



OPEN Spatiotemporal evolution and driving mechanisms of habitat quality on the Loess Plateau

Zheng Cui¹, Haixu Jiang²✉, Liang Zou², Juemei Wu¹ & Hua Yin³

The Loess Plateau, serving as a crucial ecological barrier in China, plays a pivotal role in maintaining and enhancing the habitat quality. This study focuses on the Loess Plateau as a research area. It innovatively integrates the InVEST with the XGBoost-SHAP model to systematically investigate the spatio-temporal characteristics and driving mechanisms of habitat quality changes. A decline in habitat quality has been observed in the study area over the past three decades. The degree of degradation shows a distinct northwest-southeast spatial gradient. High-quality habitat areas are predominantly located in mountainous ecological zones such as the Lüliang Mountains and the Taihang Mountains, as well as on the southern Shanbei Plateau and the western Longzhong Plateau. Terrain ruggedness index (TRI), population density (Pop), and normalized difference vegetation index (NDVI) are the primary factors driving habitat quality spatiotemporal differentiation. The study provides a scientific foundation for ecosystem services management on the Loess Plateau.

Keywords Loess Plateau, Habitat quality, InVEST model, XGBoost-SHAP

The Loess Plateau is one of the most significant ecological barriers in China and globally¹. The sustainability of its ecosystem services is closely linked to the lifeblood of the region's environmental security and socioeconomic development. Habitat quality, as a core indicator of ecosystem services, directly reflects the region's ability to maintain biodiversity and the overall health of its ecosystems². However, the Loess Plateau has long been plagued by soil erosion, land degradation, and ecological fragility³. This has resulted in a steady decline in habitat quality and the deterioration of ecosystem services^{4,5}. In recent years, climate change, rapid urbanization, and intensified human activities have made the Loess Plateau's environmental problems more prominent⁶. It is therefore urgent to investigate the spatial and temporal evolution of habitat quality and its driving mechanisms in the region.

China has always attached significant importance to the Loess Plateau's environmental management. In the 1950s–1970s, the focus of management was on “soil and water conservation,” where soil erosion was mitigated by planting trees and building terraces, etc^{1,7}. In the 1980s–1990s, the policy direction gradually shifted to “ecological restoration, with reforestation, grassland restoration, and environmental migration becoming the main tasks⁸. In the 21st century, the management strategy was further upgraded, and the concepts of “sustainable civilization construction” and “green development” became guidelines. The promotion of the national strategy of “environmental protection and high-quality development of the Yellow River Basin” has injected new policy impetus into the ecological restoration of the Loess Plateau^{8–10}. Despite increasing policy support, the natural environment in this region still faces many difficulties. The spatial and temporal evolution of habitat quality and its driving mechanisms has not been thoroughly studied.

Habitat quality is influenced by natural and human factors. Natural factors, including precipitation, temperature, and topography, play a significant role¹¹. Precipitation affects water conditions for vegetation growth, but it is also closely related to soil erosion¹². Adequate and evenly distributed precipitation promotes vegetation growth, which enhances habitat quality. Temperature affects vegetation growth cycles and species distribution¹³. Extreme temperature events threaten organism survival and reduce the habitat quality¹⁴. Topography determines land drainage, light and heat distribution, and different terrains, such as mountains, hills, and plains, have different impacts on habitat quality¹⁵. For example, vertical climatic variations in mountainous areas create diverse ecological environments. These environments are conducive to biodiversity but may also exacerbate soil erosion risk due to the steepness of the terrain¹⁶. Human activities also affect habitat quality. Urbanization invades natural habitats, leading to habitat fragmentation and living space compression¹⁷. Agricultural activities,

¹School of Geographical Sciences, Liaoning Normal University, Dalian 116029, China. ²Research Institute for Integrated Development of Culture and Tourism, Liaoning University of International Business and Economics, Dalian 116052, China. ³School of Business Administration, DongBei University of Finance and Economics, Dalian 116025, China. ✉email: wl201608092024@163.com

such as excessive cultivation and irrational irrigation methods, can lead to soil degradation and desertification and reduce habitat suitability¹⁸. Pollution caused by industrial activities, such as air, water, and soil pollution, directly jeopardizes organism survival and reproduction and seriously damages habitat quality.

In previous studies, scholars used traditional methods such as correlation coefficient analysis, multiple regression analysis, spatial regression analysis, and geo-detectors^{19,20}. However, in practice, there are often complex nonlinear interactions between factors²¹. When dealing with this nonlinearity and multidisciplinary datasets, they face the challenges of insufficient explanation power and limited ability to capture complex relationships²². To overcome these limitations, domestic and international studies have begun to turn to the application of more advanced machine learning models, including machine learning models such as elastic network regression (ENR) and random forest (RF), as well as geographically weighted regression (GWR) and multiscale geographically weighted regression (MGWR)^{23–25}. XGBoost, as an advanced machine learning algorithm, has unique advantages in ecological research. Based on the gradient boosting framework, XGBoost integrates multiple decision trees to optimize prediction performance step by step^{26,27}. This handles large-scale data efficiently but also has excellent prediction accuracy²⁸. When dealing with a large amount of data in the study of habitat quality on the Loess Plateau, XGBoost can quickly analyze complex data about environmental factors²⁹. It can also explore potential patterns therein. Using the SHAP (Shapley Additive Explanations) algorithm, which is based on the Shapley value in game theory, this study further enhanced the interpretability of the model^{30,31}. It quantifies each feature's contribution to the model output and provides local and global explanations for the XGBoost model³². Therefore, combining the XGBoost and SHAP algorithms can provide insights into the key factors affecting habitat quality.

There are mainly two innovations in this study: first, the InVEST model is combined with XGBoost-SHAP using a multi-model fusion approach. In addition to assessing habitat quality with the InVEST model, the XGBoost-SHAP model analyzes the driving mechanism in depth, thereby compensating for the shortcomings of a single model. Secondly, the XGBoost-SHAP model can provide a more scientific basis for ecological governance than the linear method by revealing the complex nonlinear relationship between factors affecting habitat quality³³. The SHAP value can clarify the key influencing factors and their degree of contribution by quantifying their contributions. It helps target environmental management strategies. This study aims to provide a solid theoretical foundation and practical guidance for scientific management and ecological restoration on the Loess Plateau. Using high-precision spatial and temporal analyses, the study reveals the dynamic characteristics of habitat quality changes on the Loess Plateau; identifies the key drivers of habitat quality and their respective contributions using the XGBoost-SHAP model; and analyzes the driving mechanisms of dominant factors on habitat quality in depth so as to provide a scientific basis for ecological management policies. It not only contributes to a deeper understanding of the Loess Plateau's ecosystem service function. They also provide valuable references for eco-sensitive areas' sustainable development.

Data sources and methods

Study area

Located in central China (103°24'–114°33' E, 33°43'–41°16' N), the Loess Plateau covers approximately 640,000 km² across Shaanxi, Gansu, Ningxia, Shanxi, Henan, Qinghai, and Inner Mongolia (Fig. 1). It is the largest loess deposit globally, characterized by loess hills, plateaus, ridges, and mounds. Elevations range from 800 to 3,000 m, ascending westward. The deep loess layer, tens to hundreds of meters thick, consists of loose, porous aeolian loess with well-developed vertical joints. With temperatures averaging 6–14 °C and precipitation ranging from 300 to 600 mm, the region experiences a temperate continental monsoon climate. Water resources are limited, and the climate is arid. There is a lack of fertility and organic matter in the soil, primarily yellow loessal and dark loessal. The ecosystem is very fragile due to the sparse forest cover of only 15.29%²⁴. In China, the Loess Plateau is a key area for soil erosion control and environmental restoration. Natural factors and human activities influence habitat quality here, varying significantly over time and space. The ability to understand these dynamics and their driving mechanisms is critical to sustainable development.

Data sources

The primary data utilized in this study are as follows:

- (1) Land use data: Raster datasets of land use for the Loess Plateau for the years 1980, 1990, 2000, 2010, and 2020, sourced from the Center for Resource and Environmental Science and Data (<https://www.resdc.cn>). These datasets are categorized into 7 major land types and 24 minor land types with a spatial resolution of 30 m.
- (2) Climate and Environmental Data: Digital Elevation Model (DEM) data were obtained from the Geospatial Data Cloud (<https://www.gscloud.cn/>), with slope and Terrain Ruggedness Index (TRI) data derived from subsequent calculations. All these datasets have a spatial resolution of 30 m. ① Temperature (Tem), soil erosion intensity and precipitation (Pre) data were provided by the Institute of Mountain Hazards and Environment, Chinese Academy of Sciences (<https://imde.cas.cn/>), both featuring a spatial resolution of 30 m. ② Normalized Difference Vegetation Index (NDVI) data were sourced from the Resource and Environment Science and Data Center (<https://www.resdc.cn>), with a spatial resolution of 250 m.
- (3) Socio-economic data: ① Gross Domestic Product (GDP) data were obtained from the Resource and Environment Science and Data Center (<https://www.resdc.cn>), with a spatial resolution of 1 km. ② Nighttime Light (NTL) data were extracted from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) Version 4 Nighttime Light Time Series (<https://eogdata.mines.edu/>), with a spatial re

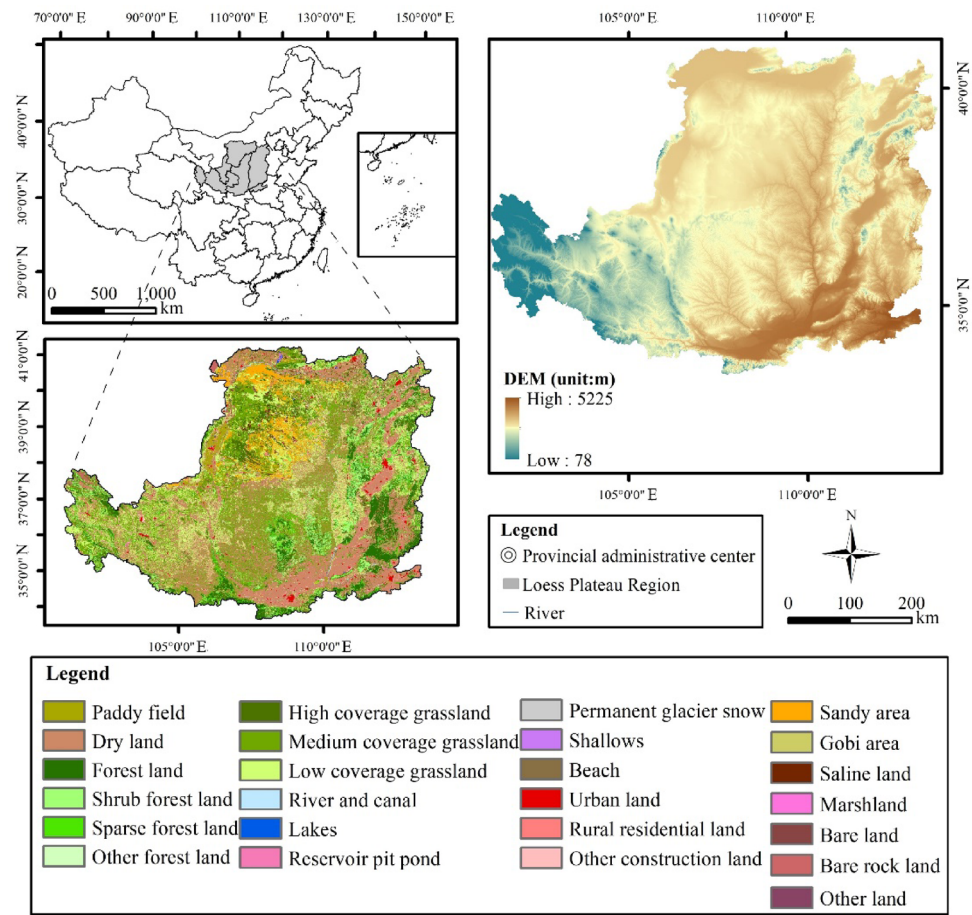


Fig. 1. The study area. Administrative boundary data of the Loess Plateau and 2020 land use data with a spatial resolution of 30 m were obtained from the Resource and Environmental Science Data Platform (<https://www.resdc.cn/>); DEM data were obtained from the Geospatial Data Cloud (<https://www.gscloud.cn/>) with a spatial resolution of 30 m. The map data of China were downloaded from GS (2022) 1873 standard maps from the National Administrative Division Information Query Platform (<https://www.gscloud.cn/>). The Chinese map data are based on the GS (2022) 1873 standard map downloaded from the National Administrative Division Information Query Platform (<http://xzqh.mca.gov.cn/map>). ArcGIS 10.7 generated the above data; the boundaries of the base map remain unchanged.

solution of 500 km. ③ Population Density (Pop) data were derived from the WorldPop project (<https://www.worldpop.org/>), providing a high-resolution dataset of 100 m.

- (4) Other data: ① China map data is based on the national administrative division information query platform (<http://xzqh.mca.gov.cn/map>). The downloaded GS (2022) 1873 standard map. ② The river data is sourced from the National Geographic Information Resource Catalog Service System (<https://www.webmap.cn/>). ③ The administrative division data of the Loess Plateau is sourced from the Resource and Environmental Science Data Platform (<https://www.resdc.cn/>).

The mentioned data were harmonized within the predicted coordinate system in preparation for the upcoming study: Krasovsky_Albers_1940.

Research methods

Habitat quality assessment

This study uses the Habitat Quality module of the InVEST model to evaluate the habitat quality of the Loess Plateau. The habitat quality index, ranging from 0 to 1, reflects the ecosystem's capacity to sustain species²³. The degradation of habitat quality is quantified using the following formula:

$$D_{xj} = \sum_{r=1}^R \sum_{y=1}^{Y_r} \left(\frac{\omega_r}{\sum_{r=1}^R \omega_r} \right) r_y i_{rxy} \eta_x S_{jr} \quad (1)$$

Where: D_{xj} Is the habitat quality stress intensity index of grid x in land use type j ; R represents threat factor; R represents the number of threat factors; ω_r is the weight of the threat factor. The range is between 0 and 1. The

closer the weight is to 1, the greater the influence on the habitat quality; S_{jr} refers to the sensitivity of land use type j to threat factor r . The range is between 0 and 1. The greater the value, the stronger the sensitivity. i_{rxy} represents the threat source value r_y of grid y . The threat level of y to grid x . The model also proposes two ways to calculate i_{rxy} , with the following formula:

$$i_{rxy} = \begin{cases} 1 - \left(\frac{d_{xy}}{d_{r \max}}\right) & (Linear \text{ decay}) \\ \exp \left[-\left(\frac{2.99}{d_{r \max}}\right) d_{xy}\right] & (Exponential \text{ decline}) \end{cases} \tag{2}$$

In the formula: The distance between grids x and y is denoted by d_{xy} , and the maximum influence range of the threat factor is shown by $d_{r \max}$. A higher estimated value of D_{xj} denotes a higher degree of habitat degradation as a result of the threat factor's larger impact on habitat quality. The habitat quality assessment formula is shown as follows²⁵:

$$Q_{xj} = H_j \left(1 - \frac{D_{xj}^z}{D_{xj}^z + K^z}\right) \tag{3}$$

Where: Q_{xj} is the habitat quality index of grid x in land use type j ; H_j is the habitat suitability of land use type j , the value range is between 0 and 1, and the closer to 1, the stronger the suitability is; Z is the normalized constant; K is the semi-saturation constant, generally 1/2 of the maximum value of D_{xj} ;

According to the parameter settings in the relevant literature^{23,26,27} the InVEST model manual the reference value range for this model was determined, with specific parameter settings as shown in Tables 1 and 2, The semi-saturation parameter K adopts the default value, which is set to 0.5.

Selection of influencing factors

In this study, we selected topography, climate, vegetation, soil and socio-economic factors to construct a comprehensive evaluation system, aiming to systematically analyze the spatial differentiation mechanism of habitat quality and its driving factors in the Loess Plateau. From the dimension of natural carrying capacity, DEM, slope and terrain relief (TRI) regulate the soil and water conservation capacity and material-energy redistribution process through the topographic gradient effect, which directly affects the stability of the ecosystem and the threshold of resistance to disturbance; Temperature (Tem) and precipitation (Pre) are important climate factors that influence how plants grow and how ecosystems function by affecting the balance of heat and water; NDVI, which measures how much vegetation is present, is crucial for understanding the quality of habitats on the Loess Plateau and plays a significant role in ecological processes. NDVI, as a quantitative characterization of vegetation cover, together with soil erosion intensity, constitutes a dual pointer for ecosystem health diagnosis, with the former reflecting the level of primary productivity and the latter revealing the risk of degradation of surface cover. From the dimension of anthropogenic coercion, GDP, nighttime lighting (NTL), and population density (Pop) map the intensity of human activities through spatial heterogeneity, and their gradient characteristics are significantly coupled with the distribution of ecological footprints, which can quantitatively characterize the coercive effects of industrialization, urbanization and agricultural development on natural habitats; and the river data, as the carrier of hydrological connectivity, has a special geographic value in the function of ecological corridors and the protection of biodiversity. River data, as a carrier of hydrological connectivity, has special geographic value in maintaining ecological corridor function and biodiversity conservation. This organized combination of different sources can show how the delicate ecosystem of the Loess Plateau reacts to natural conditions, explain how human activities and the environment affect each other, and offer theoretical help in understanding the complex reasons behind the differences in habitat quality in the region.

XGBoost regression

By using XGBoost-SHAP, this study evaluated the factors affecting habitat quality on the Loess Plateau. With XGBoost, multiple decision trees are integrated iteratively to improve prediction accuracy. It is also capable of handling complex nonlinear relationships and high-dimensional data¹⁹. Compared with traditional gradient

Threat factors	Maximum impact distance/km	Weight	Type of recession
Paddy field	3	0.3	Linear
Dry land	5	0.6	Linear
Urban land	10	1	Exponential
Rural residential land	8	0.8	Exponential
Other construction land	9	0.9	Exponential
Sandy area	10	0.7	Linear
Gobi area	6	0.6	Linear
Saline land	5	0.6	Linear
Bare land	3	0.5	Linear
Bare rock land	3	0.5	Linear

Table 1. Threat factor parameter setting.

Land-use type	Habitat suitability	Sensitivity									
		Paddy field	Dry land	Urban land	Rural residential land	Other construction land	Sandy area	Gobi area	Saline land	Bare land	Bare rock land
Paddy field	0.4	0	0.3	0.5	0.6	0.5	0.3	0.3	0.3	0.3	0.3
Dry land	0.3	0.4	0	0.6	0.7	0.6	0.5	0.4	0.5	0.4	0.4
Forest land	1	0.6	0.7	0.8	0.6	0.7	0.5	0.3	0.4	0.2	0.2
Shrub forest land	0.9	0.6	0.6	0.7	0.8	0.5	0.4	0.2	0.3	0.3	0.3
Sparse forest land	0.8	0.6	0.7	0.8	0.7	0.6	0.7	0.3	0.5	0.2	0.2
Other forest land	0.7	0.7	0.7	0.8	0.7	0.7	0.6	0.3	0.4	0.2	0.2
High coverage grassland	0.9	0.5	0.6	0.6	0.5	0.6	0.9	0.3	0.6	0.1	0.1
Medium coverage grassland	0.8	0.6	0.7	0.7	0.7	0.7	0.8	0.4	0.5	0.2	0.2
Low coverage grassland	0.7	0.5	0.6	0.7	0.7	0.8	0.7	0.4	0.5	0.2	0.2
River and canal	1	0.8	0.7	0.9	0.7	0.6	0.7	0.4	0.6	0.5	0.5
Lakes	0.9	0.8	0.7	0.9	0.8	0.5	0.9	0.4	0.6	0.5	0.5
Reservoir pit pond	0.8	0.7	0.7	0.9	0.8	0.5	0.9	0.4	0.6	0.5	0.5
Permanent glacier snow	0.6	0.6	0.6	0.6	0.6	0.4	0.7	0.5	0.5	0.4	0.4
Shallows	0.6	0.7	0.6	0.9	0.8	0.6	0.6	0.4	0.4	0.5	0.5
Beach	0.6	0.7	0.6	0.9	0.8	0.6	0.6	0.4	0.4	0.5	0.5
Urban land	0	0	0	0	0	0	0	0	0	0	0
Rural residential land	0	0	0	0	0	0	0	0	0	0	0
Other construction land	0	0	0	0	0	0	0	0	0	0	0
Sandy area	0	0	0	0	0	0	0	0	0	0	0
Gobi area	0.1	0.1	0.1	0.1	0.1	0.4	0.6	0	0.1	0.1	0.1
Saline land	0.2	0.1	0.1	0.3	0.2	0.3	0.4	0.2	0	0.1	0.1
Marshland	0.7	0.6	0.6	0.9	0.8	0.6	0.6	0.4	0.3	0.2	0.2
Bare land	0.1	0.1	0.1	0.2	0.2	0.2	0.5	0.3	0.2	0	0.1
Bare rock land	0.1	0.1	0.1	0.2	0.2	0.2	0.5	0.3	0.2	0.1	0
Other land	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Table 2. Sensitivity of different landscape type to habitat threat factors.

boosting algorithms, XGBoost employs weighted data allocation in data processing to prioritize key data points, while using regularization techniques to effectively prevent overfitting, and supports parallel computation to significantly accelerate the training speed, and has a built-in function to deal with missing values, and excels in large-scale, high-dimensional data processing scenarios by virtue of its powerful learning capability and excellent generalization performance^{28–30}.XGBoost has the following formulas¹⁷:

$$\hat{Y}_q^K = \sum_{k=1}^K f_k(C_q) = \hat{Y}_q^{K-1} + f_K(C_q) \tag{4}$$

Here, \hat{Y}_q^K represents the predicted result of the q -th input variable after k iterations, f_K is the objective function of the k -th tree model, and C is the input value of the q -th input variable.

SHAP algorithm

The resultant model exhibits the characteristics of an inscrutable “black box.” To conduct an in-depth analysis of the XGBoost regression model and to gain a more comprehensive understanding of the impact of various influencing factors on habitat quality, the SHAP model was employed to interpret the XGBoost model, thereby elucidating the specific contribution of each driving factor¹⁷. Furthermore, the SHAP model has the capability to unveil the precise manner in which each driver influences habitat quality^{21,30,31}.

$$g(h) = \phi_0 + \sum_{j=1}^L \phi_j \tag{5}$$

$$\phi_j = \sum_{C \in \mathcal{E} \setminus \{j\}} \frac{|C|!(|E|-|C|-1)!}{|E|!} \{f(C \cup \{j\}) - f(C)\} \tag{6}$$

where g is the SHAP model, ϕ_0 is the average of all predicted values, L is the number of features, and ϕ_j is the SHAP value of the j th feature. E signifies the set of all features. C iterates over all subsets of E excluding j . The core term, $f(C \cup \{j\}) - f(C)$, computes the difference in model predictions before and after adding feature j . The weighting coefficient, $\frac{|C|!(|E|-|C|-1)!}{|E|!}$, assigns weights to subsets C of different sizes.

Results and analysis

Spatio-temporal patterns of habitat quality

Habitat quality spatiotemporal characteristics

The habitat quality index serves as a crucial indicator of habitat suitability within a region. Overall, the mean habitat quality index for the Loess Plateau from 1990 to 2020 was recorded as 0.527, 0.525, 0.526, and 0.522, consistently maintaining a favorable level (Fig. 2). From a temporal perspective, habitat quality in the region declined from 1990 to 2000. However, a notable recovery was observed from 2000 to 2010, primarily attributed to the implementation of ecological conservation policies, such as the Grain-for-Green Program initiated in 1999, along with the continuous enhancement of regional environmental protection and management. Nevertheless, post-2010, with the rapid progression of urbanization, the expansion of construction land significantly increased, continually encroaching upon agricultural and ecological lands, resulting in a renewed decline in the mean habitat quality index of the Loess Plateau.

To more intuitively illustrate the evolution of habitat quality, this study draws on relevant literature and, in consideration of the specific conditions of the Loess Plateau^{24,32} classifies the habitat quality index into five levels (Fig. 3): mediocre (0.0–0.2), average small (0.2–0.4), medium (0.4–0.6), moderate moderate (0.6–0.8), and excellent (0.8–1.0). Analysis of area changes (Fig. 4) reveals significant differences in the quantitative structure of habitat quality levels from 1990 to 2020. Overall, high-quality and low-quality habitat areas have increased, with increases of 1046.38 km² and 113.89 km², respectively. In contrast, low, medium, and relatively high-quality habitats have generally decreased, with reductions of 9829.24 km², 106.85 km², and 2615.36 km², respectively. Spatially, the distribution generally exhibits a pattern of higher quality in the southeast and lower quality in the northwest. High-quality habitat areas are primarily located in the southeastern regions of the Luliang Mountains and Taihang Mountains, as well as in the southern part of the Shanbei Plateau and the western part of the Longzhong Plateau. Relatively high-quality habitat areas are concentrated in the northern and central parts of the Loess Plateau and the eastern grassland regions of Inner Mongolia. Low-quality areas are predominantly found in the Mu Us Sandy Land and Kubuqi Desert regions, with scattered patches in the southeast. Generally, poor-quality areas are mainly distributed in a strip-like pattern in the north and southeast, while medium-quality areas are distributed around relatively low-quality areas.

Habitat quality degradation characteristics

Habitat degradation is classified into five levels (Figs. 3 and 4): No degradation, Mild degradation, Moderate degradation, High degradation, and Severe degradation. From the perspective of the area-to-area ratio, the trend of habitat degradation from 1990 to 2020 indicates that the areas of non-degraded and severely damaged habitats decreased by 6.97% and 0.14%, respectively. In contrast, mild, moderate, and high degradation increased by 4.05%, 2.49%, and 0.57%, respectively. As shown in Fig. 4, in 1990, habitat quality degradation on the Loess Plateau was primarily concentrated in the northwest and central regions, characterized by moderate and high degradation. Over time, by 2000, the extent of degradation expanded, particularly in the central and eastern regions. This is where the intensity of degradation increased, and extremely degraded areas began to emerge. By 2010, the degradation trend further intensified, with severely degraded areas extending to the southeast and covering more extensive regions. By 2020, habitat quality degradation on the Loess Plateau reached its peak. This was due to severely damaged areas dominating most of the region, especially in the eastern and southeastern parts, where degradation was most pronounced. Overall, the degradation of habitat quality on the Loess Plateau has intensified from the northwest to the southeast over the studied period.

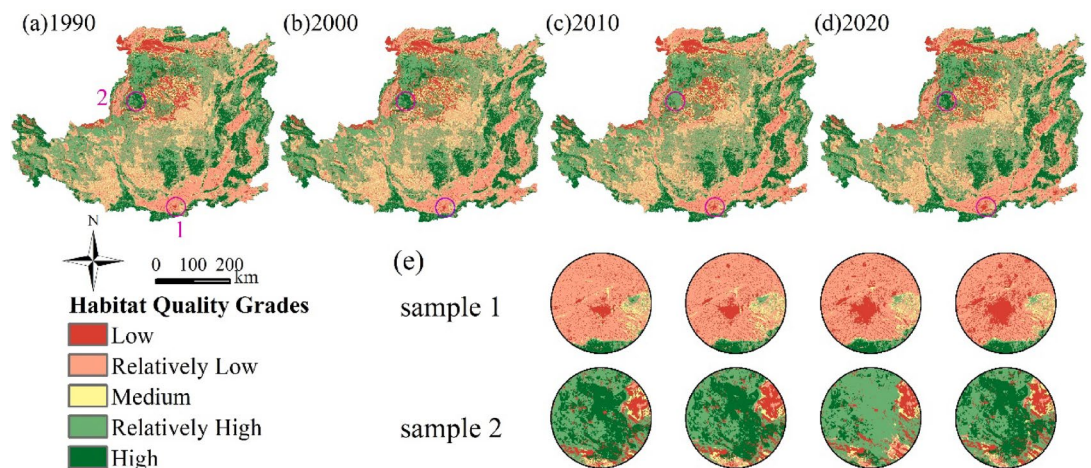


Fig. 2. Spatiotemporal patterns of habitat quality in different periods. (a) 1990, (b) 2000, (c) 2010, (d) 2020, (e) sample of spatial variation in habitat quality. This map was created using the Habitat Quality module in InVEST 3.10.2 and ArcGIS 10.7, and it is based on the results obtained by running the land use classification result data.

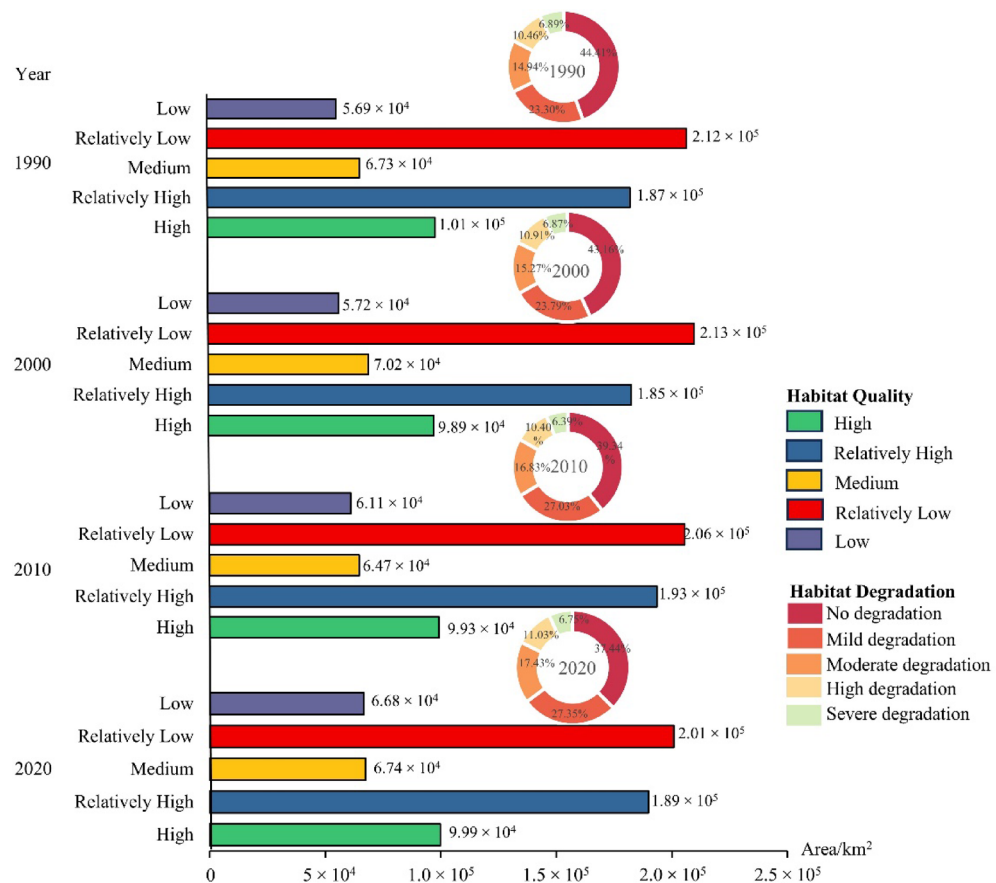


Fig. 3. Degree of habitat quality degradation and quantitative changes in habitat quality.

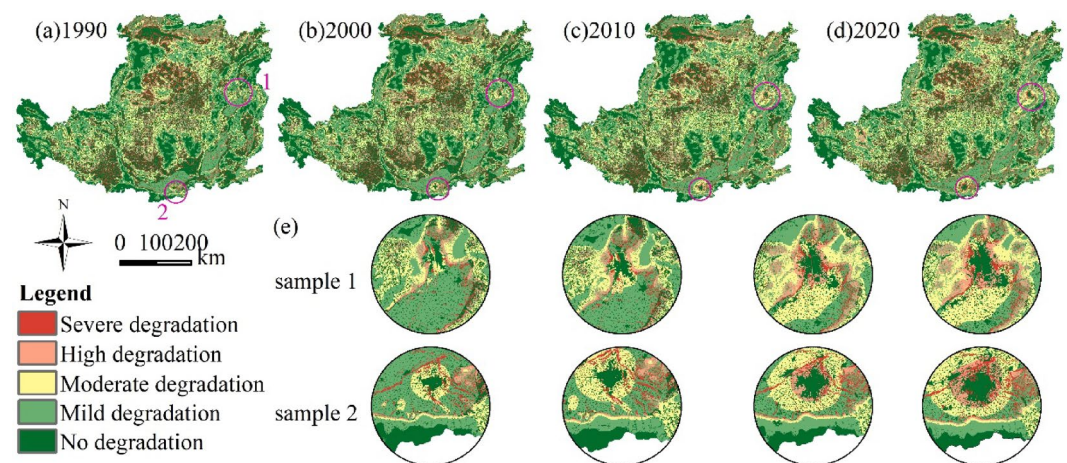


Fig. 4. Degree of habitat quality degradation. (a) 1990, (b) 2000, (c) 2010, (d) 2020, (e) sample areas of significant degradation. This map was created using the Habitat Quality module in InVEST 3.10.2 and ArcGIS 10.7, and it is based on the results obtained by running the land use classification result data.

Analysis of the driving factors of habitat quality

The regression results of the XGBoost-SHAP model

The XGBoost regression model, after hyperparameter tuning, achieved an optimal parameter combination as follows: nrounds = 186, max_depth = 5, eta = 0.05, gamma = 0.3, colsample_bytree = 1, min_child_weight = 1, and subsample = 0.8. The test results demonstrated that the XGBoost regression model exhibits strong generalization capabilities ($R^2=0.802$, $MSE=0.0287$, $RMSE=0.169$), indicating that the XGBoost model has high accuracy and

reliability in explaining and predicting the impact of habitat quality. The high R^2 value and low error metrics suggest that the selected influencing factors in this study can largely explain habitat quality spatial variability.

SHAP-based model interpretation results

XGBoost model prediction results were complemented with SHAP values to determine the extent to which influencing factors affected habitat quality. In terms of the selection of influencing factors, the study covered multi-dimensional factors such as topography, climate, vegetation, human activities, etc., and included a total of 11 characteristic variables. As shown in Fig. 5, the factors affecting the habitat quality of the Loess Plateau are listed in descending order of importance as follows: X3 (TRI) > X8 (Pop) > X6 (NDVI) > X4 (Tem) > X11 (Soil Erosion Intensity) > X5 (Pre) > X1 (DEM) > X2 (Slope) > X10 (Distance to River) > X9 (GDP) > X7 (NTL).

The X3 (TRI), X8 (Pop), X6 (NDVI), X4 (Tem), X11 (Soil Erosion Intensity) and X5 (Pre) all showed positive correlations with habitat quality (Fig. 5). There was a significant contribution from TRI with 8.12%, indicating that ecological diversity brought about by complex topography significantly enhanced habitat quality; Population contributed with 6.16%, suggesting that high-density urban areas might contribute to habitat improvement by implementing ecological protection measures or greening initiatives; vegetation cover influenced habitat quality positively with a 3.85% contribution, as indicated by NDVI; However, the predicted value of habitat quality is positively correlated with temperature and precipitation recharge through suitable thermal conditions and precipitation recharge.

In contrast, the X1 (DEM), X2 (slope), X10 (distance from river), X9 (GDP), and X7 (NTL) results were highly correlated. With increasing distance from the river (X10), the negative effect of reduced water recharge on habitat quality increased; GDP and nighttime lighting (NTL) contributed 0.78% and 0.54%, respectively, indicating a slight negative effect on habitat quality caused by human economic activity intensity and light pollution; slope (X2) and digital elevation model (X1) have an inverse correlation. Overall, TRI, Pop, and NDVI are the core positive factors driving changes in habitat quality. In contrast, anthropogenic factors such as NTL, although weaker in magnitude, still exert a non-negligible negative perturbation on habitat quality.

The dependency relationship between dominant factors and model prediction results

In order to better elucidate the effects of dominant factors on habitat quality, the top six factors in order of SHAP visualization importance - X3 (TRI) > X8 (Pop) > X6 (NDVI) > X4 (Tem) > X11 (Soil Erosion Intensity) > X5 (Pre) - were selected for analysis in this study. This approach is intended to visualize the effects of these factors on habitat quality, thereby increasing the credibility of the model (Fig. 6).

The dependence of each factor on habitat quality showed significant non-linear characteristics. X3 (TRI) SHAP value increased with the increase of terrain ruggedness index, and its effect on habitat quality was weak at low TRI values, and its positive effect was significantly enhanced at high values; x8 (Pop) SHAP value increased and then decreased with the increase of population density, and its effect was small at low density, positively promoted at medium density, and turned to negative effect after exceeding the threshold value; x6 (NDVI) had limited effect on habitat quality when vegetation cover was sparse, and its effect was limited when NDVI values

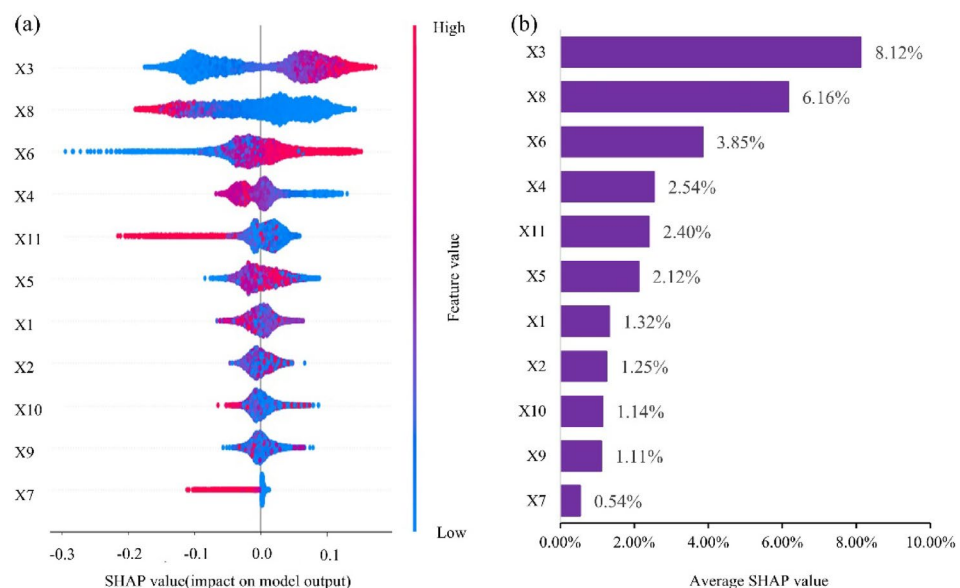


Fig. 5. SHAP value importance swarm chart and bar chart. (a) X1: Digital elevation model (DEM); X2: Slope; X3: Terrain ruggedness index (TRI); X4: Temperature (Tem); X5: Precipitation (Pre); X6: Normalized difference vegetation index (NDVI); X7: Night-time light (NTL); X8: Population density (Pop); X9: GDP; X10: Distance to Rivers; X11: Soil erosion intensity; (b) The percentages on the right represented the contribution of each feature to the model's predicted results, i.e. the magnitude of the impact of changes in each feature on the predicted results.

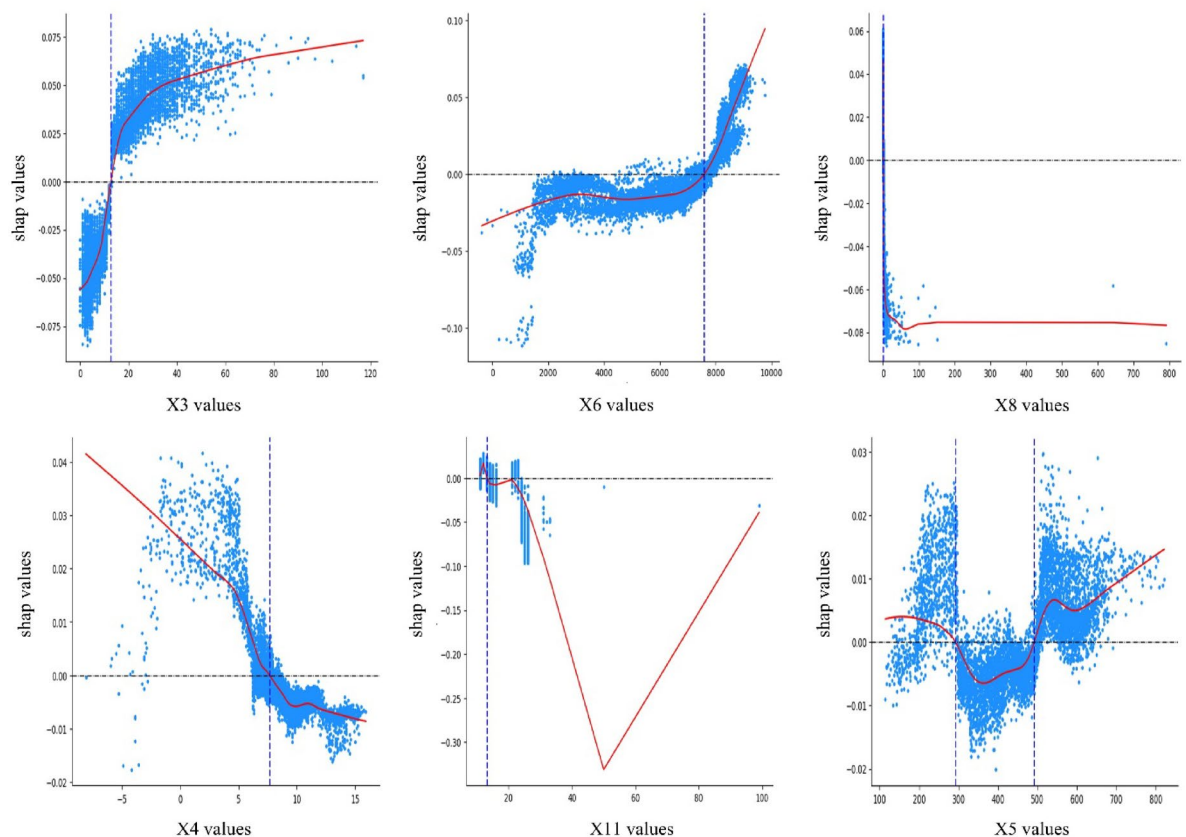


Fig. 6. Dependence relationship between the dominant factors and the model prediction results.

increased, and its effect was limited at high values. Threshold, then turned into a negative effect; X6 (NDVI) had a limited effect on habitat quality when vegetation cover was sparse, and SHAP values fluctuated and climbed as NDVI values increased, with positive enhancement of habitat quality being particularly prominent at high value stages. In addition, X4 (Tem), X5 (Pre) and X1 (DEM) gradually turned from low fluctuation to significant increase in SHAP values with increasing temperature, precipitation and elevation, respectively, suggesting that low temperature, little rainfall and low elevation areas have a weak effect on habitat quality, while its positive effect on habitat quality is more pronounced under high temperature, abundant precipitation and high elevation conditions. Together, these results revealed that the effects of the factors on habitat quality varied in strength across the range of values, and that they were all nonlinearly correlated.

Conclusions and discussion

Conclusion

This study innovatively integrates the InVEST-XGBoost-SHAP framework to assess the spatiotemporal evolution and driving mechanisms of habitat quality on the Loess Plateau from 1990 to 2020. The key findings are as follows:

- (1) Habitat quality on the Loess Plateau exhibited a declining trend from 1990 to 2000. It had a distinct spatial distribution pattern characterized by higher quality in the southeast and lower quality in the northwest.
- (2) The degradation of habitat quality on the Loess Plateau has intensified, spreading from the northwest to the southeast over the study period.
- (3) The habitat quality of the Loess Plateau was positively driven by the core factors of terrain complexity (TRI), population density (Pop), and vegetation cover (NDVI), which contributed 8.12%, 6.16%, and 3.85%, respectively, whereas human economic activities (GDP, NTL) and terrain elevation (DEM) showed a negative inhibitory effect.

Discussion

Spatiotemporal distribution of habitat quality

Habitat quality on the Loess Plateau declined from 1990 to 2000, followed by a significant improvement from 2000 to 2010. Overall, however, habitat quality on the Loess Plateau demonstrated a downward trend from 1990 to 2020. This aligns with existing research findings^{24,26,33}. Policy interventions significantly influence habitat quality and ecosystem functions on the Loess Plateau, as these changes indicate. Ecological restoration initiatives, such as the “Grain for Green” program and the “Three-North Shelterbelt

Forest” project, have played a pivotal role in enhancing habitat quality. The recent decline in habitat quality may be attributed to the Land Management Law’s increasingly stringent requirements for quality management of cultivated land, forests, and grasslands. Moving forward, it is imperative to strengthen ecological protection and restoration efforts, enhance vegetation coverage and soil quality, and promote the rational utilization and planning of land resources. The goal is to mitigate environmental degradation caused by overexploitation and unsustainable practices.

Habitat quality driving factors analysis

During the period 1990–2020, the habitat quality of the Loess Plateau as a whole showed a decreasing trend, and its spatial distribution was characterized by relatively high habitat quality in the southeast and low habitat quality in the northwest, with the degree of degradation gradually increasing from the northwest to the southeast. This phenomenon is mainly influenced by positive factors such as terrain ruggedness index (TRI), population density (Pop), normalized vegetation index (NDVI), temperature (Tem), precipitation (Pre), and soil erosion intensity (X11). The Terrain Ruggedness Index (TRI) had the highest contribution of 8.12%, and the complex terrain provided diverse habitat conditions for a wide range of organisms, promoting ecosystem stability and biodiversity. Population density (Pop) contributed 6.16%, and habitat quality can be improved by ecological protection inputs or urban greening projects within a certain range. However, excessive density can also drive land use change and ecological degradation, especially in the southeast, where greater population pressure exacerbates habitat degradation. With a Normalized Vegetation Index (NDVI) contribution of 3.85%, increased vegetation cover directly enhances ecosystem stability and biodiversity. Temperature (Tem) and precipitation (Pre) had a moderate positive contribution to habitat quality through suitable heat and precipitation recharge, but the contribution was relatively low. Soil erosion intensity (X11), although positively correlated with habitat quality, may increase to reflect reduced vegetation cover and other ecological degradation processes, whose improvement could enhance habitat quality. In contrast, factors such as digital elevation model (DEM), slope (X2), distance from river (X10), gross domestic product (GDP), and nighttime lighting (NTL) had a weaker effect on habitat quality, which was mainly inhibitory. Among them, high altitude and steep slope areas had weaker vegetation cover and ecosystem function due to natural conditions; increased distance from rivers led to reduced water recharge and weakened ecosystem stability; GDP and nighttime lighting index contributed 0.78% and 0.54%, respectively, reflecting the slight negative effects of human economic activity intensity and light pollution on habitat quality, which mostly originated from overdevelopment and accelerated urbanization, leading to habitat fragmentation and degradation. Overall, the Terrain Ruggedness Index (TRI), Population Density (Pop), and Vegetation Cover (NDVI) were the core positive factors driving habitat quality changes, whereas anthropogenic factors such as Night Time Lighting (NTL), although weaker, still exert negative perturbations on habitat quality.

To improve the ecological environment quality of the Loess Plateau, ecological protection zones should be demarcated around the areas with high Terrain Ruggedness Index (TRI) values. The “ecology-economy” synergistic development model will be implemented in the urban-rural transition zones. This model will help communities get involved in restoring plants by providing ecological compensation, while also limiting the uncontrolled growth of areas with too many people. Enhance environmental resilience in climate-sensitive areas by integrating vegetation configuration and temperature and precipitation thresholds. Reduce slope development environmental risks by returning farmland to grass and implementing comprehensive soil erosion management in high-altitude zones. A “environmental red line + industrial access” system should be used in areas where economic development is rapid and the nighttime light index (NTL) is increasing, thereby limiting the layout of highly polluting industries and promoting green infrastructure to reduce light pollution and land hardening. Further, policy support for restoration and protection of the environment should be strengthened, land-use planning should be optimized, and urbanization and population density should be regulated to prevent the destruction of natural habitats. In order to balance southeastern economic and environmental development, the green development model should be promoted to balance the environmental and economic development of the southeastern region; it is necessary to strengthen climate change adaptation measures, and rainfall must be used efficiently to enhance ecosystem resilience; an ecological compensation mechanism should be developed to motivate local governments and residents to take part in the protection of the environment and to build a cooperative relationship across society. Encourage local governments and residents to participate actively in ecological protection, and develop a framework for collaborative governance. Through these comprehensive strategies, the habitat quality of the Loess Plateau region will be gradually restored and improved. The foundation will be laid for the sustainable development of its economy, environment, and social sectors.

Limitations and prospects

The InVEST-XGBoost-SHAP framework adopted in this study offers valuable insights into habitat quality and its drivers in the Loess Plateau, but it also has certain limitations. While the XGBoost model is capable of capturing complex nonlinear relationships, its “black-box” nature may make it difficult to understand its internal mechanisms. Although the SHAP values explain feature importance and dependencies, in some cases, this explanation may still have limitations. The SHAP values, for instance, reflect the average impact of features on the predicted results of the model. The complex interactions between features are not fully revealed by them. Additionally, functional independence may not be fully valid. Models may fail to capture interdependencies and synergies among influences in real ecosystems. Although soil physicochemical properties (e.g., organic carbon,

texture) may influence local habitats, their effects are predominantly overridden by erosion-deposition processes at the landscape scale of the Loess Plateau. For example, vegetation cover (NDVI) and soil erosion intensity (X11) may have a mutual feedback mechanism. An increase in vegetation cover reduces soil erosion, and a decrease in soil erosion favors vegetation growth—a complex dynamic relationship that may not be adequately captured by simple feature importance ranking and dependency analyses. Considering these limitations, future studies should:

- (1) Explore the framework's internal mechanisms in greater depth and combine them with other explanatory models to overcome XGBoost's "black box" nature;
- (2) Break through the assumption of functional independence by analyzing interdependencies and synergistic effects (e.g., NDVI–X11 feedbacks) through complex ecological models or multidisciplinary approaches;
- (3) Incorporate additional soil factors and improve the spatial resolution of selected factors to reduce scaling errors.

Data availability

The author confirms that all data generated or analysed during this study are included in this published article.

Received: 5 March 2025; Accepted: 15 July 2025

Published online: 29 October 2025

References

1. Hui, L. et al. Identifying priority conservation areas based on systematic conservation planning analysis in the Loess Plateau, China. *Glob. Ecol. Conserv.* **58**, e03495 (2025).
2. Zhou, J. et al. Driving mechanisms of ecosystem services and their trade-offs and synergies in the transition zone between the Qinghai-Tibet plateau and the Loess Plateau. *Ecol. Ind.* **171**, 113148 (2025).
3. Chen, Z., Chen, Y., Zhang, H., Zhang, H. & Xu, M. Understanding Spatiotemporal changes and influencing factors in the habitat quality of coastal waters: A case study of Jiangsu province, China (2006–2020). *Ecol. Ind.* **170**, 113125 (2025).
4. Lai, J., Qi, S., Chen, J. & Guo, J. Changes in habitat quality and their driving forces in the coastal zone of Hainan island: insights from 20 years of research and future projections. *Ecol. Front.* <https://doi.org/10.1016/j.ecofro.2024.12.010> (2025).
5. Xiao, Y., Wang, R., Wang, F., Huang, H. & Wang, J. Investigation on Spatial and Temporal variation of coupling coordination between socioeconomic and ecological environment: A case study of the Loess Plateau, China. *Ecol. Ind.* **136**, 108667 (2022).
6. Luo, M., Jia, X., Zhao, Y., Zhang, P. & Zhao, M. Ecological vulnerability assessment and its driving force based on ecological zoning in the Loess Plateau, China. *Ecol. Ind.* **159**, 111658 (2024).
7. Li, G., Jiang, C., Gao, Y. & Du, J. Natural driving mechanism and trade-off and synergy analysis of the Spatiotemporal dynamics of multiple typical ecosystem services in Northeast Qinghai-Tibet plateau. *J. Clean. Prod.* **374**, 134075 (2022).
8. Zhang, X. et al. Spatiotemporal evolution of ecological vulnerability in the yellow river basin under ecological restoration initiatives. *Ecol. Ind.* **135**, 108586 (2022).
9. Wei, L., Zhou, L., Sun, D., Yuan, B. & Hu, F. Evaluating the impact of urban expansion on the habitat quality and constructing ecological security patterns: A case study of Jiziwan in the yellow river basin, China. *Ecol. Ind.* **145**, 109544 (2022).
10. Yong, X. & Chuansheng, W. Ecological Protection and High-quality Development in the Yellow River Basin: Framework, Path, and Countermeasure. (2020).
11. Luo, Y. et al. Spatial and Temporal evolution of habitat quality and its shrinkage effect in shrinking cities: evidence from Northeast China. *Ecol. Ind.* **161**, 111919 (2024).
12. Wang, H. et al. Spatiotemporal dislocation of ecosystem supply and demand services from habitat quality under different development scenarios. *Ecol. Ind.* **157**, 111230 (2023).
13. Xie, B., Meng, S. & Zhang, M. Evolution of habitat quality and its response to topographic gradient effect in a karst plateau: A case study of the key biodiversity conservation project area of Wuling mountains. *IJERPH* **20**, 331 (2022).
14. Zhang, Y. et al. Identifying ecological security patterns based on the supply, demand and sensitivity of ecosystem service: A case study in the yellow river basin, China. *J. Environ. Manag.* **315**, 115158 (2022).
15. Gu, J., Wu, J. & Xie, D. Coastal salt marsh changes in china: landscape pattern, driving factors, and carbon dynamics. *Geogr. Sustain.* **100281** <https://doi.org/10.1016/j.geosus.2025.100281> (2025).
16. Chen, Y. et al. Relationship between urban Spatial form and seasonal land surface temperature under different grid scales. *Sustain. Cities Soc.* **89**, 104374 (2023).
17. Wei, J., Li, Y., Jia, L., Liu, B. & Jiang, Y. The impact of Spatiotemporal effect and relevant factors on the urban thermal environment through the XGBoost-SHAP model. *Land* **14**, 394 (2025).
18. Zeng, W. et al. A Prediction–Interaction–Driving framework for ecosystem services under climate change and human activities: A case study of Zoigè plateau. *Land* **14**, 441 (2025).
19. Wei, Z., Chen, D., Huang, Q., Chen, Q. & Wei, C. Temporal-Spatial Evolution and Driving Mechanism for Ecosystem Health Service Based on GD-MGWR-XGBOOST-SEM Model: A Case Study in Guangxi Region. <https://www.preprints.org/manuscript/202502.1539/v1> <https://doi.org/10.20944/preprints202502.1539.v1> (2025).
20. Zhou, B., Chen, G., Yu, H., Zhao, J. & Yin, Y. Revealing the nonlinear impact of human activities and climate change on ecosystem services in the karst region of southeastern Yunnan using the XGBoost–SHAP model. *Forests* **15**, 1420 (2024).
21. Ouyang, N., Rui, X., Zhang, X., Tang, H. & Xie, Y. Spatiotemporal evolution of ecosystem health and its driving factors in the Southwestern karst regions of China. *Ecol. Ind.* **166**, 112530 (2024).
22. Yuan, Y., Guo, W., Tang, S. & Zhang, J. Effects of patterns of urban green-blue landscape on carbon sequestration using XGBoost-SHAP model. *J. Clean. Prod.* **476**, 143640 (2024).
23. Yang, Y. Evolution of habitat quality and association with land-use changes in mountainous areas: A case study of the Taihang mountains in Hebei province, China. *Ecol. Ind.* **129**, 107967 (2021).
24. Shi, N. et al. Spatiotemporal variation of habitat quality and its response to fractional vegetation cover change and human disturbance in the Loess Plateau. *Front. Ecol. Evol.* **11**, 1302224 (2023).
25. Ye, H., Song, Y. & Xue, D. Multi-Scenario simulation of land use and habitat quality in the Guanzhong plain urban agglomeration, China. *IJERPH* **19**, 8703 (2022).
26. Dong, J. et al. Spatiotemporal variations and driving factors of habitat quality in the loess hilly area of the yellow river basin: A case study of Lanzhou city, China. *J. Arid Land* **14**, 637–652 (2022).
27. Zhang, X. et al. Spatiotemporal Variation and Influence Factors of Habitat Quality in Loess Hilly and Gully Area of Yellow River Basin: A Case Study of Liulin County, China. *Land* **11**, 127 (2022).

28. Hee Shim, S. & Hyun Choi, J. Building an XGBoost model based on landscape metrics and meteorological data for nonpoint source pollution management in the Nakdong river watershed. *Ecol. Ind.* **165**, 112156 (2024).
29. Qu, J., Xu, Z., Dong, B., Wang, H. & Han, Y. Analyzing trade-offs, synergies, and driving factors of ecosystem services in Anhui Province using Spatial analysis and XG-boost modeling. *Ecol. Ind.* **171**, 113098 (2025).
30. Erdem Okumus, D. & Akay, M. Quantitative assessment of non-stationary relationship between multi-scale urban morphology and urban heat. *Build. Environ.* **272**, 112669 (2025).
31. Zuo, L. et al. Assessing the destabilization risk of ecosystems dominated by carbon sequestration based on interpretable machine learning method. *Ecol. Ind.* **167**, 112593 (2024).
32. Xu, T., Tian, G., Lin, T., Li, W. & Liu, C. Multiscale analysis of ecosystem service interactions and driving factors in the Loess Plateau: implications for ecological management. *J. Clean. Prod.* **495**, 145074 (2025).
33. Hou, Y., Lü, Y., Chen, W. & Fu, B. Temporal variation and Spatial scale dependency of ecosystem service interactions: a case study on the central Loess Plateau of China. *Landsc. Ecol.* **32**, 1201–1217 (2017).

Acknowledgements

We express our gratitude to anonymous reviewers and editors for their professional comments and suggestions.

Author contributions

Conceptualization, Z.C., Y.H. and L.Z.; methodology, Z.C., Y.H. and J.W.; software, Z.C., J.W. and L.Z.; validation, Z.C. and L.Z.; formal analysis, Z.C. and Y.H.; investigation, L.Z. and H.Y.; data curation, Z.C. and H.Y.; writing—original draft preparation, Z.C., and Y.H.; writing—review and editing, Z.C., Y.H. and L.Z.; All authors have read and agreed to the published version of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025