

Simple assessment of farmland soil phosphorus loss risk at county scale with high landscape heterogeneity

Huirong Ning¹, Qimeng Liu², Shiwen Zhang^{2*}, Huichun Ye^{3*}, Qiang Shen⁴, Wentao Zhang², Zhen Li⁵

(1. People's Government of Xianshengdian Township, Lu'an 237005, Anhui, China;

2. School of Earth and Environment, Anhui University of Science and Technology, Huainan 232001, Anhui, China;

3. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;

4. School of Earth Sciences and Engineering, Hohai University, Nanjing 210098, China;

5. College of Land Science and Technology, China Agricultural University, Beijing 100193, China)

Abstract: In order to improve the existing phosphorus index assessment methods, using the interactive evaluation index (IEI) as an auxiliary variable, the geographically weighted regression (GWR) was adopted as prediction means. A method of regional soil phosphorus risk assessment was constructed by modifying phosphorus index model (MPIM). The GWR-IEI method more accurately predicted available phosphorus (AP) and soil organic matter (SOM), and the prediction precision and goodness of fit were high. Compared with the ordinary least square (OLS) method, the relative improvement of the root mean squared errors (RMSE) with the GWR-IEI method reached 28.95% for available phosphorus predicted, while that of SOM was 21.24%. The phosphorus loss risk of most of the study area (95.29%) was moderate to low. The areas featuring an extremely high phosphorus loss accounted for merely 0.33% of the total research area. Phosphorus loss depends on the effects of many factors. Areas which have strong source or transfer factors are not necessarily high-risk areas for phosphorus loss. Only the co-occurrence of transfer and source factors leads to high risk and greater potential for phosphorus loss. The GWR-IEI-MPIM method accurately reflected the degree of risk for phosphorus at the regional scale, which provides a valuable reference for risk assessment of phosphorus.

Keywords: high landscape heterogeneity, GWR, county scale, simplified assessment

DOI: 10.25165/j.ijabe.20211402.6185

Citation: Ning H R, Liu Q M, Zhang S W, Ye H C, Shen Q, Zhang W T, et al. Simple assessment of farmland soil phosphorus loss risk at county scale with high landscape heterogeneity. *Int J Agric & Biol Eng*, 2021; 14(2): 126–134.

1 Introduction

Phosphorus is one of the nutrients necessary for the growth of crops, and contributes to eutrophication in aquatic environments. Due to long-term excessive use of phosphate fertilizer, massive quantities of phosphorus have accumulated in soil. Many studies on the risk of phosphorus loss have been conducted across multiple countries and continents^[1-5]. The prerequisite for controlling phosphorus loss is to identify risk areas of phosphorus loss and the factors which influence rates of loss. The most effective way to define the areas at risk of phosphorus loss is to conduct risk evaluations, and many studies have already begun to address this. In the mid-1970s, the non-point source mechanism model began to be widely used. Lemunyon and Gilbert^[2] proposed and improved

the phosphorus index (*PI*), and it rapidly became the most widely-used phosphorus risk evaluation method in the world. However, the *PI* fails to incorporate the effect of different geographic features and conditions of study areas, which have effects on phosphorus loss. Much work has been done to improve and apply the *PI* since it was proposed^[6-10]. However, the existing research still has some shortcomings: (1) in attempts to improve the phosphorus index model, most studies increase the number of metrics used in the evaluation index, such as water quality factors, soil phosphorus retention index, and phosphorus saturation, which is more time-consuming^[11-14]; (2) the GWR model is implemented, allowing for consideration of multiple environmental impact factors as well as local changes in space, while precision is lost when a single geostatistic is used^[15-18]; (3) Most studies do not take into account the scale and spatial heterogeneity of the research area. Therefore, how to further amend the model and develop a simple, fast, and rational method to evaluate phosphorus loss in soil according to region (mountainous areas and plains) and scale (such as the county scale) is worthy of deeper exploration. The county is a basic administrative unit in China, and within these areas, there are differences among some factors which influence phosphorus loss, including rainfall, plantation structure, and soil and water conservation. In addition to landscape heterogeneity, differences in evaluation indicators, scoring, and weights across various studies make it difficult to access data. Therefore, it is highly necessary to select evaluation factors in a targeted way and develop a practical and useful method to evaluate county-level phosphorus risk loss according to county and landform.

Using the interactive evaluation index (IEI) as an auxiliary

Received date: 2020-09-26 **Accepted date:** 2021-01-30

Biographies: Huirong Ning, MS, research interests: environmental resource management, Email: 2394695197@qq.com; Qimeng Liu, PhD, Professor, Doctoral Supervisor, research interests: assessment and prevention of environmental geological hazards, Email: qmliu@aust.edu.cn; Qiang Shen, PhD, research interests: remote sensing of ecological environment, Email: qiangshen@hhu.edu.cn; Wentao Zhang, PhD, research interests: environmental science, Email: wtzhang819@163.com; Zhen Li, PhD, research interests: spatial variability of soil properties, Email: zhenly@cau.edu.cn.

***Corresponding author:** Shiwen Zhang, PhD, Professor, Doctoral Supervisor, research interests: quantitative research of soil. No. 168, Taifeng Road, Tianjiaan District, Huainan 232001, Anhui, China. Tel: +86-18949658012, Email: swzhang@aust@163.com; Huichun Ye, PhD, Associate Research Fellow, research interests: mechanism and application of agricultural remote sensing. No. 9, Dengzhuang South Road, Haidian District, Beijing 100094, China. Tel: +86-13426258830, Email: yehc@aircas.ac.cn.

variable, the geographically weighted regression (GWR) is used to predict the spatial distribution of soil organic matter (SOM) and available phosphorus (AP). The modified phosphorus index model (MPIM) was used to assess the risk of phosphorus loss with GWR and IEL. The purpose of this study was to improve the existing phosphorus risk assessment methods and provide a reference for phosphorus risk assessment and agricultural production in areas with high landscape heterogeneity.

2 Research method system

2.1 Simplified assessment model of soil phosphorus loss risk

There are many ways to evaluate the risk of phosphorus loss, including soil tests for phosphorus, the threshold of soil phosphorus environment, and various models for evaluating nitrogen and phosphorus pollution^[19-21]. The *PI* was proposed by Lemunyon and Gilbert after considering the synergistic effect of multiple factors on phosphorus loss. In the *PI*, eight factors including soil erosion, surface runoff, and the phosphorus background value of soil nutrients were used to establish an evaluation indicator system. These indicators can be divided into source and transfer factors. Source factors include the phosphorus background value of soil nutrients and the quantity of organic fertilizer, while the transfer factors include soil erosion and surface runoff.

For spatial analysis with ArcGIS, Equation (1) was adopted to obtain the index of phosphorus loss risk (*PI*).

$$PI = [\sum (S_i W_i)] \cdot \prod [T_j W_j] \quad (1)$$

where, S_i refers to the corresponding level score of the evaluation indicator i of source factors; W_i is the weight of the evaluation indicator i of source factors; T_j represents the corresponding level

score of the evaluation indicator j of transfer factors; W_j symbolizes the weight of the evaluation indicator j of transfer factors.

Phosphorus loss is influenced by such factors as the physicochemical properties of soil, hydrology, hydrodynamic features, and farmland management^[22,23]. Influencing factors vary with scale. At the county scale, the amount, timing, and method of fertilization are usually associated with land use and planting mode. Differences in rainfall and associated erosion and tillage factors are relatively small, while surface runoff is mainly affected by soil texture and surface slope^[24,25]. After a thorough consideration of highly heterogeneous landscapes, this study amends the *PI* to develop a fast and simple method to evaluate phosphorus loss risk.

(1) Amendment of the indicator system. The multi-layered evaluation indicator system of this study consists of phosphorus source (AP and fertilization factor (FF)) and transfer factors (soil erosion potential (SEP), surface runoff potential (SRP), the distance between potential pollution sources and rivers (DPPSR), and topographic factor (TF)).

In the same county, there can exist slight differences among some factors which influence phosphorus loss, especially the human ones; therefore, they were not used as evaluation factors in this study.

(2) Amendment of the factor calculation method. The fertilization factor was replaced by the type of land utilization^[26]. The surface runoff was substituted with a combination of soil texture and slope. Finally, the widely used method of forecasting soil loss was replaced with the soil erosion factor. The details of these amendments are shown in Table 1.

Parameters for the simplified *PI* index model were obtained based on characteristics of Pinggu District, Beijing, China and relevant literatures^[27-29] (Table 2 and Table 3).

Table 1 Amended treatment and standard of soil phosphorus risk assessment at county scale with high degree of heterogeneity

Evaluating indicators		Units	General computing methods	Supporting literature for general computing methods	Simplified methods
Phosphorus source factors	AP	mg kg ⁻¹	Total phosphorus or available phosphorus	[3,29]	Available phosphorus
	FF	—	Score statistics for organic and inorganic fertilizer amount and fertilization methods	[21,26]	Land use types
Migration factors	SRP	—	Runoff depth	[4,15,20]	Slope and soil texture
	SEP	—	Methods of soil erosion prediction	[28]	K value of soil erosion factor
	DPPSR	m	Linear distance function	[7]	Linear distance function
	TF	—	Slope length index	[27]	Slope length index

Table 2 Soil phosphorus risk assessment index and classification table in the study area

Evaluation indicators		Weight	Index grade and score																																	
Phosphorus source factors	AP /mg kg ⁻¹	1.0	Class	≤3				>3-5				>5-10				>10-20				>20																
			Score	0				1				2				4				8																
	FF	0.5	Class	Others				Forestland				Cropland				Orchard				Vegetable field																
			Score	0				1				2				4				8																
Migration factors	SRP	1.0	Class	Soil texture	Coarse sandy and fine sandy				Sandy loam and light loam				Medium loam and loam				Clay, heavy clay																			
			Score	Slope/(°)	0-3	>3-12	>12-20	>20	0-3	>3-12	>12-20	>20	0-3	>3-12	>12-20	>20	0-3	>3-12	>12-20	>20	0-3	>3-12	>12-20	>20												
	SEP	1.0	Class	Soil texture	Coarse sandy				Fine sandy				Sandy loam				Light loam				Medium loam				Heavy loam				Clay				Heavy clay			
			Score	The mass fraction of SOM/g kg ⁻¹	≤15	>15	≤15	>15	≤15	>15	≤15	>15	≤15	>15	≤15	>15	≤15	>15	≤15	>15	≤15	>15	≤15	>15	≤15	>15	≤15	>15	≤15	>15	≤15	>15				
DPPSR/m	1.0	Class	>2400				>1000-2400				>300-1000				>120-300				≤120																	
		Score	0.2				0.4				0.6				0.8				1.0																	
TF	1.0	Class	≤1				>1-5				>5-10				>10-20				>20																	
		Score	0.6				0.7				0.8				0.9				1.0																	

Table 3 Classification of phosphorus loss risk

Class	PI value	Scalar implicature
Low	≤1	Regional phosphorus loss has little influence on water resources
Moderate	>1-3	Regional phosphorus migration can cause a certain degree of water quality damage. Soil protection or phosphorus management measures should be strengthened to prevent the risk of large phosphorus migration.
High	>3-5	Migration of phosphorus caused serious water pollution, and new soil and water conservation measures or phosphorus management measures are required to reduce the downward movement of phosphorus.
Extremely high	>5	Potential danger to water bodies has reached its limit and urgent restoration actions must be carried out, including soil and water conservation measures and phosphorus management programs that halt the use of phosphate fertilizer.

2.2 Geographic weighted regression

When the target variable is related to an explanatory variable, the optimal linear regression relationship between the target and explanatory variables can be established by means of ordinary least squares (OLS), which is a global relationship^[30,31]. In the geographically weighted regression (GWR) model^[32,33], the regression coefficient of x_0 position target variables on the auxiliary variable no longer use global information and the OLS constant instead approaching observations are estimated locally with a weighted regression coefficient. For position x_0 , the GWR model is

$$y(x_0) = \beta_0(x_0) + \sum_{k=1}^p \beta_k(x_0)q_k(x_0) + \varepsilon(x_0) \quad (2)$$

where, $y(x_0)$ is the target variable at location x_0 ; $q_k(x_0)$ refers to the k th auxiliary variables; $\beta_0(x_0)$ is the intercept; $\beta_k(x_0)$ represents the coefficients of auxiliary variables; p is the number of auxiliary variables; $\varepsilon(x_0)$ is the error. In matrix form, this can be represented as

$$y(x_0) = q(x_0)\beta_0(x_0) + \varepsilon(x_0) \quad (3)$$

where, $\beta_0(x_0)$ is the vector of the series of regression lines at position x_0 ; $q(x_0)$ is the auxiliary variable row vector at x_0 position. At the location x_0 , the regression coefficients are estimated to be

$$\beta_0(x_0) = [q^T W(x_0) q]^{-1} q^T W(x_0) Y \quad (4)$$

where, Y is $n \times 1$ vector of the dependent variable, n is the number of adjacent stations; $q = [I, q_1^T, q_2^T, \dots, q_n^T]$ expresses the design matrix of the auxiliary variable; I is the unit vector, Equation (3) is the diagonal weight matrix at x_0 ; Equation (4) is a vector containing an intercept and the regression coefficients of p auxiliary variables.

The following Gauss function is chosen as the spatial weighting function.

$$w_i(x_0) = \exp\left(-\frac{1}{2}\left(\frac{d_i}{r}\right)^2\right) \quad (5)$$

where, d_i is the distance from the i observation station to the estimated point x_0 ; r is the bandwidth parameter, and $W_i(x_0)$ is the weight. To adapt to the different station densities, the adaptive bandwidth selection strategy and correction information criterion are used to determine the optimal bandwidth. For each point estimated, after the optimal bandwidth and weight function are determined, the local regression coefficients are estimated^[34].

2.3 GWR-IEI-MPIM method

This study established GWR-IEI-MPIM, which integrated the acquisition of basic data with risk evaluation.

(1) The GWR was used to obtain basic data, and the evaluation indicators related to the risk of phosphorus loss in soil were used as explanatory variables to ensure the precision of the data.

(2) The number of evaluation indicators was reduced, and the methods of calculating evaluation factors were amended, to establish a simple PI model for counties with a highly heterogeneous landscape. ArcGIS10.2 software was used to calculate the indicators and comprehensive levels of GWR, OLS, and the phosphorus risk evaluation form; SPSS20.0 software was used for the stepwise regression analysis.

2.4 Analysis on the effect of spatial prediction

The differences between the observed responses (values of the variable being predicted) in the given dataset and those predicted by a linear function of a set of explanatory variables. Visually, this is the sum of the squared vertical distances between each data point in the set and the corresponding point on the regression line the smaller the differences, the better the model fits the data. The resulting estimator can be expressed by a simple equation, especially in the case of a single regressor on the right-hand side^[35]. Cross verification via the root mean square error (RMSE) was used to measure the predictive precision of different methods, and to evaluate the accuracy of prediction. A lower RMSE value indicated a more accurate prediction.

The OLS method was used for comparison. OLS is a method for estimating the unknown parameters in a linear regression model, with the goal of minimizing the sum of the squares of the equation is as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (z(x_j) - \bar{z}(x_j))^2} \quad (6)$$

where, $z(x_j)$ is the measured values; $\bar{z}(x_j)$ is the predicted values; n is the number of samples in the validation set.

3 Risk assessment of regional phosphorus loss based on the GWR-IEI-MPIM

3.1 Research area

This study adopted GWR-IEI-MPIM to conduct an empirical study on the evaluation of phosphorus risk in soil in counties with a highly heterogeneous landscape, using Pinggu District in Beijing as a case study. Pinggu District is located between 116°55'20"E to 117°24'09"E and 40°01'44"N to 40°22'39"N, located in northeast Beijing (Figure 1), Pinggu District is a warm temperate continental monsoon climate zone, with four distinct seasons, dry and windy in spring, high temperature and rainy in summer. Pinggu District connects the south side of Yanshan Mountain and the north side of the North China Plain. Showing a northeast-southwest trend, it is smooth in the middle and takes the shape of a pan. With a width of 35.5 km from east to west and a length of 30.5 km from north to south, it covers a total area of 1075 km². Among them, the northern region is a mountainous, hilly region, and the southwest is plain, over 60% of the district is comprised of hills and mountains, and the vegetation coverage rate reached 36%. The landscape is highly heterogeneous. The soil in the research area was divided into four types (brown, cinnamon, moisture, and paddy soils) and 11 sub-types.

3.2 Data collection and processing

The digital elevation map (25 m×25 m), soil map (1:50000), modified map (1:10000) of land use in 2015, and the water system map of the research area were collected (Figure 2). The soil map includes such elements as soil type and texture. The testing points were scattered in the form of a grid, and stratified sampling was conducted according to the soil type and land use of the research

areas. After the sampling points were determined in the field, they were located with GPS, and the longitude and latitude were recorded. The soil samples from 5 points in a diameter of 10 m at each sampling point were mixed and taken back to the lab as the samples to be analyzed. There were altogether 1578 sampling points (Figure 1).

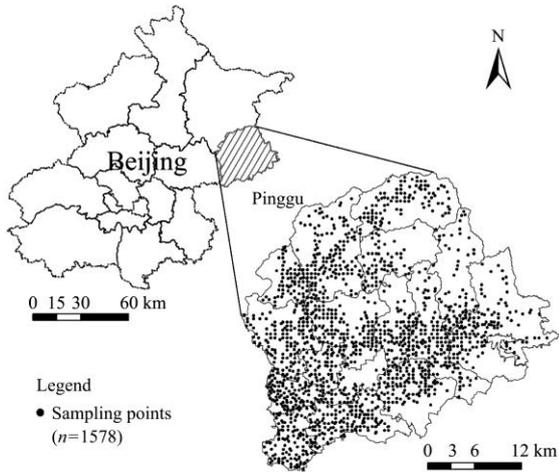


Figure 1 Location and distribution of sample locations in the study area

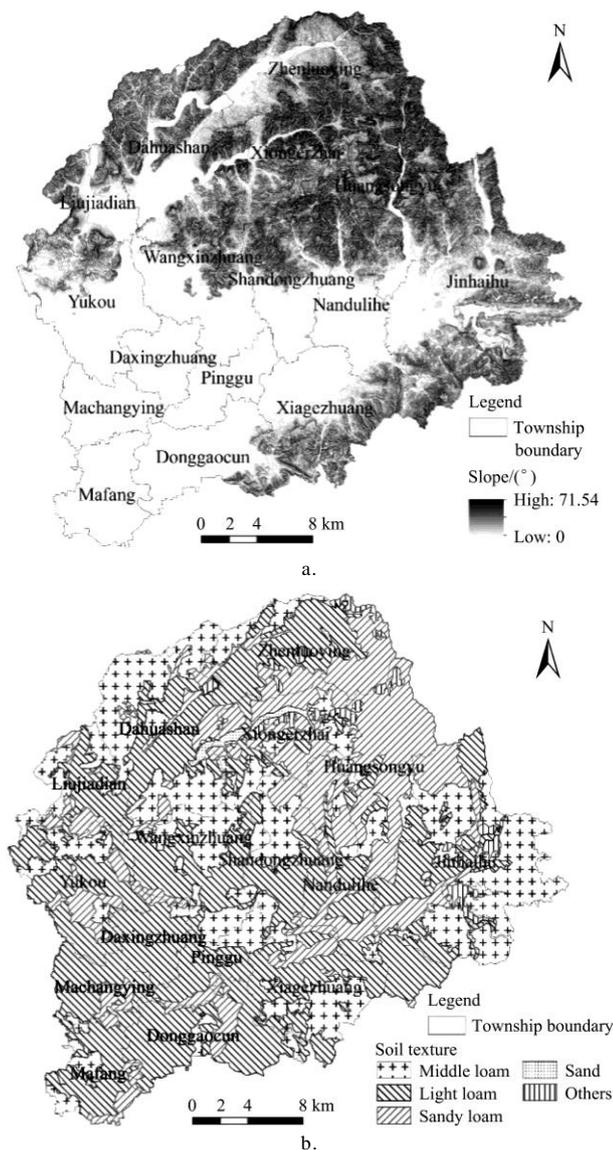


Figure 2 Slope (a) and soil texture (b) map of the study area

The Olsen-P method was used for the AP test, and the potassium dichromate volumetric method tested SOM.

3.3 Results

3.3.1 Spatial prediction of AP and SOM based on the GWR-IEI method

The GWR method was adopted for spatial prediction of AP and SOM, and the interaction was used for the GWR model. The phosphorus source and transfer factors that did not involve AP and SOM were taken as the independent variables, and their values were used for the GWR prediction. The results were compared with those of traditional OLS to test the accuracy.

As the spatial distribution of available phosphorus and organic matter in soil tends to concentrate in certain areas, the adjustable spatial core was used for the GWR model and the double square weight function was taken as the spatial weight function. Corrected akaike information criterion (AICc) was used as the measurement standard, and a golden section search was employed to search the bandwidth which led to the minimum AICc, determining the optimal bandwidth. There were four factors of phosphorus loss risk which were not influenced by AP or SOM, including fertilization, surface runoff, the distance between potential pollution sources and rivers, and landform. SPSS20.0 software (I-Corp, 2013) was used for the stepwise regression of the four indicators. According to the stepwise regression, the auxiliary variables for the GWR prediction of SOM were the distance between potential pollution sources and rivers as well as surface runoff; the auxiliary variables for the GWR prediction of available phosphorus in soil were the distance between potential pollution sources and rivers as well as fertilization. According to the results of the ANOVA, the fitted equation of the GWR method was significant ($F_{SOM}=19.63, p<0.001; F_{AP}=14.87, p<0.001$), the goodness of fit was high ($R^2_{AP}=0.61, R^2_{SOM}=0.58$).

The OLS method was taken for comparison. Root mean square error (RMSE) was used to measure the predictive precision of different methods and to evaluate the accuracy of prediction.

The GWR showed a higher prediction accuracy in the scatter fitting of different measured and predicted values and the RMSE when the prediction target was the same (Figure 3). Compared with the OLS method, the GWR method showed a relative increase of 28.95% in the RMSE of AP and a relative increase of 21.24% in the RMSE of SOM. The predicted values from the OLS method were concentrated in a small scope and showed significant central tendency, but failed to reveal the spatial instability of parameters in different spaces.

The GWR-based content of AP in Pinggu District ranged from 0.5 mg/kg to 183.28 mg/kg, and that of SOM ranged from 1.97 mg/kg to 28.73 mg/kg.

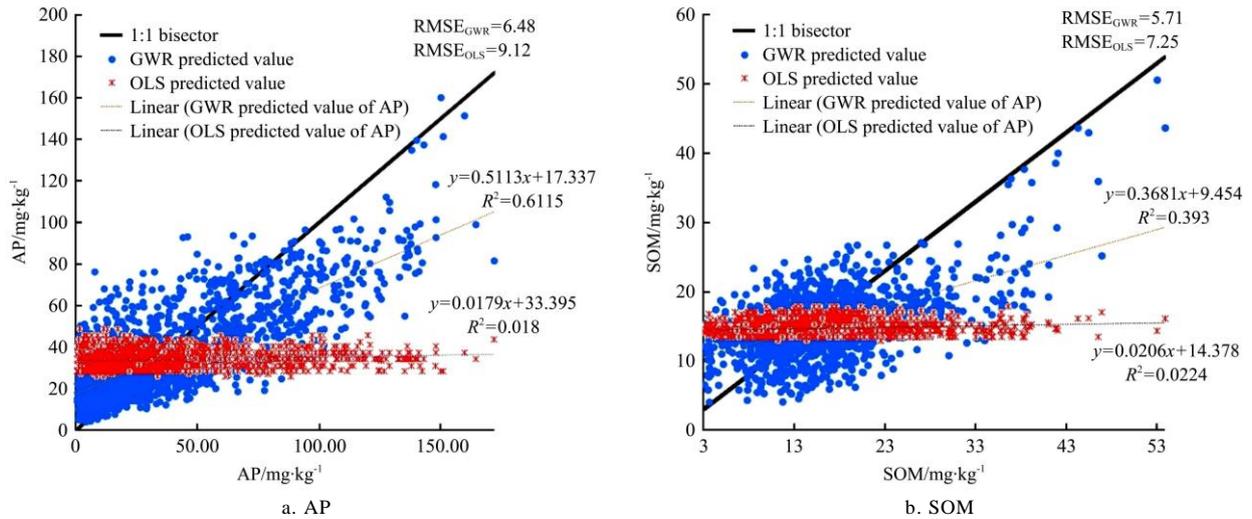
3.3.2 Acquisition of indicator value and score division

(1) AP

The GWR method, together with the information described in Table 2, was adopted to obtain the spatial distribution map of the content of available phosphorus in soil (Figure 4a). The AP content of the research area was largely high. However, the phosphorus content in the north and south was relatively low, while that of the transition zones of mountainous areas and plains was relatively high. The score for AP was 8 in most areas (65.88% of the total area), especially in the three townships and towns (Yukou, Dahuashan and Zhangjiadian) in the southwest. These areas traditionally include high concentrations of gardens and vegetable

fields, and this long-term investment in agriculture has resulted in the high content of available phosphorus in the local soil, with a correspondingly high risk of phosphorus loss. Areas with an AP score of 4 were concentrated around the towns of Machangying,

Daxingzhuang, Yukou, Shandongzhuang, and Mafang, as well as scattered areas in Zhenluoying Town, Xiong'erzhai Township, and Jinhaihu Town. Areas with a score of 4 accounted for 25.44% of the total research area (Figure 4a).



Note: RMSE is the root mean square error; OLS represents Ordinary Least Square; GWR is Geographically weighted regression.

Figure 3 Scatter plots of measured and predicted values of AP and SOM

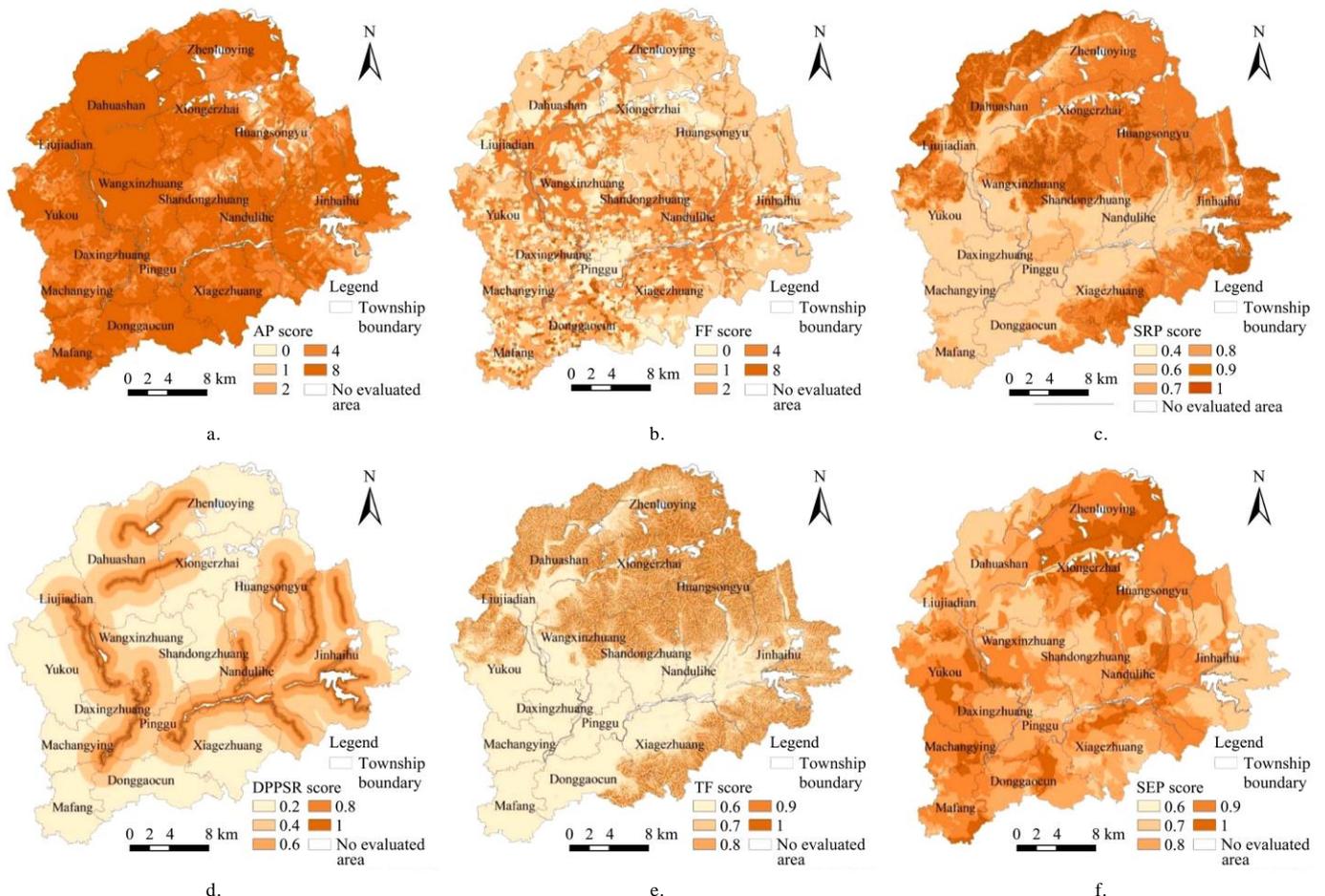


Figure 4 Evaluation results of single factor a, b, c, d, e, and f represent AP, FF, SRP, SEP, DPPSR and TF scores respectively

(2) FF

The map of land utilization in 2015 was used as the base map. Classification and integration were conducted according to the methods described in Table 2, and objects in the map were rasterized, with each lattice 25 m×25 m. Fertilization scores were standardized for various parts of the research areas (Figure 4b). As land was considered in fertilization, the spatial distribution and

land utilization of the fertilization scores were consistent. Areas with a score of 1 (forest land) occupied 36.45% of the total area; areas with a score of 2 (farmland) represented 11.29% of the area; a score of 4 (gardens) accounted for 23.70% of the area. In total, areas with low scores (0-4) occupied 63.38%, indicating that the study area is primarily areas with low scores (Figure 4b). The vegetable fields and gardens in the research area were concentrated

on the flat land between the mountainous areas and plains in Pinggu District. The fertilization score of these areas is high, and the risk of phosphorus loss is relatively higher.

(3) SRP

In the rasterized soil texture map, each grid was 25 m×25 m. Slope was extracted from the DEM, and the size of the lattice was 25 m×25 m. The soil texture map and the slope were spatially superimposed to obtain the slope and texture type of each lattice. The score map (Figure 4c) of SRP of the research areas was drawn according to Table 2.

As the topography of the research area was complex, the scores of the surface runoff potential of portions of the research area were high, ranging from 0.6 to 1.0. The areas where the score of surface runoff potential was 0.6 accounted for 38.30% of the total area, and occurred primarily in the southwest. The altitude of these areas is low, and the slope was lower than 3°. Some areas were influenced by the soil texture (medium loam), and the score for their surface runoff potentials was 0.7. The areas where the score of surface runoff potential was 0.8 were mainly in the northeast, representing 8.12% of the total area. The altitude of these areas is high, and with a slope higher than 20°, the surface runoff potential was high. The scores of the northwest, center, and east were 0.9 and 1.0, and these areas accounted for 47.22% of the total area. The surface runoff potential of these areas was the highest (Figure 4c).

(4) SEP

The GWR was adopted for the spatial prediction of SOM, and the results of the spatial prediction were re-classified with 15 g/kg as the threshold. The rasterized maps of the soil texture and the content of organic matter in soil were used to obtain the score chart (Figure 4d) of soil erosion potential of the research area through spatial superposition. On the whole, the soil erosion potential of the research area was high, and the areas where scores were 0.9 or 1.0 accounted for 47.02% of the total area. Most of these areas are in the northeast and southwest, where the soil texture is either light loam or sandy loam. The areas with scores of 0.7 or 0.8 accounted for 52.84% of the total area. Most of these areas are in the center of the study area, and the main soil texture is medium loam. The content of organic matter was low, and so was the score of soil erosion potential. The distribution of soil erosion potential was correlated with that of soil texture, but negatively related to the content of available soil phosphorus in the research area.

(5) DPPSR

The hydrological analysis tool of ArcGIS extracted the water system network of the research area. The straight-line distance function of ArcGIS was used to obtain the distance between pollution sources and rivers based on the water system map. The buffer tool of ArcGIS was applied to draw the map of buffer areas of rivers according to the river distribution map of the research area, at five levels (≤120 m, >120-300 m, >300-1000 m, >1000-2400 m, and >2400 m). The score chart (Figure 4e) of distance between potential pollution sources and rivers of the research area was generated according to standards. A shorter distance would lead to a higher score and a greater risk of phosphorus loss, while a longer distance would lead to a lower score and a smaller risk of phosphorus loss. The areas with a distance of ≤120 m, >120-300 m, >300-1000 m, >1000-2400 m and >2400 m from the rivers accounted for 3.58%, 5.10%, 18.85%, 29.05%, and 43.42% of the total research area, respectively.

(6) TF

The TF was calculated using the method proposed by

Wischmeier et al^[36]. in ArcGIS 10.2 software. Areas with low landform scores were in the center and southwest of the research area, accounting for 60.16% of the total research area (Figure 4f). These areas are plains, featuring a low slope and a flat landform. The areas with high landform scores (0.9 and 1.0) were in the northwest, north, east, and southeast of the Pinggu District, occupying 26.17% of the total research area. These areas feature moderate and low hills, steep slopes, and undulating landforms. The landform scores accurately reflected the landform of the Pinggu District.

3.3.3 Comprehensive evaluation of phosphorus risk

The score charts obtained with the phosphorus source and transfer factors, Equation (1), and Table 3 were used to obtain the phosphorus risk map and the level area proportion map (Figure 5) of the townships and towns in the research area.

GWR was used to obtain the basic data, the simplified and corrected PI model was applied to the evaluation of phosphorus loss risk in the soil of Pinggu District (Figure 5). The phosphorus loss risk was higher in the north and mountainous areas, and lower in the south and plains. Loss risk showed a belt-shaped distribution near rivers, where risk was high near water and decreased with distance (Figure 5). This distribution was the result of the effects of several factors. In the research area, the risk extent of transfer factors was negatively correlated with that of phosphorus source factors as a whole. Human intervention was reduced in the areas where landform was highly heterogeneous and transfer factor scores were high, and the scores of phosphorus source factors were low.

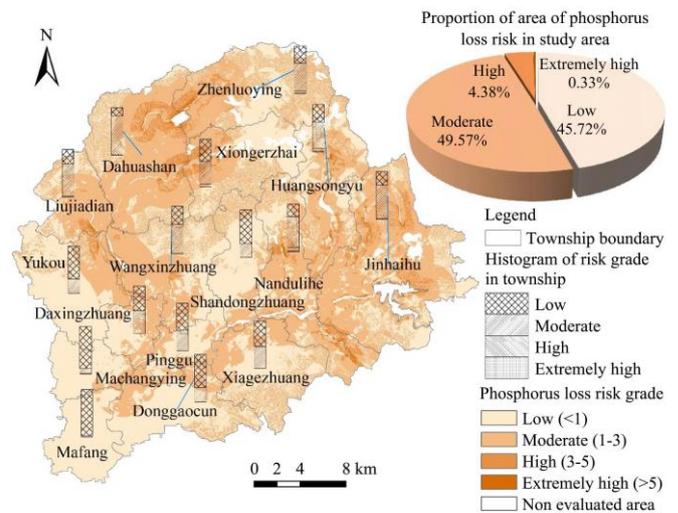


Figure 5 Grade distribution and area ratio of phosphorus loss risk in Pinggu District based on the GWR-IEI-MPIM method

The phosphorus loss risk of most of the research area was moderate to low, and the extent of the areas where phosphorus loss risk was high was small (Figure 5). The areas featuring an extremely high phosphorus loss risk accounted for merely 0.33% of the total research area, and the areas with high phosphorus loss risk occupied 4.38% and occurred in the northeast and west of the research area, including the Cuohe River Basin of Liujiadian Township and the areas around Xiyu Reservoir of Dahuashan Town. Most of the land in these areas was used for gardens and vegetable fields, and featured a high content of available phosphorus in soil. Additionally, most of these areas are characterized by hills, a high slope, serious soil erosion, and proximity to rivers, all of which could easily contribute to phosphorus loss. The areas with a moderate risk of phosphorus

loss accounted for 45.72% of the research area, and were concentrated on both sides of rivers, around reservoirs, in the upper reaches of Cuohe River and the Xunhe River of Liujiadian Township, and around Xiyu Reservoir of Dahuashan Town and Haizi Reservoir of Jinhaihu Town. Several factors contribute to this elevated risk. Areas within 2000 m of a river had high available phosphorus content, and pollutants would easily flow into the rivers. However, these areas feature plains and a low soil erosion potential; hence, the level of the phosphorus loss risk was “moderate”. The areas featuring a low risk of phosphorus loss took up the largest proportion (49.57%) of the research area.

The reliability of this method proposed in this paper was tested by OLS method and RMSE. The phosphorus index model adopts the general calculation method in Table 2. The results are shown in Figure 6.

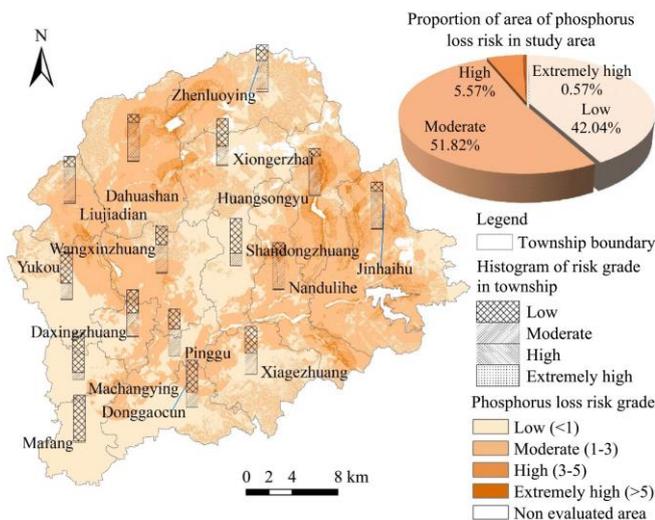


Figure 6 Grade distribution and area ratio of phosphorus loss risk in Pinggu District based on conventional methods

The results of the two methods are very similar, when used for the spatial distribution pattern, the size of the different grades, or the proportion of the different grades of each township. The GWR-IEI-MPIM method is feasible for assessing soil phosphorus risk at county level. This method can not only ensure the accuracy of the data but the calculation process is relatively simple, which saves time and material resources. However, this method is best used at the county scale, and with such a large scale, some indexes can not be simplified.

4 Discussion

4.1 Soil phosphorus risk evaluation based on the GWR-IEI method

The GWR method is a partial regression model. As the regression parameters of the independent variables change with spatial location, the model has gradually become valuable for research on the spatial distribution of soil properties^[37]. Qu et al.^[38] assert that the GWR is a well-developed spatial statistical method in both theory and practice, and that it can reveal some partial changes which may be concealed by spatial instability and reflect more reliable spatial changes in soil properties. This method has been widely applied to such fields as soil, land, and economic considerations. A rational evaluation of the regional risk of soil phosphorus loss depends on the accuracy of data. Based on the GWR, the spatial prediction was made on the basic data about available phosphorus and organic matter in soil necessary for evaluation. It was found that the GWR could be

used to obtain a more accurate spatial distribution map of available phosphorus and organic matter in soil. This result is consistent with that of similar studies. Guo^[39] employed the GWR to make a spatial prediction of soil pH, organic matter, available phosphorus, available potassium, and Alkaline-N. The study showed that the GWR model was accurate in predicting the indicators of soil properties. Yang et al.^[40] also indicate that the GWR model is more precise for predicting soil properties. Tan et al.^[41] proposed a mixed GWR for spatial analysis, and found that the mixed GWR could detect the spatial instability of spatial statistical relations. The results of that analysis were highly consistent. The traditional OLS method merely focuses on an “average” or “overall” estimation of parameters and cannot reveal the spatial instability of parameters; Traditional linear regression analysis and other overall situation-oriented methods neglect the spatial properties of geographic issues and cannot make an effective description of the spatial model. Modeling of spatial autocorrelation incorporates spatial dependence into the traditional regression analysis, while spatial regression analysis introduces items which can describe spatial autocorrelation and spatial instability and can effectively solve the problems of the traditional linear regression model. Spatial regression analysis includes the spatial autocorrelation model and the GWR model. The parameters in the spatial autocorrelation regression model do not change with spatial location and it is, in essence, an overall situation-oriented model. GWR, however, is a non-parameter partial spatial regression analysis which models the relationships among independent and dependent variables in the sub-area of different spaces according to spatial change. In this method, the regression coefficients of independent variables change with spatial location. As a simple and effective technology, GWR can show the spatial instability of parameters in different spaces. Therefore, it is strong in the partial analysis of spatial data.

An explanatory variable also plays a key role in the GWR-based prediction of a soil property. This paper proposed the interactive selection method—selectively entering the GWR model through the interaction with the indicators of phosphorus risk evaluation. In our GWR, the phosphorus source and transfer factors unrelated to available phosphorus and organic matters in soil are taken as the independent variables and the scores as the values of independent variables. In this study, phosphorus risk evaluation included six indicators: the content of available phosphorus in soil, fertilization, soil erosion potential, surface runoff potential, the distance between potential pollution sources and rivers, and landform. Of these six indicators, the four which are free from the effect of available phosphorus and organic matter in soil include fertilization, surface runoff potential, the distance between potential pollution sources and rivers, and landform. Most of these indicators are comprehensive ones and are closely related to the soil, land, and landform of the research area. They can be taken as the auxiliary variables of the spatial prediction of available phosphorus and organic matter in soil. SPSS20.0 was used for the stepwise regression of four indicators. After the stepwise regression, the auxiliary variables were the distance between potential pollution sources and rivers and surface runoff, and those utilized in the GWR prediction of available phosphorus in soil were fertilization and the distance between potential pollution sources and rivers. The stepwise regression was used to select the explanatory variables which influenced phosphorus loss. In consideration of possible spatial instability in the effect of regression explanatory variables on phosphorus loss, the GWR

model was used for analysis again.

To sum up, the GWR-interactive evaluation indicators method can provide precise statistical sources for regional evaluation of phosphorus risk.

4.2 Amendment of the soil phosphorus risk evaluation model

The *PI* is an effective method to evaluate the potential risk of phosphorus loss, by defining key source areas and rationally controlling agricultural non-point source pollution. The *PI*-based evaluation of soil phosphorus risk requires a large quantity of data; it is difficult to obtain data, and investment can be great. In recent years, much work has been done to improve and apply the *PI*. For instance, the distance between potential pollution sources and water bodies is considered at the basin scale; the risk of losing particle and solution phosphorus is used to evaluate the effect of land management on the loss of different forms of phosphorus; the leaching potential and connectivity between source areas and water are taken as indicators of transfer factors; the type of receiving water is used to recognize the importance of phosphorus loss.

Adding relevant evaluation indicators according to regional features improved the model, but has cost more time and labor than the original already complicated model computation. Therefore, how to further simplify and correct the model and develop a simple, fast, and rational method to evaluate the phosphorus loss risk in soil according to region (mountainous areas and plains) and scale (such as county) is worthy of deeper exploration. To some extent, the difference among some indicators is not very significant. As a result, these indicators are uninformative and will not offer any additional reference information to the area-based governance of regional phosphorus. Therefore, the feature and size characteristics can be combined in the actual evaluation to simplify and correct the *PI* model. After correction, this paper proposed a multi-layered evaluation indicator system consisting of phosphorus source factors (available phosphorus in soil and fertilization) and transfer factors (the potential of soil erosion, the potential of surface runoff, the distance between potential pollution sources and rivers, and landform). The calculation of fertilization, surface runoff, and soil erosion were simplified and corrected according to size and regional feature; the fertilization factor was replaced by the type of land utilization; surface runoff was replaced by the combination of soil texture and slope; the method of predicting soil loss was calculated by the soil erosion. The soil phosphorus risk of the research areas is highly differential, and the risk of phosphorus loss depends on source and transfer factors which co-occur and interact with each other (Figure 5). The areas which have strong source or transfer factors are not necessarily high-risk areas of phosphorus loss. Only the cooccurrence of high-risk transfer and source factors lead to a high risk level of phosphorus loss and greater potential for phosphorus loss.

5 Conclusions

This research amended the *PI* to develop a simple and rational method to evaluate phosphorus loss risk at the county scale in heterogeneous county-scale landscapes, and realized assessment of the risk of phosphorus loss. Conclusions are drawn as follows:

(1) The GWR showed a higher prediction accuracy in the scatter fitting of different measured and predicted values and the RMSE when the prediction target was the same. Compared with the OLS method, the GWR method showed a relative increase of 28.95% in the RMSE of AP and a relative increase of 21.24% in the RMSE of SOM. The GWR-IEI method could be used to obtain a more accurate prediction of available phosphorus and

organic matter in soil, and the prediction precision and the goodness of fit were high. However, this method is best used at the county scale, and with such a large scale, some indexes will not be simplified.

(2) The GWR-IEI-MPIM method can better distinguish the phosphorus loss risk in regional soils and offer reference information for the evaluation of phosphorus risk for the same area. The areas featuring an extremely high phosphorus loss risk accounted for merely 0.33% of the total research area, and the areas with high phosphorus loss risk occupied 4.38% and occurred in the northeast and west of the research area. The areas with a moderate risk of phosphorus loss accounted for 45.72% of the research area, Areas within 2000 m of a river had high available phosphorus content, and pollutants would easily flow into the rivers. However, these areas feature plains and a low soil erosion potential; hence, the level of the phosphorus loss risk was “moderate”.

(3) The phosphorus loss risk for the studied area was highest in the north, in mountainous areas, and near rivers (where it showed a belt-shaped distribution) and reservoirs, and lower in the south and flat areas. The soil phosphorus risk of the research areas is highly differential, and the risk of phosphorus loss depends on source and transfer factors which co-occur and interact with each other.

Acknowledgements

This research was funded by the National Key Research and Development Program of China (2016YFD0300801, 2020YFC1908601) and the National Natural Science Foundation of China (41471186).

[References]

- [1] Romkens M J, Nelson D W, Manner J V. Nitrogen and phosphorus composition of surface runoff as affected by tillage method. *Journal of Environment Quality*, 1973; 2(2): 292–295.
- [2] Lemunyon J L, Gilbert R G. The concept and need for a phosphorus assessment tool. *Journal of Production Agriculture*, 1993; 6(4): 483–486.
- [3] Sharpley A N, McDowell R W, Weld J L, Kleinman P J A. Assessing site vulnerability to phosphorus loss in an agricultural watershed. *Journal of Environmental Quality*, 2001; 30(6): 122–132.
- [4] Wang X Y, Wang Y X, Wang X F, Wang Z G, Wang Q P, Qiuju H U. The character of nutrient loss and land use in a small watershed of Miyun reservoir. *Research of Environmental Sciences*, 2003; 16(1): 30–33. (in Chinese)
- [5] McDowell R W, Sharpley A N, Brookes P C, Poulton P. Relationship between soil test phosphorus and phosphorus release to solution. *Soil Science*, 2001; 166(2): 137–149.
- [6] Jokela W E. Phosphorus index for Vermont: Background, rationale, and questions, Paper presented at Annual meeting of SERA-17 Minimizing P Loss from Agriculture, Quebec City, Canada, June 10th, 1999.
- [7] Gburek W J, Sharpley A N, Heathwaite L, Folmar G J. Phosphorus management at the watershed scale: a modification of the phosphorus index. *Journal of Environmental Quality*, 2000; 29(1): 130–144.
- [8] Kleinman P. Phosphorus and agriculture V: the New York P Index. *New York State Cooperative Extension Bulletin*, 2001; 10(3): 4–5.
- [9] Sharpley A N, Withers P J A. The environmentally-sound management of agriculture phosphorus. *Fertilizer Research*, 1994; 39: 133–146.
- [10] Coale F J, Sims J T, Leytem A B. Accelerated deployment of an agricultural nutrient management tool: the Maryland phosphorus site index. *Journal of Environmental Quality*, 2002; 31(5): 1471–1476.
- [11] Hughes K J, Magette W L, Kurz I. Identifying critical source areas for phosphorus loss in Ireland using field and catchment scale ranking schemes. *Journal of Hydrology*, 2005; 304(1-4): 430–445.
- [12] Gao C, Zhang T L, Wu W D. Phosphorus sorption and desorption of agricultural soils under different land uses. *Environmental science*, 2001; 22(4): 67–71.
- [13] Zhang S R, Chen L D, Fu B J, Li J R. The risk assessment of nonpoint

- pollution of phosphorus from agricultural lands: a case study of Yuqiao reservoir watershed. *Quaternary Sciences*, 2003; 23(3): 262–269. (in Chinese)
- [14] Zhou H P, Gao C. Identifying critical source areas for non-point phosphorus loss in Chaohu Watershed. *Environmental Science*, 2008; 29(10): 2696–2702. (in Chinese)
- [15] Wu J S, Li B, Huang X L. Spatio-temporal dynamics and driving mechanisms of resident trip in small cities. *Journal of Geo-Information Science*, 2017; 19(2): 38–46. (in Chinese)
- [16] Cre R B, Aut D Y. Geographically Weighted Regression. *Journal of the Royal Statistical Society*, 2015; 47(3): 431–443.
- [17] Fotheringham A S, Brunsdon C, Charlton M. Geographically weighted regression: The analysis of spatially varying relationships. *International Union of Crystallography*, 2004; 86(2): 554–556.
- [18] Kumar S, Lal R, Liu D S. A geographically weighted regression kriging approach for mapping soil organic carbon stock. *Geoderma*, 2012; 189–190: 627–634.
- [19] Ahuja L R, Sharpley A N, Yamamoto M, Menzel R G. The depth of rainfall-runoff-soil interaction as determined by ³²P. *Water Resource Research*, 1981; 17(4): 969–974.
- [20] Wang L H, Wang X Y, Zhang Z F. Establishment and revision of PI. *Transactions of the CSAE*, 2006; 27(6): 85–88. (in Chinese)
- [21] Hao F H, Cheng H G, Yang S T. Non-point source pollution model: theory, methods and applications. Environmental Science Press, Beijing, China, 2006; pp.154–196. (in Chinese)
- [22] Han X X, Xiao J, Wang L Q, Tian S H, Liang T, Liu Y J. Identification of areas vulnerable to soil erosion and risk assessment of phosphorus transport in a typical watershed in the Loess Plateau. *The Science of the Total Environment*, 2020; 758: 143661–143669.
- [23] Huang J, Xu C C, Ridoutt B G, Wang X C, Ren P A. Nitrogen and phosphorus losses and eutrophication potential associated with fertilizer application to cropland in China. *Journal of Cleaner Production*, 2017; 159: 171–179.
- [24] Guzha A C, Rufino M C, Okoth S, Jacobs S, Nðrega R L B. Impacts of land use and land cover change on surface runoff, discharge and low flows: Evidence from East Africa. *Journal of Hydrology-Regional Studies*, 2017; 15: 49–67.
- [25] Smith D R, King K W, Johnson L, Francesconi W, Richards P, Baker D, et al. Surface runoff and tile drainage transport of phosphorus in the midwestern United States. *Journal of Environmental Quality*, 2015; 44(2): 495–502.
- [26] Zhang S W, Wang P P, Ye H C, Huang Y F, Mu E L, Wang L B, et al. Risk simplified assessment on phosphorus loss risk based on digital soil system at county scale. *Transactions of the CSAE*, 2012; 28(11): 110–117. (in Chinese)
- [27] Wischmeier W H, Smith D D. Rainfall energy and its relationship to soil loss. *American Geophysical Union*, 1958; 39(2): 285–291.
- [28] Renard K G, Foster G R, Weesies G A, McCool D K, Yoder D C. Predicting soil erosion by water: A guide to conservation planning with the revised universal soil loss equation (RUSLE). United States Department of Agriculture. *Agricultural Handbook*: Washington D C, USA, 1997; pp.537–537.
- [29] Sharpley A N, Smith S J. Prediction of soluble phosphorus transport in agricultural runoff. *Journal of Environment Quality*, 1989; 18(3): 313–316.
- [30] Keskin H, Grunwald S, Harris W G. Digital mapping of soil carbon fractions with machine learning. *Geoderma*, 2019; 339: 40–58.
- [31] Dadashpoor H, Azizi P, Moghadasi M. Land use change, urbanization, and change in landscape pattern in a metropolitan area. *Science of the Total Environment*, 2019; 655: 707–719.
- [32] Yu H C, Fotheringham A S, Li Z Q, Oshan T, Kang W, Wolf L J. Inference in multiscale geographically weighted regression. *Geographical Analysis*, 2020; 52(1): 87–106.
- [33] Arabameri A, Pradhan B, Rezaei K. Gully erosion zonation mapping using integrated geographically weighted regression with certainty factor and random forest models in GIS. *Journal of Environmental Management*, 2019; 232: 928–942.
- [34] Shi S Q, Cao Q W, Yao Y M, Tang H J, Yang P, Wu W B, et al. Influence of climate and socio-economic factors on the spatio-temporal variability of soil organic matter: A case study of central Heilongjiang Province, China. *Journal of Integrative Agriculture*, 2014; 13(7): 1486–1500. (in Chinese)
- [35] Williams M N, Grajales C A G, Kurkiewicz D. Assumptions of multiple regression: Correcting two misconceptions. *Practical Assessment Research & Evaluation*, 2013; 18(11): 14–26.
- [36] Shi H, Zhang L, Liu J. A new spatial-attribute weighting function for geographically weighted. *Canadian Journal of Forest Research*, 2017; 36(4): 996–1005.
- [37] Qu M K, Li W D, Zhang C R, Huang B. Geographically weighted regression and its application prospect in soil and environmental sciences. *Soils*, 2014; 2014(1): 15–22.
- [38] Wu L, Li P W. Studies on rainfall process, flow-making and nitrogen and phosphorus losses. *Acta Science Circumstantiate*, 1996; 16: 111–116.
- [39] Guo L. Spatial distribution characteristics and prediction of multi-scale cultivated land based on spatial regression model. *Huazhong Agricultural University*, 2013; 3: 16–23. (in Chinese)
- [40] Yang S H, Zhang H T, Guo L, Ren Y. Spatial interpolation of soil organic matter using regression kriging and geographically weighted regression kriging. *Journal of Applied Ecology*, 2015; 26(6): 1649–1656.
- [41] Tan W Z, Wang J M, Liu M L. Algorithm for mixed geographically weighted regression model. *Geomatics & Information Science of Wuhan University*, 2007; 32(2): 115–119. (in Chinese)