Catena 269 (2026) 110145

Contents lists available at ScienceDirect

Catena

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

Assessing soil organic carbon sequestration capacity and its multi-ecosystem driving mechanisms in Northeast China

Zicheng Wang [a], Qianlai Zhuang [b], Zijiao Yang [a], Xinxin Jin [a], Shuai Wang [a,*], Yang Wang [c,**], Xinyu Zhang [a], Di Shi [a], Shouyuan Bian [a]

[a] College of Land and Environment, Shenyang Agricultural University, Shenyang, Liaoning Province 110866, China
[b] Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN 47907, USA
[c] Jilin Agricultural University, College of Resources and Environment, Changchun, Jilin Province 130118, China

ARTICLE INFO

Keywords:
Soil organic carbon sequestration potential
Boosted regression trees
Digital soil mapping
Hassink model
Elustondo model

ABSTRACT

Soil organic carbon (SOC) is a crucial component of terrestrial carbon pools and plays a fundamental role in mitigating climate change. Northeast China, a major grain-producing region and one of the country's most important black soil zones, exhibits high potential for SOC sequestration. A systematic and quantitative assessment of SOC sequestration potential (SOC_p) is urgently needed to support China's "Dual Carbon" strategy. Based on soil-forming factors and soil-landscape theory, this study developed a predictive framework by integrating the Hassink and Elustondo models. Using boosted regression trees (BRT), soil sample ($n = 293$), and 10 multi-source environmental variables including climate, topography, and vegetation, we accurately estimated SOC_p. The model was rigorously validated via 10-fold cross-validation and 100 iterative runs. Results indicated that the Hassink model showed higher stability in BRT predictions ($R^2 = 0.75$), while the Elustondo model ($R^2 = 0.74$) was more sensitive to extreme values. To retain common patterns, eliminate single-model bias, and improve prediction reliability and universality, the arithmetic mean method combines both models' outputs to generate the final SOCp estimate. The spatial distribution of SOC_p in Northeast China showed higher levels in the south and plains, lower in the north and mountainous areas, with a total of 5831.50 Tg C and an areal density of 7.38 kg C m^{-2}. Farmland, covering 39.3% of the region, had the highest sequestration density (8.20 kg C m^{-2}) and contributed 43.7% (2547.51 Tg C) of total SOC_p, indicating its dominant role. Forests (42.8% of area) and grasslands (4.3%) accounted for 38.0% and 4.2% of SOC_p, respectively. Mean annual temperature, mean annual precipitation, and net primary productivity were the primary drivers, explaining 55% of the variation, with ecosystem-specific differences observed. This study reveals SOC_p patterns and driving mechanisms in Northeast China, supporting region-specific carbon management and climate goals.

1. Introduction

Since the Industrial Revolution, anthropogenic activities (fossil fuel combustion and intensive land-use change) have severely disrupted the global carbon cycle, elevating atmospheric CO_2 concentrations, exacerbating global warming and extreme climate events, and threatening ecological stability and socioeconomic development (McBratney et al., 2014). As the largest terrestrial carbon pool; soil organic carbon (SOC) plays an irreplaceable role in climate regulation; with surface soils storing over half of the global SOC stocks. Soil carbon sequestration is thus recognized as a cost-effective and sustainably negative emissions technology (Stockmann et al., 2013); and has been prioritized by global organizations and nations as a core strategy to offset industrial emissions (Amelung et al., 2020). However, precise quantification of SOC sequestration potential and identification of its intrinsic driving mechanisms across diverse ecosystems at regional scales remain insufficiently addressed, posing a critical gap for targeted carbon sequestration practice.

Globally, research on SOC_p has made significant progress, primarily focusing on quantifying its magnitude and identifying the key factors

* Corresponding author at: College of Land and Environment, Shenyang Agricultural University, No. 120 Dongling Road, Shenhe District, Shenyang, Liaoning Province 110866, China.
** Corresponding author.
E-mail address: shuaiwang666@syau.edu.cn (S. Wang).

https://doi.org/10.1016/j.catena.2026.110145
Received 8 October 2025; Received in revised form 13 April 2026; Accepted 17 April 2026
Available online 25 April 2026

that govern SOC_p across various ecosystems and agricultural management practices. Agricultural practices, such as straw return and conservation tillage have been empirically validated to enhance SOC stocks (Follett, 2001; Lessmann et al., 2022). Furthermore; soil physicochemical properties (e.g.; texture; pH) and environmental variables (e.g.; temperature; precipitation) jointly influence soil microbial activity; which plays a central role in driving SOC cycling and stabilization processes (Liang et al., 2017; Jiang et al., 2022). The soil organic carbon saturation theory sets the upper limit for the physical storage of mineral-associated organic matter (MAOM). The amount of residue produced by microorganisms through their carbon use efficiency (CUE) serves as a key precursor for forming stable MAOM—high CUE promotes MAOM carbon sequestration when the soil is unsaturated, but loses this effect once saturation is reached (Sokol et al., 2022; Tao et al., 2023; Yang et al., 2025). Research on SOC saturation mechanisms has provided a critical theoretical framework for quantifying the upper limits of SOC sequestration capacity. Spatially explicit models, such as random forest (a machine learning approach) and the Denitrification-Decomposition (DNDC) model (a process-based model), have been widely applied to map the spatial distribution of SOC_p, thereby elucidating the influence of climate, topography, vegetation cover, and cropping systems (Zhou et al., 2003; Benítez et al., 2007). Methodologies for estimating SOC_p are increasingly interdisciplinary, integrating long-term field observational data, geographic information systems (GIS), and machine learning techniques with traditional statistical, mechanistic, and spatial analysis approaches (Falloon et al., 2002; Nair, 2011; Pechanec et al., 2018; Chen et al., 2019). Nevertheless, substantial discrepancies in SOC_p predictions across different methodologies highlight the urgent need for rigorous model validation against independent field data and the development of harmonized analytical frameworks.

Despite significant progress in studying global soil organic carbon sequestration potential across various regions, most existing research (such as North America, Europe, and other regions) has focused solely on the sequestration potential of individual ecosystems at different scales. Integrated assessment and analysis of soil carbon sequestration potential across multiple ecosystems at the regional scale remains a scientific gap. This gap is not only reflected in the lack of systematic integrated research frameworks, but also stems from limitations in existing technical methods and insufficient quantification of driving mechanisms, specifically manifested in the following two aspects. First, prevailing SOC_p estimation methods including unit carbon balance models and remote sensing inversion techniques frequently fail to adequately capture spatial heterogeneity in environmental and management conditions, resulting in substantial biases in regional-scale SOC_p assessments. Second, the synergistic and antagonistic effects of multiple drivers, encompassing climate (e.g., temperature, precipitation), agricultural management practices (e.g., tillage, fertilization), and soil microbial interactions remain insufficiently quantified, impeding the targeted and precise implementation of SOC-enhancing measures to support carbon neutrality goals. There is a growing scientific consensus that dynamic SOC_p models by integrating empirical field data and advanced computational algorithms are indispensable for improving the accuracy of SOC_p estimation. This is particularly critical for key agricultural regions such as Northeast China, which is characterized by extensive land areas, critical cropland ecosystems, and considerable SOC sequestration potential (Lal, 2004).

Although studies across farmland, forests, and grasslands have identified ecosystem-specific sequestration mechanisms and influencing variables, research depth remains uneven. For instance, forest soil systems, especially regarding micro-food web carbon conversion processes, are poorly understood. Additional limitations include insufficient multi-scale integration, lack of spatial interaction analyses, and inadequate standardization and validation of estimation methods. Moreover, the long-term efficacy and regional adaptability of anthropogenic interventions require further investigation. In Northeast China, existing research has yet adequately addressed estimation inaccuracies or elucidated the multi-factor synergistic mechanisms governing SOC sequestration. Therefore, this study aims to: (1) develop a high-resolution estimation model for topsoil SOC_p in Northeast China by integrating multi-source environmental data including soil properties, climate, terrain, and remote sensing indicators by using advanced machine learning techniques, and assess model performance and uncertainty through 10-fold cross-validation to ensure robustness and generalizability across diverse ecosystems; (2) Leverage the technical advantages of nonlinear fitting and variable importance ranking within the BRT model, comprehensively analyze the heterogeneous characteristics of key drivers across different ecosystems at the regional scale, compare the output differences between two SOCp prediction models, and systematically clarify their core performance advantages and applicability; (3) quantitatively examine the spatial heterogeneity and distribution patterns of SOC_p at regional and landscape scales; and (4) identify and prioritize key driving variables such as soil texture, vegetation productivity, and climatic variables and elucidate their individual and interactive effects on SOC_p through sensitivity analysis and partial dependence plots.

2. Materials and methods

2.1. Study area

This study focused on three northeastern provinces of China, namely Heilongjiang Province, Jilin Province, and Liaoning Province, with a geographical range of 38°43′–53°33′N latitude and 118°57′–135°03′E longitude, covering a total area of approximately 790,000 km^2. The climate in this region is characterized by a temperate monsoon pattern with distinct seasonal variations, including cold and long winters, warm and short summers, an average annual temperature ranging from −6 °C to 9 °C, and annual precipitation between 400 mm and 800 mm. The terrain of the study area is complex and diverse, primarily consisting of plains, mountains, lakes, and water bodies, with the highest elevation reaching 2618 m. Among these, the Northeast Plain is one of the largest plains in China and serves as a crucial agricultural production base. The Changbai Mountains, Xing'an Mountains, and other mountain ranges form a natural ecological barrier, while the region's water system is well-developed, with major rivers such as the Heilongjiang, Songhua, and Liao Rivers, which hold significant economic and ecological value. The soil types in the three northeastern provinces are highly diverse. According to the World Reference Base for Soil Resources (WRB) (IUSS Working Group WRB, 2006), the dominant soil types are Cambisols (52.4%) and Chernozems (34.1%), together accounting for more than 85% of the study area. The black soil region in the three northeastern provinces constitutes one of the four major black soil zones globally. This area not only serves as a key grain production base in China but also plays a vital role in timber and mineral resource supply. The total forested land area reaches 38.75 million hectares, accounting for 14.7% of the national total, with a forest coverage rate of 39.6%, and it is rich in biodiversity, including abundant wild animal and plant resources. In summary, the three northeastern provinces play a significant role in China's economic and social development. Their abundant and high-quality soil types, particularly black soil, provide a solid foundation for studying soil organic carbon (SOC) storage and carbon sequestration potential. Therefore, these three northeastern provinces of China were selected as the study area.

2.2. Soil data collection

Due to the extensive spatial coverage and diverse soil types within the research area, achieving a dense and representative set of sampling points during fieldwork is challenging. Therefore, this study employed a purposive soil sampling method (Zhu et al., 2008) to systematically determine the spatial distribution of sampling sites and completed sample collection and analysis between 2018 and 2020. First, the

primary environmental variables influencing SOCs were identified, and a multidimensional environmental variables matrix was constructed by including climatic variables (MAT and MAP), topographic variables (elevation, slope gradient, and slope aspect), and land use types. These selected environmental variables were then resampled to a uniform spatial resolution (90 m) and coordinate system (Albers Conical Equal Area projection system) using ArcGIS software. Second, the fuzzy C-means clustering method was applied to perform spatial clustering on the aforementioned environmental factors (Dunn, 1973), aiming to identify environmental landscape units that best reflect the spatial variability of SOC. The clustering process resulted in 27 distinct environmental landscape units. Based on the principle of spatial balance, 8 to 10 soil sampling sites were allocated within each landscape unit, yielding a total of 293 sampling sites. Meanwhile, to assess the representativeness of the sampling design, this study conducted a Principal Component Analysis (PCA) on the multi-dimensional environmental

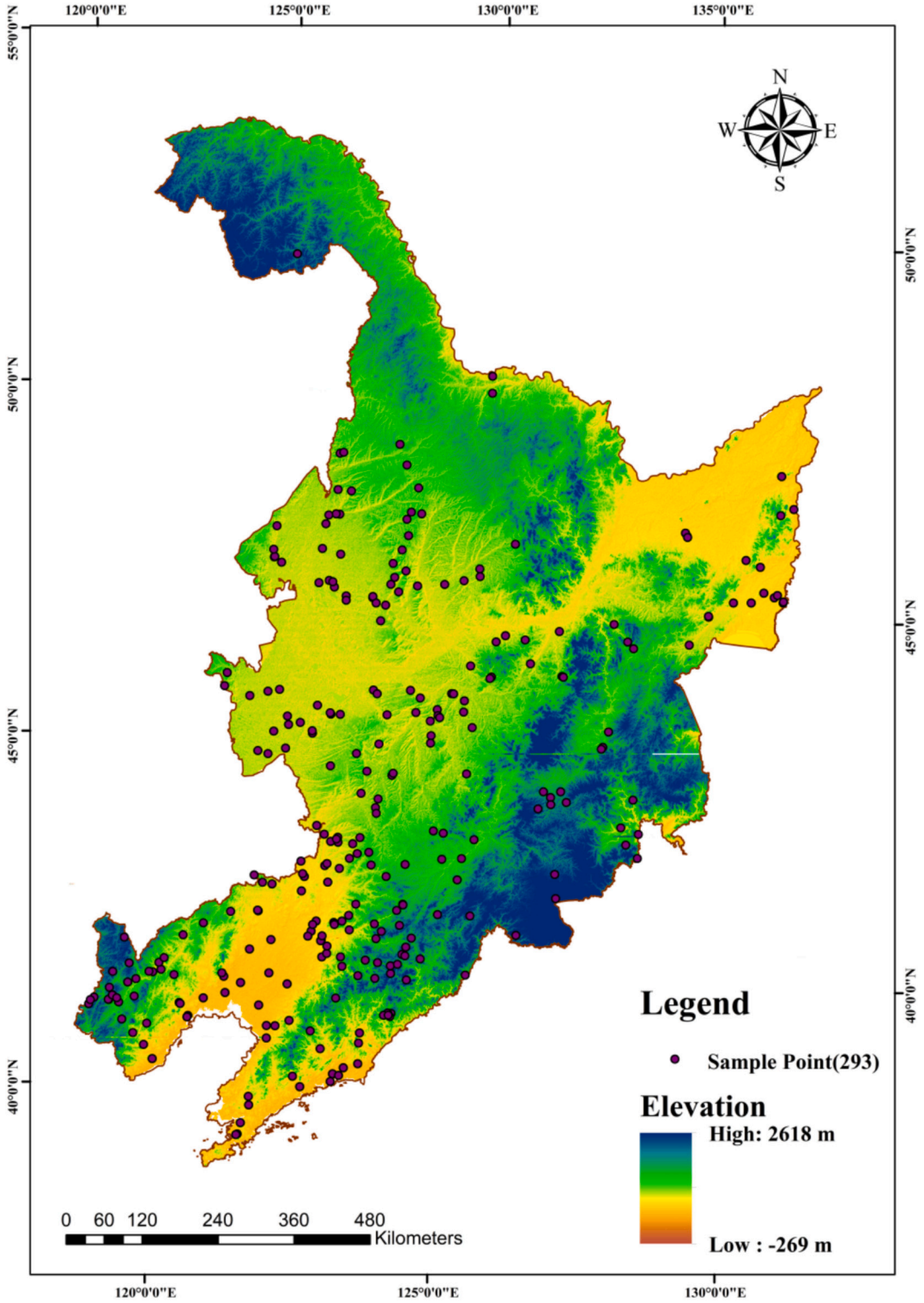

Fig. 1. Distribution of the study area and sampling points.

covariates. The purpose was to quantify and visualize the environmental space of both the entire study area and the sampling points. PCA was utilized to reduce the dimensionality of the environmental variables to two orthogonal principal components (PCA1 and PCA2). Subsequently, these two principal components were used to construct a two-dimensional environmental space and generate a scatter plot (Fig. 2). This scatter plot enabled us to evaluate the degree to which the sampling points overlapped with the environmental envelope of the study area, thus verifying the representativeness of the sampling distribution. Our analysis indicated that the sampling points covered the vast majority of the primary environmental gradients within the study area, which confirmed the robustness and representativeness of the sampling strategy adopted in this study. A handheld global positioning system (Garmin eTrex 221×, USA) was used to record the geographical coordinates of each site. Surface soil samples (0–30 cm depth) were collected following standardized protocols, including 1 kg of mixed soil samples and ring knife samples for soil bulk density determination. Finally, all soil samples were transported to the Laboratory Analysis Center of the School of Land and Environment at Shenyang Agricultural University for analysis. The SOC content was determined by wet oxidation method; soil bulk density was determined by drying method. The soil particle size distribution was determined using the suction tube method, and the proportion of fine particle fractions was calculated. Ultimately, the percentage of fine particles with a diameter < 50 μm was selected as the basis for estimating the soil organic carbon sequestration potential. (See Fig. 1.)

2.3. Environmental variables

This study selected 10 environmental variables as key drivers of SOC_p, specifically including 6 topographic factors (elevation, slope gradient, slope aspect, plan curvature, profile curvature, topographic moisture index), 2 climatic factors (annual mean temperature, annual mean precipitation), and 2 vegetation factors (normalized vegetation index, net primary productivity). Annual mean temperature (MAT) and annual precipitation (MAP) were selected because they directly regulate soil microbial activity and organic carbon decomposition processes (Fantappiè et al., 2011; Yang et al., 2023). Topographic factors such as elevation, slope gradient, aspect, plan curvature, profile curvature, and terrain moisture index (TWI) were selected because they directly alter the spatial distribution of soil organic carbon by reconfiguring regional water and heat distribution patterns (Guo et al., 2019; Li et al., 2014); vegetation factors including normalized difference vegetation index (NDVI) and net primary productivity (NPP) were selected as they directly reflect the intensity of soil organic carbon sources from vegetation biomass input (Zheng et al., 2019; Zhang et al., 2024). The integration of these 10 environmental variables comprehensively captures the primary drivers of soil organic carbon spatial distribution within the study area, avoiding research bias caused by missing core drivers and establishing a robust methodological foundation for soil organic carbon prediction.

2.3.1. Topographic variables

The topographic variables selected in this study include elevation, slope gradient (SG), slope aspect (SA), plan curvature (PLC), profile curvature (PRC), and the topographic wetness index (TWI). Elevation, SG, SA, PLC, and PRC were derived from raster data of a 90-m resolution digital elevation model (DEM) obtained from the Geospatial Data Cloud website of the Chinese Academy of Sciences (http://www.gscloud.cn/sources/accessdata/305?pid=1). TWI was calculated using the Automated Earth Science Analysis System (SAGA GIS) (Conrad et al., 2015).

Elevation refers to the vertical distance between a specific location on the Earth's surface and sea level. According to a study by Wu et al. (2024); elevation was one of the key variables influencing SOC_{stock}; with SOC content showing a significant positive correlation with elevation (Tesfaye et al., 2016). SG is a parameter that describes the steepness of surface units; which directly affects the distribution of surface materials and energy; soil development processes; vegetation patterns; and land use types (Bae and Ryu, 2015). SA refers to the direction of the projection of slope normals onto the horizontal plane; i.e.; the orientation of the slope. SA significantly influences plant growth and the spatial distribution of SOC by altering water and temperature conditions; with sunny slopes generally exhibiting higher organic carbon content (Fang et al., 2016; Jiang et al., 2020). PLC and PRC are parameters used to describe the degree of curvature in two-dimensional and three-

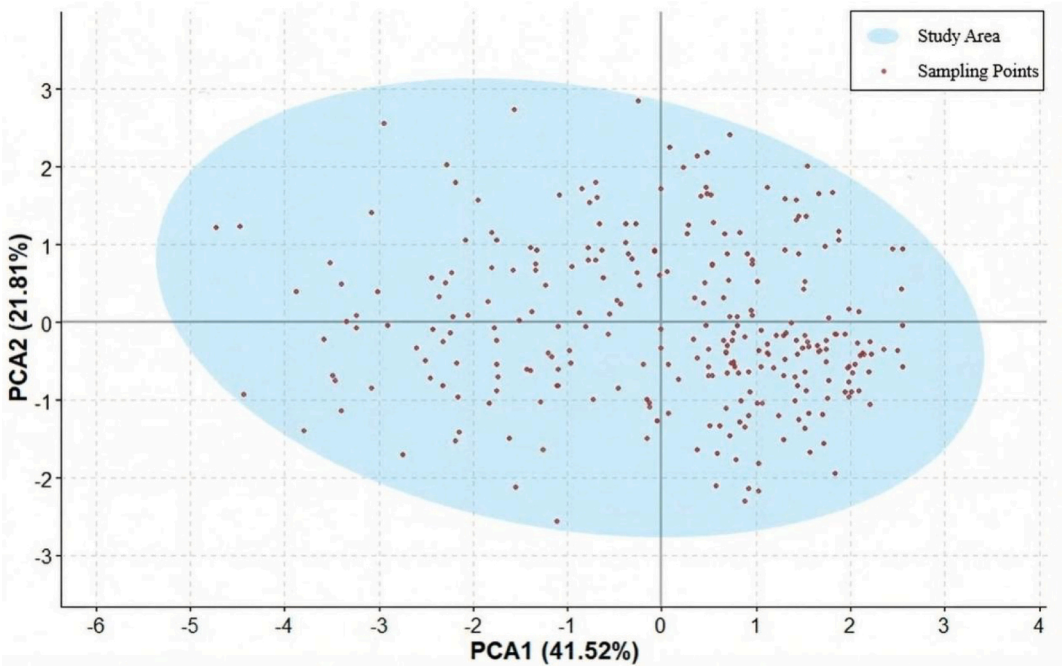

Fig. 2. Principal component analysis (PCA) scatter plot of the environmental space of sampling points and the study area based on environmental variables. PCA1 explains 41.52% of the total variance, while PCA2 explains 21.81% of the total variance.

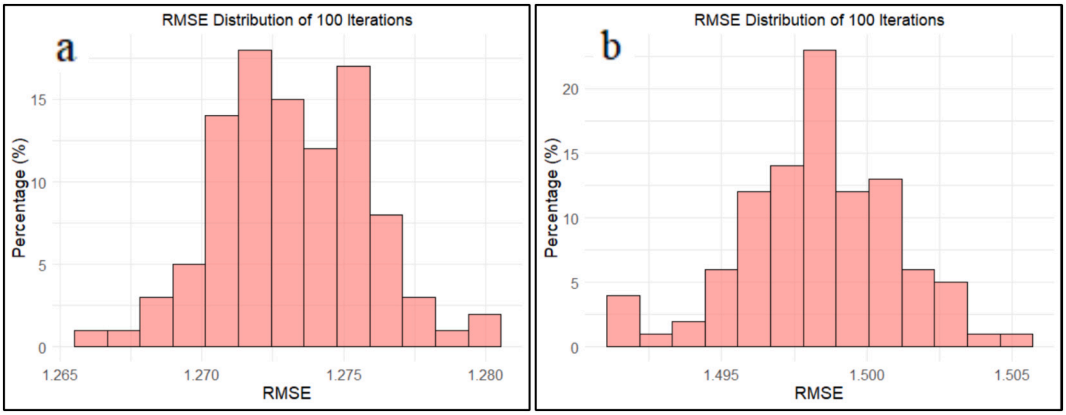

Fig. 3. Histogram of the RMSE response of the Hassink model (a) and the Elustondo model (b) across 100 iterations of the BRT model.

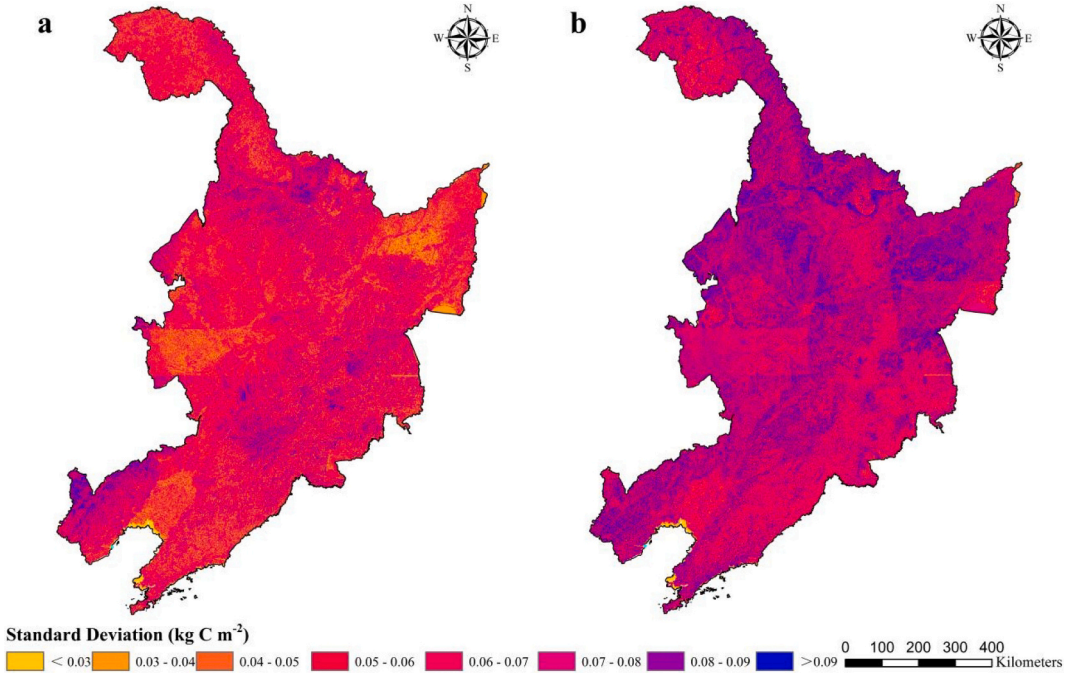

Fig. 4. Standard deviation of the 100-iteration prediction results of the BRT model for predicting soil organic carbon sequestration potential (SOCp) of surface soils in Northeast China (a for Hassink model; b for Elustondo model).

dimensional terrain surfaces, respectively. Terrain curvature has a significant impact on SOC distribution by modifying hydrological conditions and plant growth environments (Li et al., 2014; Deng et al., 2020). TWI is a physical indicator that reflects the influence of terrain on runoff direction and accumulation. It can be used to identify rainfall runoff patterns, areas with potentially higher soil moisture, and zones with water stagnation. This index quantitatively characterizes soil water retention and drainage capacity, revealing soil properties and their spatial distribution patterns under various landscape conditions (Kirkby, 1999; Zhu et al., 2008). In summary, the aforementioned topographic variables were among the primary driving forces influencing the spatial distribution of soil properties, particularly the spatial variability of SOC (Guo et al., 2019).

2.3.2. Climatic variables

This study selected mean annual temperature (MAT) and mean annual precipitation (MAP) as climate variables, with data obtained from the National Met-eorological Science Data Center (https://data.cma.cn/data/detail/dataCode/A.0012.0001.S011.html). Specifically, raster data covering a 30-year period and having a spatial resolution of 1 km were downloaded from the data center. The inverse distance weighted interpolation method was applied to generate continuous surfaces. Subsequently, the data were averaged and resampled to a 90 m resolution to meet the spatial analysis requirements of this study. Previous studies have demonstrated that climatic variables, particularly MAT and MAP, significantly influence SOC_{stock} and their spatial distribution patterns (Fantappiè et al., 2011; Yang et al., 2023).

2.3.3. Land use data

This study used the 2020 land use data obtained from the Resource and Environmental Science Data Center of the Chinese Academy of Sciences (https://www.resdc.cn/DOI/DOI.aspx?DOIid=54). The data were applied in relevant chapters to extract spatial geographic information of different ecosystems within the study area, and subsequently, spatial distribution maps of soil carbon sequestration potential across various ecosystems were generated. Previous studies have demonstrated that different land use patterns exhibit varying potentials for enhancing carbon sinks (Smith, 2008); and that changes in land use significantly

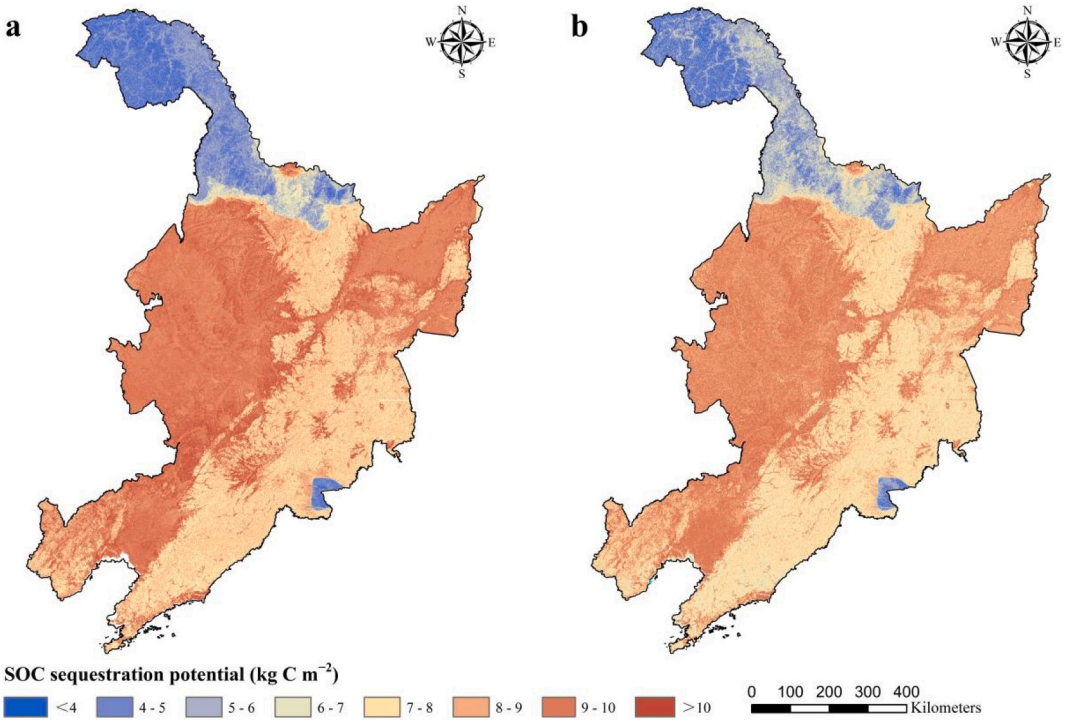

Fig. 5. Spatial distribution of soil organic carbon sequestration potential (SOCp) in Northeast China (a for Hassink model; b for Elustondo model).

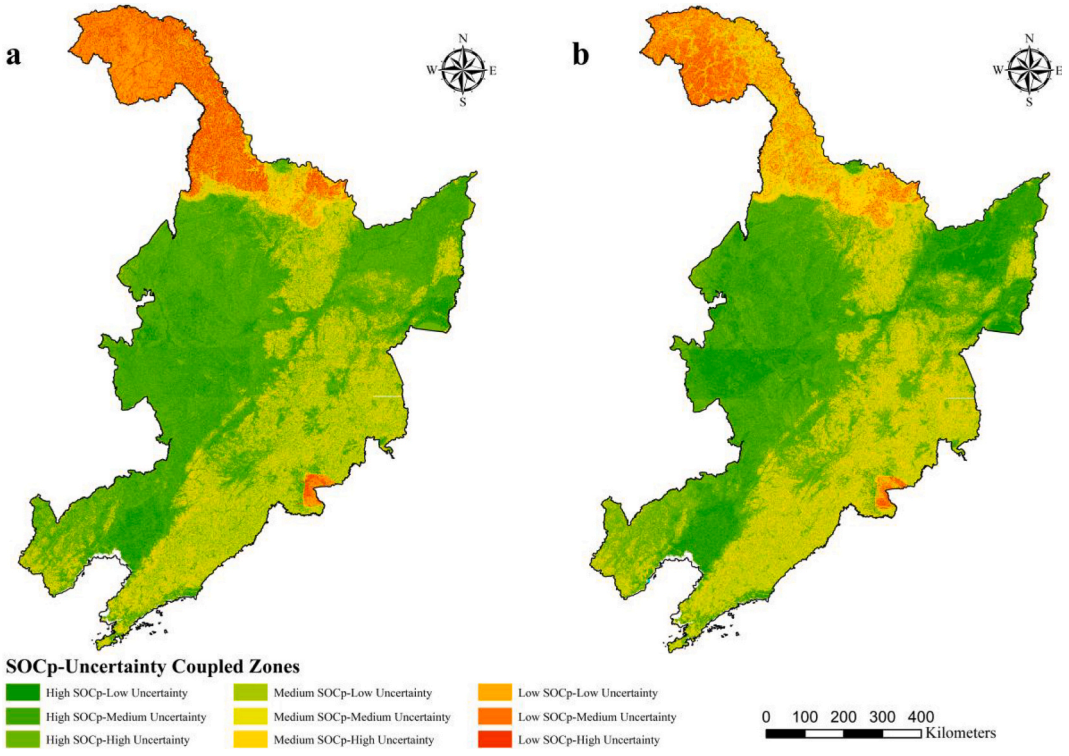

Fig. 6. Spatial distribution of soil organic carbon sequestration potential (SOCp)a and coupled prediction uncertainty zones in Northeast China (a for Hassink model; b for Elustondo model).

influence the carbon cycle of terrestrial ecosystems (Don et al., 2011).

2.3.4. *Vegetation variables*

Based on topographic and climatic variables, this study further incorporated the normalized difference vegetation index (NDVI) and net primary production (NPP), which reflected vegetation growth status and productivity levels, respectively. The NDVI data were derived from the MOD13A3 product officially released by NASA, which is based on observations from the MODIS sensor onboard the Terra satellite (https://doi.org/10.5067/MODIS/MOD13A3.061), NPP data obtained

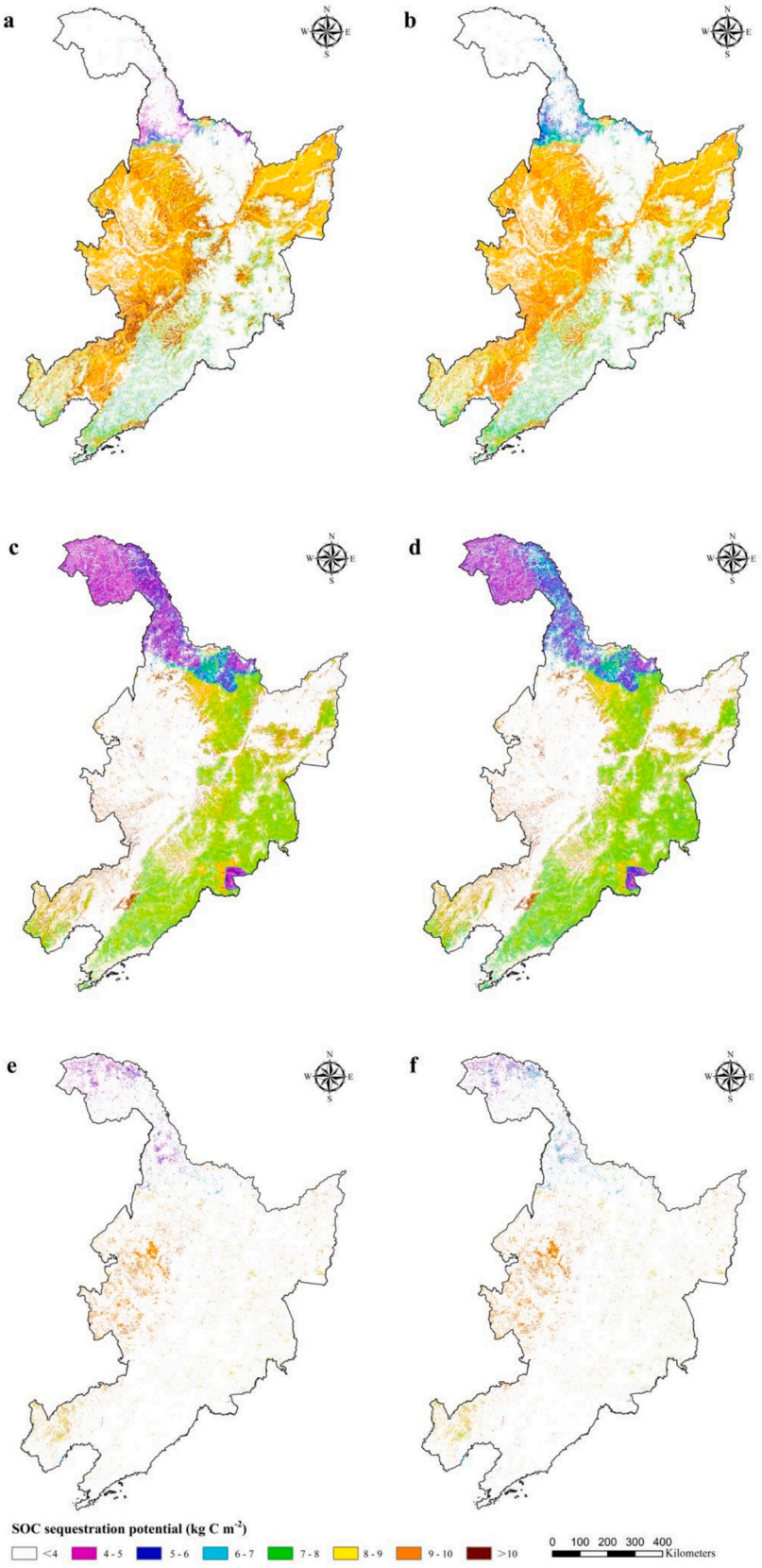

Fig. 7. Spatial distribution of soil organic carbon sequestration potential (SOCp) in different ecosystems (a for Hassink model-Farmland; b for Elustondo model-Farmland; c for Hassink model-Forest; d for Elustondo model-Forest; e for Hassink model-Grassland; f for Elustondo model-Grassland).

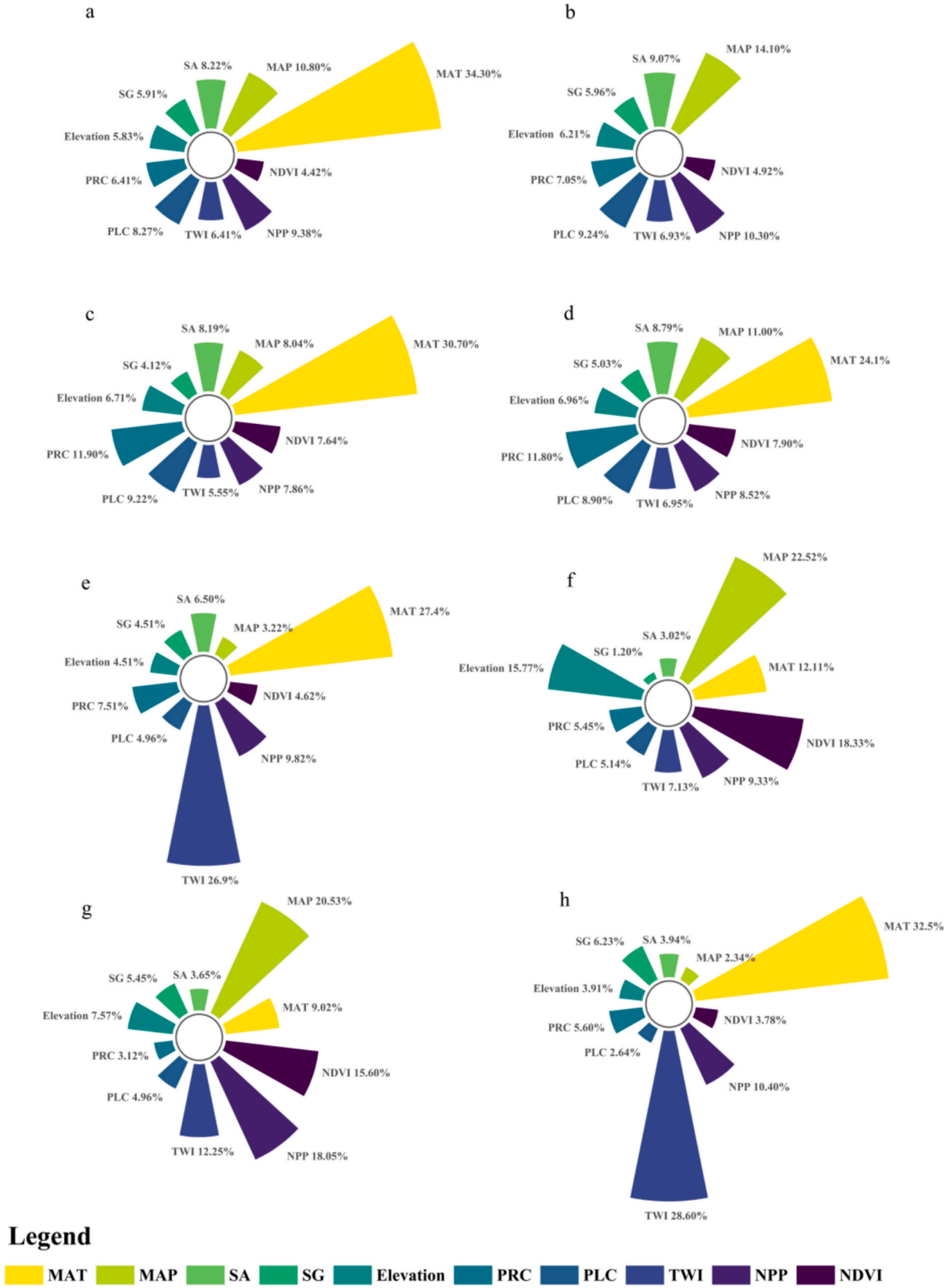

Fig. 8. The relative importance of environmental variables of soil organic carbon sequestration potential in the study area (a for Hassink model-Entire study area; b for Elustondo model-Entire study area; c for Hassink model-Farmland; d for Elustondo model-Farmland; e for Hassink model-Forest; f for Elustondo model-Forest; g for Hassink model-Grassland; h for Elustondo model-Grassland).

from the Resource and Environmental Science Data Center of the Chinese Academy of Sciences (https://www.resdc.cn/data.aspx?DATAID=204), were covering the period from 2000 to 2023, with a spatial resolution of 1 km. The study area includes the three provinces of Northeast China. A time synthesis algorithm based on the weighted average method was applied to calculate the long-term average, and the data was resampled from its original 1 km resolution to a 90 m grid resolution using bilinear interpolation. This resampling method was solely employed to meet the spatial grid data matching requirements of this study and did not actually increase the information content of the original data. NDVI serves as a key indicator for assessing vegetation coverage and growth conditions, while NPP is a core parameter for measuring vegetation productivity. The synergistic relationship between NDVI and NPP significantly influences SOCs and had been widely utilized in SOC-related research in recent years (Zheng et al., 2019; Zhang et al., 2024).

2.4. Calculation method for soil organic carbon-related data

2.4.1. Calculation of SOC_{stock}

Soil organic carbon stock (SOC_{stock}) refers to the mass of organic carbon contained in a unit area or volume of soil, and is a key indicator for measuring soil carbon sequestration capacity and carbon sequestration potential. This study uses the following formula to calculate SOC_{stock}:

$$SOC_{density} = p \times SOC \times BD \times (1 - ce) \times 10^{-2} \quad (1)$$

$$SOC_{stock} = S_i \times SOC_{density} \quad (2)$$

where $SOC_{density}$ is the soil organic carbon density (kg C m^{-2}); P is the actual thickness of the soil layer (cm); SOC is the soil organic carbon content (g C kg^{-1}); BD is the bulk density (kg m^{-3}); ce is the content of gravel with a diameter greater than 2 mm, measured in volume ratio (%); S_i is the area (m^2) corresponding to the i-th land type; SOC_{stock} is the soil organic carbon stocks (kg C).

2.4.2. Calculation of SOCs

Soil organic carbon saturation (SOCs) refers to the theoretical maximum capacity of soil to store organic carbon under stable environmental conditions. In this study, the Hassink model and the Elustondo model were employed to estimate SOCs levels, the demonic coefficient in this study strictly adheres to the findings of the original literature research. Hassink (1997) demonstrated that soil retains SOC through the binding of organic matter with clay and silt particles; where the relative proportion of clay and silt significantly influences the soil's carbon sequestration potential. A SOCs model applicable to both tropical and temperate soils was proposed. Elustondo et al. (1990) found a positive correlation between soil clay content and the storage of both carbon and nitrogen. The Hassink model defines 20 μm as the upper limit of soil particle size, whereas the Elustondo model uses 50 μm as the upper limit. Since Hassink was the first to propose the strong relationship between fine soil particles and soil organic matter, many early studies adopted the 20 μm threshold. However, subsequent research has shown that the Hassink model remains valid when using 50 μm as the upper limit. Therefore, recent studies predominantly use 50 μm as the standard particle size cutoff, which not only harmonizes the scale between the two models but also facilitates direct comparisons of organic matter content across soil particle sizes. This study applied both the Hassink and Elustondo models to predict SOC saturation values and compared the carbon sequestration potentials estimated by each model. The objective was to evaluate the performance of different models on the same dataset to more scientifically and rigorously assess the soil carbon sequestration potential in the study area. The calculation formulas for the two models are:

$$SOC_{sat-Hassink} = p \times (4.09 + 0.37 \times finefraction) \times BD \times (1 - ce) \times 10^{-2} \quad (3)$$

$$SOC_{sat-Elustondoo} = p \times (9.04 + 0.27 \times finefraction) \times BD \times (1 - ce) \times 10^{-2} \quad (4)$$

where $SOC_{sat\text{-}Hassink}$ represents the soil organic carbon saturation value calculated using the Hassink model (units: kg C m^{-2}); $SOC_{sat\text{-}Elustondo}$ denotes the soil organic carbon saturation value derived from the Elustondo model (units: kg C m^{-2}); *finefraction* indicates the proportion of fine particles in the soil with a particle size smaller than 50 μm; P is the actual thickness of the soil layer (cm); ce is the content of gravel with a diameter greater than 2 mm, measured in volume ratio (%); S_i is the area (m^2) corresponding to the i-th land type; SOC_{stock} is the soil organic carbon stocks (kg C).

2.4.3. Calculation of SOC_p

Soil organic carbon sequestration potential (SOC_p) refers to the maximum capacity of soil to capture and store additional carbon dioxide over a specific period through human-induced management practices or natural restoration processes. This potential primarily depends on the accumulation and stabilization of soil organic matter. During the decomposition of organic matter, although some carbon is released into the atmosphere as carbon dioxide, stable organic compounds are also formed, which can remain stored in the soil over long periods, thereby achieving effective carbon sequestration. Soil carbon sequestration potential not only serves as a crucial indicator for ensuring national food security, but also represents a key strategy for mitigating climate change. The core mechanism involves reducing the greenhouse effect by increasing the total SOC content, while simultaneously enhancing soil fertility and supporting ecosystem services (Lal, 2003). SOC_p is commonly quantified by calculating the difference between the current SOCs and its saturation level. This study utilizes both the Hassink model and the Elustondo model to estimate carbon sequestration potential under different assumptions, which are labeled as Hassink and Elustondo models, respectively.

$$SOC_{p-Hassink} = SOC_{sat-Hassink} - SOC_{density} \quad (5)$$

$$SOC_{p-Elustondo} = SOC_{sat-Elustondo} - SOC_{density} \quad (6)$$

where $SOC_{p\text{-}Hassink}$ denotes the soil organic carbon sequestration potential per unit area under the Hassink model (kg C m^{-2}); $SOC_{p\text{-}Elustondo}$ denotes the soil organic carbon sequestration potential per unit area under the Elustondo model (kg C m^{-2}).

2.5. Boosted regression trees

Boosted regression trees (BRT) is an ensemble machine learning method that integrates decision trees with gradient boosting techniques. The core concept involves iteratively training a sequence of weak learners, typically shallow decision trees, where each subsequent tree aims to correct the residual errors of its predecessor. The final predictive model is obtained by aggregating the outputs of all individual trees through weighted summation. BRT is particularly effective at capturing nonlinear relationships and interaction effects, while also accommodating missing data. As a result, it is widely applied in modeling and predicting complex relationships within fields such as ecology, environmental science, and geography. This study is based on the R language (4.4.2) platform and uses the BRT model to explore the nonlinear relationship between SOC_{stock}, soil carbon sequestration potential predicted by Hassink and Elustondo models, and ten environmental driving factors (including terrain factors: altitude, slope, aspect, plane curvature, profile curvature, and terrain humidity index; climate factors: annual average temperature and annual precipitation; vegetation factors: normalized vegetation index and net primary productivity). In the

process of model construction, systematic optimization and setting were carried out for key hyperparameters of BRT: tree complexity (interaction.depth) controls the ability to capture interactions between variables, avoiding the trade-off between overfitting and underfitting; The learning rate (shrinkage) affects the contribution weight of each tree, and lower values can improve stability but increase computational burden; The number of trees (n.trees) and learning rate are synergistically regulated, jointly affecting the convergence and fitting strength of the model; Bag fraction improves the model's noise resistance through random sampling at each iteration, especially for ecological observation data; The minimum number of observations (n.minobsinnode) limits the sample size of leaf node splitting to enhance generalization performance. After multiple rounds of testing and cross validation, the final Hassink model parameters were set to: n.trees = 1600 , interaction.depth = 9 , shrinkage = 0.0025 , n.minobsinnode = 10 , bag.fraction = 0.65 ; Elustondo model parameters were set to: n.trees = 1850 , interaction.depth = 8 , shrinkage = 0.0025 , n.minobsinnode = 10 , bag.fraction = 0.60. This parameter combination effectively serves the modeling needs of complex relationships between multi-source environmental variables and soil carbon indicators, while balancing model interpretability and prediction robustness.

2.6. Model verification

To ensure the scientific rigor and accuracy of model evaluation, this study employs a 10-fold cross-validation approach to comprehensively assess the performance of the BRT model across three dimensions: goodness of fit, error characteristics, and consistency. Specifically, four evaluation indicators are selected. The coefficient of determination (R^2) reflects the model's ability to explain the variability of the target variable (Chicco et al., 2021). Root Mean Square Error (RMSE) quantifies the magnitude of prediction errors and is particularly sensitive to large deviations (Hodson, 2022). Mean Absolute Error (MAE) measures the average magnitude of errors and is more robust to outliers (Hodson, 2022). Lin's Consistency Correlation Coefficient (LCCC) evaluates both the correlation and agreement between predicted and observed values (Willmott and Matsuura, 2005). By implementing ten-fold cross-validation, the influence of random data partitioning is minimized, the risk of overfitting is reduced, and the overall reliability of the evaluation results is enhanced. The specific calculation formula are as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|a_i - b_i| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(a_i - b_i)^2} \tag{8}$$

$$R^2 = \frac{\sum_{i=1}^{n}(a_i - \overline{b}_i)^2}{\sum_{i=1}^{n}(b_i - \overline{b}_i)^2} \tag{9}$$

$$LCCC = \frac{2r\partial_a\partial_b}{\partial_a^2 + \partial_b^2 + (\overline{a} - \overline{b})^2} \tag{10}$$

where a_i and b_i denote the observed and predicted values of soil organic carbon sequestration potential (SOC_p) at the i sampling location, respectively. The total number of samples is represented by n. The variances of the observed and predicted values are expressed as ∂a and ∂b, respectively. The Pearson correlation coefficient, r, is utilized to quantify the linear association between the observed and predicted datasets.

3. Results

3.1. Descriptive statistics

The SOC_p calculated via the Hassink and Elustondo models, together with data of 10 environmental variables, were statistically analyzed at the sampling point scale (Table 1). Results showed that the average SOC_p from the Elustondo model was 12% higher than that from the Hassink model, with a wider extreme value range and larger standard deviation, indicating greater data volatility, higher dispersion, and a more pronounced right-skewed distribution with concentrated high values. Both models had a kurtosis of 2.57, presenting flatter distributions than the normal distribution with more scattered tails. Overall, the Elustondo model was more suitable for predicting high-potential or extreme scenarios but with higher uncertainty, while the Hassink model exhibited higher stability and was better adapted for conservative estimations.

Pearson correlation analysis (Table 2) revealed minimal correlation differences between the two models, supporting their combination for subsequent analysis. Key environmental variables showed distinct correlations with SOC_p: MAT had a significant positive correlation with both models ($|r| \geq 0.1$), and NDVI was significantly negatively correlated with elevation. NPP, SG, and MAP displayed weak positive correlations with the models, while PLC, PRC, and TWI showed weak negative correlations. Some of these relationships deviated from conventional expectations, possibly due to insufficient data standardization, high heterogeneity from multi-ecosystem inclusion, and potential indirect variable interactions. Most other environmental variables exhibited weak to moderate positive correlations. To improve model accuracy and scientific rigor, the BRT method was adopted for further analysis. (See Table 3.)

3.2. Performance analysis of model predictions

The performance comparison of the models revealed that the R^2 value of the Hassink model was marginally higher than that of the Elustondo model by approximately 0.012, explaining 75% and 74% of the spatial variability, respectively. This suggested that Hassink demonstrates a slightly better capacity in capturing the spatial variability of SOC_p. The interquartile range for both models was 0.005, with a mere difference of 0.001 between the first and third quartiles. The scores were highly concentrated, suggesting that both models exhibited exceptional stability across multiple iterations and were minimally affected by random variations. In terms of MAE, Hassink model performed better than Elustondo model, with a difference of about 0.19. The small differences between MAE and RMSE for Hassink model (0.27) and Elustondo model (0.30) further suggested a relatively concentrated error distribution without significant skewness. The LCCC values for both models were close to 0.9, with Hassink model showing a slightly higher value by 0.002, reflecting excellent consistency and predictive adaptability. RMSE analysis also showed that Hassink model had a lower error (by 0.22), indicating smaller prediction errors and stronger robustness and generalization ability, particularly for extreme values. When extended to farmland, forest, and grassland ecosystems, the Hassink and Elustondo models maintained performance trends consistent with the overall analysis, while showing ecosystem-specific nuances. Across all three systems, Hassink's R^2 (0.75) remained marginally higher (by 0.01) than Elustondo's (0.74), explaining 75% and 74% of SOC_p spatial variability respectively; core metrics (e.g., R^2, LCCC) exhibited a 0 interquartile range (identical 1st and 3rd quartile values) for both models, with highly concentrated scores signaling exceptional iterative stability and minimal sensitivity to random fluctuations. In error metrics, Hassink outperformed Elustondo across systems: its MAE was 0.33 (farmland), 0.20 (forest), and 0.18 (grassland) lower, while its RMSE was 0.23 (farmland/forest) and 0.22 (grassland) lower—indicating smaller prediction deviations and concentrated error distributions. Their LCCC values (near 0.88–0.89) were closely matched, with Hassink holding a slight consistency edge ($\leq$0.01). These results reaffirm that the Hassink model demonstrates greater stability than the Elustondo model across different terrestrial ecosystems, though their core performance metrics show little overall difference.

Table 1
Summary statistics of soil organic carbon sequestration potential (SOCp) and environmental variables at sampling sites in the study area ($n = 293$).

Property	Unit	Min.	Max.	Mean	SD	Skewness	Kurtosis
$SOC_{p\text{-}Hassink}$	kg C m^{-2}	1.10	12.68	5.64	2.57	0.42	2.57
$SOC_{p\text{-}Elustondo}$	kg C m^{-2}	0.61	14.60	6.33	2.96	0.44	2.57
Elevation	m	1.00	1682.00	225.12	169.25	2.84	21.12
PLC		−0.06	0.06	0.00	0.01	−0.12	11.45
PRC		−0.03	0.11	0.00	0.01	2.61	17.70
SG	degree	0.02	15.64	1.49	1.96	3.12	17.27
SA	degree	4.40	359.42	176.28	99.15	0.01	1.90
TWI		6.41	13.31	10.36	1.41	−0.22	2.40
MAT	Celsius degree	−5.75	8.50	4.71	1.97	−0.57	4.74
MAP	mm	339.58	1226.25	506.58	129.25	1.53	6.63
NDVI		0.28	0.90	0.81	0.10	−2.01	8.53
NPP	kg C m^{-2} $year^{-1}$	1312.00	8120.00	3760.32	951.18	0.87	5.03

Note: SG, slope gradient; SA, slope aspect; PLC, plan curvature; PRC, profile curvature; TWI, topographic wetness index; NDVI, normalized difference vegetation index; NPP, net primary production; MAT, mean annual temperature; and MAP, mean annual precipitation.

Table 2
Pearson correlation coefficient between environmental variables at the sampling points in the study area.

Property	$SOC_{p\text{-}Hassink}$	$SOC_{p\text{-}Elustondo}$	Elevation	PLC	PRC	SG	SA	MAT	MAP	NDVI	NPP
$SOC_{p\text{-}Elustondo}$	0.96										
Elevation	−0.14	−0.12									
PLC	−0.02	0.01	0.06								
PRC	−0.03	−0.04	0.12	−0.49							
SG	0.06	0.04	0.36	0.03	0.48						
SA	−0.04	−0.03	−0.13	−0.03	0.04	−0.09					
TWI	−0.02	−0.02	−0.56	−0.01	−0.32	−0.7	0.07				
MAT	0.52	0.43	−0.27	0.01	0.02	0.2	−0.01	−0.2			
MAP	0.01	0.03	0.24	0.05	0.2	0.35	0.04	−0.44	0.12		
NDVI	−0.19	−0.14	0.16	0.04	0.05	0.04	−0.02	−0.05	−0.36	0.15	
NPP	0.05	0.06	0.35	0.05	0.25	0.52	0	−0.67	0.24	0.61	0.16

Note: SG, slope gradient; SA, slope aspect; PLC, plan curvature; PRC, profile curvature; TWI, topographic wetness index; NDVI, normalized difference vegetation index; NPP, net primary production; MAT, mean annual temperature; and MAP, mean annual precipitation.

To further reveal the predictive performance of the BRT model in predicting Hassink model and Elustondo model, we plotted a histogram of RMSE (Fig. 3). The RMSE distribution histogram plot further demonstrated that Hassink model exhibited smaller errors with a more concentrated distribution, ranging from 1.26 to 1.28, and a primary peak around 1.27, indicating a symmetrical pattern. In contrast, the RMSE values for Elustondo model ranged from 1.49 to 1.50, with a main peak near 1.49 and a slightly higher proportion of elevated error values. Overall, the Hassink model demonstrated superior accuracy and stability, making it more appropriate for high-precision prediction of soil carbon sequestration potential.

The uncertainty in model prediction performance was reflected by the standard deviation of the BRT model's 100 iterative prediction results (Fig. 5). The standard deviation of the Hassink model was predominantly within the range of 0.03–0.06 kg C m^{-2}, with over 70% of the area exhibiting low fluctuations. This suggested that the Hassink model produced consistent prediction results and demonstrates high stability across the 100 iterations. Moreover, the regions with low standard deviation were concentrated in the central Songnen Plain, which aligns with the homogeneity characteristics of black soil. In contrast, the standard deviation of the Elustondo model was mainly distributed between 0.07 and 0.09 kg C m^{-2}, with more than 60% of the area showing high variability. This indicates that the Elustondo model exhibited greater fluctuations in its prediction outcomes after 100 iterations, although it still maintains a moderate level of overall stability. The fundamental difference in standard deviation between the two models stolen from their varying sensitivities to input variable changes, which were specifically reflected in differences in model slope and intercept.

To more precisely identify the response relationship between the spatial distribution of soil organic carbon sequestration potential (SOCp) and the stability of its prediction results, this study conducted a two - dimensional coupled superposition analysis of SOCp and prediction uncertainty. Based on the distribution characteristics of these two indicators, we carried out a three - level scientific classification of both indicators using the mean deviation method, established a coupled classification system of SOCp and prediction result stability, and plotted the spatial distribution map of two-dimensional coupled zoning (Fig. 6). This work realized the spatial integration of areas with high SOCp and areas with low prediction uncertainty, and accurately pinpointed high-confidence priority carbon management hotspots.

The classification results indicate that both the predicted soil organic carbon (SOCp) and the prediction uncertainty in the study area display significant spatial differentiation characteristics. The nine types of zones resulting from their coupling demonstrate an evident zonal distribution pattern. Specifically, based on the simulation results of the two models, the high-confidence priority carbon management hotspots (High SOCp—Low Uncertainty Zone) are predominantly continuously distributed in the western and southwestern regions of the study area, with sporadic distribution in the northeastern part. This type of zone simultaneously features extremely strong potential for enhancing carbon sinks, extremely high reliability of prediction results, and optimal prediction stability. It serves as the core area of focus for implementing regional soil carbon sink management policies.

3.3. Soil carbon sequestration potential and its spatial distribution across different ecosystems

Based on the BRT model predictions integrated with land use data within the ArcGIS platform, this study quantified the SOC_p and areal extent of major ecosystems—farmland, forestland, and grassland—across Northeast China, which covers a total area of 790,605 km^2. Farmland accounted for 39.3% of the region, followed by forestland (42.8%) and grassland (4.2%). Significant discrepancies were observed

Table 3
Evaluation of BRT model with 100 iterations for predicting SOCp of topsoil in Northeast China.

Model	Index	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Hassink model	R^2	0.75	0.75	0.75	0.75	0.75	0.75
	RMSE	1.27	1.27	1.27	1.27	1.27	1.28
	LCCC	0.89	0.89	0.89	0.89	0.89	0.89
	MAE	0.99	1.00	1.00	1.00	1.00	1.00
Elustondo model	R^2	0.74	0.74	0.74	0.74	0.74	0.74
	RMSE	1.49	1.49	1.49	1.49	1.50	1.50
	LCCC	0.89	0.89	0.89	0.89	0.89	0.89
	MAE	1.19	1.19	1.19	1.19	1.20	1.20
Hassink model (Farmland ecosystem)	R^2	0.75	0.75	0.75	0.75	0.75	0.75
	RMSE	1.42	1.42	1.42	1.42	1.42	1.43
	LCCC	0.89	0.89	0.89	0.89	0.89	0.89
	MAE	1.11	1.12	1.12	1.12	1.13	1.13
Elustondo model (Farmland ecosystem)	R^2	0.74	0.74	0.74	0.74	0.74	0.74
	RMSE	1.65	1.65	1.65	1.65	1.65	1.65
	LCCC	0.89	0.89	0.89	0.89	0.89	0.89
	MAE	1.32	1.32	1.32	1.32	1.32	1.32
Hassink model (Forest ecosystem)	R^2	0.75	0.75	0.75	0.75	0.75	0.75
	RMSE	1.15	1.15	1.15	1.15	1.15	1.15
	LCCC	0.88	0.88	0.89	0.89	0.89	0.89
	MAE	0.90	0.90	0.90	0.90	0.90	0.90
Elustondo model (Forest ecosystem)	R^2	0.74	0.74	0.74	0.74	0.74	0.74
	RMSE	1.38	1.38	1.38	1.38	1.38	1.38
	LCCC	0.88	0.88	0.88	0.88	0.88	0.88
	MAE	1.10	1.10	1.10	1.10	1.10	1.10
Hassink model (Grassland ecosystem)	R^2	0.75	0.75	0.75	0.75	0.75	0.75
	RMSE	1.27	1.27	1.27	1.27	1.27	1.27
	LCCC	0.87	0.88	0.89	0.89	0.89	0.89
	MAE	1.01	1.01	1.01	1.01	1.01	1.01
Elustondo model (Grassland ecosystem)	R^2	0.74	0.74	0.74	0.74	0.74	0.74
	RMSE	1.49	1.49	1.49	1.49	1.49	1.49
	LCCC	0.88	0.88	0.88	0.88	0.88	0.89
	MAE	1.19	1.19	1.19	1.19	1.19	1.19

Note: Min., minimum; 1st Qu., first quantile; 3rd Qu., third quantile; Max., maximum.

between the Hassink and Elustondo models across all ecosystems. To systematically analyze the overall soil carbon sequestration potential prediction results for the study area, this research integrated the prediction outcomes of the two models using both arithmetic and weighted averaging methods. Upon comparison, it was found that the Pearson correlation coefficients of both models were consistent, with an R-squared difference of only 0.01. Additionally, the results from the two averaging methods were similar, with the total SOC_p differing by only 2.82 Tg C. The arithmetic mean method not only eliminates the inherent systematic bias of a single model but also maximally preserves the common patterns between the two models, thereby enhancing the reliability and generalizability of the prediction results, all while ensuring no loss of core effective information (Rukhin and Vangel, 1998; Jacquier et al., 2003). Therefore, this study ultimately selected the integrated results from the arithmetic mean method as the final predicted value for the total soil organic carbon sequestration potential in the study area (Table 4). For farmland, the unit SOC_p was 7.63 kg C m^{-2} (Hassink) and 8.76 kg C m^{-2} (Elustondo), a difference of 14.9%, translating to a total SOC_p difference of 353.76 Tg C. Similarly, forestland showed unit values of 6.06 and 7.04 kg C m^{-2} (16.1% difference, 330.68 Tg C total difference), and grassland 6.68 and 7.82 kg C m^{-2} (17.1% difference, 38.3 Tg C total difference). Although both models exhibited comparable standard deviations (0.82–0.87 kg C m^{-2}), the Elustondo model demonstrated a higher coefficient of variation (9.9%) in farmland, indicating greater prediction variability. The integrated mean SOC_p across the region was 5831.50 Tg C, with an average unit potential of 7.38 kg C m^{-2}. Farmland contributed the most (43.7%, 2547.51 Tg C) and exhibited the highest unit potential (8.20 kg C m^{-2}) and spatial homogeneity, underscoring its critical role in regional carbon sequestration strategies. Forestland, covering the largest area, contributed 38.0% (2213.99 Tg C) but had the lowest unit potential (6.55 kg C m^{-2}), while grassland, though efficient per unit area (7.25 kg C m^{-2}), contributed only 4.2% (243.70 Tg C) due to its limited spatial extent. It is worth noting that due to human disturbance, the predicted high SOC_p values in agricultural ecosystems represent only the theoretical upper limit under ideal conditions, not the actual potential under current natural conditions.

Spatial distribution of SOC_p, derived from R-outputted TIF files and overlaid with land use data in ArcGIS (Fig. 7), revealed a concentric pattern aligned with topographic gradients: higher values occurred predominantly in the central and southern Songnen Plain, the Sanjiang Plain, and the northern Liaohe Plain—regions characterized by fertile black soils. In farmland, high-potential zones (>9 kg C m^{-2}) concentrated in these core black soil areas, medium-level potentials (7–9 kg C m^{-2}) formed circular patterns across central Jilin, central Liaoning, and eastern Heilongjiang, and low-potential areas (<7 kg C m^{-2}) were mainly distributed in the Greater Khingan Mountains, Changbai Mountain area, and western Liaoning hills. The Hassink model produced more fragmented spatial patterns with abrupt transitions and generally lower estimates in the north—approximately 20% below Elustondo—-while the Elustondo model yielded smoother gradients and broader boundaries, reflecting higher sensitivity to extreme values.

Farmland SOC_p was spatially dominated by black soil regions, which formed the core of high-potential distribution areas within the study region (Table 4). Coupled with the combined effects of temperature variations, differences in cropping systems, and freeze-thaw erosion in northern areas, the soil organic carbon potential showed a gradual declining trend.In the prediction of cropland SOC_p, the Hassink model focused more on macro-level prediction and exhibited more stable performance in estimating the overall carbon sequestration potential of the study region; however, the patchiness of its estimation results was more pronounced compared to the Elustondo model. In contrast, the Elustondo model emphasized small-scale plot-level prediction and demonstrated particularly outstanding performance in predicting the extremes of carbon sequestration potential. Nevertheless, it showed smoother transitions between zones with different carbon sequestration potentials and lower sensitivity in boundary prediction. Forestland SOC_p generally decreased northward, with high-value areas south of Changbai Mountain, medium values in the Lesser Khingan and surrounding Changbai regions, and low values in the northern Greater Khingan

Table 4
Soil organic carbon sequestration potential (SOC_p) of topsoil in different ecosystems in Northeast China.

Ecosystem	Area (km^2)	Hassink model		Elustondo model		(Hassink + Elustondo)/2		Hassink & Elustondo R^2-weighted	
		Mean (kg C m^{-2})	SOC_p (Tg C)	Mean (kg C m^{-2})	SOC_p (Tg C)	Mean (kg C m^{-2})	SOC_p (Tg C)	Mean (kg C m^{-2})	SOC_p (Tg C)
Farmland ecosystem	310,862	7.63	2370.63	8.76	2724.39	8.20	2547.51	8.19	2546.32
Forest ecosystem	338,116	6.06	2048.65	7.04	2379.32	6.55	2213.99	6.55	2212.88
Grassland ecosystem	33,600	6.68	224.55	7.82	262.85	7.25	243.70	7.25	243.57
Other ecosystems	108,027	7.09	766.28	8.20	886.34	7.65	826.31	7.64	825.91
Total	790,605	6.84	5410.11	7.91	6252.90	7.38	5831.50	7.37	5828.68

Note: Hassink model ($R^2 = 0.75$)、Elustondo model ($R^2 = 0.74$).

Mountains and alpine forests. The Hassink model generated conservative, patchy estimates suitable for macro-level planning, whereas the Elustondo model predicted more extensive high-value areas with gradual transitions, ideal under optimized environmental conditions. Grassland SOC_p was higher in central and western regions but spatially fragmented. High-value zones formed continuous belts in the Songnen Plain, with medium and low values interwoven in radial or strip patterns. The Hassink model emphasized natural continuity with sharp boundaries, supporting robust large-scale estimates, while the Elustondo model produced discrete, uniformly distributed units with smoothed transitions, better capturing local extremes but lacking spatial sensitivity. The high spatial agreement in core zones between models confirms overall consistency, yet model choice must align with specific objectives: the Hassink model is preferable for conservative, natural-factor-driven projections at regional scales, whereas the Elustondo model offers superior detail for grid-level ideal-condition simulations.

3.4. Main influencing variables of soil carbon sequestration potential across different ecosystems

The evaluation of the relative importance (RI) of 10 environmental variables in predicting the topsoil SOC_p, based on the BRT model (100 iterations), indicated that the MAT, MAP, and NPP were the dominant variables influencing SOC_p in Northeast China, collectively contributing approximately 55% to the total RI (Fig. 8). In the two models, the RI of MAT was significantly higher in the Hassink model (34.3%) compared to the Elustondo model (26.2%), underscoring the critical influence of temperature on SOC_p. MAP exhibited relatively high importance (14.1%) in the Elustondo model, reflecting its directed effect on carbon input. Regarding topographic variables, SA had a greater influence than SG. Additionally, PLC and PRC demonstrated higher contributions in the Elustondo model, while the TWI showed comparable influence in both models. The RI of the NDVI was relatively low (approximately 4.5%), suggesting its role in indicating vegetation cover status rather than directly influencing carbon sequestration potential.

To further identify the dominant environmental variables within each ecosystem, we applied the BRT model using data from sampling sites distributed across three distinct ecosystems, enabling us to determine the key environmental variables specific to each ecosystem. In different ecosystems, dominant variables varied significantly: farmland ecosystem was primarily influenced by MAT, PRC, and MAP, reflecting the synergistic effects of climatic and topographic variables; forestland was mainly governed by MAT and TWI, indicating the coupled response of carbon accumulation to temperature and hydrological conditions; grassland carbon sequestration potential was predominantly regulated by MAP and vegetation productivity (NDVI/NPP). The Hassink model demonstrates better predictive performance for farmland ecosystem (MAT), forestland ecosystem (TWI), and grassland ecosystem (MAP), making it suitable for temperature-sensitive and topographically complex regions. In contrast, the Elustondo model was more appropriate for scenarios where the synergistic effects of MAP and NPP were critical, such as in the fine management of farmland. These findings reveal how environmental factors differentially influence SOC_p across ecosystems, offering a scientific basis for the formulation of region-specific carbon management strategies.

4. Discussion

4.1. Spatial distribution characteristics of SOC_p

Based on the spatial distribution of soil organic carbon sequestration potential (SOC_p) in the study area (Figs. 4–5), it could be concluded that SOC_p in Northeast China exhibits significant spatial heterogeneity, which is generally characterized by the pattern of “high in the southern plains and low in the northern mountains”. Specifically, the range of SOC_p values was approximately 4–10 kg C m^{-2}. The Songnen Plain, Sanjiang Plain, and Liaohe Plain were identified as the core regions with high SOC_p. In contrast, the northern part of the Greater Khingan Range and the Hulunbuir Plateau served as the primary zones with low SOC_p, while the hilly and mountainous areas in the east and west functioned as transitional zones with moderate SOC_p. By comparing the north-south spatial gradient characteristics of SOC_p, it was evident that the high-SOC_p zones in the southern region were mainly concentrated in the black soil areas of the Songnen Plain, the central parts of the Sanjiang Plain, and the Liaohe Plain, presenting a continuous patchy spatial distribution. Conversely, the low-SOC_p zones in the northern region covered high-altitude areas (e.g., Mohe in the Greater Khingan Range). The significant north-south disparity in SOC_p is primarily attributed to the lower mean annual temperature in the north, which directly intensifies the freeze-thaw cycle and consequently accelerates the loss of soil organic carbon (Portner et al., 2010). Furthermore; as latitude increases (i.e.; moving northward); the proportion of soil clay particles decreases; leading to a gradual decline in the soil's physical capacity to sequester organic carbon (Churchman et al., 2020). In contrast; the central and southern regions—characterized by higher SOC_p—benefit from deep black soil layers and extensive arable land(Yu et al., 2006). Specifically; crop rotation systems involving corn; soybeans; and other crops enhance root-derived carbon input and increase surface SOC content (Luo et al., 2024). Additionally; the synergistic effect of water-heat coupling in these regions effectively minimizes the microbial decomposition of SOC (Li et al., 2023).

From the comparison of the SOC_p prediction results of the two models, the Hassink model demonstrates higher spatial continuity and clarity in predicting SOC_p. High-SOC_p areas predicted by this model were concentrated in the central-western region, with well-defined geographic boundaries. Medium-SOC_p areas formed a circular gradient zone surrounding the high-SOC_p areas, while radial striping (corresponding to the spatial distribution of rivers) was observed in the northeastern region. Low-SOC_p areas were characterized by continuous patchy distributions in the northern part of the study area, with an isolated low-SOC_p zone located in the southeast corner. In contrast, the Elustondo model exhibits distinct prediction characteristics. High-SOC_p areas predicted by this model were more spatially fragmented and distributed in a honeycomb grid-like pattern. Medium- and high-SOC_p areas were spatially interlaced and embedded with each other, forming relatively smooth striped boundaries. Compared to the Hassink model, the area of the northern patchy low-SOC_p zones predicted by the Elustondo model is reduced by approximately 30–40%, and these low-SOC_p zones are spatially interwoven with medium-SOCp areas. Overall, the spatial distribution map of SOCp generated by the Hassink model better reflects the spatial influence of natural geographic elements. This is because its potential distribution aligns closely with geographical element boundaries and features clear demarcations. Although the Elustondo model exhibits pronounced fragmentation and overly smooth boundaries—thereby weakening the influence of natural geographical elements—it demonstrates greater sensitivity to extreme values of soil carbon sequestration potential and is more suitable for analyzing extreme conditions caused by non-natural factors, such as human activities.

The main reasons for the above results can be attributed to the following points: (1) Long-term water erosion and tillage disturbance lead to surface carbon loss in the eroded areas of black soil slope farmland, resulting in an insufficient state of carbon saturation. However, due to the strong adsorption capacity of soil particles for organic carbon, these areas still exhibit relatively high carbon sequestration potential (Lal et al., 2015). (2) The low average annual temperature and concentrated precipitation in Northeast China cause frequent soil freezing and thawing cycles; which damage soil aggregate structure and reduce the carbon sequestration potential of SOC (Matzner and Borken, 2008; Song et al., 2017). Additionally, variations in altitude, precipitation, and temperature significantly influence the spatial distribution characteristics of soil carbon sequestration potential (Monson et al.,

2002; Campo and Merino, 2016; Huang et al., 2018). (3) Microorganisms also play a role in soil carbon sequestration potential; erosion reduces microbial carbon utilization efficiency; further decreasing carbon sequestration capacity (Kheirfam, 2020; Jiang et al., 2022). Therefore, the combined effects of natural and anthropogenic factors jointly shape the spatial distribution of soil carbon sequestration potential (Smith, 2005; Steger et al., 2019).

This study identified climatic factors (mean annual temperature, mean annual precipitation) as the dominant drivers of soil organic carbon sequestration potential (SOC_p). Based on this finding, targeted zoning management measures were developed for areas with different SOC_p levels, and the practical implications of this conclusion were further clarified: given that climatic factors dominate SOC_p dynamics, quantifying the achievable potential of enhancing SOC_p through agricultural management practices (e.g., conservation tillage, organic fertilizer application) under climatic constraints is critical, which enhances the practical guiding significance of the study's conclusions. In areas with high SOC_p, it is recommended to prioritize the protection of black soil zones by reducing tillage intensity to lower the risk of water erosion (You et al., 2020); while implementing full straw return and rational application of organic fertilizers to enhance soil carbon storage (Lessmann et al., 2022); Under the constraints of local moderate temperature and precipitation (the dominant climatic conditions here); these measures can increase SOC_p by approximately 15 to 20% (based on regional trial data); which is the main feasible space for human intervention. In the transitional zone of moderate SOC_p; where precipitation fluctuates (a key climatic constraint); constructing ecological corridors; increasing vegetation coverage (to buffer climatic instability); and optimizing climate adaptive afforestation techniques can expand SOC_p improvement space by 10 to 12%; which serves as a practical path for human regulation under climatic constraints. In areas with low SOC_p; it is necessary to focus on monitoring carbon loss during freeze thaw processes (driven by temperature changes) and enhance regional ecological stability by constructing ecological barriers (Yuan et al., 2022); This measure can reduce climatic induced carbon loss by 8 to 10%; which is the core space for human management to compensate for low SOC_p caused by harsh climatic conditions. More importantly; future risk assessment of these measures should integrate climate change scenarios (e.g.; temperature rise; precipitation variation) to track how the effectiveness of management measures changes under dynamic climatic constraints (Price et al., 2007).

4.2. *Predictive effectiveness of environmental variables in SOC_p prediction across different ecosystems*

Based on the BRT model, a quantitative analysis was conducted to identify the dominant environmental variables influencing SOC_P across different ecosystems in Northeast China. The study revealed a synergistic effect between climatic variables (e.g., temperature and precipitation) and topographic variables (e.g., curvature and humidity index), with variations observed among different models. Research findings indicated that MAT, MAP, and NPP were key environmental variables affecting SOC_p in this region, collectively accounting for approximately 55% of the total RI. Specifically, the SOC_p in farmland ecosystems was primarily driven by MAT, precipitation (PRC), and MAP, demonstrating a synergistic dominance of climate and topography. In forest ecosystems, it is mainly regulated by the TWI and MAT, reflecting the coupling effect between TWI and climatic factors. Grassland ecosystems depend on the interactive control of MAP and vegetation productivity (NDVI/NPP). Model comparisons showed that the Hassink model is more suitable for describing the influence of water and heat conditions on carbon sequestration potential in natural ecosystems (including farmland, forests, and grasslands), whereas the Elustondo model is better suited for analyzing the synergistic optimization of topography and productivity in farmland management. This study further confirms the significant impact of climate and topography on carbon cycling through mechanisms such as microbial metabolism, runoff pathways, and vegetation cover. It provides a theoretical basis for differentiated carbon management strategies in Northeast China, such as land conservation tillage, forest gully and valley protection, and grassland rotation and replanting, and can improve the accuracy of carbon sequestration potential prediction by integrating the strengths of different models. This chapter not only verifies the significant influence of synergistic effects among various environmental factors on soil carbon sequestration potential, but also offers a scientific foundation for regional ecological protection, comprehensive management, and regulatory strategies.

At the same time, the key environmental variables predicted by the SOC_p framework constructed in this study were largely consistent with previous research findings, demonstrating a certain level of scientific validity and accuracy. MAT and MAP were key variables influencing the spatial distribution of SOC_p. Moreover, MAT and MAP significantly affected the accumulation of soil microbial residual carbon by regulating microbial metabolic efficiency (e.g., carbon use efficiency, CUE) (Giardina and Ryan, 2000; Fantappiè et al., 2011; Adhikari et al., 2019). Specifically, optimal temperatures enhance microbial carbon utilization efficiency by balancing microbial biomass synthesis and decomposition: microorganisms allocate more carbon toward biomass synthesis rather than energy expenditure, thereby increasing the production of resistant microbial residues. These materials make significant contributions to soil organic carbon sequestration. The synergistic effected between natural and anthropogenic factors had a significant influence on SOC storage. Among these, human activities such as conservation tillage and irrigation indirectly affected the SOC_{stock} in farmland ecosystems by altering terrain and crop growth conditions, thereby regulating SOC content (Tang et al., 2019). Different land use patterns alter vegetation cover types and biomass accumulation rates; and consequently influence SOC input and decomposition processes; which directly determine the SOC_p across various ecosystems (Post and Kwon, 2000). NPP and NDVI fusion data significantly influenced SOC content in different ecosystems (Chuai et al., 2022; Das et al., 2023). Areas with higher NPP and NDVI values provide abundant readily decomposable carbon sources for microbial growth, thereby sustaining diverse and active microbial communities capable of producing substantial amounts of stable residual carbon. For example, temperate forests exhibit relatively high SOC_p due to their elevated NPP and NDVI values, as well as higher microbial residue carbon density (Xu et al., 2024). TWI was also a key indicator for predicting ecosystem carbon sequestration potential; with its effected mediated through soil moisture; vegetation type; and microbial activity. High TWI areas (e.g.; wetlands and moist forests) maintain elevated soil moisture; which promotes hydrophilic vegetation growth and modulates microbial communities—favoring anaerobic microbes that produce recalcitrant residues while suppressing aerobic decomposers accelerating carbon loss. However; human interventions (e.g.; drainage) convert anaerobic environments to aerobic ones; stimulating the decomposition of accumulated microbial residues and causing significant carbon loss. Thus; while high TWI areas typically exhibit high carbon sequestration capacity; caution is needed regarding potential carbon loss from extreme humidity or human disturbances. Meanwhile; PLC and PRC significantly shape the spatial heterogeneity of organic carbon by influencing surface runoff pathways (Amirian Chakan et al., 2017; Mohseni and Salar, 2021).

In terms of model construction, this study employs a unified model framework based on the BRT model to predict SOC_p across different ecosystems in Northeast China. The core objective is to conduct a systematic assessment of regional-scale SOC_p and comprehensively compare overall trends across ecosystems, rather than analyzing micro-level mechanisms within individual ecosystems. While developing dedicated models for each ecosystem could refine the driving mechanisms within a single ecosystem, it would lead to discrepancies in assessment criteria, parameter systems, and prediction accuracy across ecosystems. This would hinder the precise calculation of regional total SOC_p and the contribution ratios of individual ecosystems during

regional-scale unified assessments. The unified model framework ensures consistent scales and benchmarks for SOC_p assessments across ecosystems through standardized variable systems and predictive logic. This study further identifies ecosystem-specific driving mechanisms by analyzing the differential relative importance of environmental variables. The high consistency between these differentiated characteristics and the carbon cycling patterns of different ecosystems validates the framework's scientific rigor, demonstrating that it does not obscure unique ecosystem-specific driving mechanisms. Leveraging the nonlinear fitting capabilities of the BRT model and the advantages of relative importance analysis, this study further enhances the framework's ability to identify differentiated mechanisms. It also mitigates model fitting variations influenced by sample size, ensuring the overall reliability of regional-scale SOC_p predictions through the synergistic information from multiple ecosystem samples.

4.3. Uncertainty analysis in current research

4.3.1. Limitations of model types and sampling coverage

The current study is limited to two empirical models—Hassink and Elustondo model —and does not include algorithms based on alternative principles, such as physically-based models (e.g., RothC, Century) (Farina et al., 2013; (Nemo Klumpp et al., 2017) or deep learning approaches (McBratney et al., 2014; Wu et al., 2024). While the fusion of the two models using averages is scientifically sound, it may still weaken model-specific information and slightly amplify random valuation errors, thereby limiting the precise characterization of SOC_p features in local regions. This restriction hampers the effectiveness of model comparisons and the assessment of system uncertainties. During the environmental variable data preprocessing stage, the raster data resampling operation did not add inherent information to the original data but may artificially amplify spatial heterogeneity at fine scales. Although this study implemented stringent measures to avoid overfitting during model construction, this operation remains prone to generating minor errors in landscape transition zones. Furthermore, due to natural constraints and financial limitations, insufficient sampling in remote areas and regions with highly heterogeneous terrain may compromise the representativeness of the training data, thereby affecting the model's generalization capability. Future work should focus on developing an integrated multi-model framework that incorporates unmanned aerial vehicle remote sensing and Internet of Things sensing technologies. This approach should involve densifying the sampling layout to improve spatial coverage and data representativeness, ultimately enhancing the model's predictive robustness across diverse geographical environments.

4.3.2. Inadequate representation of environmental variable selection and dynamic mechanisms

This study primarily focuses on static environmental variables, including terrain, climate, and vegetation, while systematically excluding key process-driven variables such as human disturbances (e.g., tillage practices and crop rotation systems), soil biological factors (e.g., microbial community structure and enzyme activity), and soil erosion dynamics. This omission limits the model's explanatory power regarding underlying mechanisms. However, this is primarily constrained by data availability, and the spatial accuracy of related variables such as human disturbance remains to be further validated. Including such data with unconfirmed precision at this stage could potentially dilute the model's analytical accuracy regarding in situ carbon sequestration potential. Therefore, this study has not yet incorporated it into the model framework. In modeling analyses of various ecosystems, this study did not establish separate models for different ecosystems but instead employed a unified modeling framework covering the entire study area for simulation. This modeling strategy may struggle to adequately accommodate and resolve the heterogeneity differences among ecosystems. It may even weaken or overlook the key driving mechanisms and core influencing factors unique to each ecosystem by overemphasizing regional commonalities, thereby introducing potential uncertainties into the research findings. Furthermore, the absence of dynamic simulations under projected climate change scenarios hinders the ability to capture long-term soil carbon sequestration potential and its responses. To address these limitations, it is recommended to integrate multi-source dynamic datasets—such as high-resolution remote sensing products, metagenomic information, and long-term field observation data—with climate scenario models (e.g., SSP-RCP frameworks). This integration would enhance the model's capacity to analyze the mechanism between carbon sequestration and environmental management.

4.3.3. Deviation between actual disturbance and theoretical potential

The soil organic carbon sequestration potential quantified in this study represents in situ sequestration potential, strictly defined as the difference between the soil organic carbon saturation threshold and current measured values. This potential solely characterizes the inherent capacity for enhanced carbon sequestration achievable under existing environmental conditions through the inherent physicochemical properties of the soil itself. The additional carbon sequestration potential stimulated by altering soil utilization patterns and vegetation cover characteristics—such as through land-use conversion—thereby exceeding the upper limit of the in situ potential threshold, is not included in the assessment scope of this study. In this context, the predicted results of the model reflect the maximum carbon sequestration potential under ideal conditions. However, the actual carbon sequestration process is significantly influenced by field management practices, extreme climatic events, and socio-economic factors, leading to potential discrepancies between theoretical estimates and real-world conditions. Therefore, future research should focus on developing an integrated “management ecology” coupling model that incorporates indicators of human activity intensity and disturbance response algorithms, and conduct a systematic assessment of the additional carbon sequestration potential induced by land-use conversion and other practices, thereby enhancing the model's applicability and predictive accuracy in real-world agricultural and ecological contexts.

5. Conclusions

This study systematically revealed the spatial pattern and driving mechanisms of SOC_p in Northeast China. The core findings are as follows. First, the SOC_p in Northeast China exhibits a distinct spatial pattern characterized by “higher in the south and lower in the north, higher on plains and lower in mountainous areas”. High-value zones are concentrated in the black soil regions of the Songnen Plain and Sanjiang Plain, while low-value zones are distributed in the northern Greater Khingan Range and high-altitude regions. Second, there are significant differences in sequestration potential exist among ecosystems. Croplands demonstrate the highest SOC_p per unit area, primarily attributed to deep humus layers and management practices like straw returning; Forest carbon sequestration processes are mainly constrained by freeze-thaw erosion and limited vegetation productivity. Third, climatic factors (MAT, MAP) and NPP are the dominant drivers of SOC_p, collectively explaining the majority of its variation; Terrain factors exert indirect influences by regulating hydrothermal conditions and vegetation growth, confirming the synergistic “Climate-Terrain-Vegetation” driving mechanism. Fourth, model comparison revealed that the Hassink model provides more robust estimates under conservative scenarios, while the Elustondo model is more sensitive to extreme values, offering guidance for model selection based on different management objectives. Overall, this research identifies key carbon sequestration hotspots and critical controlling factors in Northeast China, providing essential scientific support for developing differentiated carbon management strategies tailored to ecosystem types and regional characteristics.

CRediT authorship contribution statement

Zicheng Wang: Writing – original draft, Data curation. **Qianlai Zhuang:** Writing – review & editing. **Zijiao Yang:** Investigation, Data curation. **Xinxin Jin:** Validation, Resources. **Shuai Wang:** Writing – review & editing. **Yang Wang:** Writing – review & editing. **Xinyu Zhang:** Investigation. **Di Shi:** Software, Resources. **Shouyuan Bian:** Methodology, Investigation.

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.

Acknowledgment

This work was funded by the National Key Research and Development Program of China (Grant No. 2023YFD1501300), the National Science Foundation of China (Grant No. 42207289), the Natural Science Foundation of Jilin Province (Grant No. YDZJ202401480ZYTS), and the Project of Liaoning Provincial Department of Science and Technology (Grant No. 2025-MSLH-612).

References

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.
Amelung, W., Bossio, D., de Vries, W., Kögel-Knabner, I., Lehmann, J., Amundson, R., Chabbi, A., 2020. Towards a global-scale soil climate mitigation strategy. Nat. Commun. 11 (1), 5427.
Amirian Chakan, A., Taghizadeh-Mehrjardi, R., Kerry, R., Kumar, S., Khordehbin, S., Yusefi Khanghah, S., 2017. Spatial 3D distribution of soil organic carbon under different land use types. Environ. Monit. Assess. 189 (3), 131.
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.
Benítez, P.C., McCallum, I., Obersteiner, M., Yamagata, Y., 2007. Global potential for carbon sequestration: geographical distribution, country risk and policy implications. Ecol. Econ. 60 (3), 572–583.
Campo, J., Merino, A., 2016. Variations in soil carbon sequestration and their determinants along a precipitation gradient in seasonally dry tropical forest ecosystems. Glob. Chang. Biol. 22 (5), 1942–1956.
Chen, S., Arrouays, D., Angers, D.A., Chenu, C., Barré, P., Martin, M.P., Walter, C., 2019. National estimation of soil organic carbon storage potential for arable soils: A data-driven approach coupled with carbon-landscape zones. Sci. Total Environ. 666, 355–367.
Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination r-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. Peerj Comput. Sci. 7, e623.
Chuai, X., Xia, M., Xiang, A., Miao, L., Zhao, R., Zuo, T., 2022. Vegetation coverage and carbon sequestration changes in China's forest projects area. Global Ecol. Conserv. 38, e02257.
Churchman, G.J., Singh, M., Schapel, A., Sarkar, B., Bolan, N., 2020. Clay minerals as the key to the sequestration of carbon in soils. Clay Clay Miner. 68 (2), 135–143.
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.
Das, M., Mandal, A., Das, A., Inácio, M., Pereira, P., 2023. Mapping and assessment of carbon sequestration potential and its drivers in the eastern himalayan region (India). Case Stud. Chem. Environ. Engineer. 7, 100344.
Deng, Q., Yu, Y., Wang, H., 2020. Effects of terrain curvature on soil organic carbon distribution in sloping farmland. Soil Tillage Res. 198, 104521.
Don, A., Schumacher, J., Freibauer, A., 2011. Impact of tropical land-use change on soil organic carbon stocks–a meta-analysis. Glob. Chang. Biol. 17 (4), 1658–1670.
Dunn, J.C., 1973. A Fuzzy Relative of the ISODATA Process and its Use in Detecting Compact Well-Separated Clusters.
Elustondo, J., Laverdière, M.R., Angers, D.A., N'dayegamiye, A., 1990. Etude comparative de l'agregation et de la matiere organique associee aux fractions granulometriques de sept sols sous culture de mais ou en prairie. Can. J. Soil Sci. 70 (3), 395–402.
Falloon, P., Smith, P., Szabo, J., Pasztor, L., 2002. Comparison of approaches for estimating carbon sequestration at the regional scale. Soil Use Manag. 18 (3), 164–174.
Fang, H., Zhang, Y., Zhao, Y., 2016. Effects of slope aspect on soil organic carbon and nitrogen stocks in an alpine region. Catena 149, 589–597.
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.
Farina, R., Coleman, K., Whitmore, A.P., 2013. Modification of the RothC model for simulations of soil organic c dynamics in dryland regions. Geoderma 200, 18–30.
Follett, R.F., 2001. Soil management concepts and carbon sequestration in cropland soils. Soil Tillage Res. 61 (1–2), 77–92.
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.
Hassink, J., 1997. The capacity of soils to preserve organic c and n by their association with clay and silt particles. Plant Soil 191 (1), 77–87.
Hodson, T.O., 2022. Root mean square error (RMSE) or mean absolute error (MAE): when to use them or not. Geosci. Model Dev. Discuss. 2022, 1–10.
Huang, J., Minasny, B., McBratney, A.B., Padarian, J., Triantafilis, J., 2018. The location- and scale-specific correlation between temperature and soil carbon sequestration across the globe. Sci. Total Environ. 615, 540–548.
IUSS Working Group WRB, 2006. World reference base for soil resources 2006. In: World Soil Resources Reports No. 103. FAO, Rome.
Jacquier, E., Kane, A., Marcus, A.J., 2003. Geometric or arithmetic mean: A reconsideration. Financ. Anal. J. 59 (6), 46–53.
Jiang, L., Yang, G., Liu, Y., Wu, Z., 2020. Slope aspect effects on soil organic carbon and its influencing factors in a mountainous area. J. Soils Sediments 20 (5), 2249–2259.
Jiang, P., Xiao, L.Q., Wan, X., Yu, T., Liu, Y.F., Liu, M.X., 2022. Research progress on microbial carbon sequestration in soil: A review. Eur. Soil Sci. 55 (10), 1395–1404.
Kheirfam, H., 2020. Increasing soil potential for carbon sequestration using microbes from biological soil crusts. J. Arid Environ. 172, 104022.
Kirkby, M.J., 1999. Catchment (1000 km^2) 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., 2003. Global potential of soil carbon sequestration to mitigate the greenhouse effect. Crit. Rev. Plant Sci. 22 (2), 151–184.
Lal, R., 2004. Soil carbon sequestration impacts on global climate change and food security. science 304 (5677), 1623–1627.
Lal, R., Negassa, W., Lorenz, K., 2015. Carbon sequestration in soil. Curr. Opin. Environ. Sustain. 15, 79–86.
Lessmann, M., Ros, G.H., Young, M.D., de Vries, W., 2022. Global variation in soil carbon sequestration potential through improved cropland management. Glob. Chang. Biol. 28 (3), 1162–1177.
Li, X., Zhang, K., Liu, W., 2014. Impact of topographic curvature on soil organic carbon in a hilly region: A case study from the loess plateau, China. Catena 123, 100–107.
Li, D., Chu, Q., Wang, J., Qian, C., Chen, C., Feng, Y., Xue, L., 2023. Effect of hydrothermal carbonization aqueous phase on soil dissolved organic matter and microbial community during rice production: A two-year experiment. Agric. Ecosyst. Environ. 356, 108637.
Liang, C., Schimel, J.P., Jastrow, J.D., 2017. The importance of anabolism in microbial control over soil carbon storage. Nat. Microbiol. 2 (8), 1–6.
Luo, B., Zhou, J., Yao, W., Wang, Y., Guillaume, T., Yuan, M., Zang, H., 2024. Maize and soybean rotation benefits soil quality and organic carbon stock. J. Environ. Manag. 372, 123352.
Matzner, E., Borken, W., 2008. Do freeze-thaw events enhance c and n losses from soils of different ecosystems? A review. Eur. J. Soil Sci. 59 (2), 274–284.
McBratney, A.B., Stockmann, U., Angers, D.A., Minasny, B., Field, D.J., 2014. Challenges for soil organic carbon research. In: Soil carbon. Springer International Publishing, Cham, pp. 3–16.
Mohseni, N., Salar, Y.S., 2021. Terrain indices control the quality of soil total carbon stock within water erosion-prone environments. Ecohydrol. Hydrobiol. 21 (1), 46–54.
Monson, R.K., Turnipseed, A.A., Sparks, J.P., Harley, P.C., Scott-Denton, L.E., Sparks, K., Huxman, T.E., 2002. Carbon sequestration in a high-elevation, subalpine forest. Glob. Chang. Biol. 8 (5), 459–478.
Nair, P.R., 2011. Methodological challenges in estimating carbon sequestration potential of agroforestry systems. In: Carbon Sequestration Potential of Agroforestry Systems: Opportunities and Challenges. Springer Netherlands, Dordrecht, pp. 3–16.
Nemo Klumpp, K., Coleman, K., Dondini, M., Goulding, K., Hastings, A., Smith, P., 2017. Soil organic carbon (SOC) equilibrium and model initialisation methods: an application to the rothamsted carbon (RothC) model. Environ. Model. Assess. 22 (3), 215–229.
Pechanec, V., Purkyt, J., Benc, A., Nwaogu, C., Štěrbová, L., Cudlín, P., 2018. Modelling of the carbon sequestration and its prediction under climate change. Eco. Inform. 47, 50–54.
Portner, H., Bugmann, H., Wolf, A., 2010. Temperature response functions introduce high uncertainty in modelled carbon stocks in cold temperature regimes. Biogeosciences 7 (11), 3669–3684.
Post, W.M., Kwon, K.C., 2000. Soil carbon sequestration and land-use change: processes and potential. Glob. Chang. Biol. 6 (3), 317–327.
Price, P.N., McKone, T.E., Sohn, M.D., 2007. Carbon Sequestration Risks and Risk Management.
Rukhin, A.L., Vangel, M.G., 1998. Estimation of a common mean and weighted means statistics. J. Am. Stat. Assoc. 93 (441), 303–308.
Smith, P., 2005. An overview of the permanence of soil organic carbon stocks: influence of direct human-induced, indirect and natural effects. Eur. J. Soil Sci. 56 (5), 673–680.

Smith, P., 2008. Land use change and soil organic carbon dynamics. Nutr. Cycl. Agroecosyst. 81 (2), 169–178.
Sokol, N.W., Whalen, E.D., Jilling, A., Kallenbach, C., Pett-Ridge, J., Georgiou, K., 2022. Global distribution, formation and fate of mineral-associated soil organic matter under a changing climate: A trait-based perspective. Funct. Ecol. 36 (6), 1411–1429.
Song, Y., Zou, Y., Wang, G., Yu, X., 2017. Altered soil carbon and nitrogen cycles due to the freeze-thaw effect: A meta-analysis. Soil Biol. Biochem. 109, 35–49.
Steger, K., Fiener, P., Marvin-DiPasquale, M., Viers, J.H., Smart, D.R., 2019. Human-induced and natural carbon storage in floodplains of the central valley of California. Sci. Total Environ. 651, 851–858.
Stockmann, U., Adams, M.A., Crawford, J.W., Field, D.J., Henakaarchchi, N., Jenkins, M., Zimmermann, M., 2013. The knowns, known unknowns and unknowns of sequestration of soil organic carbon. Agric. Ecosyst. Environ. 164, 80–99.
Tang, H., Liu, Y., Li, X., Muhammad, A., Huang, G., 2019. Carbon sequestration of cropland and paddy soils in China: potential, driving factors, and mechanisms. Greenhouse Gases 9 (5), 872–885.
Tao, F., Huang, Y., Hungate, B.A., Manzoni, S., Frey, S.D., Schmidt, M.W., Luo, Y., 2023. Microbial carbon use efficiency promotes global soil carbon storage. Nature 618 (7967), 981–985.
Tesfaye, M.A., Bravo, F., Ruiz-Peinado, R., Pando, V., Bravo-Oviedo, A., 2016. Impact of changes in land use, species and elevation on soil organic carbon and total nitrogen in ethiopian central highlands. Geoderma 261, 70–79.
Willmott, C.J., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim. Res. 30 (1), 79–82.
Wu, G., Huang, G., Lin, S., Huang, Z., Su, Y., 2024. Changes in soil organic carbon stocks and its physical fractions along an elevation in a subtropical mountain forest. J. Environ. Manag. 351, 119823.
Xu, S., Song, X., Zeng, H., Wang, J., 2024. Soil microbial necromass carbon in forests: A global synthesis of patterns and controlling factors. Soil Ecol. Lett. 6 (4), 240237.
Yang, R.M., Huang, L.M., Zhang, X., Zhu, C.M., Xu, L., 2023. Mapping the distribution, trends, and drivers of soil organic carbon in China from 1982 to 2019. Geoderma 429, 116232.
Yang, Y., Gunina, A., Cheng, H., Liu, L., Wang, B., Dou, Y., Chang, S.X., 2025. Unlocking mechanisms for soil organic matter accumulation: carbon use efficiency and microbial necromass as the keys. Glob. Chang. Biol. 31 (1), e70033.
You, M., Han, X., Hu, N., Du, S., Doane, T.A., Li, L.J., 2020. Profile storage and vertical distribution (0–150 cm) of soil inorganic carbon in croplands in Northeast China. Catena 185, 104302.
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.
Yuan, H.H., Wang, Z., Xu, W.G., You, G.Y., Zhang, J.L., 2022. Vegetation dynamics and influence factors in forest-steppe transition ecozone: the case of daxing'an mountains, Northeast China. Acta Ecol. Sin. 42, 7321–7335.
Zhang, F., Liu, Y., Wu, S., Liu, J., Luo, Y., Ma, Y., Pan, X., 2024. Prediction and spatial–temporal changes of soil organic matter in the huanghuaihai plain by combining legacy and recent data. Geoderma 450, 117031.
Zheng, K., Wei, J.Z., Pei, J.Y., Cheng, H., Zhang, X.L., Huang, F.Q., Ye, J.S., 2019. Impacts of climate change and human activities on grassland vegetation variation in the chinese loess plateau. Sci. Total Environ. 660, 236–244.
Zhou, C., Zhou, Q., Wang, S., 2003. Estimating and analyzing the spatial distribution of soil organic carbon in China. AMBIO 32 (1), 6–12.
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: Digital Soil Mapping with Limited Data. Springer Netherlands, Dordrecht, pp. 233–245.