The Moran′s I and PCA Analysis Study on Evaluation Index of Soil Fertility Based on the Tilling Layer of Northeast Slope in China

Nandi[1, a], Xu Sun[1, b], Yanan Li[1, c] and Xiaoguang Wu[2, d]

[1]School of Inner Mongolia Agricultural University, Huhhot 010020, China;

[2]Inner Mongolia Autonomous Region Land Regulation Center, Huhhot 010019, China.

[a]nandi1115@163.com, [b]1637772590@qq.com, [c]314379711@qq.com, [d]379911627@qq.com

Abstract

195 samples were collected in Inner Mongolia Autonomous Region of China based on the survey of northeast slope tillage layer. Using Moran index analysis Moran's I and principal component analysis PCA the minimum data set (MDS) of slope tillage layer in Northeast China was established. The characteristic index of tillage layer was analyzed by using MDS soil fertility comprehensive index (Integrated fertility index-Moran's, IFI-MR and Integrated fertility index-PCA, IFI-PCA) and total data set soil fertility comprehensive index (Integrated fertility index, IFI-TDS). There were obvious differences in soil quality characteristics in this study area, and the thickness of tillage layer (25.27 ± 0.45) cm, soil bulk density (1.37 ± 0.01) g·cm-3, pH (7.10 ± 0.08), organic matter (36.04 ± 1.44)g·kg-1, total nitrogen(3.25 ± 0.22)g·kg-1,available phosphorus(54.90±4.00)mg·kg-1,available potassium (132.09 ± 6.35) mg·kg-1. The whole area is moderately barren. After linear regression analysis, IFI-TDS and IFI-MR, IFI-PCA are positively correlated, R2 is 0.330 and 0.702, Nash effective coefficients are 0.857 and 0.900, relative deviation coefficients are 0.014 and 0.003, and the variation range IFI-MR> IFI-T> IFI-PCA, the average value IFI-T> IFI-PCA> IFI-MR, the coefficient of variation is IFI-MR> IFI-PCA> IFI-T, so choose MDS-PCA instead of TDS pair The soil quality of the cultivated layer is evaluated, and the indicators selected by MDS-PCA are pH, organic matter, soil bulk density, cultivated layer thickness, and available phosphorus. According to the comprehensive index of soil fertility, the number of optimal I samples is 4, accounting for 2.05%, the number of excellent II samples is 39, accounting for 20%, the number of good III samples is 42, accounting for 21.54%, the number of moderate IV samples is 66, accounting for 33.5%, the number of poor V samples is 28, accounting for 14.36%, and the number of extremely poor VI samples is 16, accounting for 8.21%.

Keywords

Cultivated land; Soil quality; Comprehensive index of soil fertility; Minimum data set.

1. Introduction

Under the pressure of our large agricultural country to feed a large number of people with a small amount of land resources, the outstanding contradictions in agricultural development have intensified. Under this background, it is necessary to protect the ecological mechanism, adapt measures to local conditions, and protect production, fertilizer and water. The development and utilization of cultivated land is too large and the quantity is gradually decreasing, which leads to the decline of soil quality and serious soil erosion. In order to solve the main contradiction between grain supply and demand, it is necessary to adjust and improve soil structure and water-gas distribution reasonably and scientifically [1]. The main scientific method for improving soil fertility is to establish an evaluation system for the suitability of cultivated land. The key is to select the optimal evaluation index. On the basis of the original variables collected, the data set with higher correlation degree and more direct reflection of cultivated land quality, that is, the minimum data set (MDS) should be selected.

A minimum set of data was proposed by the Larson in monitoring land dynamics, and some independent indicators were screened from the basic data on soil environment and nutrients [2]. To establish data sets based on the extent of soil impact on cultivated land, Andrews S.S., Sparling G.P. and Bram Govaerts to USDA-NRCS the development of data [3,4,5] Such data sets have also been mentioned in studies of soil fertility assessment and quantitative expression through soil physical, chemical properties and topographic factors. It can be seen that as a method of land quality evaluation, the evaluation indexes of different soil species, different soil genera and different climatic conditions and environmental conditions are also different for different regions. This method is widely used in the study of land use patterns at home and abroad. the characteristics of land fertility were analyzed by cluster analysis, combined with minimum data set[6] , total data set and soil fertility comprehensive index[7] . so on ,to carry on the quantitative appraisal to the northeast sand soil establishes the appraisal index system, has provided the basis for the reasonable scientific cultivation measure[8]. based on the analysis of variance, principal component analysis and other relevant statistical analysis, the impact on land use mode and soil quality is used as a measure of the minimum data set, and the proposed method is suitable for many regions with high stability, reliability and repeat-ability. Therefore, this paper studies the selection of the evaluation index and the establishment of the minimum data set for the soil fertility of the tillage layer on the northeast slope, and establishes the corresponding minimum data set by the global Moran index analysis and the principal component analysis [9].To further verify the accuracy of the data set, to explore the feasibility of the method from a reasonable and scientific point of view, to provide a reliable theoretical basis for the northeast slope cultivated land to regulate the soil structure, improve soil fertility, concentrate on the intensive way out.

2. Materials and Methods

2.1 Servey area

Inner Mongolia, as the main study area, is an important ecological barrier in the north and one of the 13 major grain-producing provinces (districts) in the country. The land area is about 1.18 million square kilometers. The annual precipitation decreases from east to west, not more than 400 mm, the annual average temperature is 6.7˚, cultivated land area is about ninety thousands km^2, per capita cultivated land is about 5.5%, resident population 25.35 million people, due to the influence of plateau type geomorphological area, mainly temperate continental monsoon climate, from the Da xing a ling mountain and the northern foot of Yin Shan Bei Lu mountain as the climatic boundary, the difference is large, mainly planting corn, soybean, potato and so on, the main types are dark chestnut soil, typical meadow soil, typical black soil, and sandy soil, as shown in Table 1.

Table 1 Distribution of sampling points of tillage layer on northeast slope

Union City	Longitude and latitude	Slope/(°)	Altitude/m	Soil type	Geomorphic type	Topographic site	Sampling points
Huhhot City	111.431°-111.799°E, 39.633°-40.098°N	1-2	1072-1413	Chestnut Brown	Hills	Slope	5
Hulunbeir City	122.125°-125.186°E, 47.381°-50.028°N	0-13	214.2-534.3	Black soil, meadow soil	Hilly, plain	Midland, midslope, Downside	62
Xingan League	120.155°-122.819°E, 45.295°-47.853°N	0-3	226-780	Black calcium soil, meadow soil	Plain	Downslope, midslope	25
Tongliao City	120.448°-123.394°E,	0-2	113.6-542.8	Sand soil, chestnut soil	Hilly, plain	Flat land, valley	35

	42.452°-44.716°N						
Bayannur City	107.270°-109.549°E, 40.880°-41.450°N	0-1	1014.3-1613.6	Irrigation silt, chestnut soil	plains, mountains	Flat land	10
Ordos City	107.916°-110.634°E, 38.151°-40.789°N	0-15	900-1367	Tidal soil	Plain	Flat land	10
Alashan League	101.880°-106.484°E, 37.554°-40.492°N	0-1	1028-1594	Ash desert, brown calcium	Plain, basin	Flat land	5
Chifeng City	117.219°-119.528°E, 42.714°-43.769°N	1-30	462.8-1685	Chestnut calcareous soil	Hilly, plateau	Slope, mid-slope	10
Xilingol League	115.023°-116.779°E, 41.611°-42.376°N	0-2	1246-1500	Chestnut calcium	Hilly, plain	Flat land	15
Wulanchabu City	112.264°-113.444°E, 40.396°-41.073°N	2-12	1235-1582	Chestnut cinnamon	Hilly, plain	Flat land	13
Baotou City	110.498°-110.937°E, 40.378°-40.597°N	1-2	990-998	Meadow soil	Plain	Flat land	5

For many years, under the pressure of the supply of agricultural products, the cultivated land has been too much consumed, soil and water loss, the abuse of chemical fertilizer, and the soil environment has been overwhelmed[10], which seriously restricted the sustainable development of agriculture, until 2018, the total planting area of crops in the whole region was about 8.83 million hectares, down 2% from the previous year. The sampling points are mainly distributed in the Tumuchuan plain, the Western Liaohe Plain Irrigation area in the underground funnel area, the eastern and western foothills of the Daxingan Mountains, and the hilly area in the ecologically serious degraded area. Because of natural climate, topography, and plant factors, groundwater exploitation is too large in the underground funnel area, and the ecologically fragile area is the heterogeneity of the environment, the dry and barren soil, the degradation of cultivated land quality and salinization.

2.2 Data sources and processing methods

The data come from the Inner Mongolia Autonomous Region Land Regulation Center, fixed-point investigation, strict control, each mixed soil sample sampling points not less than 15, pilot tillage layer thickness of 15-30 cm, sampling depth of 0-30 cm, after mixing the soil samples, remove the soil surface cover, using four-point method to select one kg bagging monitoring. Determination of slope, precipitation, soil nutrient status, such as total nitrogen content (semi-micro-open method), organic matter content (potassium dichromate plus heat), available phosphorus content (by anti-colorimetric method), available potassium content (ammonium acetate extraction-flame photometry), pH value (level method) and other trace elements, effective accumulated temperature, effective soil layer thickness, plowing layer thickness, annual precipitation, plowing layer bulk weight, etc[11]. The longitude and latitude coordinates of 195 pilot points are introduced into the Arcmap, which is transformed into a plane coordinate system (in meters), superimposed on the national, municipal, provincial and county boundary lines, and formed the pilot distribution map of the autonomous region, as shown in Fig. 1.the data results were obtained from SPSS19.0、MATLAB and Arcgis 10.4 analysis.

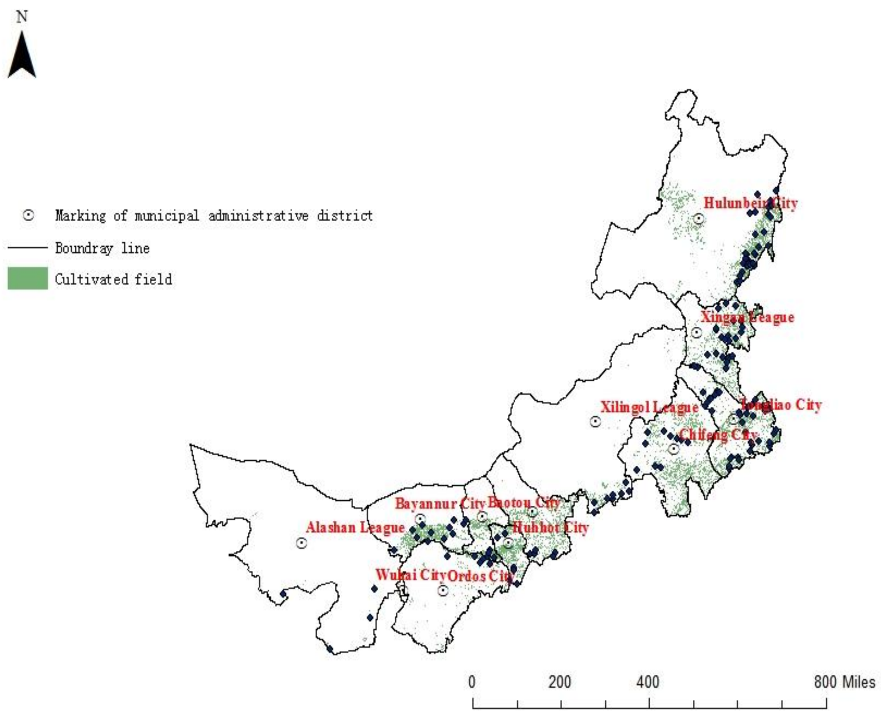


Fig. 1 Distribution of Sampling Points in Inner Mongolia Autonomous Region of China

2.3 Establishment of minimum data sets

2.3.1 spatial auto-correlation analysis

An important measure of spatial correlation of cultivated land in a region is the global Moran index (Global Moran's I). Based on Arcgis spatial statistics, the first law of geography has also mentioned that the index of -1.0 to 1.0 is obtained by normalization of variance, to judge the difference of nutrient content of cultivated land under the influence of soil environment, structural factors and site conditions[12].At present, the correlation interpretation of the spatial difference of cultivated land quality by using Moran index in China, and the global Moran index can effectively measure the spatial correlation, so this study chooses this angle to analyze. Moran index has positive correlation and negative correlation, Moran's I>0 is a positive correlation, the larger the score is, the stronger the correlation, the numerical value of the data set tends to produce clustering in space, otherwise, Moran's I <0 is a negative correlation, the smaller the score, the greater the difference, and the greater the degree of dispersion.

The I index formula for spatially auto-correlated Global Moran's I as follows:

$$I = \frac{n}{R} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} s_i s_j}{\sum_{i=1}^{n} s_i^2} \quad (1)$$

Among them S_i is the deviation of the observed value of the index i from its average value $x_i - \bar{x}$, and w_{ij} refers to the spatial weight between i and j , n is the total number of all indicators, R is the sum of all spatial weights.

$$R_i^2 = \frac{\sum_{j=1, j\neq i}^{n} (x_j - \bar{x})^2}{n-1} \tag{2}$$

2.3.2 Principal Component Analysis

The principal component analysis is the key to establish the minimum data set. based on the SPSS19.0 to derive a small number of principal components from the dimension reduction of the original variables, to reveal the internal structure of each index, and to extract relatively independent indicators to explain. The principal component of eigenvalue ≥1 is obtained, and the normal mode Norm value of each evaluation index is calculated for the sub-component below the principal component. The geometric meaning of this value is the spatial scale of the vector normal mode in the mufti-dimensional space of the component. The larger the Norm value, the larger the load quantity of the index in the principal component, and the larger the explanatory quantity naturally[8,13]. The formula for calculating the Norm values is as follows:

$$N_{ik} = \sqrt{\sum_{i=1}^{k} (u_{ik}^2 \cdot \lambda_k)} \tag{3}$$

Where N_{ik} is the load of the i index on the top k principal components with eigenvalues greater than 1, u_{ik} is the load of the i index on the k principal component; λ_k is the k principal component The characteristic value of the composition. Obtain the norm value of each index. The rule to enter the minimum data set is that the norm value is large, the index with high correlation enters MDS, and the correlation coefficient> 0.3 is highly correlated.

2.3.3 Comprehensive Soil Fertility Index

For qualitative and quantitative evaluation of land, the soil quality of the study area depends on the soil fertility comprehensive index IFI The greater the soil fertility index, the better the soil productivity, the more reasonable the structure of the soil layer, the more reliable theoretical basis for the rational use of soil and the integration of intensive land[6,14].as shown in Table 2. The index of positive action for soil tilling layer is tilling layer thickness, effective soil layer thickness, soil organic matter, total nitrogen, available phosphorus, available potassium and available phosphorus. Therefore, the index of negative action for soil tilling layer is the slope, using ring-down membership function; soil bulk density, pH value and ≥10° effective accumulated temperature and soil tilling layer have suitable critical value range, so the peak-type membership function is chosen from upper limit value and inflection point value[15].

Table 2 Membership Function of Soil Fertility Evaluation Index in Northeast China

Membership function type	Calculation formula	Parametric expression	indicators
Preferential membership function	$y(x)=\begin{cases}1, x\ge b\\ \frac{x-a}{b-a}, a<x<b\\ 0, x\le a\end{cases}$	$y(x)$：Membership function x：Index measured value a：Upper threshold b: Lower threshold	Tilling layer thickness, effective soil layer thickness, soil organic matter, total nitrogen, available potassium, available phosphorus
Unrested membership function	$y(x)=\begin{cases}1, x\le a\\ \frac{x-b}{a-b}, a<x<b\\ 0, x\ge b\end{cases}$	$y(x)$：Membership function x：Index measured value a：Upper threshold	Slope

		b: Lower threshold	
Peak membership function	$y(x)=\begin{cases}1, b_2 \ge x \ge b_1 \\ \frac{x-a_1}{b_1-a_1}, a_1 < x < b_1 \\ \frac{x-a_2}{b_2-a_2}, a_2 > x > b_2 \\ 0, x \le a_1 或 x \ge a_2\end{cases}$	*y(x)*： Membership function *x*： Index measured value a_1： Upper threshold a_2: Lower threshold b_1： Upper limit of optimum value b_2： Lower limit of optimum value	Soil bulk density, pH、≥10° effective accumulated temperature

the fertility synthesis index of each data set is calculated based on the weight and membership degree obtained from the principal component analysis. The formula is as follows:

$$IFI=\sum_{i=1}^{n} F_i \cdot C_i \tag{4}$$

In the formula F_i is the weight of the i indicator; C_i is the degree of membership of the i indicator, and n is the number of indicators.

2.3.4 Tilling layer

The accuracy of global Moran `s I index and PCA principal component analysis is verified by Nash effective coefficient and relative deviation coefficient[16]. The formula is:

$$E_f = 1-\frac{\sum(R_t-R_{MDS})^2}{\sum(R_t-\bar{R}_t)^2} \tag{5}$$

$$E_R=\frac{\left|\sum_{i=1}^{n} R_{ti}-\sum_{i=1}^{n} R_{MDS}\right|}{\sum_{i=1}^{n} R_{ti}} \tag{6}$$

R_t is the fertility synthesis index of the global data set TDS (Total data set), $\bar{R}_t$ the average fertility synthesis index of the TDS, and R_{MDS} is the fertility synthesis index of MDS . The closer the calculated Nash's effective coefficient E_f is to 1, the more accurate the fertility synthesis index based on MDS and the closer the score， the relative deviation coefficient E_R is to 0, it is proved that the smaller the deviation of fertility synthesis index based on MDS, the higher the accuracy.

3. Results and analysis

3.1 Soil Fertility in Tilling Layer on Northeast Slope

The study of soil fertility in the north-east slope requires various indexes, such as physical and chemical properties of the soil, and descriptive statistics are carried out on the indexes of the SPSS19.0, such as Table 3, the average thickness of the effective soil layer (54.81 ± 1.71) cm, the thickness of the tillage layer (25.27 ± 0.45) cm, and the sampling thickness of the tillage layer is basically between 20 cm and 40 cm[7,17], effective accumulated temperature above 10° (2513.20 ± 34.7), the decrease of rainfall and unreasonable farming methods will also increase the evaporation of soil moisture and reduce the effective accumulated temperature needed for crops; soil bulk density (1.37 ± 0.01) g·cm^{-3}. To consider soil moisture movement and nutrient transfer it is necessary to add soil bulk density, which is a measure of soil structure and compactness, so the soil is tight because there is no long-term tillage deep loosening; pH (7.10±0.08), soil pH is neutral, organic matter (36.04 ± 1.44) g·kg^{-1}, total nitrogen (3.25 ± 0.22) g·kg^{-1}, available phosphorus (54.90 ± 4.00) mg·kg^{-1}, available potassium (132.09 ± 6.35)$^{-1}$, soil potassium content is low mainly due to unreasonable fertilization and artificial tillage [18].

Table 3 Statistical Characterization of Soil Parameters for Quality Evaluation of Cultivated Layer on Northeast Slope

Projects	Minimum value	Maximum value	Average	Standard deviation	Skewness	Kurtosis	Coefficient of variation(%)	K-S test
Effective soil thickness （ESL）	18	70	54.8103	1.71105	0.699	0.199	43.593%	0.153
Thickness of tillage layer （CH）	14	45	25.2718	0.45409	0.828	-0.109	25.092%	0.305
effective accumulated temperature over 10° （AT）	1417	3400	2513.1949	34.66134	0.008	-0.826	19.259%	0.121
Bulk （VW）	1	1.8	1.3708	0.01112	0.394	-0.247	11.326%	0.097
pH	4.64	8.82	7.1071	0.08692	-0.286	-1.498	17.078%	0.17
organic matter （OM）	11.4	145	36.0443	1.44176	1.287	3.391	55.857%	0.104
Total nitrogen（TN）	0.26	26.3	3.2471	0.22357	4.718	26.184	96.149%	0.404
Available phosphorus （AP）	2.2	255.8	54.9036	3.99644	2.121	4.268	101.646%	0.243
Available potassium （AK）	30	571	132.0923	6.34989	1.661	4.476	67.128%	0.112
Slope （S）	0	30	5.6821	0.34012	2.915	10.459	83.588%	0.286

The coefficient of variation (CV) is the criterion for the variability of the physical and chemical properties of the soil in the region. all data must obey the normal distribution. test- S concludes that the data are all normal distribution ($p>0.05$). To sum up, the coefficient of variation of total nitrogen and available phosphorus is the highest (≥90%), 96.149% and 101.646%, which may be affected by fertilization and tillage patterns in some areas. The available potassium, organic matter, slope and effective soil layer thickness are the moderate variation (≥40%,≤90%), pH、soil bulk density, effective accumulated temperature and tillage layer thickness (40%) are weak variation without significant variation.

3.2 Establishment of a minimum data set based MDS Moran′s I index

Based on the global Moran index analysis, the inverse distance is used to explain the correlation of the elements by decreasing the relationship of distance, and the Euclidean shortest distance is adopted, that is, the shortest distance between two points based on space. The Moran index analyzed is effective soil layer thickness is 0.1011, the tillage layer thickness is 0.2089, effective accumulated temperature above 10° is 0.4897, soil bulk density is 0.1196, pH is 0.7382、organic matter is 0.4764, total nitrogen is 0.59245, available phosphorus is 0.0931, available potassium is 0.2329,the slope is 0.0650, so as to express the characteristics of each region about the spatial distribution of each index.

The p value represents the probability of the random distribution of the data in space. the measured data p-values are all less than 1%. the spatial pattern is unlikely to be randomly generated, showing a significant state, so the null hypothesis is rejected. Z score was multiple of standard deviation, and the aggregation state was divided from large to small into pH> effective accumulated temperature above 10° > organic matter>available potassium>tillage layer thickness>soil bulk density>effective soil layer thickness>available phosphorus>slope and total nitrogen. pH Moran`s index has the strongest maximum correlation and the Z score is the largest, so it is been selected of the minimum data set, followed by the effective accumulated temperature above 10°, organic matter, available potassium, layer thickness, total nitrogen Moran index, although those of them have large correlation, the Z score is the smallest, scattered state, cannot meet the requirements of the selected minimum data set, the others Moran`s index is less than 0.2, not considered. Therefore, more than pH≥of effective accumulated temperature10°, organic matter, available potassium and thickness of tilling layer were selected to the MDS.

3.3 Minimum Data Set Based on Principal Component Analysis MDS Establishment

The principal component analysis based on SPSS19.0 shows that there are four principal components of the resulting eigenvalue >1, and the cumulative contribution rate is 76.253%, which meets the conditions of principal component analysis, and uses the four principal components to explain each component of its index(Table 4). Organic matter, pH, effective accumulated temperature above 10°, the thickness of tilling layer and available potassium were divided into the first group, pH was highly correlated with an effective accumulated temperature above 10° of ≥ (0.477**) and organic matter (-0.574**) [19].The Norm value of organic matter is up to 1.295, as shown in Table 5,which automatically enters the MDS layer thickness with effective accumulated temperature (-0.363*) and the slope (0.351*) above 10°, so the thickness of tilling layer is selected into the MDS. The effective soil layer thickness and slope are the second groups, the effective soil layer thickness is compared with the Norm value of the slope, and the effective soil layer thickness can be entered MDS, However, the effective soil layer thickness is less frequently used than the tillage layer thickness, and the two explanations are similar, so the elimination. Effective phosphorus is the third group and automatically enters the MDS. For the fourth group of total nitrogen and soil bulk density, soil bulk density can explain the characteristics of soil compactness and structural pores, and can also reflect the degree of soil erosion, so the soil bulk density is selected to enter the MDS. Therefore, the five indexes that finally enter the MDS are pH、organic matter、thickness of tillage layer、available phosphorus、soil bulk density.

Table 4 Evaluation index person correlation coefficient matrix

Evaluation index of	ESL	CH	AT	VW	pH	OM	TN	AP	OK	S
ESL	1									
CH	0.324*	1								
AT	-0.05	-0.363	1							
VW	-0.139	-0.13	0.236	1						
pH	0.094	-0.336	0.477**	-0.041	1					
OM	-0.175	0.232	-0.325	-0.063	-0.574**	1				
TN	-0.015	0.136	-0.062	0.198	-0.257	0.137	1			
AP	-0.062	0.19	-0.196	0.223	-0.48	0.234	0.712**	1		
A K	0.022	0.199	-0.248	-0.081	-0.338	0.32	0.105	0.349*	1	
S	0.277	0.351*	-0.225	-0.069	-0.096	0.048	-0.087	0.011	0.046	1

Table 5 Load matrix and Norm value of soil fertility assessment index of tillage layer on northeast slope

Evaluation indicators	1	2	3	4	Norm values
OM	0.735	-0.091	-0.411	0.114	1.294
pH	-0.727	0.362	-0.002	0.113	1.285
AT	-0.715	-0.172	0.235	0.031	1.219
CH	0.618	0.404	0.289	0.030	1.186
AK	0.581	-0.135	-0.043	0.185	0.985
S	0.331	0.541	0.324	0.141	0.973
AP	0.450	-0.52	0.457	0.134	1.132
TN	0.110	-0.467	0.399	-0.705	1.050
VW	-0.197	-0.438	0.433	0.627	1.029

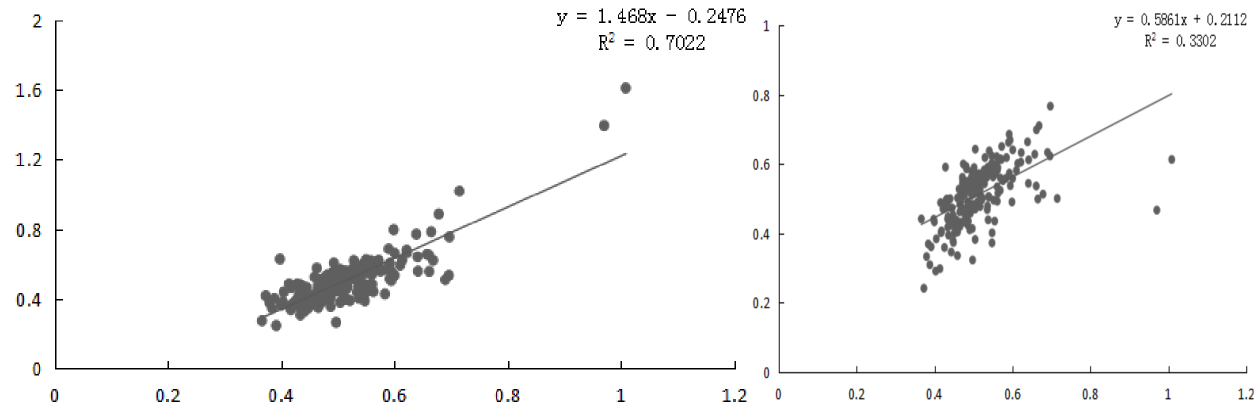
3.4 Fertility composite index MDS the minimum data set

The most intuitive reflection of soil production capacity and quality is the comprehensive index of soil fertility. The weighted square sum of the indicators is used to obtain the common factor variance of each index, that is, the contribution value of each index to the comprehensive fertility of the soil, and then the common factor variance is used to obtain their respective weights, as shown in Table 6. Using the membership function of the fuzzy mathematics method, the membership degree of each sample point is obtained, and the soil fertility comprehensive index IFI value of each minimum data set is calculated by combining membership degree and weight, and the IFI value of each group is compared. The variation range of soil fertility comprehensive index IFI-T based on full volume data set is 0.366-0.770, the average value is 0.513 ± 0.005, and the coefficient of variation is 15.358%, which belongs to weak variation. The variation range of soil fertility comprehensive index IFI-MR based on the minimum data set of Moran`I index analysis is 0.246~0.887, the average value is 0.495±0.007, and the coefficient of variation is 20.420%, which belongs to moderate variation. The variation range of soil fertility comprehensive index IFI-PCA based on the minimum data set of principal component analysis is 0.242~0.767, the average value is 0.512±0.006, and the coefficient of variation is 16.743%, which is a weak variation. IFI-PCA、IFI-MR and IFI-T are closer to full-volume data sets in terms of range, mean and coefficient of variation. The linear regression analysis of the IFI-T with IFI-MR and IFI-PCA, as shown in Fig. 2, shows a positive correlation, R^2 is 0.330 and 0.702, respectively, which proves that the linear fitting of IFI-PCA and IFI-T is the best. The Nash effective coefficient of IFI-MR is 0.857, the relative deviation coefficient is 0.014, the Nash effective coefficient of IFI-PCA is 0.900, and the relative deviation coefficient is 0.003, so it is more accurate to judge the soil fertility interpretation of northeast slope tillage layer IFI-PCA, so the soil fertility comprehensive index selected to replace the whole data set to evaluate the soil quality of northeast slope tillage layer.

Table 6 Total and minimum data sets based on soil fertility composite index

Indicators	Total data sets TDS		Principal Component Minimum Data Set MDS-PCA		Moran Index Minimum Data Set MDS-MR	
	Communality	Weight	Communality	Weight	Communality	Weight
ESL	0.680	0.103				
CH	0.629	0.095	0.459	0.142	0.545	0.158
AT	0.598	0.090			0.626	0.180
VW	0.811	0.122	0.793	0.244		
pH	0.673	0.101	0.767	0.236	0.713	0.205

OM	0.730	0.110	0.582	0.179	0.639	0.184
TN	0.887	0.134				
AP	0.699	0.106	0.646	0.199		
AK	0.391	0.059			0.949	0.273
S	0.527	0.08				



a.MDS-PCA b.MDS-MR

Fig. 2 Correlation between MDS-PCA based on TDS and MDS comprehensive index of soil fertility

3.5 IFI Grading

Based on the comprehensive index of IFI-PCA soil fertility calculated by the MDS index, the soil fertility is graded and graded. Through the SPSS19.0 Wad method, the interval uses square-European distance to carry on the cluster analysis, and each sample is graded according to the pedigree map of cluster analysis. As shown in Table 7, the IFI-PCA values are divided into six grades: optimal I, excellent II, good III, moderate IV, poor V, extreme poor VI, optimal I (0.710-0.767) sample number is 4, proportion 2.05%, excellent II (0.589-0.650) sample number is 39, proportion 20%, good III (0.551-0.573) sample number .54%, medium IV (0.460-0.500) sample number is 66, accounting for 33.5%, V (0.410-0.432) sample number is 28, accounting for 14.36%, and VI (0.242-0.359) sample number is 16, accounting for 8.21% (Fig. 3).

Table 7 Frequency distribution (%) of soil fertility grades based on MDS-PCA in the Northeast Slope China

City	I	II	III	IV	V	VI
Tongliao City	1	3.6	4.1	8.2	1.5	2
Hulunbeir City	0.5	7.7	6.7	10.8	4.1	3.1
Wulanchabu City	0.5	0	2	3.1	0	0.5
Xilingol League	0	3.1	0.5	0.5	1	0
Xingan League	0	2.5	3.1	2.6	3.1	1.5
Chifeng City	0	1.5	1	2	0.5	0
Baotou City	0	1	0.5	0.5	0	0
Ordos City	0	0	1.5	3.1	1	0
Bayannur City	0	0	1	3.1	1	0
Hohhot City	0	0	0	1	0.5	0
Alashan League	0	0	0	0.5	0	0

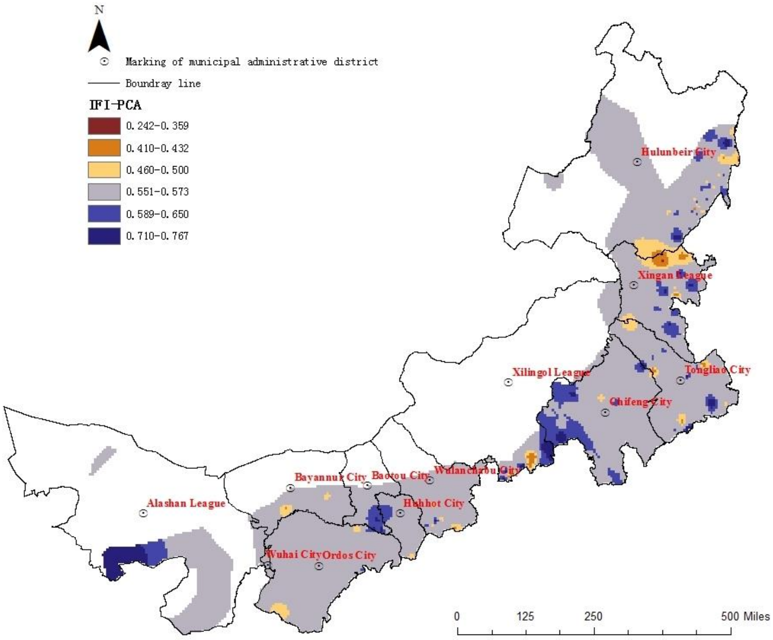


Fig. 3 Distribution of IFI-PCA index in Inner Mongolia Autonomous Region

4. Conclusions

In this study, the Inner Mongolia Autonomous Region was used as the sampling area for the northeast slope farmland. This area has a wide spatial scale and is greatly affected by man-made effects. Soil nutrients, meteorological factors, and environmental factors are all dynamic changes. The restrictive factors in different regions are also different. Soil quality is also relatively different. A lot of work has been carried out at home and abroad to study the establishment of a minimum data set. This article uses the analysis method and evaluation method as a new entry point, and couples the global Moran index to the comprehensive index of soil fertility of the full data set and the minimum data set. And principal component analysis to select the smallest data set and select the best method for analyzing the smallest data set. The full data set includes the thickness of the plow layer, the effective plow layer thickness, the soil bulk density, the effective accumulated temperature above ≥10°, pH, slope, organic matter, and quick effect. Among the 10 indicators including potassium, available phosphorus and total nitrogen, the minimum data and soil bulk density, pH, organic matter, available phosphorus, and arable layer thickness were selected. The frequency of use is also among the top ten in soil fertility evaluation at home and abroad.

(1) The difference of soil fertility evaluation index in northeast slope tillage layer is still obvious, the thickness of tillage layer (25.27 ± 0.45) cm, the thicker the tillage layer, the higher the land productivity, the thinner the tillage layer thickness, and the cm of soil bulk density $(1.37 \pm 0.01)^{-3}$.The soil is relatively compact and requires deep loosening and tillage. pH(7.10 ± 0.08) The soil pH is neutral and organic matter (36.04 ± 1.44) is $g\cdot kg^{-1}$, total nitrogen (3.25 ± 0.22) $g\cdot kg^{-1}$, available phosphorus (54.90 ± 4.00) $mg\cdot kg^{-1}$, available potassium $(132.09 \pm 6.35\ mg\cdot kg)^{-1}$ soil nutrient content is abundant.

(2) The soil fertility evaluation indicators of the arable layer on the northeast slope include the arable layer thickness, pH, soil bulk density, organic matter and available phosphorus. The soil fertility comprehensive index of the full data set, the minimum data set under the Moran index and the minimum data set under principal component analysis, the change range is IFI-MR (0.246-0.887)> IFI-T (0.366-0.770) > IFI -PCA (0.242-0.767), average IFI-T (0.513 ± 0.005)> IFI-PCA (0.512 ± 0.006) > IFI-MR (0.495 ± 0.007), coefficient of variation is IFI-MR (20.430%) > IFI- PCA (16.743%)> IFI-T (15.358%), the Nash effective coefficients of IFI-T and IFI-MR、IFI-PCA are 0.857 and 0.900, the relative deviation coefficients are 0.014 and 0.003, and the linear regression R^2 are 0.330 and

0.702. It can be seen that the minimum data set based on principal component analysis of MDS-PCA can replace the full data set for soil fertility analysis.

(3) According to the comprehensive index of soil fertility in the study area, the soil fertility was graded ,2.05% of the 195 samples belonged to the comprehensive index of soil fertility 0.710~0.767, and the land quality grade was the optimal I, mainly distributed in Tongliao City, Hulunbeir City, Ulanchab City ,20% sample soil fertility comprehensive index 0.589~0.650, land quality grade II, Hulunbeir City accounted for the best 7.7%, followed by Tongliao City and Xilingol League ,21.54% soil fertility comprehensive index 0.551~0.573,soil quality grade good III, the same Hulunbeir City accounted for the highest 6.7%,33.5% sample soil fertility comprehensive index 0.460-0.500 Soil quality grade medium IV, Hulunbeir City, Tongliao City, Ulanchab City, Xingan League, Chifeng City, Ordos City, Bayannur City have different proportion ,14.36% sample site soil fertility comprehensive index is 0.410-0.432, land quality grade is poor V, mainly distributed in Hulunbeir City and Xingan League ,8.21% sample site soil fertility comprehensive index 0.242-0.359, land quality grade is extremely poor VI, Ulanchabu City and Xingan League. Baotou belongs to the middle and upper sensitive level, Hohhot belongs to the middle and lower sensitive level, Alashan League belongs to the middle sensitive level.

(4) According to the different soil nutrient status in the study area, the corresponding soil fertility improvement measures were adopted, and conservation tillage was carried out, which could be rotated, fallow, deep tillage, straw returning to the field, and other high tillage, to adapt to local conditions and achieve the goal of water, fertilizer and production conservation, and different farming methods have different improvement effects [20,21].The minimum organic matter 11.4g·kg $^{-1}$.We can take straw return measures to improve soil structure, increase organic matter, total nitrogen minimum of 0.26 g·kg^{-1} Area, can add nitrogen fertilizer to adjust the carbon-nitrogen ratio, straw return and fallow can raise soil porosity to vary degrees, water stability agglomeration, reconcile pH value, let land get recuperation, also can increase crop yield, nitrogen content, activate soil phosphorus, the salinized area can add lime rice to adjust pH, high slope area can adopt horizontal slope tillage along the contour line, control runoff, ensure soil moisture content to reduce soil erosion.

Acknowledgments

Natural Resources Science and Technology Innovation Project

References

[1] Lai R, Long Y, Cheng C, et al. Optimization Construction and Technology Integration of Interplanting Patterns in Kiwifruit Orchard. Chinese Journal of Ecological Agriculture 2019; 27(09): 1430-1439.

[2] Larson W E, Pierce F J. The dynamics of soil quality as a measure of sustainable management. Defining soil quality for a sustainable environment 1994; 35: 37-51.Larson W E,Pierce F J.The dynamics of soil quality as a measure of sustainable management.In： Doran J W,et al.ed.Defining Soil Quality for a Sustainable Environment.Soil Science Society of American Publication No.35.Inc, Madison,Wisconsin,USA,1994.37~52

[3] Andrews S S, Karlen D L, Mitchell J P. A comparison of soil quality indexing methods for vegetable production systems in Northern California. Agriculture, ecosystems & environment 2002; 90(1): 25-45. S.S.Andrews,D.L.Karlen, J.P.Mitchell, A comparison of soil quality indexing methods for vegetable production systems in Northern California[J].. Agriculture,Ecosystems and Environment,2002,90(1).

[4] Sparling G P, Schipper L A, Bettjeman W, et al. Soil quality monitoring in New Zealand: practical lessons from a 6-year trial. Agriculture, Ecosystems & Environment 2004; 104(3): 523-534. G.P.Sparling,L.A.Schipper, W.Bettjeman,R.Hill.Soil quality monitoring in New Zealand：practical lessons from a 6-year trial[J]. Agriculture,Ecosystems and Environment,2004,104(3).

[5] Govaerts B, Sayre K D, Deckers J. A minimum data set for soil quality assessment of wheat and maize cropping in the highlands of Mexico. Soil and Tillage Research 2006; 87(2): 163-174.Bram Govaerts,Ken D.Sayre,Jozef Deckers.A minimum data set for soil quality assessment of wheat and maize cropping in the highlands of Mexico[J]. Soil &Tillage Research,2005,87(2).

[6] Jin H, Shi D, Chen Z, et al. Soil Quality Evaluation Index of Cultivated Land Based on Cluster and PCA Analysis on Red Soil Slope. Journal of Agricultural Engineering 2018; 34(07): 155-164.

[7] Miao S, Zhao H, Qiao Y, et al. Evaluation of the effect of tillage measures on soil quality of sandy soil in Northeast China based on tillage index. Chinese Soil and Fertilizer 2019; 34(04): 9-15.

[8] Li G, Chen J, Tan M, et al. Establishment of minimum data set for soil quality assessment based on land use change. Journal of Soil 2008; 45(01): 16-25.

[9] Liu H, Wang H, Xie Y, et al. Temporal and Spatial Evolution Characteristics and Influencing Factors of Cultivated Land Eco-efficiency in Concentrated Special Hardship Areas--Taking Lv Liang Mountain Area as an Example. Soil and Water Conservation Research 2020; 27(02): 323-329.

[10] Fan Y, Yan X, Zhang H, et al. Evaluation of crop suitability and allocation of water-saving irrigation methods in cultivated sandy land in Inner Mongolia. Journal of Agricultural Engineering 2017; 33(21): 115-127.

[11] Baustan. Soil Agrochemical Analysis (3rd Edition). Beijing: China Agricultural Press, 2007.

[12] Huang J, Liu Y. Measurement of Agricultural Ecological Efficiency in Three Gorges Reservoir Area and Analysis of its Influencing Factors. Statistics and Decision Making 2018; 34(7): 123-127.

[13] Li P, Zhang X, Hao M, et al. Evaluation of Soil Quality in Reclamation in Loess Plateau Mining Area Based on Minimum Data Set. Journal of Agricultural Engineering 2019; 35(16): 265-273.

[14] Ma Q, Zhang G, Geng Y, et al. Analysis on Soil Fertility in Slope Cultivated Areas in China. Journal of Soil and Water Conservation 2016; 30(05): 190-204.

[15] Li S, Liu H, Bai W, et al. Analysis on the Evolution of Cultivated Land and Its Relationship with Planting Structure, Soil Bulletin 2016; 47(06): 1300-1305.

[16] Nash J E, Sutcliffe J V. River flow forecasting through conceptual models part I—A discussion of principles. Journal of hydrology 1970; 10(3): 282-290.

[17] Liu Z, Shi D, Jin H, et al. A Study on Soil Quality Characteristics and Barrier Factors of Cultivated Land on Red Soil Slope. China Soil and Fertilizer 2018; 03: 7-41.

[18] Han T, Wang B, Zhang H, et al. Effects of Long-term Fertilization and Lime Effect on Potassium in Rhizosphere of Corn at Different Growth Period. Journal of Soil 2017; 54(06): 1497-1507.

[19] Deng S, Zeng L, Guan Q, et al. Evaluation of soil quality in cold-impregnated fields in southern China based on minimum data set. Journal of Soil 2016; 53(05): 1326-1333.

[20] Li C, Tang Y, Wu X, et al. Effects of Different No-tillage Rotation Patterns on Soil Properties in Paddy Fields. Journal of Agricultural University of China 2019; 24(05): 20-29.

[21] Yao D, Pei J, Wang J. A Study on the Temporal and Spatial Changes of Cultivated Land Quality in Typical Black Soil. Northeast China Chinese Journal of Ecological Agriculture (Chinese and English) 2020; 28(01): 104-114.