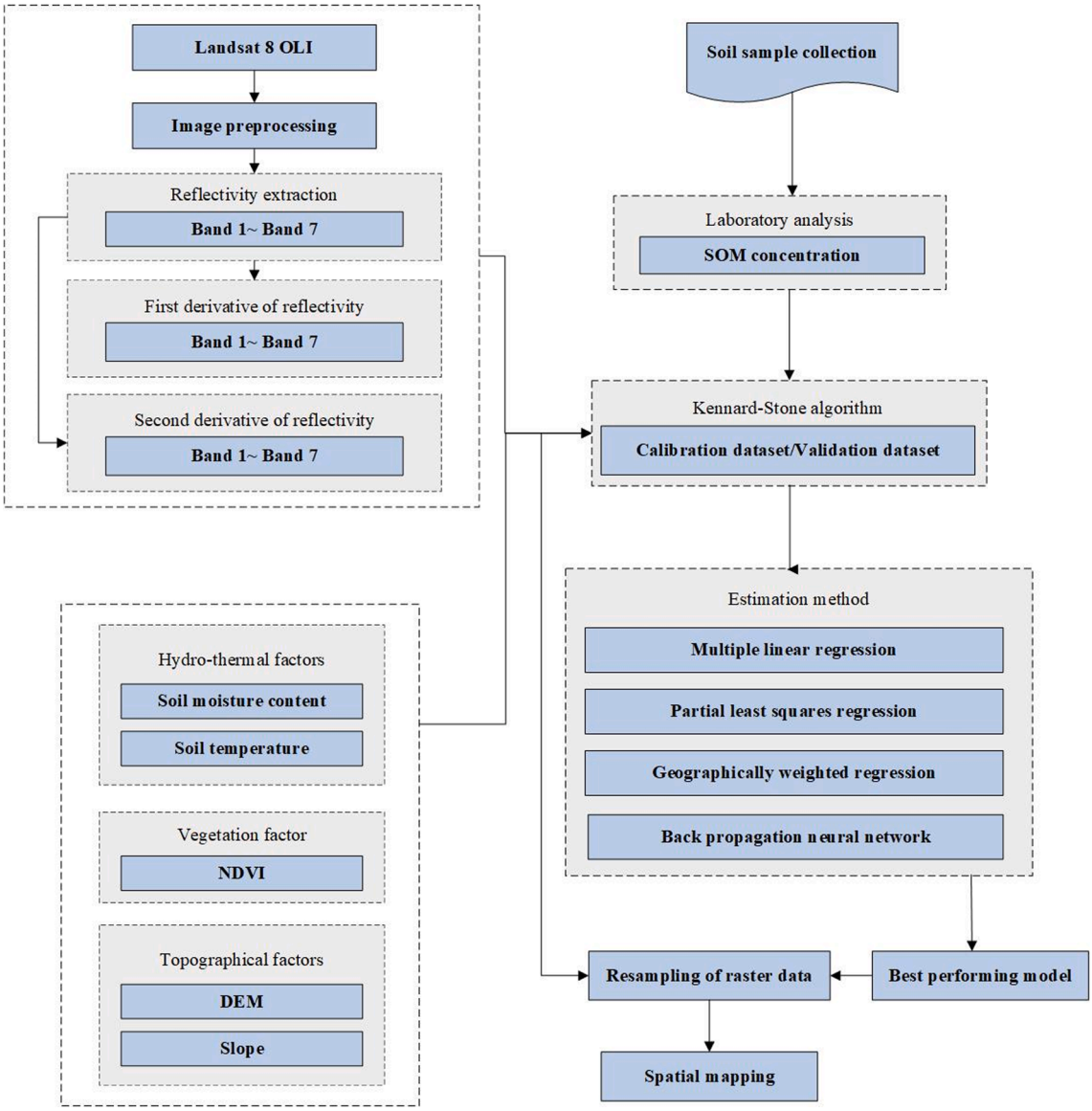


Fig. 3. Technical flow chart of this study.

Table 2
The data used for estimation in this study.

Data type	Variables	Spatial resolution (m)	Source
Landsat 8 OLI image	OR (1–7 band)	30	http://www.gscloud.cn/
	FDR (1–7 band)	30	http://www.gscloud.cn/
	SDR (1–7 band)	30	http://www.gscloud.cn/
Soil hydrothermal factors	Soil moisture Content	30	Field measurement, spatial interpolation, resampling
	Soil temperature	30	Field measurement, spatial interpolation, resampling
Terrain factors	Elevation	30	http://www.gscloud.cn/
	Slope	30	http://www.gscloud.cn/
Vegetation factor	NDVI	30	http://www.gscloud.cn/

3. Results

3.1. Statistical analysis of SOM content

The statistical characteristics of calibration and validation datasets were similar (Table 3), and the interval distribution was reasonable, which proved the reliability of the model to be built. Table 3 also showed the comparison of the SOM content in the study area with that of other typical regions in the world. The average content in the study area was 11.41 $g{\cdot}kg^{-1}$ higher than that in the whole southwest QTP, with the highest value being nearly twice as high. Compared with other areas of the plateau, the river valley has lower altitude, flat terrain, higher temperature and better precipitation conditions, which benefits the accumulation of SOM (Dai et al., 2014; Grunwald et al., 2017; Li et al., 2019). The average SOM content in the study area was 6.42 $g{\cdot}kg^{-1}$ lower than that in Fujian Province. The average SOM content was also higher than that in Beijing and Hubei Province, particularly in the Loess Plateau with severe soil erosion and poor soil nutrients where the average content was only 12.83 $g{\cdot}kg^{-1}$. Compared with other regions over the world, it was 8.11 $g{\cdot}kg^{-1}$ higher than Calabria of Italy, similar to Santa Catarina of Brazil, but much lower than Viborg of Denmark, Hawaii and

Table 3
Comparison of SOM content in the study area with other typical areas.

	Mean (g·kg^{-1})	Minimum (g·kg^{-1})	Maximum (g·kg^{-1})	Standard deviation (g·kg^{-1})	CV (%)	Reference
Entire dataset	34.31	11.21	96.37	20.70	60.33	—
Calibration dataset	34.34	11.55	96.37	20.57	59.90	—
Validation dataset	34.26	11.21	93.08	21.32	62.23	—
Southwest QTP, China	22.90	2.40	53.66	—	—	Dai et al., 2014
Beijing, China	14.88	0.49	53.05	5.16	35.00	Hu et al., 2014
Fujian Province, China	40.73	16.42	78.74	14.07	34.50	Liu et al. 2020
Hubei Province, China	27.30	6.10	71.8	12.0	43.81	Hong et al., 2018
Loess Plateau, China	12.83	6.10	20.20	2.46	19.18	Chen et al., 2019
Calabria, Italy	26.20	3.00	65.00	14.30	—	Conforti et al., 2015
Florida, America	57.29	3.36	854.47	—	—	Knox et al., 2015
Viborg, Denmark	77.92	12.59	307.73	—	—	Kuang et al., 2015
Hawaii, America	185.33	2.59	953.199	13.89	—	McDowell et al., 2012
Santa Catarina, Brazil	33.79	3.62	118.44	10.60	—	Dotto et al., 2017
Eastern Canada	42.00	23.00	73.00	—	26.00	Mabit and Bernard, 2010

Florida of United States and Eastern Canada.

3.2. Correlation analysis of band spectral reflectance and SOM content

The basis of remote sensing estimation is to find the relationship between remote sensing data and ground feature parameters. The images processed by the first and second derivatives of each band are shown in Figs. 4 and 5, respectively. In order to clarify the relationship between SOM content and the multispectral band, the correlation analysis between the multispectral reflectance data and the measured SOM content was used in this study (Table 4).

It could be seen that band 1–5 of the original image had a weak negative correlation with SOM content in the study area, while band 6 and band 7 had extremely significant correlation with SOM content, and the correlation coefficients of 0.366 and 0.342. After the first derivative transformation of reflectivity, it was found that band 2 and band 5 showed significant correlation with SOM content. Band 6 showed significant correlation with SOM content and band 7 showed extremely significant correlation with SOM content. Only band 6 and band 7 had correlation with SOM content after the second derivative transformation of reflectivity, and other bands had weak correlation with SOM content. It can be found that the correlation between the band reflectance processed by the first-order derivative and soil organic matter content was obviously enhanced. However, the band reflectance processed by the second-order derivative did not show a good effect, which may be the lack of spectral information after the second-order derivative transformation.

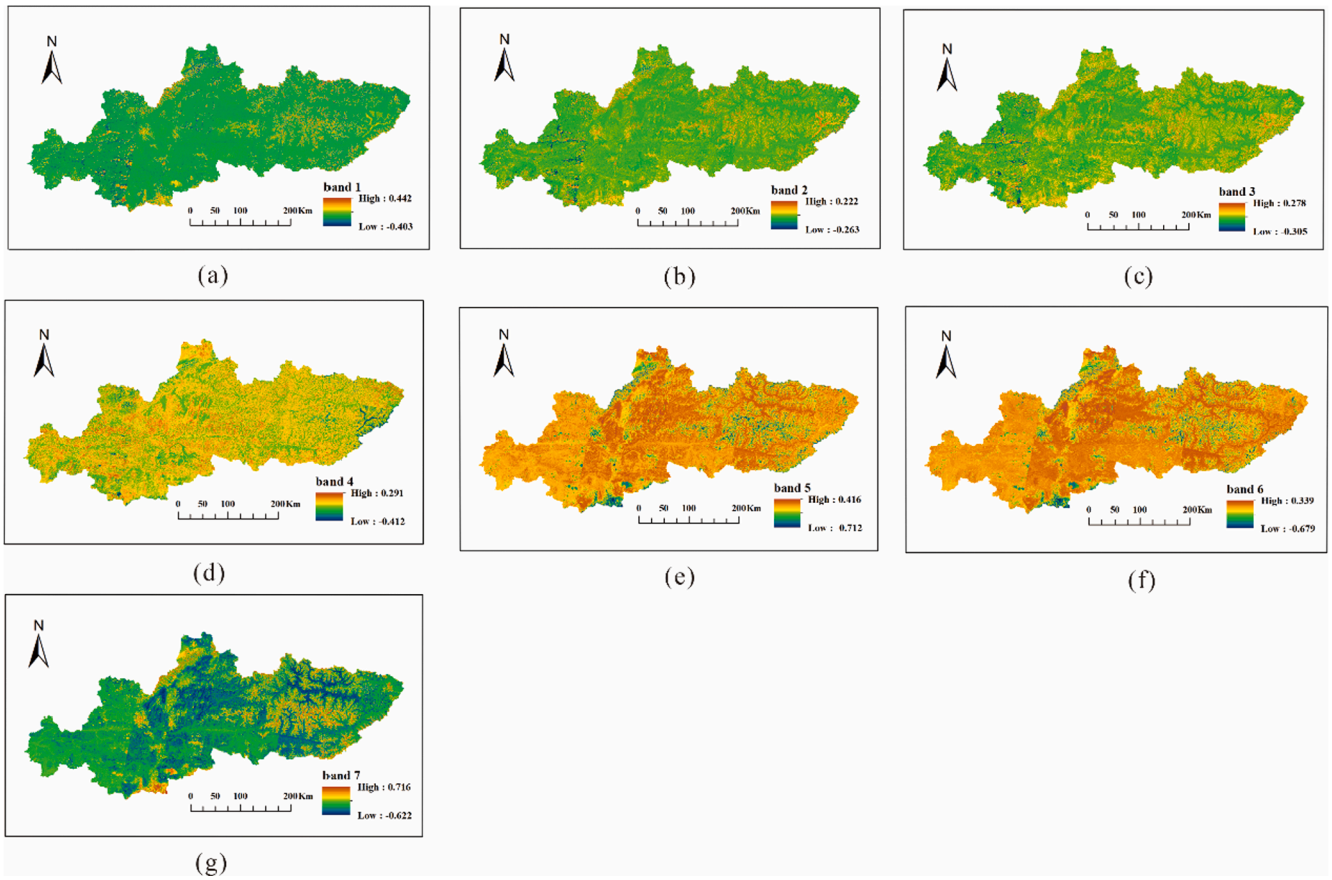


Fig. 4. The first-derivative of reflectivity in each band.

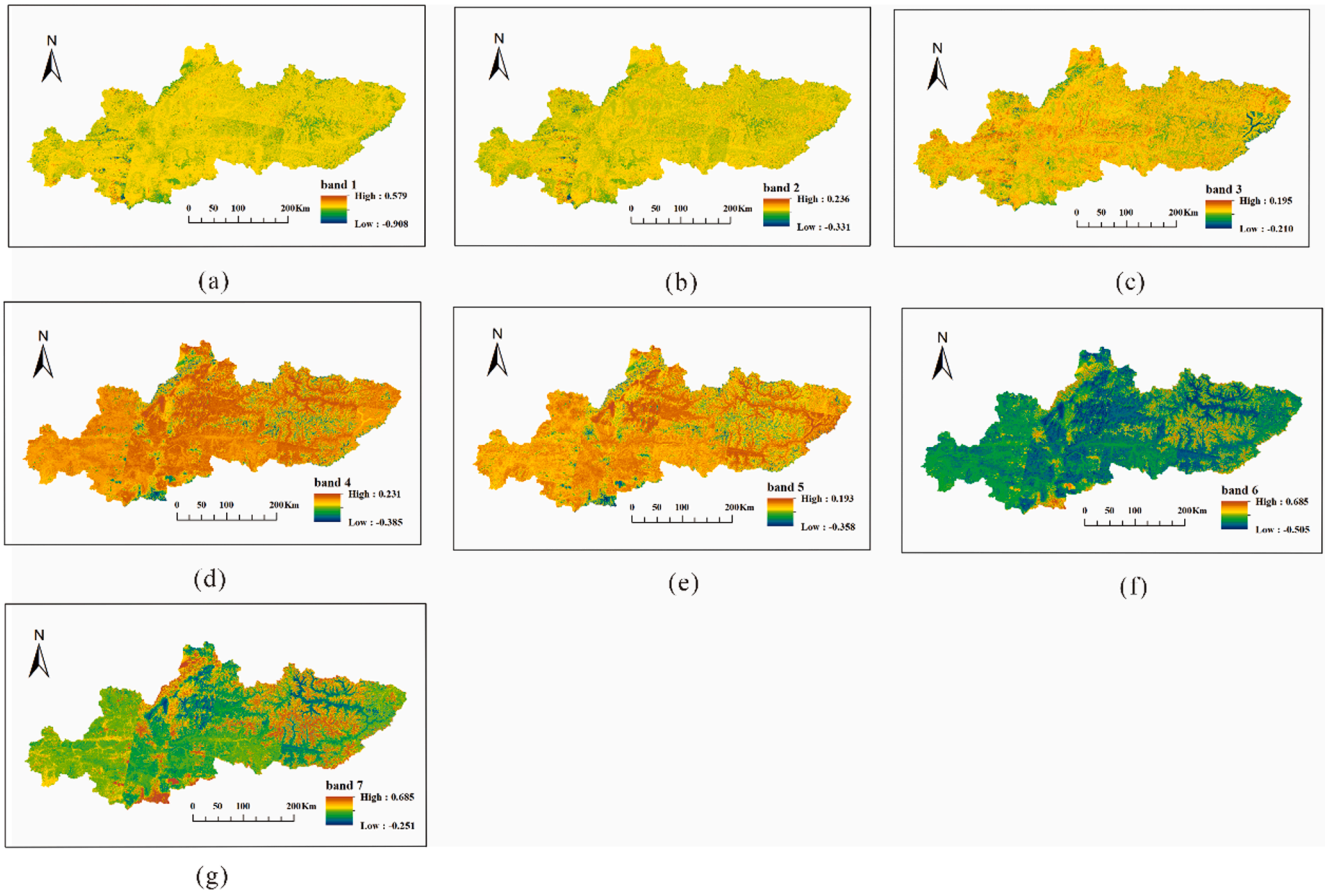


Fig. 5. The second-derivative of reflectivity in each band.

Table 4
Correlation between the reflectance of each band and SOM content.

Pretreatment method	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7
OR	−0.156	−0.137	−0.093	−0.056	−0.065	0.366**	−0.342**
FDR	0.001	0.220*	0.208	0.067	0.256*	0.247*	−0.313**
SDR	0.160	0.167	−0.054	0.182	0.198	−0.294**	−0.296**

** Correlation is significant at the 0.01 level.
* Correlation is significant at the 0.05 level

3.3. Descriptive statistics of environmental factors in the study area

Environmental factors have a certain impact on the spatial distribution of SOM content, but the impact is different among different environmental factors. Table 5 is the descriptive statistics of environmental factors in the study area. The elevation of the sample sites was between 2940 m and 5680 m, and the average elevation was 4139.53 m, suggesting the terrain of the study area was very high. The maximum slope of the study area was 43.39° and the minimum slope was 1.01°.

Table 5
Statistics of environmental factors in the study area.

Environmental factor	Minimum	Maximum	Mean	Standard deviation	CV (%)
Soil temperature (℃)	9.00	32.30	21.14	5.16	24.41
Soil moisture content (%)	12.80	56.50	29.36	8.86	30.18
NDVI	−0.55	0.76	0.19	0.24	126.31
Slope (°)	1.01	49.39	13.52	11.38	84.17
Elevation (m)	2940.00	5680.00	4139.53	627.92	15.169

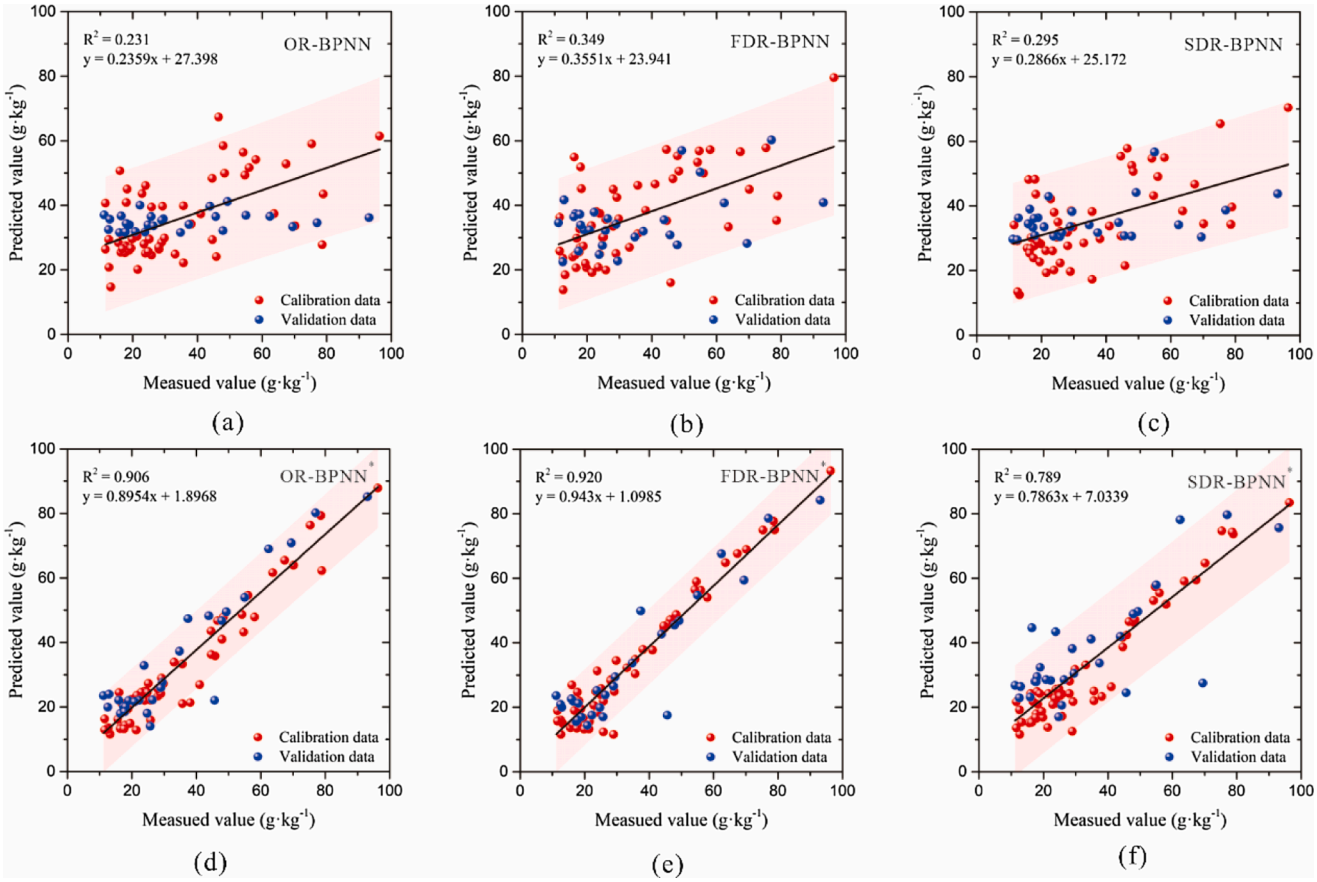
NDVI was between −0.55 and 0.76, with an average value of 0.19, and its CV was 126.31%, indicating an uneven vegetation distribution in the study area. The change of soil moisture content in each sample site was also large, ranging from 12.80% to 56.50%, with an average of 29.36%. The average soil temperature was 21.14 °C. Because the study area is located in the valley area, the soil temperature is high, which is suitable for the growth of crops.

3.4. Model calibration and validation

The calibration and validation results of various estimation models are shown in Table 6. For calibration model, the estimation accuracy of the spectral model with environmental factors was much higher than that with multispectral bands only. It was obvious that the R^2 of BPNN model built by adding environmental factors was greater than 0.9, which was about 0.5 higher than that of BPNN model built only by multi spectral bands. The FDR-BPNN model with environmental factors was the best one. Its R^2 was 0.947, RMSEC was 4.701 $g{\cdot}kg^{-1}$, MAEC was 3.254 $g{\cdot}kg^{-1}$, R^2 was higher than all other models, and RMSEC and MAEC were lower than other models, which showed that the FDR-BPNN model with environmental factors had the best estimation effect. In

Table 6
The result statistics of calibration models and validation models.

Method	Model		Calibration dataset			Validation dataset	
			R^2	RMSEC ($g \cdot kg^{-1}$)	MAEC ($g \cdot kg^{-1}$)	RMSEV ($g \cdot kg^{-1}$)	MAEV ($g \cdot kg^{-1}$)
Multispectral band	OR	MLR	0.323	16.775	12.590	18.460	13.902
		PLSR	0.269	17.424	13.544	20.762	16.751
		GWR	0.507	14.305	11.110	16.987	12.895
		BPNN	0.327	16.725	12.400	20.474	16.566
	FDR	MLR	0.317	16.853	12.668	18.549	14.179
		PLSR	0.255	17.597	14.004	19.619	14.833
		GWR	0.439	15.258	11.712	16.714	12.267
		BPNN	0.417	15.557	11.494	18.727	14.906
	SDR	MLR	0.286	17.225	13.195	19.942	15.800
		PLSR	0.225	17.949	13.934	20.636	16.453
		GWR	0.302	17.030	12.883	20.782	15.039
		BPNN	0.387	15.959	12.180	19.633	15.835
Multispectral band added environmental data	OR	MLR	0.411	15.646	11.967	17.873	13.421
		PLSR	0.310	16.937	12.917	19.463	15.030
		GWR	0.666	11.781	8.862	13.032	9.578
		BPNN	0.909	6.119	4.332	7.293	5.236
	FDR	MLR	0.410	15.644	11.719	17.888	13.358
		PLSR	0.311	16.916	13.201	18.972	14.327
		GWR	0.540	13.815	10.221	14.879	11.190
		BPNN	0.947	4.701	3.254	7.832	5.485
	SDR	MLR	0.392	15.901	12.459	18.707	14.044
		PLSR	0.255	17.596	13.622	19.639	15.212
		GWR	0.385	15.984	12.412	22.837	15.915
		BPNN	0.917	6.057	4.345	13.884	10.511



Red shadow represents the 95% confidence interval

* Indicates the estimation model adding environmental data

Fig. 6. Scatter plots of measured and predicted SOM contents of various BPNN models. (Red shadow represents the 95% confidence interval * Indicates the estimation model adding environmental data). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

addition to the strong prediction ability of BPNN model with environmental factors, OR-GWR model and FDR-GWR model with environmental factors also showed good prediction effect. Their R^2 were 0.666 and 0.540, and RMSEC were 13.815 and 11.781 g·kg^{-1}, respectively. However, the estimation ability of MLR model and PLSR model was poor regardless of whether environmental factors were added or not, and regardless of which spectral processing methods were used. The validation results were consistent with the calibration results, i.e., the validation model with good calibration performance also showed high accuracy. From the aspect of verification model, the FDR-BPNN model with environmental factors was still the best one. All kinds of BPNN models with environmental factors had excellent prediction ability. For GWR models with environmental factors, the prediction results of the models based on OR and FDR are better, but the models based on SDR do not show good estimation results. The prediction ability of MLR model and PLSR model is very poor.

Taking the BPNN modeling method as an example, the scatter diagrams of predicted and measured values of various BPNN models are drawn in Fig. 6. It was obvious that the estimation effect after adding environmental factors was much better than that without environmental factors, and the predicted value was in good agreement with the measured value (close to y = x line, Fig. 6(d)–(f)). In Fig. 6(e) the predicted value and measured value had the best fitting effect, indicating the FDR-BPNN model with environmental factors added showed the strongest estimation ability.

In general, no matter which model, the estimation ability of the model became stronger after adding environmental factors. For the spectral processing method, the first derivative estimation was the best, followed by the estimation of the original spectral band. The effect of spectral bands after second-order transformation was the worst, which was not as good as the estimation effect of original spectral band. The BPNN model had the highest estimation ability, followed by GWR model, while the MLR and PLSR models established had poor performance and could not achieve accurate estimation of SOM content. The FDR-BPNN model with environmental factors was the best model for estimating organic matter content. Fig. 7 shows the fitting diagram of predicted and measured SOM content of FDR-BPNN model with environmental factors added. It could be seen that the fitting effect was good and the prediction of extreme value was also good.

3.5. Selection of optimal model and spatial mapping of SOM content

In this study, we calculated d-factor for various spectral calibration models and validation models, and carried out uncertainty analysis. The uncertainty analysis results of each model are shown in Table 7. Whether it was the original spectral band, the band after the first-order derivative processing, or the band after the second-order derivative processing, the average *d –factor* value of the BPNN model built with environmental factors was small. The *d –factor* values of the established calibration models were all around 1.0. The *d –factor* values of these models were much smaller than those of other estimation models. However, for the validation model, the d-factor of BPNN model processed by SDR was greater than 2, which led to high uncertainty and instability of the model. The OR-BPNN model and FDR-BPNN model built by adding environmental factors had less uncertainty and higher model stability. The *d –factor* values of the calibration model were 0.991 and 0.943, and the *d –factor* values of the validation model were 1.465 and 1.539, respectively.

According to the results of model error analysis and uncertainty analysis, the FDR-BPNN model with environmental factors had high fitting accuracy and stability, which could be used to estimate SOM content. Therefore, we used the FDR-BPNN model with environmental factors to estimate SOM content of the entire study area. The SOM content distribution map is shown in Fig. 8. It could be seen that SOM content in the study area was between 2.014 and 132.732 g·kg^{-1}. The SOM content in the middle and southwest of the study area was higher than that in other areas.

Table 7
Uncertainty measuring parameter of various models.

Model		Multispectral band		Multispectral band added environmental data	
		Calibration model	Validation model	Calibration model	Validation model
CR	MLR	1.950	1.832	2.063	1.968
	PLSR	1.850	1.559	1.928	1.508
	GWR	1.912	1.952	1.745	1.779
	BPNN	2.046	1.538	0.991	1.465
FDR	MLR	1.939	1.809	2.058	1.968
	PLSR	1.817	1.440	1.657	1.687
	GWR	1.747	2.039	2.039	2.439
	BPNN	2.125	1.670	0.943	1.539
SDR	MLR	1.884	1.276	2.046	1.788
	PLSR	1.740	1.730	1.817	1.430
	GWR	1.794	2.620	2.354	3.692
	BPNN	2.026	1.146	1.037	2.269

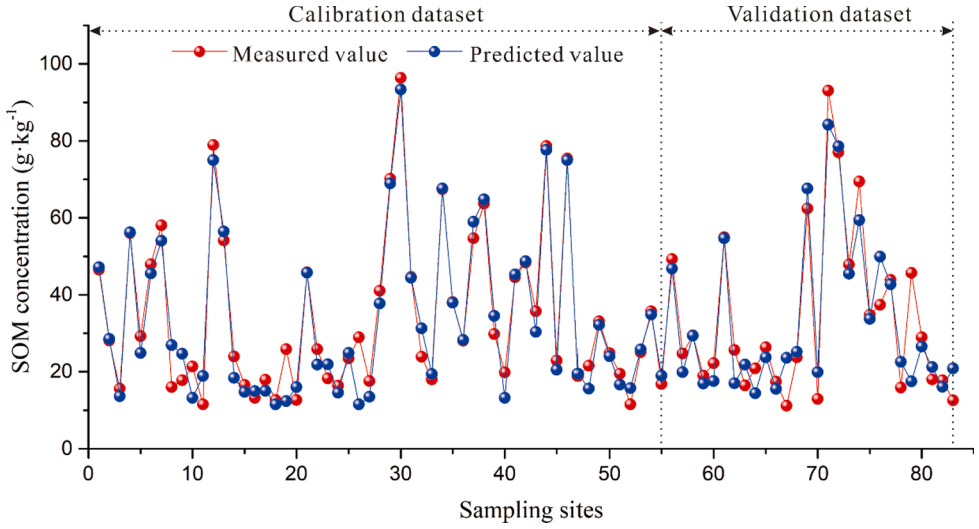


Fig. 7. The fitting diagram of the predicted and measured SOM contents of the optimal model.

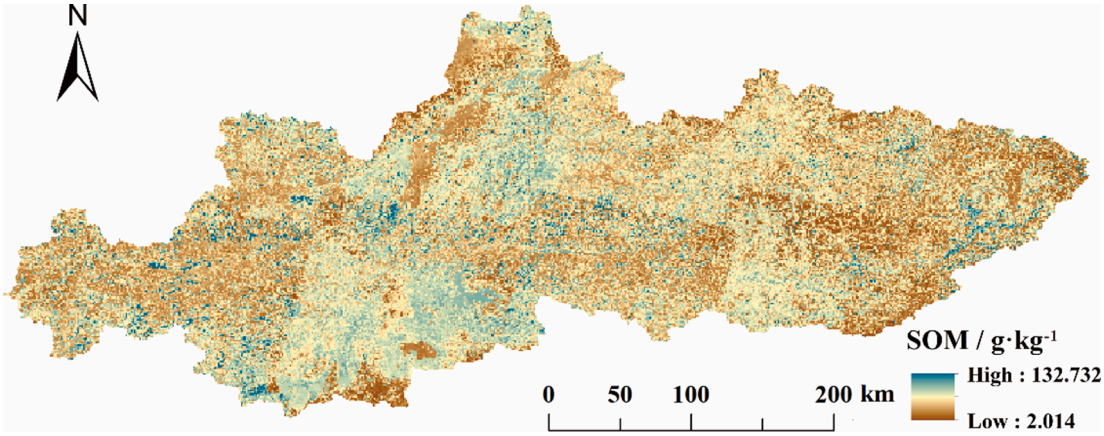


Fig. 8. The spatial distribution of SOM content estimated by the optimal model.

4. Discussion

In this study, derivative processing was performed on the band reflectance of Landsat 8 OLI images. Derivative processing was rarely used in Landsat images, but it was the most widely used method in SOM estimation of measured near-ground hyperspectral. A large number of studies had showed that the first-order derivative processing could smooth the influence of background interference as well as make the profile of spectral data clearer (Dai et al., 2014; Nawar et al., 2016; Yu et al., 2016). Our study showed the estimation ability of FDR was the strongest, while the estimation effect of SDR was not good and even smaller than OR. After derivative processing, the image changed obviously, and the image noise with the first derivative processing was reduced, which highlighted the hidden information of SOM in the image. However, the image processed by the second derivative could suppress the information of SOM and could not accurately estimate SOM content. Among these modeling methods, the BPNN model had the best estimation effect on the SOM content, could better explain the nonlinear relationship between variables. GWR model took the spatial heterogeneity into account. Therefore, GWR model could also meet the requirement of estimating SOM content. Among the 24 estimation models used in this study, FDR-BPNN model with environmental factors had the strongest estimation ability.

The SOM content is affected by many environmental factors. Soil moisture content, soil temperature and NDVI are the main factors affecting the SOM content, and elevation and slope also significantly affect the distribution of SOM content. Lu et al. (2018) used DEM as an auxiliary variable combined with Landsat image to estimate the SOM content in a hickory plantation region, showing a good estimation effect. Jaber and Al-Qinna (2017) confirmed that soil moisture, temperature, and other variables could affect the accumulation of soil organic carbon when studying SOM estimation in the Zarqa Basin in Jordan. Wang et al. (2010) found that models with elevation, slope, and vegetation index would be more accurate when using hyperspectral images to map the SOM content of land degraded areas. In our study, we used topographic factors, hydrothermal factors, and vegetation factors to analyze and got the same results. As shown in Fig. 2, it could be seen that the central and southwestern parts of the study area had low elevation, small slopes, and high vegetation coverage. This was consistent with the distribution of areas with high contents of SOM. The above confirmed that environmental factors had certain influence on the spatial distribution of SOM content, although this effect showed differences among different environmental factors.

The factors affecting the spatial distribution of SOM content are complex. In addition to environmental factors, other factors such as human factors and land use status should be also considered in the next research, so as to further improve the accuracy of remote sensing monitoring of SOM content.

5. Conclusions

In the field environment, the error of remote sensing image in the estimation of soil internal components is large due to the complex surface properties. To address this defect, soil hydrothermal factors, terrain factors, and vegetation factors that are significantly related to SOM content were collected in the river valley of the southern QTP. We presented a methodology to improve the estimation of SOM content by using Landsat 8 OLI images combined with these environmental factors. The band reflectance of Landsat 8 OLI images was processed by first-derivative and second-derivative. The modeling methods of MLR, PLSR, GWR and BPNN were used, and 24 models were established with or without environmental factors. It was clear from the results that the accuracy of various models to estimate SOM content had been improved by adding environmental factors. The estimation effect of the FDR processing was the best, but the SDR processing was poor. Various BPNN models built by adding environmental factors had strong estimation ability. Among them, the FDR-BPNN model with environmental factors was the best estimation model, with R^2 being 0.947, RMSEC being 4.701 $g{\cdot}kg^{-1}$ and MAEV being 5.485 $g{\cdot}kg^{-1}$, and the stability of the model was high. These outcomes will provide a scientific reference for satellite remote sensing images combined with environmental factors to estimate SOM content, and is of great significance for precision agriculture and sustainable agricultural development in the QTP.

Research involving human participants and/or animals

This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent

This article does not contain any studies with human participants.

CRediT authorship contribution statement

Qing Yu: Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing. **Tianci Yao:** Methodology, Investigation, Writing - original draft, Writing - review & editing. **Hongwei Lu:** Conceptualization, Writing - review & editing, Funding acquisition. **Wei Feng:** Methodology, Software, Writing - original draft. **Yuxuan Xue:** Investigation, Methodology. **Binxiao Liu:** Methodology, Software.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences [Grant No. XDA20040301], the Second Tibetan Plateau Scientific Expedition and Research Program (STEP) [Grant No. 2019QZKK1003] and the National Key Research and Development Program of China [Grant No.2019YFC0507801].

References

Allory, V., Cambou, A., Moulin, P., Schwartz, C., Cannavo, P., Vidal-Beaudet, L., Barthès, B.G., 2019. 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. 686, 764–773. https://doi.org/10.1016/j.scitotenv.2019.05.192.

Caddeo, A., Marras, S., Sallustio, L., Spano, D., Sirca, C., 2019. Soil organic carbon in Italian forests and agroecosystems: Estimating current stock and future changes with a spatial modelling approach. Agric. For. Meteorol. 278, 107654. https://doi.org/10.1016/j.agrformet.2019.107654.

Chacón Iznaga, A., Rodríguez Orozco, M., Aguila Alcantara, E., Carral Pairol, M., Díaz Sicilia, Y.E., de Baerdemaeker, J., Saeys, W., 2014. Vis/NIR spectroscopic measurement of selected soil fertility parameters of Cuban agricultural Cambisols. Biosyst. Eng. 125, 105–121. https://doi.org/10.1016/j.biosystemseng.2014.06.018.

Chen, D., Chang, N., Xiao, J., Zhou, Q., Wu, W., 2019. Mapping dynamics of soil organic matter in croplands with MODIS data and machine learning algorithms. Sci. Total Environ. 669, 844–855. https://doi.org/10.1016/j.scitotenv.2019.03.151.

Conforti, M., Castrignanò, A., Robustelli, G., Scarciglia, F., Stelluti, M., Buttafuoco, G., 2015. Laboratory-based Vis-NIR spectroscopy and partial least square regression with spatially correlated errors for predicting spatial variation of soil organic matter content. Catena 124, 60–67. https://doi.org/10.1016/j.catena.2014.09.004.

Croft, H., Kuhn, N.J., Anderson, K., 2012. On the use of remote sensing techniques for monitoring spatio-temporal soil organic carbon dynamics in agricultural systems. Catena 94, 64–74. https://doi.org/10.1016/j.catena.2012.01.001.

Dai, F., Zhou, Q., Lv, Z., Wang, X., Liu, G., 2014. Spatial prediction of soil organic matter content integrating artificial neural network and ordinary kriging in Tibetan Plateau. Ecol. Indic. 45, 184–194. https://doi.org/10.1016/j.ecolind.2014.04.003.

Dotto, A.C., Dalmolin, R.S.D., Grunwald, S., ten Caten, A., Pereira Filho, W., 2017. Two preprocessing techniques to reduce model covariables in soil property predictions by Vis-NIR spectroscopy. Soil Tillage Res. 172, 59–68. https://doi.org/10.1016/j.still.2017.05.008.

Fotheringham, A.S., Charlton, M.E., Brunsdon, C., 1998. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. Environ. Plan. A 30, 1905–1927. https://doi.org/10.1068/a301905.

Funes, I., Savé, R., Rovira, P., Molowny-Horas, R., Alcañiz, J.M., Ascaso, E., Herms, I., Herrero, C., Boixadera, J., Vayreda, J., 2019. Agricultural soil organic carbon stocks in the north-eastern Iberian Peninsula: Drivers and spatial variability. Sci. Total Environ. 668, 283–294. https://doi.org/10.1016/j.scitotenv.2019.02.317.

Gruba, P., Socha, J., Błońska, E., Lasota, J., 2015. Effect of variable soil texture, metal saturation of soil organic matter (SOM) and tree species composition on spatial distribution of SOM in forest soils in Poland. Sci. Total Environ. 521–522, 90–100. https://doi.org/10.1016/j.scitotenv.2015.03.100.

Grunwald, D., Kaiser, M., Junker, S., Marhan, S., Piepho, H.P., Poll, C., Bamminger, C., Ludwig, B., 2017. Influence of elevated soil temperature and biochar application on organic matter associated with aggregate-size and density fractions in an arable soil. Agric. Ecosyst. Environ. 241, 79–87. https://doi.org/10.1016/j.agee.2017.02.029.

Guo, L., Zhao, C., Zhang, H., Chen, Y., Linderman, M., Zhang, Q., Liu, Y., 2017a. Comparisons of spatial and non-spatial models for predicting soil carbon content based on visible and near-infrared spectral technology. Geoderma 285, 280–292. https://doi.org/10.1016/j.geoderma.2016.10.010.

Guo, N., Shi, X., Zhao, Y., Xu, S., Wang, M., Zhang, G., Wu, J., Huang, B., Kong, C., 2017b. Environmental and anthropogenic factors driving changes in paddy soil organic matter: A case study in the middle and lower Yangtze river plain of China. Pedosphere 27, 926–937. https://doi.org/10.1016/S1002-0160(17)60383-7.

Guo, P.T., Wu, W., Sheng, Q.K., Li, M.F., Liu, H. Bin, Wang, Z.Y., 2013. Prediction of soil organic matter using artificial neural network and topographic indicators in hilly areas. Nutr. Cycl. Agroecosyst. 95, 333–344. https://doi.org/10.1007/s10705-013-9566-9.

Hong, Y., Chen, S., Zhang, Y., Chen, Y., Yu, L., Liu, Yanfang, Liu, Yaolin, Cheng, H., Liu, Yi, 2018. Rapid identification of soil organic matter level via visible and near-infrared spectroscopy: Effects of two-dimensional correlation coefficient and extreme learning machine. Sci. Total Environ. 644, 1232–1243. https://doi.org/10.1016/j.scitotenv.2018.06.319.

Hu, K., Wang, S., Li, H., Huang, F., Li, B., 2014. Spatial scaling effects on variability of soil organic matter and total nitrogen in suburban Beijing. Geoderma 226–227, 54–63. https://doi.org/10.1016/j.geoderma.2014.03.001.

Jaber, S.M., Al-Qinna, M.I., 2017. TM-based SOC models augmented by auxiliary data for carbon crediting programs in semi-arid environments. Photogramm. Eng. Remote Sens. 83, 447–457. https://doi.org/10.14358/PERS.83.6.447.

Jin, X., Du, J., Liu, H., Wang, Z., Song, K., 2016. Remote estimation of soil organic matter content in the Sanjiang Plain, Northest China: The optimal band algorithm versus the GRA-ANN model. Agric. For. Meteorol. 218–219, 250–260. https://doi.org/10.1016/j.agrformet.2015.12.062.

Jin, X., Song, K., Du, J., Liu, H., Wen, Z., 2017. Comparison of different satellite bands and vegetation indices for estimation of soil organic matter based on simulated spectral configuration. Agric. For. Meteorol. 244–245, 57–71. https://doi.org/10.1016/j.agrformet.2017.05.018.

Knox, N.M., Grunwald, S., McDowell, M.L., Bruland, G.L., Myers, D.B., Harris, W.G., 2015. Modelling soil carbon fractions with visible near-infrared (VNIR) and mid-infrared (MIR) spectroscopy. Geoderma 239–240, 229–239. https://doi.org/10.1016/j.geoderma.2014.10.019.

Kuang, B., Tekin, Y., Mouazen, A.M., 2015. 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. 146, 243–252. https://doi.org/10.1016/j.still.2014.11.002.

Li, S., He, F., Zhang, X., Zhou, T., 2019. Evaluation of global historical land use scenarios based on regional datasets on the Qinghai-Tibet Area. Sci. Total Environ. 657, 1615–1628. https://doi.org/10.1016/j.scitotenv.2018.12.136.

Li, S., Shi, Z., Chen, S., Ji, W., Zhou, L., Yu, W., Webster, R., 2015. In situ measurements of organic carbon in soil profiles using vis-NIR spectroscopy on the qinghai-tibet plateau. Environ. Sci. Technol. 49, 4980–4987. https://doi.org/10.1021/es504272x.

Liu, E., Liu, J., Yu, K., Wang, Y., He, P., 2020. A hybrid model for predicting spatial distribution of soil organic matter in a bamboo forest based on general regression neural network and interative algorithm. J. For. Res. 31, 1673–1680. https://doi.org/10.1007/s11676-019-00980-3.

Lu, W., Lu, D., Wang, G., Wu, J., Huang, J., Li, G., 2018. Examining soil organic carbon distribution and dynamic change in a hickory plantation region with Landsat and ancillary data. Catena 165, 576–589. https://doi.org/10.1016/j.catena.2018.03.007.

Mabit, L., Bernard, C., 2010. Spatial distribution and content of soil organic matter in an agricultural field in Eastern Canada, as estimated from geostatistical tools. Earth Surf. Process. Landforms 35, 278–283. https://doi.org/10.1002/esp.1907.

McDowell, M.L., Bruland, G.L., Deenik, J.L., Grunwald, S., Knox, N.M., 2012. Soil total carbon analysis in Hawaiian soils with visible, near-infrared and mid-infrared diffuse reflectance spectroscopy. Geoderma 189–190, 312–320. https://doi.org/10.1016/j.geoderma.2012.06.009.

Mirzaee, S., Ghorbani-Dashtaki, S., Mohammadi, J., Asadi, H., Asadzadeh, F., 2016. Spatial variability of soil organic matter using remote sensing data. Catena 145, 118–127. https://doi.org/10.1016/j.catena.2016.05.023.

Muñoz, J.D., Kravchenko, A., 2011. Soil carbon mapping using on-the-go near infrared spectroscopy, topography and aerial photographs. Geoderma 166, 102–110. https://doi.org/10.1016/j.geoderma.2011.07.017.

Nawar, S., Buddenbaum, H., Hill, J., Kozak, J., Mouazen, A.M., 2016. Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. Soil Tillage Res. 155, 510–522. https://doi.org/10.1016/j.still.2015.07.021.

Nawar, S., Mouazen, A.M., 2018. Optimal sample selection for measurement of soil organic carbon using on-line vis-NIR spectroscopy. Comput. Electron. Agric. 151, 469–477. https://doi.org/10.1016/j.compag.2018.06.042.

Ondrasek, G., Bakić Begić, H., Zovko, M., Filipović, L., Meriño-Gergichevich, C., Savić, R., Rengel, Z., 2019. Biogeochemistry of soil organic matter in agroecosystems & environmental implications. Sci. Total Environ. 658, 1559–1573. https://doi.org/10.1016/j.scitotenv.2018.12.243.

Pouladi, N., Møller, A.B., Tabatabai, S., Greve, M.H., 2019. Mapping soil organic matter contents at field level with Cubist, Random Forest and kriging. Geoderma 342, 85–92. https://doi.org/10.1016/j.geoderma.2019.02.019.

Schmidt, M.W.I., Torn, M.S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I.A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D.A.C., Nannipieri, P., Rasse, D.P., Weiner, S., Trumbore, S.E., 2011. Persistence of soil organic matter as an ecosystem property. Nature 478, 49–56. https://doi.org/10.1038/nature10386.

Tian, Y., Zhang, J., Yao, X., Cao, W., Zhu, Y., 2013. Laboratory assessment of three quantitative methods for estimating the organic matter content of soils in China based on visible/near-infrared reflectance spectra. Geoderma 202–203, 161–170. https://doi.org/10.1016/j.geoderma.2013.03.018.

Vohland, M., Harbich, M., Ludwig, M., Emmerling, C., Thiele-Bruhn, S., 2016. Quantification of soil variables in a heterogeneous soil region with VIS-NIR-SWIR data using different statistical sampling and modeling strategies. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 https://doi.org/10.1109/JSTARS.2016.2572879.

Wang, H.J., Shi, X.Z., Yu, D.S., Weindorf, D.C., Huang, B., Sun, W.X., Ritsema, C.J., Milne, E., 2009. Factors determining soil nutrient distribution in a small-scaled watershed in the purple soil region of Sichuan Province, China. Soil Tillage Res. 105, 300–306. https://doi.org/10.1016/j.still.2008.08.010.

Wang, J., He, T., Lv, C., Chen, Y., Jian, W., 2010. Mapping soil organic matter based on land degradation spectral response units using Hyperion images. Int. J. Appl. Earth Obs. Geoinf. 12, S171–S180. https://doi.org/10.1016/j.jag.2010.01.002.

Wang, X., Zhang, F., te Kung, H., Johnson, V.C., 2018. 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. 218, 104–118. https://doi.org/10.1016/j.rse.2018.09.020.

Wu, C., Wu, J., Luo, Y., Zhang, L., DeGloria, S.D., 2009. Spatial prediction of soil organic matter content using cokriging with remotely sensed data. Soil Sci. Soc. Am. J. 73, 1202–1208. https://doi.org/10.2136/sssaj2008.0045.

Yu, X., Liu, Q., Wang, Y., Liu, Xiangyang, Liu, Xin, 2016. Evaluation of MLSR and PLSR for estimating soil element contents using visible/near-infrared spectroscopy in apple orchards on the Jiaodong peninsula. Catena 137, 340–349. https://doi.org/10.1016/j.catena.2015.09.024.

Zhao, Y., Xu, X., Tian, K., Huang, B., Hai, N., 2016. Comparison of sampling schemes for the spatial prediction of soil organic matter in a typical black soil region in China. Environ. Earth Sci. 75, 1–14. https://doi.org/10.1007/s12665-015-4895-4.

Zhu, P., Ren, J., Wang, L., Zhang, X., Yang, X., MacTavish, D., 2007. Long-term fertilization impacts on corn yields and soil organic matter on a clay-loam soil in Northeast China. J. Plant Nutr. Soil Sci. 170, 219–223. https://doi.org/10.1002/jpln.200620635.