Ecological Indicators 178 (2025) 114041

Contents lists available at ScienceDirect

Ecological Indicators

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

Original Articles

Estimating regional terrestrial ecosystem carbon sinks on multi-model coupling approach

Qingzhou Lv [a], Hui Yang [a,*], Jia Wang [b], Gefei Feng [c], Wanzeng Liu [d], Yunhui Zhang [e], Wenfeng Wang [f], Cheng Wang [g], Xinfeng Huang [a], Liu Cui [a], Yina Qiao [a], Huaiwei Fan [a], Yuejing Yao [a], Yin Qin [a], Wenkai Zhang [h], Yang Han [a]

[a] School of Resources and Geosciences, China University of Mining and Technology, Xuzhou 221116, China
[b] School of Computing and Mathematical Sciences, University of Greenwich, London, United Kingdom
[c] School of Linguistic Sciences and Arts, Jiangsu Normal University, Xuzhou 221009, China
[d] National Geomatics Center of China, Beijing 100830, China
[e] Xinjiang Uygur Autonomous Region Meteorological Observatory, Ürümqi 830002, China
[f] College of Geology and Mining Engineering, Xinjiang University, Ürümqi 830017, China
[g] National 305 Project Office, The People's Government of Xinjiang Uygur Autonomous Region, Ürümqi 830000, China
[h] College of Earth and Environmental Sciences, Lanzhou University, Lanzhou 730020, China

A R T I C L E I N F O

Keywords:
Terrestrial ecosystem carbon storage
Carbon cycle
Model coupling
Analysis of carbon sink driving mechanisms

A B S T R A C T

The regional terrestrial ecosystems serve as primary carbon sinks, characterized by strong spatial heterogeneity and significant interannual fluctuations. In Xinjiang, one of China's five autonomous regions, carbon storage increased from 12967.897 Tg C (1Tg = 10^{12} g) to 14262.31 Tg C. Traditional carbon sink assessment methods struggle to fully account for the combined impacts of human activities and environmental factors, impeding accurate depiction of the spatial distribution and evolution of regional carbon stocks. This study proposes a regional terrestrial ecosystem carbon density estimation method based on an ARIMA-CatBoost-RNN coupled model. Firstly, the ARIMA model forecasts carbon density time series, the CatBoost model reduces the impacts of spatial heterogeneity, and the RNN model estimates ecosystem carbon density values. Secondly, terrestrial carbon storage is estimated using an parameter-updated InVEST model, with an accuracy of up to 78.4 %. Finally, the Geodetector model quantifies the influence of nine driving factors on carbon sink capacity. The results reveal that soil carbon stocks comprise 55 %–61 % of total carbon storage, making them the main component of Xinjiang's terrestrial ecosystems. Annual average carbon sequestration is 39.02 T C/km^2, with forests showing the highest capacity at 103.33 T C/km^2. NDVI (Normalized Difference Vegetation Index) has the most significant impact on Xinjiang's carbon sink capacity, contributing up to 0.615.

1. Introduction

Terrestrial ecosystems are the second-largest carbon reservoir on Earth, after ocean ecosystems. They play a crucial role in global carbon storage and sequestration, mitigating the greenhouse effect (Zheng & Zheng, 2023). Studies have shown that terrestrial ecosystems absorbed approximately 31 % of anthropogenic CO_2 emissions globally between 2010 and 2019 (Zhang et al., 2023). Currently, the global terrestrial carbon sink is primarily estimated using the carbon budget approach, which indirectly calculates terrestrial carbon uptake by subtracting the amount of CO_2 absorbed by the atmosphere and oceans from the total emissions generated by fossil fuel combustion and land-use change (Piao et al., 2009). This indirect method has the advantage of circumventing the challenges associated with the spatial heterogeneity of terrestrial ecosystems. However, due to the influence of atmospheric transport processes and uncertainties in regional emission inventories, this method is not suitable for estimating terrestrial carbon stocks at the regional scale. Accurately quantifying terrestrial carbon storage at regional scales remains a significant challenge in the scientific community.

In support of regional carbon sink estimation, researchers have proposed a variety of approaches, which can be broadly categorized into

* Corresponding author.
E-mail address: TB23010015A41@cumt.edu.cn (H. Yang).

https://doi.org/10.1016/j.ecolind.2025.114041
Received 29 March 2025; Received in revised form 11 August 2025; Accepted 11 August 2025
Available online 29 August 2025

bottom-up and top-down frameworks (Piao et al., 2022). The bottom-up approach refers to the extrapolation of site-level or grid-scale observational or simulated data to a regional extent. Common methods under this category include the inventory method (Pan et al., 2004), eddy covariance method (Zeng et al., 2020), and ecosystem process modeling (Sun & Liu, 2020). In contrast, the top-down approach estimates terrestrial carbon sinks by inverting atmospheric CO_2 concentration changes. Each of these methods has its strengths and limitations. Fang *et al.* employed the inventory method using 50 years of forest resource inventory data to assess changes in forest biomass carbon storage across the mid-latitudes of the Northern Hemisphere in China (Fang et al., 2001). Although the inventory method provides high-accuracy estimates, it relies heavily on extensive ground-based data collection, making large-scale applications costly and logistically difficult. Rodda *et al.* utilized long-term eddy covariance measurements to evaluate the amount of carbon sequestration of tropical dry deciduous forests in central India (Rodda et al., 2021). While this method offers fine temporal resolution and reduces the influence of climatic variability, it only captures static conditions and lacks the ability to predict future carbon dynamics.

In response, researchers have developed various ecosystem carbon budget models, such as the Carnegie-Ames-Stanford Approach (CASA) model (Wang et al., 2017), the Dynamic Global Vegetation Model (DGVM)(Chen et al., 2020), and the Integrated Valuation of Ecosystem Services and Tradeoffs (InVEST) model integrated with land use models (Wang et al., 2024; Yang et al., 2017), have been developed to explore factors influencing carbon storage and predict future trends. Wang *et al.* proposed a new framework combining system dynamics (SD) models, the Patch-generating Land Use Simulation (PLUS) model, and the InVEST model to predict future carbon stock changes in Bortala, China using land use and climate data (Wang et al., 2022). However, these approaches often rely on the assumption of fixed carbon density and fail to account for the variability in carbon storage among land parcels of the same category, which limits the generalizability of the results. In summary, current research on terrestrial carbon storage has evolved from focusing on single ecosystems (e.g., soil, grasslands, forests, wetlands) to broader spatial scales such as watersheds, urban agglomerations, provinces, and countries. Methodologically, studies have shifted from traditional field surveys to integrated approaches combining remote sensing and model coupling. Nonetheless, prevailing methods still face limitations, including high estimation costs, sensitivity to parameter uncertainties, insufficient accuracy, and limited scalability.

As one of the world's largest emitters of greenhouse gases, China bears a significant responsibility for both emission reduction and carbon sequestration (Wang et al., 2020). Xinjiang, the largest provincial-level administrative region in China by land area—accounting for one-sixth of the nation's territory—hosts diverse ecosystems, including forests, grasslands, deserts, and wetlands, and thus possesses considerable potential for carbon sequestration (Tong et al., 2023; Bahtebay et al., 2021). The variations in carbon stock in this region have profound implications for regional sustainable development strategies. As a core area of the "Belt and Road Initiative," Xinjiang has experienced rapid economic growth, continuous urban expansion, and a rising population. The progressive development of supporting infrastructure in transportation, water conservancy, and power supply has brought significant socioeconomic benefits (Cheng et al., 2023). However, it has also imposed profound pressure on the local ecosystems. Increasing water scarcity, land degradation, heightened ecosystem vulnerability, and biodiversity loss have become increasingly severe issues. Under the influence of intensive human activities, carbon storage in Xinjiang's terrestrial ecosystems exhibits marked fluctuations and strong spatial heterogeneity, making it a representative region for studying the driving mechanisms of carbon stock changes (Yang et al., 2025; Yan et al., 2019).

With the advancement of remote sensing technology and ecosystem modeling, significant progress has been made in the study of carbon storage in the Xinjiang region. Jin et al. employed the Markov–-FLUS–InVEST model to estimate the spatiotemporal distribution of carbon storage under land use changes in Xinjiang, revealing the substantial impacts of urban expansion and grassland degradation on carbon stocks. Zhang *et al.* applied the Ensemble Empirical Mode Decomposition (EEMD) method to analyze the temporal dynamics of terrestrial ecosystem carbon sink capacity in Xinjiang and identified key natural drivers such as Temperature and Precipitation (PRE). Despite these achievements, current research on terrestrial ecosystem carbon storage in Xinjiang still faces several limitations. Most studies focus on the spatiotemporal distribution of carbon storage in single ecosystem types—such as forests and grasslands—while paying insufficient attention to other ecosystem types, including croplands, urban areas, and deserts (Liu et al., 2024; Song et al., 2024;Gong et al., 2012). Moreover, constrained by traditional estimation methods, existing carbon storage datasets for Xinjiang suffer from large uncertainties, insufficient spatial and temporal resolution, incomplete coverage, and difficulty in timely updates (Ghosh et al., 2021; Pan et al., 2020; Zhang et al., 2022). In addition, Xinjiang is characterized by complex climate variability and intensive human activities, resulting in highly volatile carbon storage dynamics. Traditional attribution methods struggle to fully account for the regulatory effects of multiple factors such as TEM, PRE, and land use change on carbon storage. These limitations hinder a comprehensive understanding of carbon storage driving mechanisms and pose challenges to the formulation of effective carbon compensation policies and the realization of regional sustainable development.

Therefore, this study selected the Xinjiang Uygur Autonomous Region of China as the study area and proposed a multi-model coupled framework to assess regional terrestrial ecosystem carbon sink capacity, with the goal of identifying carbon sequestration differences among areas with the same land use type. This approach enhances the accuracy of carbon stock estimation compared to traditional methods and addresses the current limitations of high uncertainty, incomplete spatial coverage, and poor temporal resolution in existing carbon stock datasets. First, the Auto Regressive Integrated Moving Average (ARIMA) model was applied to estimate annual carbon density values for aboveground biomass, belowground biomass, and soil components across the region. Next, the CatBoost model was trained to classify the region into three carbon density levels (low, medium, high), thereby reducing the effect of spatial heterogeneity on subsequent modeling. Then, a Recurrent Neural Network (RNN) model was constructed to explore the nonlinear relationships between carbon density and eight environmental and anthropogenic driving factors, under the constraints of ARIMA and CatBoost outputs, enabling fine-scale (grid-level) carbon density evaluation. Based on the multi-model coupling results, the carbon density parameters in the InVEST model were optimized to achieve high-precision estimation of carbon stocks at the grid-cell scale. Finally, the Geographical Detector Model was employed to investigate the spatial driving mechanisms of natural conditions and human activities on carbon sink capacity. This integrated modeling framework supports China's goal of achieving carbon neutrality by 2060 and advances methodological innovation in the assessment of regional terrestrial ecosystem carbon storage.

2. Materials and methods

2.1. Data acquisition and processing

The datasets utilized in this study are categorized into three types: (1) driving factor data for model training, (2) carbon density data for assessing model performance, and (3) statistical data for estimating total carbon storage. The detailed information on the data can be found in Table 1 and Fig. 1.

2.1.1. Driving factor data

The driving factor dataset includes eight variables: annual mean PRE

Table 1
Information on Basic Data.

	Data Type	Data Source	Coverage Period	Units	Spatial Resolution
Impact Factor	PRE	Big Earth Data Platform	1901–2021	mm	1 km
	TEM	Big Earth Data Platform	1901–2021	℃	1 km
	SM	Big Earth Data Platform	2000–2020	m^3/m^3	1 km
	NDVI	Resource and Environment Science	2000–2020	——	1 km
	NPP	National Environmental Science	2000–2020	$KgC/m^2/y$	1 km
	POP	Resource and Environment Science	2000–2020	p/km^2	1 km
	GDP	Resource and Environment Science	2000–2020	10 $k¥/km^2$	1 km
	Land Cover Data	Resource and Environment Science	2000、2005、2010、2015、2020	——	1 km
Carbon density	Aboveground	NASA	2010	Mg C/ha	300 M
	Belowground	NASA	2010	Mg C/ha	300 M
	Soil Carbon	National Glacier Permafrost	2010	kg C/ m^2	30 M
	Dead Organic Carbon	National Glacier Permafrost	2010	kg C/m^2	1 km
Statistics	Carbon density statistics	Paper Database	2000–2014	kg C/m^2	——

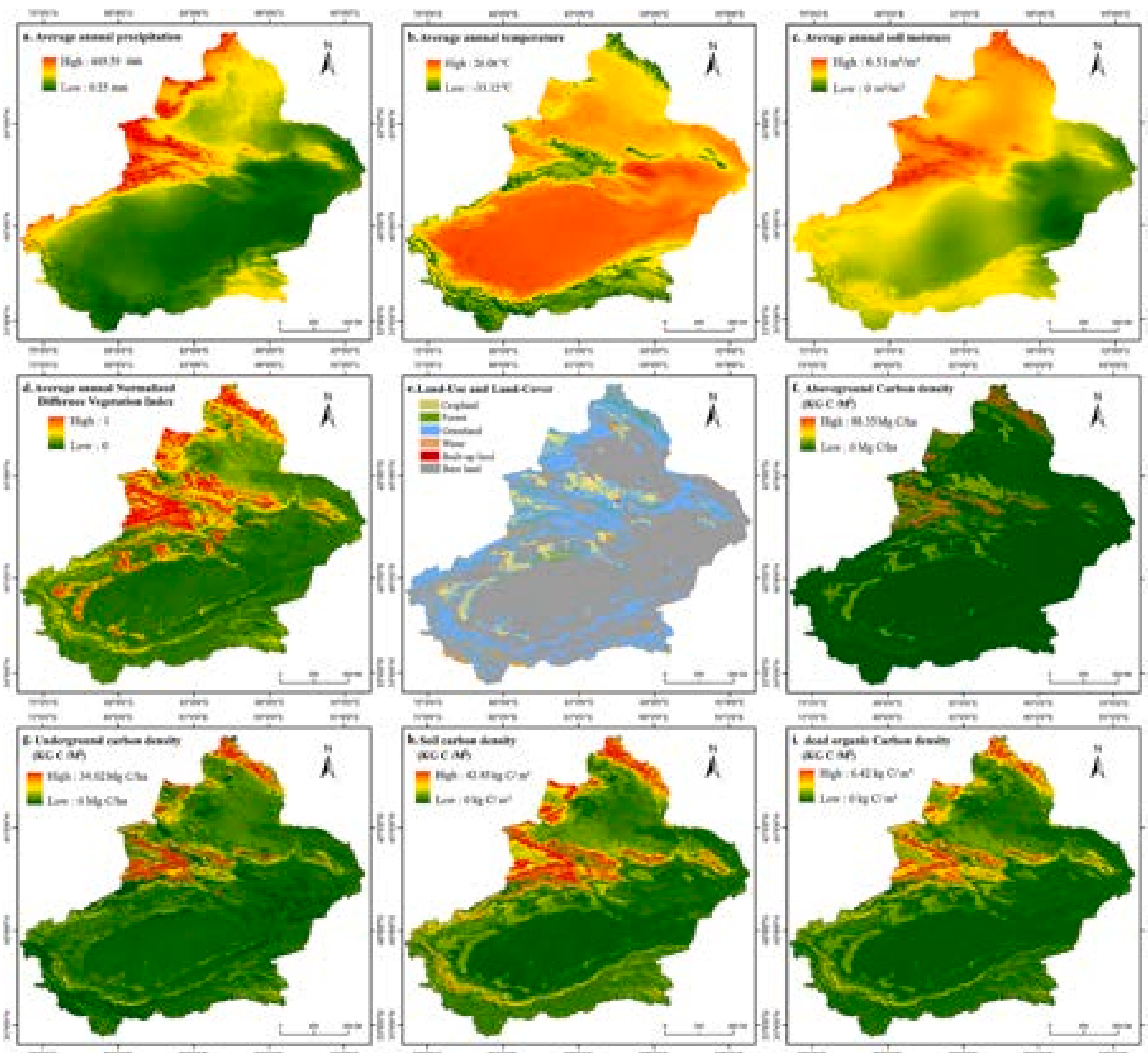

Fig. 1. Impact factor and validation data preprocessing.

(Peng, 2020), annual mean TEM (Peng, 2019), annual mean soil moisture (SM)(Shangguan et al., 2022), NDVI (Li et al., 2024), Net Primary Productivity (NPP)(He et al., 2021), population density (POP)(Yang et al., 2009), per capita Gross Domestic Product (GDP)(Liu et al., 2005), and land cover data (Xu et al., 2018), all with a spatial resolution of 1 km. PRE, TEM, and SM are monthly averaged data obtained from the China Environmental Data Platform. NDVI, POP, GDP, and land cover data are sourced from the Resource and Environment Science and Data Center of the Chinese Academy of Sciences. NPP data are provided by the National Ecosystem Science Data Center of China.

2.1.2. Carbon density data

Carbon density data include belowground biomass carbon density and aboveground biomass carbon density (Spawn et al., 2020), soil organic carbon density (Dong et al., 2025), and dead organic matter carbon density (Liu et al., 2021). The aboveground and belowground biomass carbon density datasets are obtained from the official NASA website, with a spatial resolution of 300 m and units expressed in Mg C/ha. These datasets are derived from remote sensing observations and field survey data using machine learning models. Soil organic carbon density is sourced from the National Tibetan Plateau/Glacier and Frozen Earth Desert Scientific Data Center of China, with a spatial resolution of 30 m and units in kg C/m^2. Due to limited availability of dead organic matter carbon data, this study estimates dead organic carbon density for the Xinjiang region using a coefficient-based method derived from the other three carbon components.

2.1.3. Statistical data

The statistical dataset is based on terrestrial carbon storage data—covering aboveground, belowground, soil, and dead organic matter carbon—from 2000 to 2014, published by the Chinese Academy of Sciences (Xu et al., 2019). A total of 15,610 carbon density records were collected from published literature between 2000 and 2014. Of these, 3,921 records related to Xinjiang and its surrounding regions were selected for model training and validation.

2.1.4. Data preprocessing

Annual mean values were derived from daily and monthly data using MATLAB. ArcGIS was then used to project, resample, and clip all datasets to a 1 km resolution under the WGS1984 coordinate system. To enhance model efficiency, reduce overfitting risk, and improve understanding of relationships between input features and carbon density, the data were preprocessed through three main steps: (1) Extraction: Raster data within the study area were masked and converted to vector points using ArcGIS, generating 1,635,396 samples with attributes including PRE, TEM, NDVI, and carbon density values. (2) Normalization: All variables were normalized to a [0,1] range using Min-Max scaling to ensure comparability. (3) Dataset Construction: Data were stratified by carbon density into equal intervals. A Python script sampled 1,000 points per interval, yielding 100,000 spatially distributed records. These were split into training, testing, and validation sets (7:2:1) and stored in CSV format to support RNN model training.

2.2. Research approach

This study proposes a novel method for estimating regional terrestrial ecosystem carbon density, which serves as a parameter foundation for high-precision carbon storage assessment at the grid scale. The approach addresses current limitations in carbon storage datasets, such as high uncertainty, infrequent updates, and unclear driving mechanisms. Focusing on the Xinjiang region, the methodology integrates multi-source remote sensing data and involves three key steps to analyze the influence of various driving factors on different ecosystem carbon pools:

First, the terrestrial ecosystem is categorized into four major carbon pools: aboveground biomass carbon, belowground biomass carbon, soil organic carbon, and dead organic matter carbon. A coupled ARIMA–CatBoost–RNN model is employed to estimate carbon density, generating spatially continuous datasets at 1 km resolution for the period 2000–2020 in five-year intervals.

Second, based on the model's estimation results, the carbon density parameters of the inVEST model are optimized to achieve high-resolution (1 km) carbon storage assessments for the terrestrial ecosystems of Xinjiang over the same period.

Third, carbon sink capacity is evaluated using the carbon accumulation rate. The spatial distribution patterns and temporal evolution of carbon storage across seven terrestrial sub-ecosystems are analyzed. Additionally, a geographic detector is applied to examine the impacts of human activities and natural factors on carbon sink capacity. The overall research framework is illustrated in Fig. 2.

2.3. Coupled carbon density estimation model

This study develops a carbon density estimation approach based on a coupled ARIMA–CatBoost–RNN model, in which the RNN model is constrained by both the temporal trends predicted by ARIMA and the spatial stratification derived from CatBoost, thereby improving the accuracy of carbon density estimation. The specific methodology is as follows:We first use statistical data on aboveground biomass, belowground biomass, and soil carbon density in Xinjiang from 2000 to 2014 to train the ARIMA model, which is well-suited for time series forecasting. The model predicts the annual average carbon densities for the years 2015–2020 (Equation (1), and these predictions are used as temporal constraints to guide the output restoration process in the subsequent RNN model. Second, we incorporate various driving factors and 2010 carbon density data to train the CatBoost model, which excels at classification tasks. This model classifies the study area into three zones with distinct carbon density levels (high, medium, low) based on input features. These zones are used to perform partitioned RNN training, which mitigates the impact of spatial heterogeneity on the accuracy of carbon density estimation by enabling localized learning. Third, we apply the RNN model, which is particularly effective for capturing complex nonlinear relationships in multidimensional data, to learn the latent mapping between pixel-level carbon density values and multiple influencing factors (Equation (2). During estimation, the RNN's outputs are constrained by both:the temporal trends predicted by the ARIMA model, and the spatial classification results from the CatBoost model. This dual constraint mechanism allows the RNN model to estimate carbon density for each grid cell with high temporal and spatial fidelity, achieving fine-scale, long-term carbon density estimation across the study region (Equation (3). The multi-model coupling process is detailed in the Supplementary Fig. 3.

2.3.1. ARIMA model

ARIMA is a statistical model used for time series forecasting. The model consists of three parts: Autoregressive (AR) part: Represents the relationship between the current observation and past observations. Typically denoted by p, which is the number of autoregressive terms. Integrated (I) part: Represents the differencing of the time series data to achieve stationarity. The number of times the series is differenced is known as the order of integration, denoted by d. Moving Average (MA) part: Represents the dependency between an observation and a residual error from a moving average model applied to lagged observations. Typically denoted by q, which is the number of lagged forecast errors in the prediction equation (Dorogush et al., 2018; Mitchard et al., 2013; Roy et al., 2021), The equation is as follows:

$$C_{dm} = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \tag{1}$$

In the formula: C_{dm} represents the prediction result, p is the number of autoregressive terms, q is the number of lagged forecast errors, e is the natural constant, t is the sample time, ϕ is the autoregressive coefficient, and θ is the moving average coefficient.

Based on the annual average statistical data of aboveground, belowground, and soil carbon densities from 2000 to 2014, we analyzed the trends in carbon densities from 2000 to 2020. Taking soil carbon density estimation as an example, based on the sample data, we used the SPSS data analysis software to determine the differencing order d for the data. We applied the Akaike Information Criterion and Bayesian Information Criterion methods to explore the values of p and q, selecting the values that minimized both criteria (p =1, q =1), which provided the best fitting parameters for constructing the ARIMA model

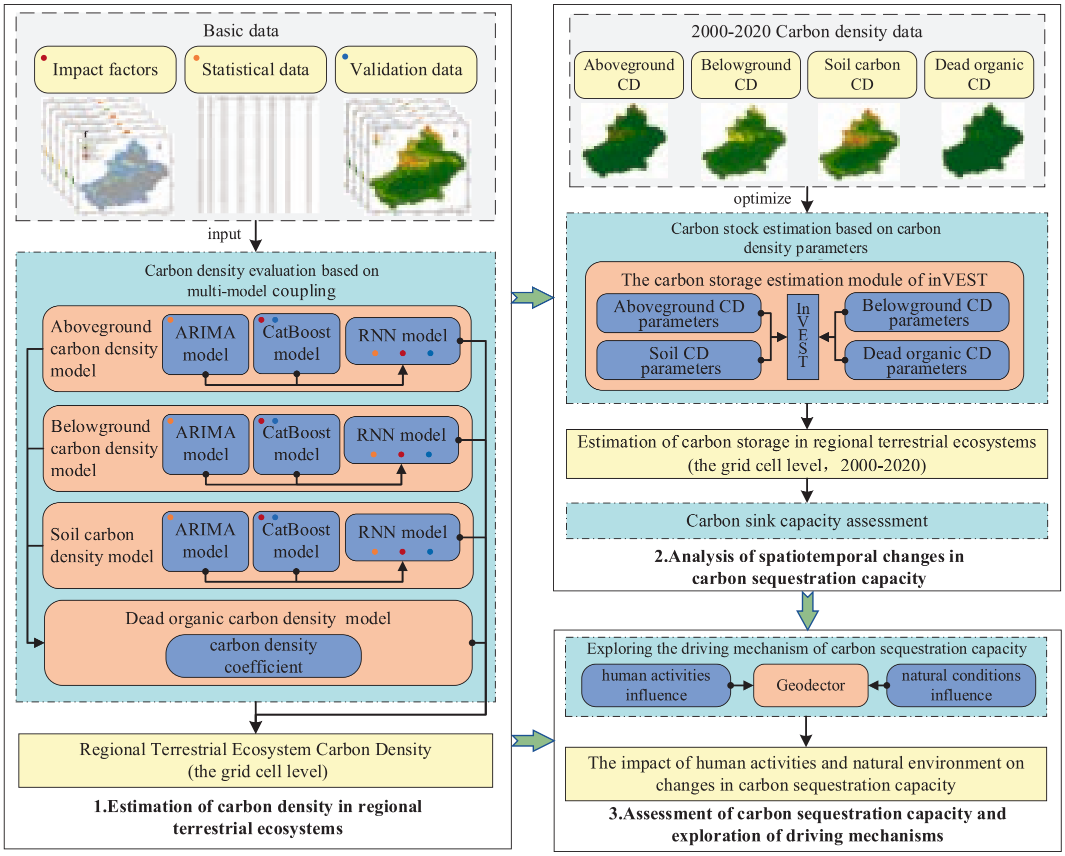

Fig. 2. Research framework. The grey boxes represent the foundational data used in the study; the blue boxes denote the models and methods; the yellow boxes indicate the research findings.

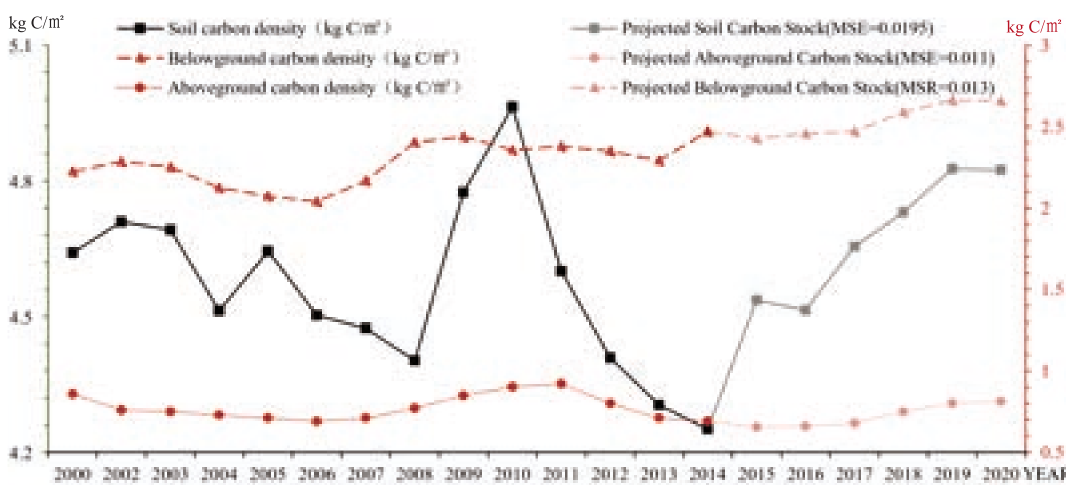

Fig. 3. ARIMA model training and prediction results. Rectangles represent the mean soil carbon density, triangles represent the mean underground carbon density, and circles represent the mean aboveground carbon density. Values of the black line correspond to the left y-axis, while values of the red line correspond to the right y-axis. Dark-colored lines denote training data, and light-colored lines denote predicted results.

(Supplementary Fig. 4). We evaluated the predictive accuracy of the model based on the Mean Squared Error metric (Fig. 3).

2.3.2. *CatBoost model*

The CatBoost model, a classification algorithm based on Gradient Boosted Decision Trees, constructs a symmetric tree structure, allowing the model to more reasonably analyze categorical feature data and minimize overfitting to the greatest extent possible (Chen et al., 2022; Du et al., 2014; Ji et al., 2019). The CatBoost model performs well when handling data with a large number of categorical features. Therefore, in this study, the CatBoost model was trained to extract the relationships between various influencing factors and carbon density, achieving the classification of aboveground, underground, and soil carbon density levels in Xinjiang in 2010. The classification accuracy of various carbon density levels was validated based on existing data (Supplementary Fig. 5). The prediction results of the CatBoost model at large and regional scales are highly consistent with the actual situation, and can effectively depict the true distribution of carbon density levels. The

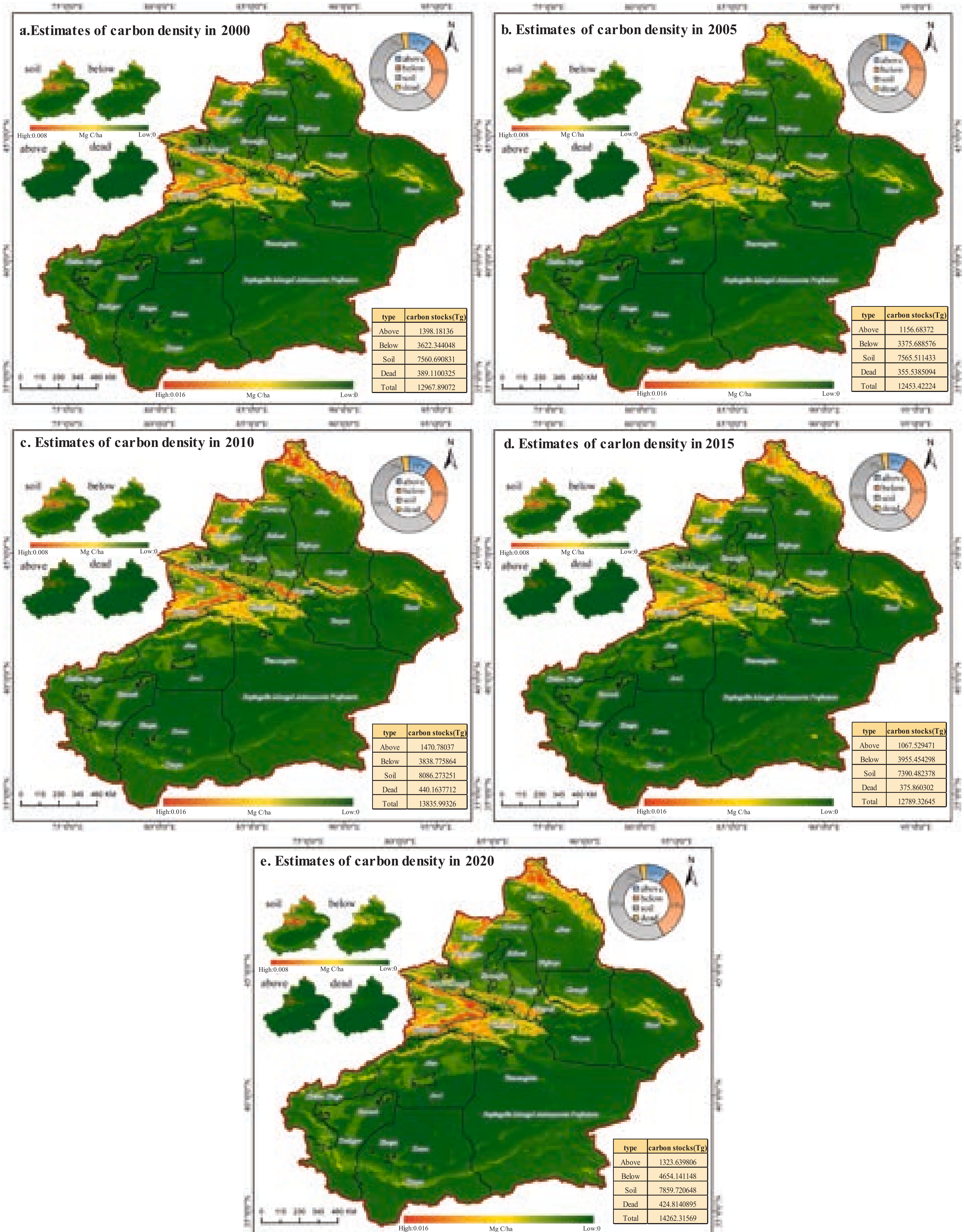

Fig. 4. Carbon density assessment results based on multi-model coupling. a-e The large figure shows the results of the assessment of terrestrial ecosystem carbon density in Xinjiang from 2000 to 2020; The four small charts in the upper left position represent the assessment results of the carbon density of soil, aboveground, underground and dead organic matter in the year, with red indicating the location of high carbon density and green indicating the location of low carbon density, and the carbon density unit is kg C/m^2; The fan chart on the upper right represents the proportion of carbon stocks in the four types of carbon pools; The statistical chart at the bottom right shows that the sum of the carbon stocks of the four types of carbon pools is in Mg C/ha.

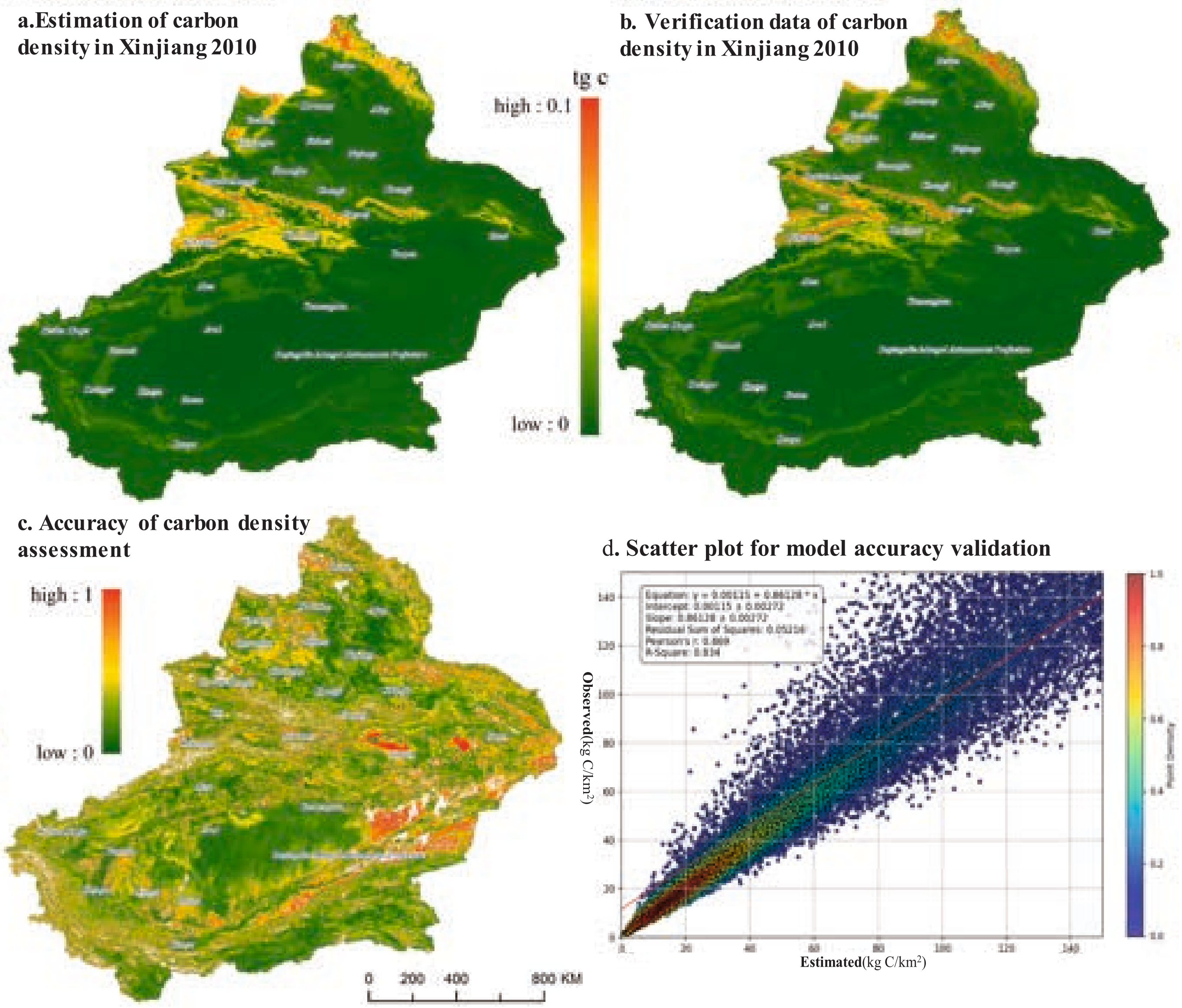

Fig. 5. Verification of carbon density accounting accuracy. Panel a represents the carbon density results of terrestrial ecosystems in the Xinjiang region for 2010 as evaluated by the coupled model; Panel b depicts the carbon density distribution of the 2010 validation data; Panel c shows the relative error between a and b, used to assess model accuracy ; Panel d is a scatter plot comparing the estimated results with the observed values, objectively reflecting the overall accuracy of the model.

accuracy metrics for soil, aboveground, and underground carbon density models are 0.92, 0.86, and 0.85, respectively, with a mean of 0.876, indicating the CatBoost model accurately distinguishes carbon density levels across grid cells (Supplementary Table 1).

2.3.3. RNN model

The RNN model typically includes five parts: the input layer, hidden layer, recurrent structure, hidden state, and output layer. The hidden layer is the key part of the model, responsible for extracting feature information from the data and transmitting it in the form of hidden states through the recurrent structure (Jung et al., 2020; Shen et al., 2009). In this study, an RNN model was trained based on the known aboveground, underground, and soil carbon density data from 2010, along with six types of influencing factor data (the model accuracy is shown in Fig. 5). The model was used to fit the functional relationship between carbon density and influencing factors, and to estimate the carbon density at a 1 km spatial resolution in Xinjiang for the years 2000–2020, based on the influencing factors data from 2000 to 2020. It is particularly crucial to note that, since the input data are all normalized, the output results need to be combined with the carbon density mean predicted by the ARIMA model and the carbon density levels predicted by the CatBoost model to calculate the actual carbon density values, according to the following formula:

$$C_{prei} = f(X_1X_2\cdots X_n) \tag{2}$$

$$C_{di} = C_{prei} * \frac{C_{dm}}{C_{dj}} * (B_{max} - B_{min}) + B_{min} \tag{3}$$

In the formulae (2) and (3): C_{prei} represents the normalized carbon density value predicted by the RNN model for the i grid cell; f is the fitting function of the RNN model; X_i is the independent variable or the value of the influencing factor; C_i represents the actual carbon density value of the i grid cell; C_{dm} is the mean carbon density of the j carbon pool predicted by the ARIMA model; C_{dj} is the mean carbon density of the j carbon pool in 2010; and B_{max}, B_{min} are the maximum and minimum values (thresholds) of carbon density levels obtained by the CatBoost model.

2.4. Carbon stock accounting and carbon sink capacity assessment

2.4.1. Parameter-Optimized InVEST model

The carbon storage in regional terrestrial ecosystems mainly includes four categories: aboveground, underground, soil, and dead organic matter carbon storage. The InVEST model is currently a commonly used carbon storage estimation model. The traditional InVEST model's carbon storage accounting module primarily calculates regional carbon storage using carbon density coefficients based on land use type data. This method overly relies on the classification accuracy of land use data and cannot distinguish the carbon density differences within the same land use type, making it difficult to specifically analyze the changes in ecosystem carbon storage of different subtypes. This significantly impacts the accuracy of regional carbon storage estimation.

Based on land use data, the terrestrial ecosystems in Xinjiang were divided into seven categories: farmland ecosystem, forest ecosystem, grassland ecosystem, water wetland ecosystem, settlement ecosystem, and desert ecosystem. Using a multi-model coupling approach, the carbon density values of four types of carbon pools for each sub-ecosystem category were estimated. The parameters of the InVEST model's carbon storage accounting module were optimized to the grid cell level, allowing it to identify the carbon density value of each grid cell rather than being limited by land use carbon coefficients. The accounting formula is as follows:

$$C = C_{above} + C_{below} + C_{soil} + C_{dead} \tag{4}$$

$$C_i = \sum_{k=1}^{n} C_d{*}S_k\ (i = 1, 2, 3, 4)(k = 1, 2, \cdots n) \tag{5}$$

In the formula 4: $C$is the total carbon storage of the terrestrial ecosystem; $C_{above}$is the total aboveground carbon storage; C_{below} is the total underground carbon storage; C_{soil} is the total soil carbon storage; $C_{dead}$is the total dead organic matter carbon storage; C_i is the carbon storage value of a specific grid cell. In formula 6 C_d is the carbon density value of a specific grid cel; S_k is the area of the grid cell; n is the total number of grid cells in the study area; k is the current grid cell number.

2.4.2. Carbon sink capacity assessment

Based on high-precision carbon storage data from 2000 to 2020 (every 5 years), trend analysis of the annual changes in total carbon storage of regional ecosystems was conducted to assess the carbon storage capacity of each grid cell. The carbon sink capacity of the Xinjiang region was evaluated based on the carbon accumulation rate. Carbon sink capacity refers to the net carbon accumulation of an ecosystem over a given period, representing the difference between carbon uptake and carbon release within one year, which is equivalent to Net Ecosystem Production (NEP) (Randerson et al., 2002; Casas-Ruiz et al., 2023), as shown in the formula below.

$$CAR_k = \frac{\Delta C_k}{\Delta T{*}S_k} \tag{6}$$

In the formula 6:$CAR_k$is the carbon accumulation rate of the k grid cell;ΔC_k is the change in carbon storage of the k grid cell over time T;ΔT is the time interval.

Due to the lack of long-term continuous data on respiratory fluxes and anthropogenic emission disturbances, we adopted NEP as a proxy indicator to represent the net carbon budget status of regional ecosystems, consistent with previous approaches to regional-scale carbon balance assessment (Huang et al., 2023; Xu et al., 2024). Over a certain period of time, a positive carbon accumulation rate indicates a carbon sink area, while a negative value indicates a carbon source area. A larger positive value indicates a stronger carbon sink capacity.

2.5. Analysis of drivers of carbon sink capacity changes

The Geodetector model is a method used for geographic data mining and spatial analysis, which combines geographic information systems, spatial statistical methods, and machine learning techniques. It is commonly used to discover correlations and influencing factors between geographic phenomena (Li et al., 2024; Xu et al., 2021). This study employs single-factor detection and interaction detection in Geodetector to analyze the relationship between carbon sink capacity (Y) and its driving factors (X). The formula for Geodetector calculation is:

$$q = 1 - \frac{1}{N\delta^2}\sum\nolimits_{h=1}^{L} N_h\delta_h^2 \tag{7}$$

In the formula (7):q is the impact degree of the factor; N, N_h are the total number of samples and the total number of samples in layer h; δ^2, δ_h^2 are the variance of Y in the entire region and in layer h. The range of q is [0, 1], with a larger value indicating stronger explanatory power of the influencing factor on the spatiotemporal differentiation of carbon sink capacity, and vice versa.

This study quantitatively evaluates the influence of multiple driving factors on the carbon sink capacity of seven types of terrestrial ecosystems in Xinjiang from two perspectives: natural environmental conditions and human activity intensity. Natural environmental factors include average annual TEM, PRE, and NDVI from 2000 to 2020. Human activity intensity factors include land use intensity (LUI), per capita GDP, and POP.

3. Results

3.1. Multi-model coupling for regional terrestrial ecosystem carbon density estimation

3.1.1. Carbon density trend prediction based on the ARIMA model

Based on multi-year carbon density statistics, an ARIMA model was trained to predict the carbon density range of aboveground, underground, and soil in Xinjiang from 2015 to 2020. The results are shown in Fig. 3. The solid line represents the annual average soil carbon density, the dashed line represents the annual average aboveground biological carbon density, and the dotted line represents the annual average underground biological carbon density. The dark lines represent the training data, while the light lines represent the model's prediction results.

The forecast results indicate that from 2000 to 2020, carbon densities in all types of ecosystems in Xinjiang have shown varying degrees of fluctuation, generally exhibiting an increasing trend. The trends of aboveground biological carbon density and underground biological carbon density are similar, showing relatively small fluctuations and a slow upward trend, suggesting a strong correlation between the two. Soil carbon density shows larger fluctuations.

In terms of average carbon density, The soil carbon pool in Xinjiang has the highest carbon density. Soil carbon density reached its maximum of 4.93 kg C/m^2 in 2010 and its minimum of 4.25 kg C/m^2 in 2014. The proportion of aboveground and underground carbon densities is relatively low, and their changes are not significant.

3.1.2. Carbon Density Classification Based on CatBoost Model

Taking into account the spatial heterogeneity of natural geographic conditions, based on various influencing factors data from 2000 to 2020 and known carbon density data, the CatBoost model was utilized to classify the aboveground, underground, and soil carbon density levels in Xinjiang. The carbon density levels of terrestrial ecosystems in Xinjiang were divided into three levels: low, medium, and high. Before estimating carbon density, regional level classification was conducted to differentiate spatial heterogeneous units, which reduces the mutual influence between different feature grid cells and ensures assessment accuracy. The carbon density classification results based on the CatBoost model can be found in Supplementary Fig. 1.

3.1.3. Multi-Model Coupled Carbon Density Estimation

For the three different types of carbon densities (aboveground, underground, and soil) and the three levels of carbon density (low, medium, and high), multiple RNN models were trained. These models were combined with ARIMA model mean predictions and CatBoost model type predictions to estimate carbon density values in Xinjiang from 2000 to 2020. Carbon density was roughly estimated based on grid cell area. The assessment results are shown in Fig. 4.

The estimation results of the coupled models show that the carbon density distribution in Xinjiang generally exhibits a pattern of higher values in the north and lower values in the south. In northern Xinjiang, carbon is predominantly concentrated in areas with high forest vegetation cover such as Ili, Tacheng, and Altay. Soil, aboveground, underground, and dead organic matter carbon densities in these areas are significantly higher than in other regions. Conversely, desertification is more severe in the eastern and southern parts of Xinjiang, leading to lower carbon densities.

From 2000 to 2020, the overall trend of carbon density in Xinjiang shows a pattern of initial decline followed by an increase. The areas with the most pronounced fluctuations in carbon density are concentrated in the northeast, where carbon density is high. The spatial distribution of annual carbon density changes is not significant; most areas with high carbon density are likely to maintain high carbon levels in the long term, while areas with low carbon density are unlikely to exceed high-density regions during the evolution process.

In terms of the proportion of different types of carbon density, soil is the primary carbon pool in Xinjiang with the highest carbon density, followed by the aboveground carbon pool, then the underground carbon pool, while the dead organic matter carbon pool has the lowest density.

3.2. Model accuracy validation

To validate the accuracy of carbon density estimates derived from the multi-model coupling approach, the differences between the estimated results and the reference dataset were analyzed. A scatter plot was used to visualize the estimation accuracy for each sample. Evaluation metrics, including Relative Error (RE), Pearson's correlation coefficient (r), and the coefficient of determination (R2), were employed to assess model performance. The corresponding formulas are as follows:

$$RE_i = \frac{abs(C_{2010} - C_i)}{C_{2010}} * 100\% \tag{8}$$

In the formula (8), RE_i is the relative error of the i grid cell, C_{2010} is the carbon density validation data for 2010, and C_i is the predicted carbon density for 2010.

As shown in the scatter plot in Fig. 5d, the model's fitted curve closely aligns with the 1:1 reference line when carbon density is below 80 kg C/km^2, indicating high estimation accuracy in this range. However, for carbon densities exceeding 80 kg C/km^2, the scatter points become more dispersed, and the predicted values tend to underestimate the observations. This underestimation may be attributed to the relatively limited number of high carbon density samples, which reduces the model's generalization capability in that range. The scatter plot also clearly illustrates a decreasing sample density with increasing carbon density (color gradient shifting from red to blue). The coefficient of determination (R2) between predicted and observed (Spawn et al., 2020) values is 0.834, and the Pearson's correlation coefficient (r) is 0.869, indicating a strong linear relationship and good model fit.

The relative error validation results (Fig. 5.a-c) indicate that the coupled model can accurately fit the mapping relationship between the driving factors and carbon density. The mean relative error between the predicted results and the validation data is 0.216, indicating that the overall prediction accuracy of the model can reach 78.4 %.

3.3. Carbon stock estimation and carbon sink capacity assessment

In Xinjiang, the desert ecosystem covers the largest area, reaching 683,700 square kilometers, which accounts for 41.16 % of Xinjiang's total area. The grassland ecosystem is the second largest, covering 479,900 square kilometers, accounting for 28.89 % of Xinjiang's total area. Other ecosystems cover an area of 293,400 square kilometers, accounting for 17.67 % of Xinjiang's total area. Forests, farmland, water bodies, and settlement ecosystems cover smaller areas, specifically 33,520 square kilometers (2.01 %), 8,950 square kilometers (5.52 %), 3,400 square kilometers (2.05 %), and 910 square kilometers (0.55 %), respectively.

As shown in Fig. 6b, the carbon storage assessment results indicate that grassland ecosystems in Xinjiang have the highest carbon storage, followed by forests, other ecosystems, farmland, deserts, aquatic systems, and settlement ecosystems. Over the 20-year period, the carbon storage of aquatic systems and wetlands decreased by 2.49 Tg C while the carbon storage of other ecosystems showed an increasing trend. Farmland ecosystems increased by 205.45 Tg C, forests by 164.57 Tg C, grassland ecosystems by 767.11 Tg C, settlement ecosystems by 1.91 Tg C, desert ecosystems by 144.32 Tg C, other ecosystems by 14.54 Tg C, and the total carbon storage of Xinjiang's terrestrial ecosystems increased by 1295.39 Tg C. Except for the carbon storage of aboveground biomass, which decreased, all other types showed an increasing trend, with the underground biomass carbon storage showing the most significant increase of 1032 Tg C. Grassland ecosystems contributed the most to the increased underground biomass carbon storage.

Based on the interannual variation in carbon storage across different ecosystem types, the carbon sink capacity of various ecosystems in Xinjiang was evaluated. Due to numerous results, the evaluation results of grassland and farmland ecosystems, which have a relatively large proportion of carbon storage, are presented in Fig. 7. The carbon sink capacity evaluation results of all ecosystems are shown in Table 2. The study shows that The average annual carbon sequestration capacity of terrestrial ecosystems in Xinjiang is 39.02 T C/KM^2(tonnes of carbon per square kilometer). Forest ecosystems have the highest carbon sink capacity, with an average annual carbon sequestration capacity of 103.33 T C/KM^2, followed by farmland ecosystems at 72.32 T C/KM^2, slightly higher than grassland ecosystems at 64.5 T C/KM^2. The carbon sink capacity of aquatic systems, settlements, deserts, and other ecosystems is weaker, with average annual carbon sequestration capacity of 9.57 T C/KM^2, 12.5 T C/KM^2, 3.71 T C/KM^2, and 7.01 T C/KM^2, respectively. Therefore, the main sources of carbon sequestration capacity in Xinjiang are forests, farmland, and grassland ecosystems.

The evaluation results of carbon sink capacity reveal significant spatial heterogeneity in the sequestration performance of grid cells within the same ecosystem type. This indicates that the use of a fixed-coefficient method for carbon storage accounting is inaccurate. carbon sink capacity Even if the ecosystem type remains unchanged over the long term, factors such as changes in vegetation cover and land use due to natural environmental changes and human activities will have a significant impact on the carbon sink capacity of ecosystems, potentially reducing carbon storage from carbon sinks to carbon sources.

4. Discussion

4.1. Comparison with traditional methods

Process-based models are currently the mainstream approach for estimating regional terrestrial ecosystem carbon storage. Representative models include the CASA, DGVM, and InVEST model. To validate the advantages of the proposed multi-model coupling approach in terms of accuracy and spatiotemporal resolution, this study conducted a comparative analysis by estimating carbon storage in Xinjiang's terrestrial ecosystems using three methods: a CASA model (Wang et al., 2024), the InVEST model, and the proposed multi-model coupling approach.

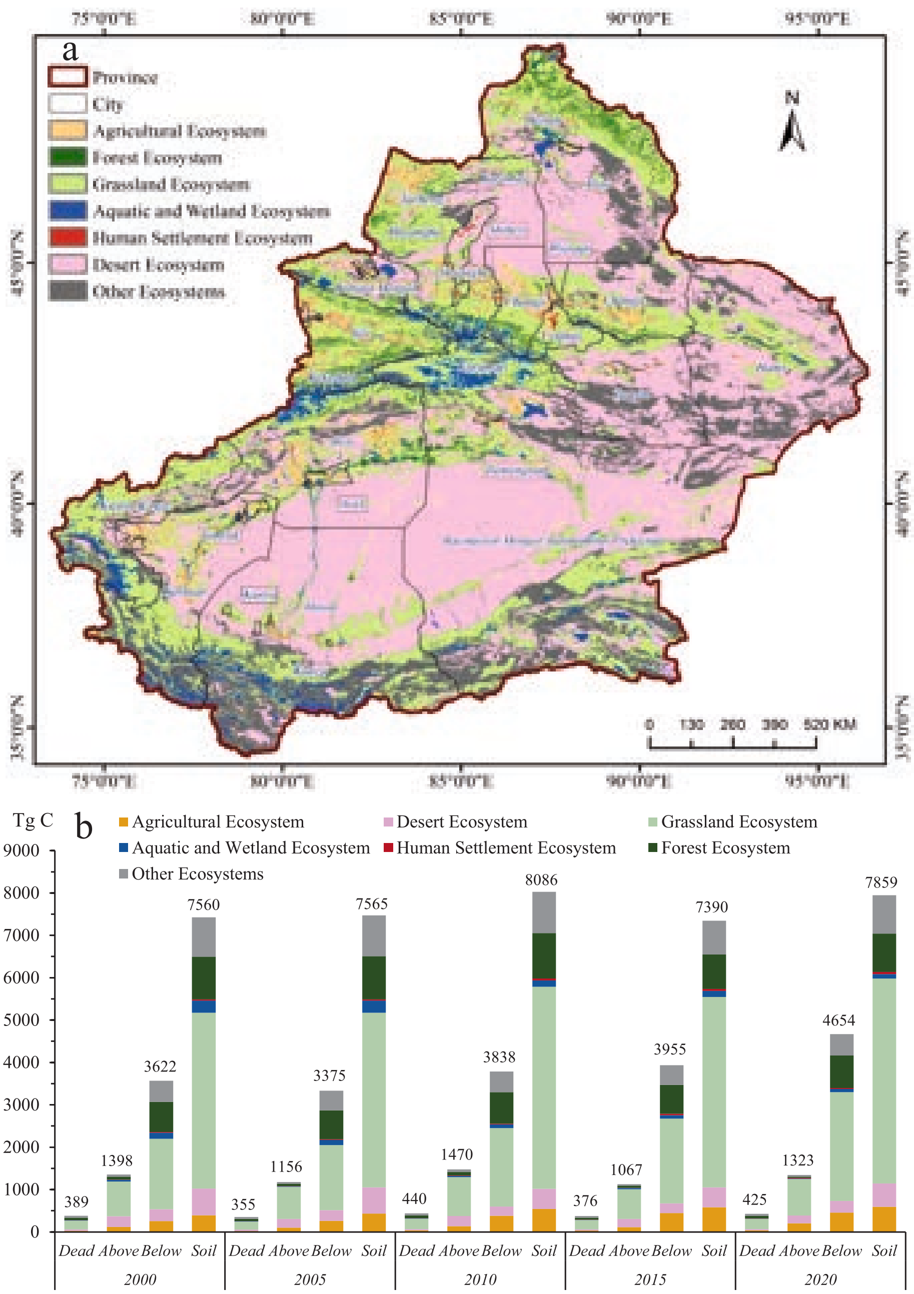

Fig. 6. Classification of terrestrial ecological subsystems and carbon storage changes in Xinjiang from 2000 to 2020. a Distribution of seven types of terrestrial ecological subsystems in Xinjiang; b Interannual variation of carbon storage in seven terrestrial ecosystems in Xinjiang.

The results were then compared with NASA-released carbon storage datasets (Spawn et al., 2020). Model performance was evaluated using three metrics: coefficient of determination (R2), mean absolute error (MAE), and root mean square error (RMSE), in order to quantify differences among the models. The evaluation results are presented in Table 3. The experimental results show that the multi-model coupling approach outperforms both the traditional CASA and InVEST models, achieving the lowest values for R2, MAE, and RMSE among the three. This indicates superior performance in terms of goodness-of-fit, estimation accuracy, and model stability.

As illustrated in Fig. 8, all three models demonstrate the capability to estimate the spatial distribution of regional ecosystem carbon storage and carbon fluxes with reasonable accuracy. However, each model has inherent limitations. The InVEST model is based on fixed carbon density coefficients assigned to land cover types, meaning its accuracy heavily depends on the scientific validity of those coefficients. In this study, we adopted carbon density parameters for Xinjiang derived from published research (Yang et al., 2025). While these values are relatively accurate, the model is still unable to capture within-class heterogeneity in carbon stocks. For example, forests with different dominant species or age structures may have vastly different carbon storage capacities, yet the InVEST model assigns a uniform carbon density to all forest types, resulting in spatially homogeneous outputs that fail to reflect intra-type variability. The CASA model, on the other hand, incorporates a broader set of parameters, including NDVI, photosynthetically active radiation (PAR), TEM, PRE, and potential evapotranspiration (PET). As a process-based model that simulates NPPusing remote sensing and ecological processes, CASA is capable of providing spatiotemporally continuous carbon stock estimates at regional or global scales. However, its accuracy is limited by uncertainties in light use efficiency (LUE) and its lack of sensitivity to anthropogenic influences on carbon dynamics.

In contrast, the proposed multi-model coupling method-trained on

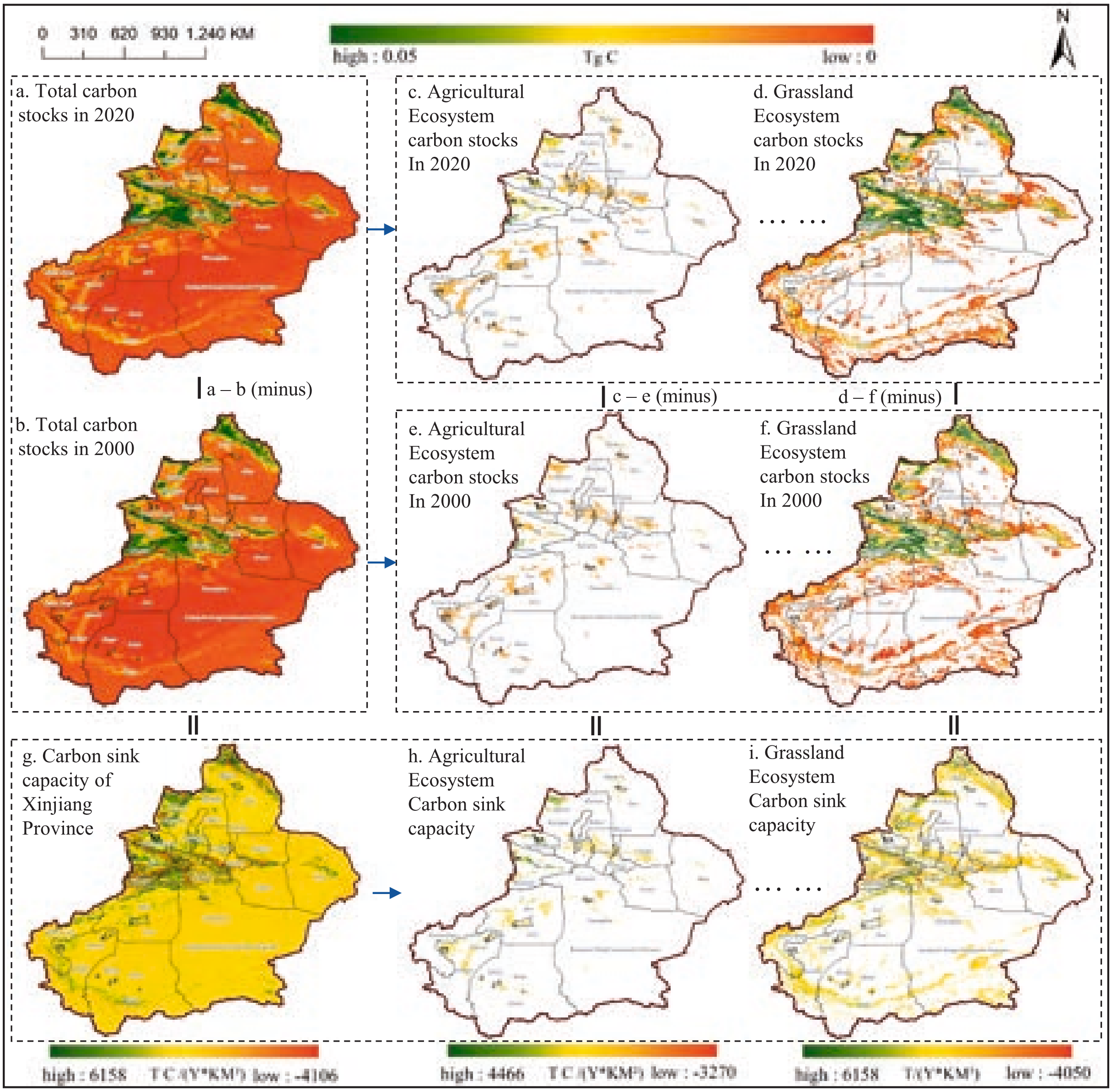

Fig. 7. Carbon stock evolution analysis and carbon sink capacity assessment. a-b are the total carbon storage distribution maps of Xinjiang Autonomous Region in 2020 and 2000, respectively, in (Tg C); c-d are the carbon storage distribution maps of Xinjiang farmland ecosystem and grassland ecosystem in 2020, respectively, in (Tg C); e-f are the carbon storage distribution maps of Xinjiang farmland ecosystem and grassland ecosystem in 2000, respectively, in (Tg C); Panel g shows the distribution of carbon sink capacity of terrestrial ecosystems in Xinjiang, in units of (T C/Y/KM^2); h-I show the distribution of carbon sink capacity of farmland ecosystems and grassland ecosystems in Xinjiang, in units of (T C/Y/KM^2).

high-quality NASA datasets and ground-truth observations—integrates a wide range of natural and anthropogenic drivers (Xu et al., 2019). By establishing a nonlinear mapping between carbon stocks and multiple influencing factors, it enables grid-scale estimation of carbon storage. This approach overcomes a key limitation of traditional models: the inability to distinguish differences in carbon stocks within the same land cover class. Consequently, it significantly improves both the precision and stability of carbon stock estimation for regional terrestrial ecosystems.

4.2. Driving factors of carbon storage in terrestrial ecosystems

The exploration of driving mechanisms behind carbon storage has long been a key focus in the field of carbon cycling. Based on existing research and considering the unique cultural and landscape characteristics of the Xinjiang Autonomous Region, this study discusses the driving mechanisms of carbon sinks from two perspectives: human activities and the natural environment. Six categories of driving factors were selected, including PRE, TEM, NDVI, POP, GDP, and land use intensity (LUI).

Annual PRE, as a key factor controlling ecosystem productivity and carbon storage, is one of the main drivers of spatial patterns in terrestrial carbon storage in Xinjiang. Fig. 4 shows that carbon storage in Xinjiang exhibits a spatial pattern of being lower in the south and higher in the north, which correlates strongly with the spatial distribution of annual PRE gradients in the region. Water not only limits the total growth of vegetation but also indirectly affects soil microbial processes, thereby regulating the input, transformation, and storage of carbon within ecosystems. In the northern regions of Xinjiang, where PRE is more abundant, vegetation coverage is higher, and the efficiency of photosynthetic carbon fixation and biomass accumulation is significantly greater than in the southern regions. Moreover, sufficient moisture

Table 2
Changes in carbon storage and carbon sink capacity assessment of terrestrial sub-ecosystems in Xinjiang.

Type	Carbon sinks (Tg C) [a]	Area change (KM^2)[b]	RAC(T C /Y/ KM^2)[c]
Agricultural Ecosystem	205.45	30,591	72.32
Forest Ecosystem	164.57	−10586	103.33
Grassland Ecosystem	767.11	5404	64.50
Aquatic and Wetland Ecosystem	−2.49	−17229	9.57
Human Settlement Ecosystem	1.91	4663	12.50
Desert Ecosystem	144.32	−6711	3.71
Other Ecosystems	14.54	−6131	7.01

Carbon sinks(Tg C) [a] indicate the changes in carbon storage across different terrestrial ecosystems in Xinjiang from 2000 to 2020, with positive values representing an increase and negative values representing a decrease.
Area change(KM^2)[b] denotes the variations in the area of different terrestrial ecosystems in Xinjiang from 2000 to 2020, with positive values indicating an increase and negative values indicating a decrease.
RAC(T C /Y/KM^2)[c] represents the average carbon sink capacity of different terrestrial ecosystems in Xinjiang from 2000 to 2020. It is equivalent to NEP and is used to indicate carbon sink capacity.

Table 3
Model Accuracy Comparison.

Method	Accuracy Metrics		
	R^2	MAE(Tg C)	RMSE(Tg C)
CASA	0.75	0.0014	0.0039
inVEST	0.48	0.0021	0.0046
Multi-model Coupling	0.83	0.0006	0.0015

R^2 (coefficient of determination): measures the goodness of fit between model estimates and actual values, ranging from 0 to 1, with lower values indicating better fit.
MAE (Tg C, mean absolute error): reflects model accuracy; smaller values indicate higher accuracy.
RMSE (Tg C, root mean square error): reflects model stability; smaller values indicate greater stability.

promotes microbial activity in the soil, which helps enhance carbon stability. Conversely, in the arid southern regions, although sunlight is intense, the limiting effect of water significantly reduces carbon storage.

The annual average TEM in Xinjiang has shown significantly increased interannual variability under the influence of global climate warming. As a key factor controlling carbon inputs and outputs in ecosystems, abnormal TEM changes have bidirectional impacts on carbon storage. In ecosystems dominated by grasslands and deserts, which are characterized by simple ecological structures, low biomass, and poor buffering capacity against climate change, TEM fluctuations exert a stronger influence. TEM also affects vegetation photosynthesis, leading to variations in NPP, which further intensifies interannual fluctuations in carbon storage. Since grasslands and deserts cover the largest area in Xinjiang's terrestrial ecosystems, TEM variability is a critical driver of the annual changes in regional carbon storage.

The spatial distribution of NDVI reflects vegetation coverage. Research shows that in areas of Xinjiang with complex topography—such as river valleys, lakes, and mountainous regions—vegetation coverage tends to be better, NDVI values are higher, and carbon sink potential is likely stronger. In contrast, in areas with uniform landforms—such as deserts, gobi, and meadows—vegetation coverage is poorer, NDVI values are lower, and carbon storage is weaker. Over the past two decades, NDVI values in Xinjiang have shown an upward trend, which is a major contributor to the increase in carbon storage.

This study also integrates POP, GDP, and LUI as representative indicators of human activity to explore how anthropogenic factors influence carbon storage. Firstly, population and economic growth are often accompanied by urban expansion and vegetation destruction. These processes typically alter original land use types, converting natural ecosystems such as grasslands and wetlands into construction land or farmland, thereby reducing the carbon sequestration capacity of aboveground vegetation. Meanwhile, This change in land use can disturb soil structure, accelerate the release of soil carbon, reduce the input of fresh organic matter, alter the composition of microbial communities, and affect the carbon mineralization rate, thereby impacting the long-term stability of the soil carbon pool (Don et al., 2011). Secondly, increases in GDP and land use intensity are usually accompanied by rising energy consumption and greenhouse gas emissions. This regional-scale CO_2 enrichment not only alters the local climatic environment but may also indirectly affect plant productivity and soil carbon input processes through the "CO_2 fertilization" effect. Therefore, these indicators can affect carbon pools in aboveground biomass, belowground biomass, soil organic matter, and dead organic matter to varying degrees. However, the impact of human activities is highly stochastic, and the strength of their influence on carbon dynamics in arid regions remains under debate (Schimel et al., 2001; Bradford et al., 2016).

In conclusion, the spatiotemporal variation in terrestrial ecosystem carbon storage in Xinjiang from 2000 to 2020 has been jointly regulated by multiple factors. From the natural environment perspective, annual PRE and NDVI are the main factors influencing the spatial distribution of carbon storage. In particular, regions in northern Xinjiang with sufficient moisture and good vegetation coverage show significantly higher carbon storage capacity than the arid southern areas. TEM, as a key factor regulating carbon inputs and outputs, significantly influences the interannual variability of carbon storage, especially in grassland and desert ecosystems. From the perspective of human activity, POP, GDP, and LUI—although difficult to directly correlate with carbon storage—indirectly affect changes in carbon storage by influencing carbon emissions and ecological processes. The spatial heterogeneity and dynamic nature of these indicators greatly increase the uncertainty in carbon storage assessments. Therefore, the formation and changes in Xinjiang's carbon storage patterns result from the interaction between natural and anthropogenic factors, with different drivers showing varied impacts in terms of spatial distribution, interannual variation, and data uncertainty.

4.3. Analysis of the driving mechanisms of regional carbon sink capacity

To further quantify the influence of various driving factors on carbon sink capacity in different subclasses of ecosystems, this study uses a natural breakpoint method to categorize all indicators into 10 classes (transforming continuous data into discrete data). The Geographic Detector model is employed to analyze the contributions of six influencing factors to carbon sink capacity (Fig. 9).

Single-factor assessment results indicate that NDVI contributes most to the carbon sink capacity of terrestrial ecosystems in Xinjiang, with an explanation power of 0.615. Following NDVI, the factors in descending order of contribution are LUI (0.318), mean annual TEM (0.288), mean annual PRE (0.267), GDP per capita (0.163), and POP (0.185). Overall, the average comprehensive influence of natural environment (0.39) exceeds that of human activities (0.196), indicating that natural environmental changes have a greater impact on terrestrial ecosystems in Xinjiang. This phenomenon may be related to the vast area and sparse population of Xinjiang, where deserts and bare lands cover a large area, and urban construction land occupies only about 1 % of the total area of Xinjiang. Although the Western Development Strategy and other national strategies have led to large-scale expansion of urban areas in Xinjiang from 2000 to 2020, encroaching on some grasslands and forest ecosystems and altering mean annual TEM, PRE, and carbon dioxide concentration to a certain extent, most ecosystems are far from human interference, and the primary regulating factor is still natural environmental changes.

Notably, although about 60 % of Xinjiang's land area is classified as

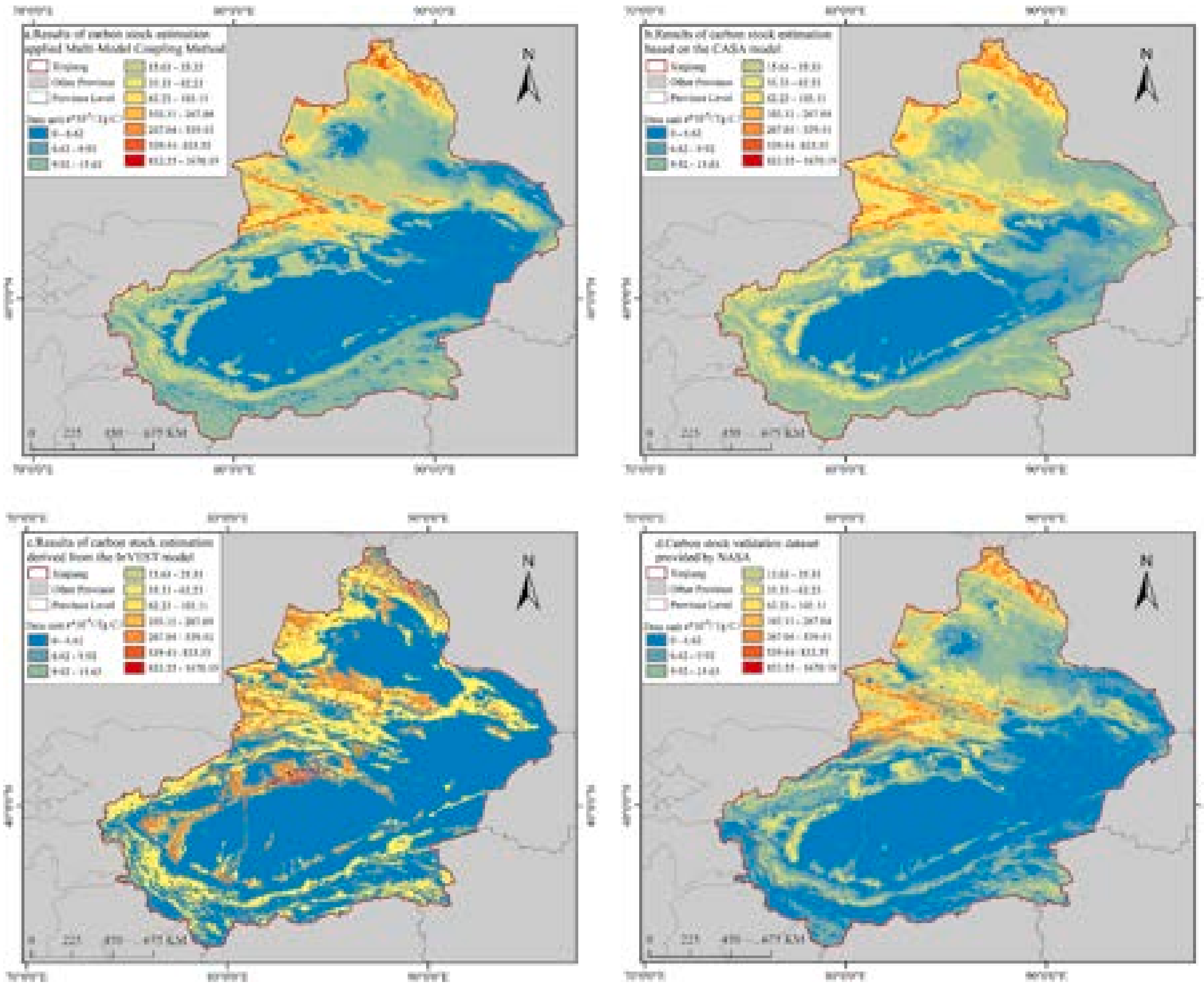

Fig. 8. A Comparison of Results from CASA, InVEST, and the Multi-Model Coupling Method.

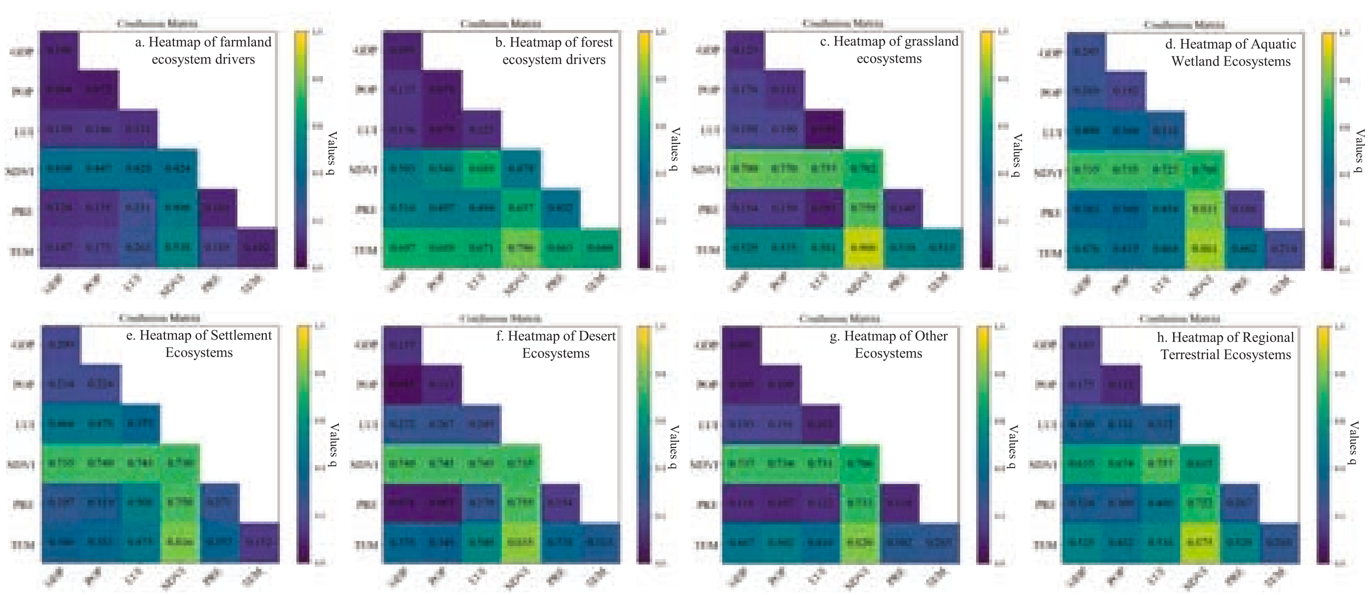

Fig. 9. Carbon sink capacity drives force analysis. a-g represent the influence of various driving factors on the carbon sink capacity of seven sub-ecosystems. Panel h represents the overall impact of different driving factors on the carbon sink capacity of terrestrial ecosystems. The diagonal values in the figure represent the quantitative results of single-factor influence, while other values represent the quantitative results of dual-factor influence.

desert ecosystems, NDVI remains the core influencing factor for carbon sink capacity. This seemingly paradoxical result can be reasonably explained from multiple perspectives. First, NDVI, as an important indicator representing vegetation productivity and coverage, can well reflect the carbon sink capacity of the aboveground part of terrestrial ecosystems, and it also has a good correlation with the carbon sink capacities of aboveground, underground, and dead organic matter (Zhang et al., 2023). Second, in arid ecosystems where vegetation is generally sparse, the ecosystem is extremely sensitive to vegetation changes. Although areas with high NDVI are limited in size, they have high carbon density, making these regions the strongest carbon sink units, which exert a concentrated and significant influence on regional carbon storage distribution. In other words, in arid regions where ecosystems are highly sensitive to vegetation changes, the effect of NDVI as a driving factor may be amplified, thereby exhibiting higher explanatory power in the Geographic Detector model. Furthermore, Xiaohong et al, developed a remote sensing estimation model for aboveground biomass in desert ecosystems in northwest China based on NDVI data, and verified a strong correlation between NDVI and carbon storage. Levent Şaylan et al, using in-situ carbon flux observations along with machine learning techniques, demonstrated that NDVI is the most significant driving factor of growing-season net ecosystem exchange (NEE) variability in Mongolian desert ecosystems. Tian et al, applied the Geographic Detector model to analyze 30-year carbon storage changes in the Tarim River Basin of Xinjiang, and found that NDVI and land surface temperature were the dominant influencing factors (Xiaohong et al., 2023; Şaylan et al., 2024; Tian et al., 2024). These findings are consistent with the present study, in which the contribution rate of NDVI to carbon sink capacity was found to be 0.615. Generally, an $R^2 \geq 0.75$ indicates a strong explanatory power, 0.50–0.75 a moderately strong level, 0.30–0.50 a moderate level, and < 0.30 a weak level; for the Geographic Detector's q-statistic, $q \geq 0.50$ indicates a strong contribution, 0.30–0.50 a moderate contribution, and < 0.30 a weak contribution. In the literature, the reported R^2 between NDVI and carbon storage in arid regions is 0.64, suggesting a moderately strong correlation. In this study, the q-value of NDVI for carbon sink capacity reached 0.615, indicating a strong contribution. The slightly higher value obtained here compared to previous studies may be attributable to differences in data resolution and ecosystem types. Overall, across different regions and methodological approaches, NDVI consistently exerts a substantial influence on the carbon sink capacity of arid ecosystems.

In summary, we consider the conclusion drawn from the Geographic Detector analysis—that NDVI is the dominant driving factor of terrestrial ecosystem carbon density in Xinjiang—to be well supported and scientifically reasonable. Nevertheless, since our study focuses on annual time scales, further investigation is needed to determine whether the influence of NDVI on Xinjiang's terrestrial ecosystems exhibits seasonal variation, which will require expanded data and analysis in future research.

Dual-factor assessment results show that the combined influence of NDVI and mean annual TEM has the strongest explanatory power for carbon sink capacity across subclasses. Following NDVI and TEM, the combined influence of NDVI and LUI also shows significant influence. The dual-factor results mostly indicate enhancement, suggesting that both natural environmental and human activities significantly affect carbon sink capacity, and these factors interact with each other. The explanatory power of multi-factor carbon sequestration capacity analysis is stronger than that of single factors. Apart from NDVI, the carbon sink capacity of agricultural and settlement ecosystems in Xinjiang is highly influenced by LUI. The carbon sink capacity of water body ecosystems, forest ecosystems, grassland ecosystems, and desert ecosystems is highly influenced by mean annual TEM. This phenomenon aligns well with the actual situation: the carbon sink capacity of agricultural ecosystems mainly comes from crops, and the expansion of crops is closely related to human activities. In areas with high LUI, the replanting of crops is more common, thereby enhancing carbon sink capacity. TEM not only affects plant photosynthesis but also influences the activities of various organisms, reflecting the intensity of solar radiation in the region, thereby obviously affecting the carbon sink capacity of water bodies, forests, grasslands, and deserts ecosystems.

The experimental results indicate that the contribution of various influencing factors to carbon sink capacity in Xinjiang Autonomous Region varies significantly. While the trends in the contributions of these factors to different ecosystem subtypes generally align with those observed for the overall ecosystem, notable regional differences exist. For example, the combined impact of TEM and PRE on grassland ecosystems is substantially greater than on agricultural ecosystems, and the combined effects of LUI and PRE on settlement ecosystems are markedly higher than on desert ecosystems. The actual spatial distribution of these factors partially explains the variability in carbon sink capacity across the region, providing important insights for developing targeted carbon sequestration management policies in the future.

4.4. Future research directions

Although this study proposed a multi-model coupling method for carbon density assessment, optimized the carbon stock assessment accuracy of the inVEST model, and discussed the driving mechanisms of various influencing factors on carbon sink capacity, there were some limitations. For instance, This study primarily relied on remote sensing data to construct data-driven models, without simulating the biogeochemical processes of ecosystems. As a result, there are certain limitations in explaining the mechanisms of the carbon cycle. In the future, the data-driven approach adopted in this study could be integrated with dynamic vegetation models (DGVMs), embedding the output of DGVMs as constraints within machine learning models. This integration would enable a more comprehensive representation of photosynthesis, soil respiration, and carbon cycling processes in arid ecosystems, thereby improving the accuracy and interpretability of carbon stock estimation.

Furthermore, while the ARIMA-CatBoost-RNN model demonstrated high assessment accuracy, its predictive ability for scenario-based discussions on carbon stocks and carbon sink capacity was limited. Future studies could integrate the construction approach of land use prediction models such as Plus and Flus to enhance their scenario-based prediction capabilities and better support the formulation of low-carbon policies.

In addition, due to constraints in training data availability and computational resources, this study estimated terrestrial ecosystem carbon stocks at a spatial and temporal resolution of 1 km. While this resolution is sufficiently fine given the large size of the study area (approximately 1,664,900 km^2), it may still introduce notable mixed-pixel effects in ecologically heterogeneous zones such as grassland–desert transition areas. For instance, pixels located at the edges of vegetation and bare land may contain both sparse vegetation and exposed soil, leading to significant errors in carbon stock estimates. Existing studies have shown that, in carbon stock estimation for ecotonal regions, such spatial scale issues may lead to estimation errors of approximately 30 % (Wang et al., 2020). The proposed multi-model coupling approach is inherently scalable and supports multi-scale spatial analysis. With the availability of higher-resolution input data and adequate computational resources in future work, the model could generate finer-resolution outputs that more accurately reflect detailed spatial patterns, including those of ecotones such as the vegetation–desert interface.

Finally, carbon cycling includes both carbon emissions and carbon sequestration, with interactions and even transformations between them. However, this study only discussed the evolution of carbon sink capacity of regional terrestrial ecosystems, with insufficient attention to carbon emissions, which are also a major factor affecting carbon sink capacity. Subsequent research could concentrate more on the relationship between carbon sequestration and carbon emissions.

Current research on regional terrestrial ecosystem carbon sources and sinks still faces many challenges, with many key issues remaining unresolved, and significant differences in carbon stock estimates between different methods. Therefore, future research should combine "multiple data, processes, scales, and methods" to build an integrated carbon accounting system for China, integrating the "sky-land-space" approach, utilizing deep learning models to fully understand the carbon cycling process, clarify carbon sources and sinks mechanisms, and provide scientific support for achieving China's carbon neutrality goal by 2060.

5. Conclusions

This study constructed an ARIMA-CatBoost-RNN model and applied a parameter-updated and spatially refined InVEST model to evaluate the spatiotemporal evolution of terrestrial ecosystem carbon stocks in Xinjiang from 2000 to 2020. Carbon sink capacity of the Xinjiang region was assessed based on carbon accumulation rates. Using a geographic detector model, driving forces of different influencing factors on carbon sink capacity of regional terrestrial ecosystem subsystems were quantified from the perspectives of human activities and natural environment, yielding the following conclusions:

1. Soil carbon reservoirs are the main carbon pool in Xinjiang's terrestrial ecosystems, accounting for approximately 55 %-61 % of the total carbon stock. This is followed by underground biomass carbon stocks (27 %–33 %), aboveground biomass carbon stocks (8 %-11 %), and dead organic matter carbon stocks (3 %). Among the seven ecosystem types, grassland ecosystems have the highest carbon stocks, accounting for 51 %-54 %, followed by forest (13 %-15 %), other types (10 %-12 %), desert (9 %-10 %), cropland (5 %-6%), water bodies (3 %-4%), and settlements (0.3 %-0.4 %)
2. From 2000 to 2020, the overall carbon density in Xinjiang exhibited a trend of initial decline followed by increase, with annual fluctuations in carbon density remaining within 0.8 kg C/m2. Over the 20-year period, the total carbon stock of Xinjiang's terrestrial ecosystems increased by 1295.39 Tg C, reaching a peak of 14262.31 Tg C in 2020. The grassland ecosystem had the largest increase in carbon stock, with an addition of 767.11 Tg C, while the water body ecosystem had the largest decrease, amounting to −2.49 Tg C. The annual average carbon sink capacity of Xinjiang's terrestrial ecosystems was 39.02 T C/KM2, with forest ecosystems exhibiting the highest annual carbon sink capacity of 103.33 T C/KM2
3. Single-factor analysis indicated that natural environmental factors exerted a greater influence on the carbon sink capacity of terrestrial ecosystems in Xinjiang than anthropogenic factors. Among these, the NDVI had the highest explanatory power (0.615). The descending order of influence was as follows: The descending order of influence was as follows: LUI (0.318), TEM (0.288), PRE (0.267), POP (0.185), and GDP (0.163). Multivariate analysis further revealed that the combined effect of NDVI and mean annual TEM exhibited the strongest explanatory power for carbon sink capacityacross ecosystem subtypes. Specifically, carbon sink capacity of cropland and settlement ecosystems were more sensitive to LUI, while those of water bodies, forest, grassland, and desert ecosystems were more strongly affected by mean annual TEM
4. The coupling model achieved a carbon density estimation accuracy of 78.4 %. Experimental results demonstrated that the multi-model coupling method could fully identify differences in carbon sink capacity of the same land types within micro-regions, better depicting detailed changes in carbon sink capacity, and advancing the exploration of carbon sequestration mechanisms

7. Code availability

The code developed for this study has been published on the GitHub platform and can be accessed from the following link: https://github.com/lqz19980404?tab=repositories.

CRediT authorship contribution statement

Qingzhou Lv: Writing – review & editing, Writing – original draft, Supervision, Software, Methodology, Conceptualization. **Hui Yang:** Writing – review & editing, Validation, Supervision, Data curation. **Jia Wang:** Writing – review & editing, Validation. **Gefei Feng:** Writing – review & editing. **Wanzeng Liu:** Writing – review & editing, Validation, Data curation. **Yunhui Zhang:** Formal analysis. **Wenfeng Wang:** Supervision. **Cheng Wang:** Supervision. **Xinfeng Huang:** Funding acquisition. **Liu Cui:** Methodology. **Yina Qiao:** Software. **Huaiwei Fan:** Software. **Yuejing Yao:** Investigation. **Yin Qin:** Writing – review & editing. **Wenkai Zhang:** Writing – review & editing. **Yang Han:** Data curation.

Funding

We appreciate the support and assistance provided by the Xinjiang Department of Natural Resources and other organizations for this research. This research was supported by National Natural Science Foundation of China (42571545), the Third Xinjiang Scientific Expedition Program (2022xjkk1006), the Xinjiang Uygur Autonomous Region Key Research and Development Program (2022B01012-1), the Science and Technology Innovation Project of Jiangsu Provincial Department of Natural Resources (2023018), the Open Research Project of The Hubei Key Laboratory of Intelligent Geo-Information Processing (KLIGIP-2023-A04).

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2025.114041.

Data availability

The carbon storage estimation data for the Xinjiang Autonomous Region of China from 2000 to 2020 (in five-year intervals) can be accessed from https://zenodo.org/records/11570044. The carbon density data can be obtained from https://zenodo.org/records/11579311.

The PRE data used in this study can be accessed from http://poles.tpdc.ac.cn/zh-hans/data/faae7605-a0f2-4d18-b28f-5cee413766a2/; the TEM data can be accessed from http://poles.tpdc.ac.cn/zh-hans/data/7e5333df-0208-4c4e-ae7e-16dcd29e4aa7/; the soil moisture (SM) data can be accessed from http://poles.tpdc.ac.cn/en/data/b0e0911b-9154-48a9-9f2d-6293ccaa3baf/ ; The NDVI data can be obtained from https://www.resdc.cn/data.aspx?DATAID=343; the Net Primary Productivity (NPP) data can be accessed from http://www.nesdc.org.cn/sdo/detail?id=612f42ee7e28172cbed3d80b; the Land Cover Data can be accessed from https://www.resdc.cn/DOI/DOI.aspx?DOIID=54. The Aboveground Biomass Carbon data can be obtained from https://data.nasa.gov/dataset/Global-Aboveground-and-Belowground-Biomass-Carbon-/7e4b-aj6f/about_data; the Belowground Biomass Carbon data can be obtained from https://data.nasa.gov/dataset/Global-Aboveground-and-Belowground-Biomass-Carbon-/7e4b-aj6f/about_data; the Soil Carbon data can be accessed from http://www.ncdc.ac.cn/portal/metadata/2064ef7f-1cb8-4f05-9ca3-3395f52617dd; and the Carbon Density Statistics data can be accessed from http://www.sciencedb.cn/dataSet/handle/603.

References

Bahtebay, J., Zhang, F., Ariken, M., Chan, N.W., Tan, M.L., 2021. Evaluation of the coordinated development of urbanization-resources-environment from the incremental perspective of Xinjiang, China. J. Clean. Prod. 325, 129309. https://doi.org/10.1016/j.jclepro.2021.129309.

Bradford, M.A., Wieder, W.R., Bonan, G.B., et al., 2016. Managing uncertainty in soil carbon feedbacks to climate change. Nat. Clim. Chang. 6 (8), 751–758. https://doi.org/10.1038/nclimate3071.

Casas-Ruiz, J.P., Bodmer, P., Bona, K.A., Butman, D., Couturier, M., Emilson, E.J., Del Giorgio, P.A., 2023. Integrating terrestrial and aquatic ecosystems to constrain estimates of land-atmosphere carbon exchange. Nat. Commun. 14 (1), 1571. https://doi.org/10.1038/s41467-023-37232-x.

Chen, J., Fan, W., Li, D., Liu, X., Song, M., 2020. Driving factors of global carbon footprint pressure: Based on vegetation carbon sequestration. Appl. Energy 267, 114914. https://doi.org/10.1016/j.apenergy.2020.114914.

Chen, M., Qiu, X., Zeng, W., Peng, D., 2022. Combining sample plot stratification and machine learning algorithms to improve forest aboveground carbon density estimation in northeast China using airborne LiDAR data. Remote Sens. (Basel) 14 (6), 1477. https://doi.org/10.3390/rs14061477.

Cheng, Y., Song, W., Yu, H., Wei, X., Sheng, S., Liu, B., Yang, D., 2023. Assessment and prediction of landscape ecological risk from land use change in Xinjiang. China. Land 12 (4), 895. https://doi.org/10.3390/land12040895.

Don, A., Schumacher, J., Freibauer, A., 2011. Impact of tropical land-use change on soil organic carbon stocks–a meta-analysis. Glob. Chang. Biol. 17 (4), 1658–1670. https://doi.org/10.1111/j.1365-2486.2010.02336.x.

Dong, Y., Wang, X., Su, W., 2025. A 1 km soil organic carbon density dataset with depth of 20cm and 100cm from 1985 to 2020 in China. Earth Syst. Sci. Data Discuss. 2025, 1–35. https://doi.org/10.5194/essd-2024-588.

Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. arXiv preprint arXiv:1810.11363. doi: 10.48550/arXiv.1810.11363.

Du, L., Zhou, T., Zou, Z., Zhao, X., Huang, K., Wu, H., 2014. Mapping forest biomass using remote sensing and national forest inventory in China. Forests 5 (6), 1267–1283. https://doi.org/10.3390/f5061267.

Fang, J., Chen, A., Peng, C., Zhao, S., Ci, L., 2001. Changes in forest biomass carbon storage in China between 1949 and 1998. Science 292 (5525), 2320–2322. https://doi.org/10.1126/science.1058629.

Ghosh, S., Behera, M., Jagadish, B., Das, A., Mishra, D., 2021. A novel approach for estimation of aboveground biomass of a carbon-rich mangrove site in India. J. Environ. Manage. 292, 112816. https://doi.org/10.1016/j.jenvman.2021.112816.

Gong, Y., Hu, Y., Fang, F., Liu, Y., Li, K., Zhang, G., 2012. Carbon storage and vertical distribution in three shrubland communities in Gurbantünggüt Desert, Uygur Autonomous Region of Xinjiang, Northwest China. Chin. Geogr. Sci. 22, 541–549. https://doi.org/10.1007/s11769-012-0564-x.

He, Q., Ju, W., Dai, S., et al., 2021. Drought risk of global terrestrial gross primary productivity over the last 40 years detected by a remote sensing-driven process mode. J. Geophys. Res. Biogeo. 126 (6), 1–15. https://doi.org/10.1029/2020JG005944.

Huang, Y., Wang, F., Zhang, L., Zhao, J., Zheng, H., Zhang, F., Zhang, W., 2023. Changes and net ecosystem productivity of terrestrial ecosystems and their influencing factors in China from 2000 to 2019. Front. Plant Sci. 14, 1120064. https://doi.org/10.3389/fpls.2023.1120064.

Ji, L., Zou, Y., He, K., Zhu, B., 2019. Carbon futures price forecasting based with ARIMA-CNN-LSTM model. Procedia Comput. Sci. 162, 33–38. https://doi.org/10.1016/j.procs.2019.11.087.

Jin, M., Duan, X., Zhang, Y., Xu, Q., 2025. Predicting the spatial pattern of land use change and carbon storage in Xinjiang: A Markov-FLUS-InVEST model approach. Public Library of Science ONE 20 (4), e0321929. https://doi.org/10.1371/journal.pone.0321929.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., 2020. Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. Biogeosciences 17 (5), 1343–1365. https://doi.org/10.5194/bg-17-1343-2020.

Li, H., Cao, Y., Xiao, J., Yuan, Z., Hao, Z., Bai, X., Liu, Y., 2024b. A daily gap-free normalized difference vegetation index dataset from 1981 to 2023 in China. Sci. Data 11 (1), 1–12. https://doi.org/10.1038/s41597-024-01936-3.

Li, Y.T., Yang, R.M., Zhang, X., Xu, L., Zhu, C.M., 2024a. Understanding drivers of the spatial variability of soil organic carbon in China's terrestrial ecosystems. Land Degrad. Dev. 35 (1), 308–320. https://doi.org/10.1002/ldr.4167.

Liu, S., Bian, Z., An, T., et al., 2021. Carbon pools of biomass and dead organic matter in typical forest ecosystems of Tibet: A new estimation based on the first forestry carbon sequestration monitoring undertaken in China. Land Degrad. Dev. 32 (10), 2877–2891. https://doi.org/10.1002/ldr.3828.

Liu, H., Jiang, D., Yang, X., Luo, C., 2005. Spatialization approach to 1 km grid GDP supported by remote sensing. Journal of Geographical Information Science 7, 120–123. https://doi.org/10.1109/ICMT.2011.6001941.

Liu, X., Xue, Z., Liu, J., Zhao, X., Fu, Y., Wang, R., Zhao, H., 2024. Spatio-Temporal Evolution and Multi-Scenario Modeling Based on Terrestrial Carbon Stocks in Xinjiang. Land 13 (9), 1–18. https://doi.org/10.3390/land13091537.

Mitchard, E.T., Saatchi, S.S., Baccini, A., Asner, G.P., Goetz, S.J., Harris, N.L., Brown, S., 2013. Uncertainty in the spatial distribution of tropical forest biomass: a comparison of pan-tropical maps. Carbon Balance Manag. 8, 1–13. https://doi.org/10.1186/1750-0680-8-10.

Pan, S., Pan, N., Tian, H., Friedlingstein, P., Sitch, S., Shi, H., Arora, V.K., Haverd, V., Jain, A.K., Kato, E., 2020. Evaluation of global terrestrial evapotranspiration using state-of-the-art approaches in remote sensing, machine learning and land surface modeling. Hydrol. Earth Syst. Sci. 24 (3), 1485–1509. https://doi.org/10.5194/hess-24-1485-2020.

Peng, S. (2019). 1-km monthly mean temperature dataset for china (1901-2021). A Big Earth Data Platform for Three Poles, DOI: 10.11888/Meteoro.tpdc.270961. CSTR: 18406.11.Meteoro.tpdc.270961.

Peng, S. (2020). 1-km monthly precipitation dataset for China (1901-2021). A Big Earth Data Platform for Three Poles, DOI: 10.5281/zenodo.3185722.

Piao, S., Fang, J., Ciais, P., Peylin, P., Huang, Y., Sitch, S., Wang, T., 2009. The carbon balance of terrestrial ecosystems in China. Nature 458 (7241), 1009–1013. https://doi.org/10.1038/nature07944.

Piao, S., He, Y., Wang, X., Chen, F., 2022. Estimation of China's terrestrial ecosystem carbon sink: Methods, progress and prospects. Sci. China Earth Sci. 65 (4), 641–651. https://doi.org/10.1007/s11430-021-9933-2.

Randerson, J.T., Chapin Iii, F.S., Harden, J.W., Neff, J.C., Harmon, M.E., 2002. Net ecosystem production: a comprehensive measure of net carbon accumulation by ecosystems. Ecol. Appl. 12 (4), 937–947. https://doi.org/10.1890/1051-0761(2002)012[0937:NEPACM]2.0.CO;2.

Rodda, S.R., Thumaty, K.C., Praveen, M., Jha, C.S., Dadhwal, V.K., 2021. Multi-year eddy covariance measurements of net ecosystem exchange in tropical dry deciduous forest of India. Agric. For. Meteorol. 301, 108351. https://doi.org/10.1016/j.agrformet.2021.108351.

Roy, S., Bhunia, G.S., Shit, P.K., 2021. Spatial prediction of COVID-19 epidemic using ARIMA techniques in India. Model. Earth Syst. Environ. 7, 1385–1391. https://doi.org/10.1007/s40808-021-01231-y.

Şaylan, L., Kimura, R., Başakin, E.E., Kurosaki, Y., 2024. Variables controlling growing season carbon dynamics in desert shrub ecosystem. Theor. Appl. Climatol. 155 (6), 4797–4812. https://doi.org/10.1007/s00704-023-04474-6.

Schimel, D., Melillo, J., Tian, H., et al., 2001. Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. Nature 414 (6860), 169–172. https://doi.org/10.1038/35102500.

Shangguan, W., Li, Q., Shi, G., 2022. China Soil Moisture Dataset (2000-2020). A Big Earth Data Platform for Three Poles. https://doi.org/10.11888/Terre.tpdc.272415. CSTR: 18406.11.Terre.tpdc.272415.

Shen, Q., Chen, Q., Tang, B.-S., Yeung, S., Hu, Y., Cheung, G., 2009. A system dynamics model for the sustainable land use planning and development. Habitat Int. 33 (1), 15–25. https://doi.org/10.1016/j.habitatint.2008.04.007.

Song, J., He, X., Zhang, F., Ma, X., Jim, C.Y., Johnson, B.A., Chan, N.W., 2024. Analyzing and Predicting LUCC and Carbon Storage Changes in Xinjiang's Arid Ecosystems Under the Carbon Neutrality Goal. Remote Sens. (Basel) 16 (23), 1–16. https://doi.org/10.3390/rs16234715.

Spawn, S.A., and H.K. Gibbs. 2020. Global Aboveground and Belowground Biomass Carbon Density Maps for the Year 2010. ORNL DAAC,Oak Ridge, Tennessee, USA. doi: 10.3334/ORNLDAAC/1763 doi: 10.3334/ORNLDAAC/1763.

Tian, Z. Y., Gao, F., He, B., & Han, F. H. (2024). Land use Change and Its Impact on Carbon Stock in the Tarim River Basin from 1990 to 2020. Huan jing ke xue= Huanjing kexue, 45(11), 6514-6526. https://doi: 10.13227/j.hjkx.202311201.

Wang, J., Feng, L., Palmer, P.I., Liu, Y., Fang, S., Bösch, H., Xia, C., 2020. Large Chinese land carbon sink estimated from atmospheric carbon dioxide data. Nature 586 (7831), 720–723. https://doi.org/10.1038/s41586-020-2747-7.

Wang, X., Jiao, H., 2020. Spatial scaling of forest aboveground biomass using multi-source remote sensing data. IEEE Access 8, 178870–178885. https://doi.org/10.1109/ACCESS.2020.3026300.

Wang, Z., Li, X., Mao, Y., Li, L., Wang, X., Lin, Q., 2022. Dynamic simulation of land use change and assessment of carbon storage based on climate change scenarios at the city level: A case study of Bortala, China. Ecol. Indic. 134, 108499. https://doi.org/10.1016/j.ecolind.2021.108499.

Wang, X., Tan, K., Chen, B., Du, P., 2017. Assessing the spatiotemporal variation and impact factors of net primary productivity in China. Sci. Rep. 7 (1), 44415. https://doi.org/10.1038/srep44415.

Xiaohong, J., Jia, L., Jingyun, Y., Bingqiang, F., Fang, B., Xiaotian, X., Bo, W., 2023. Estimating carbon storage of desert ecosystems in China. Int. J. Digital Earth 16 (2), 4113–4125. https://doi.org/10.1080/17538947.2023.2257416.

Xu, L., He, N., Yu, G., 2019. A dataset of carbon density in Chinese terrestrial ecosystems (2010s). China Science Data 4 (1), 86–92.

Xu, L., Du, H., Zhang, X., 2021. Driving forces of carbon dioxide emissions in China's cities: An empirical analysis based on the geodetector method. J. Clean. Prod. 287, 125169. https://doi.org/10.1016/j.jclepro.2020.125169.

Xu, M., Guo, B., Zhang, R., 2024. Variations of terrestrial net ecosystem productivity in China driven by climate change and human activity from 2010 to 2020. Forests 15 (9), 1484. https://doi.org/10.3390/f15091484.

Xu, X., Liu, J., Zhang, S., Li, R., Yan, C., Wu, S., 2018. Remote Sensing Monitoring Dataset of Land Use in Different Periods in China (CNLUCC). Resource and Environment Science and Data Center, Chinese Academy of Sciences. https://doi.org/10.12078/2018070201.

Yan, A., Li, B., Huang, F., Zhang, W., Jiang, P., Sheng, J., 2019. Distribution and storage of soil organic and inorganic carbon under different ecological zones in Xinjiang, China. Int. J. Agric. Biol. Eng. 12 (1), 116–125. https://doi.org/10.25165/j.ijabe.20191201.3872.

Yang, X., Huang, Y., Dong, P., Jiang, D., Liu, H., 2009. An updating system for the gridded population database of China based on remote sensing. GIS and spatial database technologies. Sensors 9 (02), 1128–1140. https://doi.org/10.3390/s90201128.

Yang, H., Li, S., Chen, J., Zhang, X., Xu, S., 2017. The standardization and harmonization of land cover classification systems towards harmonized datasets: A review. ISPRS Int. J. Geo Inf. 6 (5), 154. https://doi.org/10.3390/ijgi6050154.

Yang, S.F., Zan, M., Yuan, R.L., Chen, Z.Z., Kong, J.J., Xue, C., Zhai, L.L., 2025. Carbon Stock Changes and Forecasting in Xinjiang Based on PLUS and InVEST Model Approach. Huan jing ke xue=. Huanjing Kexue 46 (1), 378–387.

Zhang, Q., Chen, Y., Li, Z., Sun, C., Xiang, Y., Liu, Z., 2023b. Spatio-temporal development of vegetation carbon sinks and sources in the arid region of Northwest China. Int. J. Environ. Res. Public Health 20 (4), 1–18. https://doi.org/10.3390/ijerph20043257.

Zhang, F., Tian, X., Zhang, H., Jiang, M., 2022. Estimation of aboveground carbon density of forests using deep learning and multisource remote sensing. Remote Sens. (Basel) 14 (13), 3022. https://doi.org/10.3390/rs14133022.

Zhang, D., Zhao, Y., Wu, J., 2023a. Assessment of carbon balance attribution and carbon storage potential in China's terrestrial ecosystem. Resour. Conserv. Recycl. 189, 106748. https://doi.org/10.1016/j.resconrec.2023.106748.

Zhang, Y., Zheng, J., Zhang, J., Mu, C., Han, W., Liu, L., 2024. Characteristic Analysis of Carbon Sink Capacity Changes in Xinjiang's Terrestrial Ecosystem Based on EEMD. Sustainability 16 (6), 1–20. https://doi.org/10.3390/su16062542.

Zheng, H., Zheng, H., 2023. Assessment and prediction of carbon storage based on land use/land cover dynamics in the coastal area of Shandong Province. Ecol. Ind. 153, 110474. https://doi.org/10.1016/j.ecolind.2023.110474.