Construction of a multifactor barrier evaluation system and classification of barrier types for arable land on the northern Huang-Huai-Hai Plain

Xiangyun Guo[1], Chi Zhang[2], Zihao Liu[1], Baozhong Yin[3], Hong Wang[2,4], Xuguang Li[5], Ruifang Zhang[2*], Yongwei Cui[6*]

(1. *College of Agronomy, Hebei Agricultural University, Baoding 071001, Hebei, China*;
2. *College of Land and Resources, Hebei Agricultural University, Baoding 071001, Hebei, China*;
3. *College of Plant Protection, Hebei Agricultural University, Baoding 071001, Hebei, China*;
4. *College of Resources and Environment, Hebei Agricultural University, Baoding 071001, Hebei, China*;
5. *Hebei Cultivated Land Monitoring and Protection Center, Shijiazhuang 050000, China*;
6. *Academy of Agricultural Planning and Engineering, Ministry of Agriculture and Rural Affairs (MARA), Beijing 100125, China)*)

Abstract: Soil physicochemical properties, climate, and human activities can create barriers to arable land in varying degrees, affecting land quality. The Huanghuaihai Plain (HHHP) is an important agricultural region in China. To clarify the factors influencing the formation of barriers to arable land in this area and their spatial distribution characteristics, this study took the northern part of the HHHP as the research object, screened and quantified the factors influencing barriers to arable land, constructed a multifactor-based arable land barrier evaluation index system, and used the index system to spatially classify the barriers to arable land. The results showed that 1) 16 evaluation indicators including the five dimensions of chemical indicators, physical indicators, biological indicators, management measures, and plot environment were screened out through the random forest model; 2) the average rating of the multifactor barrier for arable land in the northern part of the HHHP was 5.3, exhibiting a medium level, and the area of grade 5 and grade 6 land accounted for the highest percentage, at up to 30%; 3) the order of barrier degree of main barrier factors from high to low was organic matter>salt content>available phosphorus>available potassium>irrigation capacity>soil texture class>soil bulk density; and 4) according to the idea of ranking barrier factors, 15 types of barriers were obtained and then divided into the three major barrier area categories of organic matter, irrigation capacity, and salinity, and the prioritization of cropland quality improvement was determined according to the sequential order of the combination of barrier factors. A preliminary multifactor barrier index system for croplands was constructed, which can provide a reference for cropland barrier abatement and the precise improvement of cropland quality in the HHHP area.

Keywords: Huanghuaihai Plain, cultivated land, barrier factors, evaluation system, barrier types
DOI: 10.25165/j.ijabe.20251802.9428

Citation: Guo X Y, Zhang C, Liu Z H, Yin B Z, Wang H, Li X G, et al. Construction of a multifactor barrier evaluation system and classification of barrier types for arable land on the northern Huang-Huai-Hai Plain. Int J Agric & Biol Eng, 2025; 18(2): 189–196.

1 Introduction

High-quality land, as a valuable natural resource, is an important safeguard for maintaining ecological balance, food production, and human health[1]. However, the degradation of ecosystem services, climate change, biodiversity loss, water scarcity, and reduction in arable lands[2-6] pose significant threats to soil quality[7]. These unfavorable conditions affect crop growth, development, and yield formation and influence the quality of crops, with adverse effects on the ecological environment. In this context, the systematic identification of the factors hindering the cultivation of arable land, building a mechanism to reduce barriers, and realizing the precise improvement of the quality of arable land have become research focuses in the field of global land science. This issue is particularly urgent in China, which is undergoing rapid urbanization and industrialization. Due to the special national conditions of a large population base and scarce arable land resources, China faces the dual challenges of arable land protection and food security, which are more complicated than those of other countries[8].

Soil is a complex ecosystem in which various physical, chemical, and biological traits can affect soil quality to varying degrees, creating barriers[9,10]. Much useful research has been conducted on the factors that influence soil quality. For example, Yang, He, Yan, and Qu et al.[11-14] proposed several factors closely related to barriers to arable land quality, such as heavy metals, soil water content, AP, pH, and cropland efficiency, in the North China

Received date: 2023-10-14 **Accepted date:** 2024-10-09

Biographies: Xiangyun Guo, MS, research interest: evaluation of farmland quality and spatial planning, Email: g827123@163.com; **Chi Zhang**, Assistant researcher, research interest: plant nutrition, Email: 49621561@qq.com; **Zihao Liu**, MS, research interest: evaluation of farmland quality and spatial planning, Email: 724741277@qq.com; **Baozhong Yin**, Associate Professor, research interest: agricultural ecology, Email: yinbaozhong@hebau.edu.cn; **Hong Wang**, researcher, research interest: soil ecology and integrated soil management, Email: wanghong@hebau.edu.cn; **Xuguang Li**, researcher, research interest: plant nutrition and fertilization, Email: 314489681@qq.com.

***Corresponding author: Ruifang Zhang**, Associate researcher, research interest: plant nutrition. College of Land and Resources, Hebei Agricultural University, Baoding 071001, Hebei, China. Tel: +86-312-7528158, E-mail: zhrf@hebau.edu.cn; **Yongwei Cui**, Professor, research interest: agriculture economic impact of natural resources. Academy of Agricultural Planning and Engineering, Ministry of Agriculture and Rural Affairs (MARA), Beijing 100020, Beijing, China. Tel: 13681582627, E-mail: ciyongwei@126.com.

Plain, Loess Plateau, Nanxiong Basin, and Northeastern Black Soil Region of China, respectively. The aforementioned studies have revealed that the majority of existing research concentrates on a single factor in the formation of barriers to arable land. However, studies on the synergistic or interlocking effects of these factors are scarce. In the soil ecosystem, no factors exist in isolation; rather, they operate through mutual constraints or promotions to impact soil quality. Furthermore, the findings of the aforementioned studies indicate that the barrier factors affecting arable land exhibit significant variation across different soil types, posing considerable challenges in determining the specific barrier factors affecting arable land. Consequently, the development of a multifactor-based evaluation index system for barriers to croplands and the quantification of the interrelationships among the factors represent crucial research avenues for future identification and mitigation of barriers to croplands.

Many useful studies have been conducted by previous authors to screen and evaluate barrier factors. For example, in the screening of evaluation indices, studies have been conducted to assess the quality of arable land using the random forest method[15,16] to assess the importance of features and the principal component analysis (PCA) method to construct a minimum dataset (MDS)[17-20]. The majority of existing research in this field has employed the hierarchical analysis method (AHP)[21] and the entropy weight method[22] to determine the weights and degree of affiliation. The affiliation function reflects the degree to which an element belongs to a certain fuzzy set, which can be broadly classified into three main categories: triangular, trapezoidal, and Gaussian functions[23,24], with regional variations influencing their performances[25]. The final physical and chemical properties of the soil from the soil health status, utilization conditions, ecological security, arable land productivity, and other latitudes were used to construct an evaluation index system[26-28]. Nevertheless, existing research on the construction of an index system for evaluating the quality of arable land lacks a systematic evaluation from the perspective of exploring the synergy of barrier factors from the common departure of multiple dimensions. Therefore, it is important to construct a multifactor barrier evaluation index system for arable land from the perspective of the mechanisms and paths of arable land barrier formation. This will facilitate a more scientific and accurate evaluation of the index system of arable-land barriers and enable targeted conservation and the upgrading of land strength.

The Huanghuaihai Plain (HHHP) is one of the main grain-producing areas in China and plays a key role in guaranteeing China's food security[29,30]. However, in recent years, with the deterioration of climatic conditions, ecological pollution, and decline in soil fertility, the quality of arable land in the HHHP has been seriously affected, further threatening regional food security. Therefore, accurately identifying the barriers to croplands in the HHHP and improving the quality of croplands have become key problems that need to be solved in this district. In this study, the barriers to arable land in the northern part of the HHHP were taken as the object of study, and the screening and evaluation of the barriers to arable land and the classification of the degree of barriers to arable land in the district were carried out using mathematical models, such as random forest and barrier degree models. This study aims to clarify the factors and combinations of driving mechanisms of barriers to arable land quality in the northern part of the HHHP and to quantify the spatial distribution of the degree of barriers to arable land quality in the district. This study provides a reference for the construction of technical systems to improve the quality of arable land and reduce barriers to arable land in HHHP and similar regions.

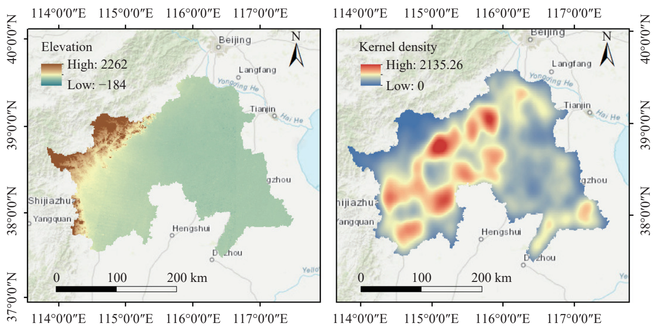
2 Material and methods

2.1 Study area and data sources

The data collection area for this study was the northern part of HHHP (113°88 ′E-117°44 ′E, 37°50 ′N-39°58 ′N), comprising 49 county-level administrative districts and a total land area of approximately 8×10^6 hm². This region is the most important winter wheat-summer maize one-year grain-producing area in China. The fundamental data for this study were sourced from the five-year mean data provided by the Department of Agriculture and Rural Affairs of Hebei Province from 2017 to 2021 (Table 1). Topographic data were primarily sourced from the geospatial data cloud (http://www.gscloud.cn/) to obtain 30×30 m DEM data, which were then integrated with the grading criteria for a comprehensive assessment, as shown in Figure 1.

Table 1 Names of indicators in the study area and corresponding abbreviations

Indicator name	Abbreviation	Indicator name	Abbreviation	Indicator name	Abbreviation
Altitude/m	ALT	Barrier Layer Type	BLT	Total Nitrogen/g·kg^{-1}	TN
Longitude	LON	Barrier Layer Depth	BLD	Available Phosphorus/mg·kg^{-1}	AP
Latitude	LAT	Barrier Layer Thickness/cm	BL	Available Potassium/mg·kg^{-1}	AK
Soil Type	STL	Irrigation Capacity	IC	Slowly Available Potassium/mg·kg^{-1}	SRK
Soil Subtype	STB	Irrigation Method	IM	Soil pH	pH
Soil Group	STG	Water Source Type	TWS	Available Sulfur/mg·kg^{-1}	AS
Soil Series	STS	Drainage Capacity	DC	Available Copper/mg·kg^{-1}	ACu
Parent Material	PM	Cropping System	CS	Available Zinc/mg·kg^{-1}	AZn
Landform Type	LFT	Annual Cropping System	PCS	Available Iron/mg·kg^{-1}	AFe
Soil Texture Class	TC	Cleanliness Degree	AP	Available Manganese/mg·kg^{-1}	AMn
Topographic Position	TP	Main Crop Name	AK	Available Boron/mg·kg^{-1}	AB
Slope Gradient/(°)	FS	Annual Yield/kg·hm^{-2}	SRK	Available Molybdenum/mg·kg^{-1}	AMo
Groundwater Depth/m	DGW	Biodiversity	pH	Available Silicon/mg·kg^{-1}	ASi
Effective Soil Depth/cm	ESD	Farmland Afforestation	AS	Lead	Pb
Plough Layer Thickness/cm	PLD	Salinization Type	ACu	Chromium/mg·kg^{-1}	Cr
Plough Layer Texture	PLT	Degree of Salinity	AZn	Cadmium/mg·kg^{-1}	Cd
Soil Bulk Density/g·cm^{-3}	SBD	Salt Content	AFe	Mercury/mg·kg^{-1}	Hg
Obstacle Factor	OF	Cleanliness Degree	AP	Arsenic/mg·kg^{-1}	As



a. Elevation map b. Point density map

Note: The study area is located in the northern Huanghuaihai Plain.

Figure 1 Geographic location of the study area

2.2 Research methods

2.2.1 Screening of evaluation indicators

The random forest model was created using the "Random Forest" package in R software version 4.1.1 to identify 53 different factors that influence the yield (GY). The out-of-bag error (OOB error) was used to calculate the importance of the characteristic variables (I). The significance of the variable (X^j) to the j is expressed as [$I(X^j)$]. Coefficient of determination (R^2), Relative prediction deviation (RPD), Root mean square error (RMSE), and Mean absolute error (MAE) were used as metrics to evaluate and validate model accuracy[31]. In this study, the random forest model was constructed with an R^2 value of 0.789, RPD of 2.030, RMSE of 61.925, and MAE of 34.997.

$$I(X^j) = \frac{1}{n}\sum_{i=1}^{n}\left(E_i^j - E_i\right) \tag{1}$$

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{3}$$

$$\mathrm{RPD} = \mathrm{SD}/\mathrm{RMSE} \tag{4}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{5}$$

2.2.2 Constructing indicator weights and affiliations

1) AHP method

AHP, proposed in the early 1970s, integrates quantitative and qualitative aspects to determine the weights of individual evaluation indicators (Table 2 and Table 3). To maintain perceived consistency, this study verified the consistency by calculating consistency ratios. These ratios were used to gauge the level of consistency between pairwise comparisons of different criteria, which is a key strength of the AHP method. The fundamental steps of the AHP approach are as follows[32]:

(1) Building judgment matrix:

$$U = (a_{uv})_{m\times m} \tag{6}$$

U is an $m\times m$ judgment matrix that represents the relative importance of factors in the same hierarchy; the matrix element a_{uv} represents the importance of factor u relative to factor v, and its value ranges from 1 to 9.

(2) Compute the maximum eigenvector of the judgment matrix.

$$U_w = \lambda_{\max} w \tag{7}$$

U_w is the product of the judgment matrix U and the eigenvector w. $\lambda_{\max}$ is the maximum eigenvalue of the judgment matrix U, which is used to measure the consistency of the matrix. w is the eigenvector corresponding to the maximum eigenvalue $\lambda_{\max}$, which represents the weight of each factor.

(3) Using the consistency test, the consistency indicator CR is calculated as follows:

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \tag{8}$$

where, CI is the consistency index and RI is the randomized average consistency index; when CR<0.1, the judgment matrix passes the consistency test.

(4) The CI was calculated as follows:

$$\mathrm{CI} = \frac{\lambda_{\max} - n}{n-1} \tag{9}$$

(5) The weight W_i of each indicator was obtained after normalization.

2) Calculation of the degree of the affiliation of each indicator

To establish a relationship between the selected evaluation indicators and the quality of cultivated land, the national standard for Cultivated Land Quality Grade (GB/T33469-2016) was referenced in conjunction with a fuzzy mathematical method based on the characteristics of the northern HHHP. This resulted in the classification of the relationships into above-the-fold, peak, conceptual, and selected numerical categories. The conceptual indicators are non-numerical and partially numerical and are not linearly related to the quality of cultivated land. Therefore, an affiliation function is not necessary to establish a linear relationship between these indicators and the arable land quality.

Table 2 Cropland barrier evaluation indicator affiliation function

Membership function type	Formula	Marginal notes	Evaluation indicator
Above-the-fold function	$y_i = \begin{cases} 0,\ u_i \le u_t \\ 1/\left[1+a_i(u_i-c_i)^2\right],\ u_t < u_i < c_i \\ (i=1,\ 2,\ \dots,\ m) \\ 1,\ c_i < u_i \end{cases}$	y_i is the affiliation of the ith factor; u_i is the measured value of the sample; c_i is the standardized indicator; a_i is the indicator coefficient; u_t is the lower limit value of the indicator	SOM, AP, AK, TN
Peak function	$y_i = \begin{cases} 0,\ ui > u_{t1} \text{ or } u_i < u_{t2} \\ 1/\left[1+a_i(u_i-c_i)^2\right],\ u_{t1} < u_i < u_{t2} \\ 1,\ u_i = c_i \end{cases}$	u_{t1} and u_{t2} are the upper and lower limit values of the indicator, respectively	pH, SBD
Conceptual indicators	-	-	TC, PLT, BD, TP, IC, DC
Selected numerical indicators	-	-	ESD, PLD, DGW, SC

Table 3 Affiliation of indicators for evaluating barriers to cropland

Degree of affiliation	Dimensions									
	Conceptual index						Numerical index			
	PLT	TC	TP	BD	IC	DC	PLD	ESD	DWG	SC
1.00	Medium loam	-	Low altitude alluvial plain/Low altitude alluvial flood plain/Low altitude alluvial lacustrine plain/Plain low grade	Abundant	Fully satisfied	Fully satisfied	≥20	≥100	≥10	≤0.02
0.95	-	Sticky type	Low altitude erosion and denudation plains	-	-	-	-	-	-	-
0.94	Light loam	-	-	-	-	-	-	-	-	-
0.93	-	Loose on top and tight on bottom type	-	-	-	-	-	-	-	-
0.92	Heavy loam			-	-	-	-	-	-	-
0.90		Loamy type	Low altitude alluvial floodplains/Low altitude alluvial depressions	-	-	-	-	-	-	-
0.88	Clay loam			-	-	-	-	-	-	-
0.85	-	Firm type	Low altitude alluvial floodplain low terrace/Low altitude fluvial low terrace	-	Satisfied	Satisfied	-	-	-	-
0.80	Sandy loam	Sandwich type	Low altitude marine alluvial plains/Plains middle order	Average			15-20	60-100	10-5	0.02-0.05
0.75	-	Sponge type/Tight on top and loose on bottom type	-	-	-	-	-	-	-	-
0.70	-	Loose type	Low altitude marine depressions/Low altitude river high terraces	-	Largely satisfied	Largely satisfied	-	-	-	-
0.65	-	-	Erosion and denudation of low altitude low hills	-	-	-	-	-	-	-
0.60	-	Through-body sand	Plains higher	Not abundant	-	-	≤15	30-60	≤5	0.05-0.08
0.50	Sandy Soil	-	-	-	Not satisfied	Not satisfied	-	-	-	-
0.40	-	Thin layer type	High hills at altitude in erosion and denudation	-	-	-	-	<30	-	≥0.08
0.35	-	-	Erosion and denudation of small undulating mesas	-	-	-	-	-	-	-
0.20	-	-	Erosion and denudation of the great rolling hills	-	-	-	-	-	-	-

2.3 Evaluation of cropland barrier levels

The cropland quality was divided into ten grades using the equidistant method, with grade 1 being the lowest obstacle and grade 10 being the highest. The data were then spatially interpolated using ArcGIS 10.2 software to obtain the evaluation results of the multifactor obstacle grades of the cropland in the study area.

$$Q = \sum (M_i \times N_i) \tag{10}$$

where, Q is the multifactor barrier composite index for the grain fields, M_i is the weight of the i^{th} indicator, and N_i is the degree of affiliation of the j^{th} indicator.

Average grade of barriers to cultivation=

$$\frac{\sum(\text{Cropland obstacle class} \times \text{area of that class})}{\text{Total area of cultivated land}} \tag{11}$$

2.3.1 Obstacle diagnosis model

This study referred to the calculation method of Cui et al.[33] and impairment degree model to explain the primary and secondary relationships among impairment factors. This study categorized the degree of indicator impairment into four levels: no obstacle (0), mild obstacle (0%–10%), moderate obstacle (10%–15%), and severe obstacle (≥15%).

$$O_{ij} = \frac{(1-F_{ij}) \times W_{ij}}{\sum_{i=1}^{j}\left[(1-F_{ij}) \times W_{ij}\right]} \times 100\% \tag{12}$$

where, O_{ij} is the barrier value of the j^{th} indicator in the i^{th} study area, F_{ij} is the factor contribution, and W_{ij} is the weight of the indicator. To further clarify the synergistic relationship between barrier factors, evaluation indicators with a barrier degree greater than 5% were selected as key indicators for the analysis.

2.3.2 Barrier factor combination ideas

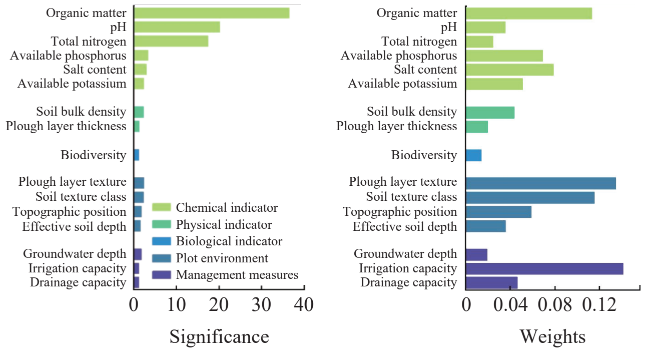
The idea of barrier zoning for the study area is as follows: use ArcGIS 10.2 software to rank the seven barrier factors and select the first-, second-, and third-level barriers. When the barrier degree of the first barrier factor reaches a severe barrier (≥15%) and the second and third barrier factors are moderate or less (≤15%), the evaluation unit for barriers to arable land is determined to be an area of the "first barrier factor" type. When the first and second barrier factors reach the level of severe barrier (≥15%) and the third barrier factor is at the level of moderate barrier or below (≤15%), the evaluation unit of barriers to arable land is determined to be an area of the type "first barrier factor - second barrier factor". When the primary, secondary, and tertiary barrier factors all reach the level of severe barrier (≥15%), the arable land barrier evaluation unit is determined to be a "primary barrier factor - secondary barrier factor - third barrier factor" type area. When the first, second, and third barrier factors are all at the level of moderate barriers or less (≤ 15%), the evaluation unit for barriers to arable land is determined to be an area of the "first barrier factor" type.

3 Results

3.1 Screening of factors affecting cultivated land quality

In Figure 2a, the first 16 variables of the random forest model are listed according to the OOB error, which is more sensitive to the evaluation results. The five most important variables were SOM, pH, TN, AP, and SC. The following characteristics of each indicator were considered: physical soil indicators (PLD and SBD), chemical indicators (SOM, pH, TN, AP, AK, and SC), biological indicators (BD), plot environments (PLT, TP, TC, and ESD), and management measures (IC, DC, and DGW). Among these, soil chemical indicators accounted for the highest percentage of the variables (83%), followed by plot environment (8%), physical soil indicators

(3.6%), management measures (4.2%), and biological soil indicators (1.2%). Among the weights assigned to the indicators, irrigation capacity had the highest weight, followed by tillage texture and texture configuration. Among the chemical indicators, organic matter had the highest weight and biodiversity had the lowest weight (Figure 2b).



a. Random forest importance ranking b. Indicator weighting

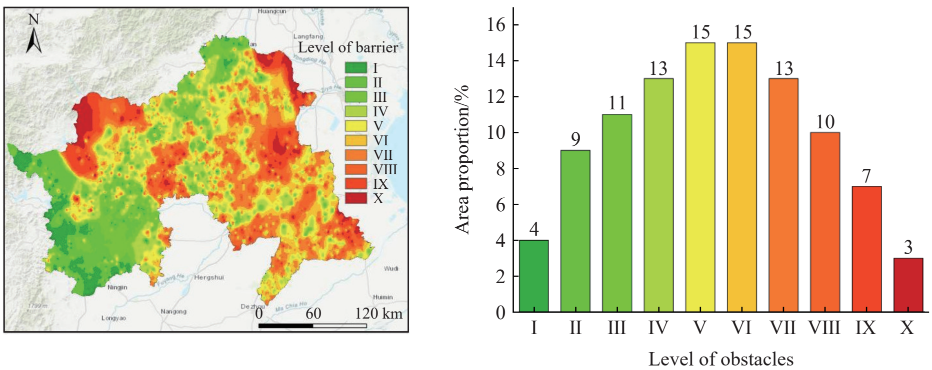
Figure 2 Ranking of importance and weighting of indicators for evaluating barriers to cropland

3.2 Barrier class of arable land in the study area

The range of the composite index of barriers to arable land in the northern part of HHHP was 0.620 736-0.975 018, which yielded an average grade of 5.3 for multifactorial barriers to arable land, indicating a medium level of severity. In the entire range of graded classes, 10%-15% were classified as classes III, IV, V, VI, VII, and VIII. The highest percentage was observed in classes V and VI, which together accounted for 30%. Meanwhile, 0%–10% were classified as class I, class II, class IX, and class X. Class X constituted the lowest proportion (3%). Lands classified as classes I–IV were distributed across a vast expanse of the southwestern and northern regions of the study area. Lands designated as classes V–VII exhibited a more dispersed pattern and were situated in the western, central, northeastern, and southeastern parts of the study area. Conversely, high-grade lands, classified as classes IX–X, were concentrated in the northwestern and eastern regions (Figure 3).

3.3 Diagnosis of single-factor barriers

As shown in Figure 4, chemical indicators exhibited the highest barrier degree (42.21%). Among chemical indicators, SOM and SC had the highest barriers, at 16.37% and 15.79%, respectively. The level of management measures had IC with the highest barrier degree (9.52%). The plot environment had the highest TC barrier degree at 6.89%, physical indicators had the highest BD barrier degree at 5.44%, and biological indicators had the lowest barrier degree at 0.97%. The seven indicators with obstacle degrees greater than 5% were SOM, SC, AP, AK, BD, TC, and IC, which were the major barriers to croplands in the northern HHHP.



a. Spatial distribution of cropland barrier levels map b. Percentage of area of cropland barrier levels

Figure 3 Map of multifactorial obstacle classes for cropland and the percentage of area in obstacle classes

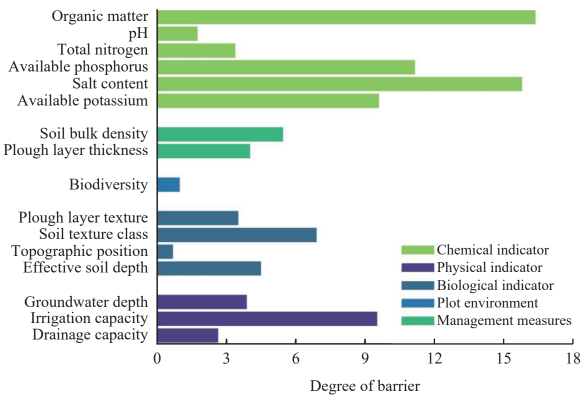


Figure 4 Diagnostic results of cropland barrier factors

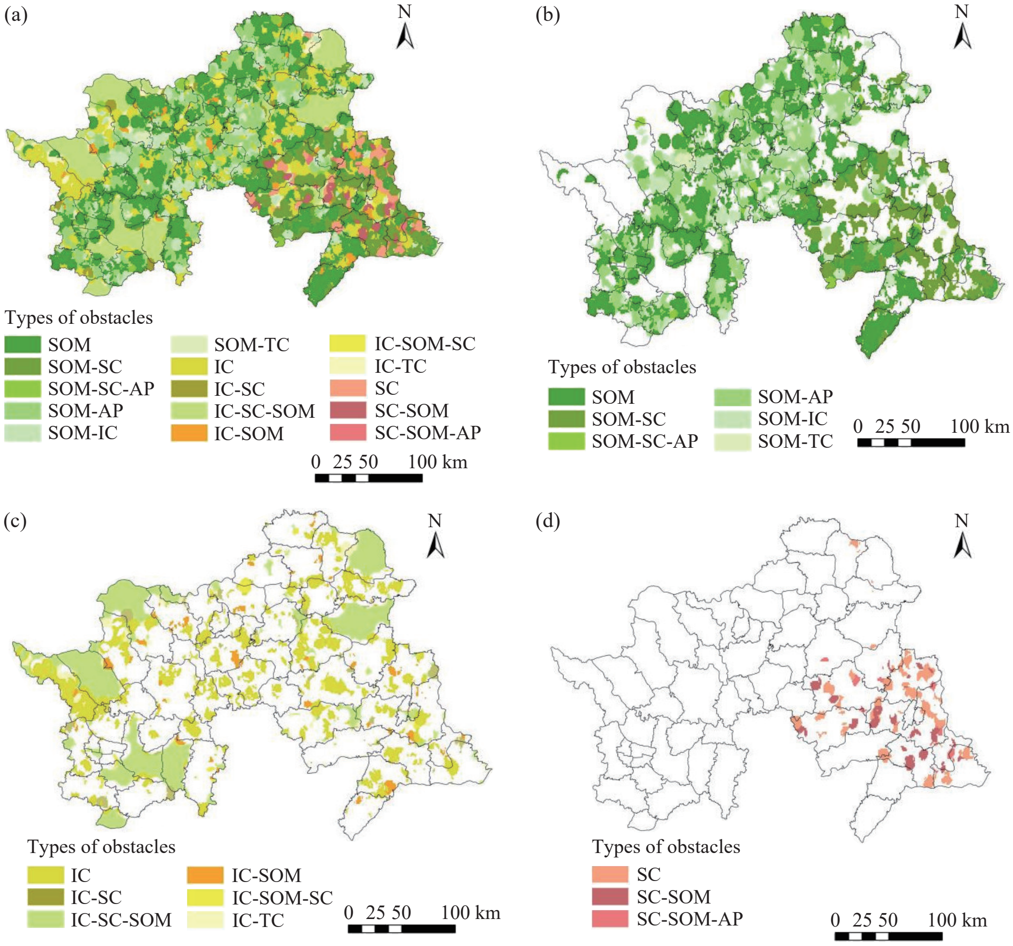
3.4 Multifactor barrier types

In accordance with the concept of barrier factor ordering, 15 barrier types were identified and subsequently classified into three principal categories of barrier areas: SOM, IC, and SC. The SOM barriers are divided into six types of barriers: SOM, SOM-IC, SOM-SC, SOM-SC-AP, SOM-AP, and SOM-TC. The IC barriers are categorized as IC, IC-SC, IC-SC-SOM, IC-SOM, IC-SOM-SC, and IC-TC. The SC obstacle zone is divided into three types of obstacles: SC, SC-SOM, and SC-SOM-AP. The SOM barrier type area has the largest area of 2 297 783.27 hm², accounting for 61.71% of the total area of the study area, with a relatively decentralized distribution. The IC barrier type area covers an area of 1 244 699.89 hm², accounting for 33.43% of the total area of the northern part of the HHHP, and it is mainly located in the western, southwestern, and northeastern parts of the study area. The SC barrier area is the smallest, which is 181 334.23 hm², accounting for 4.87% of the total area of the study area, and is mainly distributed in the southeastern coastal area of the study area.

4 Discussion

4.1 Construction of the evaluation index system for barriers to arable land

In the screening process for evaluation indicators, a random forest model was employed to identify 16 representative evaluation indicators. This approach effectively identifies key variables through feature importance ranking, and its objectivity has been extensively validated in soil quality assessment research[34-36]. For instance, Wang et al.[37] utilized random forest modeling to calculate



Note: Figure 5a shows the general distribution of barrier types, Figure 5b shows the distribution of organic matter barrier types, Figure 5c shows the distribution of irrigation capacity barrier types, and Figure 5d shows the distribution of salinity barrier types.

Figure 5 Types of barriers to cropland

the feature importance of red-edge parameters, whereas Cao et al.[38] incorporated multiple environmental factors as downscaling variables for indicator screening. These methodological approaches are aligned with the algorithmic design employed in this study. Although random forest modeling objectively quantifies indicator contributions through information gain ratios, inherent methodological uncertainties persist. First, the model demonstrates sensitivity to initial parameter configurations, potentially affecting the feature importance rankings[39]. Second, the default assumption of additive relationships among input variables may neglect nonlinear interactions affecting cultivated land degradation[40].

Significant variations emerged across distinct topographic units regarding the regional adaptability of the indicator systems. To illustrate this, in the case of China's freshwater lake wetland region[41], a comprehensive evaluation index system was constructed using SOM, TN, AP, AK, pH, TP, and TK. In contrast, for the hilly regions of East China [42], indicators such as SOM, AP, AK, and TP were selected. In a study on the US federal state wetland region[43], indicators for the evaluation of MDS were screened, including pH, TN, BD, and total organic carbon (TOC). From this, it can be seen that chemical indicators such as SOM, pH, TN, AP, and AK have been widely studied by various scholars, which is consistent with the results of this study. Further analysis revealed potential uncertainties in constructing the indicator system. At the data collection level, the density of the spatial distribution of soil samples may not be able to fully capture the microtopographic variability. At the temporal dynamics level, data collected at a single time point makes it difficult to characterize the interannual fluctuations in soil properties. Future research should enhance system robustness through spatiotemporal interpolation modeling and dynamic threshold adjustment mechanisms.

Notably, indicator selection discrepancies stem from fundamental differences in research scales and objectives. While the principal component analysis-derived MDS[44] simplifies the evaluation systems, the large-scale study encompassing diverse landforms requires multilevel indicator frameworks. This approach aligned with the recommendations of Bünemann et al.[45] that hierarchical systems should be employed at scales exceeding 1: 100 000. Furthermore, crop-specific studies, such as that of Sánchez-Guzmán et al. on maize cultivation[46], introduced specialized indicators (SMR and qCO_2) in addition to conventional parameters. By incorporating regional indicators (SC, PLT, and TC) while preserving the comparability of core parameters (SOM, pH, and TN), this study enhanced the systemic interpretability. Future investigations should adopt context-specific indicator selections based on spatial scales and research objectives.

4.2 Classification of types of barriers to arable land

Soils are the basis of human survival; however, they are complex and fragile ecosystems in which small changes in certain indicators can lead to a decline in the productive capacity of the soil, thereby reducing its ability to self-regulate and recover. TN, SOM, pH, available nitrogen, and CEC were found to be the primary degradation factors in the Hetao Plain[47], whereas the Yangtze River Basin[42] showed limitations in terms of ESD, SOM, pH, IC, OF, and tillage layer texture. Jiangxi's red soil sloping fields[48] demonstrate high PLD constraints, in contrast to the elevation and IC dominance in the northern Tianshan Mountains[49]. This study revealed that SOM, SC, AP, AK, IC, TC, and BD are critical constraints in the northern HHHP, aligning with previous findings on the significance of SOM and AK. These results not only confirm the pivotal role of SOM in soil quality assessment but also guide targeted soil improvement strategies.

Further analysis of the above revealed that several barrier factors were specific to croplands in different regions. TN and pH constraints in the Hetao Plain reflect arid salinization characteristics, whereas northern HHHP SOM and AK depletion correlate with high cropping intensity. The ESD and tillage layer issues of the Yangtze Basin are linked to hilly erosion processes, in contrast to the flat terrain of the northern HHHP, which minimizes such effects. A comparative analysis showed that Jiangxi's PLD constraints are associated with acidic soil structural instability, whereas Xinjiang's IC issues stem from irrigation practices. The northern HHHP exhibits compound salinity-nutrient constraints due to groundwater overexploitation.

Notably, the northern HHHP exhibited significant spatial heterogeneity, with 15 identified degradation types (Figure 3a). This variation stems from a combination of the internal topography, climate, and pedological diversity. Specifically, the northern plains experience increased soil compaction due to urban-induced fragmentation, whereas the central and southwestern grain belts face depletion of SOM and AK as a result of intensive cultivation. Irrigation limitations are a prominent issue in western mountainous areas, whereas the eastern riverine zones contend with salt accumulation. Furthermore, spatiotemporal variations in cropping indices and management practices have exacerbated spatial disparities.

The results of the aforementioned studies provide a wealth of insight that can inform subsequent research on the barriers to cropland abatement in diverse ecological settings. However, a notable limitation of these studies is their exclusive focus on single-factor barrier assessment. From the complexity of soil ecosystems, it can be inferred that the quality of arable land itself contains several intertwined influencing factors. In this regard, this study proposed a multifactor barrier evaluation index system for arable land based on a combination of barrier factors using an obstacle diagnosis model and the concept of barrier factor combination. Furthermore, this study quantitatively classified barriers to arable land in the northern part of the HHHP in China. However, this study was deficient in its exploration of the synergistic relationship between barrier factors; the concept of establishing a multifactor barrier evaluation index system through a combination of barrier factors is of reference value. Further research will be conducted to gain a deeper understanding of the synergistic mechanisms and driving pathways among barrier factors. The technical pathways of multifactor barrier abatement will be integrated, validated, and optimized, providing a basis for the precise enhancement of the quality of arable land in the HHHP area.

5 Conclusions

A random forest model was employed to screen 16 evaluation indicators affecting the quality of arable land in the northern part of the HHHP. These included SOM, AP, and AK. The contributions of these indicators were ranked. Furthermore, based on the affiliations and weights of the aforementioned evaluation indicators, the average rank of the multifactor barriers to arable land in the northern part of the HHHP was determined. The extent of the barriers posed by the evaluation indicators of arable land in the northern part of the HHHP was quantified and a multifactor evaluation index system for barriers to arable land was proposed based on a combination of primary and secondary barrier factors. The aforementioned indicator system was employed to categorize the barriers to arable land in the northern part of the HHHP into three principal categories: SOM, IC, and SC. The spatial distributions of these barrier areas were determined.

Based on the results of this research, it is recommended that the relevant decision-makers make precise resource inputs, allocate agricultural resources according to the dominant factors in different obstacle zones, classify management by zones, formulate differentiated protection and utilization policies for different obstacle zones, establish a long-term dynamic monitoring system to monitor and evaluate the quality of arable land regularly so that measures can be adjusted promptly, strengthen the technical training of farmers, and popularize appropriate agricultural technologies to improve the quality of arable land in the northern part of the HHHP in various aspects.

Acknowledgements

This study was funded by the National Key Research and Development Program of China (Grant No. 2021YFD1901001-02), "Construction of Multifactor Barrier Evaluation Indicator System for Grain Fields in the Huanghuaihai."

[References]

[1] Giovannetti M, di Fossalunga A S, Stringlis I A, Proietti S, Fiorilli V. Unearthing soil-plant-microbiota crosstalk: Looking back to move forward. Front Plant Sci, 2023; 13: 1082752.

[2] Zheng W W, Ke X L, Zhong T, Yang B H. Trade-offs between cropland quality and ecosystem services of marginal compensated cropland - A case study in Wuhan, China. Ecol Indic, 2019; 105: 613–620.

[3] Tsesmelis D E, Karavitis C A, Kalogeropoulos K, Zervas E, Vasilakou C G, Skondras N A, et al. Evaluating the degradation of natural resources in the mediterranean environment using the water and land resources degradation index, the case of Crete Island. Atmosphere, 2022; 13(1): 135.

[4] Webb N P, Marshall N A, Stringer L C, Reed M S, Chappell A, Herrick J E. Land degradation and climate change: building climate resilience in agriculture. Front Ecol Environ, 2017; 15(8): 450–459.

[5] Abdelrahman M, Afifi A, Scopa A. A time series investigation to assess climate change and anthropogenic impacts on quantitative land degradation in the North Delta, Egypt. ISPRS Int J Geo-Inf, 2021; 11(1): 30.

[6] Simmonds J S, Sonter L J, Watson J E M, Bennun L, Costa H M, Dutson G, et al. Moving from biodiversity offsets to a target-based approach for ecological compensation. Conserv Lett, 2020; 13(2): e12695.

[7] Paz A M, Castanheira N, Miloczki J, Carrasco M, Vicente C, Carranca C, et al. Collected knowledge on the impacts of agricultural soil management practices in Europe. Eur J Soil Sci, 2024; 75(2): e13468.

[8] Cheng X, Long R Y, Chen H. Obstacle diagnosis of green competition promotion: a case study of provinces in China based on catastrophe progression and fuzzy rough set methods. Environ Sci Pollut Res, 2017; 25(5): 4344–4360.

[9] Zornoza R, Acosta J A, Bastida F, Dominguez S G, Toledo D M, Faz A. Identification of sensitive indicators to assess the interrelationship between soil quality, management practices and human health. Soil, 2015; 1(1): 173–185.

[10] Bran Nogueira Cardoso E J, Figueiredo Vasconcellos R L, Bini D, Horta Miyauchi M Y, dos Santos C A, Lopes Alves P R, et al. Soil health: looking for suitable indicators. What should be considered to assess the effects of use and management on soil health? Sci Agric, 2013; 70(4): 274–289.

[11] Yang J, Wang J Y, Qiao P W, Zheng Y M, Yang J X, Chen T B, et al. Identifying factors that influence soil heavy metals by using categorical regression analysis: A case study in Beijing, China. Front Environ Sci Eng,

2020; 14(3): 37.
[12] He S L, Ma L F, Zhao K T, Yang J J, Chang Y W. Survey analysis of soil physicochemical factors that influence the distribution of Cordyceps in the Xiahe Region of Gansu Province. Open Life Sci, 2017; 12(1): 76–81.
[13] Yan P, Peng H, Yan L B, Zhang S Y, Chen A M, Lin K R. Spatial variability in soil pH and land use as the main influential factor in the red beds of the Nanxiong Basin, China. Peer J, 2019; 7: e6342.
[14] Qu Y, Pan C L, Guo H P. Factors affecting the promotion of conservation tillage in black soil-The case of northeast China. Sustainability, 2021; 13(17): 9563.
[15] Zhang H, Ren X N, Chen S K, Xie G Q, Hu Y S, Gao D D, et al. Deep optimization of water quality index and positive matrix factorization models for water quality evaluation and pollution source apportionment using a random forest model. Environ Pollut, 2024; 347: 123771.
[16] Zhao M J, Yan C H, Feng Y Z, Xue J, Tang H S, Zhang T L, et al. A novel strategy for quantitative analysis of soil pH via laser-induced breakdown spectroscopy coupled with random forest. Plasma Sci Technol, 2020; 22(7): 074003.
[17] Zhang X N, You Y, Wang D C, Zhu L. Quality evaluation of the soil-root composites layer of Leymus chinensis grassland based on different degradation degrees. CATENA, 2022; 215: 106330.
[18] Abd-Elwahed M S. Effect of long-term wastewater irrigation on the quality of alluvial soil for agricultural sustainability. Ann Agric Sci, 2019; 64(2): 151–160.
[19] Mahajan G, Das B, Morajkar S, Desai A, Murgaokar D, Kulkarni R, et al. Soil quality assessment of coastal salt-affected acid soils of India. Environ Sci Pollut Res, 2020; 27(21): 26221–26238.
[20] Teng L D, Jiang G H, Ding Z L, Wang Y, Liang T B, Zhang J Z, et al. Evaluation of tobacco-planting soil quality using multiple distinct scoring methods and soil quality indices. J Clean Prod, 2024; 441: 140883.
[21] Shinde S P, Barai V N, Gavit B K, Kadam S A, Atre A A, Pande C B, et al. Assessment of groundwater potential zone mapping for semi-arid environment areas using AHP and MIF techniques. Environ Sci Eur, 2024; 36(1): 87.
[22] Ye J. Multicriteria fuzzy decision-making method using entropy weights-based correlation coefficients of interval-valued intuitionistic fuzzy sets. App.Math Model, 2010; 34(12): 3864–3870.
[23] Chhibber D, Bisht D C S, Srivastava P K. Pareto-optimal solution for fixed-charge solid transportation problem under intuitionistic fuzzy environment. App.Soft Comput, 2021; 107: 107368.
[24] Khairuddin S H, Hasan M H, Hashmani M A, Azam M H. Generating clustering-based interval fuzzy Type-2 triangular and trapezoidal membership functions: A structured literature review. Symmetry, 2021; 13(2): 239.
[25] Guo L L, Sun Z G, Ouyang Z, Han D R, Li F D. A comparison of soil quality evaluation methods for Fluvisol along the lower Yellow River. CATENA, 2017; 152: 135–143.
[26] Tan Y, Chen H, Lian K, Yu Z. Comprehensive evaluation of cultivated land quality at county scale: A case study of Shengzhou, Zhejiang Province, China. Int J Environ Res Public Health, 2020; 17(4): 1169.
[27] Zhao C, Zhou Y, Jiang J H, Xiao P N, Wu H. Spatial characteristics of cultivated land quality accounting for ecological environmental condition: A case study in hilly area of northern Hubei province, China. Sci Total Environ, 2021; 774: 145765.
[28] Wang M C, Liu X N, Liu Z W, Wang F Y, Li X Y, Hou G L, et al. Evaluation and driving force analysis of cultivated land quality in black soil region of Northeast China. Chin Geogr Sci, 2023; 33(4): 601–615.
[29] Yue Y J, Yang W Q, Wang L. Assessment of drought risk for winter wheat on the Huanghuaihai Plain under climate change using an EPIC model-based approach. Int J Digit Earth, 2022; 15(1): 690–711.
[30] Shi X Y, Zhao J, Jia H, Zhao J C, Lu J, Zhao M Y, et al. Seeking sustainable pathway of crop production by optimizing planting structures and management practices from the perspective of water footprint. Sci Total Environ, 2022; 843: 157091.
[31] Nozari S, Pahlavan-Rad M R, Brungard C, Heung B, Borůvka L. Digital soil mapping using machine learning-based methods to predict soil organic carbon in two different districts in the Czech Republic. Soil Water Res, 2024; 19(1): 32–49.
[32] Li Y, Sun L, Wu Z N, Liu H. Evaluation and obstacle factors of marine resources and environment carrying capacity in Beibu Gulf Urban Agglomeration. Front Mar Sci, 2023; 10: 1196196.
[33] Cui X F, Cai T, Deng W, Zheng R, Jiang Y H, Bao H J. Indicators for evaluating high-quality agricultural development: Empirical study from Yangtze River economic belt, China. Soc Indic Res, 2022; 164(3): 1101–1127.
[34] Fei H, Fan Z H, Wang C K, Zhang N N, Wang T, Chen R G, et al. Cotton classification method at the county scale based on multi-features and random forest feature selection algorithm and classifier. Remote Sens, 2022; 14(4): 829.
[35] Yuan X G, Liu S R, Feng W, Dauphin G. Feature importance ranking of random forest-based end-to-end learning algorithm. Remote Sens, 2023; 15(21): 5203.
[36] Pu H B, Yu J X, Sun D W, Wei Q Y, Shen X L, Wang Z. Distinguishing fresh and frozen-thawed beef using hyperspectral imaging technology combined with convolutional neural networks. Microchem J, 2023; 189: 108559.
[37] Wang Z, Li Y X, Wu F, Zou X C. Estimation of winter wheat chlorophyll content by combing canopy spectrum red edge parameters with random forest machine learning. Transactions of the CSAE, 2024; 40(4): 166–176. (in Chinese)
[38] Cao Y, Fang X Q, Yang L L, Jiang X Y, Liao M Y, Ren L L. Downscaling of CCI soil moisture in the Xiliaohe River basin based on random forest. J Geo-Inf Sci, 2023; 25(8): 1669–1681. (in Chinese)
[39] Genuier R, Poggi J-M, Tuleau-Malo C. Variable selection using random forests. Pattern Recognit Lett, 2010; 31(14): 2225–2236.
[40] Cutler D R, Edwards Jr T C, Beard K H, Cutler A, Hess K T, Gibson J, et al. Random forests for classification in ecology. Ecology, 2007; 88(11): 2783–2792.
[41] Li Z W, Huang J Q, Li Y Y, Guo W, Zhu J F. Assessment on soil fertility of Dongting Lake wetland area (China) based on GIS and fuzzy evaluation. J Cent South Univ Technol, 2011; 18(5): 1465–1472.
[42] Tang M M, Wang C T, Ying C Y, Mei S, Tong T, Ma Y H, et al. Research on cultivated land quality restriction factors based on cultivated land quality level evaluation. Sustainability, 2023; 15(9): 7567.
[43] Chang T T, Feng G, Paul V, Adeli A, Brooks J P, Jenkins J N. Soil health assessment for different tillage and cropping systems to determine sustainable management practices in a humid region. Soil Tillage Res, 2023; 233: 105796.
[44] Xue R, Wang C, Liu M L, Zhang D, Li K, Li N. A new method for soil health assessment based on Analytic Hierarchy Process and meta-analysis. Sci Total Environ, 2019; 650: 2771–2777.
[45] Bünemann E K, Bongiorno G, Bai Z G, Creamer R E, De Deyn G, de Goede R, et al. Soil quality-A critical review. Soil Biol Biochem, 2018; 120: 105–125.
[46] Sánchez-Guzmán A, Bedolla-Rivera H I, Conde-Barajas E, Negrete-Rodríguez M d l L X, Lastiri-Hernández M A, Gámez-Vázquez F P, et al. Corn cropping systems in agricultural soils from the Bajio Region of Guanajuato: Soil quality indexes (*SQIs*). App.Sci, 2024; 14(7): 2858.
[47] Zhang J H, Huang H Y, Ding Q D, Jia K L. Soil quality evaluation and obstacle diagnosis of saline-alkali cultivated land in Hetao Plain. Environ Sci, 2024: 1–17. doi: 10.13227/j.hjkx.202404226 (in Chinese)
[48] Jin H F, Shi D M, Song G, Duan T, Ye Q, Jiang N. Cultivated-layer quality characteristics and obstacle types division of red soil sloping farmland. Transactions of the CSAM, 2019; 50(12): 313–321, 340. (in Chinese)
[49] Tang D, Yang J J, Cheng P. Comprehensive evaluation of soil substrate improvement based on the minimum data set method. Sustainability, 2022; 14(7): 3939.