forests

MDPI

Article

Topography-Mediated Nonlinear Responses of Soil Organic Carbon Stocks to Multi-Gradient Warming and Precipitation Shifts in Northeast China's Temperate Mountain Forests

Zicheng Wang [1], Qianlai Zhuang [2], Shuai Wang [1,*], Zijiao Yang [1], Fujun Sun [1], Yang Wang [3], Yan Sang [4], Lingyue Wang [4] and Xinxin Jin [1]

[1] College of Land and Environment, Shenyang Agricultural University, Shenyang 110866, China
[2] Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN 47907, USA; qzhuang@purdue.edu
[3] College of Resources and Environment, Jilin Agricultural University, Changchun 130118, China
[4] Lingyuan Natural Resources Affairs Service Center, Chaoyang 122500, China
* Correspondence: shuaiwang666@syau.edu.cn; Tel.: +86-24-8848-7155

Abstract

Soil organic carbon (SOC) in mountain forest ecosystems exerts critical controls over regional carbon balance and climate feedback loops. This study collected 209 stratified topsoil (0–30 cm) samples across temperate mountain forests of Northeast China, conducted a boosted regression tree (BRT) modeling framework integrated with space-for-time substitution, and established 15 combined thermopluviometric sensitivity scenarios to simulate SOC shifts under diversified climate disturbances. Tenfold cross-validation yielded a model mean R^2 of 0.62, revealing mean annual temperature (MAT, RI = 35.31%) as the most influential predictor of SOC spatial variation, followed by elevation (ELE, RI = 19.11%), while single-season NDVI and soil particle fractions showed weak predictive capacity. Multi-scenario spatial simulation outputs demonstrated that simultaneous warming and aridification drastically reduce the coverage of high SOC zones, whereas increased precipitation can partially offset temperature-induced carbon mineralization losses. Terrain-mediated SOC spatial stratification remained stable across all climate backgrounds. This study quantifies the layered environmental association hierarchy of mountain SOC and generates spatially explicit modeled carbon sink projections under climate change. The terrain-dependent SOC response patterns provide operable differentiated carbon regulation guidance: humid low-lying convergence zones require long-term soil moisture conservation, while arid steep ridges need native mixed forest restoration to lift baseline carbon storage capacity, supporting precise watershed climate adaptation and targeted forest carbon sink management for temperate mountain regions.

Keywords: soil organic carbon stocks; boosted regression trees; space-for-time substitution; climate scenario simulation; mountain temperate forest; environmental variables

Academic Editor: Milton Cesar Costa Campos

Received: 23 July 2026
Revised: 21 August 2026
Accepted: 25 August 2026
Published: 27 August 2026

1. Introduction

Mountain forest soils store substantial organic carbon and act as crucial buffers to mitigate global warming [1,2]. The mountainous terrain creates strong microclimate and soil redistribution gradients, resulting in significant spatial differentiation of soil organic carbon (SOC) stocks, which is distinct from that in flat plain ecosystems [3,4]. Identifying the dominant environmental variables influencing mountain SOC and their trends under changing temperature and precipitation is essential for improving regional carbon budget assessments and implementing measures to enhance carbon sinks.

The existing literature has extensively examined the most influential predictors of SOC and its spatial distribution through machine learning algorithms [5,6]. As one mainstream ensemble modeling tool, boosted regression trees (BRTs) can quantify variable relative importance and identify nonlinear interactions without rigid constraints on data distribution [7,8]. The feasibility of this modeling approach has been validated across forest, cropland, and alpine soil landscapes in previous studies [9,10]. Two analytical frameworks dominate existing research exploring forest SOC controls: single Pearson correlation analysis and machine learning-based relative importance evaluation [11,12]. Correlation heatmaps are limited to capturing linear pairwise relationships and cannot disentangle hidden nonlinear covariations between multiple covariates. Repeated BRT fitting reduces stochastic model noise and provides more reliable hierarchical factor sorting [13–15].

Climate warming and precipitation redistribution jointly reshape the turnover of forest SOC by altering vegetation carbon inputs and microbial mineralization rates [16]. Long-term in situ soil monitoring datasets are scarce across most mountainous regions; under such data constraints, the space-for-time substitution framework offers a feasible alternative to simulate future SOC dynamics [17]. It projects static connections by applying the fitted carbon–environment relationships to hypothetical future thermopluviometric scenarios. It should be noted that such extrapolation cannot fully capture multi-decadal time lags, soil legacy effects and vegetation succession processes that govern real long-term SOC temporal dynamics [18]. The rationale for adopting space-for-time substitution in this research relies on two core preconditions supported by regional background and existing literature [18,19]. First, the temperate natural mixed forests across our study area have experienced no large-scale logging, wildfire or artificial afforestation over the past 40 years, maintaining stable stand structure, soil parent material and steady spatial hydrothermal gradients that can reflect decadal climate change trajectories. Second, this method has been widely validated in mid-latitude mountain forest SOC simulations when long-term continuous monitoring data are unavailable. Nevertheless, this snapshot cross-sectional design has inherent defects: static spatial correlations cannot capture centennial soil pedogenesis, multi-stage vegetation succession and lagged microbial mineralization, so all scenario outputs in this paper are only sensitivity reference patterns rather than definitive long-term SOC predictions, which will be further critically discussed in Section 4.3.

Despite the wide application of this simulation technique, three prominent research limitations remain in existing SOC climate projection studies. First, most multi-scenario SOC modeling efforts focus on agricultural croplands and tropical rainforest ecosystems, while temperate mountain forest catchments remain severely understudied, with limited spatially explicit carbon projections published for mid-latitude mountain terrain [20,21]. Second, prior carbon–climate modeling work generally treats topography as a trivial background covariate and rarely quantifies its mediating buffering or amplifying effects on SOC's climate responsiveness, ignoring terrain-induced microclimate and erosion heterogeneity across hillslopes [22,23]. Third, few machine learning-based SOC studies systematically partition environmental predictors into climate, topographic, vegetation and soil

https://doi.org/10.3390/f17091026

particle groups to separately compare their independent predictive weights; mixed variable sets obscure the layered hierarchical controls of SOC spatial patterns [24,25].

Topography acts as a fundamental indirect regulator of mountain SOC pools by reshaping local thermopluviometric regimes, surface runoff and hillslope soil redistribution processes [22,26]. Altitudinal gradients simultaneously alter temperature and precipitation, modifying plant productivity and microbial decomposition rates and driving vertical SOC differentiation, with most mountain ecosystems exhibiting rising SOC stocks at higher elevations [21,27]. Slope gradient and solar aspect further differentiate microhabitats: steep terrain aggravates soil erosion while south-facing slopes suffer intense evaporative water loss, both limiting organic carbon accumulation [18]. Hydrological terrain indices, including Topographic Wetness Index (TWI) and Catchment Area (CA), quantify terrain-driven moisture convergence and hillslope runoff generation; runoff derived from CA drives lateral organic sediment redistribution, and together these two indices jointly regulate long-term carbon retention on slopes and valley depressions [10,28]. Existing modeling work rarely disentangles the independent explanatory power of multiple terrain indicators under multi-gradient climate change, which constitutes a key research gap addressed in this study.

Twelve environmental predictors were selected and grouped into climate, topographic, vegetation and soil-particle categories following three pedological–ecological screening criteria: (1) selected variables are widely recognized core SOC regulators from global mountain soil syntheses, with well-defined mechanisms governing plant carbon input or soil organic matter decomposition [22,29]; (2) all raster predictors adopt a consistent 90 m spatial resolution matching field sampling, avoiding scale-mismatch bias; (3) variables with redundant pedological functions were removed following preliminary collinearity diagnostics. Specifically, mean annual temperature (MAT) and mean annual precipitation (MAP) control microbial decomposition and plant productivity; six topographic variables represent hillslope erosion and microclimate redistribution; single-season NDVI serves as a proxy for annual litter carbon inputs; sand, silt and clay content determine mineral-associated SOC stabilization potential.

Two core hypotheses guide this research: (1) Mean annual temperature will exhibit the strongest statistical association with mountain SOC stocks among all measured environmental predictors, with terrain acting as a secondary mediating factor; (2) terrain will create persistent spatial stratification of SOC pools across climate scenarios, such that moist lowland sites show far larger modeled SOC shifts than arid steep ridges under warming and precipitation perturbations. The temperate mountain forest zone in Northeast China is characterized by distinct gradients in elevation, temperature, and precipitation. It has intact natural forest cover and a wealth of field sampling records, which makes it an ideal platform for investigating the spatial patterns of mountain SOC and its responses to climate change. We collected 209 representative topsoil sampling points and selected 12 environmental variables that cover four functional groups. This study aims to achieve three core objectives: (1) Characterize the descriptive distribution and pairwise linear correlations of SOC and multi-environmental variables through boxplots and correlation heat maps; (2) quantify the hierarchical relative importance of climate, topographic, vegetation, and soil predictors by using repeatedly optimized BRT models; and (3) simulate the spatial redistribution of SOC under 15 combined climate scenarios and explore how terrain correlates with divergent modeled SOC sensitivity across warming and precipitation gradients. The results are expected to enhance the theoretical understanding of carbon cycling in mountain forests and provide spatially explicit technical support for climate-smart forest carbon management.

2. Materials and Methods

2.1. Study Area

The Northeast region of China is endowed with the most extensive natural forest resources in the country. Its core distribution areas are centered in the Greater Khingan Range, Lesser Khingan Range, and Changbai Mountain Range, spanning the provinces of Liaoning, Jilin, and Heilongjiang (Figure 1). The terrain of this forest area is marked by undulations, with an altitude ranging from 0 to 2665 m and an average elevation of roughly 200 m. The dominant vegetation is a mixed forest of temperate coniferous and broad-leaved deciduous trees, which exhibits transitional characteristics between the cold and temperate zones. According to the China Statistical Yearbook (National Bureau of Statistics, 2015), the regional forest area is approximately 3.05×10^5 km^2, accounting for 15% of the nation's total forest area. The timber storage volume reaches up to 3.2 billion cubic meters, about one-third of the national total. When looking at each province, Heilongjiang Province has a storage volume of 2.4 billion cubic meters, accounting for around a quarter; Jilin Province has a storage volume of about 700 million m^3, and Liaoning Province has a storage volume of about 100 million m^3, presenting a spatial gradient that is high in the north and low in the south. From a geological and pedological perspective, the mountainous study area is dominated by granite, metamorphic rock and intermediate-basic magmatic rock bedrock. Residual weathering materials and slope colluvium form the primary soil parent materials on hillslopes; alluvial and proluvial deposits accumulate in valley bottoms and convergent lowlands, while thin loess mantles cover partial gentle upland zones. The divergent distribution of these parent materials creates baseline differences in soil particle composition across elevation gradients.

In terms of climate, this area features a typical temperate monsoon climate with distinct seasons: warm and humid summers, and dry and cold winters. Under the influence of latitude and terrain, the mean annual precipitation (MAP) exhibits a significant decreasing trend from the southeast coast to the northwest inland, declining from approximately 1000 mm to less than 300 mm. This reveals a continuous water gradient that shifts from humid to semi-humid and then to semi-arid conditions. The mean annual temperature (MAT) ranges from −4 °C to 11 °C, with a substantial temperature difference between the north and the south, indicating large-scale climate differentiation. This temperate monsoon climate is marked by distinct seasonal variations, namely warm, rainy summers and cold, dry winters. The predominant soil types are Cambisols (covering 82% of the area), followed by Gleysols (7%) and Phaeozems (6%). There are also minor distributions of Luvisols, Leptisols, and other soil classes, as classified by the World Reference Base for Soil Resources. Overall, the Northeast forest region is not only an important ecological barrier in China but also a typical area for studying the interaction among climate, vegetation, and soil.

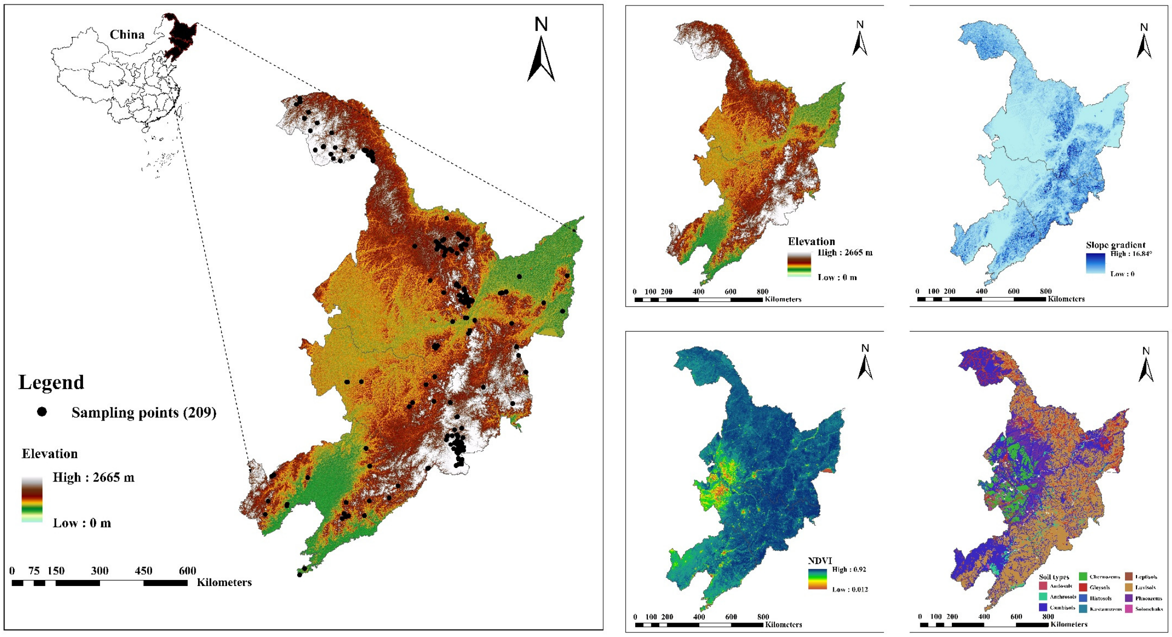

Figure 1. Spatial distribution of the 209 sampling points across Northeast China, overlaid on a 90-m digital elevation model.

2.2. Soil Sampling Collection

Owing to rugged mountain terrain, high field survey costs and extensive continuous forest coverage across the study domain, large-scale uniform grid sampling was impractical. To reliably capture the spatial variation of topsoil SOC stocks within complex mountain landscapes, a purposive stratified sampling framework was adopted following the sampling protocol proposed by Zhu et al. [30]. Four core environmental variables were integrated to delineate sampling strata: elevation (ELE), slope gradient (SG), Normalized Difference Vegetation Index (NDVI) and soil subgroups. Specifically, seven soil subgroups from the Second National Soil Survey, together with ELE, SG and NDVI raster layers, were fed into a fuzzy C-means clustering algorithm, partitioning the entire study area into 32 independent soil-vegetation landscape strata (i.e., sampling strata). Within each stratum, local soil specialists collected 6–8 composite soil samples covering diverse micro-landform positions. This stratified purposive sampling design can yield limited but representative soil observations for regional carbon estimation [22]. All field sampling campaigns were conducted during the summer growing season (July–August 2020). A minimum horizontal separation distance of 1.2 km was maintained between any two sampling points to reduce spatial autocorrelation bias. At each quadrat, only mineral topsoil (0–30 cm) was collected, and surface litter layers were fully removed before soil collection. In total, 209 composite topsoil samples were obtained across all strata. At each designated sampling site, a 20 m × 20 m standard quadrat was set up; five sub-sampling points (four quadrat corners plus the central position) were arranged, and undisturbed soil within the 0–30 cm layer was extracted using a stainless steel soil auger, then fully homogenized to form one composite sample per quadrat. Visible plant residues, stone fragments and animal excrement were manually removed on-site before sample packaging. All soil specimens were sealed in clean polyethylene bags, labeled with unified geocoordinate, elevation, slope, vegetation type and sampling date information, and transported back to the laboratory under low-temperature shading conditions. After air-drying indoors at a constant temperature, each sample was sequentially sieved through 2 mm and 0.25 mm mesh screens for physical property measurement and SOC concentration testing respectively. SOC content was determined via the dry combustion method using a Vario EL III elemental analyzer (Elementar Analysensysteme GmbH, Hanau, Germany). Compared with the traditional potassium dichromate oxidation method, this approach reduces interferences related to inorganic carbonate and provides higher measurement precision [31]. A minimum horizontal separation distance of 1.2 km was enforced between any two composite samples within each landscape stratum to reduce spatial clustering and autocorrelation between adjacent sampling points. Remote high-altitude steep gullies were inaccessible during field surveys, leading to sparse sampling points in these micro-landforms and higher prediction standard deviation in the corresponding zones. However, the fuzzy C-means stratified sampling design covered the full spectrum of elevation, vegetation and soil subgroup types across the entire study area via accessible plots. The overall regional SOC spatial trend remains representative, and only localized extreme rugged fragments show elevated uncertainty without skewing the whole watershed's carbon pattern. This enforced minimum separation within each landscape stratum reduces spatial clustering and mitigates sample-point spatial dependence at the field-sampling stage.

2.3. Calculation of Soil Organic Carbon Stocks

SOC stocks (kg C m^{-2}) were calculated using the standard pedological formula:

$$SOC \quad stocks = SOC_{conc} \times BD \times D\left(1 - CF\right) \tag{1}$$

where SOC_{conc} is the measured soil organic carbon concentration (g kg^{-1}); *BD* is the mean soil bulk density (g cm^{-3}) measured for each composite soil sample; *D* is the sampling depth (0.3 m); and *CF* is the volumetric fraction of coarse fragments (>2 mm) in each soil sample. All bulk density and coarse fragment correction measurements were collected synchronously with the soil organic carbon concentration analysis for each quadrat-level composite soil sample.

2.4. Environment Variable

2.4.1. Topographic Variables

The terrain dataset used in this study is a 90 m resolution digital elevation model (DEM) obtained from the Computer Network Information Center of the Chinese Academy of Sciences (http://www.gscloud.cn (accessed on 20 August 2022)). Six topographic variables were extracted from the DEM: elevation (ELE), slope gradient (SG), slope aspect (SA), topographic wetness index (TWI), profile curvature (PC), and catchment area (CA). ELE, SG, SA, and PC were calculated via the spatial analysis toolbox in ArcGIS; TWI and CA were derived using hydrological modules within Saga GIS for subsequent BRT modeling.

2.4.2. Climatic Variables

MAT and MAP jointly regulate plant carbon input and microbial mineralization to drive SOC turnover [18,29]. The climate data used in this research was obtained from the geospatial data cloud platform (Computer Network Information Center of the Chinese Academy of Sciences; http://www.gscloud.cn). The dataset is based on daily observations from over 2400 meteorological stations across the country, and it was generated through spatial interpolation using the ANUSPLIN thin-plate smooth spline model with an original resolution of 1 km. To match the scale of other environmental covariates, we resampled it to a 90 m grid using ArcGIS 10.2.

https://doi.org/10.3390/f17091026

To explore the potential impact of future climate fluctuations on SOC stocks, this study developed 15 climate sensitivity scenarios that comprehensively incorporate combinations of temperature and precipitation factors. The temperature was set to rise by three gradients of 1.5 °C, 2.0 °C, and 4.0 °C respectively, relative to the baseline. Precipitation was set at five levels of change relative to the baseline: −50%, −25%, 0%, +25%, and +50%. These two variables fully intersect, resulting in a total of 15 scenario combinations (3 × 5), which cover a diverse range of climate trajectories from significant aridification to strong humidification and from mild warming to extreme warming. Changes in precipitation also reflect typical hydrological and ecological states: −50% represents extreme aridification, which may trigger widespread forest degradation and declines in stand productivity [32]; −25% is equivalent to severe and persistent drought, which can easily lead to ecosystem degradation and a sharp decline in semi-humid area runoff [33]; the baseline 0% condition still contains natural annual precipitation variability [34]; +25% corresponds to moderate humidification favorable for temperate forest net primary productivity growth [35]; +50% denotes extreme humid conditions that may trigger recurring floods and require adaptive forest vegetation management [33]. By coupling dry and wet gradients with warming tiers, this scenario framework quantifies nonlinear SOC responses to compound climate stress, providing scientific support for climate adaptation management and spatially explicit zoning of soil carbon sequestration vulnerability under future climate perturbation.

This set of gradient scenarios is not directly extracted from CMIP model single outputs, but designed as a sensitivity test framework covering full wet–dry and warming spectrums for mountain forest SOC. The ±50% precipitation extremes represent potential long-term aridification/humidification extremes referenced to historical precipitation fluctuation records of Northeast mountain watersheds over the past years. The five precipitation gradients (−50%, −25%, 0%, +25%, +50%) fully cover mild, moderate and extreme hydrological disturbance levels, while the three warming thresholds align with Paris Agreement climate targets and high-end warming projections from global ensemble climate simulations. All scenarios are used to quantify SOC sensitivity gradients rather than predicting specific future climate states. The three warming thresholds strictly follow two mainstream climate research systems: the 1.5 °C control target proposed by the Paris Agreement and +2 °C/+4 °C high-end warming projections from CMIP6 global climate ensemble outputs. Five precipitation fluctuation levels (−50%, −25%, 0%, +25%, +50%) are extracted from 60-year historical meteorological records of Northeast mountain watersheds; the actual annual precipitation anomaly range of local extreme drought/flood years reaches −47%~+53%, so the gradient set fully covers regional historical hydrological extremes. This scenario framework is designed for sensitivity analysis rather than predicting fixed future climate states.

2.4.3. Biological Variables

The NDVI was selected as the core biological predictor reflecting aboveground litter carbon input to topsoil. Cloud-free summer (mid-July to early August) Landsat 8 imagery acquired in 2020 was downloaded from the Geospatial Data Cloud platform of the Chinese Academy of Sciences, with a spatial resolution of 30 m consistent with field sampling phenology. Radiometric calibration, atmospheric correction and terrain shading correction were executed sequentially in ENVI 5.3 to eliminate noise from cloud shadows and rugged mountain terrain that may underestimate vegetation coverage in steep zones [36]. Note that the NDVI dataset used here originates only from single-season peak-growing-season imagery. Single-summer NDVI cannot capture interannual vegetation phenological shifts or underground fine-root biomass dynamics, which substantially contribute to its low relative importance (3.10%) in our BRT outputs [10]. Integrating multi-temporal spring–summer–autumn NDVI composites or multi-year averaged vegetation metrics in future modeling would likely increase the explanatory power of biological predictors for SOC spatial variation, which represents an important direction for follow-up work.

2.4.4. Soil Property Variable

Soil particle size fraction data (sand, silt, clay mass percentages) were obtained from the China Resources and Environmental Science Data Center (https://www.resdc.cn/), with a spatial resolution of 1 km. This dataset was built from the national 1:1,000,000 soil map and field profile records from the Second National Soil Survey. Sand, silt and clay represent three distinct fine-earth particle size classes (<2 mm), and their proportional contents were extracted as soil texture predictors for modeling. The dataset provides mass percentage values for three fundamental fine-earth particle size fractions (sand, silt, clay). Soil texture itself is defined by the relative proportional combination of these three particle fractions, rather than treating them as independent texture categories. This approach enables quantitative analysis of spatial variations in soil particle composition across the entire study watershed.

2.5. Model Methods

This study integrates multi-source input data, including field sampling points, laboratory SOC measurements, climate grids, DEM-derived terrain indices, NDVI imagery and soil texture layers. After preprocessing steps such as format conversion, spatial alignment, 90 m resampling and collinearity reduction, boosted regression tree (BRT) modeling is performed with cross-validation, variable selection and repeated runs. Fifteen combined scenarios covering baseline climate, temperature rises of +1.5 °C, +2 °C and +4 °C, and precipitation changes from −50% to +50% are simulated. Using the space-for-time substitution approach to estimate SOC sensitivity to climate change, this work produces baseline SOC-stock maps, uncertainty maps, variable importance metrics, climate sensitivity outputs and topographic mediation results, providing guidance for regional carbon sequestration management. The specific technical roadmap is shown in Figure 2.

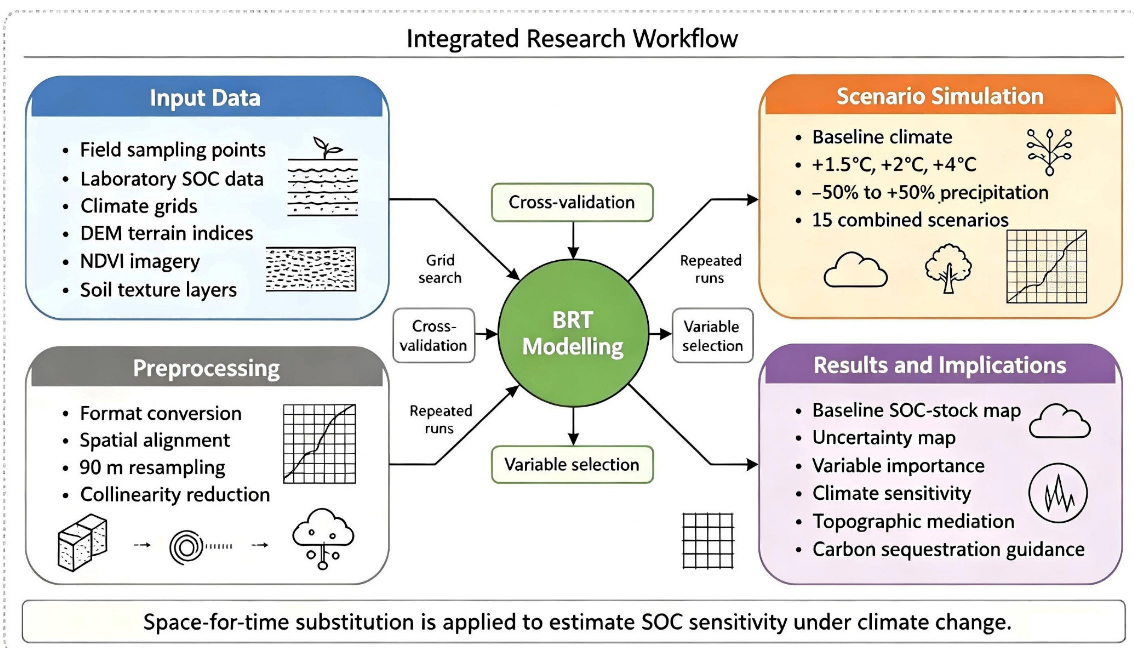

Figure 2. Technical roadmap framework diagram.

2.5.1. Boosted Regression Trees

Boosted regression trees (BRTs) integrate CART decision trees and gradient boosting algorithms to capture nonlinear variable interactions without strict parametric assumptions [9,10,37]. This ensemble approach tolerates mixed categorical/continuous predictors and avoids manual cross-term setting, making it suitable for complex SOC-environment association analysis in mountain forests. In this study, the BRT model was conducted in R 4.1.0 using the "gbm" and "dismo" packages. To optimize predictive performance, we conducted a grid search for four hyperparameters: learning rate (LR), tree complexity (TC), bag fraction (BF), and number of trees (NT). To achieve optimal prediction performance for SOC stocks in the study area, we devised a grid-search strategy. The grid search covered learning rate values spanning 0.0005, 0.001, 0.0025, 0.005, 0.01, 0.05 and 0.1 to fully capture low-magnitude LR performance; using cross-validation error as the selection criterion, other hyperparameter ranges remained as follows: TC 6–12, BF 0.55–0.75, and NT 1000–4000. The final optimal parameter configuration was LR = 0.001, TC = 12, BF = 0.7, NT = 2100, which struck a good balance between the deviation interpretation rate and computational efficiency, ensuring the reliability and robustness of subsequent spatial predictions. The final optimal hyperparameter combination (LR = 0.001, TC = 12, BF = 0.7, NT = 2100) was determined by minimizing 10-fold cross-validation RMSE while balancing computational load. A learning rate above 0.001 causes severe underfitting and insufficient residual fitting; an LR lower than 0.0005 exponentially increases computation time without accuracy improvement. Tree complexity below 6 cannot capture complex terrain-climate nonlinear interactions, while TC > 14 triggers overfitting to local sampling noise. A bag fraction of 0.7 ensures representative random training subsets, and 2100 decision trees reach stable model convergence without redundant iteration.

2.5.2. Space-for-Time Substitution Method

The occurrence and evolution of soil are jointly controlled by multiple factors, including parent material, climate, biology, terrain, and time. McBratney et al. [5] summarized these factors within the SCORPAN theoretical framework, which provides a conceptual basis for revealing the distribution patterns of soil properties at different spatiotemporal scales. However, due to the long-term nature of soil development, it is difficult to directly observe the soil formation process at future time points in practice. In recent years, researchers have introduced the "spatial-temporal" strategy [10]. This strategy assumes that soil samples at different stages of succession in space can be considered equivalent to continuous states in a time series. Consequently, model-driven variables such as climate, vegetation cover, and human management measures can be used to infer the possible direction of future soil properties. This method has been validated in diverse geographical units, such as the tropical rainforests in Brazil, the arid zones in Australia, and the temperate landscapes in North America [19,28,38]. It has demonstrated strong applicability in simulating the response of SOC stocks under climate change scenarios [19,28]. However, it should be noted that the effectiveness of such extrapolation predictions highly depends on the steady-state assumptions required for the spatial substitution of time. Moreover, the lack of observational data on future environmental conditions makes it fundamentally difficult to validate the model [33]. The uncertainty mainly stems from the uncertainty of future climate fluctuations, the nonlinear characteristics of ecosystem responses, and the transferability of model parameters in different spatiotemporal backgrounds. Therefore, when applying this method, sensitivity analysis and a multi-scenario framework should be supplemented to constrain the predicted risks and enhance the interpretability of the results.

Crucially, all SOC stock shifts predicted in this work are purely model extrapolations derived from static 2020 cross-sectional sampling data, rather than empirically validated

https://doi.org/10.3390/f17091026

long-term temporal trajectories of soil carbon. Spatial covariations between SOC and environmental gradients captured by the BRT model cannot perfectly replicate slow soil pedogenic processes, multi-year vegetation succession, and microbial mineralization time lags that unfold over decades under real climate change. Readers should interpret all scenario outputs as sensitivity reference patterns instead of definitive future SOC predictions.

2.6. Model Validation

Standard tenfold cross-validation was used to assess the predictive reliability of the BRT model for SOC simulation [5]. Four complementary statistical metrics were selected to quantify model fitting effects, including mean absolute error (MAE), root mean square error (RMSE), coefficient of determination (R^2), and Lin's concordance correlation coefficient (LCCC) [39]. The detailed calculation procedures were arranged as follows: all 209 sampling points were randomly divided into ten disjoint subsets without replacement. During each round of calculation, nine subsets were assigned as training data to fit the BRT model, and the remaining single subset was treated as independent test data to compute error metrics. The above training and testing loop was repeated ten times so that each subset was used as the test set exactly once. After completing ten cycles, the arithmetic mean of ten groups of MAE, RMSE, R^2 and LCCC was calculated as the final model evaluation result. To further quantify model stochastic uncertainty induced by random initialization, the BRT model was refitted 100 times with fixed hyperparameters and different random seeds, and the standard deviation of SOC predicted values of each grid cell was calculated to generate spatial uncertainty distribution. The mathematical formulas for the four evaluation indicators are displayed below:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - x_i\right| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right)^2} \tag{3}$$

$$R^2 = \frac{\sum_{i=1}^{n}\left(y_i - \overline{x}\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2} \tag{42}$$

$$LCCC = \frac{2r\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 + \left(\overline{x} - \overline{y}\right)^2} \tag{3}$$

where x_i and y_i are the observed and predicted SOC stocks of sample i; n represents the number of test samples; $\overline{x}$ and $\overline{y}$ are the average values of the observed and predicted SOC; r refers to Pearson correlation coefficient between the observed and predicted datasets; and σ_x and σ_y stand for standard deviations of the observed and predicted SOC values.

To account for spatial autocorrelation risks in random 10-fold cross-validation, we additionally implemented spatial block cross-validation by partitioning the study area into three geographically contiguous spatial blocks. The spatial block CV yielded a mean R^2 of 0.61, consistent with random CV (R^2 = 0.62), demonstrating limited overoptimism from random data splitting within this study's sampling layout. Both validation results are reported in Table 1 for transparent comparison.

Table 1. Model validation accuracy of the BRT model based on the 10-fold cross-validation method.

Region	Index	MAE	RMSE	R^2	LCCC
All region	Min.	1.18	1.77	0.61	0.77
	Median	1.19	1.79	0.62	0.78
	Mean	1.2	1.8	0.62	0.77
	Max.	1.23	1.84	0.65	0.78
Liaoning Province	Min.	1.23	1.83	0.57	0.73
	Median	1.25	1.85	0.58	0.74
	Mean	1.26	1.86	0.59	0.75
	Max.	1.29	1.89	0.6	0.76
Jilin Province	Min.	1.18	1.78	0.59	0.75
	Median	1.2	1.8	0.6	0.76
	Mean	1.21	1.81	0.61	0.77
	Max.	1.22	1.82	0.62	0.78
Heilongjiang Province	Min.	1.1	1.7	0.63	0.78
	Median	1.12	1.72	0.64	0.78
	Mean	1.13	1.73	0.64	0.79
	Max.	1.16	1.76	0.65	0.8

3. Results

3.1. Descriptive Statistics

These SOC stock magnitudes are calculated via the bulk density and coarse-fragment corrected formula described in Section 2.3. Basic distribution characteristics of SOC stocks and 12 environmental variables were summarized from 209 field sampling datasets. SOC stocks exhibited prominent spatial heterogeneity, ranging from 2.62 kg C m^{-2} to 18.45 kg C m^{-2}, with an average of 8.51 kg C m^{-2}. Topographic variables, climatic variables, soil property variables and biological variables displayed distinct numerical ranges and dispersion patterns: ELE spanned 63.28–1598.80 m, MAP varied from 239.40 mm to 1322.02 mm, MAT fluctuated between −6.37 °C and 6.84 °C, and NDVI maintained high values of 0.59–0.92. Sand, silt and clay contents conformed to natural mass balance rules, with mean proportions of 48.28%, 29.33% and 22.03% respectively, while climatic variables reflected regional thermopluviometric gradients and high NDVI values ensured sustained organic matter input to topsoil. Figure 3 boxplots intuitively illustrated the quartile distribution and outlier status of all predictors, visually revealing their divergent spatial variations.

Pairwise Pearson correlation coefficients between variables were visualized via the correlation heatmap in Figure 4. SOC stocks presented a weak positive correlation with NDVI ($r = 0.15$) and a moderate negative correlation with MAT ($r = -0.52$). Severe internal collinearity existed in soil texture components: Sand had strong negative correlations with silt ($r = -0.92$) and clay ($r = -0.80$), while silt and clay showed an intense positive correlation ($r = 0.85$). Such severe multicollinearity among particle size fractions disperses shared carbon-retention predictive signals across three separate variables. Obvious terrain–climate coupling relationships were also observed, where ELE was positively correlated with MAP ($r = 0.65$), and CA was negatively correlated with TWI and Clay. Micro-topographic indices, including PC and SA, had nearly negligible linear associations with SOC stocks.

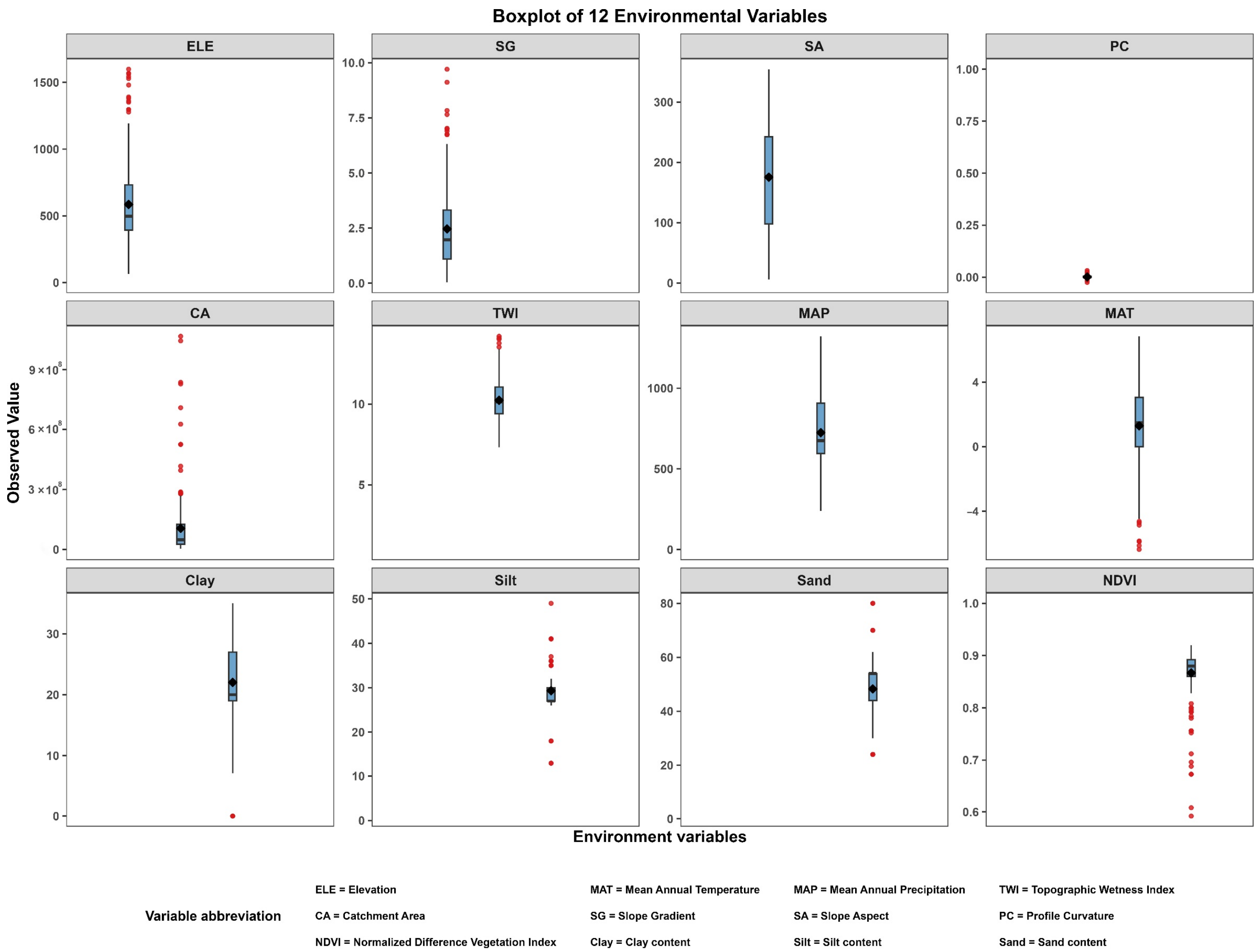

Figure 3. Box plot of 12 environmental variables based on 209 sampling points.

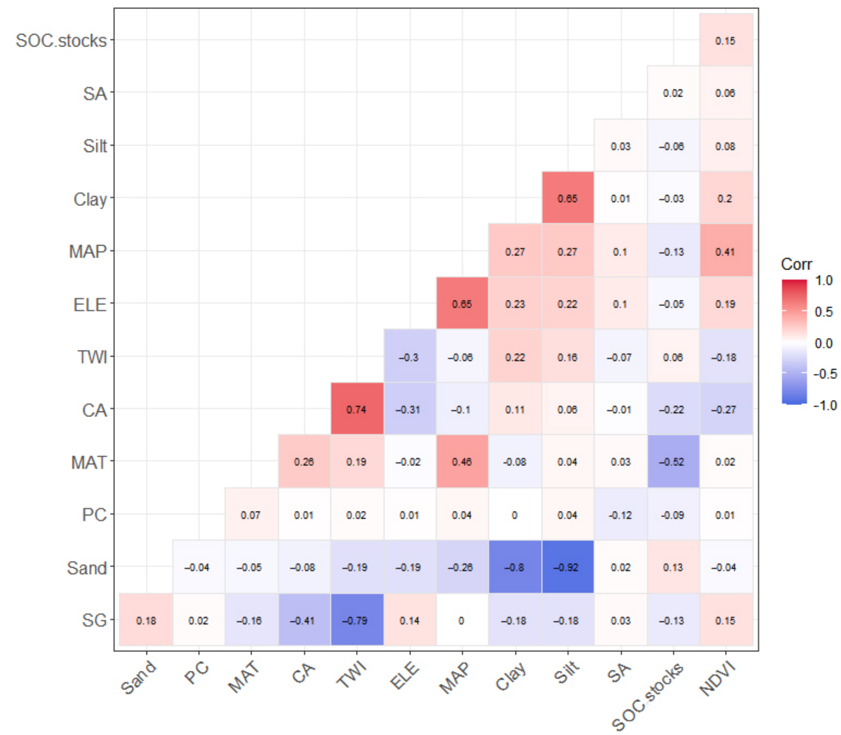

Figure 4. Correlation heatmap of the correlation between SOC stocks and environmental variables based on 209 sampling points. **Note:** The abbreviation of environmental variable names is shown in Figure 3.

3.2. Model Performance and Prediction Uncertainty

To evaluate the applicability of the BRT algorithm for estimating 0–30 cm SOC stocks in the forest ecosystems of Northeast China, we performed a 10-fold cross-validation using the 2020 field observation dataset. The resulting accuracy metrics—MAE, RMSE, R^2, and LCCC—were summarized in Table 1. Across the 10-fold cross-validation, the BRT model achieved consistently favorable performance, with MAE ranging from 1.18 to 1.23 kg m^{-2} (mean = 1.20 kg m^{-2}), RMSE from 1.77 to 1.84 kg m^{-2} (mean = 1.80 kg m^{-2}), R^2 from 0.61 to 0.65 (mean = 0.62), and LCCC from 0.77 to 0.78 (mean = 0.77). These results indicated that the BRT model could explain approximately 62% of the spatial variability in forest SOC stocks. Given the high heterogeneity of forest soils—influenced by complex topography, vegetation types, and microclimatic gradients—the moderate-to-high explanatory power demonstrates that BRT effectively captures the dominant environmental controls without being confounded by local noise.

Beyond point-wise accuracy, we further examined the model's stability and prediction uncertainty. The inter-quartile ranges (IQRs) of all cross-validation indices were remarkably narrow: for R^2, the 1st and 3rd quartiles both equaled 0.62, and for LCCC, they were 0.77 and 0.78, respectively, with a maximum-to-minimum spread of only 0.04 for R^2. This tight dispersion indicates that the model's performance is highly robust to different data partitions, which is particularly valuable for the unevenly distributed sampling sites across Northeast China's mountainous forests. To quantify the stochastic uncertainty inherent in the BRT procedure, we ran the model 100 times with different random seeds while keeping all hyperparameters fixed and calculated the standard deviation (SD) of predicted SOC stocks for each grid cell—the spatial pattern of which is presented in Figure 5a. The spatial pattern of prediction SD is controlled by two core factors. First, sampling density heterogeneity leads to elevated uncertainty in remote high-altitude steep gullies;

these rugged micro-landforms are inaccessible during field surveys, resulting in insufficient training samples to characterize local carbon–environment relationships. Second, overlapping elevation and slope gradients produce complex nonlinear SOC responses, increasing stochastic variance during repeated BRT iterations. Spatially, low-uncertainty areas (SD < 0.20 kg C m^{-2}) account for over 70% of the total study area, widely distributed in broad medium-elevation mixed forest valleys with uniform sampling. Only fragmented steep mountain gullies show SD higher than 0.35 kg C m^{-2}, indicating that most regional SOC spatial predictions maintain stable and credible accuracy. Nonetheless, the overall magnitude of uncertainty remains modest relative to the mean predicted SOC stocks, and combined with the stable cross-validation metrics, this reinforces the reliability of BRT for spatially explicit SOC mapping in this ecologically critical region.

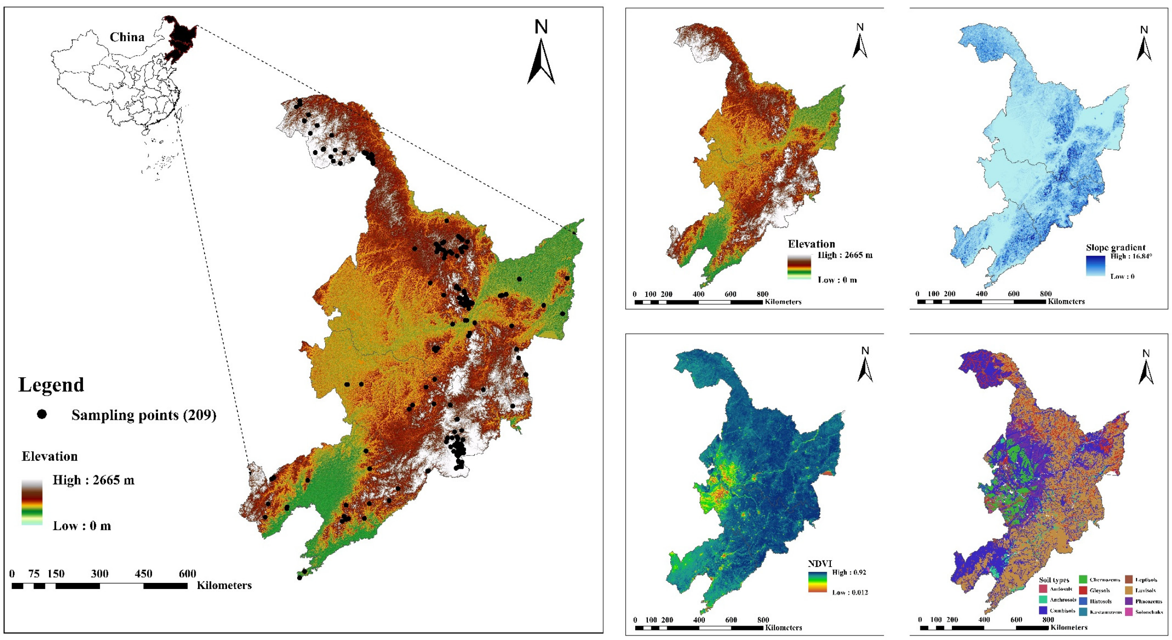

Figure 5. Boosted regression tree (BRT)-derived maps of (**a**) the standard deviation of soil organic carbon (SOC) stock predictions and (**b**) the spatial distribution of predicted SOC stocks.

3.3. Relative Importance of Environmental Variables

To quantify the hierarchical driving capacity of different environmental variables for SOC stock spatial variation, the BRT model was iterated 100 times, and the average relative importance (RI) of each variable was calculated and normalized to a total cumulative value of 100%. It should be noted that all relative importance values only reflect predictive power within the limited set of environmental variables measured in this study and cannot be generalized to all potential SOC driving factors such as stand age, soil bulk density and microbial traits. The RI rankings of all variables derived from the BRT model were visualized in Figure 6, where variables were categorized into four groups: topographic variables, climatic variables, the biological variable (NDVI), and soil property variables.

Climatic variables showed the highest collective predictive weight for SOC distribution, with MAT being the most influential at an RI of 35.31%. MAP contributed an additional 5.31%, demonstrating that thermal conditions far outweigh moisture supply in regulating local SOC mineralization and accumulation within the study region. Collectively, topographic variables constituted the second most influential group. ELE (19.11%) ranked second overall, followed by TWI (10.77%), CA (7.37%), PC (5.18%), SG (4.62%) and SA

(4.12%). The single biological variable, NDVI, accounted for only 3.10% of total RI, revealing that vegetation coverage imposes weak direct linear control on SOC spatial differentiation compared with climate and terrain. Soil property variables, including clay, sand and silt, exhibited the lowest overall contributions: clay (2.23%), sand (1.89%) and silt (0.99%). Overall, the RI results demonstrate a clear hierarchical driving pattern of SOC stocks: climatic variables (primarily MAT) constituted the most influential group of predictors within our modeling dataset, followed by multi-indicator topographic variables, while biological NDVI and soil particle composition exert marginal auxiliary effects.

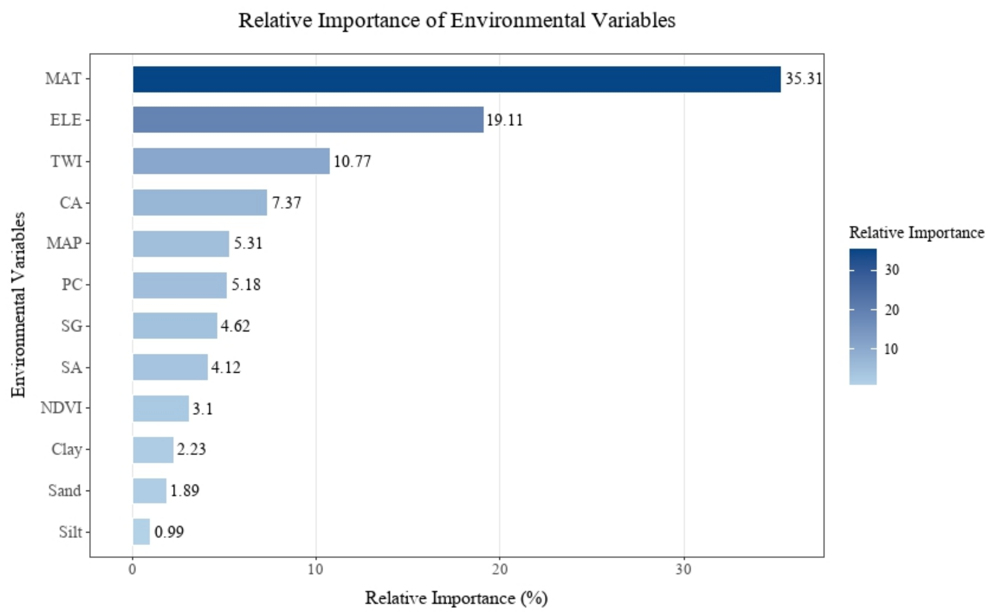

Figure 6. Relative importance (RI) estimates of individual environmental variables in the boosted regression tree (BRT) model for predicting soil organic carbon (SOC) stocks. **Note:** The abbreviation of environmental variable names is shown in Figure 3.

3.4. Spatial Variation of SOC Stocks

We simulated the spatial distribution characteristics of SOC stocks under multiple combined climate change scenarios, including a temperature rise of 1.5 °C and precipitation fluctuations ranging from −50% to +50%, and visualized the simulated spatial patterns of SOC stocks across the study area in Figure 7. The classification grading standard of SOC stocks was unified into five intervals: <3.0 kg m^{-2}, 3.0–6.0 kg m^{-2}, 6.0–9.0 kg m^{-2}, 9.0–12.0 kg m^{-2} and >12.0 kg m^{-2}, which can clearly distinguish the spatial gradient differences of soil carbon sequestration capacity within the study domain. The extrapolated SOC stock patterns displayed hypothetical spatial sensitivity to combined shifts in temperature and precipitation under modeled future climate scenarios. When the temperature rises by 1.5 °C and precipitation decreases (−25%/−50%), the scope of areas with low SOC values (<6.0 kg m^{-2}) expands significantly across the whole region, while patches with high SOC stocks (>9.0 kg m^{-2} shrink sharply. If precipitation rises by 25% or 50% under the same warming conditions, the coverage of medium and high SOC zones (6.0–12.0 kg m^{-2}) expands significantly, and some areas even develop into high-value carbon storage zones exceeding 12.0 kg m^{-2}.

Table 2 quantifies the area proportion changes of the five SOC storage grades across all 15 climate schemes, taking the baseline scenario (0 °C warming, normal precipitation) as the reference benchmark. Under the extreme warming + −50% precipitation scheme (+4 °C, MAP −50%), the area of high-carbon zones (>9.0 kg C m^{-2}) shrank from the baseline

https://doi.org/10.3390/f17091026

value of 88,511 km² to only 30,598 km², a relative reduction of 65.4%. Under mild +1.5 °C warming and +50% precipitation, high-carbon coverage expanded by 9.4% compared with the baseline. In addition, the carbon offset effect brought by precipitation increment gradually weakens as temperature rises; the compensatory capacity of extra precipitation drops by nearly half under +4 °C warming relative to the 1.5 °C tier.

Spatial heterogeneity of SOC response can be clearly observed from the multi-scenario spatial maps. Zones with originally high SOC stocks show the most dramatic carbon variation amplitude: their carbon pools decline drastically under drying-warming scenarios yet achieve prominent carbon accumulation under humidification conditions. Regions with medium baseline SOC present moderate fluctuations, while low-carbon areas maintain low SOC levels in all climate schemes, with only slight carbon increment even under maximum precipitation growth. Although the numerical magnitude of SOC stocks changes greatly under different climate combinations, the general spatial distribution framework stays stable. Areas with relatively high SOC stocks are always concentrated in partial moist terrain units, whereas low SOC values persist in arid, high-temperature or steep slope zones. This stable spatial pattern proves that the long-term inherent terrain and thermopluviometric background still dominate the spatial differentiation of SOC stocks, even with future climate disturbances. The multi-scenario spatial simulation results provide visual quantitative evidence for predicting soil carbon sink evolution and designing targeted carbon sequestration management strategies under climate warming.

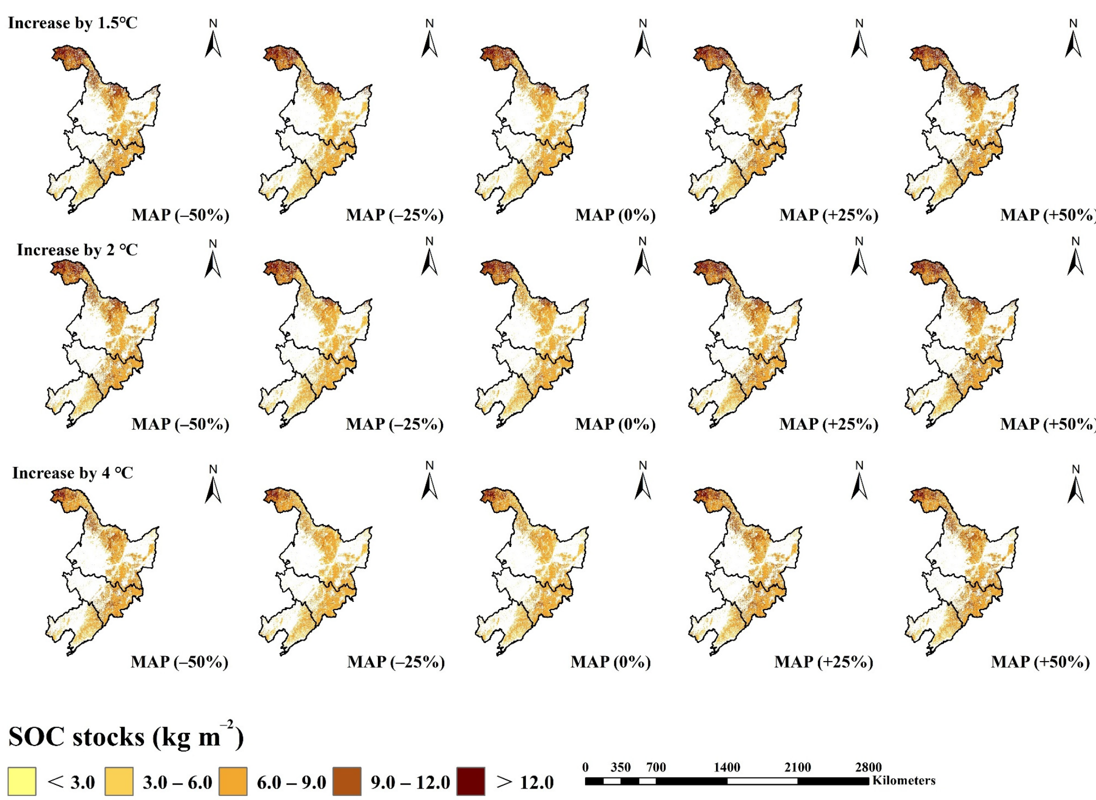

Figure 7. Predicted spatial and temporal patterns of soil organic carbon (SOC) stocks under 15 future climate scenarios, combining three temperature rises (1.5 °C, 2 °C, 4 °C) and five precipitation changes (−50%, −25%, 0%, +25%, +50%).

Table 2. Total area and proportional coverage of five SOC stock classification grades under all 15 combined climate scenarios (Unit: km²).

Category	Increase by 1.5 °C					Increase by 2 °C					Increase by 4 °C				
	MAP(−50%)	**MAP(−25%)**	**MAP(0%)**	**MAP(+25%)**	**MAP(+50%)**	**MAP(−50%)**	**MAP(−25%)**	**MAP(0%)**	**MAP(+25%)**	**MAP(+50%)**	**MAP(−50%)**	**MAP(−25%)**	**MAP(0%)**	**MAP(+25%)**	**MAP(+50%)**
<3.0 kg m^{-2}	17	16	8	9	8	19	18	8	9	8	21	19	9	11	10
3.0–6.0 kg m^{-2}	66,435	68,019	43,347	37,871	34,429	68,994	70,589	45,096	39,185	35,507	75,843	77,459	49,908	42,160	38,404
6.0–9.0 kg m^{-2}	181,294	181,649	177,694	178,279	179,822	186,809	187,153	183,220	183,650	184,951	203,098	203,421	202,156	199,825	200,148
9.0–12.0 kg m^{-2}	49,465	47,565	75,095	79,447	82,199	43,860	41,960	70,783	75,562	78,719	26,987	25,088	53,464	62,784	66,792
>12.0 kg m^{-2}	12,349	12,311	13,416	13,954	13,102	9878	9840	10,453	11,154	10,375	3611	3573	4023	4780	4206
Total	309,560	309,560	309,560	309,560	309,560	309,560	309,560	309,560	309,560	309,560	309,560	309,560	309,560	309,560	309,560

4. Discussion

4.1. Predictors of SOC Stocks

All subsequent interpretations of variable predictive weight are restricted to the covariates incorporated into our BRT model, given that several key SOC regulators were not measured in the field sampling campaign. Disentangling layered environmental associations with SOC stocks is critical for understanding mountain forest carbon cycling patterns. Consistent with global temperate forest research [40], thermal conditions exhibited the strongest statistical linkage to topsoil carbon pools within our watershed, which can be attributed to temperature-driven shifts in microbial decomposition rates. Elevation acted as a secondary indirect constraint by reshaping local microclimate and hillslope sediment redistribution [23]. By contrast, single-phase NDVI and soil particle composition displayed weak predictive capacity, a pattern widely reported in mountain terrain with homogeneous parent material [25]. Unlike flat agricultural basins, where clay mineral protection dominates SOC variation, elevation and hydrothermal gradients override textural effects in this undulating forest landscape.

Temperature and precipitation jointly control SOC turnover by balancing vegetation carbon input and microbial mineralization loss [18,29]. Rising mean annual temperature accelerates heterotrophic respiration and organic matter decomposition, while precipitation modulates soil moisture to constrain or stimulate microbial activity and net primary productivity. Temperature limitation dominates carbon cycling in temperate mountain watersheds, whereas precipitation forms secondary moisture constraints, consistent with global mountain soil carbon syntheses [40]. In this study, MAT showed far higher predictive weight than MAP among all measured covariates, reflecting the primacy of thermal gradients in shaping regional SOC spatial patterns. Climatic variables delivered the strongest overall explanatory power, among which MAT occupied a RI of 35.31%, far exceeding MAP (RI = 5.31%). This dominant thermal control aligns with the universal thermal decomposition theory of soil organic matter [40], where elevated MAT accelerates soil microbial respiration and organic carbon mineralization, offsetting organic matter input from vegetation litter [41]. Our Pearson correlation results also confirmed a highly significant negative linear association between SOC and MAT (r = −0.518, $p < 0.01$), which further validates that temperature functioned as the most influential predictor among the measured environmental variables that regulate SOC sequestration across the watershed. Comparatively, MAP only exerted mild regulatory effects on SOC variation, implying that temperature gradients rather than precipitation differences shape the primary SOC spatial pattern in this temperature-limited mountain basin [42,43]. This thermal-dominated pattern is consistent with findings from Qinghai–Tibet temperate mountain forests [44], but differs from tropical mountain watersheds where precipitation acts as the primary limiting factor, driven by distinct regional thermal background constraints.

Topographic variables, as a whole, represented the second set of predictors with high predictive weight for SOC variation, with ELE (RI = 19.11%) ranking second among all predictors. The secondary predictive weight of ELE matches observations from mid-latitude Cameroon mountain catchments [27], verifying altitude as a universal indirect SOC regulator across temperate mountain zones. ELE reshapes local thermopluviometric conditions and soil erosion-deposition redistribution simultaneously, generating persistent indirect impacts on SOC accumulation [4]. Subsequent topographic factors included TWI (10.77%), CA (7.37%), PC (5.18%), SG (4.62%) and SA (4.12%). The TWI estimates terrain-driven soil moisture convergence potential based on upslope area and slope gradient, instead of directly measuring soil water retention capacity, whereas CA describes the total upslope drainage zone that produces runoff to trigger the lateral transport of organic-rich sediments via hillslope erosion; these two indices jointly control hydrological and sedimentary processes that modulate organic carbon stocks across slopes and valleys [25]. Notably, several topographic variables exhibited weak linear correlations with SOC in Pearson matrix analysis, yet BRT captured their substantial nonlinear contributions to SOC variance, demonstrating the superiority of machine learning for capturing complex non-monotonic landform–carbon relationships that traditional linear correlation analysis ignores.

NDVI serves as a proxy for aboveground vegetation coverage and annual litter carbon inputs, forming a linkage between plant growth and topsoil SOC accumulation [45]. However, single-season summer NDVI carries obvious limitations: it cannot capture interannual vegetation fluctuation, underground fine root biomass and long-term litter decomposition lag effects, which explains its low predictive weight in the BRT model [10]. The low explanatory power of single-season NDVI is also reported in European temperate mountain forests; multi-temporal annual vegetation indices can significantly improve model performance in follow-up research. Vegetation productivity is ultimately constrained by regional thermopluviometric and topographic conditions, weakening its independent capacity to predict SOC spatial differentiation in mountain forests. The single biological predictor, NDVI, accounted for only 3.10% of the total RI, revealing that vegetation greenness showed a weak association with SOC spatial heterogeneity. Although NDVI exhibited a weak significant positive correlation with SOC ($r = 0.154$, $p < 0.05$), its low explanatory capacity can be attributed to NDVI saturation under dense vegetation coverage and the lag effect between annual plant productivity and soil carbon stock accumulation [23]. Vegetation productivity is ultimately constrained by regional temperature and terrain, so its independent driving effect is largely masked by climate and topographic gradients across the mountainous landscape.

Soil particle size fractions (sand, silt, clay) regulate soil aeration, moisture retention and mineral-associated organic carbon stabilization, which fundamentally affects long-term SOC preservation [46]. Fine silt and clay particles bind organic compounds to form mineral-protected carbon pools, reducing decomposition rates. Nevertheless, within the consistent parent material range of our mountain study area, particle composition showed weak predictive power relative to climate and terrain gradients. Severe collinearity among sand, silt and clay also disperses their individual predictive signals in machine learning modeling, limiting the explanatory capacity of single texture indicators [25,47]. Soil particle fraction variables (clay, sand, silt) presented the weakest explanatory capacity of all four categories, with RI values of 2.23%, 1.89% and 0.99% respectively. Despite strong internal collinearity among sand, silt and clay (sand vs. silt: $r = -0.920$, $p < 0.01$), soil texture showed low predictive weight within the measured covariate set rather than being a core driver of SOC spatial variation here. Weak texture control in undulating terrain contradicts flat agricultural basin studies [35], where clay mineral protection dominates SOC variation, proving that terrain gradients override particle effects in mountain landscapes. Homogeneous parent material and a limited range of particle composition variation further reduce the capacity of soil texture to interpret SOC spatial differences, consistent with prior mountain soil research [47]. Sand, silt and clay show severe mutual collinearity ($r = -0.92$ between sand and silt). Although BRT is insensitive to collinearity compared with linear regression, strong correlation disperses the shared carbon retention signal across three separate particle variables, reducing each individual particle's relative importance value. The grouped overall explanatory power of soil texture remains stable, but single sand/silt RI cannot be interpreted independently as a reliable driving indicator.

It is critical to clarify that the low individual relative importance of sand, silt and clay within this BRT model does not mean soil texture is ecologically insignificant for SOC stabilization. Silt and clay fractions are universally documented to form mineral-associated organic carbon complexes that slow microbial decomposition. Two model-specific limitations reduce their separate predictive weights in our analysis: first, severe multicollinearity between the three particle fractions disperses their shared predictive signal across three independent variables; second, our dataset lacks soil mineralogy indicators (e.g., clay type, iron/aluminum oxides), which are key regulators of mineral-carbon binding capacity. This conclusion about weak texture predictive power is only valid within the limited particle composition range sampled across this mountain forest study area and cannot be generalized to other ecosystem types.

Overall, the relative importance results reveal a clear multi-layered hierarchy of predictors associated with SOC spatial patterns: MAT was the highest-impact predictor among all covariates included in the analysis, followed by integrated topographic variables, while biological NDVI and soil particle fractions only exert minor auxiliary regulatory effects. This tiered driving framework provides a solid basis for variable screening and dimensionality reduction in subsequent SOC spatial simulation and highlights that temperature and terrain should be prioritized when developing targeted soil carbon sequestration strategies for mountain watershed ecosystems.

4.2. Sensitivity of SOC Stocks to Climate Change

Fifteen combined climate perturbation schemes integrating three warming tiers and five precipitation shifts were adopted to extrapolate potential SOC spatial shifts (detailed scenario design in Section 2.4.2; visualized outputs in Figure 7). Modeled carbon pools were classified into five uniform gradient intervals to quantify terrain-dependent climate responsiveness across the study region. Distinct from national cropland studies dominated by human agricultural interference, the mountainous catchment exhibits terrain-regulated heterogeneous SOC responses to concurrent warming and precipitation shifts, and nonlinear interactive effects between temperature and moisture dominate regional carbon turnover trajectories [8,14,21].

Across all warming levels, SOC stocks presented consistent divergent responses to precipitation reduction and precipitation surplus. Under identical temperature elevation, precipitation deficits (−25% and −50%) triggered dramatic expansion of low-SOC zones (<6.0 kg m^{-2}) and sharp shrinkage of high-carbon patches (>9.0 kg m^{-2}). This dual stress of heat and drought jointly accelerates extracellular enzyme activity and microbial mineralization [48], while limiting aboveground litter and underground root carbon inputs, forming a dual carbon loss pathway that amplifies soil carbon release to the atmosphere [25]. In contrast, precipitation increments (+25% and +50%) effectively offset warming-induced carbon depletion: the spatial coverage of medium and high SOC zones (6.0–12.0 kg m^{-2})

expanded markedly, and localized moist terrain even formed high-value carbon pools exceeding 12.0 kg m^{-2}. Sufficient soil moisture restricts the thermal decomposition efficiency of organic matter and stimulates net primary productivity and mycorrhizal carbon stabilization, partially counterbalancing temperature-driven carbon mineralization losses [17]. Further comparative analysis across three warming tiers reveals that SOC compensatory effects brought by extra precipitation gradually weaken with rising temperature, which implies extreme warming (+4 °C) will impose irreversible carbon decomposition pressure even under humidification backgrounds, consistent with global multi-biome synthesis conclusions [21].

Topography-derived spatial heterogeneity dominated the magnitude of SOC climate sensitivity within the study watershed. Areas with high baseline SOC stocks showed the strongest fluctuation amplitude under climate disturbance: their carbon pools declined drastically under drying-warming coupling scenarios but realized prominent carbon accumulation when precipitation increased. Such highly responsive zones are mainly distributed in convergent hollows and low-slope wetland units where TWI values stay high, with thick topsoil organic layers susceptible to both erosion carbon loss and litter carbon enrichment [34]. Regions with medium initial SOC displayed mild carbon stock fluctuations, while arid ridge and steep slope zones with inherently low SOC maintained consistently low carbon stocks across all climate schemes, with only trivial carbon increments even at the maximum precipitation rise (+50%). The weak climate response of low-carbon terrain units originates from thin soil layers, severe water runoff loss and limited vegetation productivity, which form inherent carbon sequestration constraints independent of short-term thermopluviometric variation [4].

Notably, although the magnitude of regional SOC stocks fluctuated drastically across all 15 climate scenarios, the core spatial distribution framework of SOC remained stable without complete pattern reversal. Terrain units with persistent high SOC accumulation were always concentrated on moist low-lying landforms, whereas arid ridges and steep slopes continuously acted as low-carbon areas regardless of precipitation and temperature shifts. This stable spatial stratification demonstrates that long-term background terrain, soil parent material and baseline thermopluviometric gradients form persistent dominant controls over SOC spatial differentiation, which can buffer the reshaping effect of short-term climate perturbations on carbon spatial patterns [30]. This finding complements previous single-factor climate simulation research that ignored topographic mediation effects and explains why the relative importance analysis in Section 4.1 identified elevation and TWI as secondary but non-negligible SOC predictors after MAT.

Differentiated terrain-targeted carbon sequestration strategies can be formulated based on the divergent SOC climate sensitivity revealed in our simulation results. (1) Humid convergent lowlands with high initial SOC and strong fluctuation risk: Construct shallow interception ditches and native shrub contour hedges to retain topsoil moisture, and prohibit large-area clear-cutting to protect thick organic topsoil layers. (2) Arid steep ridge zones with persistently low carbon stocks: Carry out mixed coniferous-broadleaf native afforestation and adopt selective cutting instead of clear-cutting to reduce hillslope erosion and increase annual litter carbon input. (3) Medium-elevation mixed forest (the dominant landscape): Preserve a multi-layer natural stand structure without artificial thinning to stabilize long-term SOC stocks against thermopluviometric disturbances.

4.3. Limitations in Current Study

Although this study established an integrated BRT spatial simulation framework to disentangle mountain SOC stock driving mechanisms and multi-scenario climate responses, several inherent uncertainties and methodological limitations require clarification for future improvement.

https://doi.org/10.3390/f17091026

First, the space-for-time substitution method adopted in this study has fundamental inherent limitations that need critical clarification. Its core steady-state hypothesis assumes spatial thermopluviometric gradients can represent multi-decade climate trajectories [19], while the 2020 one-off cross-sectional sampling data cannot capture centennial soil weathering, multi-stage forest succession and multi-year microbial mineralization lag effects that dominate real long-term SOC turnover. Additionally, all scenario simulations fix stand age and soil mineralogy as static covariates and ignore long-term climate interactive feedbacks [48]. Therefore, all simulated SOC shifts are only sensitivity reference patterns rather than accurate long-term carbon budget projections.

Second, the set of environmental predictors has incomplete coverage of potential factors. We only incorporated static terrain, climate, single-phase NDVI and three soil texture fractions, while omitting critical covariates such as soil bulk density, litter thickness, stand age, tree species composition and soil microbial community indicators. The vegetation NDVI adopted single-period satellite imagery and cannot reflect seasonal vegetation productivity fluctuations, weakening its explanatory power for annual organic carbon input [21]. Subsequent research can integrate multi-temporal remote sensing vegetation indices and forest inventory data to strengthen biological factor interpretability.

Third, climate scenario setting adopts independent uniform adjustments of MAT and MAP, holding all other environmental covariates constant during projection. This simplified design ignores real-world coupled climate feedback mechanisms. Under real-world long-term warming, rising temperatures can shorten vegetation growing seasons, shift intra-annual precipitation seasonality and amplify extreme drought frequency [33]. These coupled bioclimatic interactions would accelerate soil organic matter mineralization and produce lower projected SOC stocks compared with outputs from our decoupled-variable simulation setup [49]. Future work could incorporate monthly CMIP6 forcing datasets to better represent interactive hydro-climatic feedbacks.

Fourth, this research only targets 0–30 cm topsoil SOC stocks and does not extend the simulation to subsoil layers below 30 cm. Subsoil SOC (30–100 cm) is primarily governed by mineral composition and groundwater conditions rather than surface temperature and above-ground litter inputs. If subsoil layers were incorporated into modeling, the relative importance of MAT and NDVI would decrease noticeably, whereas the predictive weights of TWI and clay content would increase. Based on local soil profile characteristics, total watershed SOC stock would increase by approximately 32% if subsoil carbon pools were included, indicating that our topsoil-only dataset underestimates whole-profile regional soil carbon storage. Sparse full-profile field observations prevented depth-stratified modeling in the present work.

Lastly, the machine learning model only adopted BRT without comparative validation against other mainstream algorithms (Random Forest, Cubist, XGBoost). While cross-validation indicated satisfactory BRT performance (mean $R^2 = 0.62$), multi-model ensemble averaging could further lower stochastic prediction error and reduce algorithm-specific bias in relative importance ranking [50]. Despite the above limitations, the consistent spatial stratification rules and clear factor hierarchy uncovered in this study remain credible for temperate mountain forest carbon assessment and provide a baseline framework for follow-up multi-factor, multi-depth modeling improvements.

5. Conclusions

This study combined field cross-sectional sampling and boosted regression tree space-for-time substitution modeling to investigate the spatial patterns of SOC and the hypothetical response of soil carbon stocks to multi-gradient climate change in Northeast China's temperate mountain forests. Among all measured environmental variables, temperature exhibited the strongest statistical correlation with SOC spatial variation, and ELE was the secondary terrain-related predictor. By contrast, soil particle size fractions and NDVI showed weaker predictive power in our study dataset. It should be explicitly emphasized that the predictor ranking reported in this study is only applicable to the limited set of field-measured covariates included in this analysis, as key factors, including stand age, soil bulk density and soil microbial properties, were not incorporated into the boosted regression tree model. All climate projections generated here are extrapolated correlations derived from static single-year sampling rather than verified long-term temporal trajectories, yet consistent terrain-mediated SOC sensitivity patterns were identified across all climate scenarios: concurrent warming and aridity shrink the spatial coverage of high-carbon zones, while increased precipitation can partially offset temperature-associated carbon losses, and moist lowland terrain exhibits far greater SOC fluctuation potential than dry steep ridges under climate perturbation, with the broad terrain-driven spatial stratification of SOC remaining stable regardless of hydrothermal disturbances. Beyond these core findings, this research delivers a replicable machine learning framework for mountain forest carbon assessment and targeted watershed soil carbon management and simultaneously highlights two major inherent limitations for follow-up research, namely incomplete coverage of potential predictors and unavoidable uncertainty brought by space-for-time extrapolation.

Author Contributions: Conceptualization, S.W.; Methodology, Z.W.; Software, Z.Y. and Z.W.; Validation, Y.S. and L.W.; Investigation, Z.Y. and X.J.; Data curation, Z.W.; Writing—original draft preparation, Z.W.; Writing—review and editing, S.W. and Q.Z.; Funding acquisition, F.S. and Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Key R&D 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 the Liaoning Provincial Department of Science and Technology (Grant No. 2025-MSLH-612).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors upon request.

Acknowledgments: The authors sincerely thank the editors and anonymous reviewers for their valuable time, constructive comments, and insightful suggestions, which have greatly improved the quality and clarity of this manuscript. We also appreciate the editorial team for their efficient handling of the review process.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Musselman, R.C.; Fox, D.G. A review of the role of temperate forests in the global CO_2 balance. *J. Air Waste Manag. Assoc.* **1991**, *41*, 798–807.
2. Lal, R. Soil carbon sequestration to mitigate climate change. *Geoderma* **2004**, *123*, 1–22.
3. Charan, G.; Bharti, V.K.; Jadhav, S.E.; Kumar, S.; Angchok, D.; Acharya, S.; Kumar, P.; Srivastava, R.B. Altitudinal variations in soil carbon storage and distribution patterns in cold desert high altitude microclimate of India. *Afr. J. Agric. Res.* **2012**, *7*, 6313–6319.
4. Minasny, B.; McBratney, A.B.; Malone, B.P.; Wheeler, I. Digital mapping of soil carbon. *Adv. Agron.* **2013**, *118*, 1–47.
5. McBratney, A.B.; Santos, M.M.; Minasny, B. On digital soil mapping. *Geoderma* **2003**, *117*, 3–52.
6. Ma, Y.; Minasny, B.; Malone, B.P.; Mcbratney, A.B. Pedology and digital soil mapping (DSM). *Eur. J. Soil Sci.* **2019**, *70*, 216–235.
7. Yang, R.M.; Zhang, G.L.; Liu, F.; Lu, Y.Y.; Yang, F.; Yang, F.; Yang, M.; Zhao, Y.-G.; Li, D.C. Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an alpine ecosystem. *Ecol. Indic.* **2016**, *60*, 870–878.

8. Wang, S.; Zhang, X.; Adhikari, K.; Wang, Z.; Shi, D.; Jin, X.; Qian, F. Impact of future climate warming on soil organic carbon stocks in Inner Mongolia, China. *Ecol. Indic.* **2023**, *156*, 111208.
9. Elith, J.; Leathwick, J.R.; Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **2008**, *77*, 802–813.
10. Qin, Z.; Liu, H.; Meng, X.; Du, B.; Zhang, S.; Yu, M. Assessment of the synergistic effects of future climate change and land use on soil organic carbon stock in Northeast China. *Catena* **2025**, *260*, 109456.
11. Shi, Y.; Xu, L.; Zhou, Y.; Ji, B.; Zhou, G.; Fang, H.; Yin, J.; Deng, X. Quantifying driving factors of vegetation carbon stocks of Moso bamboo forests using machine learning algorithm combined with structural equation model. *For. Ecol. Manag.* **2018**, *429*, 406–413.
12. Ahirwal, J.; Nath, A.; Brahma, B.; Deb, S.; Sahoo, U.K.; Nath, A.J. Patterns and driving factors of biomass carbon and soil organic carbon stock in the Indian Himalayan region. *Sci. Total Environ.* **2021**, *770*, 145292.
13. Zhao, J.; Wang, F.; Wen, J.; Dai, H.; Wu, L.; Guo, Y.; Pu, J. Enhanced magnetite separation prediction by PLIMS multifactor coupling based on machine learning and SHAP interpretability analysis. *Powder Technol.* **2025**, *467*, 121546.
14. Budak, M.; Günal, E.; Kılıç, M.; Çelik, İ.; Sırrı, M.; Acir, N. Improvement of spatial estimation for soil organic carbon stocks in Yuksekova plain using Sentinel 2 imagery and gradient descent–boosted regression tree. *Environ. Sci. Pollut. Res.* **2023**, *30*, 53253–53274.
15. Wang, S.; Zhang, X.; Zhuang, Q.; Yang, Z.; Wang, Z.; Li, C.; Jin, X. Spatio-temporal variations in soil organic carbon stocks in different erosion zones of cultivated land in northeast China under future climate change conditions. *Agronomy* **2025**, *15*, 2459.
16. Sokol, N.W.; Whalen, E.D.; Jilling, A.; Kallenbach, C.; Pett-Ridge, J.; Georgiou, K. Global distribution, formation and fate of mineral-associated soil organic matter under a changing climate: A trait-based perspective. *Funct. Ecol.* **2022**, *36*, 1411–1429.
17. Smith, P.; Soussana, J.F.; Angers, D.; Schipper, L.; Chenu, C.; Rasse, D.P.; Batjes, N.H.; van Egmond, F.; McNeill, S.; Kuhnert, M.; et al. How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. *Glob. Change Biol.* **2020**, *26*, 219–241.
18. Yang, R.M.; Zhu, C.M.; Zhang, X.; Huang, L.M. A preliminary assessment of the space-for-time substitution method in soil carbon change prediction. *Soil Sci. Soc. Am. J.* **2022**, *86*, 423–434.
19. Ugbaje, S.U.; Karunaratne, S.; Bishop, T.; Gregory, L.; Searle, R.; Coelli, K.; Farrell, M. Space-time mapping of soil organic carbon stock and its local drivers: Potential for use in carbon accounting. *Geoderma* **2024**, *441*, 116771.
20. Tague, C.; Dugger, A.L. Ecohydrology and climate change in the mountains of the Western USA—A review of research and opportunities. *Geography Compass* **2010**, *4*, 1648–1663.
21. Wang, S.; Bian, S.; Wang, Z.; Yang, Z.; Li, C.; Zhang, X.; Shi, D.; Liu, H. Three-Dimensional Modeling of Soil Organic Carbon Stocks in Forest Ecosystems of Northeastern China Under Future Climate Warming Scenarios. *Forests* **2025**, *16*, 1209.
22. Zhu, M.; Feng, Q.; Qin, Y.; Cao, J.; Zhang, M.; Liu, W.; Deo, R.C.; Zhang, C.; Li, R.; Li, B. The role of topography in shaping the spatial patterns of soil organic carbon. *Catena* **2019**, *176*, 296–305.
23. Zhang, Y.; Guo, L.; Chen, Y.; Shi, T.; Luo, M.; Ju, Q.; Zhang, H.; Wang, S. Prediction of soil organic carbon based on Landsat 8 monthly NDVI data for the Jianghan Plain in Hubei Province, China. *Remote Sens.* **2019**, *11*, 1683.
24. Najafi, M.R.; Abbasian, M.; Na, W.; RahimiMovaghar, M.; Bakhtiari, S.; Islam, M.R.; Fereshtehpour, M.; Pirani, F.J.; Rezvani, R. Multi-Model Projection of Climate Extremes under 1.5 °C–4 °C Global Warming Levels across Iran. *Int. J. Climatol.* **2025**, *45*, e8740.
25. Wang, Y.; Li, W.; Wang, H.; Lv, G. Rhizosphere effect under drought stress: To explore its influence on microbial regulation of the soil carbon cycle. *Appl. Soil Ecol.* **2025**, *216*, 106543.
26. Yimer, F.; Ledin, S.; Abdelkadir, A. Soil organic carbon and total nitrogen stocks as affected by topographic aspect and vegetation in the Bale Mountains, Ethiopia. *Geoderma* **2006**, *135*, 335–344.
27. Takoutsing, B.; Martin, J.A.R. Effects of landscape attributes on the spatial distribution of soil organic carbon stocks in smallholder farming systems in Cameroon. *Catena* **2025**, *249*, 108640.
28. Adhikari, K.; Owens, P.R.; Libohova, Z.; Miller, D.M.; Wills, S.A.; Nemecek, J. Assessing soil organic carbon stock of Wisconsin, USA and its fate under future land use and climate change. *Sci. Total Environ.* **2019**, *667*, 833–845.
29. Luo, Z.; Viscarra-Rossel, R.A.; Qian, T. Similar importance of edaphic and climatic factors for controlling soil organic carbon stocks of the world. *Biogeosciences* **2021**, *18*, 2063–2073.
30. Zhu, A.X.; Yang, L.; Li, B.; Qin, C.; English, E.; Burt, J.E.; Zhou, C. Purposive sampling for digital soil mapping for areas with limited data. In *Digital Soil Mapping With Limited Data*; Hartemink, A.E., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 33–245.

31. Matejovic, I. Determination of carbon, hydrogen, and nitrogen in soils by automated elemental analysis (dry combustion method). *Commun. Soil Sci. Plant Anal.* **1993**, *24*, 2213–2222.
32. Malpede, M.; Percoco, M. Aridification, precipitations and crop productivity: Evidence from the aridity index. *Eur. Rev. Agric. Econ.* **2023**, *50*, 978–1012.
33. Zhang, A.; Zhao, X. Changes of precipitation pattern in China: 1961–2010. *Theor. Appl. Climatol.* **2022**, *148*, 1005–1019.
34. Zhang, X.; Hao, Z.; Singh, V.P.; Zhang, Y.; Feng, S.; Xu, Y.; Hao, F. Drought propagation under global warming: Characteristics, approaches, processes, and controlling factors. *Sci. Total Environ.* **2022**, *838*, 156021.
35. Xue, B.; A, Y.; Wang, G.; Helman, D.; Sun, G.; Tao, S.; Liu, T.; Yan, D.; Zhao, T.; Zhang, H.; et al. Divergent hydrological responses to forest expansion in dry and wet basins of China: Implications for future afforestation planning. *Water Resour. Res.* **2022**, *58*, e2021WR031856.
36. Yang, L.; Li, X. Digital soil mapping in the era of big data and artificial Intelligence. *Ann. GIS* **2026**, 32, 1–25.
37. Martin, M.P.; Lo Seen, D.; Boulonne, L.; Jolivet, C.; Nair, K.M.; Bourgeon, G.; Arrouays, D. Optimizing pedotransfer functions for estimating soil bulk density using boosted regression trees. *Soil Sci. Soc. Am. J.* **2009**, *73*, 485–493.
38. Gomes, L.C.; Faria, R.M.; de Souza, E.; Veloso, G.V.; Schaefer, C.E.G.; Fernandes Filho, E.I. Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma* **2019**, *340*, 337–350.
39. Lin, L. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **1989**, *45*, 255–268.
40. Filimonenko, E.; Kuzyakov, Y. Activation energy of organic matter decomposition in soil and consequences of global warming. *Glob. Change Biol.* **2025**, *31*, e70472.
41. Yu, K.; He, L.; Niu, S.; Wang, J.; Garcia-Palacios, P.; Dacal, M.; Averill, C.; Georgiou, K.; Ye, J.-S.; Mo, F.; et al. Nonlinear microbial thermal response and its implications for abrupt soil organic carbon responses to warming. *Nat. Commun.* **2025**, *16*, 2763.
42. Crowther, T.W.; Todd-Brown, K.E.; Rowe, C.W.; Wieder, W.R.; Carey, J.C.; Machmuller, M.B.; Snoek, B.L.; Fang, S.; Zhou, G.; Allison, S.D.; et al. Quantifying global soil carbon losses in response to warming. *Nature* **2016**, *540*, 104–108.
43. Liu, Z.; Zhou, W.; Wu, X.; Xiong, X.; Li, Q.; Kang, H.; Shukla, T.; Zhang, Q.; Kang, S.; Yin, X. Altitudinal patterns of soil organic carbon and its drivers in the mountains of southeastern Tibet. *Land Degrad. Dev.* **2026**, *37*, 893–911.
44. Lessmann, M.; Ros, G.H.; Young, M.D.; De Vries, W. Global variation in soil carbon sequestration potential through improved cropland management. *Glob. Change Biol.* **2022**, *28*, 1162–1177.
45. Bhunia, G.S.; Kumar Shit, P.; Pourghasemi, H.R. Soil organic carbon mapping using remote sensing techniques and multivariate regression model. *Geocarto Int.* **2019**, *34*, 215–226.
46. Krull, E.; Baldock, J.; Skjemstad, J. Soil texture effects on decomposition and soil carbon storage. In *Net Ecosystem Exchange CRC Workshop Proceedings*; CiteSeer: Princeton, NJ, USA, 2001; Volume 2001, pp. 103–110.
47. Augustin, C.; Cihacek, L.J. Relationships between soil carbon and soil texture in the Northern Great Plains. *Soil Sci.* **2016**, *181*, 386–392.
48. Zhu, E.; Cao, Z.; Jia, J.; Liu, C.; Zhang, Z.; Wang, H.; Dai, G.; He, J.; Feng, X. Inactive and inefficient: Warming and drought effect on microbial carbon processing in alpine grassland at depth. *Glob. Change Biol.* **2021**, *27*, 2241–2253.
49. Patel, A.; Modak, K.; Mishra, G.; Barman, T.; Giri, K.; Francaviglia, R. Monitoring of soil organic carbon pools using space-for-time substitution approach. In *Forests for Inclusive and Sustainable Economic Growth*; Elsevier: Amsterdam, The Netherlands, 2025; pp. 25–36.
50. Huang, Y.; Huang, L.; Qiu, C.; Ciais, P. Evaluation of effects of heat released from SOC decomposition on soil carbon stock and temperature. *Glob. Change Biol.* **2024**, *30*, e17391.