

Assessing the carbon sequestration potential and identifying influential factors of cultivated soils in Northeast China

Shuai Wang^{a,b,c}, Zicheng Wang^a, Qianlai Zhuang^d, Kabindra Adhikari^e, Roland Bol^{b,f}, Yan Wang^g, Xingyu Zhang^a, Xinxin Jin^{a,*}, Fengkui Qian^{a,*}

^a College of Land and Environment, Shenyang Agricultural University, Shenyang 110866, Liaoning Province, China

^b Institute of Bio- and Geosciences, Agrosphere (IBG-3), Forschungszentrum Jülich GmbH, Wilhelm-Johnen-Straße, 52428 Jülich, Germany

^c Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

^d Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN, USA

^e USDA-ARS, Grassland, Soil and Water Research Laboratory, Temple, TX 76502, USA

^f School of Natural Sciences, Environment Centre Wales, Bangor University, Bangor LL57 2UW, UK

^g College of Tourism and Geography, Jiujiang University, Jiujiang 332005, China

ARTICLE INFO

Keywords:

Soil organic carbon potential sequestration
Cultivated soil
Environmental factors
Digital soil mapping
Cambisols
Phaeozems

ABSTRACT

Cultivated lands play a crucial role in terrestrial carbon cycle, and enhancing soil organic carbon (SOC) sequestration in these areas can help effectively mitigate the rise of atmospheric CO₂ concentration. In this study, topsoil (0–20 cm) saturated SOC and its density of cultivated soil in Northeast China were mapped using a boosted regression trees (BRT) model. The distribution of the SOC sequestration potential was also calculated based on the difference between saturated SOC and SOC density in ArcGIS. Nine environmental factors including climate, topography and lengths of cultivation data (LCD) and 197 soil samples are used. A 10-fold cross-validation technique is applied to derive four statistical indices - mean absolute prediction error (MAE), root mean square error (RMSE), coefficient of determination (R²) and Lin's consistent correlation coefficient (LCCC) to verify the model performance. The model explains 81% and 85% of the spatial variation of saturated SOC and SOC density, respectively. Mean annual temperature and mean annual precipitation are key factors controlling SOC density and saturated SOC distribution. In addition, LCD showed a similar spatial pattern to SOC sequestration potential, influencing the distribution of SOC density and saturated SOC in the study area. We recommend LCD as an important factor to consider in saturated SOC and SOC density predictions, especially in the farmland ecosystem with a long reclamation history. Accurate mapping SOC sequestration potential and identifying environmental factors will help manage land use and promote soil quality evaluation and improve soil carbon sequestration in this region.

1. Introduction

Climate change has a great impact on human living environment and sustainable social and economic development (Mikhaylov et al., 2020). Reducing greenhouse gas emissions and increasing carbon sequestration can minimize climate change impacts to the society and environment (Lal, 2004). Cultivated lands have considerable carbon sequestration potential and it can play an irreplaceable role in the global carbon cycle (Schulp and Verburg, 2009; Amelung et al., 2020). However, human activities have a significant impact on cultivated land ecosystems,

resulting in large temporal and spatial variations of soil organic carbon (SOC) (Ramankutty and Foley, 1999; Cui et al., 2011). Improper land management measures, such as reduction of organic matter input, frequent tilling, and leaving soils uncovered lead to the net loss of carbon in soils and the whole system becomes a carbon source (Sumfleth and Duttman, 2008; Amelung et al., 2020). On the contrary, implementing reasonable management measures such as minimum or no tillage can effectively enhance carbon content in cultivated soils (Hontoria et al., 1999), thereby this system may become a potential soil carbon sink (Farage et al., 2007; Denich et al., 2019; Walling and

* Corresponding authors at: College of Land and Environment, Shenyang Agricultural University, No. 120 Dongling Road, Shenhe District, Shenyang, Liaoning Province 110866, China.

E-mail address: Jinxinxin0218@syau.edu.cn (X. Jin).

<https://doi.org/10.1016/j.geodrs.2023.e00655>

Received 9 January 2023; Received in revised form 17 May 2023; Accepted 22 May 2023

Available online 25 May 2023

2352-0094/© 2023 Elsevier B.V. All rights reserved.

Vaneekhaute, 2020).

The storage and composition of SOC in cultivated lands determine SOC levels in these agroecosystems. Reasonably increasing carbon storage in cultivated land ecosystems is of great significance for ensuring national food security and mitigating global climate change (Lal, 2004; Kane and Solutions, 2015; Wang et al., 2019a, 2019b). Saturated SOC of cultivated soil represents the maximum stable capacity for carbon storage under local environmental conditions (Rogasik et al., 2004; Singh and Lal, 2005; Lorenz and Lal, 2005). It is the balance between carbon input determined by transforming organic carbon into soils and output due to decomposition (Akpa et al., 2016). The SOC sequestration potential can be calculated by subtracting the SOC density from the saturated SOC (Hu et al., 2022). Therefore, precise and efficient prediction of SOC sequestration potential in cultivated land ecosystems across large regions holds practical significance for the sustainable management of cultivated land (Rogasik et al., 2004; Novara et al., 2021).

There are several methods to estimate the regional or global SOC sequestration potential. One is based on long-term field experimental data under optimized management condition and be extrapolated to regional levels, for example, the studies of Lal (2002) and Lu et al. (2009) who estimated SOC sequestration potential of cultivated soils in China. The other method is to use simulation based on process-based model and scenario hypothesis (Heenan et al., 1995), such as the Carbon Exchange between Vegetation, Soil, and Atmosphere (CEVSA) (Yan et al., 2007), Century (Kelly et al., 1997), Denitrification-decomposition (DNDC) (Li et al., 2016), RothC (Coleman et al., 1997; Li et al., 2016), and Erosion-Productivity Impact Calculator (EPIC) (Izaurre et al., 2006). Another approach is the Dexter-ratio method that quantifies SOC sequestration potential empirically using SOC and clay content, and it has been successfully tested in Denmark (Schjøning et al., 2009). The aforementioned approaches have their own limitations, especially when considering the spatial heterogeneity of soil and the regional climate variability. As a result, there may be significant uncertainties in estimating SOC sequestration potential in cultivated lands (Podwojewski et al., 2011; Wang et al., 2011; Martin et al., 2011; Dorji et al., 2014). For example, the extrapolation of SOC sequestration potential rate from the optimized land management condition to other areas may not be congruent (Allmaras et al., 2000; Jiang et al., 2014; Novara et al., 2021), as it ignores the local SOC sequestration potential influencing factors leading to high uncertainty in estimation (Ogle et al., 2003; Qin et al., 2013; Wang et al., 2019a, 2019b). Most simulation models require several inputs and parameters to optimize, and it is rather difficult to model output variability and uncertainty in the spatial context (Padarian et al., 2019; Wadoux et al., 2020). Moreover, some models perform well in some local conditions but may not perform well in other areas (Yan et al., 2007). For example, the Dexter-ratio method can only be used to identify area with critical low organic matter contents (Schjøning et al., 2009). The limitations of these models are significant and large-scale extrapolation cannot be conducted (Smith et al., 1997).

Therefore, this study introduced digital soil mapping (DSM) technology and combined it with long-term field experiment data. DSM is a precise, efficient, and economical soil attribute mapping technique that can quickly obtain high-precision and high-resolution spatial change information of soil attributes (McBratney et al., 2003; Malone et al., 2011; Adhikari et al., 2013; Wang et al., 2022). Among many DSM models, the boosted regression trees (BRT) model does not need prior data conversion or outlier removal, can fit complex nonlinear relationships, and can automatically handle the interaction between predictors (McCaffrey et al., 2004; Elith et al., 2008). The process of fitting multiple trees using BRT can maximally compensate for the weak predictive ability of a single tree model, and has stronger predictive ability than most traditional model methods. It can handle a large number of practical problems in model fitting (Martin et al., 2011; Yang et al., 2016; Wang et al., 2022).

In this study, we developed an efficient SOC sequestration potential

model using a long-term field experiment data and the BRT model, and mapped the spatial SOC sequestration potential distribution across the cultivated soil in Northeast China. Our study objectives were to: 1) accurately map the saturated SOC and SOC density of cultivated soil in Northeast China; 2) understand the mechanism of soil carbon sequestration and its influencing factors; and 3) objectively evaluate the soil carbon sequestration potential of cultivated soil at the regional scale, which is the premise for formulating reasonable carbon management measures.

2. Material and methods

2.1. Study area

The study area is located in the Northeast China Plain (118.53°–135.05° E, 38.43°–53.33° N) (Fig. 1) that includes Songnen Plain, Liaohe plain and Sanjiang Plain, and it covers an area of about 217,000 km² accounting for 20% of the total cultivated land in China. Most part of the region belongs to temperate monsoon climate, but due to the high latitude of some regions, the winter is cold and long, the summer is warm and short, the snow falls in winter, the evaporation is small, the climate is humid, and the lowlands are swampy (cold and wet). The mean annual temperature (MAT) ranges from −4 °C to 11 °C, and the mean annual precipitation (MAP) ranges from 350 mm to 1100 mm. The terrain is mainly plain and mountainous, with the highest elevation of 2665 m and the average elevation of 200 m (Fig. 1). According to the World Reference Base (WRB) for Soil Resources (IUSS Working Group WRB, 2006), the dominant soil types are Cambisols (52.4%) and Phaeozems (34.1%) both covering >85% of the study area, and the remaining area covered by other soil types, Anthrosols (4.1%), Gleysols (3.8%), Leptosols (3.7%), and Luvisols (1.0%).

With the growth of population, the demand for grain and other agricultural products increased, and the area of reclamation increased year by year. The earliest written record of cultivated land reclamation in the Northeast Plain began in the Qing Dynasty of China (1636–1912), and gradually moved from the south to the north of China. At present, Northeast China is an important grain region in China, mainly planting corn, soybean and sugar beet. According to China's National Bureau of statistics (<http://www.stats.gov.cn/english/>), the total grain output of this region in 2021 was 144.456 million tons, accounting for 21.15% of the total grain output of the country. The study area presents an ideal area to investigate the SOC sequestration potential in cultivated land ecosystems.

2.2. Soil sampling and analysis

Because of the diverse nature of the study site and vast area covered, we applied a purposeful sampling design (Zhu et al., 2008) to identify soil sampling locations. We first obtained the quantitative information of synergistic environmental factors reflecting the spatial change of SOC. The main factors considered were climate (temperature and precipitation), topography (elevation, slope gradient, slope aspect, etc.) and soil reclamation history. Through fuzzy c-means clustering (FCM) the environmental factors were clustered to obtain an optimum combination of environmental factors corresponding to the spatial change pattern of soil organic carbon. We obtained 27 different clusters or environmental-landscape units, and selected 8–10 sampling points in each unit to collect soil samples. Finally, a total of 197 sampling sites were identified, and from each site one kilogram of soil sample was collected from the topsoil (0–20 cm) depth in 2013. The geographic coordinates of each site were recorded using a handheld global positioning system (Garmin, eTrex221 x, America). A separate core sample was collected for bulk density (BD) estimation with three replicates. Soil samples were sent to the Liaoning Key Laboratory of agricultural resources and environment, Shenyang Agricultural University for laboratory analysis. SOC content was determined using wet oxidation method

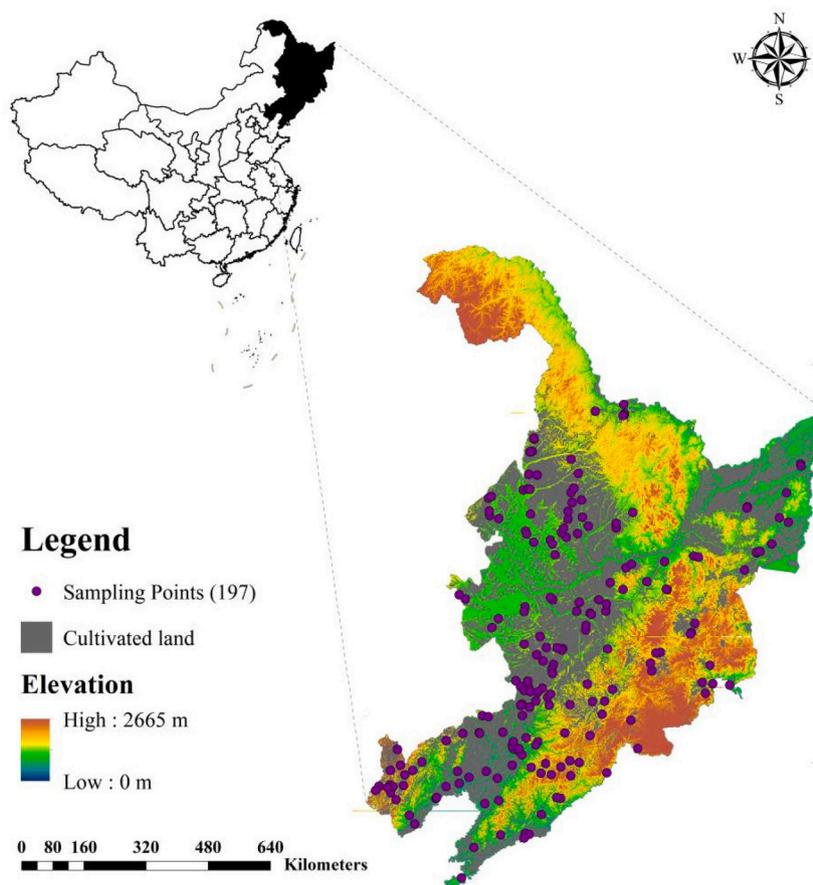


Fig. 1. Location of the study area and 197 sampling sites superimposed on a 90-m resolution digital elevation model.

(Nelson and Sommers, 1996), BD by oven-dry method, pH using a pH meter (Ohaus, AB33pH, USA), and soil texture by hydrometer method (McKeague, 1978). In addition, because the soil cores were not collected at 37 sampling sites leading those sampling sites missing BD data, which were estimated according the Pedo-Transfer Functions (PTFs) (Wang et al., 2019a, 2019b). The specific formula is as follows:

$$BD = 1.52 - 0.08 \times \sqrt{SOC} \quad (R^2 = 0.74, p < 0.001) \quad (1)$$

2.3. Environmental factors

2.3.1. Topographic factors

A digital elevation model (DEM) with a 90-m grid resolution was downloaded from the Geospatial Data Cloud site of the Chinese Academy of Sciences (<http://www.gscloud.cn>). Five topographic factors, namely elevation, slope gradient (SG), slope aspect (SA), plan curvature (PLC) and profile curvature (PRC) were derived from the DEM in ArcGIS 10.2 (ESRI Inc., USA). Topographic wetness index (TWI) was calculated in System for Automated Geoscientific Analyses (SAGA GIS) (Conrad et al., 2015). Elevation is the height difference between a certain place and the sea level. SG is the steepness and gentleness of the surface unit, which is the deviation of a topographic surface from the horizontal plane. SG directly affects the distribution of surface material and energy, the development of soil, the distribution of vegetation and the type of land use (Bae and Ryu, 2015). SA affects the flow direction of surface runoff and the distribution of surface light and heat resources (Muñoz-Rojas et al., 2015). Curvature can be used to characterize the movement of surface materials and also affect the distribution of soil attributes (Conforti et al., 2016). TWI can reveal the soil attributes and spatial distribution under different landscapes through the quantitative

reflection of soil water storage and drainage (Kirkby, 1999; Zhu et al., 2008). These topographic factors are known to influence the spatial distribution of soil attributes including SOC distribution (Fissore et al., 2017; Guo et al., 2019; Zhang et al., 2021).

2.3.2. Climatic factors

A 30-yr annual average of MAT and MAP data were downloaded from China Meteorological Data Service Center (<http://data.cma.cn/en>), and were interpolated using inverse distance weighting to generate continue maps. The downloaded climate factors (MAT and MAP) were 1-km resolution raster data, which were resampled to a grid of 90 m to be used in this study. High altitude areas have higher MAP and lower MAT, with opposite trends in MAP and MAT.

2.3.3. Data on cultivation history

In the past 300 years, Northeast China has been a typical area where land use has changed a lot, especially through land reclamation (Wang et al., 2019a, 2019b). Due to changes in political power, policies and migration, the land development has undergone several major changes and the reclamation boundary has been moving northward. This has created a great impact on social and economic development of the whole region since the last 300 years. To quantify the influence of cultivation history on SOC distribution (Wang et al., 2022), we combined the detailed historical data on land reclamation and applied a factor correction method to update and delimit the lengths of cultivation data (LCD) of Ramankutty and Foley (1999). The updated LCD was further divided into seven time periods, namely 0–10 years (T1), 10–30 years (T2), 30–70 years (T3), 70–120 years (T4), 120–200 years (T5), 200–300 years (T6) and >300 years (T7). The time period LCD were later converted into integer values 1, 2, 3, 4, 5, 6 and 7, respectively, to

facilitate modeling. The entire procedure to derive the LCD is documented in Wang et al. (2019a, 2019b).

2.4. Prediction model

Prediction model was built using Boosted regression tree (BRT) algorithms which have been widely used in soil attributes predictions with promising results (Martin et al., 2011; Yang et al., 2016; Wang et al., 2018). BRT was proposed by Elith et al. (2008), which has a good interpretation and prediction characteristics, such as it can handle nonlinear relationship between variables and is not typically affected by missing values (McCaffrey et al., 2004). BRT combines regression tree and boosting methods (Martin et al., 2011). Regression tree part combines dependent variables and their predictors with recursive binary splits (Yang et al., 2016) while boosting combines multiple simple models and improve the prediction ability (Wang et al., 2022). This method does not need prior data conversion or outlier removal, can fit complex nonlinear relationships, and can automatically deal with the interaction between prediction factors (Elith et al., 2008). It has stronger prediction ability than most traditional tree-based models and can handle a large number of practical problems in model fitting (McCaffrey et al., 2004; Martin et al., 2011; Yang et al., 2016).

The BRT model requires to set four model parameters, namely learning rate (LR), tree complexity (TC), bagging fraction (BF), and number of trees (NT). LR determines the relative contribution rate of a single decision tree during model building (Yang et al., 2016). TC controls the number of branch nodes of a single decision tree or the number of splits, representing the depth of the tree and controls interaction between prediction variables (Wang et al., 2018). BF represents the proportion of data set be used in the model, that is, the proportion of random observations used to obtain a single decision tree in the total sample data set, which is generally recommended to be 0.5–0.75 (McCaffrey et al., 2004). NT is obtained by the best fitting between LR and TC (Yang et al., 2016). Finally, we set LR, TC, BF, and NT to 0.025, 9, 0.55, and 1500 in predicting saturated SOC, respectively. In predicting SOC density, LR, TC, BF, and NT were set to 0.025, 9, 0.65, and 1750, respectively.

2.5. Calculation of saturated soil organic carbon

The calculation of soil saturated SOC follows the method proposed by Qin and Huang (2010), which used long-term field experimental data from 95 cultivated lands worldwide to establish a comprehensive test database. The database included information about the geography, climate (temperature and precipitation), soil attributes including SOC and management information. They used Levenberg marquardt (LM) and universal global optimization (UGO) algorithms to establish the saturated SOC model. Among 95 long-term field experimental data, 76 long-term field experimental data from all over the world except China were used for correlation analysis and model parameter determination, and the other 19 long-term field experimental data from China were used for model validation. The calculation formula of saturated SOC was as follows:

$$\begin{aligned} \text{Saturated SOC} = & 140.5 \times e^{(-0.021 \times \text{MAT})} - 98.8 \times e^{(-0.42 \times \text{MP})} - 39.6 \times e^{(-0.1 \times \text{CL})} \\ & - 4.1 \times \text{pH} - 27.7 \end{aligned} \quad (2)$$

$$(R^2 = 0.58, n = 76)$$

where saturated SOC is the soil organic carbon sequestration potential; MAT is the mean average temperature (°C); MP refers to the mean annual water supply (100 mm, paddy field is the sum of mean annual precipitation and mean annual irrigation; dry farm is only the mean annual precipitation); CL is the soil clay content (%). The model coefficient is determined by 76 foreign long-term field experimental data.

2.6. Calculation of soil organic carbon density

Soil organic carbon (SOC) density was calculated according to Eq. 3:

$$\text{SOC density} = \text{SOC} \times \text{BD} \times D \times (1 - F) \div 10 \quad (3)$$

where SOC density is soil organic carbon density (t ha^{-1}); SOC is soil organic carbon content (g kg^{-1}); D represents the thickness of the soil layer (cm), and this study focused on the topsoil (20 cm) with a sampling depth of 10 cm. F represents the fraction of >2 mm fragments in soil; BD represents soil bulk density (g cm^{-3}).

2.7. Model validation

The BRT model performance was evaluated with mean absolute error (MAE), root mean square error (RMSE), coefficient of determination (R^2) and Lin's concordance correlation coefficient (LCCC) (Lin, 1989) indices (Eq. 4–7) using a 10-fold cross-validation technique. In the R environment (R Development Core Team, 2013), the BRT model comes with a ten fold cross validation technology, which can be implemented and verified during the BRT model construction process (Elith et al., 2008). The saturated SOC model was verified by the long-term field experimental data of 19 stations in China with MAE, RMSE, and R^2 . The calculation formula are as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |X_i - Y_i| \quad (4)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - Y_i)^2} \quad (5)$$

$$R^2 = \frac{\sum_{i=1}^n (X_i - \bar{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y}_i)^2} \quad (6)$$

$$\text{LCCC} = \frac{2r\partial_X\partial_Y}{\partial_X^2 + \partial_Y^2 + (\bar{Y} - \bar{X})^2} \quad (7)$$

where X_i and Y_i are the observed and predicted saturated SOC and SOC density values at site i , respectively; n is the number of samples; ∂_X and ∂_Y are the variances of observed and predicted values; and r is the Pearson correlation coefficient between the observed and predicted values.

3. Results

3.1. Exploratory data analysis

The summary statistics of saturated SOC, SOC density and 9 environmental variables at sampling locations are shown in Table 1. Saturated SOC ranged from 35.30 t ha^{-1} to 89.04 t ha^{-1} , with an average of 65.87 t ha^{-1} . The average SOC density was 46.8 t ha^{-1} . The skewness coefficient of saturated SOC and SOC density were -0.38 t ha^{-1} and 0.26 t ha^{-1} , respectively. The dataset presents a non-generalized skew distribution.

The linear relationship between saturated SOC and SOC density with environmental variables is shown in Table S1. Elevation (0.39 vs. 0.32), TWI (0.33 vs. 0.25) and MAP (0.46 vs. 0.17) were significantly positively correlated with saturated SOC and SOC density. Correspondingly, saturated SOC and SOC density were significantly negatively correlated with SG (-0.17 vs. -0.16), MAT (-0.77 vs. -0.68) and LCD (-0.59 vs. -0.62). In addition, multi-collinearity among the environmental factors was checked with variance inflation factor (VIF). The results showed that the VIF of each factor was <3 , indicating there was no

Table 1
Summary statistics of saturated SOC, SOC density and environmental variables at sampling sites.

Property	Unit	Min.	Max.	Mean	SD	Skewness	Kurtosis
Saturated SOC	t ha ⁻¹	35.30	89.04	65.87	11.72	-0.38	-0.54
SOC density	t ha ⁻¹	16.15	83.22	46.81	14.50	0.26	-0.63
Elevation	m	2.00	673.00	207.57	128.66	1.03	1.39
SG	degree	0.00	17.87	1.90	2.40	3.00	12.50
SA	degree	0.00	358.26	163.63	103.84	0.10	-1.14
PRC		-0.28	0.15	0.00	0.04	-1.62	13.81
PLC		-0.15	0.24	0.00	0.04	0.93	9.03
TWI		6.86	12.99	10.57	1.19	-0.29	-0.31
MAT	degree Celsius	-0.30	10.60	5.34	2.33	-0.02	-0.73
MAP	mm	409.30	1094.40	593.22	130.81	1.35	2.06

Notes: SOC, soil organic carbon; SG, slope gradient; SA, slope aspect; PRC, profile curvature; PLC, plan curvature; TWI, topographic wetness index; MAT, mean annual temperature; MAP, mean annual precipitation; SD, standard deviation; Min., minimum value; Max., maximum value.

multicollinearity problem (Table S2).

3.2. Performance evaluation and uncertainty

RMSE, MAE and R² were selected to test the accuracy of the saturated SOC model simulation, which were 7.0 t ha⁻¹, 5.7 t ha⁻¹ and 0.74, respectively. The validation results showed that the saturated SOC model could better simulate the saturated SOC of cultivated land soil in China based on local climate and soil factors. The regression slope and intercept between saturated SOC simulation value and observation value were 0.82 and 5.2 t ha⁻¹, respectively.

The summary statistics of the validation indices based on 100 iterations of the BRT model to predict saturated SOC and SOC density are shown in Table 2. The model had a lower systematic MAE (3.63 t ha⁻¹ vs. 4.40 t ha⁻¹) and RMSE (4.99 t ha⁻¹ vs. 5.68 t ha⁻¹) but a higher R² (0.79 vs. 0.85) and LCCC (0.90 vs. 0.92) to predict saturated SOC and SOC density, respectively. The R² value ranged between 0.73 and 0.82 for saturated SOC, and between 0.82 and 0.86 for SOC density prediction. Similarly, LCCC ranged from 0.88 to 0.90 for saturated SOC and from 0.91 to 0.92 for SOC density predictions. Results showed that the BRT model was able to explain nearly 80% and 85% of the saturated SOC and SOC density variations in the study area, respectively.

The uncertainty in the model prediction was represented by the standard deviation (SD) generated by 100 iterations of the BRT model (Fig. 2c and d). The model had an average SDs of 0.64 and 0.97 t ha⁻¹, respectively, in predicting saturated SOC and SOC density confirming a robust prediction. In addition, we also mapped the histograms of RMSE for 100 iterations of the BRT model (Fig. 3). The mean RMSE for saturated SOC and SOC density were 4.99 t ha⁻¹ and 5.68 t ha⁻¹, and the RMSE for each BRT model run was distributed on the mean RMSE, indicating that the BRT model has robust performance in predicting saturated SOC and SOC density. Although other sources of uncertainty such as sampling error, laboratory analysis error and model error that might have reduced the model performance, quantification of such uncertainty was not included in the present study.

Table 2
Predictive quality of the boosted regression tree (BRT) models for saturated SOC and SOC density.

Property	Index	Min.	1st quartile	Median	Mean	3rd quartile	Max.
Saturated SOC	MAE	3.51	3.59	3.63	3.63	3.66	3.82
	RMSE	4.84	4.93	4.98	4.99	5.03	5.31
	R ²	0.73	0.78	0.80	0.79	0.81	0.82
	LCCC	0.88	0.90	0.90	0.90	0.90	0.90
	MAE	4.33	4.36	4.38	4.40	4.45	4.47
SOC density	RMSE	5.60	5.61	5.68	5.68	5.72	5.81
	R ²	0.82	0.84	0.85	0.85	0.85	0.86
	LCCC	0.91	0.92	0.92	0.92	0.92	0.92

Notes: SOC, soil organic carbon; MAE, mean absolute error; RMSE, root mean squared error; R², coefficient of determination; LCCC, Lin's concordance correlation coefficient; Min., minimum value; Max., maximum value.

3.3. Relative importance of environment factors

The average relative importance (RI) of environmental factors in predicting topsoil saturated SOC and SOC density was calculated through 100 BRT model iterations, and the RI values were standardized to 100% (Fig. 4). We found MAT, MAP, elevation, LCD and TWI (accounting for 81.6% of RI) as the main environmental factors affecting the spatial variability of topsoil saturated SOC, and MAT, MAP, TWI, elevation, SG and LCD that accounted for 83.6% of RI for SOC density in Northeast China. It was worth noting that climatic factors (MAT and MAP) played an important role in the spatial variation of saturated SOC and SOC density.

3.4. Spatial distribution of SOC sequestration potential

Predicted map of saturated SOC and SOC density in the study area are shown in Fig. 2a and b. Both saturated SOC and SOC density maps had similar spatial distribution characteristics- predicted values gradually decreased from southwest to northeast. The average of saturated SOC and SOC density were 68.28 t ha⁻¹ and 49.70 t ha⁻¹, respectively.

Fig. 5a is the SOC sequestration potential map derived as a difference between the saturated SOC and SOC density map. SOC sequestration potential refers to the maximum capacity of soil to increase SOC under certain environmental conditions, which is jointly affected by human activities, soil characteristics and natural environment. Average SOC sequestration potential in the study area was 18.47 t ha⁻¹ with a SD of 5.57 t ha⁻¹. The SOC sequestration potential values were mainly concentrated in three different levels: 10–15 t ha⁻¹, 15–20 t ha⁻¹ and 20–25 t ha⁻¹, accounting for about 80.4% of the total area (Fig. 5b). We found that the spatial distribution pattern of SOC sequestration potential was similar to the LCD (Fig. S1), and the SCO sequestration potential was mainly concentrated at three LCD levels: T4 (70–120 years), T5 (120–200 years) and T6 (200–300 years) (Table 3). In order to further reveal the spatial distribution characteristics of SOC sequestration potential in the region, we summarized the SOC sequestration potential under different soil types (Table 4). The results showed that the SOC sequestration potential of Cambisols and Phaeozems was the largest,

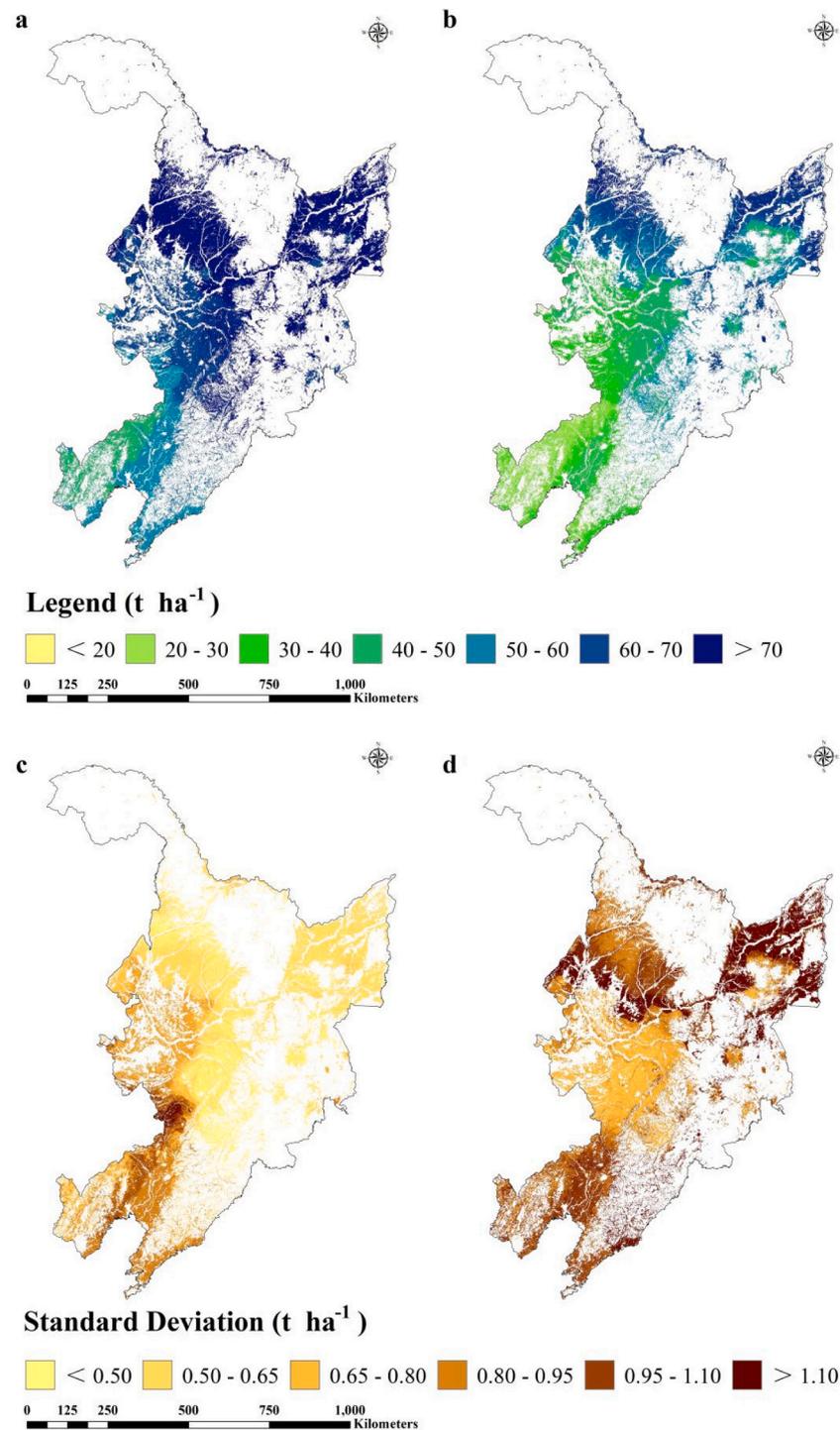


Fig. 2. Standard deviation (SD) and spatial distribution of the saturated soil organic carbon ($t\ ha^{-1}$) (a,c) and the soil organic carbon density ($t\ ha^{-1}$) (b,d) predicted from 100 runs of the boosted regression trees (BRT) model.

accounting for 52.5% and 34.9% of the total SOC sequestration potential in the study area, respectively. The corresponding SOC sequestration potential was $29.9 \times 10^7\ t$ and $19.9 \times 10^7\ t$, respectively.

4. Discussion

4.1. Role of environmental factors in saturated SOC

Among all the environmental factors, MAT and MAP were identified as the key variables affecting the spatial distribution of saturated SOC in

the study area (Fig. 4), as also reported in previous studies been obtained (Giardina and Ryan, 2000; Fantappiè et al., 2011; Adhikari et al., 2019). Changes in climatic factors influence SOC levels through precipitation, temperature and atmospheric CO_2 concentration, (Wiesmeier et al., 2012; Wang et al., 2019a, 2019b; Mikhaylov et al., 2020) and subsequently influenced the spatial variations in saturated SOC. Availability of water and heat control the nature and quantity of plant residues, an important source of soil organic matter (Lal, 2004), and it impacts microbial population, their activity and biomass (Luo et al., 2017). The rise of temperature would cause drought, exacerbate transpiration and

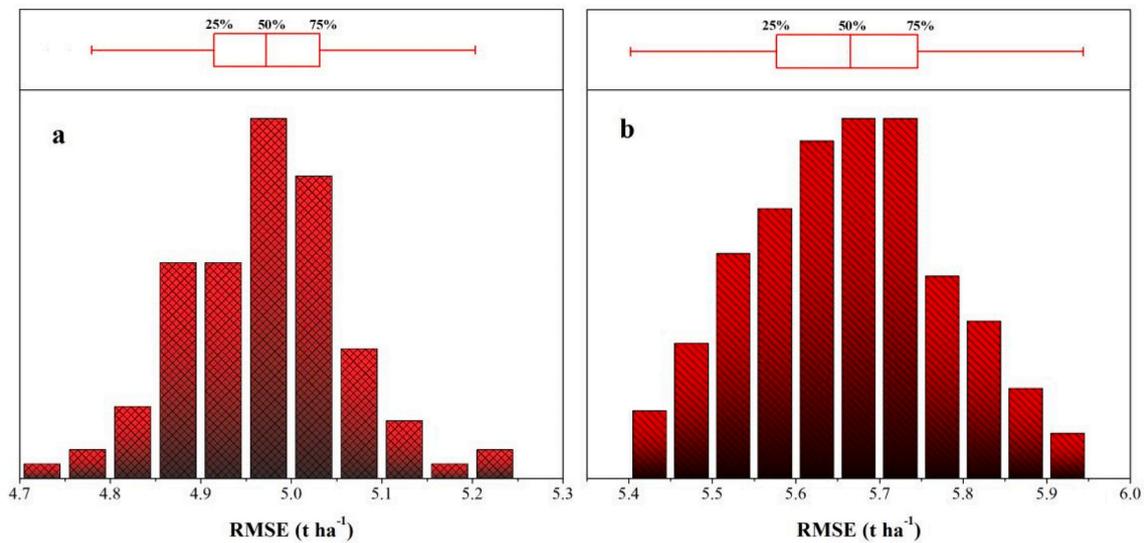


Fig. 3. Histogram showing the root mean square error (RMSE) response to 100 iterations of BRT model for saturated soil organic carbon (a) and the soil organic carbon density (b) predictions (b).

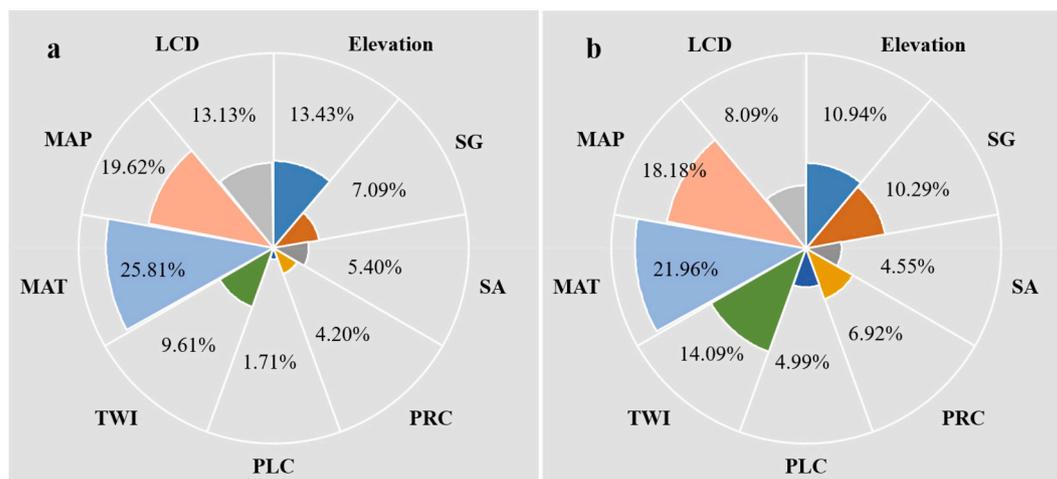


Fig. 4. Relative importance of each predictor in predicting a) saturated soil organic carbon and b) soil organic carbon density using the boosted regression tree (BRT) model. SG, slope gradient; SA, slope aspect; PRC, profile curvature; PLC, plan curvature; TWI, topographic wetness index; MAT, mean annual temperature; MAP, mean annual precipitation.

evaporation, promoting the release of SOC (Fantappiè et al., 2011; Aguilera et al., 2013; Wang et al., 2018; Reyes Rojas et al., 2018). Similarly, too high or too low water would limit soil respiration and the release of SOC. In Northeast China, temperature and precipitation affect the spatial differentiation of saturated SOC by changing crop productivity and litter decomposition (Wang et al., 2019a, 2019b).

As one of the five major soil forming factors, topographic factors had a vital impact on the spatial distribution of soil moisture, temperature conditions and other soil forming processes (Webster et al., 2011; Brus et al., 2016). Therefore, the spatial variability of soil properties was closely related to topographic factors (McBratney et al., 2003; Adhikari et al., 2013; Yang et al., 2016). We found elevation as the most important topographic factors in saturated SOC prediction, followed by TWI, SG and SA, while PRC and PLC had the weakest influence. The study area is highly variable in terms of topography, and elevation could form different hydrothermal conditions or micro-climatic pockets, which indirectly affected microbial activity, resulting in the variation of saturated SOC. Similar conclusion has been verified by Wiesmeier et al. (2014a, 2014b) and Akpa et al. (2016) in their studies. TWI could reflect the influence of topography on soil water distribution, and might have

influenced the spatial variation of saturated SOC in the region (Allmaras et al., 2000). SG had a negative effect on the spatial distribution of saturated SOC, which could be attributed to the low vegetation biomass and limited retention of precipitation on the steep slope (Novara et al., 2021).

In cultivated land ecosystems, a higher correlation of saturated SOC and SOC density with LCD was expected (Bedard-Haughn et al., 2006; Wang et al., 2017; Wang et al., 2022). We found LCD as one of the main factors affecting saturated SOC in the study area. Tillage measures had a positive impact on SOC, especially on microorganisms, which could explain the higher RI of LCD in saturated SOC prediction (Ramankutty and Foley, 1999). Once the natural soil is reclaimed and converted to a cultivated land, surface vegetation and soil microbial community changes accelerating a faster SOC decomposition during the early stage of reclamation. After that, a gradual downward trend in decomposition is expected with the increase of reclamation years (Post and Kwon, 2000; Jiang et al., 2014; Xu et al., 2020). Therefore, we proposed to use LCD in the spatial prediction of saturated SOC and SOC density, especially in areas with a long farming history.

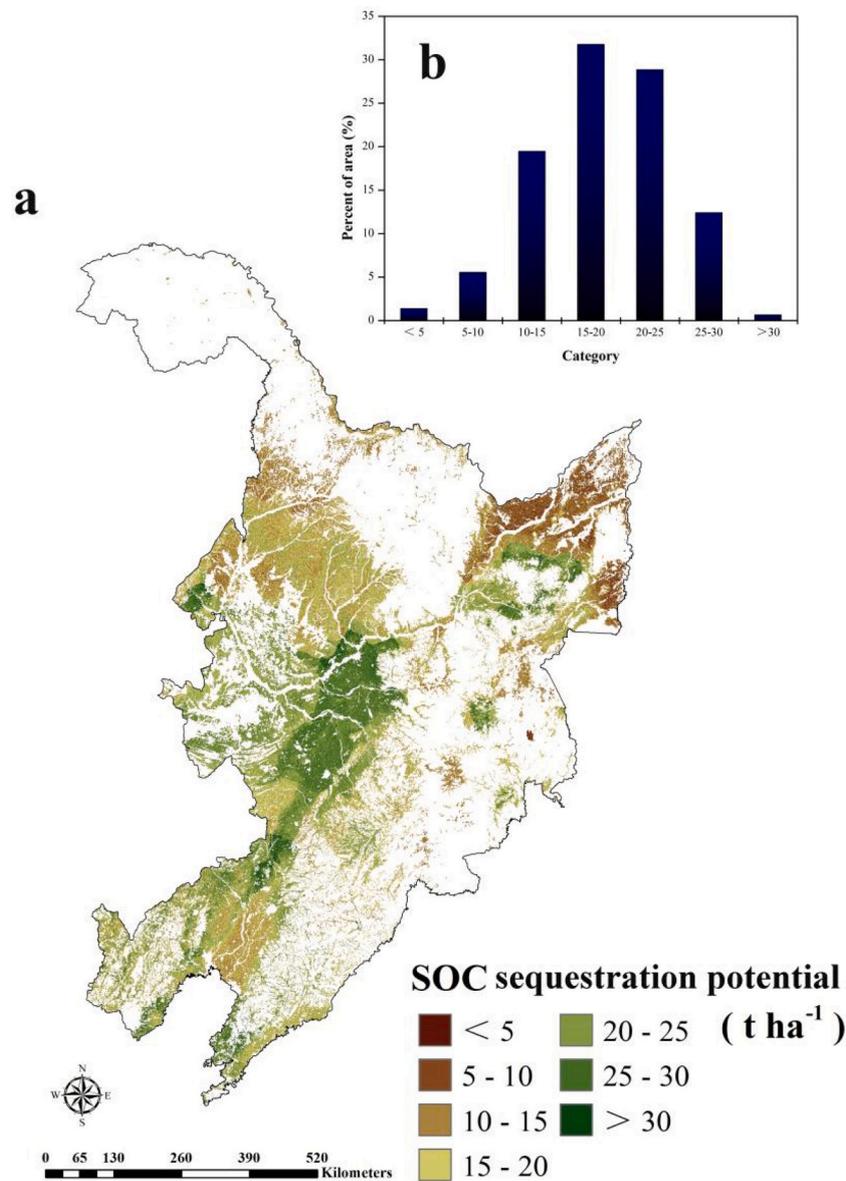


Fig. 5. Spatial variation map (a) and percentages map (b) of soil organic carbon sequestration potential in the future.

Table 3
Soil organic carbon sequestration potential under different Lengths of cultivation.

Lengths of cultivation	Area (km ²)	Saturated SOC		SOC density		SOC sequestration potential (t)
		Mean ± SD (t ha ⁻¹)	Summary (t)	Mean ± SD (t ha ⁻¹)	Summary (t)	
T1	14,001.37	69.6 ± 25.17	10.70 × 10 ⁷	57.69 ± 21.69	8.88 × 10 ⁷	1.82 × 10 ⁷
T2	7850.37	64.12 ± 21.67	5.63 × 10 ⁷	51.06 ± 18.94	4.60 × 10 ⁷	1.03 × 10 ⁷
T3	37,871.08	72.3 ± 20.23	29.03 × 10 ⁷	59.78 ± 18.14	24.04 × 10 ⁷	4.99 × 10 ⁷
T4	131,269.38	65.97 ± 18.58	91.90 × 10 ⁷	48.16 ± 16.02	66.96 × 10 ⁷	24.94 × 10 ⁷
T5	50,297.47	63.42 ± 17.53	33.14 × 10 ⁷	44.54 ± 16.01	23.18 × 10 ⁷	9.96 × 10 ⁷
T6	53,196.65	54.33 ± 16.1	30.55 × 10 ⁷	34.78 ± 11.98	19.40 × 10 ⁷	11.15 × 10 ⁷
T7	15,885.68	49.84 ± 12.67	8.34 × 10 ⁷	31.26 ± 9.37	5.23 × 10 ⁷	3.11 × 10 ⁷
Total	310,372.00	-	209.29 × 10 ⁷	-	152.29 × 10 ⁷	57.00 × 10 ⁷

Note: SOC, soil organic carbon; SD, standard deviation; T1, 0–10 years; T2, 10–30 years; T3, 30–70 years; T4, 70–120 years; T5, 120–200 years; T6, 200–300 years; T7, above 300 years.

Table 4
Soil organic carbon sequestration potential under different soil type.

Soil type	Area (km ²)	Saturated SOC		SOC density		SOC sequestration potential (t)
		Mean ± SD (t ha ⁻¹)	Summary (t)	Mean ± SD (t ha ⁻¹)	Summary (t)	
Andosols	54.51	68.1 ± 30.88	0.04 × 10 ⁷	56.55 ± 26.47	0.03 × 10 ⁷	0.01 × 10 ⁷
Anthrosols	13,013.03	62.93 ± 15.38	8.40 × 10 ⁷	45.83 ± 13.05	6.09 × 10 ⁷	2.31 × 10 ⁷
Luvisols	3144.21	68.71 ± 22.33	2.3 × 10 ⁷	53.76 ± 19.64	1.78 × 10 ⁷	0.52 × 10 ⁷
Cambisols	162,720.47	61.13 ± 20.72	105.14 × 10 ⁷	44.28 ± 18.57	75.21 × 10 ⁷	29.94 × 10 ⁷
Gleysols	11,888.15	70.68 ± 22.66	8.99 × 10 ⁷	58.74 ± 20.28	7.49 × 10 ⁷	1.51 × 10 ⁷
Solonchaks	1437.4	54.2 ± 19.22	0.85 × 10 ⁷	34.56 ± 12.84	0.54 × 10 ⁷	0.31 × 10 ⁷
Histosols	533.71	66.94 ± 21.72	0.38 × 10 ⁷	52.62 ± 19.54	0.30 × 10 ⁷	0.08 × 10 ⁷
Phaeozems	106,000.64	70.06 ± 14.97	76.38 × 10 ⁷	52.02 ± 14.9	56.48 × 10 ⁷	19.90 × 10 ⁷
Leptosols	11,579.88	54.97 ± 17.49	6.80 × 10 ⁷	35.61 ± 12.87	4.38 × 10 ⁷	2.42 × 10 ⁷
Total	310,372.00	–	209.29 × 10 ⁷	–	152.29 × 10 ⁷	57.00 × 10 ⁷

Note: SOC, soil organic carbon; SD, standard deviation.

4.2. Estimates of SOC sequestration potential and its distribution

The map of saturated SOC and SOC density showed a similar spatial distribution pattern (Fig. 2a and b), gradually decreasing saturated SOC and SOC density from southeast to northwest. The larger saturated SOC and SOC density were usually concentrated in the north of the study area where the main soil type was Phaeozems, an area with the most abundant SOM in China. Once these areas are disturbed by human activities and reclaimed into cultivated land, the organic carbon in the soil will be lost rapidly. In Fig. S1, we could see that the LCD showed an opposite spatial distribution pattern with saturated SOC and SOC density (Fig. 2a and b). In areas with long reclamation years, saturated SOC and SOC density were usually low. As reported in Wang et al. (2019a, 2019b), this study further verified the usefulness of LCD to represent the degree of human interference in cultivated land, especially in areas with long farming history. We found that saturated SOC and SOC density decreased gradually with the increase of reclamation years. Although the value of saturated SOC should not change with the change of LCD in theory, it was difficult to reach the theoretical maximum value with the influence of human activities. It was only a dynamic balance, which was different from the saturated SOC of natural soil. Therefore, human activities were the main inducement for the reduction of SOC in cultivated land soil. In Wafangdian District of Dalian, Wang et al. (2018) reached a similar conclusion. They concluded that human activities were the main reason for the reduction of topsoil SOC stocks in the main agricultural region that supplied agricultural products to annex cities. Similar conclusions were reached in Seoul Forest Park, Seoul, Republic of Korea by Bae and Ryu (2015).

Among all soil types, Camposols and Phaeozems were the main soil types in the region (Table 4). The saturated SOC and SOC density under the Camposols were 105.14 × 10⁷ t and 75.21 × 10⁷ t, accounting for 50.24% and 49.39% of all saturated SOC and SOC density, respectively. In the Northeast cultivated land ecosystem, the variation of SOM in the Camposols was large, and the change of soil erosion rate caused by human activities and tillage were the main reasons (Liu et al., 2006). Phaeozems was mainly distributed in Heilongjiang provinces and Jilin provinces, which was the most fertile land and the largest commercial grain production base in China (Yu et al., 2006). The areas with Phaeozems are characterized by warm and rainy summer, with lush vegetative growth causing more organic matter entering the soil (Zhang et al., 2016). Correspondingly, the cold winter makes the soil frozen and thus the microbial decomposition is inhibited, leading to the accumulation of organic matter in the soil mainly in the form of humus (Yu et al., 2006). Therefore, a thicker humus layer is formed in this area, with the humus content gradually decreasing from the top to the bottom of the soil profile. Under natural conditions, the humus layer of Phaeozems could be up to 1 m thick, with rich nutrient content and high fertility level. However, due to long-term reclamation and overutilization, and climate change, Phaeozems became thinner and

harder, threatening the sustainable development of agriculture (Liu et al., 2006; Ou et al., 2017). Therefore, in order to manage these soils accurately and reasonably appropriate land management practice policies should be taken.

In addition, we found that the SOC sequestration potential had similar spatial distribution characteristics with the LCD. The largest SOC sequestration potential was usually concentrated in the area with the longest LCD period, which was mainly distributed in the central part of the study region (Fig. 5). With the increase of reclamation years, the SOC content in soil showed a decreased trend. In cultivated soil, the effects of tillage and other agricultural activities would affect soil water status and microbial activity (Ramankutty and Foley, 1999), leading to the destruction of physical protective layer and accelerating organic matter decomposition. Due to the increase of microbial activity, increasing soil respiration, SOC decomposition rate and SOC mineralization rate decreases SOM content in cultivated soils (Post and Kwon, 2000; Jiang et al., 2014). The SOC sequestration potential of different LCD was 1.03 × 10⁷ t ha⁻¹ to 24.94 × 10⁷ t ha⁻¹, while the SOC sequestration potential of T2 was the lowest and T4 was the highest (Table 3). We found that with the increase of LCD, the SOC sequestration potential increased first and then decreased. Our analysis showed that T4 (70–120 years) was the stage before and after the establishment of new China. To date, a large number of forest land and grassland has been reclaimed for cultivated land, and the corresponding supplement to organic substances was relatively small, resulting in a large loss of SOC and the largest SOC sequestration potential. With the implementation of conservation tillage and black land protection policies in recent years, this trend has been effectively minimized (Zhang et al., 2016; Ou et al., 2017).

4.3. Model performance

This study used an empirical formula based on long-term field experimental data combined with machine learning algorithms—BRT model to predict SOC sequestration potential in topsoil cultivated land ecosystems in Northeast China. The calculation of SOC sequestration potential was based on the difference between saturated SOC and SOC density. The BRT model has high predictive performance in predicting saturated SOC and SOC density, which explains 79% and 85% of the spatial variations of saturated SOC and SOC density within the region (Table 2). Traditional SOC sequestration potential estimation was often based on a large amount of long-term field experimental data, which were used to estimate carbon sequestration rate under one or several management measures, then extrapolating it to estimate regional or global-scale SOC sequestration potentials (Heenan et al., 1995; Mishra and Riley, 2012). These methods often ignore the impact of environmental factors on SOC sequestration potential, leading to increased uncertainty in potential estimation (Qin and Huang, 2010). On the other hand, our study introduces a DSM technology that can reduce the

quantification uncertainty. Chen et al. (2018) also used the same strategy to predict the topsoil (0–30 cm) and subsoil (30–50 cm) of the SOC sequestration potential in different ecosystems using the Hassink equation and the regression kriging method in France. Our study focused on cultivated land ecosystems, using long-term agricultural experimental data to estimate saturated SOC, and ultimately obtain the topsoil SOC sequestration potential in Northeast China.

4.4. Uncertainties in the present study

Although the extrapolation algorithm based on long-term field experimental data and BRT model well predicted saturated SOC and SOC density, there were some uncertainties mainly coming from five different sources: 1) the saturated SOC model was obtained from Qin and Huang (2010), which used 76 long-term field experimental data covering most agricultural areas in the world to build the model. There might be some systematic differences between experimental conditions, research methods, and data quality reported in the literature. Although some measures were taken to improve the data quality, there were still some uncertainties; 2) the spatial climate data were obtained from interpolation of meteorological station data, which does not truly represent the actual meteorological conditions; 3) due to the lack of annual irrigation data in Northeast China, only the MAP rather than the annual water supply was used to calculate saturated SOC, which might lead to the underestimation of saturated SOC; 4) although the 90 m spatial resolution is relatively fine, it still cannot capture the terrain details in the area with large terrain fluctuation, which would inevitably lead to the deviation of saturated SOC simulation; 5) due to the lack of bulk density data at some sampling points, this study used the Pedo-Transfer Functions (PTFs) to complete the bulk density data at these points. Because the data were not from the actual measurement, it might cause the error of SOC density simulation.

5. Conclusions

Combining the extrapolation algorithm, long-term field experimental data, and a BRT model, we predicted saturated SOC, SOC density, and SOC sequestration potential of cultivated topsoil (0–20 cm) in Northeast China. The model presented lower systematic MAE and RMSE, higher R^2 and LCCC. With terrain, climate and LCD as predictors, the models had high prediction performance and explained the spatial variation of 81% saturated SOC and 85% SOC density, respectively. MAT and MAP were the most important environmental factors affecting saturated SOC and SOC density. In addition, LCD was an important environmental factor. With the increase of reclamation period, saturated SOC, SOC density and SOC sequestration potential all increased first and then decreased. The spatial distribution pattern of saturated SOC and SOC density was similar, and gradually decreased from southeast to northwest. SOC sequestration potential had similar spatial distribution characteristics to the LCD. At present, due to over reclamation and utilization, climate change and other factors, the topsoil of cultivated land has become thinner and harder posing challenges to sustainable agricultural development. Scientific and accurate spatial data of SOC sequestration potential is the premise for formulating reasonable carbon management measures. Similarly, accurately predicting the spatial distribution of topsoil SOC sequestration potential and identifying its controlling factors are of great significance in studying the regional carbon cycle, soil fertility maintenance, which will facilitate decision-making for cultivated land ecosystem environmental planning in the region.

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.

Data availability

The authors do not have permission to share data.

Acknowledgment

This work was funded by the National Key R&D Program of China (2021YFD1500200), National Natural Science Foundation of China (Grant No.42077149), China Postdoctoral Science Foundation (Grant No. 2019M660782 and 2020M680983), Basic Scientific Research Projects of Liaoning Provincial Department of Education (Grant No. LJKMZ20220996), National Science and Technology Basic Resources Survey Program of China (Grant No. 2019FY101300), Doctoral research start-up fund project of Liaoning Provincial Department of Science and Technology (Grant No. 2020-BS-136 and 2021-BS-136), and China Scholarship Council (201908210132).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.geoder.2023.e00655>.

References

- Adhikari, K., Kheir, R.B., Greve, M.B., Böcher, P.K., Malone, B.P., Minasny, B., Greve, M. H., 2013. High-resolution 3-D mapping of soil texture in Denmark. *Soil Sci. Soc. Am. J.* 77 (3), 860–876.
- Adhikari, K., Owens, P.R., Libohova, Z., Miller, D.M., Wills, S.A., Nemecek, J., 2019. Assessing soil organic carbon stock of Wisconsin, USA and its fate under future land use and climate change. *Sci. Total Environ.* 667, 833–845.
- Aguilera, E., Lassaletta, L., Gattinger, A., Gimeno, B.S., 2013. Managing soil carbon for climate change mitigation and adaptation in Mediterranean cropping systems: ametaanalysis. *Agric. Ecosyst. Environ.* 168, 25–36.
- Akpa, S.I., Odeh, I.O., Bishop, T.F., Hartemink, A.E., Amapu, I.Y., 2016. Total soil organic carbon and carbon sequestration potential in Nigeria. *Geoderma* 271, 202–215.
- Allmaras, R.R., Schomberg, H.H., Douglas, C.L., Dao, T.H., 2000. Soil organic carbon sequestration potential of adopting conservation tillage in US croplands. *J. Soil Water Conserv.* 55 (3), 365–373.
- Amelung, W., Bossio, D., de Vries, W., Kögel-Knabner, I., Lehmann, J., Amundson, R., Bo, R., Collins, C., Lal, R., Leffeld, J., Minasny, B., Pan, G., Paustian, K., Rumpel, C., Sanderman, J., van Groenigen, J.W., Mooney, S., van Wesemael, B., Wander, M., Chabbi, A., 2020. Towards a global-scale soil climate mitigation strategy. *Nat. Commun.* 11 (1), 1–10.
- Bae, J., Ryu, Y., 2015. Land use and land cover changes explain spatial and temporal variations of the soil organic carbon stocks in a constructed urban park. *Landsc. Urban Plan.* 136, 57–67.
- Bedard-Haughn, A., Jongbloed, F., Akkerman, J., Uijl, A., De Jong, E., Yates, T., Pennock, D., 2006. The effects of erosional and management history on soil organic carbon stores in ephemeral wetlands of hummocky agricultural landscapes. *Geoderma* 135, 296–306.
- Brus, D.J., Yang, R.M., Zhang, G.L., 2016. Three-dimensional geostatistical modeling of soil organic carbon: a case study in the Qilian Mountains, China. *Catena* 141, 46–55.
- Chen, S., Martin, M.P., Saby, N.P., Walter, C., Angers, D.A., Arrouays, D., 2018. Fine resolution map of top-and subsoil carbon sequestration potential in France. *Sci. Total Environ.* 630, 389–400.
- Coleman, K., Jenkinson, D.S., Crocker, G.J., Grace, P.R., Klir, J., Körschens, M., Richter, D.D., 1997. Simulating trends in soil organic carbon in long-term experiments using RothC-26.3. *Geoderma* 81, 29–44.
- Conforti, M., Lucà, F., Scarciglia, F., Matteucci, G., Buttafuoco, G., 2016. Soil carbon stock in relation to soil properties and landscape position in a forest ecosystem of southern Italy (Calabria region). *Catena* 144, 23–33.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.* 8 (7), 1991–2007.
- Cui, Y., Ma, J., Liu, X., Duan, Z., 2011. Effects of human activities on soil organic carbon pool. *J. Desert Res.* 31 (2), 407–414 (In Chinese, English abstract).
- Denich, M., Kanashiro, M., Vlek, P.L.G., 2019. The potential and dynamics of carbon sequestration in traditional and modified fallow systems of the eastern Amazon region, Brazil. In: *Global Climate Change and Tropical Ecosystems*. CRC Press, pp. 213–229.
- Dorji, T., Odeh, I.O., Field, D.J., Baillie, I.C., 2014. Digital soil mapping of soil organic carbon stocks under different land use and land cover types in montane ecosystems, Eastern Himalayas. *For. Ecol. Manag.* 318, 91–102.
- Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. *J. Anim. Ecol.* 77 (4), 802–813.
- Fantappiè, M., L'Abate, G., Costantini, E.A.C., 2011. The influence of climate change on the soil organic carbon content in Italy from 1961 to 2008. *Geomorphology* 135 (3–4), 343–352.

- Farage, P.K., Ardö, J., Olsson, L., Rienzi, E.A., Ball, A.S., Pretty, J.N., 2007. The potential for soil carbon sequestration in three tropical dryland farming systems of Africa and Latin America: a modelling approach. *Soil TillRes.* 94 (2), 457–472.
- Fissore, C., Dalzell, B.J., Berhe, A.A., Voegtli, M., Evans, M., Wu, A., 2017. Influence of topography on soil organic carbon dynamics in a Southern California grassland. *Catena* 149, 140–149.
- Giardina, C.P., Ryan, M.G., 2000. Evidence that decomposition rates of organic carbon in mineral soil do not vary with temperature. *Nature* 404 (6780), 858–861.
- Guo, Z., Adhikari, K., Chellasamy, M., Greve, M.B., Owens, P.R., Greve, M.H., 2019. Selection of terrain attributes and its scale dependency on soil organic carbon prediction. *Geoderma* 340, 303–312.
- Heenan, D.P., McGhie, W.J., Thomson, F.M., Chan, K.Y., 1995. Decline in soil organic carbon and total nitrogen in relation to tillage, stubble management, and rotation. *Aust. J. Exp. Agr.* 35 (7), 877–884.
- Hontoria, C., Saa, A., Rodríguez-Murillo, J.C., 1999. Relationships between soil organic carbon and site characteristics in peninsular Spain. *Soil Sci. Soc. Am. J.* 63 (3), 614–621.
- Hu, Y., Zhang, Q., Hu, S., Xiao, G., Zhang, L., Han, L., Ye, P., 2022. Crop yield and carbon sink potential with precipitation in maize and potato cropland ecosystems over the summertime monsoon transition zone of China. *Soil Use Manag.* 39 (2), 12878.
- IUSS Working Group WRB, 2006. World reference base for soil resources 2006. In: *World Soil Resources Reports No. 103*. FAO, Rome.
- Izaurrealde, R.C., Williams, J.R., McGill, W.B., Rosenberg, N.J., Jakas, M.Q., 2006. Simulating soil C dynamics with EPIC: model description and testing against long-term data. *Ecol. Model.* 192, 362–384.
- Jiang, G., Xu, M., He, X., Zhang, W., Huang, S., Yang, X., Murphy, D.V., 2014. Soil organic carbon sequestration in upland soils of northern China under variable fertilizer management and climate change scenarios. *Global. Biogeochem. Cy.* 28 (3), 319–333.
- Kane, D., Solutions, L.L.C., 2015. Carbon sequestration potential on agricultural lands: a review of current science and available practices. In: *National Sustainable Agriculture Coalition Breakthrough Strategies and Solutions*. LLC, pp. 1–35.
- Kelly, R.H., Parton, W.J., Crocker, G.J., Graced, P.R., Klir, J., Körschens, M., Richter, D. D., 1997. Simulating trends in soil organic carbon in long-term experiments using the CENTURY model. *Geoderma* 81, 75–90.
- Kirkby, M.J., 1999. catchment (1000 km²) scales. In *Regionalization in Hydrology: Proceedings of an International Conference Held at the Technical University of Braunschweig, Germany, 10-14 March 1997* (No. 254, p. 1). IAHS.
- Lal, R., 2002. Soil carbon sequestration in China through agricultural intensification, and restoration of degraded and desertified ecosystems. *Land Degrad. Dev.* 13, 469–478.
- Lal, R., 2004. Soil carbon sequestration impacts on global climate change and food security. *Science* 304, 1623–1627.
- Li, S., Li, J., Li, C., Huang, S., Li, X., Li, S., Ma, Y., 2016. Testing the RothC and DNDC models against long-term dynamics of soil organic carbon stock observed at cropping field soils in North China. *Soil Tillage Res.* 163, 290–297.
- Lin, L., 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45, 255–268.
- Liu, D., Wang, Z., Zhang, B., Song, K., Li, X., Li, J., Duan, H., 2006. Spatial distribution of soil organic carbon and analysis of related factors in croplands of the black soil region. *Northeast China Agr. Ecosyst. Environ.* 113 (1–4), 73–81.
- Lorenz, K., Lal, R., 2005. The depth distribution of soil organic carbon in relation to land use and management and the potential of carbon sequestration in subsoil horizons. *Adv. Agron.* 88, 35–66.
- Lu, F., Wang, X., Han, B., Ouyang, Z., Duan, X., Zheng, H., Miao, H., 2009. Soil carbon sequestrations by nitrogen fertilizer application, straw return and no-tillage in China's cropland. *Glob. Chang. Biol.* 15, 281–305.
- Luo, Z., Feng, W., Luo, Y., Baldock, J., Wang, E., 2017. Soil organic carbon dynamics jointly controlled by climate, carbon inputs, soil properties and soil carbon fractions. *Glob. Chang. Biol.* 23 (10), 4430–4439.
- Malone, B., McBratney, A., Minasny, B., 2011. Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes. *Geoderma* 160 (3–4), 614–626.
- Martin, M.P., Wattenbach, M., Smith, P., Meersmans, J., Jolivet, C., Boulonne, L., Arrouays, D., 2011. Spatial distribution of soil organic carbon stocks in France. *Biogeosciences* 8, 1053–1065.
- McBratney, A.B., Santos, M.L.M., Minasny, B., 2003. On digital soil mapping. *Geoderma* 117, 3–52.
- McCaffrey, D.F., Ridgeway, G., Morral, A.R., 2004. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychol. Methods* 9 (4), 403.
- McKeague, J.A., 1978. Manual on soil sampling and methods of analysis. *Can. Soc. Soil. Sci.* 212.
- Mikhaylov, A., Moiseev, N., Aleshin, K., Burkhardt, T., 2020. Global climate change and greenhouse effect. *Entrep. Sustain. Iss.* 7 (4), 2897.
- Mishra, U., Riley, W.J., 2012. Alaskan soil carbon stocks: spatial variability and dependence on environmental factors. *Biogeosciences* 9 (9), 3637–3645.
- Muñoz-Rojas, M., Jordán, A., Zavala, L.M., De la Rosa, D., Abd-Elmabod, S.K., AnayaRomero, M., 2015. Impact of land use and land cover changes on organic carbon stocks in Mediterranean soils (1956–2007). *Land Degrad. Dev.* 26 (2), 168–179.
- Nelson, D.W., Sommers, L.E., 1996. Total Carbon, Organic Carbon, and Organic Matter. *Methods of Soil Analysis Part 3—Chemical Methods, (methodsofsoil3)*, pp. 961–1010.
- Novara, A., Sarno, M., Gristina, L., 2021. No till soil organic carbon sequestration could be overestimated when slope effect is not considered. *Sci. Total Environ.* 757, 143758.
- Ogle, S.M., Jay Breidt, F., Eve, M.D., Paustian, K., 2003. Uncertainty in estimating land use and management impacts on soil organic carbon storage for US agricultural lands between 1982 and 1997. *Glob. Chang. Biol.* 9 (11), 1521–1542.
- Ou, Y., Rousseau, A.N., Wang, L., Yan, B., 2017. Spatio-temporal patterns of soil organic carbon and pH in relation to environmental factors—a case study of the black soil region of Northeastern China. *Agric. Ecosyst. Environ.* 245, 22–31.
- Padarian, J., Minasny, B., McBratney, A.B., 2019. Using deep learning for digital soil mapping. *Soil* 5 (1), 79–89.
- Podwojewski, P., Poulenard, J., Nguyet, M.L., de Rouw, A., Nguyen, V.T., Pham, Q.H., Tran, D.T., 2011. Climate and vegetation determine soil organic matter status in an alpine inner-tropical soil catena in the Fan Si Pan Mountain. *Vietnam Catena.* 87, 226–239.
- Post, W.M., Kwon, K.C., 2000. Soil carbon sequestration and land-use change: processes and potential. *Glob. Chang. Biol.* 6 (3), 317–327.
- Qin, Z., Huang, Y., 2010. Quantification of soil organic carbon sequestration potential in cropland: a model approach. *Sci. China Life Sci.* 53 (7), 868–884.
- Qin, Z., Huang, Y., Zhuang, Q., 2013. Soil organic carbon sequestration potential of cropland in China. *Global Biogeochem. Cy.* 27 (3), 711–722.
- R Development Core Team, 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna (Austria).
- Ramankutty, N., Foley, J.A., 1999. Estimating historical changes in global land cover: croplands from 1700 to 1992. *Glob. Biogeochem. Cycles* 13 (4), 997–1027.
- Reyes Rojas, L.A., Adhikari, K., Ventura, S.J., 2018. Projecting soil organic carbon distribution in Central Chile under future climate scenarios. *J. Environ. Qual.* 47 (4), 735–745.
- Rogasik, J., Schroetter, S., Funder, U., Schnug, E., Kurtinez, P., 2004. Long-term fertilizer experiments as a data base for calculating the carbon sink potential of arable soils. *Arch. Agron. Soil Sci.* 50, 11–19.
- Schjønning, P., de Jonge, L.W., Olesen, J.E., Greve, M.H., 2009. Organic farming effects on clay dispersion in carbon-exhausted soils. *ICROFS News* 4, 8–10.
- Schulp, C.J.E., Verborg, P.H., 2009. Effect of land use history and site factors on spatial variation of soil organic carbon across a physiographic region. *Agric. Ecosyst. Environ.* 133 (1), 86–97.
- Singh, B.R., Lal, R., 2005. The potential of soil carbon sequestration through improved management practices in Norway. *Environ. Develop. Sust.* 7, 161–184.
- Smith, P., Smith, J.U., Powlson, D.S., McGill, W.B., Arah, J.R.M., Chertov, O.G., Whitmore, A.P., 1997. A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments. *Geoderma* 81, 153–225.
- Sumfleth, K., Duttman, R., 2008. Prediction of soil property distribution in paddy soil landscapes using terrain data and satellite information as indicators. *Ecol. Indic.* 8 (5), 485–501.
- Wadoux, A.M.C., Minasny, B., McBratney, A.B., 2020. Machine learning for digital soil mapping: applications, challenges and suggested solutions. *Earth Sci. Rev.* 210, 103359.
- Walling, E., Vaneckhaute, C., 2020. Greenhouse gas emissions from inorganic and organic fertilizer production and use: a review of emission factors and their variability. *J. Environ. Manag.* 276, 111211.
- Wang, S., Wang, X., Ouyang, Z., 2011. Effects of land use, climate, topography and soil properties on regional soil organic carbon and total nitrogen in the upstream watershed of Miyun Reservoir. *North China J. Environ. Sci.* 24 (3), 387–395.
- Wang, S., Zhuang, Q., Wang, Q., Jin, X., Han, C., 2017. Mapping stocks of soil organic carbon and soil total nitrogen in Liaoning Province of China. *Geoderma* 305, 250–263.
- Wang, S., Adhikari, K., Wang, Q., Jin, X., Li, H., 2018. Role of environmental variables in the spatial distribution of soil carbon (C), nitrogen (N), and C: N ratio from the northeastern coastal agroecosystems in China. *Ecol. Indic.* 84, 263–272.
- Wang, Y., Wang, S., Adhikari, K., Wang, Q., Sui, Y., Xin, G., 2019a. Effect of cultivation history on soil organic carbon status of arable land in northeastern China. *Geoderma* 342, 55–64.
- Wang, S., Zhuang, Q., Yang, Z., Yu, N., Jin, X., 2019b. Temporal and spatial changes of soil organic carbon stocks in the forest area of northeastern China. *Forests* 10 (11), 1023.
- Wang, S., Zhou, M., Adhikari, K., Zhuang, Q., Bian, Z., Wang, Y., Jin, X., 2022. Anthropogenic controls over soil organic carbon distribution from the cultivated lands in Northeast China. *Catena* 210, 105897.
- Webster, K.L., Creed, I.F., Beall, F.D., Bourbonnière, R.A., 2011. A topographic template for estimating soil carbon pools in forested catchments. *Geoderma* 160 (3–4), 457–467.
- Wiesmeier, M., Spörlein, P., Geuß, U.W.E., Hangen, E., Haug, S., Reischl, A., Kögel-Knabner, I., 2012. Soil organic carbon stocks in Southeast Germany (Bavaria) as affected by land use, soil type and sampling depth. *Glob. Chang. Biol.* 18 (7), 2233–2245.
- Wiesmeier, M., Barthold, F., Spörlein, P., Geuß, U., Hangen, E., Reischl, A., Schilling, B., Angst, G., von Lützwow, M., Kögel-Knabner, I., 2014a. Estimation of total organic carbon storage and its driving factors in soils of Bavaria (Southeast Germany). *Geoderma. Reg.* 1, 67–78.
- Wiesmeier, M., Hübner, R., Spörlein, P., Geuß, U., Hangen, E., Reischl, A., Kögel-Knabner, I., 2014b. Carbon sequestration potential of soils in Southeast Germany derived from stable soil organic carbon saturation. *Glob. Chang. Biol.* 20 (2), 653–665.
- Xu, E., Zhang, H., Xu, Y., 2020. Exploring land reclamation history: soil organic carbon sequestration due to dramatic oasis agriculture expansion in arid region of Northwest China. *Ecol. Indic.* 108, 105746.
- Yan, H., Cao, M., Liu, J., Tao, B., 2007. Potential and sustainability for carbon sequestration with improved soil management in agricultural soils of China. *Agric. Ecosyst. Environ.* 121, 325–335.

- Yang, R.M., Zhang, G.L., Liu, F., Lu, Y.Y., Yang, F., Yang, F., Li, D.C., 2016. Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an alpine ecosystem. *Ecol. Indic.* 60, 870–878.
- Yu, G., Fang, H., Gao, L., Zhang, W., 2006. Soil organic carbon budget and fertility variation of black soils in Northeast China. *Ecol. Res.* 21 (6), 855–867.
- Zhang, Y., Zhu, L., Wang, J., Wang, J., Su, B., Zhang, C., Li, C., 2016. Biodegradation of endosulfan by bacterial strain *alcaligenes faecalis* JBW4 in argi-udic ferrosols and Hapli-udic Isohumosols. *Water Air Soil Pollut.* 227 (11), 1–10.
- Zhang, X., Li, X., Ji, X., Zhang, Z., Zhang, H., Zha, T., Jiang, L., 2021. Elevation and total nitrogen are the critical factors that control the spatial distribution of soil organic carbon content in the shrubland on the Bashang Plateau. *China Catena.* 204, 105415.
- Zhu, A.X., Yang, L., Li, B., Qin, C., English, E., Burt, J.E., Zhou, C., 2008. Purposive sampling for digital soil mapping for areas with limited data. In: Hartemink, A.E. (Ed.), *Digital Soil Mapping with Limited Data*. Springer, pp. 33–245.