https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

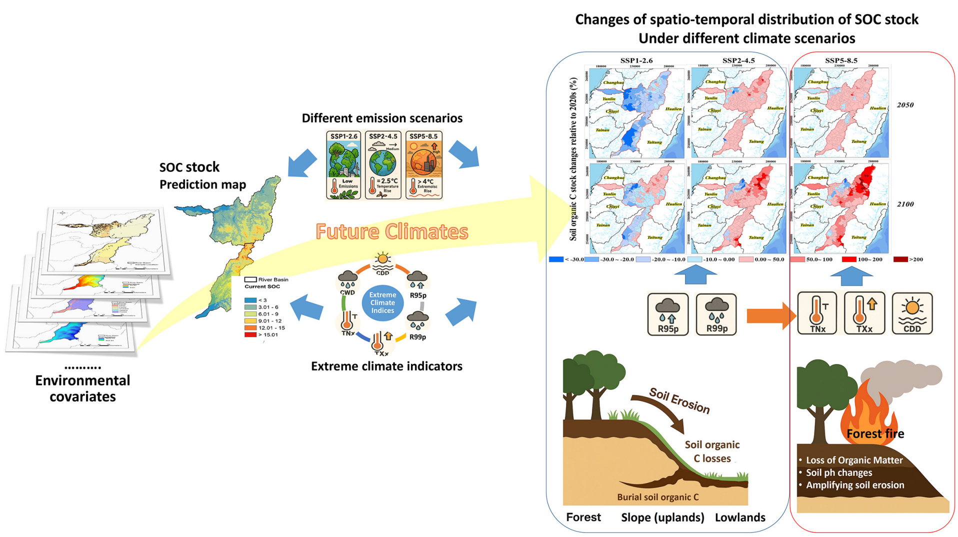
Soil Organic Carbon Projections and Climate Adaptation Strategies across Pacific Rim Agro-ecosystems

Chien-Hui Syu[1], Chun-Chien Yen[1,2], Selly Maisyarah[2], Bo-Jiun Yang[1], Yu-Min Tzou[2], and Shih-Hao Jien[2]*

[1] Agricultural Chemistry Division, Taiwan Agricultural Research Institute, Taichung City 40227, Taiwan, ROC

[2] Department of Soil and Environmental Sciences, National Chung Hsing University, Taichung 40227, Taiwan, ROC

Corresponding author: Shih-Hao Jien; E-mail: shjien@nchu.edu.tw



https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

Abstract

In Pacific Rim regions highly exposed to climate variability, accurate projections of soil organic carbon (SOC) are critical for furture effective land management and climate adaptation strategies. This study integrated digital soil mapping with CMIP6-based climate projections to estimate the spatiotemporal distribution of SOC stocks in subtropical (Zhuoshui River) and tropical (Laonong River) watersheds in Taiwan. We collected 1377 soil samples and data on 18 environmental covariates and modeled SOC stocks at a 20-m resolution through the Cubist and random forest algorithms, which were also combined with regression kriging. The Cubist-based kriging model was discovered to achieve the highest performance in SOC stock prediction. Forested areas were found to contain >80% of SOC stocks, and tropical zones were discovered to store substantially less carbon than subtropical zones. Future emission scenarios revealed spatial heterogeneity in SOC stock dynamics. In scenario SSP1-2.6, a maximum SOC stock decline of approximately 20.9% was predicted, particularly for uplands, because of erosion induced by extreme rainfall events (R95p and R99p), whereas in scenarios SSP2-4.5 and SSP5-8.5, increases of 7.9% to 58% were predicted, respectively; particularly corresponded to forested areas because of enhanced productivity caused by increased TNx and TXx (extremes of minimum and maximum temperature). Partial least squares path modeling revealed a climate–topography interaction in SOC stocks, dominated by topography and followed by prolonged dry spells. Examining the interactions between climatic extremes, landscape types, and SOC stocks is essential for enhancing soil resilience and ensuring stable SOC stocks in the future.

Keywords: soil organic carbon, stock, digital soil mapping, climatic extreme, emission scenario, land type

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

1. Introduction

Soil organic carbon (SOC) is one of the largest carbon pools in the global carbon cycle (Grace, 2004) and is a key concern of agricultural and environmental policies (Johnston et al., 2004). SOC also has a crucial influence on the carbon cycle at the local and global levels (Singh et al., 2018). Multiple studies have examined whether carbon storage in agricultural soils can offset global warming, and many scales have been developed for evaluating the dynamics of SOC. Of these scales, the landscape scale has enabled researchers to consider the interplay between natural processes, human patterns, and SOC dynamics (Viaud et al., 2010). Therefore, this scale is the most appropriate for assessing environmental and agricultural ecosystems (Li et al., 2021).

Changes in climatic conditions such as temperature, carbon dioxide concentration, and precipitation may influence the dynamics of SOC by affecting the rates of soil processes such as mineralization, decomposition, leaching, and total carbon loss. In areas prone to climatic extremes—such as floods, droughts, and heat waves—these conditions may further affect the dynamics of SOC (Li et al., 2021; Chalchissa et al., 2022). In addition, extreme climate events may strongly affect the content of SOC, with subsequent effects on agricultural productivity and ecosystem services. Therefore, before the dynamics of SOC can be evaluated at the landscape scale in response to climate change, a spatiotemporal technique is required. Zhu and Lin (2010) argued that in areas with major terrain variation and low sampling density, utilizing a non-geostatistical approach or a combination of geostatistical and non-geostatistical approaches can improve prediction ability.

In digital soil mapping (DSM), soil properties in unsampled or partially sampled areas are predicted through numerical models developed using various statistical methods or algorithms; these models primarily rely on soil observational data and corresponding environmental factors (Grunwald, 2009). According to the literature, spatial variations in soil properties play a key role in model construction (Zhu and Lin, 2010). Two modeling approaches are commonly used to predict soil properties in unsampled or partially sampled sites: non-geostatistical approaches and geostatistical approaches. Non-geostatistical approaches are based on the SCORPAN model (Jenny, 1941; McBratney et al., 2003) and include multiple linear regression (MLR) models, generalized additive models, Cubist models (Quinlan, 1992), and random forest (RF) models (Breiman, 2001). Geostatistical approaches account for spatial autocorrelation in data and include ordinary, simple, and universal kriging. In addition, regression kriging is a hybrid spatial interpolation approach that combines the results of a regression model (such as Cubist

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

and RF) with the spatial interpolation of its residuals (Ma et al., 2017). The performance of a model is influenced by factors such as spatial scale (Poggio et al., 2010), observation density (Tsui et al., 2016; Keskin and Grunwald, 2018), and terrain (Zhu and Lin, 2010). Machine learning algorithms, including MLR, RF, and Cubist, have been widely used for mapping SOC content and SOC stocks (Lamichhane et al., 2019; Siewert, 2018; Yang et al., 2016; Akpa et al., 2016; Gray and Bishop, 2016; Rudiyanto et al., 2018). Although regression kriging provides favorable predictions for RF (Guo et al., 2015) and Cubist (Dorji et al., 2014; Ma et al., 2017) models, Vaysse and Lagacherie (2015) argued that it does not offer any advantages in specific scenarios. Therefore, given the absence of a universal model, Lamichhane et al. (2019) emphasized the importance of meta-analytical evaluations.

Taiwan is located on the frontline of the Pacific Rim and is highly prone to the combined effects of climate change and El Niño–Southern Oscillation events. The frequency and spatial variability of extreme climate events in this region are expected to dramatically increase in the future. Therefore, understanding the combined effects of extreme climate variability and long-term climate change on regional climate and SOC variation is essential for evaluating the vulnerability of regional agriculture, water resources, and ecosystems. In Taiwan, the Zhuoshui River Watershed (ZRW) and Laonong River Watershed (LRW) are the two largest and most crucial agricultural eco-watersheds. The ZRW includes fluvial plains and is one of the most essential agricultural areas in Taiwan. In the ZRW, rice, vegetables, and other crops are extensively cultivated. This high agricultural activity underscores the importance of SOC in sustaining soil fertility and agricultural production. In the LRW, SOC plays an essential role in supporting the limited amount of agriculture that is practiced. These two watersheds are located in different climatic zones, which may affect their SOC dynamics. The ZRW is located in central Taiwan and has a subtropical climate, whereas the LRW is located in southern Taiwan and has a tropical monsoon climate. Future climate-change-related changes in temperature and precipitation may substantially affect the content and total stocks of SOC in these two regions, altering their agricultural production and land use patterns.

Therefore, this study applied digital soil mapping approaches to generate high-resolution maps of SOC stock distribution in the surface layer (0–30 cm) of the ZRW and LRW in Taiwan, with the aim of better understanding the spatiotemporal dynamics of SOC under different emission scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5) and various extreme climate indicators projected for 2050 and 2100.

2. Materials and Methods

2.1 Research area

The Zhuoshui River watershed (ZRW) is located in central Taiwan. Its basin covers Changhua County, Nantou County, Yunlin County, and Chiayi County. The watershed has an area of 3156.9 km^2 and is located in a subtropical climate zone (Fig. 1a and 1b). This area has an average annual temperature of 8.6–23.6 °C and receives annual cumulative rainfall of 834.6–3693.4 mm. The elevation of the study area ranges from 0 to 3844.2 m above sea level, and it has diverse topography consisting of mountains (2060.7 km^2, 65.2%), hills (843.3 km^2, 26.6%), and plains (261.7 km^2, 8.2%). In terms of soil classification, the upstream areas are primarily characterized by stony soils, whereas the western plains are predominantly characterized by silty alluvial soils. The Laonong River watershed (LRW) is located in southern Taiwan, representing an upper mainstream area of the Gaoping River. Its basin covers Nantou County, Kaohsiung City, Pingtung County, and Taitung County. The watershed has an area of 2038 km^2 and is located in a tropical climate zone. This area has an average annual temperature of 19.5 °C, and it received annual cumulative rainfall of 3222.6 mm during the period 2011–2020. According to the 2015 Land Cover Survey, the upper reaches of the basin are predominantly forested areas (73.3%), whereas the downstream gentle slopes and plains are predominantly agricultural areas (10.6%; Fig. S1).

2.2 Soil samples and analyses

Soil survey data for the period 2012–2020 were obtained from the Taiwan Agricultural Research Institute. A total of 1377 topsoil samples (0–30 cm) were obtained. Each sample's location was recorded using a handheld global positioning system device. After the samples had been air-dried at room temperature, they were sieved through a 35-mesh screen and stored in plastic containers. They were then analyzed using the loss-on-ignition (LOI) method (Nelson and Sommers, 1996). Because the LOI method typically overestimates SOC (Li et al., 2021), a correction function was applied to adjust SOC content from LOI values to those obtained using a TOC analyzer (solid TOC cube, Elementar). The correction equation is as follows:

$$TOC = 0.7084 * LOI - 0.0986\ (R^2 = 0.94;\ P < 0.001) \qquad [1]$$

where TOC and LOI refer to the SOC contents (%) determined by the TOC analyzer and the LOI method, respectively. A sample's total bulk density was determined using the clod

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

method or soil core method (Blake and Hartge 1986). Finally, the soil organic carbon stock (SOC_{stock}, kg m^{-2}) was calculated using the following equation:

$$SOC_{stock} = TOC * \rho * D/10 \quad [2]$$

where TOC is the SOC content (%), ρ is the bulk density of soil (g cm^{-3}), and D is the soil depth (cm). Owing to the substantial variability in coarse fragment content, this study excluded them from the calculation of SOC stocks.

2.3 Environmental covariates

Environmental covariates were categorized on the basis of factors pertaining to soil formation, including topographical data such as digital elevation models (DEM), satellite remote sensing imagery, meteorological data, land use survey data, and soil order (Table 1). All environmental covariates were resampled at a spatial resolution of 20 m by using R software version 4.0.5 (R Foundation for Statistical Computing, Vienna, Austria).

The DEM was derived from a 20-m grid numerical terrain model established by the Taiwanese Ministry of the Interior. To create an elevation map, the "Fill Sinks" function of SagaGIS 8.0.1 was used to smooth the discontinuities in the model. These elevation data were employed to generate relevant topographical attributes. These attributes included the slope, aspect, terrain ruggedness index (TRI), terrain position index (TPI), topographic wetness index (TWI), multiresolution index of valley bottom flatness (MrVBF), multiresolution ridge top flatness (MrRTF), curvature, flow accumulation, and stream power index (SPI). Ma et al. (2017) argued that topographical parameters can serve as environmental covariates in organic carbon prediction models.

The normalized difference vegetation index (NDVI) was calculated using infrared (b4) and near-infrared (b8) satellite imagery data (Sentinel 2) for the period 2016–2020 to determine the proportion of space covered by vegetation, which was produced by the Google Earth Engine at a resolution of 20 m. Climate is one of the key soil-forming factors and, according to Wiesmeier et al. (2019), a major driver influencing SOC storage. This study used climatic data from 2011 to 2020, including mean annual temperature (MAT) and total annual precipitation (TAP), obtained from Taiwan's Central Weather Bureau. The original resolution of these data was 1 km. In addition to these factors, land cover type also influences SOC storage (Edmondson et al., 2014). Therefore, this study used a 2015 land cover map produced by TARI,

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

classified into five categories: (1) paddy fields; (2) upland farming (including miscellaneous grains, tea trees, betel nuts, and bamboo); (3) orchards; (4) forests (including plantations, primary forests, and high-mountain arrow bamboo forests); and (5) others (miscellaneous and riverine lands).

2.4 Climate data in various emission scenarios and with extreme climate indices

Future climate predictors were obtained from CMIP6-based global climate models, including shared socioeconomic pathways (SSPs) established by the Intergovernmental Panel on Climate Change (IPCC). Model MIROC6—developed by the Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, University of Tokyo—was selected for predicting SOC stocks from future climate data. The historical and projected extreme climate indices of CMIP6 were employed for different socioeconomic pathways, specifically for scenarios SSP1-2.6, SSP2-4.5, and SSP5-8.5. These data were used to examine climate patterns and trends and establish models for predicting the impact of future climate change on various environments.

In this study, we identified the following six extreme climatic indices: consecutive dry days (CDD), consecutive wet days (CWD), very wet day precipitation (R95P), extremely wet day precipitation (R99P), minimum value of daily maximum temperature (TNn), and maximum value of daily maximum temperature (TXx). These indicators provided valuable insights into the effects of extreme weather events across the study area. According to Chalchissa and Kuris (2024), the correlation between these indicators and soil health factors may offer a comprehensive understanding of soil health and the potential for carbon sequestration in agricultural systems.

2.5 Predictive models

Because of research advancements in the field, the techniques used in DSM have evolved from simple linear models to comprehensive machine learning technologies (Minasny and McBratney, 2016). In this study, two widely used data mining models, namely Cubist and RF models, were employed. Both models were further combined with regression kriging to account for both geographical and non-geographical effects, resulting in the Regression Kriging with Cubist and Regression Kriging with Random Forest models. Their capabilities in predicting the spatial distribution of SOC were compared.

The Cubist model is a rule-based classification algorithm proposed by Quinlan (1992). It

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

was developed on the basis of the M5 tree model. The Cubist model segregates data into several subsets on the basis of "if–then" patterns and identifies linear relationships between the target variables and environmental covariates in each subset. In the present study, we used the Cubist package in R software version 4.0.5 for model development. We adopted the following parameters: (1) rules for data classification based on rule count, (2) extrapolation for determining the model's degree of extrapolation, and (3) committees that generate multiple committee models on the basis of the number of samples to be processed in the model (i.e., the amount of data) to refine the previous prediction and output collective results. We did not set specific rules or extrapolations but instead relied on Cubist defaults. Committees were calculated using the caret package and set at 20.

The RF model is an ensemble learning algorithm introduced by Breiman (2001). This model reconstructs a data set into multiple new sets with identical sample size through random resampling during model training. For each data set, environmental covariates are randomly selected for constructing classification or regression trees. In the case of continuous variables, the model's predicted value is the average output of all regression trees. In this study, we used the "randomForest" package of R software version 4.0.5 for model development. We adopted the following parameters: (1) mtry, which determines the number of environmental covariates extracted for each new data set in regression tree construction, and (2) ntree, which determines the number of regression trees in the RF. We also used the caret package and set these parameters to mtry = 7 and ntree = 500.

2.6 Model training and validation

Before model development, we employed the "rpart" package in R software version 4.0.5 to extract 70% of the data as the training data set (calibration set), totaling 600 samples. We used the remaining 301 samples (30%) as the validation data set (validation set) and used it to determine the model's predictive performance. The distribution of the two data sets is depicted in Fig. 2. Model performance was evaluated by comparing the predicted values with the observed values in the validation group. Root mean square error (RMSE) and coefficient of determination (R^2) values were used as assessment indicators and calculated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - p_i)^2}, \qquad [3]$$

$$R^2 = \frac{\sum_{i=1}^{n}(p_i - \mu_o)^2}{\sum_{i=1}^{n}(o_i - \mu_o)^2}, \qquad [4]$$

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

where p_i and o_i represent the predicted and observed values, respectively, and μ_o represents the mean observed value. Because of the spatial variability of soil properties, effectively quantifying uncertainty in predictive map outputs is crucial. Therefore, in this study, the prediction interval was calculated using 90% quantiles of upper and lower limits of prediction through the bootstrap method (Malone et al., 2014).

2.7 Geographic information and data analyses

In this study, maps were created and spatial statistical analyses were conducted using the geographic information system (GIS) software ArcMap 10.8. Data processing and statistical analyses were conducted using Microsoft Excel 2016 (Microsoft, Redmond, WA, USA) and R software version 4.0.5.

Redundancy analysis is a multivariate or multiresponse technique similar to regression. In this study, redundancy analysis was conducted to determine whether extreme climatic indices and SOC stock changes were associated with the matched grid cells (different land types) evaluated in our previous study (Jien et al., 2025). All statistical analyses, unless indicated otherwise, were conducted using SPSS version 18.0 (SPSS, Chicago, IL, USA). A p-value of <0.05 was considered statistically significant. Partial least squares path modeling (PLS-PM) was employed to identify the pathways underlying the study variables, including emission scenarios, extreme climate indices, and SOC stock and land types. A PLS-PM model was constructed using the “innerplot” function of the “plspm” package. The model’s quality and performance were evaluated using the goodness-of-fit (GOF) test. Finally, the “ggplot2” package in R software was used for redundancy analysis and plot generation (Villanueva and Chen, 2019).

3. Results

3.1 Statistical description of SOC stock

The sampling sites of this study are presented in Fig. S1(a). The average topsoil SOC stock across all sampling points was 4.36 kg m^{-2}. In the ZRW, the topsoil SOC stock ranged from 0.19 to 31.8 kg m^{-2}, with an average of 4.51 kg m^{-2}. In the LRW, the topsoil SOC stock ranged from 0.41 to 14.4 kg m^{-2}, with an average of 3.80 kg m^{-2}. Although the data for the LRW were more concentrated, the overall data exhibited positive skewness, with a skewness value of 0.15. Therefore, a natural logarithm transformation was applied to improve model

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

performance by approximating a normal distribution.

3.2 Model performance in SOC stock prediction

This study constructed SOC stock predictive models using the Cubist, RF, and regression kriging with the training data set and environmental covariates. Coefficient of determination (R^2) and RMSE values were used to evaluate the performance of these models. Among the evaluated models, the RF model demonstrated the highest predictive performance in the training data set. Notably, Shaik and Srinivasan (2019) highlighted the likelihood of overfitting in RF models, indicating that while the model may predict training data accurately, it may perform poorly when applied to unseen data outside the training set. Therefore, we focused on model performance in the prediction of validation data set prior to model selection. In this respect, the performance indicators of the Cubist model were R2 = 0.43 and RMSE = 0.45 kg m-2, while those of the RF model were R2 = 0.46 and RMSE = 0.43. After incorporating regression kriging, the indicators improved to R^2 = 0.48 and RMSE = 0.42 for the Cubist model, and remained at R^2 = 0.46 and RMSE = 0.43 for the RF model (Fig. 2).

3.3 Importance analysis of environmental covariate

For the variable importance analysis in the RF model, the increase in mean squared error (MSE) was calculated when each covariate was excluded during the random selection process. In the Cubist model, the usage ratios of various environmental covariates were computed. These indicators revealed the key role of environmental covariates in the prediction of SOC stocks (Fig. 3a). In the RF model, covariates that led to an increase in mean squared error (IncMSE) greater than 15% included elevation (23%), soil order (20%), annual mean temperature (19%), and precipitation (15%). In the Cubist model, the primary classification factors were annual mean temperature (45%), soil order (18%), and elevation (17%). For the construction of grouped regression equations, all continuous covariates—except for slope aspect, curvature, and flow accumulation—were utilized by the Cubist model. Among these, more than half of the data incorporated covariates such as elevation (98%), annual mean temperature (62%), NDVI (59%), TRI (54%), K-value (53%), and slope (52%). In summary, the two models identified soil order, elevation, and annual mean temperature as the factors representing the influence of soil, topography, and climate, respectively, on the SOC stock in the study areas.

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

3.4 Predicted map of SOC stock

The predicted spatial distribution of SOC stock is presented in Fig. 3b. According to the statistical analysis of the prediction map, the mean SOC stock in the ZRW and LRW was 5.51 and 6.38 kg m^{-2}, respectively. The first quartile, median, and third quartile were 3.85, 5.31, and 6.88 kg m^{-2}, respectively, for the ZRW and 3.80, 5.96, and 8.85 kg m^{-2}, respectively, for the LRW. A reduction in SOC stock from forested areas to plain areas (lowlands) was found for both watersheds. In addition, high SOC stock (approximately 15 kg m^{-2}) was discovered in southeastern areas in the ZRW and in northeastern areas in the LRW. Lower storage values of <3 kg m^{-2} were found in downstream areas in the LRW and near the estuary in the ZRW. The areas located along the downstream plains of ZRW exhibited low SOC stock (<2.5 kg m^{-2}) near the river, with SOC stock higher farther from the river.

3.5 Uncertainty analysis for predictive models

For each sampling point in the training data set, prediction residuals were established through leave-one-out cross-validation for the regression kriging and Cubist models. The study area was then classified by landscape in accordance with the classification rules of the Cubist model (Table 2). Fig. 4 presents 90% confidence interval maps drawn using data segmentation and cross-validation techniques. These prediction limit intervals can be regarded as indicators of the model's uncertainty. In the downstream areas of the study region, the confidence interval widths were generally below 6 kg m^{-2}, whereas in the mountainous regions they were substantially higher, with some areas reaching up to 40 kg m^{-2}.

3.6 SOC stock distribution with various landscape types and land uses

The SOC stock spatial distribution was categorized on the basis of topography to demonstrate the distribution of SOC stocks under various landscape types. As shown in Fig. 5, in the lowlands of the ZRW, the lowest average SOC stock was identified in dry farming areas (1.93 kg m^{-2}), whereas the highest average SOC stock was identified in paddy fields (3.08 kg m^{-2}). In the lowlands of the LRW, the lowest average SOC stock was identified in "other" land cover types (1.89 kg m^{-2}), whereas the highest average SOC stock was identified in forested areas (3.17 kg m^{-2}). In the uplands of both catchments, the lowest cover was identified in paddy fields (3.01 and 2.16 kg m^{-2}), whereas the highest cover was identified in forests (4.22 and 3.6 kg m^{-2}). In terms of landscape type, the highest SOC stock was identified in forested areas, with 6.54 kg m^{-2} in the ZRW and 8.02 kg m^{-2} in the LRW, whereas the lowest SOC stock

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

was identified in orchard areas, with 4.82 kg m^{-2} in the ZRW and 5.6 kg m^{-2} in the LRW.

In all emission scenarios, major spatial heterogeneity and temporal increases were found in SOC stocks (Table 3, Figs. 6 and 7), particularly under high-emission conditions. These findings underscore the importance of modifying the management practices of land use in the future, especially if climate change is severe. In forested areas in both watersheds, significant SOC accumulation was predicted. Areas with an SOC accumulation value of >15 Mg C ha^{-1} were expected to exhibit an increase in SOC accumulation from <5% (2020, baseline) to more than 25% by 2100 in scenario SSP5-8.5. By contrast, lowland agricultural zones are expected to maintain relatively low SOC stocks (<9 Mg C ha^{-1}), with minor gains across scenarios. Scenario SSP5-8.5 was found to result in the greatest projected increase in SOC stocks as a result of elevated CO_2 and potential biomass input, although spatial disparities are expected to increase, particularly in erosion-prone or intensively cultivated lands (Fig. S2).

3.7 Extreme climate index parameter estimates in three emission scenarios

Extreme climate indices in three SSPs were compared: SSP1-2.6 (sustainable development), SSP2-4.5 (middle of the road), and SSP5-8.5 (fossil-fuel-based development). Projections were evaluated for mid-century (2050) and end-century (2100) time points at units of sub-catchment for each watershed. These units were classified as whole area, lowlands, uplands, and forested areas, denoted W, L, U, and F, respectively, in Tables S1 and S2.

The CDD, CWD, R95p/R99p, and TNx/TXx were analyzed as extreme climate indices. In all SSPs, the increases in temperature- and precipitation-related extremes in the two watersheds were significant. In scenario SSP5-8.5, the magnitude and spatial heterogeneity of these changes were predicted to intensify toward 2100 compared to the 2020 baseline (Fig. 6a). In the ZRW, scenario SSP5-8.5 was predicted to result in a prominent increase in CDD, especially by 2100, with uplands and forest areas projected to experience CDD increases of 145% and 188%, respectively (Table S1). By contrast, for scenario SSP1-2.6, the CDD was predicted to decrease slightly by 26.3%, particularly in lowlands and plains. For both watersheds, CWD was predicted to increase in the emission scenarios for 2050 and 2100, but different trends were discovered for the ZRW and LRW. Regarding the ZRW, CWD was predicted to significantly increase in lowlands and decrease in forested areas. However, in the LRW, it was predicted to increase only in uplands (Fig. 6a). These results indicated the polarization of wet–dry periods, particularly under high-emission conditions. A major increase in rainfall extremes was predicted, with R95p increasing by 1558 mm in the ZRW (entire area, scenario SSP5-8.5 for

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

the year 2100) (Table S1; S2). For the same scenario and time frame, R99p was predicted to reach 3829 mm, with uplands and forests receiving rainfall of 2634 and 2250 mm, respectively. Temperature extremes were also predicted to increase, especially in scenario SSP5-8.5. Regarding the ZRW, TXx was predicted to increase by up to 32.5% in forested areas (scenario SSP5-8.5 for the year 2100), whereas TNx was predicted to increase by 76.6% in uplands, indicating pronounced warming.

Regarding the LRW, which is characterized by a tropical monsoon climate, high warming and precipitation extremes were predicted, particularly in uplands and forested areas, indicating climate spatial heterogeneity. For scenario SSP5-8.5, an approximately 211% CDD increase for the entire watershed was predicted by 2100 (Table S2), with the largest increase predicted for forested areas (236%). In these forested areas, R95p and R99p were predicted to reach 772.8 and 2442 mm, respectively. Moreover, CWD was predicted to increase in uplands by up to 85.7%, suggesting prolonged wet conditions. Notably, TXx and TNx were predicted to substantially increase in highlands and forests, reaching up to 34.7% and 83.0%, respectively, emphasizing the intensification of heat extremes. Overall, this scenario may present another type of threat: long-term droughts with torrential downpours and extreme heat. These climate conditions may overwhelm current agricultural systems and infrastructure and undermine the current ecological carrying capacity (Zhang et al., 2020).

3.8 Relationships between extreme climate indices and SOC stocks

Principal component analysis (PCA) was conducted to examine the relationships between extreme climate indicators and SOC stock and to determine their topographic distribution characteristics for three scenarios. Regarding scenario SSP1-2.6, SOC stock variation exhibited negative correlations with R95p and R99p, indicating that extreme precipitation may be detrimental to the maintenance of SOC stocks (Fig. 8a). Pearson's correlation analysis revealed significant negative correlations of SOC stocks with R95p ($r = -0.32$, $p < 0.05$) and R99p ($r = -0.29$, $p < 0.01$; Fig. 8b). Regarding scenario SSP2-4.5, SOC stock variation exhibited a positive correlation with CWD ($r = 0.21$, $p < 0.05$; Fig. 8c and 8d), indicating that stable wet conditions may promote SOC accumulation under moderate emission conditions. Regarding scenario SSP5-8.5, SOC stock variation was not significantly correlated with most of the extreme climate indicators, indicating that SOC responses may be influenced by complex interactions under strong emission conditions (Fig. 8e and 8f). For all scenarios, strong positive R95p–R99p and TXx–TNx correlations were found (e.g., $r = 0.87$ between

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

R95p and R99p in scenario SSP5-8.5). Taken together, these findings suggest uniform increases in the frequencies of extreme rainfall and extreme heat events in terms of both spatial distribution and climatic mechanisms.

Our results revealed clear topographic effects, with distinct spatial variations in SOC stock dynamics predicted for the different emission scenarios. For scenario SSP1-2.6, SOC stocks were projected to decrease by 3% to 21%, whereas for scenarios SSP2-4.5 and SSP5-8.5, SOC stocks were projected to increase by 7.91% to 58.3%, particularly in forested areas (because of their enhanced net primary productivity). In addition, the variance and coefficient of variation (CV) of SOC stock percentages show a significant increase in interquartile range and CV over time under moderate (SSP1-2.6, SSP2-4.5) and high (SSP5-8.5) emission scenarios relative to the 2020 baseline. For both the moderate-emission (SSP1-2.6 and SSP2-4.5) and high-emission (SSP5-8.5) scenarios, the variance and coefficient of variation of the SOC stock percentage distribution were predicted to increase (Table 3). Collectively, these results indicate that future climatic extremes are projected to significantly increase the spatial heterogeneity of the percentage distribution of SOC stocks. In the majority of scenarios, particularly in scenario SSP5-8.5, uplands and forested areas are expected to exhibit a drastic response to extreme climate indicators, including major increases in CDD, R99p, TXx, and TNx. In uplands, SOC stocks are expected to respond strongly to the extreme climate, suggesting the susceptibility of slope soils to extreme rainfall and thermal destabilization. Furthermore, more extreme values are predicted for the LRW than for the ZRW, which is likely attributable to the topographic elevation distribution and baseline tropical monsoon climate in the LRW.

In our PLS-PM analysis, we discovered a goodness of fit (GOF) value ranging from 43.0 to 45.7, indicating the high explanatory power of our findings (Fig. 9). The PLS-PM results also revealed distinct differences in the controls for SOC stocks between the ZRW and LRW. For the ZRW (GOF = 45.7%), topographic variables (elevation and slope gradient) were found to have the strongest positive total effect on SOC stocks (standardized total effect = 0.579), followed by consecutive dry and wet periods (CDD and CWD, total effect = 0.238). Despite these findings, extreme rainfall events and temperature did not appear to have a direct effect on SOC stocks. For the LRW (GOF = 43.0%), stronger topographic control of SOC stocks was discovered (total effect = 0.753), with the direct path being positive (0.84, $p < 0.01$), indicating that the SOC accumulation patterns observed predicted this watershed will be closely linked to its land type. Regarding scenarios SSP2-4.5 and SSP5-8.5, most of the predicted increases in SOC stocks are concentrated in forested areas. Despite these results, the effects of temperature

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

extremes on SOC stocks will presumably be weakened by topographic and hydrological stability, particularly in the LRW. For both watersheds, prolonged dry spells were predicted to indirectly increase SOC stocks, whereas rainfall extremes were predicted to reduce SOC stocks, particularly in uplands (slope lands; Sarstedt et al., 2014). Taken together, these findings underscore the importance of incorporating topographic and extreme climate variables into SOC modeling and climate-resilient soil management strategies.

4. Discussion

4.1 Ability of machine learning models to predict SOC stocks

Lamichhane et al. (2019) argued that several predictive models can be used to predict the spatial distribution of SOC. In the present study, common machine learning models—the Cubist, RF, and regression kriging models—were used to predict SOC stock in the sampling area. Some of the collected data were randomly selected for validation (Fig. 2). The results indicated that integrating regression kriging into the Cubist model yielded the highest predictive performance (R^2 = 0.48, RMSE = 0.50), which was significantly higher than that achieved by the Cubist model alone (R^2 = 0.43, RMSE = 0.45). However, incorporation of regression kriging into the RF model resulted in limited improvements, consistent with the results of Vaysse and Lagacherie (2015). This limited improvement may be attributable to the inherently low residuals in the RF model, indicating that even standardizing these residuals and adding them to the model would have minimal effects on predictions. As shown in Fig. 4, SOC stocks were underestimated for several samples from mountainous regions with high organic carbon stock due to plant residues, suggesting that this type of variance was not captured by the models or residuals. Other studies have demonstrated various predictive disparities. These studies include those conducted by Lacoste et al. (2014), who applied a Cubist model to predict organic carbon stock in a France-based study (R^2 = 0.12, RMSE = 12.64); Adhikari et al. (2014), who employed regression kriging in a Denmark-based study (R^2 = 0.41, RMSE = 0.24); and Ma et al. (2017), who combined regression kriging with a Cubist model in a China-based study (R^2 = 0.25, RMSE = 0.12). Although the models used in the present study exhibited several predictive disparities, they nonetheless exhibited high reliability in terms of their overall predictive ability.

In SOC stock forecasting models, empirical estimations of uncertainty involve geographic spatial segmentation. In Cubist models, input data are divided into groups on the basis of a series of rule-based classifications. Therefore, examining the empirical distribution of the regression kriging residuals in each category is appropriate (Malone et al., 2014). According to classification rules, if low-elevation areas such as plains, foothills, and valleys involve young, weakly developed soils or miscellaneous lands, these areas are considered to exhibit a wide distribution of residuals. By contrast, other soil order categories are considered to exhibit a more concentrated distribution of residuals. In mountainous regions, classification is based on the NDVI, where a higher NDVI indicates more vegetation, whereas a low NDVI suggests the presence of water bodies or bare soils. In this study, analysis of 90% confidence interval maps (Fig. 4) revealed that for low-altitude areas with high sampling density, abundant data were available for model construction, leading to generally low prediction residuals. Even with various landscape classifications, the prediction limit intervals were low. By contrast, in mountainous areas, the existence of few samples and substantial variability in environmental covariates increased the difficulty of prediction. Generally, when establishing empirical divisions with insufficient data, outliers can easily influence the residual distribution in a given category. Therefore, future sample planning in these areas can be guided by such map data.

4.2 Effects of environmental covariates on SOC stocks

In this study, SOC stocks were influenced by the following key topographical attributes: elevation, the multiresolution index of valley bottom flatness, slope, and the topographic wetness index (Fig. 3a). As reported by Mishra and Riley (2015), who conducted an Alaska-based study, elevation is a crucial predictor of SOC stock, regardless of resolutions. According to Adhikari et al. (2014), the multiresolution index of valley bottom flatness (MrVBF) and the topographic wetness index (TWI) are important covariates, ranking just below precipitation. The MrVBF is used to identify flat valley bottoms and thereby indicate potential areas of erosion or deposition, whereas the TWI is used to indicate terrain's control over soil moisture, reflecting wet or dry conditions (Lamichhane et al., 2019). Slope affects SOC stocks by influencing solar radiation and moisture retention. Regarding meteorological covariates, annual cumulative precipitation and mean annual temperature are crucial in determining SOC stock. Gray et al. (2015) identified positive correlations of the SOC content of topsoil in New South Wales, Australia, with precipitation and relative humidity. According to Lamichhane et al. (2019), high precipitation may enhance vegetation growth or create anoxic conditions that slow soil carbon oxidation. Rial et al. (2017) reported a negative correlation between temperature and SOC content in Europe, with higher altitude and latitude found to correspond to slower SOC decomposition. In the present study, the effect of elevation was attributable to

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

temperature, particularly because the original resolution of the temperature data was 1 km, and high-resolution elevation data were necessary to obtain a detailed spatial distribution. Of the two biotic factors used in this study, the NDVI was the primary environmental covariate. Wang et al. (2018) highlighted the role of the NDVI as an indicator of vegetation cover, which is strongly correlated with SOC. They emphasized that convincing results could be obtained if long-term remote sensing data could be obtained to calculate NDVI values for multiple time periods.

4.3 Influence of land cover on the spatial distribution of SOC stocks

In this study, a trend of increasing SOC stock with increasing elevation was discovered, and this effect is likely driven by elevation and temperature (Fig. 3b, Fig. S2). The SOC stock in differing land coverage types was found to exhibit variation for different terrains. For instance, for lowlands, SOC stocks are predicted to rapidly decrease as a result of intensive cultivation, leading to low SOC stocks in agricultural production lands (Fig. 3b). Regarding the LRW, which has a tropical climate, SOC stock in farmlands was considerably lower than that in forests. This finding may be attributable to the farmlands in the LRW being frequently tilled for triple cropping or the mean annual temperature in the area being higher than that on the Zhuoshui River plains. In slope lands, the rice fields in the two basins had SOC stock levels similar to those observed for plains, although an increase was discovered for orchard and forested lands. In mountainous areas, the SOC stock predictions were higher than those for plains and slope lands across all types of land cover. According to the literature, the eastern region of the LRW experiences high precipitation and low temperatures, which result in higher organic carbon storage than that observed in the ZRW, particularly in forested areas (Fig. 3b; Guo et al., 2019). In addition, in mountainous areas, forests are the main space available for SOC stock.

4.4. Adaptation strategies for the management of SOC stocks in various emission scenarios

In this study, future climatic variables calculated from global climate models (GCMs) were used as inputs to estimate the spatiotemporal variation in global topsoil organic carbon stocks in 2020, 2050, and 2100. All GCM were obtained from the Coupled Model Intercomparison Project. CMIP6 was specifically selected for estimating future topsoil organic carbon stocks. Tables S1 and S2 list the extreme climate indicators involved in the considered scenarios (scenarios SSP1-2.6, SSP2-4.5, and SSP5-8.5) on the basis of CMIP6 data.

Our results indicated increasingly pronounced spatial heterogeneity in SOC stocks in the scenario involving severe greenhouse gas emissions (Table 3, Fig. 6). Severe warming in the future will cause increasingly regional climate variability, resulting in greater spatial heterogeneity in SOC stocks. This phenomenon will be particularly evident in countries with complex topography, such as those located on the frontline of the Pacific Rim, which is directly exposed to the threats of rapid climate change. Generally, the combination of steep terrain and intricate terrain complicates management of the SOC distribution because spatial variability driven by extreme climatic events is difficult to predict and control.

According to our results, extreme climate and land type are the most crucial determinants of SOC stocks in regions near the Pacific Rim (Dialynas et al., 2016; Wei et al., 2024; Chen et al., 2024). As shown in Figs. 6 and 7c, in scenario SSP1-2.6, SOC stocks are projected to be depleted by 2050 (−21%) and 2100 (−3.75%), with this depletion most severe in lowlands and uplands (slope lands). A significant increase in R95p, R99p, or CWD may increase soil erosion, leading to major losses in SOC stocks. Intense rainfall events may also cause topsoil erosion and the leaching of dissolved organic carbon, and episodic carbon export may exceed respiratory losses (Olaya-Abril et al., 2017; Rillig et al., 2021). In certain forested areas, localized SOC gains are predicted, even for the low emission projections, which may be related to spatially uneven warming and rainfall, leading to enhanced vegetation productivity and underground carbon input (Fig. 6; Guo et al., 2019). These findings are consistent with the region-specific SOC responses to temperature and precipitation anomalies in previous modeling studies (Wang et al., 2023).

In scenario SSP2-4.5, which involves CO_2 emissions that approach the current levels until the mid-century time point before declining but do not reach net zero by 2100, the SOC stocks are predicted to be controlled by R95p, R99p, and TXx (not statistically significant). For all study areas, slight warming and increased extreme rainfall events (smallest increase among all emission scenarios) will facilitate vegetation growth, which will in turn increase SOC stocks. Despite these findings, losses in SOC stocks are still predicted to occur in certain upland regions (Fig. 6) as a result of erosion events caused by increases in CWD, particularly in the LRW. Overall, these results underscore the importance of drainage devices and specific agricultural management practices in uplands (Vereecken et al., 2022; Wang et al., 2023).

In contrast to previous findings, our results indicate that SOC stocks will likely increase by an average of 45.4% to 58.3% in the study area. They will even exceed 200% in certain forested areas in scenario SSP5-8.5 (Fig. 7c). Under this scenario, TNx and TXx are crucial

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

factors influencing SOC stocks, particularly due to increased net primary productivity driven by warming temperatures, CO_2 fertilization, and extended growing seasons that promote vegetation growth. (Elbasiouny et al., 2022). When the amount of biomass, including wood debris and roots, increases, more biomass can be transformed into stable SOC for storage. This finding is consistent with the microbial efficiency matrix stabilization framework, which treats root-derived microbial residues as the main precursors to mineral-associated organic matter, a stable SOC pool (Cotrufo et al., 2013; Sokol and Bradford, 2019). In the high-emission scenario, severe warming and prolonged droughts in forested areas may result in wildfires or drought-induced dieback, which may reverse previous carbon gains in cases of ecological equilibrium within forests. Although an increase in SOC stocks is projected for scenario SSP5-8.5, some uplands will experience clear losses in SOC stocks. This effect indicates the vulnerability of uplands to erosion caused by runoff under extreme precipitation, thereby threatening long-term carbon retention. In forested areas and some uplands, significant increases in CDD are projected for scenario SSP5-8.5. However, the increases in CDD might decrease vegetation growth and soil moisture, which in turn will lead to less organic input and greater carbon losses. This condition is common in uplands, not forested areas, because the effects of temperature extremes on SOC stocks may be weakened by topography. These findings were confirmed by our PLS-PM analysis (Fig. 9). For scenario SSP5-8.5, the interaction between the Birch effect and erosion may be the reason underlying the losses predicted in SOC stocks in uplands (Birch, 1958; Schimel et al., 2007). In terms of strategies for adaptation in scenario SSP5-8.5, firebreak corridors or buffer zones in forested areas and drainage constructions in uplands should be prioritized.

In addition to climatic factors, land type plays a major role in SOC stock responses. In this study, forested areas were found to have higher levels of SOC and to be more sensitive to climate change compared with other land types. The varied responses across different land types emphasize the need to include topography, climate, and land management practices in SOC stock models and the importance of developing carbon mitigation strategies (IPCC, 2019).

However, several studies have highlighted that SSP5-8.5 is increasingly regarded as an implausible scenario for future climate projections and the likelihood of such a trajectory materializing could be negligible (Pielke & Ritchie, 2021; Burgess et al., 2021). Originally designed as a high-end "stress test" pathway, SSP5-8.5 assumes exceptionally high fossil-fuel use, rapid population growth, and minimal mitigation—conditions that diverge significantly from current global trends in energy transition, technology adoption, and policy

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

implementation. Therefore, in this study, SSP5-8.5 is included only as a computational benchmark to illustrate the response of soil carbon dynamics under an extreme forcing scenario, rather than as a realistic projection of the future.

5. Conclusions

Through DSM, this study established effective models for predicting SOC stock, achieving an R^2 range of 0.43–0.50. It also highlighted key environmental covariates, such as topography, climate, remote sensing parameters, and the prediction interval maps for identifying areas not covered in the sampling distribution. This study demonstrated that projected topsoil SOC stocks exhibit substantial spatio-temporal variability across emission scenarios, with clear sensitivity to landscape type and climate extremes. In the scenario of severe emissions, the sensitivity of SOC dynamics to extreme climate events was found to be high. Land type was also found to have a key influence on SOC stocks. These effects pose both location- and time-specific challenges for SOC management in studies on mid- to late-century time points. In the low-emission scenario (scenario SSP1-2.6), extreme rainfall events are predicted to induce a significant reduction in SOC stocks though erosion in upland areas. However, in the moderate- and high-emission scenarios (scenarios SSP2-4.5 and SSP5-8.5), warming (TNx and TXx) and extreme rainfall events (R95p and R99p) may simultaneously increase biomass input and increase soil erosion risks. These results indicate that SOC management strategies should be highly specific to the site and time. In both the ZRW and LRW, even though the SOC stock dynamics in forested areas are likely to be affected by extreme rainfall events, heat waves, and prolonged droughts, future mitigation strategies should focus on reducing warming and preventing wildfires. Adaptive strategies such as the planting of heat-tolerant tree species may also be necessary.

In upland areas in both the ZRW and LRW, SOC stock changes are predicted to be mainly driven by R95p, R99p, and CWD. Significant SOC losses will occur in certain upland areas for all emission scenarios. Therefore, management strategies should emphasize soil and water conservation to ensure that excess rainfall can be infiltrated into the soil without triggering erosion. These strategies should include the implementation of eco-engineering techniques on slope lands, maintaining vegetation cover and soil permeability, and establishing effective drainage systems. Overall, clarifying the interactions between climatic extremes, land types, and SOC stocks to develop site-specific management practices is key to enhancing soil's resilience and ensuring that SOC stocks continue to service ecosystems despite climate change.

Funding

This work was supported by the National Science and Technology Council of Taiwan (NSTC 113-2321-B-005-004 and NSTC 112-2313-B-005-059-MY2).

CRediT authorship contribution statement

Shih-Hao Jien: Writing—Review & Editing, Writing—Original Draft, and Data Curation. **Chien-Hui Syu and Chun-Chien Yen:** Methodology, Investigation, and Conceptualization. **Selly Maisyarah and Bo-Jiun Yang:** Visualization and Methodology. **Yu-Min Tzou:** Writing—Review & Editing.

Declaration of competing interests

The authors confirm that they have no financial or personal relationships that could have appeared to influence the research presented in this paper.

Data availability

The data used or analyzed in this study are available from the corresponding author upon reasonable request.

References

Adhikari, K., Hartemink, A. E., Minasny, B., Bou Kheir, R., Greven, M. B., and Greve, M. H.: Digital mapping of soil organic carbon contents and stocks in Denmark, PLoS ONE 9, e105519, https://doi.org/10.1371/journal.pone.0105519, 2014.

Akpa, S. I., Odeh, I. O., Bishop, T. F. A., Hartemink, A. E., and Amapu, I. Y.: Total soil organic carbon and carbon sequestration potential in Nigeria, Geoderma, 271, 202–215, https://doi.org/10.1016/j.geoderma.2016.02.021, 2016.

Birch, H. F.: The effect of soil drying on humus decomposition and nitrogen availability, Plant and Soil, 10, 9–31, https://doi.org/10.1007/BF01343734, 1958.

Blake, G. R., and Hartge, K. H.: Bulk density, in: Klute, A. (Ed.), Methods of soil analysis. Part 1. Physical and Mineralogical Methods, 2nd ed. ASA, SSSA, Madison, WI, 363–375, https://doi.org/10.2136/sssabookser5.1.2ed.c13, 1986.

Burgess, G. M., Ritchie, J., Shapland, J., and Pielke Jr. Roger,: IPCC baseline scenarios have over-projected CO_2 emissions and economic growth, Environmental Research Letter, 16, 014016, https://doi.org/10.1088/1748-9326/abcdd2, 2021.

Chalchissa, F. B., and Kuris, B. K.: Modeling the impacts of extreme climate scenarios on soil acidity (pH and exchangeable aluminum) in Abbay River Basin, Ethiopia, Heliyon, 10, 12, e33448, https://doi.org/10.1016/j.heliyon.2024.e33448, 2024.

Chalchissa, F. B., Diga, G. M., Feyisa, G. L., and Tolossa, A.R.: Impacts of extreme agroclimatic indicators on the performance of coffee *(Coffea arabica L.)* aboveground biomass in Jimma Zone, Ethiopia, Heliyon, 8, 8, e10136, https://doi.org/10.1016/j.heliyon.2022.e10136, 2022.

Chen, F., Wang, S. J., Dong, Q. J., Esper, J., Büntgen, U., Meko, D., Linderholm, H. W., Wang, T., Yue, W. P., Zhao, X. E., Hadad, M., González-Reyes, Á., and Chen, F.H.: Role of Pacific Ocean climate in regulating runoff in the source areas of water transfer projects on the Pacific Rim. npj Clim, Atmos. Sci., 7, 153, https://doi.org/10.1038/s41612-024-00706-1, 2024.

Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Denef, K., and Paul, E.: The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter?, Glob. Chan. Bio., 19, 988–995, https://doi.org/10.1111/gcb.12113, 2013.

Dialynas, Y. G., Bastola, S., Bras, R. L., Billings, S. A., Markewitz, D., and Richter, D. d B.: Topographic variability and the influence of soil erosion on the carbon cycle, Glob. Biogeochem. Cycles, 30, 644–660, https://doi.org/10.1002/2015GB005302, 2016.

Dorji, T., Odeh, I. O., Field, D. J., and Baillie, L. C.: Digital soil mapping of soil organic carbon stocks under different land use and land cover types in montane ecosystems, Eastern Himalayas, For. Eco. and Manag., 318, 91–102, https://doi.org/10.1016/j.foreco.2014.01.003, 2014.

Edmondson, J. L., Davies, Z. G., McCormack, S. A., Gaston, K. J., and Leake, J.R.: Land–cover effects on soil organic carbon stocks in a European city, Sci. Total Environ, 472, 444–453, https://doi.org/10.1016/j.scitotenv.2013.11.025, 2014.

Elbasiouny, H., El-Ramady, H., Elbehiry, F., Rajput, V. D., Minkina, T., and Mandzhieva, S.: Plant Nutrition under Climate Change and Soil Carbon Sequestration, Sustainability, 14, 914, https://doi.org/10.3390/su14020914, 2022.

Grace, J.: Understanding and managing the global carbon cycle, J. Ecol., 92, 189–202, https://doi.org/10.1111/j.0022-0477.2004.00874.x, 2004.

Gray, J. M., and Bishop, T. F. A.: Change in soil organic carbon stocks under 12 climate change projections over New South Wales, Australia, Soil Sci. Soc. Am. J., 80, 1296–1307, https://doi.org/10.2136/sssaj2016.02.0038, 2016.

Gray, J. M., Bishop, T. F. A., and Wilson, B. R.: Factors controlling soil organic carbon stocks with depth in eastern Australia, Soil Sci. Soc. Am. J., 79, 1741–1751, https://doi.org/10.2136/sssaj2015.06.0224, 2015.

Grunwald, S.: Multi–criteria characterization of recent digital soil mapping and modeling approaches, Geoderma, 152, 195–207, https://doi.org/10.1016/j.geoderma.2009.06.003, 2009.

Guo, L., Shi, T. Z., Linderman, M., Chen, Y. Y., Zhang, H. T., and Fu, P.: Exploring the Influence of Spatial Resolution on the Digital Mapping of Soil Organic Carbon by Airborne Hyperspectral VNIR Imaging, Remote Sens, 11, 1032, https://doi.org/10.3390/rs11091032, 2019.

Guo, P. T., Li, M. F., Luo, W., Tang, Q. F., Liu, Z. W., and Lin, Z. M.: Digital mapping of soil organic matter for rubber plantation at regional scale: An application of random forest plus residuals kriging approach, Geoderma, 237, 49–59, https://doi.org/10.1016/j.geoderma.2014.08.009, 2015.

IPCC: Climate Change and Land: Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems, https://doi.org/10.1017/9781009157988, 2019.

Jenny, H.: Factors of soil formation: a system of quantitative pedology, McGraw Hill, New York, 1941.

Jien, S. H., Minasny, B., Yang, B. J., Liu, Y. T., Yen, C. C., Ocba, M. A., Zhang, Y. T., and Syu, C. H.: Enhancing Soil Carbon Storage: Developing high-resolution maps of topsoil organic carbon sequestration potential in Taiwan, Geoderma, 459, 117369, https://doi.org/10.1016/j.geoderma.2025.117369, 2025.

Johnston, C. A., Groffman, P., Breshears, D. D., Cardon, Z. G., Currie, W., Emanuel, W., Gaudinski, J., Jackson, R. B., Lajtha, K., Nadelhoffer, K., Jr., D. N., and Post, W. M.: Carbon cycling in soil. Front, Ecol. Environ, 2, 522–528, https://doi.org/10.1890/1540-9295(2004)002[0522:CCIS]2.0.CO;2, 2004.

Keskin, H., and Grunwald, S.: Regression kriging as a workhorse in the digital soil mapper's toolbox, Geoderma, 326, 22–41, https://doi.org/10.1016/j.geoderma.2018.04.004, 2018.

Lacoste, M., Minasny, B., McBratney, A., Michot, D., Viaud, V., and Walter, C.: High resolution 3D mapping of soil organic carbon in a heterogeneous agricultural landscape, Geoderma, 213, 296–311, https://doi.org/10.1016/j.geoderma.2013.07.002, 2014.

Lamichhane, S., Kumar, L., and Wilson, B.: Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review, Geoderma, 352, 395–413, https://doi.org/10.1016/j.geoderma.2019.05.031, 2019.

Li, H. W., Wu, Y. P., Chen, J., Zhao, F., Wang, F., Sun, Y. H., Zhang, G. C., and Qiu, L. J.: Responses of soil organic carbon to climate change in the Qilian Mountains and its future projection, J. Hydrol, 596, 126110, https://doi.org/10.1016/j.jhydrol.2021.126110, 2021.

Ma, Y. X., Minasny, B., and Wu, C.: Mapping key soil properties to support agricultural production in Eastern China, Geoderma Regional, 10, 144–153, https://doi.org/10.1016/j.geodrs.2017.06.002, 2017.

Malone, B. P., Minasny, B., POdgers, N., and McBratney, A. B.: Using model averaging to combine soil property rasters from legacy soil maps and from point data, Geoderma, 232, 34–44, https://doi.org/10.1016/j.geoderma.2014.04.033, 2014.

McBratney, A. B., Santos, M. M., and Minasny, B.: On digital soil mapping, Geoderma, 117, 3–52, https://doi.org/10.1016/S0016-7061(03)00223-4, 2003.

Minasny, B., and McBratney, A.B.: Digital soil mapping: A brief history and some lessons, Geoderma, 264, 301–311, https://doi.org/10.1016/j.geoderma.2015.07.017, 2016.

Mishra, U., and Riley, W. J.: Scaling impacts on environmental controls and spatial heterogeneity of soil organic carbon stocks, Biogeosciences, 12, 3993–4004, https://doi.org/10.5194/bg-12-3993-2015, 2015.

Nelson, D. A., and Sommers, L.E.: Total carbon, organic carbon, and organic matter, in: Sparks, D.L. (Ed.), Methods of soil analysis. Part 3, Chemical Methods. ASA, SSSA, Madison, WI, 961–1010, https://doi.org/10.2136/sssabookser5.3.c34, 1996.

Olaya-Abril, A., Parras-Alcántara, L., Lozano-García, B., and Obregón-Romero, R.: Soil organic carbon distribution in Mediterranean areas under a climate change scenario via multiple linear regression analysis, Sci. Total Enviro., 592, 134–143, https://doi.org/10.1016/j.scitotenv.2017.03.021, 2017.

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

Pielke Jr. Roger, and Ritchie J.: Distorting the view of our climate future: The misuse and abuse of climate pathways and scenarios, Energy Research & Social Science 72, 101890, https://doi.org/10.1016/j.erss.2020.101890. 2021.

Poggio, L., Gimona, A., Brown, I., and Castellazzi, M.: Soil available water capacity interpolation and spatial uncertainty modelling at multiple geographical extents, Geoderma, 160, 175–188, https://doi.org/10.1016/j.geoderma.2010.09.015, 2010.

Quinlan, J.R.: Learning with Continuous Classes, in: Proceedings of Australian Joint Conference on Artificial Intelligence, Hobart, 1992.

Rial, M., Cortizas, A. M., and Rodríguez-Lado, L.: Understanding the spatial distribution of factors controlling topsoil organic carbon content in European soils, Sci. Total Enviro., 609, 1411–1422, https://doi.org/10.1016/j.scitotenv.2017.08.012, 2017.

Rillig, M. C., Leifheit, E., and Lehmann, J.: Microplastic effects on carbon cycling processes in soils, PLoS Biol., 19, e3001130, https://doi.org/10.1371/journal.pbio.3001130, 2021.

Rudiyanto, Minasny, B., Setiawan, B. I., Saptomo, S. K., and McBratney, A. B.: Open digital mapping as a cost-effective method for mapping peat thickness and assessing the carbon stock of tropical peatlands, Geoderma, 313, 25–40, https://doi.org/10.1016/j.geoderma.2017.10.018, 2018.

Sarstedt, M., Ringle, C. M., Smith, D., Reams, R., and Hair, J. F.: Partial least squares structural equation modeling (PLS-SEM): A useful tool for family business researchers, J. Fam. Bus. Strategy, 5, 105–115, https://doi.org/10.1016/j.jfbs.2014.01.002, 2014.

Schimel, J. P., Balser, T. C., and Wallenstein, M.: Microbial stress-response physiology and its implications for ecosystem function, Ecology, 88, 1386–1394, https://doi.org/10.1890/06-0219, 2007.

Shaik, A. B., and Srinivasan, S.: A Brief Survey on Random Forest Ensembles in Classification Model, in: Bhattacharyya, S., Hassanien, A., Gupta, D., Khanna, A., Pan, I. (Eds.), International Conference on Innovative Computing and Communications. Lecture Notes in Networks and Systems, 56. Springer, Singapore, 253–260, 2019.

Siewert, M. B.: High-resolution digital mapping of soil organic carbon in permafrost terrain using machine learning: a case study in a sub-Arctic peatland environment, Biogeosci., 15, 1663–1682, https://doi.org/10.5194/bg-15-1663-2018, 2018.

Singh, B. P., Setia, R., Wiesmeier, M., and Kunhikrishnan, A.: Chapter 7 - Agricultural management practices and soil organic carbon storage, Soil carbon storage, 207–244, https://doi.org/10.1016/B978-0-12-812766-7.00007-X, 2018.

Sokol, N. W., and Bradford, M. A.: Microbial formation of stable soil carbon is more efficient from belowground than aboveground input, Nat. Geosci., 12, 46–53, https://doi.org/10.1038/s41561-018-0258-6, 2019.

Tsui, C. C., Liu, X. N., Guo, H. Y., and Chen, Z.S.: Effect of Sampling Density on Estimation of Regional Soil Organic Carbon Stock for Rural Soils in Taiwan, Geospatial Technology–Environmental and Social Applications, 35–53, https://doi.org/10.5772/64210, 2016.

Vaysse, K., and Lagacherie, P.: Evaluating digital soil mapping approaches for mapping GlobalSoilMap soil properties from legacy data in Languedoc-Roussillon (France), Geoderma Reg., 4, 20–30, https://doi.org/10.1016/j.geodrs.2014.11.003, 2015.

Vereecken, H., et al.: Soil hydrology in the Earth system, Nat. Rev. Earth Environ., 3, 573–587, https://doi.org/10.1038/s43017-022-00324-6, 2022.

Viaud, V., Angers, D. A., and Walter, C.: Toward landscape-scale modeling of soil organic matter dynamics in agroecosystems, Soil Sci. Soc. Am. J., 74, 1847–1860, https://doi.org/10.2136/sssaj2009.0412, 2010.

Villanueva, R. A. M., and Chen, Z. J.: ggplot2: Elegant Graphics for Data Analysis (2nd ed.), Measurement-Interdisciplinary Research and Perspectives, 17, 160-167, https://doi.org/10.1080/15366367.2019.1565254, 2019.

Wang, B., Waters, C., Orgill, S., Gary, J., Cowie, A., Clark, A., and Liu, D. L.: High resolution mapping of soil organic carbon stocks using remote sensing variables in the semi-arid rangelands of eastern Australia, Sci. Total Enviro., 630, 367–378, https://doi.org/10.1016/j.scitotenv.2018.02.204, 2018.

Wang, M. M., Zhang, S., Guo, X. W., Xiao, L. J., Yang, Y. H., Luo, Y. Q., Mishra, U., and Luo, Z. K.: Responses of soil organic carbon to climate extremes under warming across global biomes, Nat. Clim. Change, 14, 98–105, https://doi.org/10.1038/s41558-023-01874-3, 2023.

Wei, Y. C., Wang, M. M., Viscarra Rossel, R. A., Chen, H., and Luo, Z. K.: Extreme Climate as the Primary Control of Global Soil Organic Carbon Across Spatial Scales, Global Biogeochem. Cycles, 38, e2024GB008277, https://doi.org/10.1029/2024GB008200, 2024.

Wiesmeier, M., Urbanski, L., Hobley, E., Lützow, M. V., Marin-Spiotta, E., Wesemael, B. V., Rabot, E., Ließ, M., Garcia-Franco, N., Wollschläger, U., Vogel, H. J., and Kögel-Knabner, I.: Soil organic carbon storage as a key function of soils–A review of drivers and indicators at various scales, Geoderma, 333, 149–162, https://doi.org/10.1016/j.geoderma.2018.07.026, 2019.

Yang, R. M., Zhang, G. L., Liu, F., Lu, Y. Y., Yang, F., Yang, F., Yang, M., Zhao, Y. G., and Li, D. C.: Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an alpine ecosystem, Ecol. Ind., 60, 870–878, https://doi.org/10.1016/j.ecolind.2015.08.036, 2016.

Zhang, G. Y., Luo, S., Jing, Z. W., Wei, S., and Ma, Y. H.: Evaluation and Forewarning Management of Regional Resources and Environment Carrying Capacity: A Case Study of Hefei City, Anhui Province, China, Sustainability, 12, 1637, https://doi.org/10.3390/su12041637, 2020.

Zhu, Q., and Lin, H. S.: Comparing ordinary kriging and regression kriging for Soil properties in contrasting landscapes, Pedosphere, 20, 594–606, https://doi.org/10.1016/S1002-0160(10)60049-5, 2010.

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

835

836 **Table 1.** Environmental covariates.

Type of data	Environmental covariates	Soil forming factor	Type
Remote sensing	Normalized difference vegetation index (NDVI)	o; t	Q
Digital elevation model	Elevation	r	Q
	Slope	r	Q
	Aspect	r	Q
	Terrain ruggedness index (TRI)	r	Q
	Topographic wetness index (TWI)	r	Q
	Terrain position index (TPI)	r	Q
	Multiresolution Index of Valley Bottom Flatness (MrVBF)	r	Q
	Multiresolution Ridge Top Flatness (MrRTF)	r	Q
	Stream power index (SPI)	r	Q
	Curvature	r	Q
	Flow accumulation	r	Q
	R-value		Q
	K-value		Q
Climate	Mean annual temperature (MAT)	c; t	Q
	Total annual precipitation (TAP)	c; t	Q
Land cover	Land cover	o; t	C
Soil	Soil Order	s	C

837 s: soil, r: relief, c: climate, t: time, o: organism, Q: quantitative, C: categorical.

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

Table 2. Data partitioning rules and 5th/95th percentiles of leave-one-out validation residuals.

Class	Conditions	Residual percentile	
		P_5	P_{95}
1	Soil Order in Other	-1.08	0.80
2	Topographic Position Index ≤ -10.8889	-1.99	0.95
3	MAT > 17.09029, Topographic Position Index > -10.8889, Soil Order in Inceptisol, Entisol, Alfisol, Spodosol, Ultisol	-0.85	0.70
4	MAT ≤ 17.09029	-0.67	0.80

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

Table 3. Variance and coefficient of variance (CV) of the spatiotemporal distribution of SOC stocks for various land uses in the three emission scenarios.

		2020				2050				2100			
		T	P	SL	F	T	P	SL	F	T	P	SL	F
	Variance	6.45	0.66	1.35	5.11								
	CV(%)	44.3	32.3	30.7	32.5								
SSP1	Variance					5.52	0.35	0.85	4.37	8.24	0.76	1.56	6.45
2.6	CV(%)					50.4	29.8	33.9	36.0	51.3	33.5	39.2	36.4
SSP2	Variance					10.1	1.14	2.25	7.95	21.4	1.74	2.53	18.7
4.5	CV(%)					48.5	35.0	37.5	34.9	59.7	35.2	40.4	43.7
SSP5	Variance					8.53	0.71	1.82	6.92	36.2	1.96	2.86	33.3
8.5	CV(%)					47.1	31.8	33.5	34.7	65.8	34.1	40.0	48.5

CV: coefficient of variance; T: total area; P: plain regions; SL: slope land regions; F: forest regions.

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

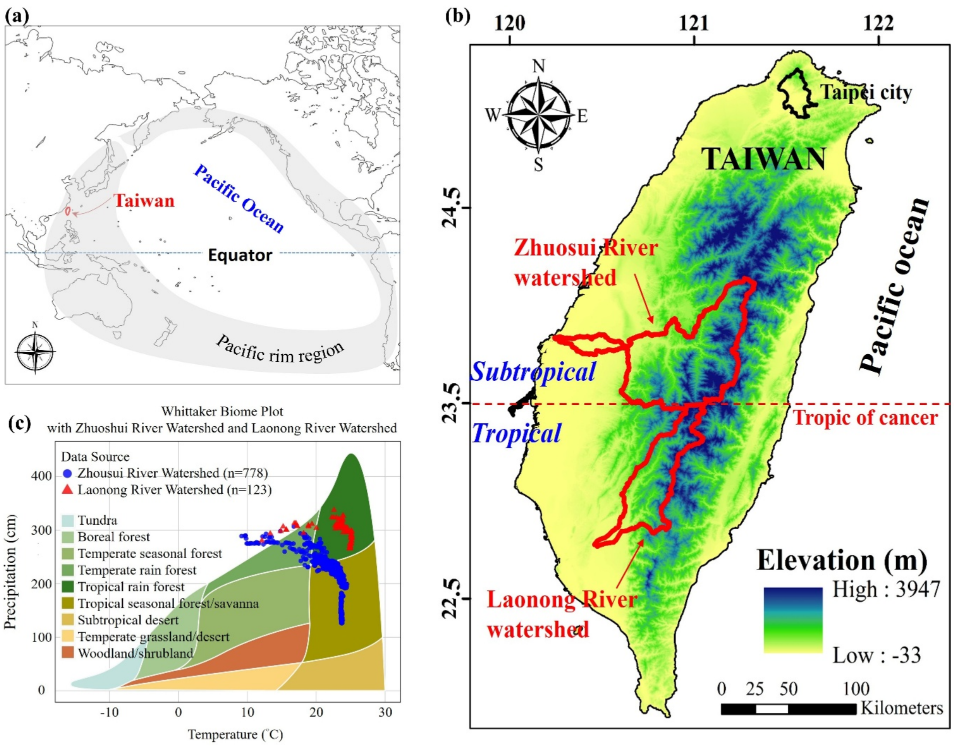
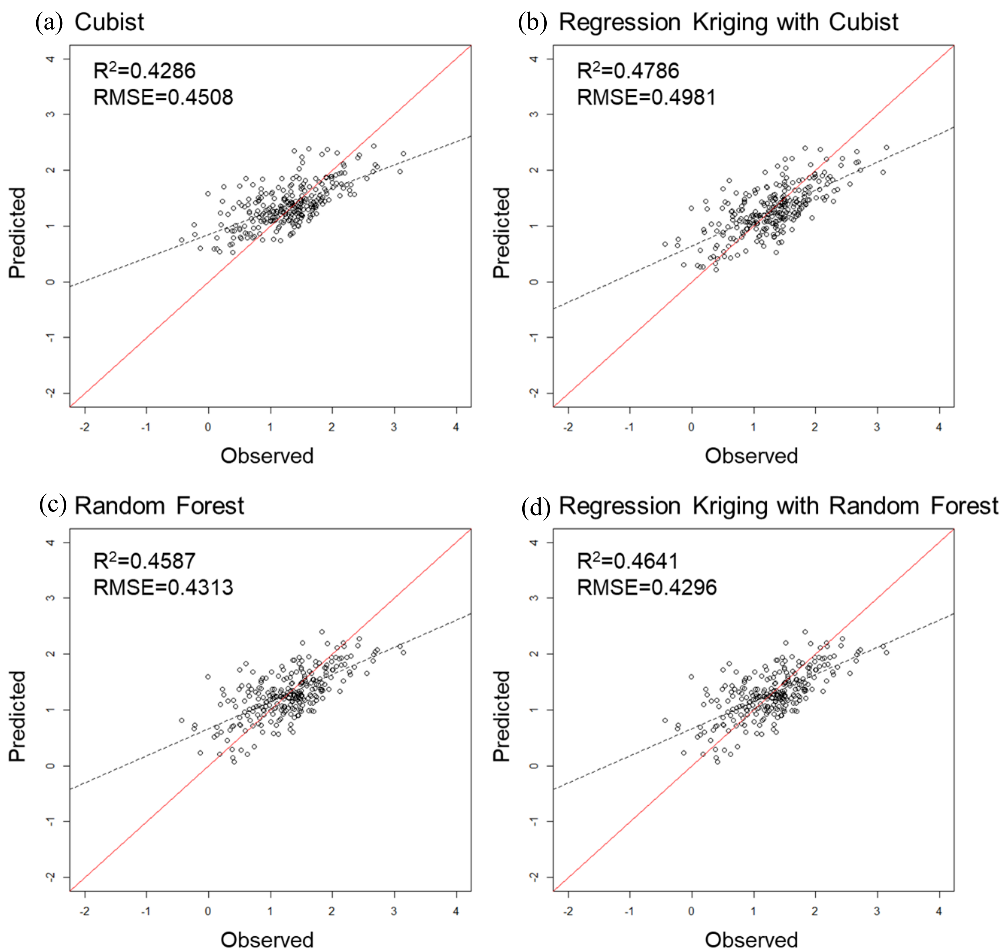


Fig. 1. Location of Taiwan in Pacific Ocean regions (a); location of Zhoushui River watershed (ZRW) and Laonong River watershed (LRW) in Taiwan; (c) Whittaker Biome Plot of ZRW and LRW.

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025



855 **Fig. 2.** Scatter plots of predicted versus observed soil organic carbon (SOC) stock in the topsoil
856 (0–30 cm) where predictions were obtained on the basis of validation data and by using the (a)
857 Cubist, (b) regression kriging (RK) with Cubist, (c) Random forest (RF), and (d) regression
858 kriging with RF models. The *x*-axis represents the observed values, and the *y*-axis represents
859 the predicted values. The solid line is the fitted line.

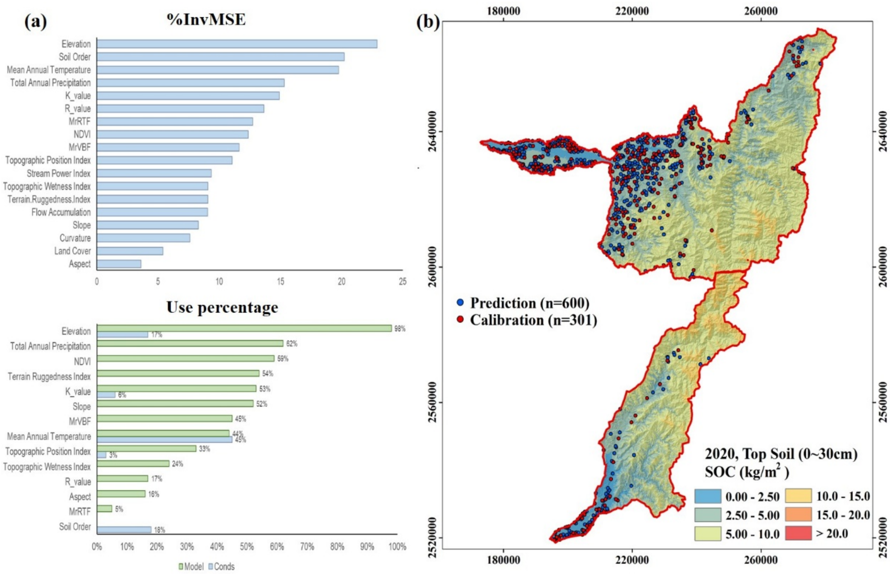


Fig. 3. Fig. 3. (a) Variable importance of Random Forest and Cubist models for SOC stock in surface soils (0–30 cm) and (b) predictive map of SOC stock in Zhuoshui River watershed and Laonong River watershed.

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

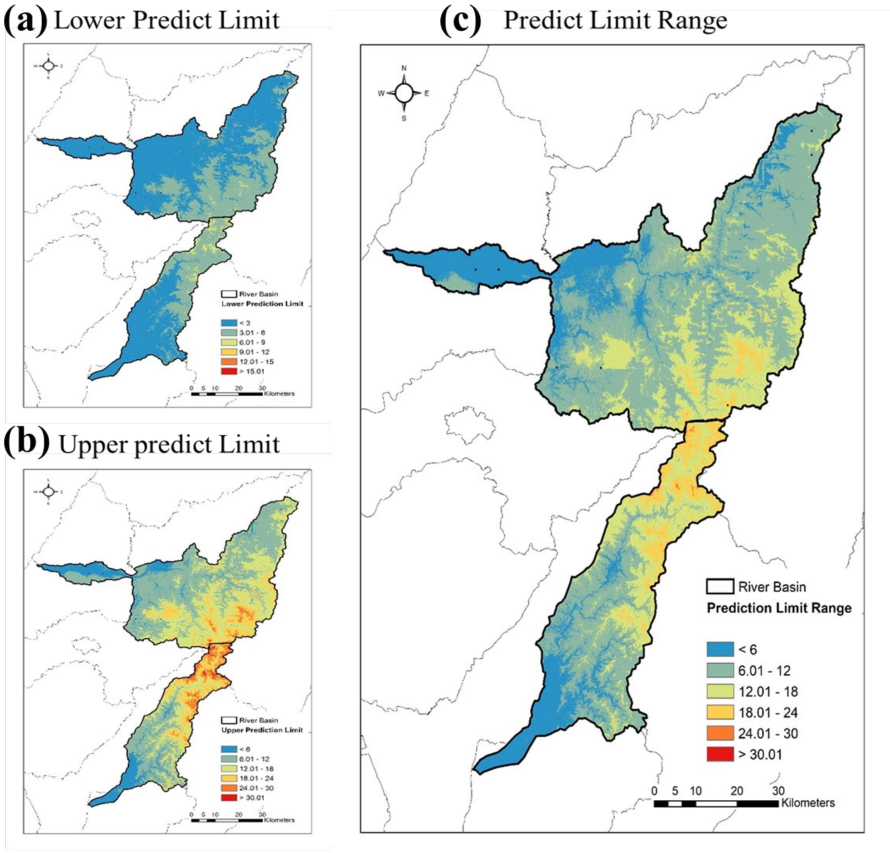


Fig. 4. Topsoil (0–30 cm) soil organic carbon (SOC) stock maps of the (a) 90% lower prediction limit, (b) 90% upper prediction limit, and (c) prediction limit range derived using bootstrapping.

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

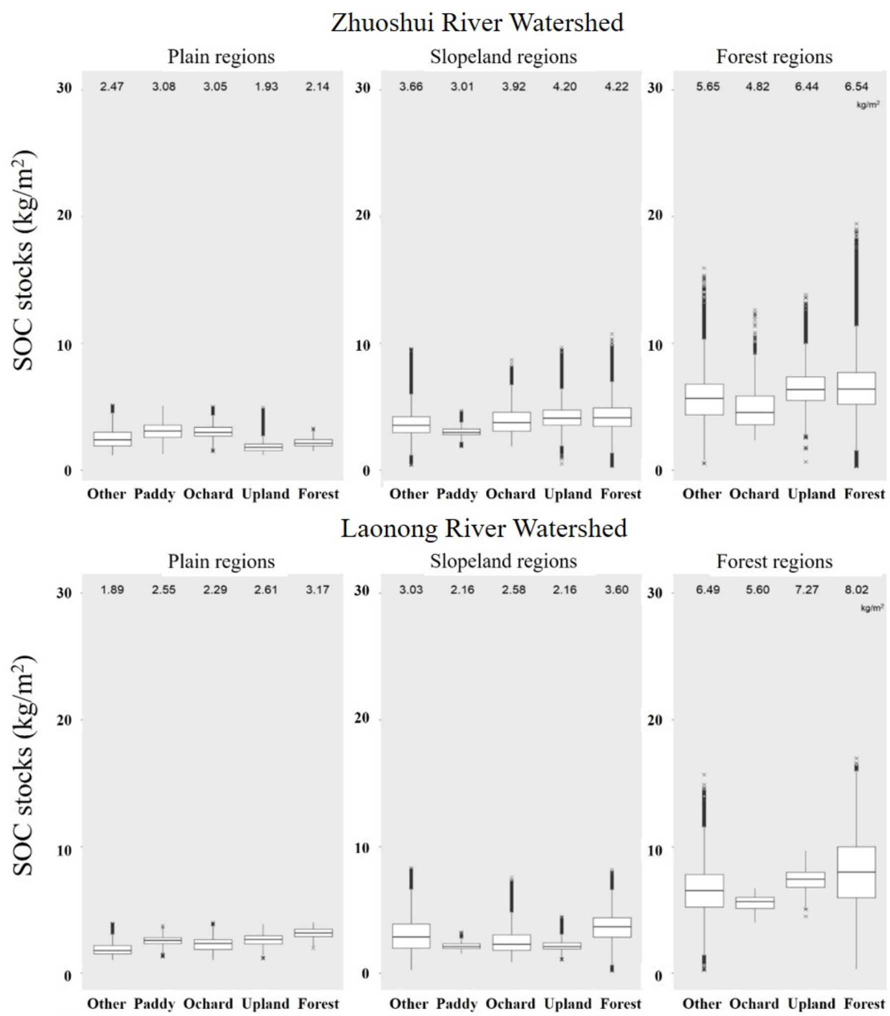


Fig. 5. Boxplots of topsoil (0–30 cm) soil organic carbon (SOC) stocks for various land cover: (left) plain regions (<100 m in elevation), (middle) slopeland regions (100–1000 m in elevation), and (right) forested regions (>1000 m in elevation) at Zhuoshui River watershed and Laonong River watershed.

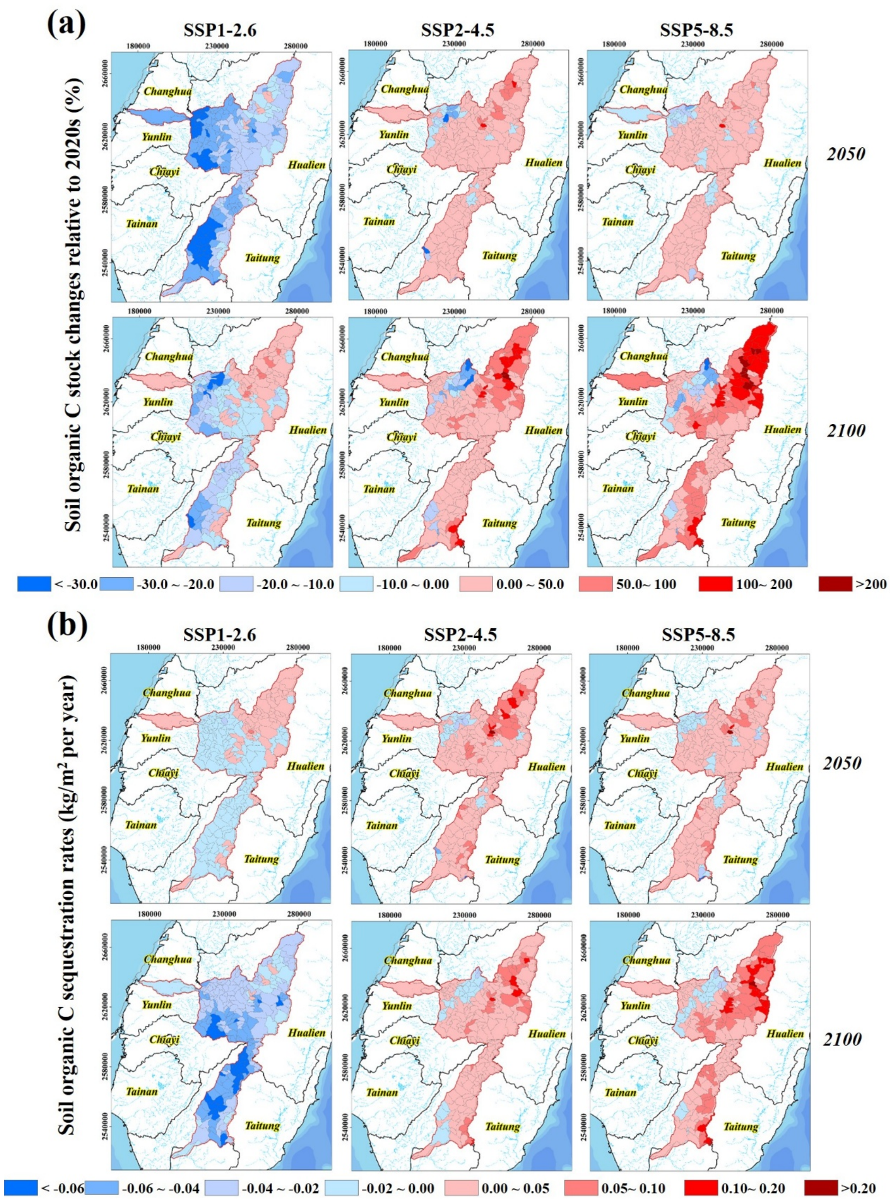


Fig. 6. Spatiotemporal predictions of (a) SOC stocks (kg m^{-2}) and (b) SOC sequestration rates (kg m^{-2} per year) relative to the 2020s under three emission scenarios. The mapping unit is sub-catchments in Taiwan.

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

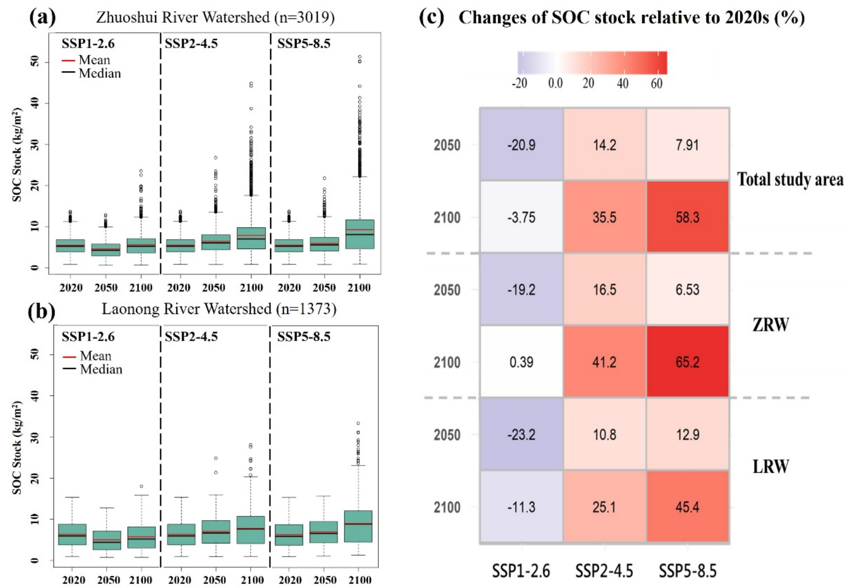


Fig. 7. Boxplots showing the temporal trends in predicted SOC stocks across three emission scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5) for 2020, 2050, and 2100 in (a) Zhousui River watershed; and (b) Laonong River watershed, and (c) Increase in the ratio of SOC stocks relative to the 2020s in the ZRW and LRW for the three emission scenarios for 2050 and 2100.

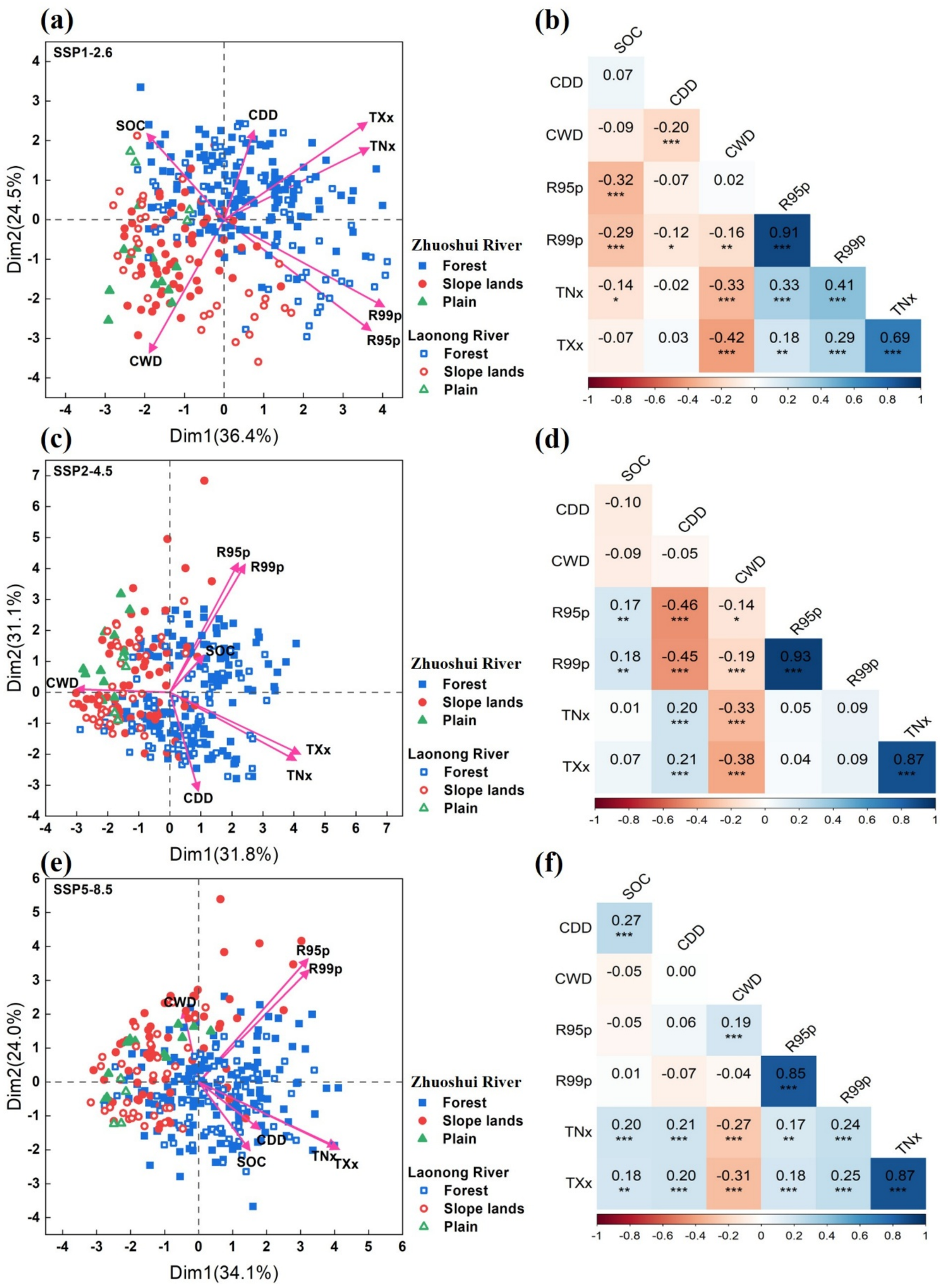


Fig. 8. Principal component analysis and Pearson's correlation coefficient of extreme climate indices and SOC stocks: (a, b) SSP1-2.6, (c, d) SSP2-4.5, and (e, f) SSP5-8.5.

https://doi.org/10.5194/egusphere-2025-4258
Preprint. Discussion started: 8 October 2025

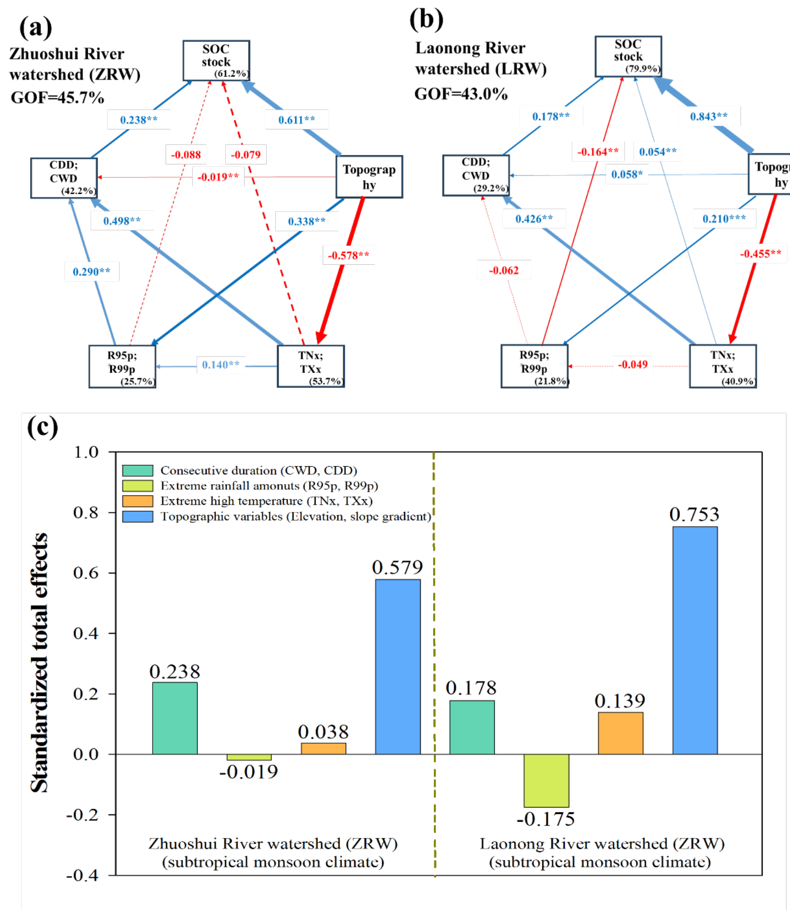


Fig. 9. Partial least squares path modeling (PLS-PM) analysis of the relationships among SOC stocks, consecutive durations of extreme climatic events (CDD and CWD), extreme rainfall amounts (R95p and R99p), extreme temperatures (TNx and TXx), and topographic variables (elevation and slope gradient) (a) Zhuoshui River watershed; (b) Laonong River watershed; (c) standardized total effects. Positive and negative effects are represented by blue and red arrows, respectively. Path coefficients that do not significantly differ from zero are depicted as gray dashed lines: $^{*}p < 0.05$ and $^{**}p < 0.01$. The percentages in the boxes represent the explanatory power of the variables. The goodness-of-fit was used to assess the model.