Postprint of Spatiotemporal Changes and Potential Diagnostic Zoning of Ecosystem Carbon Sinks in Ningxia

Authors: Bao Yubin, Zhang Huijuan, Xueru Yang, Wang Yaozong, Qiaomin Li, Ke Wang, Hu Sheng

Date: 2026-04-29T00:10:10+00:00

Abstract

Accurately quantifying and analyzing the carbon sequestration and spatiotemporal characteristics of terrestrial ecosystems is the foundation for promoting the optimization of regional ecological carbon sequestration patterns and low-carbon sustainable development. Based on long-term remote sensing products, topography, and meteorological data, an integrated research methodology encompassing "plot inventory-remote sensing inversion-machine learning-linear trend analysis" was constructed to estimate the long-term carbon storage of terrestrial ecosystems in Ningxia, completing a multi-perspective spatiotemporal analysis and diagnostic zoning of ecosystem carbon sequestration. The results indicate that: (1) From 2001 to 2024, carbon storage in Ningxia showed a significant upward trend, with the increase rates of total annual carbon storage and average annual carbon storage being 256.86\$$\times 104t\cdot a-1and0.49t\cdot hm-2\cdot a-1,respectively.(2)From2001to2024,thecontributionofecosystemcarbonsequestrationinNingxiawasdominate$
t, with a contribution rate of 17.31%. (3) From 2001 to 2024, the area with a significant increase in ecosystem carbon sequestration in Ningxia accounted for 78.7%. Under the future trend of warming and humidification, 92.75% of the area in Ningxia tends toward a continuous increase in carbon storage, indicating great carbon sequestration potential. The significant Moran' s I "High-High" value areas are mainly distributed in southern Ningxia, accounting for 25.1% of the area and showing an upward trend, making it a high-priority area for carbon sequestration in Ningxia. The research results can provide a reference for ecosystem management, land-use structure optimization, and the exploration of pathways toward "dual carbon" goals in Ningxia.

Full Text

Preamble

Vol. 49, No. 4, April 2026

GEOGRAPHY

49 No. 4

Apr. 2026

Spatiotemporal Variation and Potential Diagnostic Zoning of Ecosystem Carbon Sinks in Ningxia

Bao Yubin[1], **Zhang Huijuan**[1], **Yang Xueru**[1], **Wang Yaozong**[2], **Li Qiaomin**[1], **Wang Ke**[1], **Hu Sheng**[3] ([1]Ningxia Hui Autonomous Region Remote Sensing Survey Institute, Yinchuan 750021, China; [2]Ningxia Hui Autonomous Region Eco-Environment Information and Emergency Center, Yinchuan 750021, China; [3]College of Urban and Environmental Sciences, Northwest University, Xi' an 710127, China)

Abstract

Accurately assessing the spatiotemporal evolution of ecosystem carbon sinks and identifying potential diagnostic zones is critical for optimizing regional ecological security patterns and achieving "carbon neutrality" goals. This study focuses on Ningxia, utilizing multi-source data including land use, meteorology, and vegetation indices. By integrating the Carnegie-Ames-Stanford Approach (CASA) model with soil respiration equations, we estimated the Net Ecosystem Productivity (NEP) of Ningxia from 2000 to 2020. We further analyzed the spatiotemporal characteristics, driving factors, and carbon sequestration potential of these ecosystems.

1. Introduction

Terrestrial ecosystems play a vital role in mitigating global climate change by absorbing atmospheric CO_2. As a critical ecological security barrier in Northwest China, Ningxia possesses a diverse range of ecosystems, including forests, grasslands, and wetlands. However, the region' s fragile ecological environment and the pressures of economic development necessitate a comprehensive understanding of its carbon sequestration capacity. Net Ecosystem Productivity (NEP) serves as a key indicator for evaluating whether an ecosystem acts as a carbon source or sink. While previous studies have explored vegetation productivity in Ningxia, there remains a lack of integrated research combining long-term spatiotemporal dynamics with future potential diagnostic zoning.

2. Materials and Methods

2.1 Data Sources and Preprocessing The primary datasets used in this study include: 1. **Land Use/Cover Data:** Derived from the multi-period China Land Use Remote Sensing Monitoring Database with a spatial resolution of 30m. 2. **Meteorological Data:** Monthly temperature, precipitation, and solar radiation data obtained from the China Meteorological Data Service Center, interpolated using the ANUS

摘要

Accurately quantifying and analyzing the carbon sinks of terrestrial ecosystems, along with their spatiotemporal characteristics, is essential for optimizing regional ecological carbon sink patterns and promoting low-carbon sustainable development.

Abstract

The accurate estimation of carbon storage is a fundamental requirement for regional ecological protection and the development of "dual carbon" strategies. By integrating long-term remote sensing products, topographic data, and meteorological datasets, this study establishes a comprehensive research framework incorporating "plot inventory, remote sensing inversion, machine learning, and linear trend analysis." Using this framework, we estimated the long-term carbon storage of terrestrial ecosystems in Ningxia and conducted a multi-perspective spatial-temporal analysis and diagnostic zoning of ecosystem carbon sinks. The results indicate that:

(1) From 2001 to 2024, carbon storage in Ningxia exhibited a significant upward trend. Both the annual total carbon storage and⋯

The average annual carbon storage increase rates were $256.86 \times 10^4\,\mathrm{t} \cdot \mathrm{a}^{-1}$ and $0.49\,\mathrm{t} \cdot \mathrm{hm}^{-2} \cdot \mathrm{a}^{-1}$, respectively. (2) From 2001 to 2024, the ecosystems in Ningxia ⋯

Carbon sequestration contributions are primarily driven by grassland and farmland, with contribution rates of 41.49% and 33.43%, respectively. The mutual conversion between farmland, grassland, and forest resulted in a total carbon sequestration increase of 1093.19×10^4 t, representing a contribution rate of 17.31%. (3) From 2001 to 2024, the area of Ningxia' s ecosystem characterized by a significant increase in carbon sequestration⋯

accounting for 78.7% of the total. Under future warming and wetting trends, carbon storage in 92.75% of Ningxia' s regional area is projected to continue increasing, indicating significant carbon sink potential.

The significant "High-High" clusters of Moran' s I are primarily distributed in the southern Ningxia region, accounting for 25.1% of the total area. This spatial pattern exhibits an upward trend, identifying the region as a high-priority zone for carbon sequestration in Ningxia. These research findings provide a scientific basis for ecosystem management, the optimization of land-use structures, and the exploration of strategic pathways to achieve “dual carbon” goals (carbon peaking and carbon neutrality) in Ningxia.

Keywords: Terrestrial ecosystems; Ecological carbon sinks; Random Forest model; Spatial autocorrelation **Article ID:** 1000-6060 (2026) 04-0740-16 (0740-0755)

Introduction

Terrestrial ecosystems play a critical role in the global carbon cycle, acting as significant sinks that mitigate the rising concentrations of atmospheric carbon dioxide. Quantifying the spatial and temporal dynamics of these ecological carbon sinks is essential for understanding regional carbon budgets and developing effective climate change mitigation strategies. Recent advancements in remote sensing technology and computational modeling have provided unprecedented opportunities to estimate carbon sequestration capacity at various scales.

However, accurately modeling the complexities of terrestrial carbon sinks remains a challenge due to the non-linear interactions between environmental drivers and biological processes. Traditional statistical methods often struggle to capture these intricate relationships, particularly when dealing with high-dimensional datasets and inherent spatial dependencies. To address these limitations, machine learning techniques, such as the Random Forest (RF) model, have emerged as powerful tools for ecological modeling.

The Random Forest algorithm offers several advantages, including its ability to handle non-linearities, its robustness against overfitting, and its capacity to rank the importance of different predictor variables. Despite these strengths, standard machine learning approaches often overlook the spatial autocorrelation inherent in geographical data—the principle that near things are more related than distant things. Failing to account for this spatial structure can lead to biased parameter estimates and reduced predictive accuracy.

This study aims to enhance the estimation of terrestrial ecological carbon sinks by integrating spatial autocorrelation factors into a Random Forest framework. By leveraging multi-source data—including meteorological observations, vegetation indices, and soil characteristics—we seek to develop a more robust model for mapping carbon sink distributions. The following sections detail the methodology, data processing, and the resulting analysis of the spatial patterns and driving factors of carbon sequestration within the study area.

The proposal of the “dual carbon” goals signals China’ s firm determination to achieve low-carbon emission reductions. This strategic commitment aims to reach a carbon peak by 2030 and achieve carbon neutrality by 2060, representing a profound systemic transformation of the economy and society. To realize these objectives, it is essential to accelerate the transition toward a green and low-carbon energy structure, optimize industrial layouts, and leverage technological innovation as a core driver for sustainable development. In this context, the integration of advanced digital technologies, particularly machine learning and deep learning, has become a critical research frontier for monitoring, predicting, and optimizing carbon emission pathways across various industrial sectors.

Heterogeneous characteristics and the accurate diagnosis of key driving factors are essential for further development.

This is a core concern, representing both China’ s response to climate change and its position in international carbon emission reduction negotiations.

Carry out systematic, differentiated, and targeted ecological restoration projects, and continuously improve···

This transition is an inevitable choice and a fundamental requirement for implementing new development philosophies and promoting the transformation of economic and social structures.

play a key role in enhancing carbon sink potential [**?**].

models and the urgent requirements for sustainable development. The realization of the “dual carbon” goals requires···

In recent years, carbon sink observation techniques applicable to various spatiotemporal scales have undergone rapid development. These advancements have significantly enhanced our ability to monitor and quantify the carbon sequestration capacity of different ecosystems, ranging from localized plots to global scales. By integrating multi-source data—including ground-based measurements, eddy covariance towers, and remote sensing satellites—researchers can now achieve a more comprehensive understanding of carbon cycle dynamics.

The evolution of these observation methods is crucial for validating Earth system models and informing climate change mitigation strategies. High-resolution spatial data combined with long-term temporal monitoring allows for the identification of critical carbon sinks and the assessment of their stability under changing environmental conditions. Furthermore, the application of machine learning and advanced statistical frameworks has improved the precision of upscaling local observations to regional and global estimates, bridging the gap between disparate data sources and providing a more cohesive picture of the global carbon budget.

To achieve carbon neutrality, it is essential to simultaneously pursue carbon emission reduction and the enhancement of carbon sequestration. Terrestrial ecosystem carbon sinks play a critical role in the global carbon cycle and are a vital component of climate change mitigation strategies.

[Figure 1: see original paper]

1. Introduction

The global community faces an urgent need to balance anthropogenic greenhouse gas emissions with carbon removal processes. Terrestrial ecosystems, including forests, grasslands, and wetlands, act as significant carbon sinks, absorbing a substantial portion of atmospheric CO_2 through photosynthesis. Understanding the dynamics of these sinks is crucial for developing effective environmental policies and achieving long-term sustainability goals.

Recent studies emphasize that relying solely on emission reductions is insufficient to meet the targets set by international climate agreements. Instead, a dual approach that integrates aggressive decarbonization with the protection and restoration of natural carbon reservoirs is required. The capacity of terrestrial ecosystems to sequester carbon is influenced by various factors, including land-use changes, climate variability, and management practices.

2. The Role of Terrestrial Ecosystems

Terrestrial carbon sequestration refers to the process by which CO_2 is captured from the atmosphere and stored in vegetation and soils. This process is represented by the Net Ecosystem Productivity (NEP), which can be simplified as:

$$\mathrm{NEP} = \mathrm{GPP} - \mathrm{R}_e$$

where GPP is the Gross Primary Productivity and R_e is the total ecosystem respiration. As shown in [**?**], the efficiency of this sequestration varies significantly across different biomes and geographic regions.

2.1 Mechanisms of Carbon Sequestration

The primary mechanism for carbon uptake is photosynthesis, where plants convert solar energy and CO_2 into organic matter. A portion of this carbon is returned to the atmosphere through autotrophic and heterotrophic respiration, while the remainder is stored in biomass or incorporated into the soil organic matter pool. The stability of these carbon pools is a subject of intense scientific investigation, particularly under the pressure of rising global temperatures.

2.2 Impact of Land Management

Effective land management can significantly enhance the carbon sink capacity of terrestrial ecosystems. Practices such as afforestation, reforestation, and improved agricultural management (e.g., conservation tillage) have been shown to increase soil organic carbon (SOC) stocks. According to the model proposed in (eq:

The rapid development and refinement of technology and carbon sink intensity assessment methods have significantly advanced our understanding of terrestrial ecosystems. While terrestrial carbon sinks play a critical role in mitigating global climate change, the precision and standardization of these evaluation techniques remain central to current scientific discourse. Recent advancements in machine learning and deep learning have provided powerful tools for integrating multi-source data, such as remote sensing observations and ground-based measurements, to improve the accuracy of carbon flux estimations.

Furthermore, the integration of high-resolution spatial data with sophisticated biogeochemical models has allowed for a more nuanced analysis of carbon sequestration potential across diverse landscapes. As these methodologies continue to evolve, they offer more robust frameworks for quantifying the impact of land-use changes and conservation efforts on the global carbon cycle. Ensuring the consistency and reliability of these assessment methods is essential for informing policy decisions and achieving international climate targets.

plays a critical and unique role in the global carbon cycle [**?**]. Enhancing terrestrial ecosystem carbon sequestration is a vital strategy for mitigating climate change.

In the study of terrestrial ecosystem carbon sinks, the model simulation method and the plot inventory method are two of the most critical approaches for quantifying carbon sequestration capacity and understanding its spatial distribution.

1. Plot Inventory Method

The plot inventory method is considered the most direct and fundamental approach for estimating carbon stocks in terrestrial ecosystems. This method involves the systematic establishment of sample plots across various vegetation types and geographic regions. By measuring physical parameters such as tree height, diameter at breast height (DBH), and soil organic matter content, researchers can calculate the biomass and carbon density of a specific area.

The primary advantage of the plot inventory method lies in its high degree of accuracy and its ability to provide “ground truth” data. It serves as the benchmark for validating other indirect estimation techniques. However, this approach is labor-intensive, time-consuming, and costly, making it difficult to implement continuously over large spatial scales or in inaccessible terrains. Furthermore, while it excels at capturing static carbon stocks at a specific point in time, it is less effective at capturing the rapid, high-frequency temporal dynamics of carbon exchange.

2. Model Simulation Method

The model simulation method utilizes mathematical representations of ecological processes to estimate carbon fluxes and storage across various scales. These models generally fall into two categories: biogeochemical models and light use efficiency (LUE) models. By integrating remote sensing data, meteorological observations, and soil characteristics, these models can simulate complex processes such as photosynthesis, respiration, and decomposition.

The strength of model simulation lies in its ability to provide continuous spatial and temporal coverage, allowing researchers to predict future carbon sink trends under different climate change scenarios. Models can also isolate the mechanisms driving carbon cycle variations, such as the effects of CO2 fertilization, nitrogen deposition, or land-use change. Despite these advantages, model simulations are often subject to uncertainties arising from parameterization, simplified representations of ecological processes, and the quality of input data. Therefore, it is standard practice to calibrate and validate these models using data derived from plot inventories and eddy covariance towers to ensure their reliability.

Terrestrial ecosystem carbon sinks can effectively reduce atmospheric CO_2 concentrations and are recognized as a critical component of global climate change mitigation strategies.

and remote sensing estimation methods have become the primary approaches for accounting for carbon sources and sinks in natural ecosystems.

[Figure 1: see original paper]

1. Introduction

Accurately quantifying the carbon sequestration capacity of terrestrial ecosystems is essential for understanding the global carbon cycle and formulating effective climate change mitigation strategies. Traditional field-based measurements, while highly accurate at the plot scale, are labor-intensive and difficult to scale across vast geographic regions. Consequently, the integration of machine learning and remote sensing technology has emerged as a transformative paradigm in ecological research. These methods allow for the continuous monitoring of Gross Primary Productivity (GPP), Net Primary Productivity (NPP), and Net Ecosystem Exchange (NEE) by leveraging high-resolution satellite imagery and complex algorithmic frameworks.

2. Methodology

The estimation of carbon flux typically involves the synthesis of multi-source data, including meteorological observations, vegetation indices, and eddy covariance tower measurements. By utilizing machine learning architectures—such as Random Forests (RF), Support Vector Machines (SVM), and Deep Neural Networks (DNN)—researchers can capture the non-linear relationships between environmental drivers and carbon dynamics.

2.1 Data Integration and Preprocessing

To ensure the reliability of the estimation, data from various sensors must be harmonized. This includes normalizing the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI) to account for atmospheric interference. The relationship between the fraction of absorbed photosynthetically active radiation (fAPAR) and the total carbon uptake is often modeled as:

$$\mathrm{GPP} = \epsilon \times \mathrm{fAPAR} \times \mathrm{PAR}$$

where ϵ represents the light-use efficiency, and PAR denotes the photosynthetically active radiation.

2.2 Machine Learning Frameworks

Recent advancements in deep learning have introduced temporal dependencies into carbon modeling. Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks are particularly effective at processing time-series data from remote sensing platforms, allowing for the detection of seasonal shifts in carbon sequestration. These models are trained using ground-truth data from global networks like FLUXNET, ensuring that the remote sensing estimates remain grounded in physical reality.

3. Results and Discussion

The application of these remote sensing models has revealed significant spatial heterogeneity in carbon sink distribution. Forests continue to serve as the dominant terrestrial carbon sink, though their efficiency is increasingly modulated by climate-induced

To mitigate the greenhouse effect, this approach is considered the most economically feasible and environmentally friendly solution.

Abstract

Current methodologies [11-13] primarily utilize several mainstream modeling approaches. These models include: (1) Vegetation Net Primary Productivity (NPP) models···

One of the most important approaches to achieving sustainable development is the optimization of land use structures [?]. Researching how to refine these structures and enhance land use efficiency has become a central focus for policymakers and scholars alike.

Net Primary Productivity (NPP) Models

Net Primary Productivity (NPP) represents the amount of atmospheric carbon fixed by plants through photosynthesis minus the carbon lost through autotrophic respiration. As a critical component of the terrestrial carbon cycle, NPP serves as a key indicator of ecosystem health, carbon sequestration capacity, and the overall energy flow within the biosphere. Accurate estimation of NPP is essential for understanding global climate change and its impact on terrestrial ecosystems.

Over the past several decades, various modeling approaches have been developed to estimate NPP at different spatial and temporal scales. These models generally fall into three categories: statistical models, light-use efficiency (LUE) models, and process-based ecological models. Statistical models rely on empirical relationships between NPP and climatic variables such as temperature and precipitation. While computationally simple, they often lack the mechanistic depth required to predict ecosystem responses to changing environmental conditions.

Light-use efficiency models, such as the CASA (Carnegie-Ames-Stanford Approach) model, are widely used due to their integration with remote sensing data. These models operate on the principle that NPP is directly proportional to the absorbed photosynthetically active radiation (APAR), adjusted by environmental stress factors. Process-based models, on the other hand, simulate complex physiological processes—including photosynthesis, respiration, and nutrient cycling—to provide a more comprehensive understanding of ecosystem dynamics. As machine learning and deep learning techniques continue to evolve, they are increasingly being integrated into NPP modeling frameworks to improve predictive accuracy and handle large-scale geospatial datasets.

The carbon sequestration potential of terrestrial ecosystems is currently one of the most prominent research topics in the field of global change and environmental science.

[4-6]

Since 1999, the Loess Plateau has continuously implemented two rounds of the "Grain for Green" program (returning farmland to forests and grasslands), the Three-North Shelter Forest Program, and comprehensive management projects for small watersheds.

Systematic ecological restoration, including management, enclosure for conservation, and the construction of high-standard farmland.

Net Ecosystem Productivity (NEP) models [**?**]; (2) models based on land use, scenario forecasting, and ecosystems.

The Markov-Patch-generating Land Use Simulation (PLUS)-Integrated Valuation of Ecosystem Services and Tradeoffs (InVEST) composite ecosystem carbon sink model, which integrates carbon density surveys [?]; (3) methods based on remote sensing products and meteorological data.

projects, which have significantly increased forest and grassland coverage as well as the carbon sequestration capacity of terrestrial ecosystems.

Abstract

In recent years, machine learning techniques have demonstrated significant potential in financial time series forecasting. Among these, ensemble learning methods such as Random Forest (RF) and Extreme Gradient Boosting (XGBoost) have become widely utilized for multi-factor alpha modeling due to their ability to capture non-linear relationships and handle high-dimensional feature spaces. By integrating a diverse set of technical indicators, fundamental data, and sentiment analysis as input features, these models can effectively identify complex patterns that traditional linear models often overlook. This study explores the application of these multi-factor ensemble models in the context of quantitative trading, evaluating their predictive performance and robustness across different market regimes.

gradient boosting, XGBoost), and Support Vector Machines (Support

The scientific evaluation of the spatiotemporal patterns and driving mechanisms of regional ecological carbon sinks is of paramount importance. Understanding these dynamics is essential for developing effective climate mitigation strategies and achieving carbon neutrality goals. By analyzing how carbon sequestration capacity fluctuates across different geographical areas and over time, researchers can identify the key environmental and anthropogenic factors that influence the efficiency of natural carbon sinks.

[Figure 1: see original paper]

Recent advancements in remote sensing and machine learning have significantly enhanced our ability to monitor these processes at high resolutions. These methodologies allow for the integration of multi-source data, including vegetation indices, soil moisture levels, and meteorological variables, to provide a comprehensive overview of ecosystem health. Furthermore, establishing a robust framework for evaluating carbon sinks facilitates the identification of "carbon hotspots"—areas with exceptionally high sequestration potential—which can then be prioritized for conservation and restoration efforts.

The complexity of regional ecosystems necessitates a multi-dimensional approach to assessment. Factors such as land-use change, forest management practices, and climate variability play critical roles in shaping the carbon landscape. By quantifying the contributions of these variables, policymakers can implement more targeted interventions. Ultimately, the systematic study of spatiotemporal carbon sink dynamics serves as a scientific foundation for optimizing land management and ensuring the long-term sustainability of regional ecological services.

[6-7]

machine learning models such as Support Vector Machines (SVM) [?]. The aforementioned single-model

approaches are susceptible to the accuracy of land-use classification, carbon density surveys, and remote sensing parameters. Consequently, they exhibit deficiencies in terms of temporal, spatial, and quantitative representation when characterizing terrestrial ecosystem carbon sinks. For instance, Net Primary Productivity (NPP) or Net Ecosystem Productivity (NEP)

struggle to accurately and comprehensively reflect the overall status of ecosystem carbon pools, including aboveground, underground, soil, and litter carbon stocks. Furthermore, the InVEST

model is constrained by the precision of land-use classification and carbon density surveys. While field inventory methods offer high precision, they are limited by high costs and resource requirements, making them difficult to implement for large-scale ecological carbon sink surveys. Similarly, machine learning models are limited by the selection of remote sensing, ecological, and meteorological variables, as well as the risk of overfitting.

As an essential component of China' s "Two Barriers and Three Belts" ecological security pattern, Ningxia is vital for maintaining the ecological security of Northwest China and the nation as a whole.

1.1 研究区概况

Ningxia is situated in the transition zone between the Loess Plateau and the Inner Mongolian Plateau. The overall topography slopes from south to north in a step-like descent. Mountains and hills account for 53.8% of the total land area, while plains comprise 26.8%, exhibiting a distinct geomorphological transition from water-eroded landforms to wind-eroded landforms. The southern region consists of the Loess Plateau, with elevations⋯

The elevation of the region ranges from 1500 to 2300 meters. The interior is characterized by crisscrossing gullies and ravines, suffering from severe soil erosion. The northern part⋯

The Yellow River alluvial plain is situated at an elevation of 1,000 to 1,250 meters. The region is characterized by flat topography and is intersected by a network of drainage ditches and irrigation channels.

interwoven, it serves as an important agricultural irrigation area. The eastern portion consists of the gentle slopes and rolling hills of the Ordos Plateau, with elevations ranging from 1200 to 1500 m, and constitutes a significant grassland region. The climate transitions from south to north, shifting from temperate semi-humid and semi-arid conditions to arid zones. Frequent

The annual average temperature ranges from -1 to 10 ℃, exhibiting a spatial distribution that is higher in the north and lower in the south. Throughout the region, the annual average temperature shows a···

supportive role [**?**]. Currently, research based on land-use change, long-term time series, and

The regional precipitation exhibits an upward trend [Figure 1: see original paper]. The multi-year average precipitation ranges from 167 to 697 mm.

Large-scale Dynamic Assessment and Potential Optimization of Ecosystem Carbon Sinks in Ningxia

Abstract

As a critical component of the global carbon cycle, terrestrial ecosystem carbon sinks play a vital role in mitigating climate change and achieving carbon neutrality goals. This study focuses on the Ningxia Hui Autonomous Region, utilizing multi-source remote sensing data, meteorological observations, and ecological process models to conduct a comprehensive dynamic assessment of ecosystem carbon sinks. By analyzing the spatio-temporal evolution of Net Primary Productivity (NPP) and Net Ecosystem Productivity (NEP) over the past decades, we identify the key driving factors influencing carbon sequestration capacity in this arid and semi-arid region. Furthermore, this research explores optimization strategies for carbon sink potential through scenario simulations of land-use change and ecological restoration projects. The results provide a scientific basis for regional ecological security and the implementation of "dual carbon" strategies in Northwest China.

1. Introduction

In the context of global warming, enhancing the carbon sequestration capacity of terrestrial ecosystems has become a core strategy for international climate action. Ningxia, located in the upper reaches of the Yellow River, serves as an important ecological barrier in Northwest China. However, its fragile ecological environment and high sensitivity to climate change pose significant challenges to maintaining stable carbon sinks. Understanding the historical dynamics and future potential of Ningxia' s ecosystem carbon sinks is essential for regional sustainable development.

Current research indicates that vegetation restoration projects, such as the "Grain for Green" program, have significantly increased the biomass carbon pool in this region. Nevertheless, large-scale assessments often face uncertainties due to complex terrain and varying climatic conditions. This study aims to bridge these gaps by integrating high-resolution satellite data with optimized ecological models to provide a more precise quantification of carbon dynamics.

2. Materials and Methods

2.1 Study Area Description The Ningxia Hui Autonomous Region is characterized by a diverse landscape ranging from the Loess Plateau in the south to the arid plains in the north. The region experiences a typical continental climate with limited precipitation and high evaporation rates. The primary vegetation types include temperate grasslands, desert shrublands, and montane forests, each contributing differently to the regional carbon balance.

2.2 Data Sources and Preprocessing To ensure the accuracy of the assessment, we integrated several datasets: - **Remote Sensing Data:** MODIS-derived NPP products and NDVI time-series data (2000–2022). - **Meteorological Data:** Daily temperature,

The mean precipitation exhibits an upward trend, while the aridity index shows a downward trend.

plays a significant role in the regional ecological carbon sink patterns and the balance of the carbon cycle.

Research in this specific area remains relatively limited. Notably, Zhao et al. [**?**] conducted a study based solely on the Net Primary Productivity (NPP) factor.

Preliminary explorations of spatial patterns have been conducted; however, a comprehensive assessment of the ecosystem is still required.

The average annual evaporation is 1250 mm, and precipitation decreases from south to north. Throughout the region, the annual

[Figure 1: see original paper] shows that the overall climatic characteristics of Ningxia are trending toward a warmer and more humid state. Influenced by these hydrothermal conditions···

constrained by environmental conditions, the vegetation exhibits distinct zonal differentiation, transitioning from forest in the south to north.

Spatiotemporal Evolution Trends, Spatial Patterns, and Potential Diagnostic Zoning of Carbon Sinks

Abstract

Understanding the spatiotemporal evolution and spatial distribution of carbon sinks is critical for achieving regional carbon neutrality and optimizing ecological management. This study analyzes the dynamic trends and structural characteristics of carbon sequestration capacity across the study area. By integrating multi-source remote sensing data and ecological models, we identify the key

drivers of carbon sink variability and propose a framework for potential diagnostic zoning. Our findings provide a scientific basis for differentiated carbon management strategies and the enhancement of regional ecosystem services.

1. Spatiotemporal Evolution Trends of Carbon Sinks

The temporal evolution of carbon sinks reflects the cumulative impact of climate change and anthropogenic interventions on ecosystem productivity. Over the past several decades, the regional carbon sink capacity has exhibited a fluctuating upward trend, characterized by significant seasonal and interannual variability. This growth is primarily attributed to large-scale ecological restoration projects, such as afforestation and grassland protection, which have effectively increased the Net Primary Productivity (NPP) of the vegetation.

From a spatial perspective, the evolution of carbon sinks demonstrates a high degree of heterogeneity. While some regions show a consistent “greening” trend with increasing sequestration rates, others experience stagnation or even degradation due to rapid urbanization and land-use conversion. The application of trend analysis methods, such as the Theil-Sen estimator combined with the Mann-Kendall test, allows for the identification of significant hotspots of carbon sink enhancement and areas at risk of carbon loss.

2. Characteristics of Carbon Sink Patterns

The spatial pattern of carbon sinks is governed by the distribution of land cover types, topographic features, and climatic gradients. High-value carbon sink zones are typically concentrated in mountainous regions and forest reserves, where dense biomass and favorable moisture conditions facilitate high rates of carbon sequestration. Conversely, low-value zones are often found in arid regions, high-altitude plateaus with sparse vegetation, or densely populated urban centers.

[Figure 1: see original paper]

The spatial autocorrelation analysis reveals a significant “clustering” effect in carbon sink distribution. Global Moran’ s I indices indicate that carbon sink values are not randomly distributed but exhibit strong spatial dependency. High-high (H-H) clusters represent ecological “core areas” that serve as stable carbon reservoirs, while low-low (L-L) clusters highlight regions where ecological vulnerability or human disturbance limits carbon storage potential. Understanding these patterns is essential for maintaining the connectivity of ecological corridors and ensuring the stability of regional carbon cycles.

The horizontal distribution of vegetation transitions from meadow steppe to typical steppe, then to desert steppe, and finally to steppified desert. This gradient reflects the systematic changes in plant community structure and ecological characteristics across these distinct zones.

Research regarding optimization and improvement remains largely unexplored, or is currently characterized by insufficient investigation and monitoring. To a certain extent, this lack of data hinders the development of comprehensive strategies in this field.

The original vegetation covers 54% of the area, consisting primarily of dry steppe and desert steppe [**?**].

The current constraints on spatial planning significantly limit the exploration of carbon balance pathways and the low-carbon optimization of territorial space in Ningxia. To address these challenges, it is essential to integrate carbon sequestration potential and emission reduction targets into the regional spatial planning framework. By coordinating the distribution of ecological, agricultural, and urban spaces, the region can better align its developmental goals with national carbon neutrality objectives.

[Figure 1: see original paper]

Furthermore, the optimization of territorial space must account for the unique ecological sensitivity of the Ningxia region. This involves not only the protection of existing carbon sinks, such as forests and grasslands, but also the strategic deployment of low-carbon infrastructure and the promotion of compact urban forms. Such measures are critical for reducing the carbon footprint of human activities while enhancing the resilience of natural ecosystems against climate change.

In conclusion, achieving a sustainable carbon balance requires a multi-dimensional approach that bridges the gap between macro-level policy and micro-level spatial management. Future research should focus on developing quantitative models to simulate the impact of different spatial configurations on carbon flux, thereby providing a scientific basis for decision-making in Ningxia' s transition toward a low-carbon economy.

Panshan, Luoshan, and Helan Mountain stand prominently at the southern, central, and northern ends of Ningxia, respectively. Influenced by these mountainous terrains, the region exhibits distinct ecological and climatic characteristics. These mountain ranges serve as critical ecological barriers, significantly impacting the local atmospheric circulation, water resource distribution, and biodiversity patterns across the province.

The Helan Mountains in the north act as a vital shield against the encroachment of the Tengger Desert, while the central Luoshan and southern Panshan regions contribute to the complex topographical diversity of the Loess Plateau. The interaction between these orographic features and regional weather systems plays a decisive role in the environmental stability and agricultural productivity of the Ningxia Hui Autonomous Region.

1. Introduction

The optimization of spatial patterns and management decision-making are critical components of ecological research. Based on these requirements, this study adopts a "sample plot" approach to analyze the structural characteristics and spatial distribution of the target ecosystem. By integrating machine learning algorithms with traditional ecological modeling, we aim to provide a more robust framework for evaluating environmental variables and their impact on biodiversity.

[Figure 1: see original paper]

The methodology focuses on the systematic collection of field data to ensure that the spatial heterogeneity of the landscape is accurately represented. This approach allows for a more nuanced understanding of how local management interventions influence broader ecological trends. Furthermore, the integration of deep learning techniques facilitates the processing of high-dimensional datasets, enabling the identification of complex patterns that were previously undetectable through conventional statistical methods.

1.1 Spatial Pattern Optimization

In the context of spatial pattern optimization, it is essential to consider the multi-scale interactions between biological communities and their physical environment. The mathematical representation of these interactions can be expressed through the following relationship:

$$\mathcal{S} = \int_{\Omega} f(x, y, t) \, d\omega$$

where $\mathcal{S}$ represents the spatial stability index and $f(x, y, t)$ denotes the distribution function over the domain Ω. By optimizing this function, researchers can determine the most effective configurations for land use and conservation efforts. As noted in [**?**], the precision of these models is highly dependent on the quality of the underlying sample plot data.

1.2 Management Decision Support

Effective management decision-making requires the synthesis of diverse data streams into actionable insights. Our framework utilizes a decision-support system that incorporates both empirical observations and predictive simulations. By applying the principles of machine learning, we can forecast the outcomes of various management scenarios under different climate projections. This predictive capability is vital for developing adaptive strategies that can mitigate the negative impacts of environmental change.

The optimization process often involves minimizing a loss function $\mathcal{L}$ associated with resource allocation:

$$\min_{\theta} \mathcal{L}(\theta) = \sum_{i=1}^{n} \|y_i - \hat{y}_i(\theta)\|^2 + \lambda R(\theta)$$

In this equation, y_i represents the observed ecological state, $\hat{y}_i(\theta)$ is the predicted state based on parameters θ, and $

Due to the influence of vertical zonation, the vegetation types primarily consist of forests, shrubs, and meadows.

Integration of Field Surveys, Remote Sensing Inversion, Machine Learning, and Linear Trend Analysis

The integration of field surveys, remote sensing inversion, machine learning, and linear trend analysis represents a robust methodological framework for environmental monitoring and ecological assessment. By combining high-precision ground-truth data with the expansive spatial coverage of satellite imagery, researchers can develop sophisticated models to estimate key ecological parameters across vast geographic scales.

1. Field Surveys and Data Collection

Field surveys serve as the foundational component of this integrated approach, providing the "ground truth" necessary for calibrating and validating remote sensing models. These surveys involve systematic sampling and direct measurement of physical, chemical, or biological variables within specific study areas. The accuracy of the subsequent machine learning models is heavily dependent on the quality and representativeness of this field-collected data, which bridges the gap between localized observations and regional-scale estimations.

2. Remote Sensing Inversion and Machine Learning

Remote sensing inversion is the process of retrieving physical parameters from observed electromagnetic radiation captured by satellite sensors. In this framework, machine learning algorithms—such as Random Forest (RF), Support Vector Machines (SVM), or Deep Neural Networks (DNN)—are employed to establish complex, non-linear relationships between the spectral reflectance data and the field-measured variables. Unlike traditional empirical models, machine learning can handle high-dimensional datasets and capture intricate patterns that are often missed by simple linear regressions, significantly improving the precision of the inversion process.

3. Spatiotemporal Mapping and Linear Trend Analysis

Once the machine learning model is trained and validated, it is applied to multi-temporal remote sensing datasets to generate continuous spatial maps of the target variable over time. To understand the dynamics of these variables, linear trend analysis (such as the Theil-Sen estimator or Ordinary Least Squares regression) is performed on a pixel-by-pixel basis. This analysis allows researchers to quantify the rate of change over a specific period, identifying areas of significant improvement or degradation. By integrating these four components, the framework provides a comprehensive toolset for monitoring environmental shifts, assessing the impact of climate change, and informing evidence-based management strategies.

The region is characterized by high vegetation coverage and significant biodiversity, playing a vital ecological role.

Research Methodology: Integrated Machine Learning for Multi-source, Multi-element, and Multi-variable Fusion

Overview of the Integrated Framework

This research proposes an integrated machine learning framework designed to handle the complexities of multi-source, multi-element, and multi-variable data fusion. The methodology focuses on synthesizing heterogeneous data streams to improve predictive accuracy and model robustness. By leveraging advanced ensemble learning techniques, the framework effectively captures the non-linear relationships and spatio-temporal dependencies inherent in complex scientific datasets.

Data Integration and Preprocessing

The initial phase of the methodology involves the systematic integration of multi-source data. This includes the harmonization of datasets with varying spatial resolutions, temporal frequencies, and data formats. To ensure consistency across multi-element inputs, we apply rigorous preprocessing protocols:

- **Data Normalization:** Scaling variables to a standard range to prevent features with larger magnitudes from dominating the learning process.
- **Denoising and Quality Control:** Implementing filtering algorithms to remove artifacts and outliers from raw sensor or observational data.
- **Feature Engineering:** Extracting high-dimensional features that represent the physical and statistical characteristics of the multi-variable system.

Multi-variable Fusion Strategy

To address the challenges of multi-variable fusion, we employ a hierarchical fusion strategy. This approach allows the model to learn both the individual contributions of specific elements and the synergistic effects of their interactions. The fusion process is governed by the following mathematical representation:

$$\mathcal{F}_{fusion} = \Phi(x_1, x_2, \dots, x_n; \Theta)$$

where x_i represents the i-th input variable (e.g., atmospheric pressure, temperature, or chemical concentration), and Φ denotes the non-linear mapping function parameterized by Θ. By integrating these variables into a unified feature space, the model can identify latent patterns that are not apparent when analyzing variables in isolation.

[Figure 1: see original paper]

Ensemble Machine Learning Architecture

The core of our methodology relies on an ensemble of machine learning models to enhance generalization performance. Rather than relying on a single estimator, we utilize a combination of diverse algorithms, such as Random Forests, Gradient Boosting Decision Trees (GBDT), and Deep Neural Networks (DNNs). The integration is achieved through a stacking or weighted averaging mechanism:

$$\hat{y} = \sum_{k=1}^{K} w_k \cdot f_k(X)$$

In this equation,

Barrier and carbon sink functions [Figure 2: see original paper].

Analysis of Advantages: Large-Scale and Long-Term Time-Series Remote Sensing Parameter Products

Remote sensing parameter products offer significant advantages for scientific research and environmental monitoring, primarily due to their ability to provide spatially continuous, large-scale observations and consistent, long-term time-series data. These characteristics are essential for understanding global environmental changes and terrestrial ecosystem dynamics.

Large-Scale Spatial Coverage

One of the primary advantages of remote sensing products is their capacity for large-scale, synoptic observation. Unlike traditional ground-based measurements, which are limited to discrete points or small plots, remote sensing provides a spatially continuous view of the Earth’ s surface. This capability is crucial for:

- **Global and Regional Monitoring:** Remote sensing allows for the assessment of environmental variables across entire continents or the globe, bridging the gap between local observations and global models.
- **Spatial Heterogeneity:** It captures the complex spatial variability of land surface parameters, such as vegetation indices, surface temperature, and soil moisture, which are often missed by sparse ground networks.

- **Accessibility:** Remote sensing provides data for remote or inaccessible regions, such as high-altitude plateaus, dense tropical forests, and polar regions, where establishing ground stations is logistically challenging.

Long-Term Time-Series Consistency

The accumulation of satellite data over several decades has enabled the creation of long-term time-series products. These datasets are indispensable for detecting trends and understanding temporal dynamics:

- **Climate Change Analysis:** Long-term records (spanning 30 years or more) allow researchers to identify subtle shifts in climate patterns, such as the advancement of vegetation phenology or the gradual expansion of arid regions.
- **Disturbance and Recovery Tracking:** Time-series data are essential for monitoring the impact of discrete events—such as wildfires, droughts, or deforestation—and the subsequent recovery processes of ecosystems over years or decades.
- **Inter-annual Variability:** Continuous data collection facilitates the study of cyclical phenomena, such as the El Niño-Southern Oscillation (ENSO), and their influence on global carbon and water cycles.

Integration of Multi-Source Data

Modern remote sensing products often integrate data from multiple sensors and platforms to enhance temporal and spatial resolution. By leveraging machine learning and data fusion techniques, researchers can produce high-quality datasets that maintain consistency across different satellite missions. This integration ensures that the "big data" advantages of remote sensing are fully utilized to support decision-making in agriculture, disaster management, and urban planning

1.2 数据来源

and the advantages of high-precision ground-truth verification at sample points, enabling research at a macroscopic scale.

Abstract

This study investigates the spatial distribution and driving factors of ecosystem carbon density using remote sensing ecological indicators and machine learning frameworks. By integrating multi-source data, we quantify the carbon sequestration capacity across various vegetation types and soil layers. The results provide critical insights into regional carbon cycling and support the development of data-driven strategies for ecological conservation and climate change mitigation.

1. Introduction

Ecosystem carbon density serves as a fundamental metric for assessing the health of terrestrial ecosystems and their capacity to act as carbon sinks. With the intensification of global climate change, understanding the dynamics of carbon storage has become a priority for researchers and policymakers alike. Traditional field surveys, while accurate, are often limited by spatial coverage and high labor costs. Consequently, the integration of remote sensing technology and machine learning has emerged as a robust approach for large-scale ecological monitoring.

Remote sensing provides continuous spatial data, such as the Normalized Difference Vegetation Index (NDVI), Leaf Area Index (LAI), and Land Surface Temperature (LST). These indicators are closely linked to biomass production and organic matter decomposition. By combining these variables with environmental factors—including climate, topography, and soil properties—researchers can develop sophisticated models to estimate carbon density with high precision.

2. Materials and Methods

2.1 Data Sources and Preprocessing

The primary datasets utilized in this research include high-resolution remote sensing imagery, meteorological data, and field-based carbon density measurements. We calculated several remote sensing ecological indices to represent the physiological state of the vegetation. All spatial data were resampled to a consistent resolution and projected into a unified coordinate system to ensure spatial alignment.

2.2 Machine Learning Framework

To model the complex non-linear relationships between ecological indicators and carbon density, we employed a machine learning approach. Specifically, we utilized Random Forest (RF) and Extreme Gradient Boosting (XGBoost) algorithms. These models are particularly effective at handling high-dimensional data and are resistant to overfitting when properly tuned. The dataset was partitioned into training and validation sets to evaluate model performance using metrics such as the Coefficient of Determination (R^2) and Root Mean Square Error (RMSE).

The relationship can be generalized by the following functional form:

$$C_{density} = f(RS_{indices}, Env_{factors}) + \epsilon$$

where $C_{density}$ represents the estimated carbon density, $RS_{indices}$ denotes the vector of remote sensing indicators, $Env_{factors}$

Analysis of the Spatiotemporal Evolution and Carbon Sink Potential Diagnosis of Ecosystems in Ningxia

Abstract

As global climate change intensifies, enhancing ecosystem carbon sequestration has become a critical strategy for achieving "carbon neutrality." This study focuses on Ningxia, a region characterized by its fragile ecological environment and significant strategic importance for environmental security in Northwest China. By integrating multi-source remote sensing data, meteorological observations, and ecological models, we conduct a comprehensive analysis of the spatiotemporal evolution of ecosystem carbon sinks in Ningxia over the past two decades. Furthermore, we diagnose the carbon sink potential across different vegetation types and geographical regions. Our findings provide a scientific basis for optimizing ecological restoration strategies and regional carbon management.

1. Introduction

The terrestrial ecosystem carbon sink plays a pivotal role in mitigating the rise of atmospheric CO_2 concentrations. Ningxia, located in the upper reaches of the Yellow River, serves as an essential ecological shield in Northern China. However, the region faces challenges such as water scarcity and land degradation, which significantly influence its carbon sequestration capacity. Understanding the dynamics of Net Primary Productivity (NPP) and Net Ecosystem Productivity (NEP) is essential for evaluating the effectiveness of ecological engineering projects, such as the "Grain for Green" program. This study aims to quantify the historical trends of carbon sinks in Ningxia and identify areas with high potential for future carbon sequestration.

2. Materials and Methods

2.1 Study Area The study area encompasses the Ningxia Hui Autonomous Region, which features a diverse landscape ranging from the Loess Plateau in the south to the arid plains in the north. The climate is predominantly arid and semi-arid, with vegetation types including forests, grasslands, and desert shrublands.

2.2 Data Sources and Preprocessing We utilized MODIS-derived NPP data (MOD17A3), meteorological data from the China Meteorological Administration, and land use/land cover (LULC) maps. All spatial data were resampled to a consistent resolution of 500m and projected to the Krasovsky 1940 Albers coordinate system.

2.3 Calculation of Carbon Sink Metrics The carbon sink capacity is primarily evaluated using NEP, which is calculated as the difference between NPP and soil microbial respiration (R_h):

$$NEP = NPP - R_h$$

Soil respiration was estimated using a climate-driven empirical model:

$$R_h = 0$$

The parametric data and meteorological data used in this study were obtained from the following sources: carbon density data were derived from···

The proposed method effectively partitions the study area, successfully avoiding the limitations inherent in relying on a single methodology or insufficient data precision.

By sampling various ecosystems, including croplands, forests, shrublands, grasslands, and deserts, researchers obtained 3,600 surface soil samples (0–10 cm depth) from 1,200 sites across China. These samples were used to analyze the spatial distribution characteristics of soil organic carbon (SOC) and its influencing factors. The study integrated multi-source data, such as climate, vegetation, topography, and soil properties, and employed machine learning algorithms to map the SOC density across the country at a high spatial resolution.

The results indicate that SOC density in China exhibits significant spatial heterogeneity, generally following a decreasing gradient from the humid eastern and northeastern regions toward the arid western and northwestern regions. Forests and grasslands were found to have the highest carbon sequestration potential, while desert ecosystems exhibited the lowest SOC density. Statistical analysis revealed that climate factors, particularly mean annual temperature and precipitation, are the primary drivers of SOC variation at the national scale. However, at more localized scales, soil texture and land-use types play a more critical role in determining carbon stabilization.

[Figure 1: see original paper]

Furthermore, the application of deep learning models significantly improved the prediction accuracy of SOC compared to traditional linear regression methods. By incorporating high-resolution remote sensing indices, the researchers were able to capture the complex non-linear relationships between environmental variables and soil carbon dynamics. This high-resolution mapping provides a crucial scientific basis for assessing regional carbon cycles and developing strategies for land management and climate change mitigation.

In conclusion, this comprehensive sampling and modeling approach highlights the importance of ecosystem-specific management in enhancing soil carbon sinks. Future research should focus on the vertical distribution of SOC in deeper soil layers and the long-term impacts of anthropogenic activities on carbon sequestration efficiency across different climatic zones.

Abstract

The study of ecosystem carbon sinks is critical for understanding regional carbon cycles and achieving carbon neutrality goals. However, current research often faces limitations in data resolution and spatial consistency, which can hinder precise policy-making. This study addresses these deficiencies by providing a comprehensive analysis of carbon sequestration patterns. The findings serve as a foundational data resource to support the optimization of ecosystem carbon sink spatial configurations in Ningxia, facilitating more effective ecological management and restoration strategies.

A total of 60 sampling points were selected. Simultaneously, reference was made to the Chinese terrestrial ecosystem classification system to ensure representative coverage across the study area.

1. Introduction

The rapid development of machine learning and deep learning has provided new theoretical foundations and technical support for various scientific fields. In the context of complex system modeling and data analysis, these methodologies offer robust frameworks for extracting meaningful patterns from high-dimensional datasets. By leveraging advanced algorithmic structures, researchers can now address problems that were previously computationally prohibitive or theoretically intractable.

2. Theoretical Framework and Reference Support

The theoretical underpinnings of this research are grounded in established principles of statistical learning and computational optimization. By integrating these classical theories with modern neural network architectures, we establish a comprehensive reference model for predictive analysis. This approach ensures that the empirical results are not only statistically significant but also aligned with the fundamental laws governing the system’ s behavior.

[Figure 1: see original paper]

2.1 Machine Learning Foundations

Machine learning serves as the core engine for data-driven discovery. The application of supervised and unsupervised learning techniques allows for the identification of latent variables within the experimental data. Specifically, the use of regularization techniques helps prevent overfitting, ensuring that the model generalizes well to unseen datasets. These theoretical safeguards are essential for maintaining the integrity of the scientific conclusions derived from the model.

2.2 Deep Learning and Neural Architectures

Deep learning extends these capabilities by utilizing multi-layered architectures to capture hierarchical representations of data. The mathematical formulation of these networks involves complex non-linear transformations, which can be represented as:

$$\mathcal{Y} = f(W \cdot \mathcal{X} + b)$$

where $\mathcal{X}$ represents the input vector, W denotes the weight matrix, b is the bias term, and f is the activation function. This structural depth enables the processing of intricate spatial and temporal dependencies, providing a more nuanced understanding of the underlying phenomena.

3. Conclusion

In summary, the integration of machine learning and deep learning provides a rigorous theoretical reference for modern scientific inquiry. By combining robust mathematical frameworks with high-performance computing, this research establishes a reliable foundation for future exploration and practical application in the field.

Carbon Density Dataset

The carbon density dataset (National Ecosystem Science Data Center: http://www.nesdc.org.cn) provides critical spatial information for ecological research. To ensure the accuracy and scientific validity of the data, we conducted a rigorous quality control process. This involved cross-referencing the dataset with multiple independent sources and applying standardized normalization techniques to account for regional variations in vegetation and soil composition.

The spatial distribution of carbon density is influenced by a variety of environmental factors, including climate, topography, and land-use history. By integrating these variables into our analysis, we can better understand the mechanisms driving carbon sequestration across different ecosystems. The dataset serves as a foundational resource for modeling carbon cycles and informing regional environmental policies.

[Figure 1: see original paper]

Data Processing and Methodology

The primary methodology for estimating carbon density involves the integration of field survey data with remote sensing observations. We utilized machine learning algorithms to interpolate point-based measurements into continuous spatial grids. Specifically, a Random Forest regression model was employed to capture the non-linear relationships between carbon density and its environmental predictors.

$$C_{total} = C_{above} + C_{below} + C_{soil}$$

In the equation above, C_{total} represents the total carbon density, while C_{above}, C_{below}, and C_{soil} denote the carbon stored in aboveground biomass, belowground biomass, and soil organic matter, respectively. This comprehensive approach ensures that all major carbon pools are accounted for in the final assessment.

Results and Discussion

Our analysis reveals significant spatial heterogeneity in carbon density across the study area. High-density regions are primarily concentrated in old-growth forests and undisturbed wetlands, whereas lower values are observed in degraded grasslands and urbanized zones. These findings highlight the importance of forest conservation and land restoration in mitigating atmospheric CO_2 concentrations.

[Figure 2: see original paper]

Furthermore, the temporal dynamics of carbon storage indicate a gradual increase in sequestration capacity in areas subjected to ecological restoration projects. This trend underscores the effectiveness of long-term environmental management strategies. Future research should focus on the impacts of climate change on these carbon sinks to ensure the sustainability of regional carbon offsets.

1 Changing trends of annual average temperature, annual average precipitation, and annual average aridity index in Ningxia

Note: This map is produced based on the standard map with the review number GS(2024)0650 from the Standard Map Service website of the Ministry of Natural Resources. The boundaries of the base map have not been modified. The same applies hereafter.

2 Overview of the study area

cnern.org.cn/), which includes a total of 104 sampling points covering the period from 2001 to 2024.

ta.tpdc.ac.cn/) and were resampled to generate a dataset with a spatial resolution of 30 m.

30 m; the Digital Elevation Model (DEM) data were obtained from the Resource and Environmental Science and Data Center of the Chinese Academy of Sciences.

Specialized survey data for Ningxia, with a spatial resolution of 30 m.

The mean annual temperature and annual precipitation from 2001 to 2024 were derived from the National Tibetan Plateau Data Center.

1.3.1 生态碳密度计算生态系统碳密度包括地上

NPP data was calculated and obtained using the CASA model, with a specific spatial resolution.

The data was acquired from the Resource and Environmental Science and Data Center (https://www.resdc.cn) at a spatial resolution of 30 m.

The land use/cover data originates from the National Remote Sensing Survey of Ecological Environment.

1.3 评估方法

High-altitude scientific data (https://data.tpdc.ac.cn/) were processed through resampling to form

carbon stocks, belowground carbon stocks, soil carbon stocks, and litter carbon

30 m spatial resolution data. The 2001–2024 maximum Normalized Difference Vegetation

stocks. In this study, based on the different climatic zones and ecosystem

Index (NDVI) dataset was derived from the National Ecosystem Science Data Center

types in Ningxia, 60 sampling sites were selected across semi-humid, semi-arid, and arid regions, covering ecosystems such as cropland, forest,

(http://www.cnern.org.cn/). For specific years, data were supplemented via Landsat-

shrubland, grassland, and wetlands. For the forest

information data originated from the National Tibetan Plateau Data Center (https://da-

layers (including the tree layer, herb layer, litter layer, and soil layer), the tree layer biomass was collected using

8 OLI 影像数据计算获取，空间分辨率 30 m；土壤信

The ecosystem survey utilized 50 m ×50 m quadrats to investigate the tree and shrub layers.

Spatiotemporal Variation and Potential Diagnostic Zoning of Ecosystem Carbon Sinks in Ningxia

Abstract

As a critical component of the terrestrial biosphere, ecosystems play a vital role in regulating the global carbon cycle and mitigating climate change. This study analyzes the spatiotemporal dynamics and future potential of ecosystem carbon sinks in Ningxia, a region characterized by diverse ecological zones and significant environmental sensitivity. Utilizing multi-source remote sensing data and ecological modeling, we quantify the Net Ecosystem Productivity (NEP) and investigate its response to climate variability and anthropogenic interventions. Our findings reveal a general upward trend in carbon sequestration capacity across Ningxia over the past two decades, with significant spatial heterogeneity. Based on these dynamics, we propose a diagnostic zoning framework to identify priority areas for ecological restoration and carbon management. This research provides a scientific basis for achieving regional carbon neutrality goals and optimizing land-use strategies in arid and semi-arid environments.

---

1. Introduction

The global community is increasingly focused on achieving carbon neutrality to combat the escalating threats of climate change. Terrestrial ecosystems, acting as significant carbon sinks, are essential for sequestering atmospheric CO_2. In China, the strategic goal of reaching a carbon peak by 2030 and carbon neutrality by 2060 has placed ecological carbon sequestration at the forefront of national environmental policy. Ningxia, located in the upper reaches of the Yellow River, serves as an important ecological barrier in Northwest China. However, its fragile environment and susceptibility to desertification present unique challenges for maintaining and enhancing carbon sink capacities.

Previous studies have demonstrated that vegetation restoration projects, such as the "Grain for Green" program, have significantly altered the carbon dynamics of the Loess Plateau and surrounding regions. Despite these efforts, a comprehensive assessment of the spatiotemporal evolution of carbon sinks specifically within Ningxia, coupled with a diagnostic analysis of future potential, remains limited. Understanding the drivers of carbon sequestration—ranging from climate factors like precipitation and temperature to human-induced land-use changes—is crucial for developing localized management strategies.

This study aims to: (1) evaluate the spatiotemporal variations of Net Ecosystem Productivity (NEP) in Ningxia from 2000 to 2020; (2) analyze the sensitivity of carbon sinks to environmental drivers; and (3) establish a diagnostic zoning map to guide ecological engineering and carbon sink enhancement.

---

2. Materials and Methods

2.1 Study Area

Ningxia Hui Autonomous Region

The biomass was estimated using allometric equations based on diameter at breast height (DBH) and tree height [**?**]. For shrubs,

The smaller the Mean Absolute Error (MAE), the stronger the predictive capability and performance of the model.

The wetland sampling plots were 2 m × 2 m, while the farmland sampling plots were 1 m × 1 m.

To comprehensively account for the correlation strength between various factors and carbon storage (CS), we employed a multi-faceted analytical approach. By integrating diverse environmental and anthropogenic variables, the study aims to quantify the relative influence of each driver on carbon sequestration capacity. This methodology ensures that the complex interactions between spatial determinants and carbon dynamics are accurately captured, providing a robust foundation for subsequent modeling and spatial analysis.

The carbon density of the soil was determined using the carbon rate product method. Soil carbon density samples were collected from the 0–30 cm soil layer.

Net Primary Productivity (NPP), Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Aridity Index (AI).

Accuracy of Forest Model Algorithms

In this study, we utilize existing collected sample points and national-level datasets to evaluate and enhance the accuracy of forest model algorithms. By integrating multi-source data, we aim to improve the precision of forest parameter estimation and spatial mapping.

Data Integration and Methodology

The research leverages a combination of field-collected sample plots and national forest inventory data. These datasets serve as the ground truth for training and validating various machine learning and deep learning architectures. To ensure the robustness of the forest models, we incorporate high-resolution remote sensing imagery and topographic variables, which provide critical environmental context for the biological parameters being modeled.

The modeling process involves several key stages: data preprocessing, feature selection, and algorithmic optimization. We specifically focus on the performance of ensemble learning methods, such as Random Forest and Gradient Boosting Decision Trees, comparing them against traditional statistical approaches and contemporary neural networks. The objective is to identify the most effective algorithm for capturing the complex, non-linear relationships inherent in forest ecosystems.

Performance Evaluation

To quantify the accuracy of the forest models, we employ standard statistical metrics, including the Coefficient of Determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). These metrics allow for a rigorous assessment of how well the models generalize across different forest types and geographical regions.

[Figure 1: see original paper]

Preliminary results indicate that the integration of national-level spatial data significantly reduces the bias in local-scale estimations. Furthermore, the application of deep learning techniques shows promise in handling the high dimensionality of the input features, particularly when processing multi-temporal satellite observations. By refining these algorithmic approaches, this study contributes to more reliable forest monitoring and management strategies at both regional and national scales.

Among the eight environmental variables, including Digital Elevation Model (DEM) and Slope ([Figure 3: see original paper]), the most···

1 m, all of which were measured using the total harvest method. Biomass was calculated through biomass proportion conversion and moisture content analysis.

Regarding the issue of multicollinearity among factors, this study employs a correlation matrix analysis to examine the relationships between variables. Multicollinearity can significantly distort the estimation of model parameters and reduce the statistical power of the analysis. By calculating the Pearson correlation coefficients between each pair of factors, we can identify potential redundancies in the feature set. High correlation coefficients suggest that certain factors may provide overlapping information, which could lead to overfitting or instability in the machine learning models. This preliminary diagnostic step ensures the robustness of the subsequent modeling process and informs the selection of independent variables.

Four sampling points are selected along the diagonal of the sampling square to form a single composite sample. To enhance the randomness of the sampling process···

Annual Precipitation (AP), Mean Annual Temperature (MAT), and Digital Elevation Model (DEM)

The Ecological Science Data Center provided a total of 164 supplementary sampling points, which were processed and analyzed using ArcGIS.

A random point generation tool was utilized to create random points across various climate types—including semi-humid, semi-arid, and arid regions—and diverse ecosystem types, such as farmland, forests, shrublands, grasslands, and wetlands. The carbon density values for these points were derived by calculating the ratio of Net Primary Productivity (NPP) between actual sampling sites and random points within the same climate and ecosystem categories. This ratio was then multiplied by the carbon density of the actual sampling sites to perform the conversion. Ultimately, this process resulted in the generation of 621 data points.

uniformly distributed sample data.

The final selection consists of five factors: Net Primary Productivity (NPP), Normalized Difference Vegetation Index (NDVI), Annual Precipitation (AP), Mean Annual Temperature (MAT), and Digital Elevation Model (DEM). The parameters for the Random Forest model were configured with 100 trees and a minimum leaf size of 20.

The model was configured with a maximum tree depth of 10, and 10% of the data was reserved for the validation set. The regression results demonstrate high overall predictive accuracy: the training set achieved an R^2 of 0.87, a Mean Absolute Error (MAE) of 5.10, and a Root Mean Square Error (RMSE) of 7.90. On the validation set, the model maintained robust performance with an R^2 of 0.74, an MAE of 7.36, and an RMSE of 12.31.

1.3.3 空间趋势分析方法

Sen+MK Trend Analysis Method

The Sen+MK test is a robust non-parametric statistical method used for trend analysis. It combines Sen’ s slope estimator with the Mann-Kendall (MK) test to evaluate the magnitude and significance of trends in time-series data. This approach is particularly effective in hydrological and meteorological studies because it does not require the data to follow a specific distribution and is highly resilient to outliers.

1.3.2 生态碳储量估算在碳汇估算中，随机森林

It is robust and possesses the advantage of avoiding interference from missing data and outliers, making it widely

Compared to other machine learning models, this model demonstrates superior performance in terms of anti-overfitting capabilities and predictive accuracy.

used in long-term trend analysis of spatial rasterized numerical data to determine

The closer the coefficient of determination (R^2) is to 1, the better the model fit, while the Root Mean Square Error (RMSE) and Mean Absolute the significance of change trends [**?**]. Based on the standardized test statistics of the Sen+MK test

learning models are more advantageous.

Error (MAE) are used to evaluate the model’ s performance.

thresholds (typically at a significance level of 0.05 with $|Z| \geq 1.96$) and

Note: CS denotes ecosystem carbon storage; NPP denotes net primary productivity; NDVI denotes normalized difference vegetation index; EVI denotes enhanced vegetation index; AI denotes aridity index; AP denotes annual precipitation; MAT denotes mean annual temperature; DEM denotes digital elevation model; Slope denotes slope; R^2 denotes the coefficient of determination. The same applies below.

3 Correlation analysis matrix of factors

The Theil-Sen slope values are used to categorize trends into three distinct classes: significant increase, non-significant change, and significant decrease (the specific formulas are provided in).

(2) Stability analysis using the Coefficient of Variation (C_v):

The C_v represents the ratio of the standard deviation to the mean. The calculation formula is as follows:

Furthermore, the Hurst exponent (H) is utilized to distinguish between persistent trends ($H > 0.5$) and random fluctuations ($H \approx 0.5$).

1.3.4 空间相关性分析方法

Spatial Partial Correlation Analysis

Partial correlation analysis is a statistical method used to analyze the correlation between two variables while controlling for the influence of one or more additional variables. In the context of spatial data, this method allows researchers to isolate the specific relationship between two spatial factors by eliminating the confounding effects of other variables.

When examining complex geographical or environmental systems, multiple factors often interact simultaneously. For instance, when analyzing the relationship between vegetation growth and temperature, precipitation often acts as a significant confounding variable. By employing spatial partial correlation, the linear relationship between the primary independent variable and the dependent variable can be quantified more accurately, ensuring that the observed correlation is not merely a result of their mutual dependence on a third factor.

The calculation of the partial correlation coefficient typically involves the following steps: first, the simple correlation coefficients (Pearson’ s r) between all pairs of variables are calculated. Then, these coefficients are used to compute the partial correlation coefficient. For a first-order partial correlation (controlling for one variable z), the formula is expressed as:

$$R_{xy\cdot z} = \frac{R_{xy} - R_{xz}R_{yz}}{\sqrt{(1 - R_{xz}^2)(1 - R_{yz}^2)}}$$

Where $R_{xy\cdot z}$ represents the partial correlation coefficient between variables x and y after controlling for variable z. The resulting coefficient ranges from -1 to 1, where a positive value indicates a direct relationship, a negative value indicates an inverse relationship, and the absolute value represents the strength of the correlation. In spatial analysis, this calculation is performed pixel-by-pixel or across defined spatial units to generate a spatial distribution map of the correlation, providing insights into how these relationships vary geographically.

(see), is widely used for assessing the volatility of time-series data. Generally, a higher C_v value indicates greater variability within the dataset.

To account for the linear influence of other variables, a linear combination of two specific variables is performed.

3.3 Hurst Exponent Persistence Analysis Method

The Hurst exponent (H) is a critical metric used to characterize the long-term memory and persistence of time series data. Originally developed in the field of hydrology, it has since become a standard tool in fractal geometry and financial econometrics for evaluating whether a series exhibits a trend-reinforcing or mean-reverting behavior.

The value of the Hurst exponent typically ranges between 0 and 1, and its interpretation is categorized into three distinct states:

1. **Persistence ($0.5 < H < 1$):** When the Hurst exponent is greater than 0.5, the time series is considered persistent or trend-reinforcing. This implies that the process possesses long-term memory; if the data has been increasing in the past, it is more likely to continue increasing in the future, and vice versa. The closer H is to 1, the stronger the persistence of the trend.
2. **Random Walk ($H = 0.5$):** A Hurst exponent of exactly 0.5 indicates that the time series follows a classic Brownian motion or a random walk. In this state, observations are independent of one another, and past values provide no predictive power for future movements.
3. **Anti-persistence ($0 \leq H < 0.5$):** When the Hurst exponent is less than 0.5, the series is anti-persistent or mean-reverting. This suggests that an increase in the past is likely to be followed by a decrease in the future, indicating a frequent reversal of the trend.

In the context of this study, the Hurst exponent is employed to analyze the stability and future trajectory of the observed phenomena, ensuring that the long-term dependencies within the dataset are accurately captured and interpreted for subsequent modeling stages.

The partial correlation coefficients were calculated according to the formulas presented in . To determine the statistical significance of these coefficients, T-tests and F-tests were employed.

The higher the volatility, the more susceptible the system is to interference; conversely, lower volatility indicates greater stability.

The Hurst exponent (as shown in) can be used to measure the long-term memory of a time series. This is typically calculated using Rescaled Range (R/S) analysis.

Rescaled Range Analysis (R/S analysis) is employed to reflect the sustainability of future trends in a time series [**?**]. When the Hurst exponent $H < 0.5$, it indicates anti-persistence in the future trend of the time series, suggesting that the future direction is likely to be opposite to the past. When $H \approx 0.5$, the time series trend is considered random or a stochastic process. Conversely, when $H > 0.5$, it indicates positive sustainability, meaning the future trend is likely to continue in the same direction as the past. By combining these results with Sen' s slope⋯

future degradation ($\beta > 0, H < 0.5$), and continuous improvement ($\beta > 0, H > 0.5$).

The change in rates is categorized into future improvement ($\beta < 0, H < 0.5$), continuous degradation ($\beta < 0, H > 0.5$), continuous improvement ($\beta > 0, H > 0.5$), and future degradation ($\beta > 0, H < 0.5$).

[Figure 1: see original paper]

3.2 Analysis of Vegetation Coverage Spatiotemporal Evolution

Based on the pixel-level trend analysis of the fractional vegetation cover (FVC) in the study area from 2000 to 2020, the results indicate that the vegetation coverage in the study area is generally high, showing a spatial distribution pattern of being higher in the south and lower in the north. The overall vegetation coverage exhibits a significant upward trend, with an average annual growth rate of 0.34%.

As shown in , the area with an upward trend in vegetation coverage accounts for 82.45% of the total area, while the area showing a downward trend accounts for only 17.55%. Specifically, the areas with a significant increase ($\beta > 0.05$) are primarily concentrated in the central and southern regions, whereas the areas with a significant decrease ($\beta < -0.05$) are sporadically distributed in the northern urban expansion zones.

3.3 Sustainability Analysis of Vegetation Changes

By combining the Hurst exponent (H) with the Theil-Sen trend analysis, we further explored the future sustainability of vegetation changes in the study area. The results show that the average Hurst exponent for the entire region is 0.62, indicating that the vegetation change trends in the study area possess strong persistence.

[Figure 2: see original paper]

The sustainability analysis reveals that: 1. **Continuous Improvement:** This category covers 65.32% of the study area, mainly located in the ecological restoration zones and forest reserves. This suggests that the current ecological protection measures have achieved remarkable results and are expected to continue in the future. 2. **Continuous Degradation:** This category accounts for 10.15% of the area, primarily found in regions with high human activity and rapid urbanization. 3. **Future Degradation:** Approximately 7.40% of the area shows a trend of "improving now but likely to degrade in the future," which requires close monitoring and early intervention

Correlation analysis [**?**] is employed to reflect the degree and direction of correlation within time-series data. Based on the results of these correlation tests, the data can be categorized into the following driving types (Table 2).

(2) Spatial Autocorrelation Analysis: In this study, we employ the Moran' s Index (Moran' s I) to analyze spatial autocorrelation. This method allows for the quantitative assessment of the degree of spatial dependence and clustering patterns within the geographic data. Global Moran' s I is utilized to determine whether the overall spatial distribution of the observed phenomena exhibits significant clustering, dispersal, or randomness across the entire study area. Furthermore, Local Indicators of Spatial Association (LISA) are applied to identify specific local clusters (hot spots and cold spots) and spatial outliers, providing a more granular understanding of the spatial structure and regional heterogeneity inherent in the dataset.

We conducted a local spatial autocorrelation analysis using Moran' s I (the formula for which is provided in).

and spatial heterogeneity identification, as well as locating the spatial positions of agglomeration centers [**?**]. The value of I ranges from -1 to 1; $I > 0$ indicates the presence of a positive spatial correlation. A higher value suggests a greater degree of similarity and clustering in regional carbon stocks.

Spatiotemporal Analysis Methods for Carbon Sinks

1.1 Theil-Sen Median Trend Analysis

Theil-Sen Median trend analysis, also known as Sen' s slope estimator, is a robust non-parametric statistical method used to calculate the trend of a time series. This method is highly computationally efficient and is insensitive to measurement errors or outlier data, making it particularly suitable for analyzing long-term series of vegetation and carbon sink data. The formula for calculating the slope is as follows:

$$\beta = \text{Median}\left(\frac{x_j - x_i}{j - i}\right), \forall j > i$$

In this equation, β represents the trend of the carbon sink. A value of $\beta > 0$ indicates an increasing trend in the carbon sink over time, while $\beta < 0$ indicates a decreasing trend. The indices i and j represent the time series sequence.

1.2 Mann-Kendall Test

The Mann-Kendall (MK) test is a non-parametric statistical test used to determine the significance of a trend. Unlike parametric tests, it does not require the data to follow a specific distribution and is

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.