Environmental Science and Ecotechnology 23 (2025) 100482

Contents lists available at ScienceDirect

Environmental Science and Ecotechnology

journal homepage: www.journals.elsevier.com/environmental-science-and-ecotechnology/

Original Research

Hydrology, vegetation, and soil properties as key drivers of soil organic carbon in coastal wetlands: A high-resolution study

Mao Guo [a, b, e], Lin Yang [a, c, e, *], Lei Zhang [a], Feixue Shen [a], Michael E. Meadows [a, d], Chenghu Zhou [a, b, e]

[a] School of Geography and Ocean Science, Nanjing University, Nanjing, 210023, China
[b] Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, 511458, China
[c] Frontiers Science Center for Critical Earth Material Cycling, Nanjing University, Nanjing, 210023, China
[d] Department of Environmental & Geographical Science, University of Cape Town, Rondebosch, 7701, South Africa
[e] State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China

ARTICLE INFO

Article history:
Received 14 February 2024
Received in revised form
27 August 2024
Accepted 30 August 2024

Keywords:
Coastal wetlands
Hydrological factors
Soil carbon stock
Blue carbon

ABSTRACT

Coastal wetlands are important blue carbon ecosystems that play a significant role in the global carbon cycle. However, there is insufficient understanding of the variations in soil organic carbon (SOC) stocks and the mechanisms driving these ecosystems. Here we analyze a comprehensive multi-source dataset of SOC in topsoil (0–20 cm) and subsoil (20–100 cm) across 31 coastal wetlands in China to identify the factors influencing their distribution. Structural equation models (SEMs) reveal that hydrology has the greatest overall effect on SOC in both soil layers, followed by vegetation, soil properties, and climate. Notably, the mechanisms driving SOC density differ between the two layers. In topsoil, vegetation type and productivity directly impact carbon density as primary sources of carbon input, while hydrology, primarily through seawater salinity, exerts the largest indirect influence. Conversely, in subsoil, hydrology has the strongest direct effect on SOC, with seawater salinity also influencing SOC indirectly through soil and vegetation mediation. Soil properties, particularly pH, negatively affect carbon accumulation, while climate influences SOC indirectly via its effects on vegetation and soil, with a diminishing impact at greater depths. Using Random Forest, we generate high-resolution maps (90 m × 90 m) of topsoil and subsoil carbon density (R^2 of 0.53 and 0.62, respectively), providing the most detailed spatial distribution of SOC in Chinese coastal wetlands to date. Based on these maps, we estimate that SOC storage to a depth of 1 m in Chinese coastal wetlands totals 74.58 ± 3.85 Tg C, with subsoil carbon storage being 2.5 times greater than that in topsoil. These findings provide important insights into mechanism on driving spatial pattern of blue carbon and effective ways to assess carbon status on a national scale, thus contributing to the advancement of global blue carbon monitoring and management.

1. Introduction

Coastal wetlands, as crucial blue carbon ecosystems, store large amounts of long-term soil carbon due to their high above- and below-ground primary productivity and unique tidally inundated, depositional environment for carbon sequestration [1,2]. There is growing global interest in coastal wetlands as targets for mitigating global warming through preserving and restoring soil carbon to increase future carbon sequestration. Although there are numerous studies on blue carbon ecosystems at small spatial scales [3–5], the spatial distribution of coastal soil carbon storage and its underlying driving mechanisms is not well understood at the regional or national scale due to limited sample data, the difficulty of accurately determining coastal wetland extent, and the complex soil carbon development environment. Obtaining an accurate estimate of the spatial distribution of soil carbon in coastal wetlands and investigating its associated driving mechanisms is essential in evaluating the current soil carbon status and providing a benchmark for

* Corresponding author. School of Geography and Ocean Science, Nanjing University, Nanjing, 210023, China.
E-mail address: yanglin@nju.edu.cn (L. Yang).

https://doi.org/10.1016/j.ese.2024.100482

policymaking related to carbon sequestration.

The soil organic carbon (SOC) stock in coastal wetlands is spatially highly heterogeneous due to the complex mechanisms governing soil carbon accumulation and deposition [6,7]. Topsoil and subsoil often have different soil carbon dynamics mechanisms [8,9]. Previous research has revealed significant differences between topsoil and subsoil in terms of abundance and function in soil microbial communities [10], plant-nitrifier interactions [11], and fractions and sources of organic carbon [12]. Therefore, a comprehensive understanding of soil carbon dynamics in coastal wetlands must account for mechanisms in both topsoil and subsoil layers. Given the distinctive environment of the coastal zone at the land-sea interface, the status of soil carbon in coastal wetlands is jointly affected by multiple factors, including climate (temperature and rainfall) [4,13], vegetation (vegetation type, productivity, biological invasion) [4,14], soil physiochemical properties (soil bulk density [BD], pH, nitrogen) [3,15,16], geomorphology (elevation, slope) [17], and the unique coastal hydrological processes (tides, waves, and inland runoff) [18,19]. A great majority of studies have focused on one or a few of these factors, such as climate [13,20], vegetation [4,21], and soil properties [5,7]. However, the impact of coastal hydrological processes on soil carbon, which has been explored at smaller scales (i.e., individual estuaries or deltas [22,23]), remains unclear on a broader scale [24], especially for deep soil layers. Furthermore, limited attention has been paid to how these factors work together to affect soil carbon distribution across different soil depths [4,6]. In fact, coastal soil carbon exists within a complex ecosystem, where each factor not only directly regulates carbon dynamics but also indirectly influences it through convoluted interactions [14,16,22]. Therefore, assessing the direct and indirect effects of each factor and identifying the controlling factors in topsoil and subsoil is needed to more clearly understand driving mechanisms and more accurately map the spatial distribution of coastal soil carbon.

Several studies have attempted to map the spatial distribution of coastal soil carbon stock on national to global scales at a coarse spatial resolution. Commonly, the mean values of a limited number of samples are taken to represent the soil carbon level of a region or country — for example, in coastal China [25,26], coastal United States [27,28], Australian vegetated coastal ecosystems [29], and global blue carbon stocks [30]. Based on average SOC values, this approach underrepresents the spatial heterogeneity of coastal soil carbon, leading to estimates with high uncertainty. Hence, researchers have developed other prediction methods. Meng et al. [31] estimated the upper 0–1 m SOC density distribution using simple interpolation based on only 96 samples in China's coastal wetlands. Holmquist et al. [32] applied mean carbon density (27.0 kg C m^{-3}) to conservatively estimate the top meter soil carbon stock as 0.72 Pg C of the conterminous United States tidal wetlands. Wang et al. [33] adopted the nearest neighbor interpolation method to obtain carbon burial rate maps of global tidal wetlands based on 613 samples. Rovai et al. [34] utilized multiple regression to predict global mangrove SOC density using five environmental covariates at a resolution of 0.25°. However, coastal wetland soil carbon maps on a large scale with a fine spatial resolution are lacking, which is a constraint considering the complexity of soil carbon–driving factor relationships.

With an extensive coastline, China has several wetland types, including salt marshes, mangroves, tidal flats, and seagrass meadows. Their complex hydrological conditions and different vegetation types make China's coastal wetlands ideal for studying the mechanisms driving the spatial heterogeneity of coastal wetland SOC. Recently, several advances have been made in estimating the upper 0–1 m blue carbon stock in China, including the SOC stock of coastal wetlands [7,35]; mangroves, salt marshes, and seagrass meadows [25,31]; and tidal flats [36]. However, these estimates were obtained using the mean values of a relatively limited number of samples, leading to inconsistent conclusions. Furthermore, little is known about how hydrology (freshwater and marine) and other factors interactively impact the spatial distribution of China's coastal wetlands' SOC at different soil depths, and detailed spatial distribution maps of topsoil and subsoil carbon stock have, until now, not been available.

The objectives of this study are as follows: (1) explore the direct and indirect impacts of environmental factors driving the distribution of topsoil and subsoil organic carbon in the Chinese mainland's coastal wetlands, (2) map the topsoil and subsoil organic carbon stock of coastal wetlands in the Chinese mainland at a high spatial resolution, and (3) estimate total SOC stock based on the maps. To achieve these goals, we collated 408 samples for topsoil (0–20 cm) and 304 for subsoil (20–100 cm) SOC density from published literature and publicly available datasets and evaluated the effects of climate, hydrology, soil, vegetation, and topography on SOC density with structural equation models. Structural equation modeling (SEM) is a powerful causal approach for analyzing complex relationships among multiple variables. Thus, it can be a potentially effective method for identifying how the environmental factors interactively influence SOC in coastal wetlands and understanding the soil carbon distribution mechanisms in topsoil and subsoil layers. We then generate topsoil and subsoil organic carbon density maps at a 90m resolution using machine learning methods. The results strengthen understanding the spatial distribution and underlying mechanisms of coastal wetland soil carbon at different depths and provide a benchmark for coastal wetland management relating to carbon sequestration.

2. Materials and methods

2.1. Study area

The study area incorporates the coastal wetlands, including mangrove forests, tidal flats, and salt marshes, in the Chinese mainland that extends from the Yalu River estuary to the Beilun River estuary and covers temperate, subtropical, and tropical zones. For this study, coastal wetlands are defined as the area between mean low tide and mean high tide but excluding artificial and rocky coasts. Underwater seagrass beds were not included in the analysis due to the difficulty of observation using remote sensing.

To predict the spatial distribution of SOC across the study area, it was necessary to delineate the extent of coastal wetlands. We used time series Landsat Collection 2 Level-2 (thematic mapper, TM) surface reflectance images to map the coastal wetland in 2010, corresponding to the soil sampling time. In this study, coastal wetlands, distinguished from inland wetlands based on periodical tide information, are classified into three types based on phenological information of tidal vegetation: unvegetated tidal flats, saltmarshes, and mangroves [37]. Tides, here, refer to those observed using Landsat images (Supplementary Material Fig. S1) and not astronomical tides [38]. For classification, we generated a total of 3368 samples (mangrove: 416, salt marshes: 956, tidal flat: 869, and others: 1127) as ground truth and divided them into a training set (70%) and a validation set (30%) using a stratified random sampling strategy. The training set was used as the random forest (RF) input to construct a classification model. Post-classification processing was conducted to generate the final wetland classification map. All methods were performed using the Google Earth Engine platform. The specific methodology and results (Supplementary Material Figs. S2–S4, and Table S1) are presented in the Supplementary Materials.

2.2. Soil sample compilation in China's coastal wetlands

We collated samples from three public soil datasets and published meta-analysis literature (Table 1) to generate the most exhaustive sample set currently available for the Chinese mainland's coastal wetlands. First, we screened 75 sampling points from public datasets, including 42 points from the Coastal Carbon Research Coordination Network dataset [39], 11 coastal soil samples from the World Soil Information Service (WoSIS) profile dataset [40], and 22 from the Soil Series dataset of China [41], using the spatial extent of coastal wetlands we generated in the above section.

We then obtained 333 samples mainly from the following four papers [7,25,31,35] — which gathered sampling points in coastal areas from published sources, some of which also collected samples in the field [7,25]. These samples were originally derived from journal papers and dissertations published in the Web of Science database (http://apps.webofknowledge.com) or China National Knowledge Infrastructure (http://www.cnki.net). As the original datasets' research purposes or data collection criteria vary, we conducted a thorough quality control process. We carefully reviewed the methodologies and sample descriptions to ensure that all data met our inclusion criteria. We collated only those sampling points obtained in the field and recorded them with accurate geographical coordinates. Sampling points from a less than 20 cm soil depth and duplicate sample points were excluded. Data are restricted to samples collected between 2005 and 2017 and represent changes in soil carbon at the decadal scale. The study does not consider longer-term soil carbon dynamics.

All sample SOC density data were standardized to 0–20 and 20–100 cm. For samples with a profile depth of less than 100 cm, only an SOC density value of 0–20 cm was used in the subsequent calculation and analysis. Different standardization procedures were conducted according to the following cases. Where sample SOC density values were provided for individual soil layers, standardization for the two depths considered here was computed as follows:

$$SOC_{\mathrm{d}0-20} = \frac{\sum_{i=1}^{N-1} SOC_{\mathrm{d}i} \times (D_i - D_{i-1}) + SOC_{\mathrm{d}N} \times (20 - D_{N-1})}{20} \quad (1)$$

$$SOC_{\mathrm{d}20-100} = \frac{SOC_{\mathrm{d}N} \times (D_N - 20) + \sum_{i=N+1}^{M-1} SOC_{\mathrm{d}i} \times D_i + SOC_{\mathrm{d}M} \times (100 - D_{M-1})}{100 - 20} \quad (2)$$

where $SOC_{\mathrm{d}i}$ represents the SOC density value (kg m^{-2}) at the i th soil layer. N refers to the number of soil layers used in target depth of topsoil (0–20 cm), and M means the number of soil layers used in target depth of subsoil (20–100 cm). Equation (1) applies when $N > 1$. If $N = 1$, that is the first layer is greater than or equal to 20 cm, then $SOC_{\mathrm{d}0-20} = SOC_{\mathrm{d}N}$. Equation (2) applies when $M - N > 1$. If $M - N = 1$, then $SOC_{\mathrm{d}20-100} = (SOC_{\mathrm{d}N} \times (D_N - 20) + SOC_{\mathrm{d}M} \times (100 - D_N))/(100 - 20)$. It is assumed that $N \geq 1$ & $M \geq 2$ by default according to the sample selection criteria.

Where SOC content and BD were recorded for each soil layer but not SOC density, we followed Xu et al.'s [42] approach to calculate the SOC_{d} (kg m^{-2}) at the depths of 0–20 cm and 20–100 cm:

$$SOC_{\mathrm{d}} = \sum_{i=1}^{n} \frac{SOC_{ci} \times BD_i \times D_i \times (1 - C_i)}{10} \quad (3)$$

where SOC_{ci} represents the SOC content (%) and BD_i represents the BD (g cm^{-3}) at the i th soil layer. C_i represents gravel content (%) at the i th soil layer that defaults to 0 where no data are provided.

Where the sample data provided only organic matter content but no organic carbon, SOC content was estimated using a conversion coefficient of 0.58 [43]. For samples lacking BD measurements, BD was estimated using the Pedotransfer Functions of Xu et al. [43].

Finally, 408 samples were selected for the topsoil (0–20 cm) and 304 for subsoil (20–100 cm) in the coastal wetlands of the Chinese mainland (Supplementary Material Fig. S5). The collected samples have good coverage along the entire coastal wetlands of the Chinese mainland (Supplementary Material Fig. S5).

2.3. Environmental factors

Five types of environmental factors potentially affecting SOC in coastal wetlands were selected and collated (Supplementary Material Table S2). Climate factors, including mean annual temperature (MAT) and mean annual precipitation (MAP), were derived from the 1 km monthly temperature and precipitation dataset of China covering the period 1980–2010 [44].

Hydrological factors, including sea surface salinity (from the global ocean surface salinity grid dataset with a 0.5° spatial resolution [45]), tidal range (from the National Marine Data Information Center, http://mds.nmdis.org.cn/pages/tidalCurrent.html), the overland runoff throughflow (g m^{-2} s^{-1}) per pixel (from Averaged, Single-Level, Assimilation, Land Surface Diagnostics V5.12.4, https://disc.gsfc.nasa.gov/), and significant wave height (SWH; from ERA5 hourly data on single levels in 2010, https://cds.climate.copernicus.eu), were selected to represent coastal hydrodynamic conditions.

Vegetation factors include normalized difference vegetation index (NDVI; http://www.gscloud.cn), net primary productivity (NPP; http://www.resdc.cn), and wetland type. Wetland type (WetType), classified as different vegetation types, including mangroves, saltmarshes, and tidal flats (unvegetated), was derived from our results of coastal wetland extraction based on remote sensing imagery (see Section 2 of Supplementary Materials).

Soil properties, including soil pH, cation exchange capacity (CEC), BD, total nitrogen content, sand, silt, clay, and coarse fragments volumetric fraction (CFVO), were aggregated from Soilgrids 2.0 [46].Topographical factors include elevation (from SRTM 90m Digital Elevation Database), slope (calculated based on elevation using the slope tool in ArcGIS 10.2), and distance to the coastline (Dist2Coast). Dist2Coast was calculated as the distance to global

Table 1
Sources and numbers of coastal samples.

Dataset	Number of samples collected	Sources
Public sample datasets	42	Coastal Carbon Research Coordination Network (CCRCN) dataset (https://serc.si.edu/coastalcarbon) [39]
	11	World Soil Information Service (WoSIS) soil profile database [40]
	22	Soil Series of China [41]
Literature	333	Published references [7,25,31,35]
Total	408	-

self-consistent, hierarchical, high-resolution shorelines (from https://docs.gmt-china.org/5.4/dataset/gshhg/).

All data were resampled to 90 m using a bilinear interpolation approach in the ArcGIS tool [47].

2.4. Quantifying the effects of driving factors on coastal wetlands' SOC using SEM

Structural equation modeling was utilized to identify the pathways and relative influence of the selected environmental factors and SOC and quantify the direct and indirect impacts of each factor. SEM has emerged as a powerful causality analysis method in various research fields, including soil science [4], vegetation science [14], and microbiology science [48]. SEM enables researchers to test theoretical models with multiple variables and pathways and estimate their relationships' relative strength and direction [49,50].

Structural equation modeling has several advantages for pathway analysis. First, a variable can simultaneously serve as both a dependent and independent variable. Second, SEM can handle observed variables directly or composite variables that cannot be measured directly but are indicated by several observed variables. For example, climate can be considered a composite variable and indicated by measured variables, including temperature and precipitation. Third, SEM quantitatively identifies the direct and indirect effects of a variable. The direct effect refers to an environmental variable directly impacting the target variable, while the indirect effect manifests through a mediation variable. For example, a hydrological factor, such as tidal range, affects soil carbon by influencing soil properties (e.g., soil water content and salinity).

The application of SEM involves the following steps to establish the network of pathways. First, we hypothesized a conceptual structural equation model that considers all the selected factors (i.e., climate, hydrology, vegetation, soil, and topography) to have a direct impact on SOC density, while climate, hydrology, and topography were assumed to indirectly impact SOC density through soil properties or plant growth (Supplementary Material Fig. S6). We then parameterized the initial SEM model with normalized data using a "scale" function in R-4.1.0. We assessed its goodness of fit using Fisher's C, p-value, and degree of freedom (df) [51]. Generally, a well-fitting SEM model should satisfy the following conditions: $p > 0.05$ and Fisher's C/degree of freedom ≤ 2. If the model is not well fitted according to the above indices, it is modified by removing the insignificant pathways or variables. Finally, we generated the optimal model that achieved the highest coefficient of determination (R^2) value among the models that met the criteria.

The standardized direct, indirect, and total effects of the factors driving the spatial distribution of SOC density were calculated based on the SEM paths. The direct effect of each driver (here, the composite variable — that is, hydrology, soil property, and climate) was the path coefficient on the arrow that directly pointed from the composite variable to the target variable. The indirect effect was measured as the product of the coefficient on the arrows from a composite variable to a mediator variable and from the mediator variable to SOC. In SEM, multiple mediating variables may be from a driving factor to SOC, so the indirect effect is the sum of indirect effects produced through each meditation. The total effect of a driver is the sum of its direct and indirect effects. All analyses were conducted using the "piecewiseSEM" package in R-4.1.0, according to Jing et al. [51]. In addition to SEM, Pearson's correlation analysis was used to investigate variations in the correlation between SOC density and environmental factors across national and regional scales in R-4.1.0.

2.5. Mapping the spatial distribution of SOC density by digital soil mapping

Digital soil mapping (DSM) was used to map SOC density in the Chinese mainland's coastal wetlands. DSM is a technique used to predict the spatial distribution of soil properties based on soil-environment relationships constructed using soil samples and environmental factors [47,52]. The accuracy of DSM is determined by both the environmental covariates and the predictive methods employed. Numerous predictive methods have been developed, among which machine learning is the most popular due to its high accuracy and ability to fit nonlinear relationships [53]. While machine learning is a data-driven method, coupling soil-environment knowledge to machine learning has recently been applied to soil mapping research [54]. In our study, the fitted structural equation model revealed the underlying mechanisms of how environmental variables impact soil organic carbon. Integrating SEM into machine learning can be a potential approach for addressing challenges in soil mapping. To evaluate the effect of applying SEM relationships on SOC prediction, we introduced two extra types of predictors (Supplementary Material Fig. S7): the composite variables (including climate, hydrology, soil, vegetation, and topography) generated through the observed variables in the fitted SEM models and the gridded SOC density prediction (Y_{SEM}) directly generated from the constructed SEM models with the environmental variables as input. Using different combinations of predictors with machine learning, we demonstrated the effectiveness of employing SEM variables. We finally obtained the optimal combination with the highest prediction accuracy. The methodological framework includes three main parts: (1) feature selection, (2) model evaluation and predictive mapping, and (3) uncertainty analysis (Supplementary Material Fig. S8).

2.5.1. Development of different combinations of predictors for mapping SOC

It has been reported that using selected variables generates higher prediction accuracies than using all the candidate covariables in the DSM [52]. Thus, we used recursive feature elimination (RFE) for variable selection. RFE is an effective method of selecting the variable subset that yields the best model performance based on evaluation metrics by adding or removing each variable iteratively [55]. In this study, a cross-validation was conducted to evaluate the model performance with mean squared error at each iteration.

In addition to the initial environmental variables, the SEM variables, including the SEM composite (representing climate, hydrology, soil, vegetation) and Y_{SEM} (the predicted SOC using SEM models) based on the constructed SEM (Y_{SEM}), were utilized for modeling. To evaluate the effectiveness of different predictors, two combinations were developed, i.e., I: RFE selected from only the initial environmental variables and II: RFE selected from both the initial environmental variables and the generated SEM variables.

2.5.2. SOC density prediction and evaluation

We employed four machine learning methods to map SOC density: RF [56], Cubist [57], Support Vector Machine (SVM) [58], and Extreme Gradient Boosting (XGBoost) [59]. These data-driven methods are commonly used in soil and other geographic element predictions or classifications (detailed description provided in Section 1.3 of Supplementary Materials). All machine learning algorithms and hyperparameter tuning were executed with the "caret" package [59] in R-4.1.0.

To validate the performance of all the models, a ten-fold cross-validation repeated five times was conducted [47]. The root mean square error ($RMSE$) and coefficient of determination (R^2) of the five

validations were used to evaluate the prediction accuracy [46]. To further validate how well the predicted results fit the observed values across the entire north-to-south coastal zones, we compared the mean predicted SOC density integrated at each segment spanning three degrees of latitude with that of the observed samples. First, we divided the entire coastal zone into three-degree latitude segments. Within each segment, we calculated the average values of the predicted SOC density from 90 m grid maps and the observed values of soil samples. We then compared the mean predicted values with the corresponding observed average values for each latitude segment. This spatially explicit assessment allowed us to identify potential patterns or variations in model performance across different geographic regions.

2.5.3. Uncertainty analysis

Quantifying the uncertainty of SOC prediction is an important aspect of DSM [53]. A nonparametric bootstrapping approach was used to quantify the uncertainty in SOC density induced by the soil samples and environmental variables by randomly sampling 100 bootstrap combinations [8]. The mean value of 100 times the prediction values was used as the final prediction result. Based on the 100 generated SOC spatial distribution maps, we calculated the standard deviation of SOC density in each grid pixel and used the spatial variation of the standard deviation divided by the mean value to represent the SOC prediction uncertainty in the study area, with the formula as follows: $Uncertainty = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\overline{y})^2}}{\overline{y}}$, where the numerator is the standard deviation, N is the number of sampling times (here $N = 100$), y_i is the ith predicted value of SOC, and $\overline{y}$ refers to the predicted average value of SOC of N sampling times. Bootstrapping was conducted using the "boot" package [8] in R-4.1.0.

2.6. Estimation of total SOC storage

Based on the obtained SOC density maps, coastal SOC storage was calculated according to the following formula [42]:

$$SOCS = \sum_{i=1}^{N}\left(SOCD_{0-20_i} + SOCD_{20-100_i}\right) \times S \quad (4)$$

where $SOCS$ is the total SOC storage, $SOCD_{0-20_i}$ and $SOCD_{20-100_i}$ represent the SOC densities (kg m^{-2}) of 0–20 cm and 20–100 cm at the ith grid, respectively, and S is the area (m^2) of the grid.

Considering that carbon stock calculations are based on the 90 m resolution grid, whereas the initial wetland extent is at a 30 m resolution, this transformation from a 30–90 m resolution may have caused overestimation. To reduce this type of error, for grid cells smaller than 90 m, the area was calculated based on the actual wetland area of the original 30 m cells rather than using the 90 × 90 m area.

3. Results

3.1. Characteristics of SOC density in the Chinese mainland's coastal wetlands

The SOC density of samples according to climate zones, ecosystems, provinces, and estuaries within the study area is shown in Fig. 1. The SOC density exhibits marked spatial variation across the coast of the Chinese mainland (Supplementary Material Fig. S5), in general increasing from temperate to tropical latitudes (Fig. 1a). Notably, the samples with high SOC values, especially in the topsoil, were located mostly in subtropical regions, which is consistent with Xia et al.'s [7] findings, related to the combined effects of vegetation productivity and carbon input in the subtropical area. Among the four marine areas, SOC density on the Yellow Sea coast exhibited the lowest values. At the same time, the South China Sea had the highest (Fig. 1b). Ecosystem type strongly influenced the distribution of SOC density, especially in the topsoil (Fig. 1c). Mangroves had the highest SOC density values, followed by salt marshes. In contrast, values for the tidal flats were generally low. SOC density in both topsoil and subsoil varied greatly among different estuaries, with the lowest mean values in Laizhou Bay and the highest in Qinglan Harbor (Fig. 1d). The mean topsoil and subsoil SOC densities in Hainan Province were the highest among all the provinces, while mean values were lowest in Shandong and Jiangsu Provinces (Fig. 1e).

3.2. Correlation between SOC density and environmental factors on national and regional scales

The controls on SOC density varied noticeably between national and regional scales (Fig. 2). In general, climate, soil physicochemical properties, and hydrodynamic factors exhibited the strongest correlations with SOC density on the national scale. The most important national controls of SOC density were temperature (MAT), precipitation (MAP), vegetation type (WetType), soil pH, and soil CEC, followed by inland runoff (Runoff) and seawater salinity (Sea Salt; Fig. 2). Generally, the correlations between environmental variables and SOC in topsoil and subsoil showed similar patterns.

The highly correlated factors in the four marine zones exhibited large differences (Fig. 2). On the Bohai Sea coast, climate factors were the most important in both the topsoil and subsoil, but hydrological factors (viz., tidal range) dominated in the subsoil. Climate and topography were the most strongly correlated for the Yellow Sea coast, while the distance from the coastline was highly correlated with the subsoil SOC. Meanwhile, vegetation and hydrology were more important than other factors on the East China and South China Seas coasts. This indicates that even the sign of correlation for different factors may vary between zones. For example, MAT or MAP was positively correlated with SOC density on the East China and South China Sea coasts and on the national scale but negatively correlated on the Bohai and Yellow Sea coasts. Moreover, the hydrodynamic intensity of different sea areas varied greatly, and this had a significant impact on the spatial variation of SOC (Fig. 2). Runoff and SWH were strongly negatively correlated with SOC density on the Yellow Sea coast, while tidal range (TidalRange) was significantly positive in the case of the Bohai Sea coast. This indicates that the controlling processes for SOC decomposition and composition are regionally distinct.

3.3. Direct and indirect effects of environmental factors on SOC density

The fitted structural equation models and the direct and indirect effects of each environmental factor for topsoil and subsoil are shown in Fig. 3. The interactions among variables are well supported in SEM. The goodness-of-fit measures of the final SEM models indicate a good fit for the models. For topsoil, 41% of the variation was explained and 40% for subsoil. Since topography did not significantly affect SOC, this factor was omitted from both final models.

The results indicate that hydrology exhibited the highest total effect on topsoil and subsoil organic carbon density, followed by vegetation, soil properties, and climate (Fig. 3c and d). Although the ranking of the four factors in terms of total effect is the same for topsoil and subsoil, the two structural equation models delineate different interactions between the four factors as well as driving mechanisms governing the spatial distribution of coastal SOC

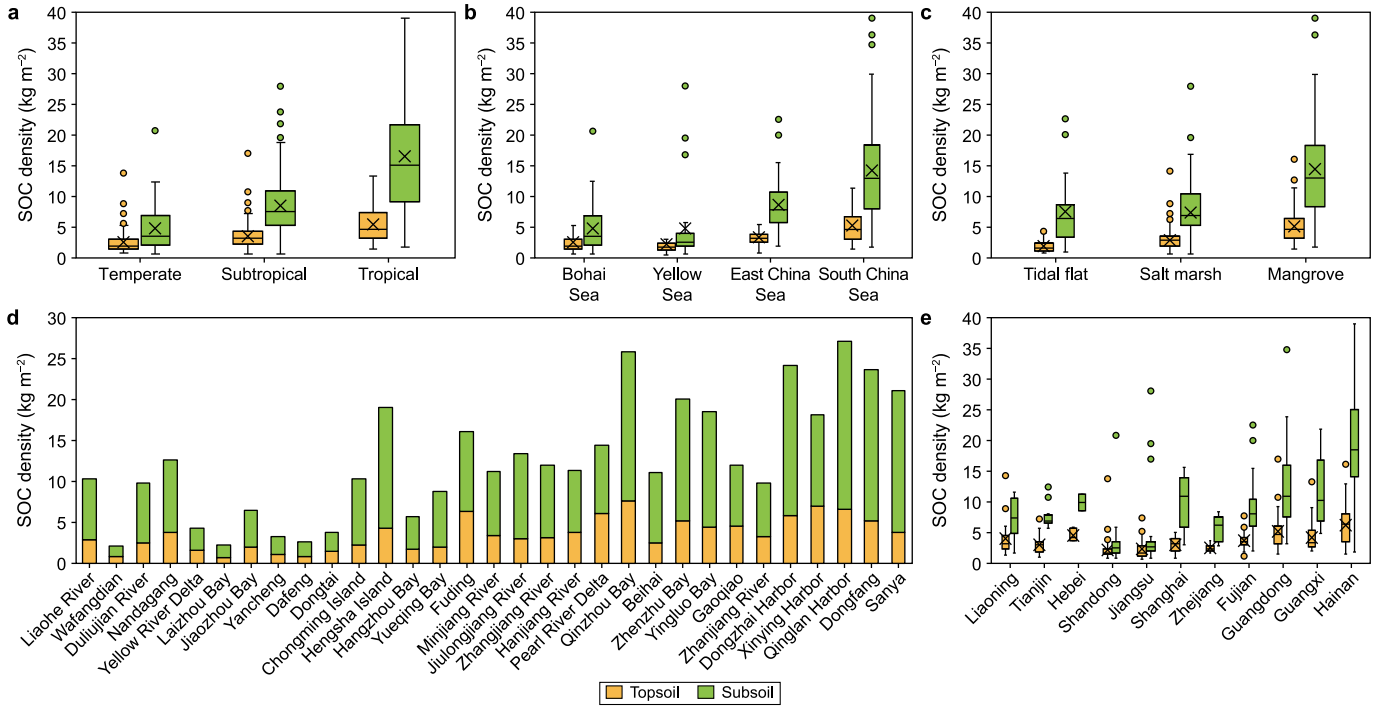

Fig. 1. The distributions of SOC density (kg m^{-2}) in topsoil and subsoil based on climate zones (**a**), marine areas (**b**), wetland types (**c**), independent wetlands (**d**), and administrative areas (**e**). " × " represents the mean and "–" represents the median. Points outside the whiskers indicate outliers, suggesting the presence of unusually high or low values.

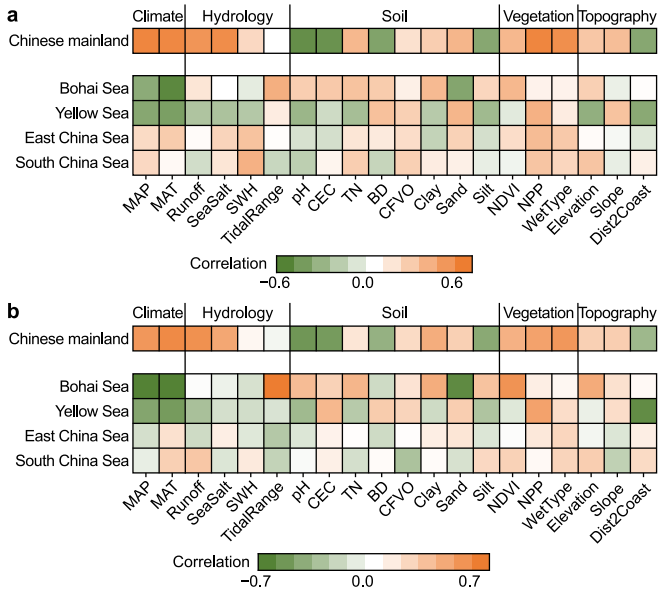

Fig. 2. Correlation between environmental variables and SOC density at national and regional scales using Pearson correlation analysis. Note: MAT: mean annual temperature, MAP: mean annual precipitation, SeaSalt: seawater salinity, SWH: significant wave height, TidalRange: tidal range, TN: total soil nitrogen, BD: bulk density, CFVO: coarse fragment volume, CEC: cation exchange capacity, WetType: wetland type, and Dist2Coast: distance to coastline.

density in the two layers (Fig. 3a and b).

In topsoil (Fig. 3a–c), vegetation had the largest direct effect (0.33) on the spatial distribution of organic carbon density. This may be because plants are the major source of topsoil carbon input and capture organic matter in nearby water bodies. Increasing NDVI and NPP were associated with increasing SOC density. Hydrology, with the second-largest direct effect, positively influenced soil carbon distribution. Greater seawater salinity (SeaSalt) and lower tidal range were associated with increasing SOC content in topsoil. Additionally, hydrological factors exhibited the largest indirect effect by affecting soil properties (with a coefficient of 0.17) and vegetation (0.04). This suggests that periodic inundation and exposure significantly impact soil pH, CEC, clay, and so forth, affecting soil carbon accumulation. Unexpectedly, climate demonstrated no statistically significant direct effect and only indirectly affected SOC density through vegetation and soil properties. In addition, the influence of climate on SOC is shown to depend more on temperature than rainfall (Supplementary Material Fig. S9).

The subsoil structural equation model (Fig. 3b–d) indicated that hydrology had the largest direct effect (0.51) on organic carbon distribution. Seawater salinity (SeaSalt) played the most important role in shaping SOC storage patterns among the hydrological variables. Hydrology emerged as the most significant factor, with seawater salinity ranking first among all factors according to the RF analysis (Supplementary Material Fig. S9). Additionally, larger surface runoff and lower significant wave height (SWH) were associated with increased SOC, as runoff transports more nutrients and lower wave energy reduces erosion and promotes a more stable soil environment. Significantly, hydrological environments and water movement also indirectly affected SOC through soil (−0.14) and vegetation (0.24). This is probably due to water flow patterns that reduce soil aeration and oxygen and increase water saturation. This special environment limits microbial activity and decomposition rates, promoting organic carbon dynamics. Moreover, hydrology brings more nutrients that enhance vegetation growth, thereby contributing to SOC through the input of organic material from plant residues and root systems. The impacts of vegetation (mainly through vegetation type), climate, and soil properties (through soil pH) on organic carbon density were weaker in subsoil (Fig. 3c and d) compared to topsoil.

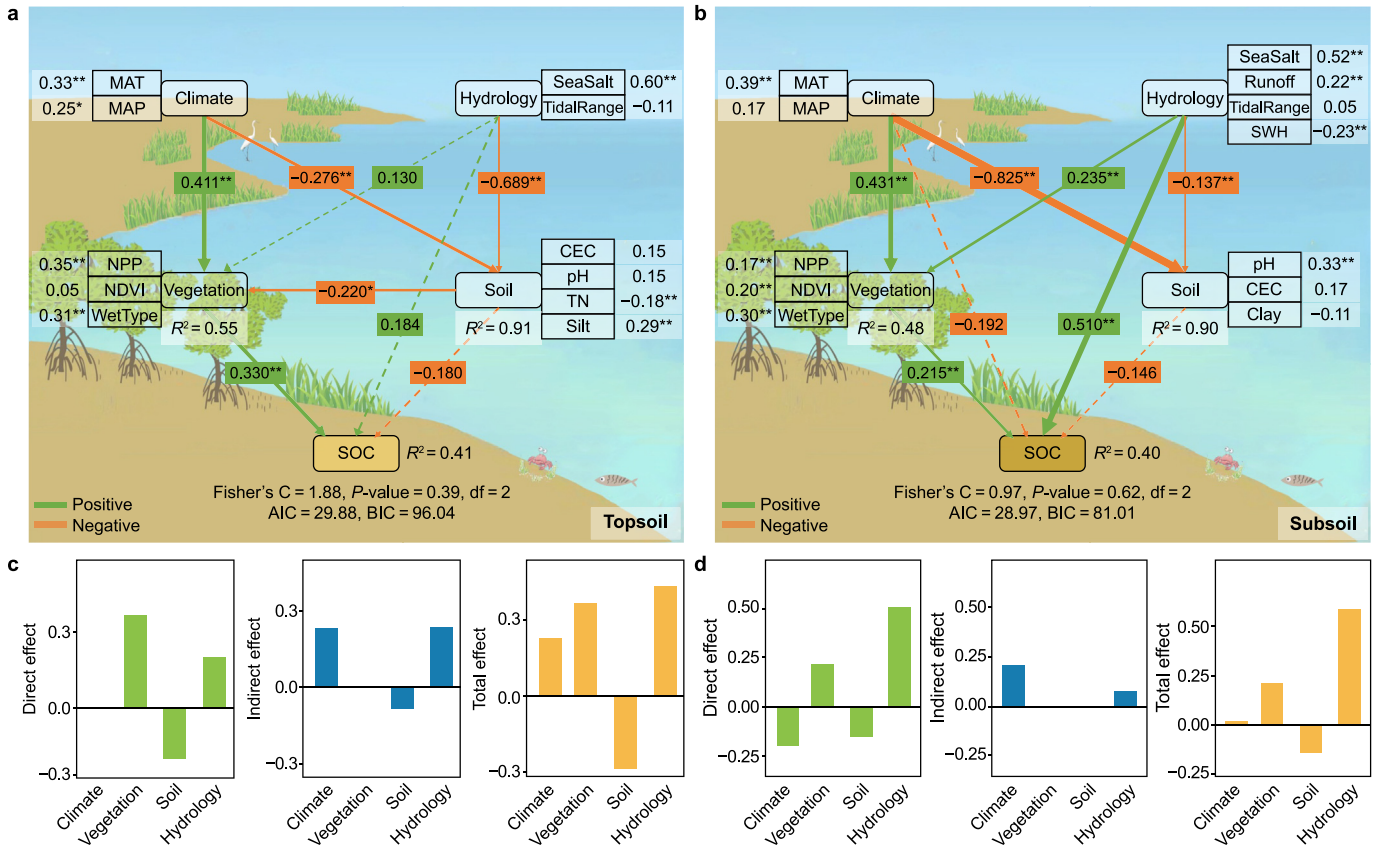

Fig. 3. **a**–**b**, SEM models of topsoil (**a**) and subsoil (**b**) SOC. **c**–**d**, The direct, indirect, and total effects of environmental factors on topsoil (**c**) and subsoil (**d**) SOC. The arrows' thickness is proportional to each arrow's standardized path coefficients (viz., direct effects). Green lines showed statistically positive pathways, while orange lines represented statistically negative pathways. Solid lines indicate statistically significant correlations, whereas dashed lines indicate non-significance. Variables within rounded rectangles are composite variables, while those in boxes are observed variables. The significance levels of each predictor on SOC are $^{**}P < 0.01$ and $^{*}P < 0.05$. The absence of a star symbol signifies no statistical significance. Note: MAT: mean annual temperature, MAP: mean annual precipitation, SeaSalt: seawater salinity, SWH: significant wave height, TidalRange: tidal range, TN: total soil nitrogen, BD: bulk density, CEC: cation exchange capacity, WetType, wetland type.

3.4. Spatial distribution of coastal SOC density

3.4.1. Model performance for SOC density prediction

The root mean square error and R^2 values of the predicted SOC densities using the four machine learning methods are displayed in Table 2. RF outperformed the other machine learning models in all cases, followed by XGBoost.

The accuracy with environmental variables, in addition to SEM variables (SEM composite variables and Y_{SEM}), was slightly improved, irrespective of any machine learning method (Table 2). This proves that adding variables representing soil-environment knowledge from the SEM model improved the mapping accuracy compared to data-driven machine learning models for both topsoil and subsoil.

Considering that the number of variables in models with variable set I was smaller than that of II (Supplementary Material Table S3), and Fig. S10 (Supplementary Material) shows that the predicted SOC density values for the entire area of 90 m grids by variable set I were closer to the observed values for the soil samples than that by variable set II (Supplementary Material Table S4), despite the slightly lower accuracy of variable set I with RF model compared to variable set II with RF model in Table 2. Hence, RF modeling with a variable set I was chosen as the optimal model for mapping topsoil and subsoil SOC density as well as estimating total carbon stock.

3.4.2. Predicted SOC density in the Chinese mainland coastal wetlands and evaluation

We used the optimal model to predict topsoil and subsoil SOC density and calculate uncertainty at a 90 m spatial resolution (Fig. 4a and b). The results indicated that the SOC density values of topsoil and subsoil generally increase from north to south, while topsoil SOC density exhibits greater spatial heterogeneity than that of subsoil, especially in the south. The predicted SOC density was the lowest in Shandong and Jiangsu Provinces and the highest in Guangxi, Guangdong, and Hainan Provinces, basically consistent with that of the samples (Fig. 5). According to the SOC density maps, the total SOC storage in the top 1 m of soil was estimated at 74.58 ± 3.85 Tg C.

The percentage uncertainty maps also exhibit marked spatial variability, ranging from 5.32% to 37.38% for topsoil and 3.18%–30.90% for subsoil (Fig. 4a). Higher levels of uncertainty (>25% on average, Fig. 4a) were observed in some areas of Liaoning Province (both topsoil and subsoil) and some areas of Fujian Province (subsoil).

The mean predicted SOC density within each three-degree latitude segment was close to the observed (Fig. 5), especially for the topsoil, suggesting the selected models were sufficiently reliable. However, there were relatively large discrepancies around latitudes 35°–37° N and 27° N, where the modeled values were overestimated. These differences may be due to a lack of samples in these areas. The results demonstrate that high SOC densities are

Table 2
Accuracy of different models with combinations of predictors.

Layer	Variable sets	Models	*RMSE*		R^2	
			Mean	S.D.	Mean	S.D.
Topsoil	I: Initial environmental variables	RF	1.67	0.36	0.53	0.13
		XGBoost	1.75	0.35	0.49	0.14
		Cubist	1.86	0.41	0.43	0.13
		SVM	1.97	0.43	0.37	0.12
	II: Initial environmental variables and SEM variables	RF	1.66	0.35	0.53	0.12
		XGBoost	1.73	0.29	0.49	0.13
		Cubist	1.85	0.40	0.44	0.13
		SVM	1.96	0.39	0.38	0.11
Subsoil	I: Initial environmental variables	RF	3.92	0.93	0.61	0.13
		XGBoost	4.33	0.96	0.56	0.15
		Cubist	4.21	0.97	0.55	0.13
		SVM	4.44	1.00	0.50	0.13
	II: Initial environmental variables and SEM variables	RF	3.87	0.94	0.62	0.13
		XGBoost	4.24	1.03	0.57	0.16
		Cubist	4.23	0.92	0.55	0.13
		SVM	4.40	0.97	0.51	0.13

Note: Two sets of variables were evaluated in the models, including I: RFE selected variables from only the initial environmental variables and II: RFE selected variables from both the initial environmental variables and the generated SEM variables. SEM variables include SEM composite (Climate, Hydrology, Soil, Vegetation) and Y_{SEM} (the predicted SOC using SEM models). For Topsoil, variable sets are I: SOC ~ pH + MAP + SeaSalt + MAT and II: SOC ~ pH + Climate + SeaSalt + MAT + Vegetation. For Subsoil, variables sets are I: SOC ~ MAT + MAP + SeaSalt + TidalRange + SWH + Runoff + NDVI + Elevation + Slope + pH + Dist2Coast + TN + NPP. II: SOC ~ MAT + MAP + SeaSalt + TidalRange + SWH + Runoff + NDVI + Elevation + Slope + Dist2Coast + TN + BD + Climate + Hydrology + Soil + Vegetation + SOCpre.

characteristic of lower latitude regions (i.e., south of 19° N), while values are lower between 32° and 37° N.

4. Discussion

4.1. Environmental factors driving the spatial distribution of SOC in coastal wetlands

This study utilized structural equation models to examine the relationship between environmental factors (including hydrology, vegetation, climate, soil properties, and topography) and coastal SOC and identify their interactions. The results demonstrate that hydrological variables and vegetation exert a more pronounced impact on the distribution of SOC density in both the topsoil and subsoil of the Chinese mainland's coastal wetlands than climate or soil properties. Furthermore, the mechanisms are different in the topsoil and subsoil layers. The topsoil usually accumulates soil carbon from plant inputs, while subsoil carbon is more easily affected by hydrology in that the environment is anoxic and conducive to carbon accumulation. This observation aligns with previous studies [24] and the feature-importance ranking results from RF models (Supplementary Material Fig. S9). Moreover, our findings highlight the importance of considering the effects of multiple environmental factors on SOC distribution by examining their interactions. Due to the distinctive land-sea interaction environment of coastal wetlands, hydrological factors (i.e., period tidal inundation and wave action) directly bring organic carbon matter in and out through water movement and indirectly influence SOC distribution by affecting vegetation growth and the soil environment. By utilizing SEM models, we could better understand the direct and indirect effects of hydrology and vegetation on coastal soil carbon dynamics.

Our results suggest that hydrological processes, such as tides, waves, seawater salinity, and inland runoff, are particularly important to coastal C and exert positive direct effects on SOC [7]. The strong positive effect of sea salinity on SOC density in our study area may be attributed to the observation that soil microbial activities decrease with salinity [5,19], which constrains carbon loss [18]. In addition, coastal wetlands receive substantial inputs of organic matter and nutrients from inland runoff, encouraging plant growth [60]. However, intensified wave activity can lead to a marked resuspension effect and greater disturbance of surface sediments [21], thereby limiting soil organic carbon accumulation. The dynamic exchange of water flow, including tides and waves, can directly import or export carbon to adjacent habitats, and the lateral transport of carbon caused by tides also affects the distribution of soil organic carbon [61]. Notably, in our study, tidal range and significant wave height (SWH) exhibit non-significant impacts on SOC, perhaps because inputs and outputs due to dynamic hydrodynamic conditions cancel each out. In addition to the direct effect of hydrological factors on SOC density, hydrological processes indirectly affect SOC stability through their impact on soil chemical and physical properties (such as pH, BD, and CEC), particularly in deep soil layers. This may be because hydrological conditions can affect subsoil anaerobic conditions closely correlated with soil microbial activity and decomposition (Fig. 3a and b). Furthermore, tidal characteristics affect vegetation type (viz., WetType), and the nutrients carried by water flow promote the growth of vegetation (viz., NPP and NDVI) [62].

Vegetation type and productivity both considerably impact SOC distribution in coastal wetlands [2], as indicated by SEM models (Fig. 3). Specific vegetation types within these categories also show notable distinctions; for instance, mangroves usually have a higher sediment carbon density than salt marshes. Moreover, a higher NPP indicates increased biomass production, leading to greater carbon incorporation into the soil through processes like litterfall, root turnover, and decomposition of organic materials [63]. Vegetation displays high productivity and consistently supplies organic carbon, contributing to elevated SOC stocks. Furthermore, although high net primary productivity induces high C input, carbon accumulation depends largely on the C stored in the soil, a product of the balance between inputs and outputs [35]. In the topsoil, vegetation plays a more significant role by, for example, preventing soil erosion and promoting sequestration, while in deeper soils, the influence of high NPP plants on coastal wetlands' SOC stocks is shown here to be relatively small (Fig. 3a and b) [60]. In general, net SOC density is influenced by primary productivity and litter input, and the interaction of vegetation with other factors, such as hydrological and soil conditions, may be key determinants of how much carbon eventually enters the soil and is preserved and

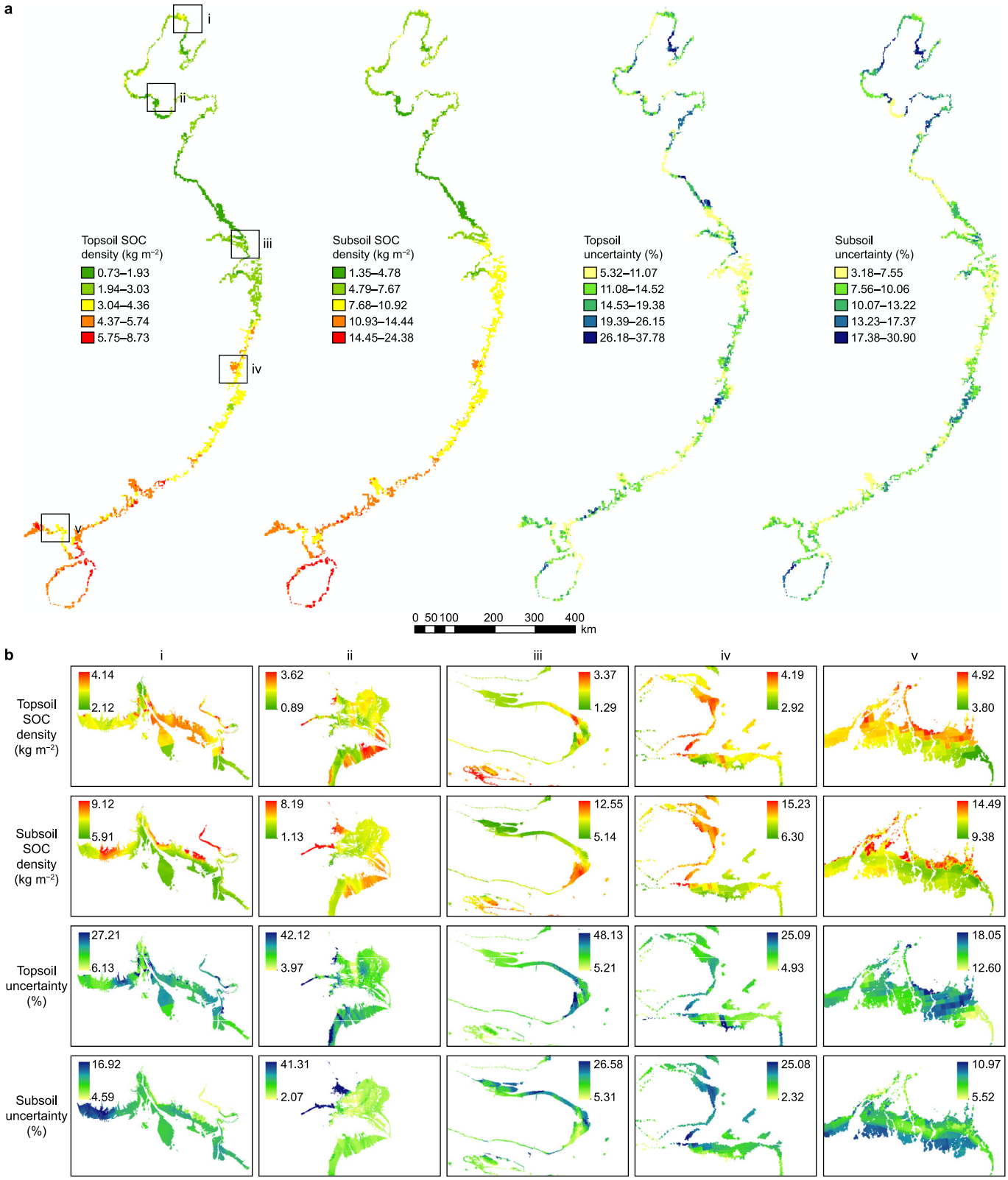

Fig. 4. Predicted SOC density maps (**a**) and uncertainty maps (**b**) in topsoil and subsoil using RF modeling with variable set I at a 90m resolution for the Chinese mainland's coastal wetlands using an Albers conic equal-area projection. For visualization purposes, maps at the national scale (**a**) are depicted in 5 km resolution, and maps in the zoom boxes (**b**) are in 90 m resolution.

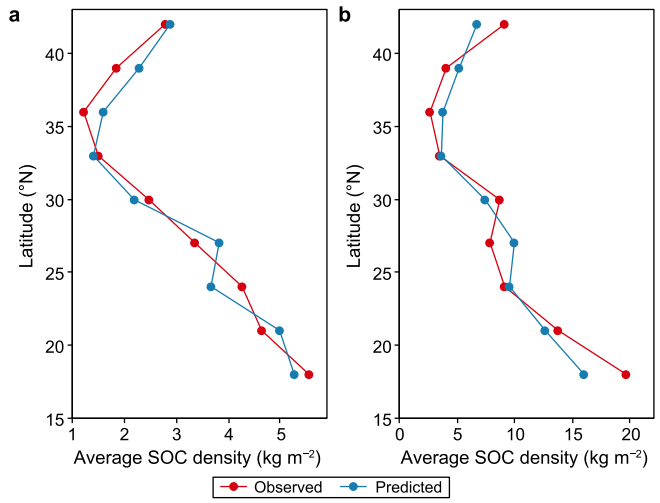

Fig. 5. Average SOC density within each three-degree latitude segment for observed samples and predicted values at 90 m resolution: **a**, topsoil; **b**, subsoil.

stabilized in situ.

There is a growing focus on the role of soil properties in soil organic carbon formation. The results here indicate that soil physicochemical properties have a negative effect on SOC in both topsoil and subsoil (Fig. 3a and b), which is consistent with the results of previous studies [8,20]. We also found that the key factors influencing SOC differ between the two soil layers. In topsoil, soil silt content emerged as the most significant factor determining SOC. This may be because soil physical factors (e.g., silt content) can stabilize carbon in the topsoil, perhaps arising from the formation of protective soil aggregates [34,64]. Meanwhile, chemical properties, including soil pH, were the most significant in the subsoil, followed by CEC. Soil pH is closely linked to microbial carbon cycle processes since greater soil acidity suppresses microbial growth, thereby limiting organic matter decomposition [18,65]. Moreover, soil respiration in coastal wetlands is limited by soil CEC [66]. Our study also revealed that soil exerts indirect effects on SOC through the mediated impact of vegetation by controlling the allocation of water and nutrients, thereby affecting the input and output of soil carbon [8].

Climate is commonly regarded as the most important factor affecting carbon storage; however, our results indicate that the direct effect of climate is not as substantial as vegetation and hydrology. Instead, climate primarily influences SOC indirectly through vegetation and soil. SEM models indicate that the direct effect of climate on SOC in the subsoil is negative, although not statistically significant. The observed negative effect is consistent with previously observed relationships due to the promotion of soil respiration and C loss with increased temperature [4,20,22]. On the other hand, the indirect influence of climate on SOC was observed as positive here, which has generally not been considered before. For example, Tang et al. [67] only reported a simple correlation between climate and SOC, and Wang et al. [33] used stepwise regression to reveal positive feedback between the climate and carbon accumulation of coastal wetlands. Moreover, the positive relationship between temperature and productivity may be partially offset by the negative relationship with decomposition [4,68]. As indicated by our study, the positive indirect effects of climate on SOC compensated for the negative direct effects and resulted in a positive total climatic effect. Our results also suggest that the effect of climate varies with soil depth (Fig. 3c and d) because, in deeper soil, C is less affected by temperature and rainfall variations, while topsoil is more strongly influenced by it [8,20]. Although climate regulates important abiotic and biotic processes that can alter carbon stocks (Fig. 3), the net effect of temperature or precipitation on coastal C varies between regions.

4.2. Estimated SOC stock of the Chinese mainland's coastal wetlands

Estimating SOC stocks in coastal wetlands is important but challenging due to the lack of accurate wetland extent and limited soil observation data. In this study, we addressed these challenges by accurately estimating the extent of coastal wetlands in China using remote sensing technology rather than statistical wetland extent data. Further, based on the currently most comprehensive soil carbon density datasets of coastal wetlands in the Chinese mainland collected in this study, we generated a 90 m SOC density map using DSM. DSM incorporating soil-environment knowledge can accurately characterize the spatial variation of soil carbon. By integrating the 90 m SOC density data, we obtained the total C stock (74.58 ± 3.85 Tg C) down to a 1 m soil depth for the entire coastal wetlands of the Chinese mainland. As seen in Table 3, previous studies estimated the soil C stock in China's coastal wetlands by multiplying the average SOC density by statistical wetland area [7,25,26,31,36], as exemplified by estimates such as 48.12–123.95 Tg C presented by Meng et al. [31], 15.4 ± 1.8 Tg C reported by Fu et al. [25], 57 Tg C estimated by Xia et al. [7], 78.07 Tg C in tidal flats topsoil by Chen et al. [36], and 41.50–114.81 Tg C reported by Wang et al. [26]. Referring to ecosystem types, Chen et al. [36] only focused on a single ecosystem, tidal flats, while Fu et al. [25] and Meng et al. [31] did not include tidal flats. Regarding wetland areas, Fu et al. [25] and Xiao et al. [35] utilized wetland survey data, excluding wetlands below 8 ha in size. Additionally, their wetland survey data encompassed marine areas with a depth not exceeding 6 m, potentially leading to overestimating the total area. Meng et al. [31] and Xia et al. [7] used wetland areas extracted from remote sensing, with the latter underestimating wetland areas. Wang et al. [26] used wetland areas extracted from remote sensing and investigation. Then they multiplied the average SOC stock reported in the literature to estimate China coastal wetlands' total 1 m soil carbon stock as 41.50–114.81 Tg. While it is difficult to directly compare the above results across methods due to variations in wetland extents used, our approach accounts for complex soil-environment relationships, employs a larger number of sampling points, and utilizes advanced soil mapping techniques, thus significantly reducing uncertainty. Our estimate is approximately 30% higher than that of Xia et al. [7], although their study includes no quantitative consideration of uncertainty. This may be due to the accurate mapping of the coastal wetland extent and the obtained soil carbon density map at a 90 m resolution generated based on a substantially greater number of samples as expressed in our study.

4.3. Implications and limitations

This study presents a framework that enhances understanding of the interaction mechanism of environmental factors driving the spatial distribution of SOC over coastal wetlands. It also enables the accurate prediction of detailed SOC maps that can be applied to other countries and regions. The relationship between soil organic carbon and hydrological factors, vegetation, soil properties, and climate and the SOC density maps provide references for wetland management to promote blue carbon storage and mitigate climate change. For example, restoring tidal channels, increasing freshwater input, and reintroducing natural vegetation are viable options.

Table 3
Comparison of existing carbon storage estimation in China's coastal wetlands.

Data sources	Wetland area	Number of soil samples	Soil sample depth	SOC grid mapping	SOC stock in 0–1 m
Meng et al. [31]	3.21×10^4 ha of mangroves, 1.20×10^5 –3.43×10^5 ha of salt marshes in 2013	96	0–1 m	Simple kriging interpolation	48.12–123.95 Tg
Chen et al. [36]	1.1×10^6 ha of tidal flats in 2016	720 for tidal flats	0–15 cm	Average	78.07 Tg
Fu et al. [25]	3.55×10^4 ha of mangroves 1.03×10^5 ha of salt marshes	144 for mangroves 98 for salt marshes	0–1 m	Average	6.3 ± 0.6 Tg 7.5 ± 0.6 Tg
Xiao et al. [35]	5.79×10^6 ha coastal wetlands in 2015	171	0–1 m	Average	490 ± 50 Tg
Xia et al. [7]	5.61×10^5 ha of coastal wetlands in 2015	142	0–40 cm, 40–100 cm	Average	57 Tg
Wang et al. [26]	1.44×10^6 ha coastal wetlands	-	0–1 m	Average	41.50–114.81 Tg (Saltmarsh 9.55 –24.86 Tg; Mangrove 4.951 Tg; Tidal flats 27 –85 Tg)
Our study	1.06×10^5 ha of coastal wetlands in 2010	408	0–20 cm, 20–100 cm	90 m resolution grid mapping	74.58 ± 3.85 Tg

Nevertheless, there are several limitations associated with this study. First, the distribution of sample points influenced the mapping results. The uncertainty of sample points in some areas, such as Jiangsu and Liaoning Provinces, was high, indicating more samples can be collected in these areas to improve prediction accuracy in the future (Fig. 4). Second, the soil properties as influencing factors of soil carbon, for the collected sample points, were not directly measured but obtained from other data sources (e.g., the SoilGrids database). Third, estimating SOC stock at a depth of 1 m for the whole study area ignores the uncertainty of soil depth variations (see Macreadie et al. [30]).

The extent of coastal wetlands constantly changes over the years, resulting in dynamic soil carbon stocks within the wetland area. To estimate the dynamic SOC stock of coastal wetlands in the future, seasonal or interannual variations in coastal extent need to be produced using multiyear satellite imagery. SOC samples should be collected continuously in the long term.

In addition to natural environmental factors, human activities (such as aquaculture and pollution) have increasingly affected soil carbon along the coastline. Thus, including covariates that reflect anthropogenic influence would enhance understanding of the mechanism driving soil carbon distribution.

Accurate assessment of soil-environment relationships or process-based models serves as a useful complement to the data-driven method. This paper provides a potential method for integrating data with machine learning models. Although adding variables generated from the fitted SEM models slightly increased the SOC prediction model accuracies at a sampling point level, the predicted SOC maps show differences between with or without the SEM variables (Supplementary Material Fig. S11). Based on the maps generated with the SEM variables, the calculated SOC stock was 93.55 ± 5.57 Tg, while models without SEM variables produced more conservative SOC stock estimations: 74.58 ± 3.85 Tg (Fig. 4). This indicates that using different predictors produces different results, similar to Holmquist et al.'s [32] and Lopatin et al.'s [54] findings. However, validation based on sampling points should not be the only criterion for choosing the final result. Further solutions to integrate knowledge with data-driven models should be investigated.

5. Conclusions

This study compiled 408 samples from publicly available literature and datasets to establish the currently most comprehensive SOC database for the Chinese mainland coastal wetlands. By doing so, we disentangled the direct and indirect effects of multiple environmental factors on SOC using SEM. We developed the first version of high-resolution (90 m) coastal wetland topsoil and subsoil SOC density grids of the Chinese mainland by integrating machine learning with the soil-environment knowledge generated by SEM. The results highlight the key role of hydrological factors (with the largest total effect) and vegetation (with the second largest total effect) in shaping the spatial distribution of coastal soil carbon in both topsoil and subsoil — that is, increased seawater salinity and higher vegetation productivity are associated with a higher SOC density. Vegetation exerts its greatest direct impact on the topsoil, but its effect decreases in the subsoil, where hydrology plays a more influential role. Soil properties, such as pH and CEC, negatively affect topsoil and subsoil SOC density. Climate has a limited direct impact and mostly affects SOC density indirectly. The generated 90 m resolution SOC density maps for topsoil and subsoil, as well as the estimated total carbon stock at the 0–1 m depth of the Chinese mainland's coastal wetlands, serve as an important benchmark. Based on the predicted 90 m resolution SOC density map, this study estimates the total blue carbon stock down to 1 m soil in coastal wetlands, including mangroves, salt marshes, and tidal flats, to be 74.58 ± 3.85 Tg C. The estimation based on the grid map was estimated to have lower uncertainty than previous studies. It improved the current understanding of the underlying mechanisms of soil carbon in the Chinese mainland's coastal wetlands. Hence, this study offers new insights for improving wetland management strategies and climate change policies related to wetlands nationally.

CRediT authorship contribution statement

Mao Guo: Writing - Review & Editing, Writing - Original Draft, Visualization, Validation, Software, Methodology, Investigation, Formal Analysis, Data Curation, Conceptualization. **Lin Yang:** Writing - Review & Editing, Writing - Original Draft, Validation, Funding Acquisition, Formal Analysis, Conceptualization. **Lei Zhang:** Writing - Review & Editing, Software, Formal Analysis. **Feixue Shen:** Writing - Review & Editing, Software, Methodology, Formal Analysis. **Michael E. Meadows:** Writing - Review & Editing, Validation, Resources. **Chenghu Zhou:** Writing - Review & Editing, Funding Acquisition, Formal Analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was financially supported by the National Key Research and Development Program Plan (Grant No. 2022YFC3800802), the Special Fund of Jiangsu Province Carbon Peak and Carbon Neutral Technology Innovation (Grant No. BK20220037), Fundamental Research Funds for the Central Universities (Grant No. 0209-14380115), and grant from State Key Laboratory of Resources and Environmental Information System.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ese.2024.100482.

References

[1] M.F. Adame, R.M. Connolly, M.P. Turschwell, et al., Future carbon emissions from global mangrove forest loss, Glob Chang Biol 27 (2021) 2856–2866.
[2] P.I. Macreadie, A. Anton, J.A. Raven, et al., The future of blue carbon science, Nat. Commun. 10 (2019) 3998.
[3] Y. Meng, R. Gou, J. Bai, et al., Spatial patterns and driving factors of carbon stocks in mangrove forests on hainan island, China, Global Ecol. Biogeogr. 31 (2022) 1692–1706.
[4] M.J. Osland, C.A. Gabler, J.B. Grace, et al., Climate and plant controls on soil organic matter in coastal wetlands, Glob Chang Biol 24 (2018) 5361–5379.
[5] Q. Zhao, J. Bai, G. Zhang, et al., Effects of water and salinity regulation measures on soil carbon sequestration in coastal wetlands of the yellow river delta, Geoderma 319 (2018) 219–229.
[6] A.C. Spivak, J. Sanderman, J.L. Bowen, et al., Global-change controls on soil-carbon accumulation and loss in coastal vegetated ecosystems, Nat. Geosci. 12 (2019) 685–692.
[7] S. Xia, Z. Song, L. Van Zwieten, et al., Storage, patterns and influencing factors for soil organic carbon in coastal wetlands of China, Glob Chang Biol 28 (2022) 6065–6085.
[8] Z. Luo, G. Wang, E. Wang, Global subsoil organic carbon turnover times dominantly controlled by soil properties rather than climate, Nat. Commun. 10 (2019) 3688.
[9] Y. Wu, Z. Guo, Z. Li, et al., The main driver of soil organic carbon differs greatly between topsoil and subsoil in a grazing steppe, Ecol. Evol. 12 (2022) e9182.
[10] S. Kramer, S. Marhan, H. Haslwimmer, et al., Temporal variation in surface and subsoil abundance and function of the soil microbial community in an arable soil, Soil Biol. Biochem. 61 (2013) 76–85.
[11] D. Liang, N. Ji, A. Kent, et al., Distinct mechanisms drive plant-nitrifier interactions in topsoil and subsoil, Soil Biol. Biochem. 192 (2024) 109370.
[12] Y. Li, C. Fu, L. Zeng, et al., Changes in organic carbon fractions and sources in deltaic topsoil and subsoil layers: autochthonous and allochthonous inputs, Eur. J. Soil Sci. 72 (2021) 2276–2291.
[13] X. Chu, G. Han, Q. Xing, et al., Dual effect of precipitation redistribution on net ecosystem co2 exchange of a coastal wetland in the yellow river delta, Agric. For. Meteorol. 249 (2018) 286–296.
[14] J. Hao, G. Xu, L. Luo, et al., Quantifying the relative contribution of natural and human factors to vegetation coverage variation in coastal wetlands in china, Catena 188 (2020).
[15] S. Doetterl, A. Stevens, J. Six, et al., Soil carbon storage controlled by interactions between geochemistry and climate, Nat. Geosci. 8 (2015) 780–783.
[16] Z. Luo, W. Feng, Y. Luo, et al., Soil organic carbon dynamics jointly controlled by climate, carbon inputs, soil properties and soil carbon fractions, Glob Chang Biol 23 (2017) 4430–4439.
[17] R.R. Twilley, A.S. Rovai, P. Riul, Coastal morphology explains global blue carbon distributions, Front. Ecol. Environ. 16 (2018) 503–508.
[18] L. Zhou, W. Yan, X. Sun, et al., Regulation of climate, soil and hydrological factors on macrophyte biomass allocation for coastal and inland wetlands in china, Sci. Total Environ. (2021) 774.
[19] G. Zhang, J. Bai, C.C. Tebbe, et al., Salinity controls soil microbial community structure and function in coastal estuarine wetlands, Environ. Microbiol. 23 (2021) 1020–1037.
[20] M. Wang, X. Guo, S. Zhang, et al., Global soil profiles indicate depth-dependent soil carbon losses under a warmer climate, Nat. Commun. 13 (2022) 5514.
[21] S.M. Mudd, A. D'Alpaos, J.T. Morris, How does vegetation affect sedimentation on tidal marshes? Investigating particle capture and hydrodynamic controls on biologically mediated sedimentation, J. Geophys. Res.: Earth Surf. 115 (2010).
[22] R. Yang, Interacting effects of plant invasion, climate, and soils on soil organic carbon storage in coastal wetlands, J. Geophys. Res.: Biogeosciences 124 (2019) 2554–2564.
[23] X. Zhang, Z. Zhang, Z. Li, et al., Impacts of spartina alterniflora invasion on soil carbon contents and stability in the yellow river delta, china, Sci. Total Environ. 775 (2021) 145188.
[24] L. Liu, H. Chen, Y. He, et al., Carbon stock stability in drained peatland after simulated plant carbon addition: Strong dependence on deeper soil, Sci. Total Environ. 848 (2022) 157539.
[25] C. Fu, Y. Li, L. Zeng, et al., Stocks and losses of soil organic carbon from Chinese vegetated coastal habitats, Glob Chang Biol 27 (2021) 202–214.
[26] F. Wang, J. Liu, G. Qin, et al., Coastal blue carbon in china as a nature-based solution toward carbon neutrality, Innovation 4 (2023) 100481.
[27] A.L. Hinson, R.A. Feagin, M. Eriksson, et al., The spatial distribution of soil organic carbon in tidal wetland soils of the continental United States, Glob Chang Biol 23 (2017) 5468–5480.
[28] J.R. Holmquist, L. Windham-Myers, B. Bernal, et al., Uncertainty in United States coastal wetland greenhouse gas inventorying, Environ. Res. Lett. 13 (2018) 115005.
[29] O. Serrano, C.E. Lovelock, B.A. T, et al., Australian vegetated coastal ecosystems as global hotspots for climate change mitigation, Nat. Commun. 10 (2019) 4313.
[30] P.I. Macreadie, M.D. Costa, T.B. Atwood, et al., Blue carbon as a natural climate solution, Nat. Rev. Earth Environ. 2 (2021) 826–839.
[31] W. Meng, R.A. Feagin, B. Hu, et al., The spatial distribution of blue carbon in the coastal wetlands of China, Estuar. Coast Shelf Sci. 222 (2019) 13–20.
[32] J.R. Holmquist, L. Windham-Myers, N. Bliss, et al., Accuracy and precision of tidal wetland soil carbon mapping in the conterminous United States, Sci. Rep. 8 (2018) 9478.
[33] F. Wang, C.J. Sanders, I.R. Santos, et al., Global blue carbon accumulation in tidal wetlands increases with climate change, Natl. Sci. Rev. 8 (2021) nwaa296.
[34] A.S. Rovai, R.R. Twilley, E. Castañeda-Moya, et al., Global controls on carbon storage in mangrove soils, Nat. Clim. Change 8 (2018) 534–538.
[35] D. Xiao, L. Deng, D.G. Kim, et al., Carbon budgets of wetland ecosystems in China, Glob Chang Biol 25 (2019) 2061–2076.
[36] J. Chen, D. Wang, Y. Li, et al., The carbon stock and sequestration rate in tidal flats from coastal china, Global Biogeochem. Cycles 34 (2020).
[37] Z. Zhang, N. Xu, Y. Li, et al., Sub-continental-scale mapping of tidal wetland composition for east asia: a novel algorithm integrating satellite tide-level and phenological features, Remote Sens. Environ. 269 (2022).
[38] M. Jia, Z. Wang, D. Mao, et al., Rapid, robust, and automated mapping of tidal flats in China using time series sentinel-2 images and google earth engine, Remote Sens. Environ. 255 (2021).
[39] CCRCN, Coastal carbon atlas. https://ccrcn.shinyapps.io/CoastalCarbonAtlas, 2019.
[40] N.H. Batjes, E. Ribeiro, A. van Oostrum, Standardised soil profile data to support global mapping and modelling (wosis snapshot 2019), Earth Syst. Sci. Data 12 (2020) 299–320.
[41] G.L. Zhang, Z.T. Gong, Soil Survey Laboratory Methods, Science Press, Beijing, 2012.
[42] L. Xu, N. He, G. Yu, A dataset of carbon density in Chinese terrestrial ecosystems (2010s), China Scientific Data 4 (2018) 90–96.
[43] L. Xu, N.P. He, G.R. Yu, et al., Differences in pedotransfer functions of bulk density lead to high uncertainty in soil organic carbon estimation at regional scales: evidence from Chinese terrestrial ecosystems, J. Geophys. Res.: Biogeosciences 120 (2015) 1567–1575.
[44] S. Peng, Y. Ding, W. Liu, et al., 1 km monthly temperature and precipitation dataset for China from 1901 to 2017, Earth Syst. Sci. Data 11 (2019) 1931–1946.
[45] L. Cheng, K.E. Trenberth, N. Gruber, et al., Improved estimates of changes in upper ocean salinity and the hydrological cycle, J. Clim. 33 (2020) 10357–10381.
[46] L. Poggio, L.M. de Sousa, N.H. Batjes, et al., Soilgrids 2.0: producing soil information for the globe with quantified spatial uncertainty, Soil 7 (2021) 217–240.
[47] F. Liu, H.Y. Wu, Y.G. Zhao, et al., Mapping high resolution national soil information grids of China, Sci. Bull. 67 (2022) 328–340.
[48] F.T. Maestre, M. Delgado-Baquerizo, T.C. Jeffries, et al., Increasing aridity reduces soil microbial diversity and abundance in global drylands, Proc Natl Acad Sci U S A 112 (2015) 15684–15689.
[49] J.B. Grace, J.E. Keeley, A structural equation model analysis of postfire plant diversity in California shrublands, Ecol. Appl. 16 (2006) 503–514.
[50] M. Guo, L. Yang, F. Shen, et al., Impact of socio-economic environment and its interaction on the initial spread of covid-19 in mainland China, Geospatial Health 17 (2022).
[51] X. Jing, N.J. Sanders, Y. Shi, et al., The links between ecosystem multifunctionality and above- and belowground biodiversity are mediated by climate, Nat. Commun. 6 (2015) 8159.
[52] A.B. McBratney, M.M. Santos, B. Minasny, On digital soil mapping, Geoderma 117 (2003) 3–52.
[53] A.M.J.C. Wadoux, Using deep learning for multivariate mapping of soil with quantified uncertainty, Geoderma 351 (2019) 59–70.
[54] J. Lopatin, T. Kattenborn, M. Galleguillos, et al., Using aboveground vegetation attributes as proxies for mapping peatland belowground carbon stocks, Remote Sens. Environ. 231 (2019) 111217.
[55] P.M. Granitto, C. Furlanello, F. Biasioli, et al., Recursive feature elimination with random forest for ptr-ms analysis of agroindustrial products, Chemometr. Intell. Lab. Syst. 83 (2006) 83–90.
[56] V. Sharma, S. Ghosh, S. Dey, et al., Modelling pm2. 5 for data-scarce zone of northwestern India using multi linear regression and random forest approaches, Spatial Sci. 29 (2023) 415–427.

[57] M. Kuhn, S. Weston, C. Keefer, et al., Cubist models for regression, R package Vignette R package version 00 18 (2012) 480.
[58] A.H. Ngandam Mfondoum, S. Hakdaoui, R. Batcha, Landsat 8bands'1 to 7 spectral vectors plus machine learning to improve land use/cover classification using google earth engine, Spatial Sci. 28 (2022) 401–424.
[59] D. Zhao, J. Wang, X. Zhao, et al., Clay content mapping and uncertainty estimation using weighted model averaging, Catena 209 (2022) 105791.
[60] D. Laffoley, G.D. Grimsditch, The Management of Natural Coastal carbon sinks, Iucn, 2009.
[61] M.J. Bogard, B.A. Bergamaschi, D.E. Butman, et al., Hydrologic export is a major component of coastal wetland carbon budgets, Global Biogeochem. Cycles 34 (2020).
[62] C.M. Duarte, I.J. Losada, I.E. Hendriks, et al., The role of coastal plant communities for climate change mitigation and adaptation, Nat. Clim. Change 3 (2013) 961–968.
[63] X. Yan, R. Wang, Z. Niu, Response of China's wetland ndvi to climate changes, Wetlands 42 (2022).
[64] K. Heckman, C.E. Hicks Pries, C.R. Lawrence, et al., Beyond bulk: density fractions explain heterogeneity in global soil carbon abundance and persistence, Glob Chang Biol 28 (2022) 1178–1196.
[65] A.A. Malik, J. Puissant, K.M. Buckeridge, et al., Land use driven change in soil ph affects microbial carbon cycling processes, Nat. Commun. 9 (2018) 3591.
[66] A.M. Hopple, S.C. Pennington, J.P. Megonigal, et al., Disturbance legacies regulate coastal forest soil stability to changing salinity and inundation: a soil transplant experiment, Soil Biol. Biochem. 169 (2022) 108675.
[67] X. Tang, X. Zhao, Y. Bai, et al., Carbon pools in China's terrestrial ecosystems: new estimates based on an intensive field survey, Proc Natl Acad Sci U S A 115 (2018) 4021–4026.
[68] M.L. Kirwan, L.K. Blum, Enhanced decomposition offsets enhanced productivity and soil carbon accumulation in coastal wetlands responding to climate change, Biogeosciences 8 (2011) 987–993.