Forest Ecosystems 10 (2023) 100101

Contents lists available at ScienceDirect

Forest Ecosystems

journal homepage: www.keaipublishing.com/cn/journals/forest-ecosystems

Spatial-temporal variations and driving factors of soil organic carbon in forest ecosystems of Northeast China

Shuai Wang [a,b,c], Bol Roland [c,d], Kabindra Adhikari [e], Qianlai Zhuang [f], Xinxin Jin [a,*], Chunlan Han [a,**], Fengkui Qian [a]

[a] College of Land and Environment, Shenyang Agricultural University, Shenyang, 110866, China
[b] Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing, 100101, China
[c] Institute of Bio- and Geosciences, Agrosphere (IBG-3), Forschungszentrum Jülich GmbH, Wilhelm-Johnen-Straße, 52428, Jülich, Germany
[d] School of Natural Sciences, Environment Centre Wales, Bangor University, Bangor, LL57 2UW, UK
[e] USDA-ARS, Grassland, Soil and Water Research Laboratory, Temple, TX, 76502, USA
[f] Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN, 47907, USA

ARTICLE INFO

Keywords:
Soil organic carbon stocks
Forest ecosystem
Spatial-temporal variation
Carbon sink
Digital soil mapping

ABSTRACT

Forest soil carbon is a major carbon pool of terrestrial ecosystems, and accurate estimation of soil organic carbon (SOC) stocks in forest ecosystems is rather challenging. This study compared the prediction performance of three empirical model approaches namely, regression kriging (RK), multiple stepwise regression (MSR), random forest (RF), and boosted regression trees (BRT) to predict SOC stocks in Northeast China for 1990 and 2015. Furthermore, the spatial variation of SOC stocks and the main controlling environmental factors during the past 25 years were identified. A total of 82 (in 1990) and 157 (in 2015) topsoil (0–20 cm) samples with 12 environmental factors (soil property, climate, topography and biology) were selected for model construction. Randomly selected 80% of the soil sample data were used to train the models and the other 20% data for model verification using mean absolute error, root mean square error, coefficient of determination and Lin's consistency correlation coefficient indices. We found BRT model as the best prediction model and it could explain 67% and 60% spatial variation of SOC stocks, in 1990, and 2015, respectively. Predicted maps of all models in both periods showed similar spatial distribution characteristics, with the lower SOC in northeast and higher SOC in southwest. Mean annual temperature and elevation were the key environmental factors influencing the spatial variation of SOC stock in both periods. SOC stocks were mainly stored under Cambosols, Gleyosols and Isohumosols, accounting for 95.6% (1990) and 95.9% (2015). Overall, SOC stocks increased by 471 Tg C during the past 25 years. Our study found that the BRT model employing common environmental factors was the most robust method for forest topsoil SOC stocks inventories. The spatial resolution of BRT model enabled us to pinpoint in which areas of Northeast China that new forest tree planting would be most effective for enhancing forest C stocks. Overall, our approach is likely to be useful in forestry management and ecological restoration at and beyond the regional scale.

1. Introduction

Forest plays an irreplaceable role in maintaining global carbon balance, and has become a central component of the earth system for global climate change research (Bradshaw and Warkentin, 2015). Forest ecosystems are the largest carbon pool on the earth except the ocean (Dixon et al., 1994; Liu et al., 2012; Harper et al., 2018), and the pool includes forest soil carbon and vegetation carbon. It accounts for about 50% of the total carbon of the whole terrestrial ecosystem, of which 2/3 of the carbon is stored in forest soils (Bockheim and Gennadiyev, 2010; Carvalhais et al., 2014; Scharlemann et al., 2014). At present, the estimation of forest soil organic carbon (SOC) stocks has great uncertainty leading to the estimation error of forest soil annual carbon emissions to the atmosphere as high as 10 Pg C, which is much higher than the total industrial emissions of 5.3 Pg C (Xu, 1994; Korhonen et al., 2001; Lynn and Peeva, 2021). Therefore, accurately quantifying the carbon pool and sequestration capacity of forest soils not only provides scientific basis and data for the global carbon balance and climate change studies, but also helps promote ecosystem services delivery, including forest carbon sink.

* Corresponding author.
** Corresponding author.
E-mail addresses: jinxinxin0218@syau.edu.cn (X. Jin), hancly@syau.edu.cn (C. Han).

https://doi.org/10.1016/j.fecs.2023.100101
Received 3 September 2022; Received in revised form 10 February 2023; Accepted 10 February 2023

Abbreviations

(SOC)	soil organic carbon
(RK)	regression kriging
(MSR)	multiple stepwise regression
(RF)	random forest
(BRT)	boosted regression trees
(MMA)	measure and multiply model
(SLM)	soil landscape modeling
(DSM)	digital soil mapping
(BD)	bulk density
(PTFs)	Pedo-Transfer Functions
(NDVI)	normalized difference vegetation index
(MAP)	mean annual precipitation
(MAT)	mean annual temperature
(TWI)	topographic wetness index
(DEM)	digital elevation model
(SAGA)	system for automated geoscientific analysis
(GIS)	geographic information system
(ELE)	elevation
(SG)	slope gradient
(SA)	slope aspect
(PC)	profile curvature
(CA)	catchment area
(MAE)	mean absolute error
(LCCC)	Lin's concordance correlation coefficient
(RI)	relative importance

The spatio-temporal variability of SOC stocks can be studied through different technologies, which can be grouped into two categories, namely, the measure and multiply model (MMA) and soil landscape modeling (SLM) model (Hengl et al., 2004; Martin et al., 2011; Wang et al., 2018). In the MMA model, the average SOC stocks of each soil type or land use type were calculated and assigned to each map unit to obtain the spatial-temporal distribution map of regional SOC stocks. However, this method makes each map unit had a constant value and cannot capture the spatial variability of SOC stocks within the map unit, and there was a large estimation error (Qi et al., 2019; Wang et al., 2019). In contrast, in order to establish and improve the basic database of forest SOC dynamics and improve the systematicness and comparability of the data, the SLM model based on digital soil mapping (DSM) technology was widely used to improve the prediction accuracy (McBratney et al., 2003). DSM techniques quantitatively describe the relationship between soil and environment and predict soil attributes based on 'scorpan' (i.e., climatic factors, biological factors, topographic factors, parent material, time factors, soil factors and spatial factors) concept of soil formation and distribution (McBratney et al., 2003). This approach has been widely used to establish various DSM methods and predict soil attributes including SOC. The methods include geographically weighted regression model (Kumar et al., 2012), support vector machine (Were et al., 2015), artificial neural network (Li et al., 2013), random forest (Were et al., 2015; Nabiollahi et al., 2019), cubist (Adhikari et al., 2014), linear hybrid model (Liu et al., 2020), similarity-based model (Wang et al., 2020a), and boosted regression trees (Martin et al., 2011; Wang et al., 2019), among others.

Among different DSM models, tree-based models are more common and have been widely used in soil properties predictions including SOC prediction and mapping (Padarian et al., 2019; Hateffard et al., 2019). Unlike traditional models, the tree-based models are more effective with better prediction performance (Wang et al., 2019; Hateffard et al., 2019; Ebrahimy et al., 2020), and boosted regression trees (BRT) is one of such examples. The BRT model is relatively new model and has been found more powerful and efficient in predicting SOC (Martin et al., 2011). It can flexibly deal with multicollinearity among variables, data types, and model over-fitting (Keskin et al., 2019). Based on these advantages, BRT model has been successfully applied in various disciplines, including epidemiology (Lampa et al., 2014), remote sensing (Colin et al., 2017), soil science (Martin et al., 2011), marine science (Navarro et al., 2020), ecology science (Elith and Leathwick, 2017), and environmental science (Ebrahimy et al., 2020). However, BRT model is rarely used to study the temporal and spatial changes of topsoil SOC in forest ecosystems (Wang et al., 2019).

This study thus aims to evaluate the application of BRT model to predict and map spatial-temporal variation of topsoil (0–20 cm) SOC stocks in forest ecosystems of Northeast China, and identify the major driving factors of such variation. The specific objectives were to: (1) predict and map the spatial distribution of SOC stocks in 1990 and in 2015; (2) analyze the spatial pattern of SOC variations in forest ecosystems in the past 25 years; and (3) identify key environmental factors affecting the spatial-temporal variations of SOC stocks.

2. Materials and methods

2.1. Study area

The study was conducted in one of the largest natural forest regions in China that mostly occupies the Greater Khingan Mountains, Lesser Khingan Mountains and Changbai Mountain (Liaoning, Jilin and Heilongjiang provinces) (Fig. 1). The area covers about 305,000 km^2 of forests that mainly consists of coniferous and deciduous broad-leaved mixed forests accounting for 37% of the total forest area in China (China National Bureau of Statistics, 2015). The timber volume reached 3.2 billion cubic meters, accounting for 1/3 of the total timber volume in China. The main tree species are Pinaceae, Taxodiaceae, Cupressaceae and Ginkgoaceae (Zhu et al., 1990). The Greater Khingan Mountains

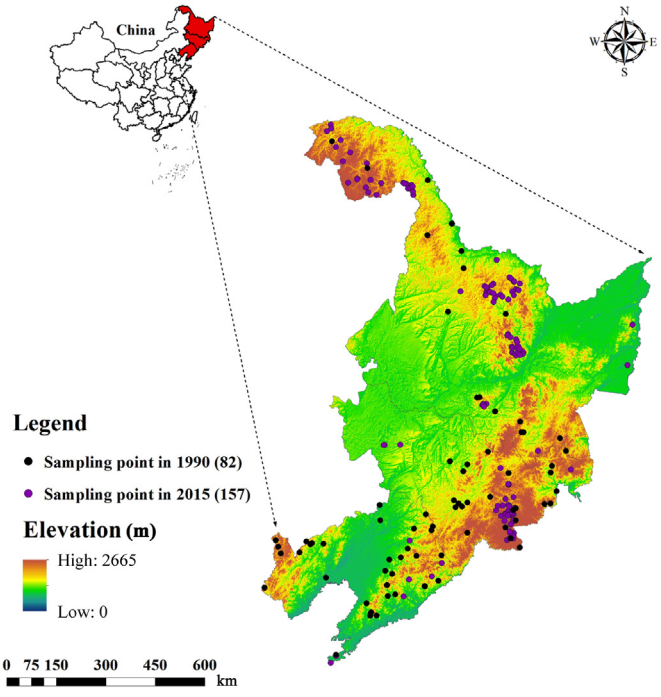

Fig. 1. Location of study area and soil sampling sites, which are superimposed on a 90-m resolution digital elevation model.

forest area was dominated by *Larix gmelinii*, accounting for 86.1% of the forest area. About 80% of Korean pine in China is distributed in the Lesser Khingan Mountains, which is the largest and most complete virgin forest in Asia. The main vegetation type in Changbai Mountain forest area is temperate coniferous and broad-leaved mixed forest, with more than 1500 species of higher plants and more than 800 species of plants with greater economic value. The climatic zone ranges from middle temperate zone to cold temperate zone from south to north with temperate monsoon climate as the main climate. However, due to the high latitude in some areas, the winter is cold, long and snowy, the summer is warm and short, with less evapo-transpiration and is humid. From the northwest to the southeast, the mean annual precipitation increased from less than 300 mm to more than 1000 mm. The average temperature in winter is about −20 °C, and the average temperature in summer is more than 25 °C, and the mean annual temperature is between −4 and 11 °C. According to the Chinese Soil Taxonomy (Chinese Soil Taxonomy Research Group, 2001), the main dominant soil types were Cambosols, Gleyosols and Isohumosols, accounting for 82%, 7% and 6% of the total area of the study area respectively. The other soil types accounting for 5% of the total area were Argosols, Primosols, Anthosols, Histosols, Andosols and Halosols.

2.2. Soil sampling

This study only obtained the regions that were forest land in 1990 and 2015 through overlay analysis in ArcGIS 10.2 (ESRI, Redlands, CA, USA), and focused on discussing the spatial and temporal changes of the topsoil forest SOC stocks in the areas where the forest cover did not change in the past 25 years. We then analyzed the main environmental factors that affected SOC spatial variation in both periods (Fig. S1). Due to the limitation in data acquisition, the spatial SOC information from different forest types has not been analyzed. In addition, we did not discuss the conversion of forests to cultivated land, grassland and other types in the past 25 years, because we did not collect soil samples from those land use types in 2015.

2.2.1. Soil survey data in 1990

Soil profile data in 1990 were obtained from the second soil survey database of Liaoning Province, Jilin Province and Heilongjiang Province respectively. It includes a total of 82 soil profile observations covering different soil types in the whole region. As this study only focused on topsoil (0–20 cm) SOC stocks, soil data from the top 20 cm depth were extracted. Not all samples had bulk density (BD) measured, so a Pedo-Transfer Functions (PTFs) (Wösten et al., 2001) (Eq. (1)) was used to supplement the missing BD (26 samples). Previous studies have shown that BD is closely related to SOC and soil texture, as well as to soil depth (Wösten et al., 2001; Yang et al., 2020). We applied a linear regression method to predict BD using SOC measurements (Eq. (1)), and supplemented the missing BD data of 26 samples. R^2 of the BD prediction model was 0.73.

$$BD = 1.57 - 0.09\sqrt{SOC} \quad (1)$$

2.2.2. Soil sampling in 2015

Since the forest area in the study area is large and the topography is rugged, soil sampling in 2015 did not follow the same location in 1990 considering cost, time, and road accessibility. Instead, a purposeful sampling method proposed by Zhu et al. (2008) was adopted. First, the environmental factors closely related to SOC in topsoil including normalized difference vegetation index (NDVI), mean annual precipitation (MAP), mean annual temperature (MAT), topographic wetness index (TWI), and elevation, were selected and pretreated to unify the coordinate system and spatial resolution. Second, the selected environmental factors were clustered with a fuzzy c-means clustering method giving a total of 33 different clusters or landscape units. Third, local soil experts were invited to locate 3–5 soil sampling sites in each landscape unit considering different factors such as road accessibility, and a total of 157 topsoil samples were obtained. A handheld GPS was used to record the longitude and latitude at each site. During the sampling, one kg mixed soil sample and 100 cm^3 undisturbed soil cores were collected at the center of 0–20 cm depth of each sample site to measure SOC and BD. The SOC content was measured with C/N analyzer (Elemental Amerivas Ins. Vario Max, Germany) and bulk density by over drying (at 105 °C for 48 h) and weighing the core samples in the Central Laboratory of Shenyang Agricultural University.

2.3. Calculation of SOC stocks

This study modelled the spatial-temporal variations of SOC stocks in 1990 and 2015. For the soil profile with k layers, Batjes (1996) formula (Eq. 2) was applied to calculate the SOC density in the profile with special depth.

$$SOC_{density} = \sum_{i=1}^{k} SOC_{content} = \sum_{i=1}^{k} SOC_{concentration} \times BD_i \times D_i \times (1 - S_i) \quad (2)$$

where $SOC_{density}$, $SOC_{content}$, and $SOC_{concentration}$ are the SOC density ($kg \cdot m^{-2}$), the SOC content ($kg \cdot m^{-2}$), and the SOC concentration ($g \cdot kg^{-1}$). BD_i, D_i and S_i are the bulk density ($Mg \cdot m^{-3}$), the thickness (m) and the volume fraction of fragments >2 mm, i represents a specific soil layer.

2.4. Environment data

In this study, 12 environmental factors related to four categories (i.e., soil, topography, climate, and biological factors) were selected as predictors of SOC stocks in both periods (1990 and 2015). Because the environmental factors were obtained from different sources or platforms, all environmental data were converted to raster data of 90-m resolution with a unified projection in ArcGIS 10.2 (ESRI, Redlands, CA, USA).

2.4.1. Soil related factors

Soils generate multitude of soil functions to provide ecosystem services (Blum, 2005; Adhikari and Hartemink, 2016), and its variability contributes to changes in biomass production and plant community diversity influencing SOC pools in forest ecosystems (Adhikari et al., 2014; Yang et al., 2020). Therefore, soils are one of the major factors influencing SOC distribution and are widely used as SOC predicting variables (Anderson, 1988; Bockheim et al., 2014; Adhikari et al., 2014; Riza et al., 2021). We used soil texture fractions (clay%, silt%, and sand%) as predictors of SOC in the study area. Soil texture fractions were compiled from a 1:1 million soil type map, and from the soil profile data obtained from the second soil survey database of China (1980s). Texture data at 1-km resolution were downloaded from the Resource and Environmental Science and Data Center (https://www.resdc.cn/), Institute of Geographical Sciences and resources, Chinese Academy of Sciences, and were resampled to 90-m spatial resolution.

2.4.2. Topographic factors

Topographic factors were derived from a 90-m gridded digital elevation model (DEM) downloaded from the Geospatial Data Cloud site, Computer Network Information Center, Chinese Academy of Sciences (http://www.gscloud.cn). A total of 6 topographic factors namely, elevation (ELE), slope gradient (SG), slope aspect (SA), profile curvature (PC), catchment area (CA), and topographic wetness index (TWI) were generated using ArcGIS 10.2 (ESRI, Redlands, CA, USA) and by system for automated geoscientific analysis (SAGA) and geographic information system (GIS) software (Conrad et al., 2015).

2.4.3. Climatic factors

The climate in Northeast China is characterized by warm and rainy summer, lush vegetation growth, and more organic matter entering the soil (Zhang et al., 2016). Accordingly, the cold winter freezes the soil, so

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, the SOC in this region is generally higher than that in other regions of China. This situation is usually caused by the special climate conditions of the region, and mean annual precipitation (MAP) and mean annual temperature (MAT) can better reflect the climate conditions of the region, while the coldest/warmest month of the year only reflect one factor and cannot reflect the climate conditions of the whole year. The MAP and MAT data of 1990 and 2015 from more than 2400 meteorological stations across China were downloaded from the China Meteorological Data Sharing Service Center (http://cdc.cma.gov.cn/), and were interpolated using splines function in ANUSPLIN software (Hutchinson, 1995). The spatial resolution of the MAP and MAT data was 1-km, and it was resampled to 90-m to use in the present study.

2.4.4. Biological factors

This study is only focused on the forest covered areas in 1990 and 2015, so the current land use map could not be used as a biological factor in SOC modeling. Moreover, the predominant forest species map in the two periods in the region is lacking. Therefore, we used normalized difference vegetation index (NDVI) data from remote sensing as a proxy to biological factor of SOC distribution. NDVI detects growth state of vegetation and is related to vegetation coverage (Kumar et al., 2016; Banday et al., 2019; Yang et al., 2020; Dong et al., 2021), the values range between −1.0 and 1.0. A negative NDVI value indicates clouds, water or snow cover, values near zero indicates rock or bare soil, and a positive value indicates vegetation coverage (0.1–0.5 for sparse vegetation, and >0.6 for dense green vegetation). The NDVI data for both were downloaded from the Resource and Environmental Science and Data Center, Institute of Geographical Sciences and resources, Chinese Academy of Sciences (http://www.resdc.cn/) at a spatial resolution of 1-km, and was resampled to 90-m grid for the analysis.

2.5. Prediction models

2.5.1. Regression kriging

Regression kriging (RK) method considers the spatial trend and influence of random factors in the model (Hengl et al., 2004). The spatial trend between SOC stocks and predictor variables is established using a linear regression, and the model residual is interpolated by ordinary kriging. The linear regression equation was established with IBM SPSS statistics-version 23.0 (Kirkpatrick, 2015), and the interpolation was completed in ArcGIS 10.2 software (ESRI, Redlands, CA, USA). The linear trend and the kriging estimate of residual were added to obtain the predicted value of SOC stocks (Eq. (3))

$$Z`(S_0) = m`(S_0) + e`(S_0) = \sum_{k=0}^{p} \beta`_k q_k(S_0) + \sum_{i=0}^{n} \lambda_i e(S_i) \Bigg) \quad (3)$$

where $m`(S_0)$ is the fitting drift; $e`(S_0)$ is the residual of interpolation; β'_k is the estimation of the drift model coefficient; λ_i is the weight of Kriging, which is determined by the independent spatial structure of the residual; $e(S_i)$ is the residual at S_i position.

2.5.2. Multiple stepwise regression

Multiple stepwise regression (MSR) is a widely used regression algorithm, and in essence, it is a multiple linear regression (Krishnan et al., 1980). Regression analysis is used to study the interdependence between multiple factors, while MSR is used to establish the optimal or appropriate regression model, so as to study the dependence between factors more deeply (Zhang et al., 2012).

The basic idea of MSR is to introduce independent variables in the model one by one, and the independent variables introduced each time have the most significant impact on the dependent variable. Each time a new independent variable is introduced, the variable previously introduced into the regression equation shall be tested one by one, and the insignificant independent variable in the current equation shall be eliminated from the independent variable with the least impact on the dependent variable until no new independent variables can be introduced (Zhang et al., 2012). Finally, the independent variables retained in the regression equation have significant impacts on the dependent variable. The MSR models developed for the two periods were:

$$\text{SOC Stocks}_{1990} = 0.992 - 0.673 \times \text{MAT} - 0.534 \times \text{SG} + 13.381 \times \text{NDVI} - 0.077 \times \text{Clay} \quad (4)$$

$$\text{SOC Stocks}_{2015} = -15.776 + 0.004 \times \text{ELE} + 0.475 \times \text{TWI} - 0.458 \times \text{MAT} - 0.534 \times \text{SG} + 18.787 \times \text{NDVI} + 0.123 \times \text{Clay} \quad (5)$$

2.5.3. Random forest

Random forest (RF) is a machine learning algorithm based on classification regression tree proposed by Breiman (2011). This method combines bootstrapping technology and feature random selection technology. RF model includes multiple classification and regression trees to ensure the diversity and stability of the model, which can be used to solve the problems related to classification and regression. RF also be widely used in prediction problems, and it is simple to apply without complex parameter adjustment. The model is optimized by adjusting the number of regression trees and the number of prediction variables at each node of the decision tree to improve prediction accuracy.

In this study, the RF modeling was completed in R language (R Development Core Team, 2013) using the "randomForest" package (Liaw and Wiener, 2002). The model requires users to define three parameters, namely the number of tree (*ntree*), the minimum size in each node (*nodesize*), and the number of variables as a predictor at each tree (*mtry*). The default value of *ntree* is 500. However, more numbers are often needed to obtain more stable prediction results. In the study, the *ntree* was set at 1000. The *nodesize* used the default value of 5, and the *mtry* was set to 6.

2.5.4. Boosted regression trees

The BRT model (Elith et al., 2008) used in this study combines both regression tree and boosting algorithms in the model to improve prediction. Regression part evaluates dependent variables and their predictors with recursive binary splits (Navarro et al., 2020), and boosting combines multiple simple models and improve model performance (Elith and Leathwick, 2017).

BRT has several advantages over other tree-based model as it which can handle different types of predictive variables and allows missing value data (Ebrahimy et al., 2020). It does not require prior data conversion or outlier removal, can fit complex nonlinear relationships, and automatically deals with the interaction between predictors (Wang et al., 2019). It has stronger prediction ability than most traditional methods, and can deal with a large number of practical problems in model fitting (Martin et al., 2011).

The BRT model requires the user to define four parameters, namely, learning rate (LR), tree complexity (TC), bag fraction (BF) and number of trees (NT). LR represents the contribution of each tree in the model to the final fitting model (Martin et al., 2011). TC is the complexity of the tree (Ebrahimy et al., 2020). BF represents the proportion of data used in the dataset (Colin et al., 2017). The optimal combination of parameter settings was tested using a 10-fold cross-validation with LR, TC, BF and NT set at 0.025, 12, 0.75 and 1200 for 1990, 0.025, 12, 0.70 and 1,000, respectively, for 2015.

2.6. Model validation

Martin et al. (2012) believed that the reasonable selection of training

set and test set could improve the prediction performance of the model, and suggested randomly dividing the data into 80% for modeling and the remaining 20% for model verification. Therefore, the prediction performance of RK, MSR, RF and BRT models was evaluated on randomly divided training (80% data) and test (20% data) data sets using four indices namely, mean absolute error (MAE), root-mean-square error (RMSE), coefficient of determination (R^2) and Lin's concordance correlation coefficient (LCCC) (Lin, 1989) (Eqs. (6)–(9)).

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|X_i - Y_i| \tag{6}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)^2} \tag{7}$$

$$R^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{Y})^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} \tag{8}$$

$$\mathrm{LCCC} = \frac{2r\partial_X\partial_Y}{\partial_X^2 + \partial_Y^2 + (\overline{X} + \overline{Y})^2} \tag{9}$$

where X_i and Y_i represent the predicted and measured values; n represents the number of samples; $\overline{X}$ and $\overline{Y}$ represent the average predicted value and the average measured value; r represents the correlation coefficient between the predicted value and the measured value; ∂_X and ∂_Y represent the variance of the prediction set and the measured set.

3. Results

3.1. Descriptive statistics

The boxplot of SOC stocks and corresponding environmental factors at sampling sites in two periods is showed in Fig. 2. In 1990, SOC stocks ranged from 0.79 to 11.28 kg·m^{-2}, with an average of 6.89 ± 1.99 kg C·m^{-2}, and in 2015 it ranged from 3.18 to 16.41 kg C·m^{-2} in 2015, with an average of 8.45 ± 2.32 kg C·m^{-2}. The Pearson's correlation coefficient between SOC stocks and environmental factors at sampling site in the two periods is listed in Table 1. SOC stocks were significantly positively correlated with ELE, MAP and NDVI, and significantly negatively correlated with MAT in the two periods. In addition, there was a significant correlation between SOC stocks and clay and sand in 1990, but no significant correlation in 2015. The climatic factors were significantly correlated with SOC stocks in both periods.

3.2. Model performance and uncertainty

The result of model evaluation (Table 2) showed a decreasing order in performance as BRT > RF > RK > MSR in both periods. Among the three models tested, BRT showed higher R^2 (0.67 vs. 0.60) and LCCC (0.75 vs.

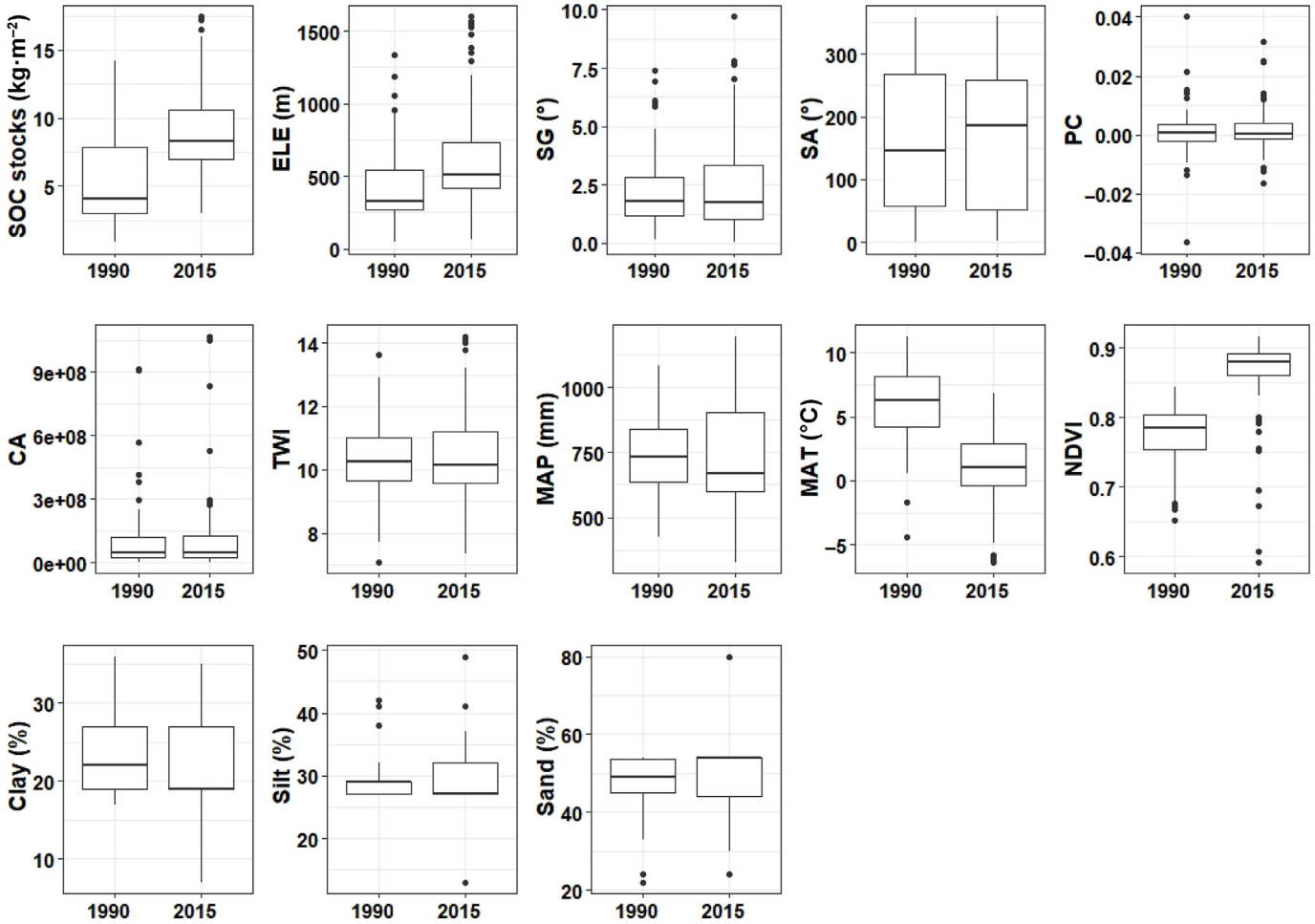

Fig. 2. Boxplot of SOC stocks in 1990 and 2015 derived for different environmental factors. SOC, Soil organic carbon; ELE, elevation; SG, slope gradient; SA, slope aspect; PC, profile curvature; CA, catchment area; TWI, topographic wetness index; MAP, mean annual precipitation; MAT, mean annual temperature; NDVI, normalized difference vegetation index.

Table 1
Relationships between SOC stocks with all environmental predictors in 1990 and 2015 surveys.

Property	SOC stocks	ELE	SG	SA	PC	CA	TWI	MAP	MAT	NDVI	Clay	Silt
1990												
ELE	0.382**											
SG	−0.036	0.400**										
SA	0.051	0.157	−0.012									
PC	−0.107	−0.044	0.412**	−0.137								
CA	−0.056	−0.309**	−0.437**	−0.013	−0.065							
TWI	0.067	−0.454**	−0.827**	0.060	−0.171	0.729**						
MAP	0.223**	−0.039	0.202	−0.202	0.180	−0.313**	−0.353**					
MAT	−0.628**	−0.344**	−0.034	−0.008	0.071	0.054	0.020	0.536**				
NDVI	0.487**	0.157	0.099	−0.110	−0.063	−0.058	−0.129	0.146	−0.442**			
Clay	0.256*	−0.013	−0.219*	−0.102	−0.006	0.221*	0.287**	−0.036	−0.005	0.065		
Silt	0.196	0.075	−0.208	0.048	−0.047	0.270*	0.324**	−0.090	0.003	0.007	0.830**	
Sand	−0.238*	−0.029	0.223*	0.034	0.026	−0.255*	−0.318**	0.064	0.001	−0.040	−0.963**	−0.950**
2015												
ELE	0.274**											
SG	−0.169*	0.164*										
SA	0.034	0.049	0.001									
PC	−0.112	0.037	0.090	−0.010								
CA	−0.198*	−0.247**	−0.416**	0.046	−0.053							
TWI	0.095	−0.265**	−0.799**	−0.015	−0.072	0.740**						
MAP	0.238**	0.688**	0	0.047	0.085	−0.117	−0.068					
MAT	−0.552**	−0.004	−0.197*	−0.016	0.097	0.275**	0.203*	0.542**				
NDVI	0.178*	0.265**	0.188*	−0.014	0.035	−0.462**	−0.280**	0.382**	−0.027			
Clay	−0.133	0.161*	−0.083	−0.102	0.031	−0.059	0.099	0.200*	0.025	−0.040		
Silt	−0.006	0.270**	−0.104	0.032	0.020	−0.089	0.073	0.324**	0.033	0.115	0.674**	
Sand	0.080	−0.233**	0.102	0.042	−0.028	0.080	−0.095	−0.282**	−0.031	−0.037	−0.924**	−0.905**

Note: $p < 0.05$ shown in "*"; $p < 0.01$ shown in "**". SOC, soil organic carbon; ELE, elevation; SG, slope gradient; SA, slope aspect; PC, profile curvature; CA, catchment area; TWI, topographic wetness index; MAP, mean annual precipitation; MAT, mean annual temperature; NDVI, normalized difference vegetation index.

Table 2
Summary statistics of the predictive performance of regression kriging (RK), multiple stepwise regression (MSR), and boosted regression trees (BRT) models in the prediction of SOC stocks in 1990 and 2015 surveys.

Year	Model	Index	Min.	1stQu.	Median	Mean	3rdQu.	Max.
1990	RK	MAE	1.63	1.63	1.63	1.63	1.63	1.63
		RMSE	2.47	2.47	2.47	2.47	2.47	2.47
		R^2	0.62	0.62	0.62	0.62	0.62	0.62
		LCCC	0.68	0.68	0.68	0.68	0.68	0.68
	MSR	MAE	1.64	1.64	1.64	1.64	1.64	1.64
		RMSE	2.49	2.49	2.49	2.49	2.49	2.49
		R^2	0.54	0.54	0.54	0.54	0.54	0.54
		LCCC	0.59	0.59	0.59	0.59	0.59	0.59
	RF	MAE	1.46	1.51	1.52	1.53	1.55	1.58
		RMSE	2.07	2.13	2.15	2.16	2.18	2.23
		R^2	0.56	0.59	0.62	0.63	0.64	0.66
		LCCC	0.63	0.68	0.69	0.71	0.73	0.74
	BRT	MAE	1.37	1.41	1.42	1.43	1.45	1.47
		RMSE	1.95	1.99	2.00	2.01	2.03	2.07
		R^2	0.59	0.66	0.68	0.67	0.68	0.69
		LCCC	0.73	0.74	0.75	0.75	0.75	0.77
2015	RK	MAE	1.55	1.55	1.55	1.55	1.55	1.55
		RMSE	2.24	2.24	2.24	2.24	2.24	2.24
		R^2	0.54	0.54	0.54	0.54	0.54	0.54
		LCCC	0.63	0.63	0.63	0.63	0.63	0.63
	MSR	MAE	1.57	1.57	1.57	1.57	1.57	1.57
		RMSE	2.30	2.30	2.30	2.30	2.30	2.30
		R^2	0.47	0.47	0.47	0.47	0.47	0.47
		LCCC	0.59	0.59	0.59	0.59	0.59	0.59
	RF	MAE	1.35	1.38	1.39	1.41	1.42	1.15
		RMSE	1.89	1.91	1.93	1.94	1.97	1.99
		R^2	0.49	0.52	0.53	0.55	0.56	0.58
		LCCC	0.69	0.71	0.73	0.74	0.75	0.77
	BRT	MAE	1.28	1.29	1.30	1.31	1.32	1.37
		RMSE	1.74	1.78	1.80	1.79	1.81	1.84
		R^2	0.56	0.57	0.59	0.60	0.62	0.65
		LCCC	0.72	0.75	0.76	0.78	0.78	0.79

Note: SOC, soil organic carbon; MAE, mean absolute error; RMSE, root mean squared error; R^2, coefficient of determination; LCCC, Lin's concordance correlation coefficient.

0.78) and lower MAE (1.43 vs. 1.31) and RMSE (2.01 vs. 1.79) explaining 67% and 60% of the spatial variability of SOC stocks in 1990 and 2015, respectively. Furthermore, with respect to the spatial SOC stocks prediction in topsoil forest ecosystem, the BRT model could fully capture it, while RK and MSR models showed the SOC spatial variation which showed a convergence trend with the characteristic of equilibrium development between regions during 1990–2015. The RK and MSR models could not fully describe the temporal variations of forest SOC stocks in the central region between 1990 and 2015, whereas BRT model did provide a more realistic scenario. Previous studies have also obtained similar conclusions (Colin et al., 2017; Ebrahimy et al., 2020; Wang et al., 2020b). Therefore, the BRT model was selected as the final prediction model to map spatial distribution of topsoil SOC stocks in the study areas in both periods. In order to further illustrate the uncertainty of BRT model, the model was iterated 100 times and the standard deviation (SD) of 100 prediction was reported as the uncertainty of model prediction in both periods (Fig. S2). Average SD was 0.36 kg C·m^{-2} in 1990, and 0.28 kg C·m^{-2} in 2015, respectively, indicating a lower prediction uncertainty in both periods.

3.3. Relative importance of environment factors

In order to identify the key environmental factors, we iterated the BRT model for both periods 100 times, and the average relative importance (RI) of each environmental factor was calculated and scaled to 100% (Fig. 3). The results showed that the key environmental factors affecting the spatial variation of SOC stocks were MAT, ELE, CA, and MAP and they accounted for nearly 67% variation in 1990. Correspondingly, MAT, ELE, TWI, and MAP were the key environmental factors affecting SOC stocks, accounting for 65% variations in 2015. We also found that MAT and ELE were the most important environmental factors among the 12 environmental factors used with corresponding RI of 42.9% and 47.5% in 1990 and 2015, respectively.

3.4. Spatial and temporal variation of SOC stocks

The average spatial distribution map of the 100 iterations of BRT models were selected as the final maps of topsoil forest SOC stocks in both periods (Fig. 4). In both periods, the SOC stocks decreased gradually from northwest to southeast of the study area in forest topsoil. In order to further reveal the spatial and temporal variation of SOC stocks, we used the grid calculation module of ArcGIS 10.2 software (ESRI, Redlands, CA, USA) to compare the spatial distribution of SOC stocks in the two periods (Fig. 5a). Then, the change of SOC stocks in six levels were obtained based on cluster analysis (Fig. 5b). Over the past 25 years, SOC stocks did show an upward trend, on average SOC stocks increased by 0.063 kg C·m^{-2}·yr^{-1} (Fig. 5a). The areas with increased SOC stocks were mainly distributed in the southwest and northeast of the studied area. The 25% of the total area with decreased SOC stocks were primary distributed in the central region, and were mostly in the −2.0 to 0 kg C·m^{-2} range of declining SOC stock and accounted for about 23% of the whole study region (Fig. 5).

In order to further explore the spatial-temporal SOC stocks distribution, we quantified the changes of SOC stocks in different soil types between the two periods (Table 3). In the past 25 years, the SOC stocks increased by 471.04 Tg C in topsoil forest ecosystem of Northeast China. The SOC stocks were mainly stored under Cambosols, Gleyosols and Isohumosols in both periods, accounting for 95.6% in 1990 and 95.9% in 2015 of the total stocks, respectively. The decrease of SOC stocks did mainly occur in Andosols and Isohumosols being −0.2 and −6.52 Tg C, respectively in the past 25 years.

4. Discussion

4.1. Estimates of SOC stocks

Forest carbon accumulation can reduce increasing atmospheric CO_2 concentration and plays an irreplaceable role in slowing down global climate change (Lal, 2004). Cao et al. (2003) estimated the carbon flux of China's terrestrial ecosystem by using CEVSA model and showed that China's forest carbon stocks increased gradually at the end of the 20th century. Xu et al. (2007) had showed that China's forest vegetation had been playing an obvious role in CO_2 sink, and the carbon sink function of China's forest vegetation would be further enhanced with the growth of carbon stocks of young and middle-aged forests since the 1980s. Zhao et al. (2009) considered the spatial distribution of SOC stocks of the forest ecosystem as closely related to the interference intensity of human activities on the forest. Similar conclusions were obtained in several other studies (e.g., Wäldchen et al., 2013; Fernández-Romero et al., 2014; Ngaba et al., 2019; Wang et al., 2020b; Su et al., 2021). Wang et al. (2020a) studied the SOC stocks of different ecosystems in Dalian of China, and found that the rapid urbanization and the transformation of land use type were the main reasons for the decline of topsoil SOC stocks. Fernández-Romero et al. (2014) considered a similar conclusion in Jaen, Spain. Our results also showed that the forest ecosystem in Northeast China played a role as carbon “sink” from 1990 to 2015, which was the

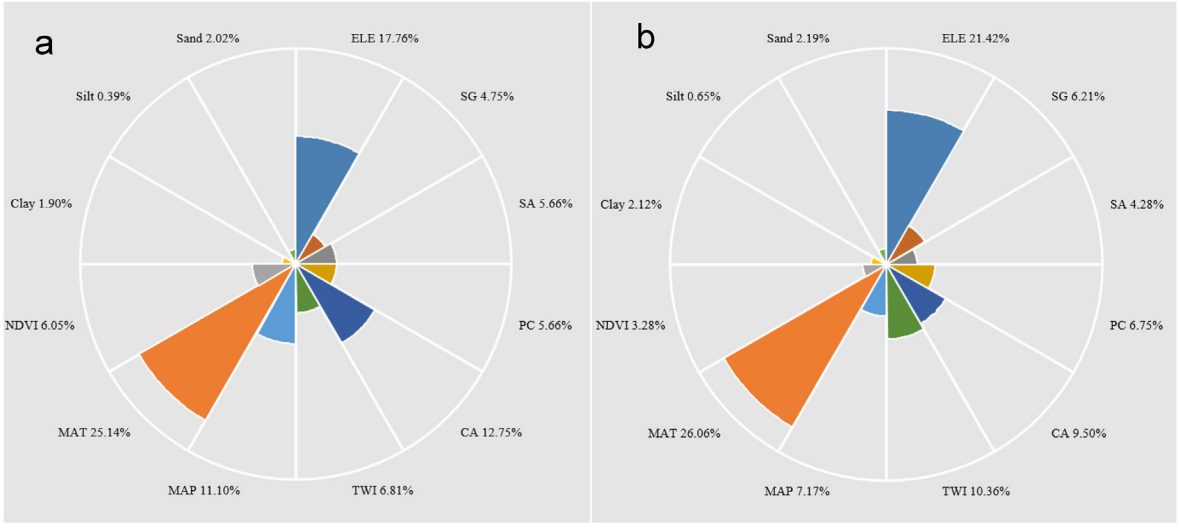

Fig. 3. Relative importance (RI) of environmental factor in SOC stocks prediction as determined from 100 runs of the boosted regression trees model in 1990 (**a**) and 2015 (**b**).

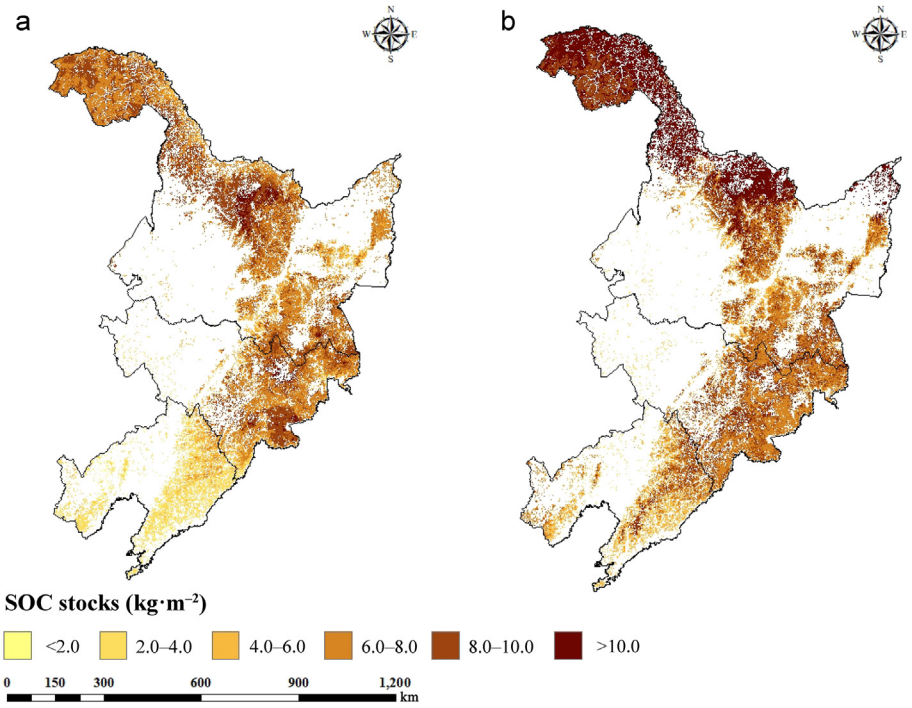

Fig. 4. Spatial distribution of soil organic carbon (SOC) stocks predicted by boosted regression trees model in 1990 (**a**) and 2015 (**b**).

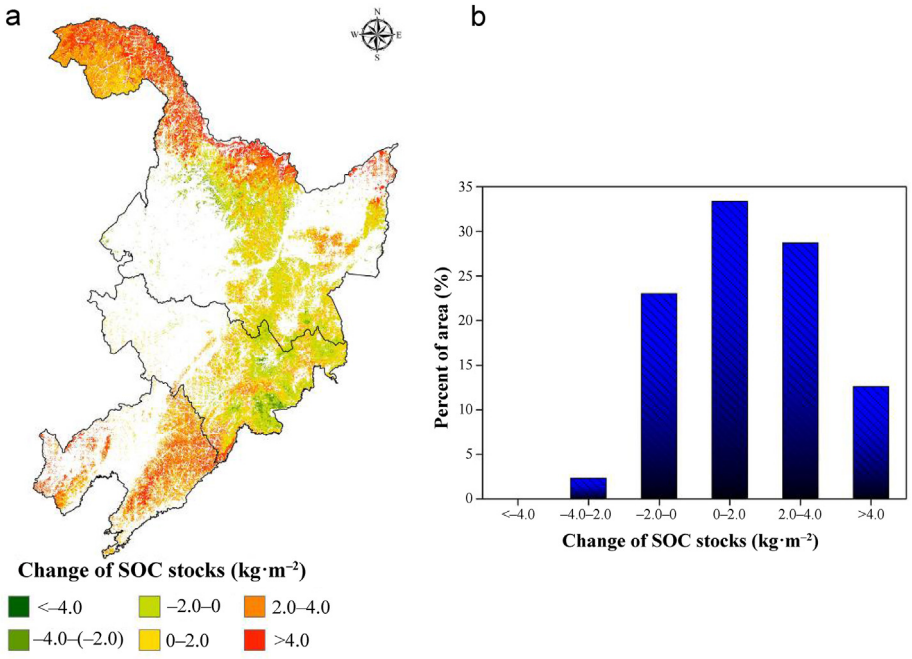

Fig. 5. Changes in SOC stocks in the study area in 1990 and 2015, and its area percentages change.

same as the overall change trend of China's forest carbon reserves in the same period (He et al., 2019).

Our results indicated that the spatial pattern of topsoil SOC stocks in the forest ecosystem in both periods showed a similar characteristics-higher SOC in the southeast and lower in the northwest, with an average of about 6.89 kg C·m^{-2} in 1990, and 8.48 kg C·m^{-2} in 2015 (Fig. 4). The highest SOC stocks appeared in the northwest part of the study area. It was mainly because of the latitudinal effect, especially, in the cold temperate zone where the MAT and evapotranspiration were lower, high soil relative humidity and long soil freezing time was conducive to the accumulation of soil organic matter (Gomes et al., 2019). With denser vegetation, SOC decomposition rate was minimum leading to higher SOC accumulation (Adhikari et al., 2019). In addition, litter accumulation for several years might have contributed to the higher soil organic matter accumulation (Wang et al., 2020b).

Our study also analyzed the spatial variation of SOC stocks under different soil types in both periods (Table 3). The reduced SOC stocks in 2015 mainly appeared on Andosols and Isohumosols, equivalent to the loss of 0.2 and 6.5 Tg C; an unexpected result. It had been shown that the two soil types have the potential to sequester carbon in the forest setting

Table 3
Soil organic carbon (SOC) stocks under different soil type in 1990 and 2015 surveys.

Soil type	Area (km^2)	Average SOC stock (kg C·m^{-2})		SOC stock (Tg C)		Change of average SOC stock (kg C·m^{-2})	Change of stocks (Tg C)
		1990	2015	1990	2015		
Andosols	253.68	7.53	7.15	2.01	1.81	−0.38	−0.20
Anthrosols	1375.03	5.84	6.69	8.75	9.96	0.85	1.21
Argosols	8340.05	7.34	8.33	63.24	69.81	0.99	6.57
Cambosols	251027.45	6.48	8.22	1693.20	2108.58	1.74	415.38
Gleyosols	22300.17	7.62	9.81	174.19	222.17	2.19	47.98
Halosols	75.95	2.73	3.16	0.24	0.28	0.43	0.04
Histosols	547.20	7.48	8.36	4.26	4.70	0.88	0.44
Isohumosols	17927.42	6.58	6.58	132.82	126.30	0	−6.52
Primosols	3111.00	3.64	5.38	13.19	19.33	1.74	6.14
Total	304957.93	–	–	2091.90	2562.94	–	471.04

(Yu et al., 2007; Huang et al., 2017; Wang et al., 2019). We think however that it was the central location in which these soil types were dominant which lead this observation. The central area was where the original forest had felled, cultivated land created, leading to an overall reduction the SOC stocks, rather than being related to the soil type itself. The inverse holds for the specific increase of SOC stocks in Cambosols and Gleyosols and Isohumosols. These soil types are mostly distributed in the more 'virgin' forest area in the northern high latitudes. Here, lower air temperatures reduce microbial activity, limiting the organic matter decomposition leading to a relative higher SOC accumulation in the region, independent of soil type (Carvalhais et al., 2014). Therefore, we suggest that the government support for returning farmland to forest under the Andosol and Isohumosol should be accelerated in the areas with relatively high elevation and steep slope. Tree planting in such areas should therefore be rapidly increased in order to speed up the recovery of the local and regional ecological environment and simultaneously to the beneficial enhancement of their forest soil carbon sequestration.

4.2. SOC stocks change and its dynamics

During the past 25 years, SOC stocks in the forest ecosystem in Northeast China has increased by 471 Tg C (Table 3), equivalent to an average SOC stock increase of 0.063 kg C·m^{-2}·yr^{-1}. It was however surprising that 25% of the study area did show a decreasing SOC trend (mostly −2 to 0 kg C·m^{-2} range) primary distributed in the central region (Fig. 5). We found that those areas might be the main timber forest production base in Northeast China. With China's reforms and opening-up in the 1990s, the growing demand for wood had increased, resulting large areas of virgin forest been felled, and the land was used for farming, urban and industrial development, mining (Zhao et al., 2009; Su et al., 2021). Subsequently, the Chinese government did realize the importance of more ecological environment protection and implemented a series of policies of returning significant areas of farmland back to forests. This greatly restored the forest coverage of these areas and significantly improved the overall ecological environment (Xu et al., 2007; Wang et al., 2020b). Qi et al. (2019) also concluded that a large number of forests and grasslands in Liaoning Province (Northeast China) had been reclaimed for farmland or converted to construction land since the 1990s, and that this was the main reason for the reduction of forest SOC stocks in that region. This conclusion was also confirmed by other studies, such as Cao et al. (2003), Zhao et al. (2009), and Wang et al. (2019).

4.3. Controlling factors of SOC stocks

For both periods, climatic factors were the most powerful and effective environmental factors in predicting SOC stocks, and the results are consistent with previous studies (Fantappiè et al., 2011; Zhong et al., 2018; Zhou et al., 2019; Gomes et al., 2019). Climate plays an important role in the accumulation and consumption of SOC (Lal, 2020). To understand the extent to which climate change can explain the SOC change in Italy from 1961 to 2008, Fantappiè et al. (2011) predicted the spatial distribution of SOC in the two periods with MAP, MAT, elevation and latitude as predictors. They found a significant interdependence between SOC and climatic variables. In our study area, SOC increased with precipitation but decreased with the increasing temperature. In Brazil, Gomes et al. (2019) used historical data and four machine learning algorithms (random forests, cubist, generalized linear model boosting and support vector machines) to estimate SOC stocks, and pointed out a significant correlation between climatic factors and SOC stocks, as also reported in this study. In North and Northeast China, Zhou et al. (2019) used environment data (land use, topographic factors, vegetation index, visible near infrared spectroscopy and climatic factors) and random forest model to predict spatial-temporal variations of SOC stocks in the 1980s and 2000s. They recognized climatic factors as the main environmental factors affecting SOC stocks but had different contributions in different regions. Adhikari et al. (2019) assessed the future climate and land use change impacts on topsoil SOC stocks in Wisconsin and found climatic factors as the main environmental factor affecting SOC stocks. Among all climatic factors, MAT and MAP were the powerful and effective environmental factors, which were widely used in the spatial simulation of SOC stocks (Gomes et al., 2019; Adhikari et al., 2019; Wang et al., 2021; Gu et al., 2022). In China, Wang et al. (2021) used nine environmental factors and BRT model to map SOC stocks in different ecosystems in the 1980s and 2010s. Their results showed that MAT and MAP were the main environmental factors affecting the spatial variation of SOC stocks. Study of Meersmans et al. (2012) reported precipitation pattern as the main reason affecting the spatial variation of regional SOC.

The RI of the topography-related factors was 53.4% in 1990, and 58.5% in 2015 indicating that topography-related factors also had an important impact on the spatial variation of SOC Stocks in the region. In both periods, ELE had the highest RI among all topographic factors. Other studies also showed an impact of elevation in SOC stocks distribution in forest ecosystem (Fernández-Romero et al., 2014; Lozano-García et al., 2016; Wang et al., 2020a; Schönauer et al., 2022). This might be related to the impact of elevation on vegetation type, climate zone, soil properties, soil microenvironment, soil layer thickness, soil microbial community structure and activity (Lozano-García et al., 2016). Tsozué et al. (2019) studied the impact of ELE on SOC stocks and found that the accumulation and stability of SOC were related to clay content, parent material, climate and vegetation, but mainly controlled by altitude gradient - increased SOC value with the increase of ELE. Prietzel and Christophel (2014) found that SOC stocks showed an upward trend in high-altitude areas with low temperature and high precipitation. Liu et al. (2011) found a significant impact of elevation on SOC across the Loess Plateau region of China, as reported in present study.

There are few studies investigating the effect of SA on SOC stocks, and most studies found a higher SOC stocks in shady slope than in sunny slope (Sun et al., 2015; Lozano-García et al., 2016; Wang et al., 2020b). Qin et al. (2017) found an increased SOC from 16.16 to 72.50 g·kg^{-1} from the south slope to the north slope, with an increase of 77.71%. SG affects soil erosion, and its impact on SOC was mainly due to the impact of SG on the

degree of soil erosion. Most studies showed a decreased SOC with increasing soil erosion. Koulouri and Giourga (2007) reported an accelerated reduction of SOC in steeper slopes. In this study, TWI was found to be a good predictor of SOC stocks. TWI represents an impact of regional topography on runoff flow direction and accumulation, and showed a positive correlation with soil moisture (Zhu et al., 2017; Schönauer et al., 2022). Craine and Gelderman (2011) considered TWI as an effective environmental factor in the prediction of SOC stocks as it relates potential moist areas in the landscape (Riihimäki et al., 2021) where a higher SOC could be expected.

NDVI and soil related factors (clay, silt and sand) showed weak RI in the spatial prediction of SOC stocks in both periods. Although studies of Wang et al. (2019), Bhunia et al. (2019), and Yang et al. (2020) reported a stronger relationship of NDVI with SOC stocks spatial distribution, we believe that the effectiveness of NDVI in our study was offset by other variables such as terrain and climate (Table 2) resulting a weaker relationship RI. Peng et al. (2015) obtained similar conclusions to ours. Among the texture fractions, clay content had a higher RI (1.9% in 1990, 2.1% in 2015) than the RI of silt and sand content. This conclusion is consistent with previous studies of Leifeld et al. (2005) and Poeplau et al. (2020). In addition, soil texture was not significantly correlated with SOC stocks in 2015, but showed a significant correlation with SOC stocks in 1990, which might be caused by the fact that the sampling sites in 2015 were mostly in economic forests and soil structure was artificially modified. In eastern China, Zhong et al. (2018) concluded that the SOC stocks and clay content at the depth of 0–10 and 10–20 cm showed a synchronous increase trend from semi-arid area to humid area, and they were correlated.

4.4. Potential research limitations

Based on low MAE and RMSE and high R^2 and LCCC, BRT model was found the best model to accurately predict the spatial distribution of SOC stocks in the study area for both periods. Apart from the higher model uncertainty in 1990 (0.36 kg C·m^{-2}) than in 2015 (0.28 kg C·m^{-2}), there were other potential sources of uncertainties associated with this prediction. Firstly, soil data in 1990 were obtained from the second soil survey database of three provinces (Liaoning, Jilin, Heilongjiang). Changing staff responsible for the data collection undertaken in 1990 might have caused some unforeseen errors during sampling in 2015. Secondly, the collection of soil samples in 2015 was not conducted in-situ in 1990, which may lead to estimation errors between the final predicted results and the actual ones. Thirdly, in 1990, some soil samples did not have BD measured. BD for those samples were estimated with PTFs that might have also added some uncertainty in the final prediction. As far as the uncertainty in both times is concerned, it might be attributed to the differences in data source, data conversion in GIS, point interpolations of weather data, and the spatial detail of the prediction grid itself. These sources of uncertainties which were rather inevitable in DSM might have led to the differences between our estimates and the true SOC stocks measured in the region.

5. Conclusions

The BRT model approach had the lowest MAE and RMSE and highest R^2 and LCCC values compared to MSR and RK in predicting spatial distribution of forest topsoil SOC stocks in Northeast China in 1990 and 2015. The average SOC stocks increased (in ca. 75% of the total area) from 6.89 kg C·m^{-2} (1990) to 8.48 kg C·m^{-2} (2015) adding a total of 471 Tg C. The SOC stocks in 1990 and 2015 had similar spatial distribution patterns, with higher stocks in northeast and lower southwest. Topsoil forest SOC stocks did decline in 25% of the study area primary in the central area of timber forest. These reduced SOC stocks were mainly found on Andosols and Isohumosols, amounting to a loss of 0.2 and 6.5 Tg C, respectively. MAT and ELE were the main controlling factors of SOC stocks spatial variation in the two periods. Accurate assessment of spatial and temporal variations of topsoil forest SOC stocks is helpful for accurate assessment of forest ecosystem carbon cycle, which is of great significance for predicting climate change and formulating strategies and measures to deal with climate change. The results of this study will also provide data support for forestry management and ecological restoration in the region.

Funding

This work was funded by the National Key R&D Program of China (Grant No. 2021YFD1500200), National Natural Science Foundation of China (Grant No. 42077149), China Postdoctoral Science Foundation (Grant No. 2019M660782), 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. 2021-BS-136), China Scholarship Council (201908210132), and Young Scientific and Technological Talents Project of Liaoning Province (Grant Nos. LSNQN201910 and LSNQN201914).

Availability of data and material

The dataset used in the current study are available from the corresponding author on reasonable request.

Authors' contributions

S.W. and F.Q. conceived the research idea and designed sampling plan. X.J., S.W., and C.H. conducted field data collection and laboratory analysis. S.W. and X.J. conduced data analysis and drafted the paper. Q.Z., K.A. and B.R. revised the paper. All authors have read and agreed to the published version of the manuscript.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fecs.2023.100101.

References

Adhikari, K., Hartemink, A.E., 2016. Linking soils to ecosystem services—a global review. Geoderma 262, 101–111.
Adhikari, K., Hartemink, A.E., Minasny, B., Bou Kheir, R., Greve, M.B., Greve, M.H., 2014. Digital mapping of soil organic carbon contents and stocks in Denmark. PLoS One 9 (8), e105519.
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.
Anderson, D.W., 1988. The effect of parent material and soil development on nutrient cycling in temperate ecosystems. Biogeochemistry 5 (1), 71–97.
Banday, M., Bhardwaj, D.R., Pala, N.A., 2019. Influence of forest type, altitude and NDVI on soil properties in forests of North Western Himalaya, India. Acta Ecol. Sin. 39 (1), 50–55.
Batjes, N.H., 1996. Total carbon and nitrogen in the soils of the world. Eur. J. Soil Sci. 47 (2), 151–163.
Bhunia, G.S., Kumar Shit, P., Pourghasemi, H.R., 2019. Soil organic carbon mapping using remote sensing techniques and multivariate regression model. Geocarto Int. 34 (2), 215–226.
Blum, W.E., 2005. Functions of soil for society and the environment. Rev. Environ. Sci. Biotechnol. 4 (3), 75–79.
Bockheim, J.G., Gennadiyev, A.N., 2010. Soil-factorial models and earth-system science: a review. Geoderma 159 (3–4), 243–251.
Bockheim, J.G., Gennadiyev, A.N., Hartemink, A.E., Brevik, E.C., 2014. Soil-forming factors and soil taxonomy. Geoderma 226, 231–237.
Bradshaw, C.J., Warkentin, I.G., 2015. Global estimates of boreal forest carbon stocks and flux. Global Planet. Change 128, 24–30.

Cao, M.K., Tao, B., Li, K.R., Shao, X.M., Prience, S.D., 2003. Interannual variation in terrestrial ecosystem carbon fluxes in China from 1981 to 1998. Acta Bot. Sin. 45 (5), 552–560.
Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., Reichstein, M., 2014. Global covariation of carbon turnover times with climate in terrestrial ecosystems. Nature 514 (7521), 213–217.
China National Bureau of Statistics, 2015. China Statistical Yearbook 2015. China Statistics Press, Beijing, China (in Chinese).
Chinese Soil Taxonomy Research Group, 2001. Keys to Chinese Soil Taxonomy. Press of University of Science and Technology of China, Hefei, China (in Chinese).
Colin, B., Clifford, S., Wu, P., Rathmanner, S., Mengersen, K., 2017. Using boosted regression trees and remotely sensed data to drive decision-making. Open J. Stat. 7 (5), 859–875.
Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1.4. Geosci. Model Dev 8 (7), 1991–2007.
Craine, J.M., Gelderman, T.M., 2011. Soil moisture controls on temperature sensitivity of soil organic carbon decomposition for a mesic grassland. Soil Biol. Biochem. 43 (2), 455–457.
Dixon, R.K., Solomon, A.M., Brown, S., Houghton, R.A., Trexier, M.C., Wisniewski, J., 1994. Carbon pools and flux of global forest ecosystems. Science 263 (5144), 185–190.
Dong, J., Zhou, K., Jiang, P., Wu, J., Fu, W., 2021. Revealing horizontal and vertical variation of soil organic carbon, soil total nitrogen and C:N ratio in subtropical forests of southeastern China. J. Environ. Manag. 289, 112483.
Ebrahimy, H., Feizizadeh, B., Salmani, S., Azadi, H., 2020. A comparative study of land subsidence susceptibility mapping of Tasuj plane, Iran, using boosted regression tree, random forest and classification and regression tree methods. Environ. Earth Sci. 79 (10), 223.
Elith, J., Leathwick, J., 2017. Boosted regression trees for ecological modeling. R Documentation. https://cran.r-project.org/web/packages/dismo/vignettes/brt.pdf. (Accessed 12 June 2011).
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.
Fernández-Romero, M.L., Lozano-García, B., Parras-Alcántara, L., 2014. Topography and land use change effects on the soil organic carbon stock of forest soils in Mediterranean natural areas. Agric. Ecosyst. Environ. 195, 1–9.
Gomes, L.C., Faria, R.M., de Souza, E., Veloso, G.V., Schaefer, C.E.G., Fernandes Filho, E.I., 2019. Modelling and mapping soil organic carbon stocks in Brazil. Geoderma 340, 337–350.
Gu, J., Bol, R., Sun, Y., Zhang, H., 2022. Soil carbon quantity and form are controlled predominantly by mean annual temperature along 4000 km North-South transect of Eastern China. Catena 217, 106498.
Harper, A.B., Powell, T., Cox, P.M., House, J., Huntingford, C., Lenton, T.M., Shu, S., 2018. Land-use emissions play a critical role in land-based mitigation for Paris climate targets. Nat. Commun. 9, 2938.
Hateffard, F., Dolati, P., Heidari, A., Zolfaghari, A.A., 2019. Assessing the performance of decision tree and neural network models in mapping soil properties. J. Mt. Sci. 16 (8), 1833–1847.
Hengl, T., Heuvelink, G.B.M., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120, 75–93.
Huang, J., Ebach, M.C., Triantafilis, J., 2017. Cladistic analysis of Chinese soil taxonomy. Geoderma Reg 10, 11–20.
Hutchinson, M.F., 1995. Interpolating mean rainfall using thin plate smoothing splines. Int. J. Geogr. Inf. Sci. 9 (4), 385–403.
Keskin, H., Grunwald, S., Harris, W.G., 2019. Digital mapping of soil carbon fractions with machine learning. Geoderma 339, 40–58.
Kirkpatrick, L.A., 2015. A Simple Guide to IBM SPSS Statistics-Version 23.0. Cengage Learning.
Korhonen, J., Wihersaari, M., Savolainen, I., 2001. Industrial ecosystem in the Finnish forest industry: using the material and energy flow model of a forest ecosystem in a forest industry system. Ecol. Econ. 39 (1), 145–161.
Koulouri, M., Giourga, C., 2007. Land abandonment and slope gradient as key factors of soil erosion in Mediterranean terraced lands. Catena 69 (3), 274–281.
Krishnan, P., Alexander, J.D., Butler, B.J., Hummel, J.W., 1980. Reflectance technique for predicting soil organic matter. Soil Sci. Soc. Am. J. 44 (6), 1282–1285.
Kumar, S., Lal, R., Liu, D., 2012. A geographically weighted regression kriging approach for mapping soil organic carbon stock. Geoderma 189, 627–634.
Kumar, P., Pandey, P.C., Singh, B.K., Katiyar, S., Mandal, V.P., Rani, M., Patairiya, S., 2016. Estimation of accumulated soil organic carbon stock in tropical forest using geospatial strategy. Egypt. J. Remote. Sens. Space Sci. 19 (1), 109–123.
Lal, R., 2004. Soil carbon sequestration impacts on global climate change and food security. Science 304, 1623–1627.
Lal, R., 2020. Managing soils for negative feedback to climate change and positive impact on food and nutritional security. Soil Sci. Plant Nutr. 66 (1), 1–9.
Lampa, E., Lind, L., Lind, P.M., Bornefalk-Hermansson, A., 2014. The identification of complex interactions in epidemiology and toxicology: a simulation study of boosted regression trees. Environ. Health 13, 57.
Leifeld, J., Bassin, S., Fuhrer, J., 2005. Carbon stocks in Swiss agricultural soils predicted by land-use, soil characteristics, and altitude. Agric. Ecosyst. Environ. 105 (1–2), 255–266.
Li, Q.Q., Yue, T.X., Wang, C.Q., Zhang, W.J., Yu, Y., Li, B., Bai, G.C., 2013. Spatially distributed modeling of soil organic matter across China: an application of artificial neural network approach. Catena 104, 210–218.
Lin, L., 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics 45, 255–268.
Liu, Z., Shao, M.A., Wang, Y., 2011. Effect of environmental factors on regional soil organic carbon stocks across the Loess Plateau region, China. Agric. Ecosyst. Environ. 142 (3–4), 184–194.
Liu, Y.H., Yu, G.R., Wang, Q.F., Zhang, Y.J., 2012. Huge carbon sequestration potential in global forests. J. Resour. Ecol. 3 (3), 193–201.
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 (5), 1673–1680.
Lozano-García, B., Parras-Alcántara, L., Brevik, E.C., 2016. Impact of topographic aspect and vegetation (native and reforested areas) on soil organic carbon and nitrogen budgets in Mediterranean natural areas. Sci. Total Environ. 544, 963–970.
Lynn, J., Peeva, N., 2021. Communications in the IPCC's sixth assessment report cycle. Clim. Change 169, 18.
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 (5), 1053–1065.
Martin, T.M., Harten, P., Young, D.M., Muratov, E.N., Golbraikh, A., Zhu, H., Tropsha, A., 2012. Does rational selection of training and test sets improve the outcome of QSAR modeling? J. Chem. Inf. Model. 52 (10), 2570–2578.
McBratney, A.B., Santos, M.M., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1–2), 3–52.
Meersmans, J., Martin, M.P., Lacarce, E., De Baets, S., Jolivet, C., Boulonne, L., Lehmann, S., Saby, N.P.A., Bispo, A., Arrouays, D., 2012. A high resolution map of French soil organic carbon. Agron. Sustain. Dev. 32, 841–851.
Nabiollahi, K., Eskandari, S., Taghizadeh-Mehrjardi, R., Kerry, R., Triantafilis, J., 2019. Assessing soil organic carbon stocks under land-use change scenarios using random forest models. Carbon Manag. 10 (1), 63–77.
Navarro, M., Hailu, A., Langlois, T., Ryan, K.L., Kragt, M.E., 2020. Determining spatial patterns in recreational catch data: a comparison of generalized additive mixed models and boosted regression trees. ICES J. Mar. Sci. 77 (6), 2216–2225.
Ngaba, M.J.Y., Hu, Y.L., Bol, R., Ma, X.Q., Jin, S.F., Mgelwa, A.S., 2019. Effects of land use change from natural forest to plantation on C, N and natural abundance of ^{13}C and ^{15}N along a climate gradient in eastern China. Sci. Rep. 9, 16516.
Padarian, J., Minasny, B., McBratney, A.B., 2019. Using deep learning to predict soil properties from regional spectral data. Geoderma Reg. 16, e00198.
Peng, Y., Xiong, X., Adhikari, K., Knadel, M., Grunwald, S., Greve, M.H., 2015. Modeling soil organic carbon at regional scale by combining multi-spectral images with laboratory spectra. PLoS One 10 (11), e0142295.
Poeplau, C., Jacobs, A., Don, A., Vos, C., Schneider, F., Wittnebel, M., Flessa, H., 2020. Stocks of organic carbon in German agricultural soils—key results of the first comprehensive inventory. J. Plant Nutr. Soil Sci. 183 (6), 665–681.
Prietzel, J., Christophel, D., 2014. Organic carbon stocks in forest soils of the German Alps. Geoderma 221, 28–39.
Qi, L., Wang, S., Zhuang, Q., Yang, Z., Bai, S., Jin, X., Lei, G., 2019. Spatial-temporal changes in soil organic carbon and pH in the Liaoning Province of China: a modeling analysis based on observational data. Sustainability 11 (13), 3569.
Qin, Y.Y., Holden, N.M., Feng, Q., 2017. Influence of slope aspect on plant community composition and its implications for restoration of a Chinese mountain range. Pol. J. Environ. Stud. 26 (1), 375–383.
Riihimäki, H., Kemppinen, J., Kopecký, M., Luoto, M., 2021. Topographic wetness index as a proxy for soil moisture: the importance of flow-routing algorithm and grid resolution. Water Resour. Res. 57 (10), e2021WR029871.
Riza, S., Sekine, M., Kanno, A., Yamamoto, K., Imai, T., Higuchi, T., 2021. Modeling soil landscapes and soil textures using hyperscale terrain attributes. Geoderma 402, 115177.
Scharlemann, J.P., Tanner, E.V., Hiederer, R., Kapos, V., 2014. Global soil carbon: understanding and managing the largest terrestrial carbon pool. Carbon Manag. 5 (1), 81–91.
Schönauer, M., Prinz, R., Väätäinen, K., Astrup, R., Pszenny, D., Lindeman, H., Jaeger, D., 2022. Spatio-temporal prediction of soil moisture using soil maps, topographic indices and SMAP retrievals. Int. J. Appl. Earth Obs. 108, 102730.
Su, F., Xu, S., Sayer, E.J., Chen, W., Du, Y., Lu, X., 2021. Distinct storage mechanisms of soil organic carbon in coniferous forest and evergreen broadleaf forest in tropical China. J. Environ. Manag. 295, 113142.
Tsozué, D., Nghonda, J.P., Tematio, P., Basga, S.D., 2019. Changes in soil properties and soil organic carbon stocks along an elevation gradient at Mount Bambouto, Central Africa. Catena 175, 251–262.
Wäldchen, J., Schulze, E.D., Schöning, I., Schrumpf, M., Sierra, C., 2013. The influence of changes in forest management over the past 200 years on present soil organic carbon stocks. For. Ecol. Manag. 289, 243–254.
Wang, S., Zhuang, Q., Jia, S., Jin, X., Wang, Q., 2018. Spatial variations of soil organic carbon stocks in a coastal hilly area of China. Geoderma 314, 8–19.
Wang, S., Zhuang, Q., Yang, Z., Yu, N., Jin, X., 2019. Temporal and spatial changes of soil organic carbon stocks in the forest area of Northeastern China. Forests 10 (11), 1023.
Wang, S., Adhikari, K., Zhuang, Q., Yang, Z., Jin, X., Wang, Q., Bian, Z., 2020a. An improved similarity-based approach to predicting and mapping soil organic carbon and soil total nitrogen in a coastal region of northeastern China. PeerJ 8, e9126.
Wang, S., Zhuang, Q., Jin, X., Yang, Z., Liu, H., 2020b. Predicting soil organic carbon and soil nitrogen stocks in topsoil of forest ecosystems in northeastern China using remote sensing data. Rem. Sens. 12 (7), 1115.

Wang, S., Xu, L., Zhuang, Q., He, N., 2021. Investigating the spatio-temporal variability of soil organic carbon stocks in different ecosystems of China. Sci. Total Environ. 758, 143644.
Were, K., Bui, D.T., Dick, Ø.B., Singh, B.R., 2015. A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. Ecol. Indicat. 52, 394–403.
Wösten, J.H.M., Pachepsky, Y.A., Rawls, W.J., 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. J. Hydrol. 251 (3–4), 123–150.
Xu, D., 1994. The effect of human management activities on the carbon in forest soils. World For. Res. 5, 26–31 (in Chinese).
Xu, X.L., Cao, M.K., Li, K.R., 2007. Temporal-spatial dynamics of carbon storage of forest vegetation in China. Prog. Geogr. 26 (6), 1–16. doi:10.11820/dlkxjz.2007.06.001 (in Chinese).
Yang, L., He, X., Shen, F., Zhou, C., Zhu, A.X., Gao, B., Li, M., 2020. Improving prediction of soil organic carbon content in croplands using phenological parameters extracted from NDVI time series data. Soil Till. Res. 196, 104465.
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.
Yu, D.S., Shi, X.Z., Wang, H.J., Sun, W.X., Warner, E.D., Liu, Q.H., 2007. National scale analysis of soil organic carbon storage in China based on Chinese soil taxonomy. Pedosphere 17 (1), 11–18.
Zhang, S., Huang, Y., Shen, C., Ye, H., Du, Y., 2012. Spatial prediction of soil organic matter using terrain indices and categorical variables as auxiliary information. Geoderma 171, 35–43.
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, 425.
Zhao, J.F., Yan, X.D., Jia, G.S., 2009. Simulation of carbon stocks of forest ecosystems in Northeast China from 1981 to 2002. J. Appl. Ecol. 20 (2), 241–249.
Zhong, Z., Chen, Z., Xu, Y., Ren, C., Yang, G., Han, X., Feng, Y., 2018. Relationship between soil organic carbon stocks and clay content under different climatic conditions in Central China. Forests 9 (10), 598.
Zhou, Y., Hartemink, A.E., Shi, Z., Liang, Z., Lu, Y., 2019. Land use and climate change effects on soil organic carbon in North and Northeast China. Sci. Total Environ. 647, 1230–1238.
Zhu, N., Jiang, H., Jin, Y., 1990. A phenology study on the common tree species of natural secondary forests in northeast China. Chin. J. Plant Ecol. 14 (4), 336–349.
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.
Zhu, Q., Zhou, Z., Duncan, E.W., Lv, L., Liao, K., Feng, H., 2017. Integrating real-time and manual monitored data to predict hillslope soil moisture dynamics with high spatio-temporal resolution using linear and non-linear models. J. Hydrol. 545, 1–11.