Farmland Soil Quality Assessment and Constraint Diagnosis Based on Minimum Dataset in Shawan City, Xinjiang (Postprint)

Authors: Gao Haifeng, Wang Xiyuan, Wu Haoying, Lei Haifeng

Date: 2024-09-25T00:00:00+00:00

Abstract

To evaluate the soil quality status of farmland in Shawan City, Xinjiang, identify key factors influencing productivity levels, and advance high-standard farmland construction in Shawan City, Principal Component Analysis (PCA) and Cluster Analysis (CA) were respectively employed to establish minimum datasets for farmland soil quality evaluation indicators, which were then combined with an obstacle factor diagnosis model to reveal the characteristics of farmland soil quality and its obstacle factors in Shawan City. The results indicated: (1) Soil quality evaluation results based on different datasets showed significant differences; the soil quality index based on cluster analysis exhibited a significant positive correlation with the full dataset (R^2=0.591, P<0.1), with a Nash-Sutcliffe Efficiency coefficient of -4.923, both values being greater than those of the minimum dataset based on principal component analysis, demonstrating that the minimum dataset based on cluster analysis is more suitable than that based on principal component analysis for replacing the full dataset in evaluating farmland soil quality. (2) The overall soil quality in the study area was at a medium or higher level, with soil quality indices ranging from 0.130 to 0.641; the soil quality index based on the minimum dataset classified farmland in Shawan City into five grades, with Grade I soils primarily distributed in the northern and northwestern parts of the study area, and Grade V soils mainly distributed in the southeastern part, presenting an overall spatial distribution pattern of high in the northwest and low in the southeast. (3) Low soil organic matter content, nitrogen deficiency, and high electrical conductivity were the primary obstacle factors affecting soil quality in the study area. The research findings can be applied to effectively improve local farmland soil quality; it is recommended that during farmland management, in addition to organic fertilizer application, measures such as deep tillage, weed cultivation, straw mulching, and plastic film mulching be adopted to amend the soil.

Full Text

Evaluation and Obstacle Diagnosis of Farmland Soil Quality in Shawan City, Xinjiang Based on Minimum Dataset

GAO Haifeng, WANG Xiyuan, WU Haoying, LEI Haifeng
(College of Ecology and Environment, Xinjiang University, Urumqi 830046, Xinjiang, China)

Abstract

To evaluate the soil quality status of farmland in Shawan City, Xinjiang, identify the key influencing factors affecting productivity levels, and promote the construction of high-standard farmland in Shawan City, this study established minimum datasets of farmland soil quality evaluation indicators using principal component analysis (PCA) and cluster analysis (CA). Combined with an obstacle factor diagnostic model, the characteristics and obstacle factors of farmland soil quality in the Shawan area were revealed. The results showed significant differences between soil quality evaluation results based on different datasets. The soil quality index based on cluster analysis showed a significant positive correlation with that based on the total dataset ($R^2 = 0.591$, $P < 0.1$), with an effective coefficient of 4.923, indicating that the minimum dataset based on cluster analysis was more suitable than that based on principal component analysis for replacing the total dataset in evaluating farmland soil quality. Based on the minimum dataset, farmland in Shawan City was divided into five classes. The overall soil quality in the study area was at a moderate or better level, with soil quality indices ranging from 0.130 to 0.641. Class I soil was mainly distributed in the northern and northwestern parts of the study area, while Class V soil was mainly distributed in the southeastern part, showing an overall spatial distribution pattern of high quality in the northwest and low quality in the southeast. Low soil organic matter content, nitrogen deficiency, and high electrical conductivity were the main obstacle factors affecting soil quality in the study area. These results can be used to effectively improve local farmland soil quality. In addition to applying organic fertilizer, it is recommended that measures such as deep tillage, weed planting, straw mulching, and plastic film mulching be adopted during farmland management.

Keywords: minimum dataset; principal component analysis; cluster analysis; soil quality evaluation; obstacle diagnosis

1. Introduction

Soil is the most fundamental natural resource for human survival and an important medium for various plants to grow. Farmland soil not only provides a place for crops to take root but also supplies the substances and energy needed for human survival []. Shawan City in Xinjiang is one of the regions for national high-standard farmland construction. As the primary industry in Shawan City, agriculture plays a vital role in local economic development. Although the cultivated land in Shawan City generally has gentle slopes, long-term intensive cultivation and unreasonable fertilizer application have led to increasingly prominent land quality issues, with soil fertility degradation, salinization, and soil compaction gradually intensifying. To comprehensively grasp the quality status of cultivated land in Shawan City, it is urgent to accurately evaluate farmland soil quality and identify the major obstacle factors affecting it. However, due to the complex and diverse properties of soil itself, single indicators are insufficient to reflect comprehensive soil quality, and experimental analysis of numerous soil indicators is difficult. Therefore, the minimum dataset method was developed to solve redundancy problems among data. It primarily uses statistical methods such as principal component analysis, grey correlation, cluster analysis, and correlation analysis to screen redundant indicators and construct a minimum dataset for soil quality evaluation []. For example, Chen et al. [] used the minimum dataset method to evaluate the tillage layer quality of sloping farmland in Yunnan Province and compared differences under different tillage conditions and soil types. Zhou et al. [] used cluster and principal component analyses to evaluate the soil quality of the tillage layer in red soil sloping farmland, verifying the applicability of the minimum dataset in soil quality evaluation. Choudhury et al. [] constructed a minimum dataset for soil quality assessment through principal component analysis to index soil properties. Thus, principal component analysis and cluster analysis have become the most widely used methods in soil quality evaluation.

However, farmland soil characteristics in different regions exhibit differences, complexity, and variability due to varying environments, and appropriate soil quality evaluation methods must be verified and compared []. Although many studies on soil quality evaluation have been conducted in China, research on farmland soil in Xinjiang is limited. By comparing the advantages and disadvantages of principal component analysis and cluster analysis methods, this study aims to select a suitable evaluation method for farmland in Shawan City, Xinjiang, and use an obstacle diagnosis model to more comprehensively understand local soil quality status, providing theoretical guidance for rational cultivation and fertilization of farmland in this region.

1.1 Study Area Overview

Shawan City is located in northwestern Xinjiang, southeast of the Tacheng region, on the southern edge of the Junggar Basin and at the northern foot of the Tianshan Mountains (84°57 ~86°09 E, 43°29 ~45°20 N). It has a continental mid-temperate arid climate with an average temperature of 6.3~6.9°C, annual precipitation of 140~350 mm, and annual evaporation of 1500~2000 mm. The main soil types are brown calcic soil, gray desert soil, saline soil, and aeolian sandy soil. Land use primarily includes irrigated farmland, dry land, saline-alkali land, sandy land, grassland, and urban residential land. The farmland area is 16.28$×10^{2}$ km^2, mainly irrigated land, with main crops including cotton, wheat, corn, and sugar beet (Fig. [Figure 1: see original paper]).

1.2 Experimental Methods

1.2.1 Sample Setting Sampling was conducted in early May 2022. Sampling points avoided ditches, roads, residential areas, and garbage dumps. Actual sampling locations were selected based on local topography and soil texture to ensure uniform and representative soil samples. Sampling followed the principles of randomness and equal quantity. Surface soil samples (0–15 cm) were collected using the five-point sampling method and mixed. The mixed soil samples were prepared into one soil sample using the quartering method [], with a mass of 1–2 kg. A total of 84 representative soil samples were collected. The soil samples were brought back to the laboratory, air-dried naturally, ground, and sieved for subsequent indicator determination.

1.2.2 Determination of Soil Physicochemical Properties Based on the spatial variability characteristics of soil physicochemical properties in the study area and following the principles of dominance, sensitivity, stability, and independence of soil quality evaluation indicators [], this study selected the following indicators: soil pH, electrical conductivity, organic matter, available phosphorus, available potassium, alkali-hydrolyzable nitrogen, total nitrogen, and total phosphorus. Soil pH and electrical conductivity were measured using a pH meter and portable conductivity meter after mixing soil with water; total nitrogen was determined by the Kjeldahl method []; total phosphorus was measured using a flow analyzer []; soil organic matter was determined by the potassium dichromate volumetric method with external heating []; alkali-hydrolyzable nitrogen was measured by the diffusion absorption method []; available phosphorus was determined by the ammonium fluoride-hydrochloric acid colorimetric method []; and available potassium was measured by flame photometry using NH_4OAC [].

1.3 Data Analysis Methods

1.3.1 Principal Component Analysis SPSS 26.0 software was used for principal component analysis of soil evaluation indicators. Indicators with eigenvalues > 1 and absolute factor loadings > 0.5 in principal components were grouped together. If a soil indicator simultaneously had factor loadings > 0.5 in two principal components, correlation analysis was conducted among indicators, and the soil indicator was assigned to the group with lower correlation with other soil indicators []. If a certain evaluation indicator had low factor loadings across all principal components, it was assigned to the group with the highest loading []. When only one indicator had high factor loading in a principal component, that indicator entered the minimum dataset. Better correlation among soil indicators indicates more similar effects, and only one highly correlated soil indicator could be selected for the minimum dataset while the rest were eliminated. The Norm value was then calculated, representing the vector norm of the indicator in the multidimensional space composed of the selected principal components. A larger Norm value indicates greater comprehensive loading of the indicator across all selected principal components []. Compared with loading on a single principal component, this indicator loses less information, has better representation, and stronger comprehensive explanatory power []. The Norm value is calculated as follows:

$$N_{ik} = \sqrt{\sum_{k=1}^{n}(u_{ik}^2\lambda_k)}$$

where N_{ik} is the comprehensive loading of the ith indicator on the first k principal components with eigenvalues > 1; u_{ik} is the loading of the ith indicator on the kth principal component; and λ_k is the eigenvalue of the kth principal component.

1.3.2 Cluster Analysis SPSS 26.0 software was used for Q-type cluster analysis of evaluation indicators using the Euclidean distance shortest method. At the appropriate clustering level, evaluation indicators were divided into several groups reflecting different aspects of farmland soil quality characteristics []. According to correlation analysis, indicators with significant correlation within each group could substitute for each other. Combined with field investigation, literature review, and previous research results, redundant indicators in farmland soil were eliminated, and representative, independent indicators were selected to enter the minimum dataset [].

1.3.3 Determination of Indicator Weights In soil quality evaluation, the calculation of indicator weights is particularly important. Currently, the common factor method based on principal component analysis is widely used in soil quality evaluation. In this study, indicator weights were obtained from the common factor variance of each soil indicator [].

1.4 Research Methods

1.4.1 Soil Quality Evaluation Method The Soil Quality Index (SQI) is a comprehensive expression of soil evaluation indicators. A larger SQI indicates better soil quality []. Selected indicators were used to calculate the soil quality comprehensive index through the following formula:

$$SQI = \sum_{i=1}^{n} W_i S_i$$

where W_i is the weight coefficient of the ith indicator; S_i is the membership value of the ith indicator; and n is the number of evaluation indicators. Based on different functional attributes of soil indicators, S-type membership functions and parabolic membership functions were selected []. The specific classification results are shown in Table . Parabolic functions have the most suitable range, while S-type functions indicate that larger indicator values represent better soil quality, and vice versa []. The turning points of parabolic membership functions corresponding to indicators were obtained through field measurement data and literature review []. The calculation formulas are as follows:

S-type membership function:

$$f(x) = \begin{cases} 0.1 + 0.9 \times \frac{x-a}{b_1-a} & \text{if } x < b_1 \\ 1.0 & \text{if } b_1 \leq x \leq b_2 \\ 1.0 - 0.9 \times \frac{x-b_2}{b-b_2} & \text{if } x > b_2 \end{cases}$$

Parabolic membership function:

$$f(x) = \begin{cases} 0.1 + 0.9 \times \frac{x-a}{b_1-a} & \text{if } a \leq x < b_1 \\ 1.0 & \text{if } b_1 \leq x \leq b_2 \\ 1.0 - 0.9 \times \frac{x-b_2}{b-b_2} & \text{if } b_2 < x \leq b \end{cases}$$

where x is the measured value of the indicator; a and b are the minimum and maximum values of the indicator, respectively; and b_1 and b_2 are the lower and upper bounds of the suitable value range for the indicator.

1.4.2 Obstacle Factor Diagnosis To improve farmland soil quality and eliminate obstacle factors affecting soil improvement, this study introduced the obstacle factor diagnosis model [], also known as the obstacle degree model. This model uses factor contribution degree, indicator deviation degree, and obstacle degree for obstacle diagnosis. Obstacle degrees were then ranked to determine the primary and secondary relationships of each obstacle factor []. Finally, the indicator obstacle degrees were divided into four categories using the equal interval method: no obstacle (0–0.1), mild obstacle (0.1–0.2), moderate obstacle (0.2–0.3), and severe obstacle (>0.3) []. The calculation formulas are as follows:

$$M_{ij} = \frac{P_{ij}W_j}{\sum_{j=1}^{n}(P_{ij}W_j)} \times 100\%$$

$$P_{ij} = 1 - S_{ij}$$

$$M_j = \frac{\sum_{i=1}^{m} P_{ij}W_j}{\sum_{j=1}^{n}\sum_{i=1}^{m} P_{ij}W_j} \times 100\%$$

where M_{ij} is the obstacle degree of the jth indicator at the ith sampling site; P_{ij} is the gap between the indicator and the maximum target, set as the standardized value of the indicator; W_j is the contribution degree of the factor to the total target, i.e., the weight of the factor; S_{ij} is the impact of the indicator on the soil quality of the study area, representing the standardized value of the jth indicator at the ith sampling site (standardized equation: $S_{ij} = \frac{x_{ij}-x_{min}}{x_{max}-x_{min}}$); and M_j is the average obstacle degree of the jth evaluation factor.

1.4.3 Accuracy Validation of Farmland Soil Quality Evaluation The coefficient of determination (R^2) and Nash-Sutcliffe efficiency coefficient (E_{NS}) were used to evaluate the accuracy of the minimum datasets based on principal component analysis and cluster analysis []. The calculation formula is:

$$E_{NS} = 1 - \frac{\sum_{i=1}^{n}(SQI_{cal,i} - SQI_{obs,i})^2}{\sum_{i=1}^{n}(SQI_{obs,i} - \overline{SQI_{obs}})^2}$$

where $SQI_{obs,i}$ and $\overline{SQI_{obs}}$ are the farmland soil quality index values and average soil quality index calculated based on the total dataset, respectively; $SQI_{cal,i}$ is the farmland soil quality index value calculated based on the minimum dataset; R^2 is the coefficient of determination. A larger E_{NS} value indicates that the farmland soil quality index calculated based on the minimum dataset is closer to the benchmark value and has higher accuracy [].

1.5 Data Processing

SPSS 26.0 software was used for principal component analysis and Pearson correlation analysis of soil data. Microsoft Office Excel was used for data processing and regression analysis. ArcGIS 10.8 software was used for spatial distribution mapping.

---

2. Results

2.1 Descriptive Statistical Analysis of Farmland Soil Quality Indicators in Shawan City

The analysis of 84 soil samples from farmland in Shawan City is shown in Table . The sensitivity of each indicator is generally measured by the coefficient of variation (CV). A larger CV indicates greater differences in the evaluation indicator's response to soil quality [], which can be divided into three types: insensitive indicators (CV $<$ 10%), moderately sensitive indicators (10% $\leq$ CV $<$ 40%), and highly sensitive indicators (40% $\leq$ CV $<$ 100%). As shown in Table , soil organic matter, available phosphorus, available potassium, total nitrogen, and alkali-hydrolyzable nitrogen are moderately sensitive indicators (CV = 40%–100%), which are the main factors for improving farmland soil quality

[]. Total phosphorus is a low-sensitivity indicator (CV $< 10\%$), which is a major factor for maintaining stable farmland soil structure [].

According to the National Second Soil Nutrient Classification Standard (Table), the mean organic matter content of farmland soil samples in Shawan City is 14.448 $g \cdot kg^{-1}$, which is at Level 6, indicating extreme nutrient deficiency. The main reasons may include: first, the high temperature and low rainfall in the study area cause severe soil compaction, which is not conducive to organic matter accumulation []; second, some areas are located at the basin edge with severe desertification and weak basic fertility; and third, during the pandemic, farmland did not receive timely nutrient supplementation, resulting in low organic matter content that requires increased organic fertilizer application in future fertilization management [].

Soil pH is generally strongly alkaline. The mean pH value is 8.28, indicating that Shawan City farmland soil is generally alkaline, which is difficult to improve through conventional tillage and fertilization measures. Therefore, attention should be paid to the application of alkaline fertilizers during agricultural production []. Soil total phosphorus and available phosphorus contents are at Level 2, with mean values of 0.975 $g \cdot kg^{-1}$ and 22.342 $mg \cdot kg^{-1}$, respectively, indicating rich phosphorus content. The high CV of soil available phosphorus may be related to unbalanced fertilizer application. Additionally, phosphorus moves slowly in soil, and its spatial distribution heterogeneity is difficult to improve over time, both contributing to the large CV of soil available phosphorus. The mean available potassium is 355.554 $mg \cdot kg^{-1}$, at Level 1, indicating relatively high potassium content.

Total nitrogen is the capacity indicator of soil nitrogen [], while alkali-hydrolyzable nitrogen reflects the short-term nitrogen supply capacity of soil. The mean total nitrogen and alkali-hydrolyzable nitrogen contents of farmland soil in Shawan City are 0.975 $g \cdot kg^{-1}$ and 64.568 $mg \cdot kg^{-1}$, respectively, both at Level 4, indicating low contents and moderate nutrient deficiency. Moreover, alkali-hydrolyzable nitrogen is significantly negatively correlated with pH, indicating that excessive nitrogen fertilizer application may cause soil acidification and intensify soil nitrification [].

2.2 Minimum Dataset Established Based on Principal Component Analysis

Among the eight selected farmland soil quality evaluation indicators, three principal components had eigenvalues > 1, with a cumulative contribution rate of 76.459%, which can explain most of the variability in soil indicators. Total nitrogen, available potassium, organic matter, alkali-hydrolyzable nitrogen, total phosphorus, and soil pH were grouped into Group 1; organic matter, soil pH, and electrical conductivity were grouped into Group 2; and total phosphorus and available phosphorus were grouped into Group 3. Since organic matter, total phosphorus, and available phosphorus simultaneously appeared in two principal components, correlation analysis was conducted (Table). The correlation coefficient between soil organic matter and total nitrogen was 0.737 ($P < 0.05$), showing high correlation. Therefore, organic matter was assigned to Group 1. Using the same method, pH, total phosphorus, and available phosphorus were sequentially analyzed. The final grouping was updated to: total nitrogen, available potassium, alkali-hydrolyzable nitrogen, and pH in Group 1; organic matter, total phosphorus, and electrical conductivity in Group 2; and available phosphorus alone in Group 3.

Since organic matter, total phosphorus, and available phosphorus simultaneously appeared in two principal components, correlation analysis was performed (Table). In Group 1, total nitrogen had the largest Norm value (0.913) and was selected for the minimum dataset, while available potassium, pH, and alkali-hydrolyzable nitrogen were eliminated because their Norm values were not within the range of total nitrogen. In Group 2, organic matter had the largest Norm value (0.891) and was selected for the minimum dataset, while electrical conductivity was eliminated because its Norm value was not within the range of organic matter. Group 3 directly reflected changes in available phosphorus, so available phosphorus was selected for the minimum dataset. In summary, total nitrogen, organic matter, available phosphorus, and electrical conductivity were selected as the four indicators entering the minimum dataset.

2.3 Minimum Dataset Established Based on Cluster Analysis

Cluster analysis using the Q-type method classified the total dataset evaluation indicators, establishing a minimum dataset for soil quality evaluation. The results showed that at the clustering level of 0.5–0.7, the eight soil indicators could be clearly divided into four categories: organic matter, total nitrogen, and alkali-hydrolyzable nitrogen were classified as Category 1; total phosphorus and available phosphorus as Category 2; pH alone as Category 3; and electrical conductivity and available potassium as Category 4.

In Category 1, the correlation coefficient between soil organic matter and total nitrogen was 0.737 ($P < 0.05$). Since most soil nitrogen comes from organic matter, increasing organic matter content can improve soil fertility retention capacity, so organic matter was selected for the minimum dataset []. In Category 2, available phosphorus is the main nutrient resource obtained by crops from soil and is basically not disturbed by external measures such as fertilization []. Jin et al. [] found that changes in available phosphorus content in red soil can reflect soil quality changes, so available phosphorus was selected for the minimum dataset. Category 3 alone was pH, the main indicator reflecting soil acidity and alkalinity, with a usage rate of 0.5–0.7 in soil quality evaluation [], so pH was selected for the minimum dataset. Salinization is generally reflected by electrical conductivity values [], so electrical conductivity was selected as an alternative for the minimum dataset. In summary, the indicators entering the minimum dataset were organic matter, available phosphorus, pH, and electrical conductivity.

2.4 Validation of Minimum Dataset for Farmland Soil Quality Evaluation

The reasonableness validation of the minimum dataset evaluation index system is an important part of farmland soil quality evaluation []. To verify the evaluation accuracy of the minimum dataset index system, this study conducted regression analysis between the soil quality indices of the minimum dataset and total dataset, comparing their consistency (Fig. [Figure 3: see original paper]).

The soil quality index based on principal component analysis varied from 0.100 to 0.528, with a mean of 0.298 and CV of 50.6%, approaching high variation. The total dataset soil quality index varied from 0.120 to 0.641, with a mean of 0.386 and CV of 22.61%, indicating low variation. The soil quality index based on cluster analysis varied from 0.130 to 0.641, with a mean of 0.369 and CV of 21.65%, also indicating low variation.

These results showed that the variation range, mean, and CV of the soil quality index based on cluster analysis were closer to those of the total dataset soil quality index than those based on principal component analysis. Scatter plots were created with regression analysis between the total dataset soil quality index and those based on principal component analysis and cluster analysis. The fitting results showed that both relationships were significantly positively correlated, with R^2 values of 0.379 and 0.591, respectively. The soil quality index based on cluster analysis showed better fitting results than that based on principal component analysis. Additionally, the effective coefficients between the total dataset soil quality index and those based on cluster analysis and principal component analysis were 4.923 and 1.108, respectively. Overall, the minimum dataset based on cluster analysis provided higher accuracy for farmland soil quality evaluation than that based on principal component analysis. Therefore, the minimum dataset based on cluster analysis was selected to replace the total dataset for evaluating farmland soil quality in Shawan City.

2.5 Analysis of Farmland Soil Quality Index and Obstacle Factors in Shawan City

The obstacle diagnosis model was used to analyze the eight soil physicochemical indicators. The obstacle degrees of pH and total phosphorus were 0.084 and 0.096, respectively, belonging to mild obstacle factors. The obstacle degrees of organic matter, electrical conductivity, available phosphorus, available potassium, total nitrogen, and alkali-hydrolyzable nitrogen were 0.221, 0.203, 0.196, 0.192, 0.189, and 0.221, respectively, all belonging to moderate obstacle factors. Therefore, the main obstacle factors for farmland soil in Shawan City were, in order, organic matter, alkali-hydrolyzable nitrogen, electrical conductivity, total nitrogen, available potassium, and available phosphorus.

Based on the minimum dataset soil quality index, the farmland soil quality index in Shawan City ranged from 0.130 to 0.641, with a mean of 0.369. According to the calculated farmland soil quality index, the spatial distribution of soil quality was realized using the ordinary Kriging interpolation model in ArcGIS 10.8 (Fig. [Figure 4: see original paper]). Higher soil quality indices indicate that the farmland soil in the study area is more suitable for crop growth and has higher crop yields. Combined with literature and soil characteristics of the study area, the soil evaluation results were divided into five classes, with agricultural production suitability decreasing progressively. The results showed that Class I and II soils were mainly distributed in the northern and northwestern parts of the study area, while Class V soil was mainly distributed in the southeastern part, showing an overall spatial distribution pattern of high quality in the northwest and low quality in the southeast. Overall, the farmland soil quality in Shawan City was at a moderate or better level, but obvious obstacle factors existed. The main obstacle factors were soil organic matter, alkali-hydrolyzable nitrogen, and electrical conductivity, indicating that nutrient deficiency and salinization are currently the main obstacle factors facing farmland in Shawan City.

Discussion

This study selected eight soil physicochemical properties as farmland soil evaluation indicators. Among the minimum dataset established through cluster analysis, four indicators (pH, electrical conductivity, organic matter, and available phosphorus) entered the minimum dataset with the highest usage frequency, indicating that the total dataset and minimum dataset evaluation index systems in this study are similar to previous research results and have certain representativeness []. By comparing the results of total dataset soil quality index and minimum dataset soil quality index, it was found that they were extremely significantly positively correlated ($R^2 = 0.591$), indicating that the minimum dataset based on cluster analysis can better replace total dataset indicators. This is consistent with the research results of Bao et al. [] on soil quality evaluation for tobacco planting in Baoshan City.

Soil organic matter can provide nutrients needed for plant growth and supply nutrients and energy for soil microorganisms and soil animal activities; pH and electrical conductivity can reflect the salinization degree of local farmland soil; and available phosphorus can intuitively reflect nutrient supply during soil quality evaluation []. The selection of these indicators is similar to most domestic and international research results []. However, some nutrient indicators in Shawan City farmland soil are not within the production suitability range of high-standard farmland. Therefore, to promote local agricultural development and accelerate the process of high-standard farmland construction, cultivation plans should be formulated according to local conditions. The application amount of indicators exceeding the production suitability range should be reduced, while the application amount of indicators below the production suitability range should be appropriately increased, combined with tillage and fertilization methods.

According to the ranking of soil indicator obstacle degrees, organic matter had the largest obstacle degree. In addition to the three reasons mentioned above for low organic matter content, there may still be individual factors not considered, such as the fact that this study focused on nutrient indicators without considering soil physical properties, leading to incomplete analysis. Moreover, in soil quality evaluation, trace element content and soil biological property indicators are also important aspects [], which were not considered in this study regarding their contribution to soil quality. Additionally, this study only examined soil during a specific time period in the study area, without investigating temporal variation characteristics of soil quality, which requires further research.

The obstacle degree of alkali-hydrolyzable nitrogen was also relatively large, and some sampling points showed excessive potassium and phosphorus fertilizer contents, possibly because sampling occurred after local farmers had begun sowing and 大量使用农药化肥. The obstacle degree of soil electrical conductivity was relatively large, with a maximum value of 4540 $\mathrm{S \cdot cm^{-1}}$ and a mean value of 587.7 $\mathrm{S \cdot cm^{-1}}$, indicating high soil salinization in Shawan City farmland. Additionally, pH and electrical conductivity showed a significant negative correlation. According to Yu [], this may be caused by fertilization factors, as the application of large amounts of chemical fertilizers aggravates the enrichment of soil salt ions, causing pH to decrease at some sampling points. During crop growth, most areas experience high temperatures and drought with little rainfall, and all water supply comes from canal water or groundwater. During irrigation, the high pH and electrical conductivity values of irrigation water aggravate farmland soil salinization, increasing the vulnerability of the local farmland ecosystem and negatively affecting crop growth.

The minimum dataset based on cluster analysis consists of four indicators: pH, electrical conductivity, organic matter, and available phosphorus. The results of total dataset and minimum dataset soil quality indices show a significant positive correlation, indicating that the minimum dataset can effectively reflect local farmland soil quality. The soil quality index ranges from 0.130 to 0.641 with a mean of 0.369, indicating that most soil quality in the study area belongs to moderate or better levels.

The main obstacle factors for farmland soil in Shawan City are low organic matter content, nitrogen deficiency, and high electrical conductivity, with obstacle degrees of 0.221, 0.221, and 0.203, respectively. Targeted tillage and fertilization measures are needed for these indicators.

References

[1] Qiao Yunfa, Zhong Xin, Miao Shujie, et al. Evaluation indicators for soil quality of cultivated layer in northeast sandy soil farmland based on minimum dataset[J]. Research of Soil and Water Conservation, 2019, 26(4): 132-138.

[2] Chen Zhengfa, Shi Dongmei, Jin Huifang, et al. Evaluation of soil quality in the cultivated layer of slope farmland in Yunnan Province based on a soil management assessment framework[J]. Transactions of the Chinese Society of Agricultural Engineering, 2019, 35(3): 256-267.

[3] Zhou Wentao, Lan Tian, Pan Yue, et al. Construction of a minimal dataset with few samples in soil quality evaluation: A case study of a typical plot in the south bank irrigation district of the Yellow River in Hangjin Banner, Inner Mongolia[J]. Journal of China Agricultural University, 2022, 27(6): 225-235.

[4] Choudhury B U, Mandal S. Indexing soil properties through constructing minimum datasets for soil quality assessment of surface and profile soils of intermontane valley (Barak, northeast India)[J]. Ecological Indicators, 2021, 123: 107369, doi: 10.1016/j.ecolind.2021.107369.

[5] Jin Huifang, Shi Dongmei, Chen Zhengfa, et al. Evaluation indicators of soil quality in the cultivated layer of red soil slope farmland based on clustering and PCA analysis[J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(7): 155-164.

[6] Li Yuchen, Xiong Aoyu, Ping Yuan, et al. Evaluation of orchard soil quality under different land preparation methods in Ningdu County based on minimum dataset[J]. Journal of Soil and Water Conservation, 2023, 37(4): 342-350.

[7] Hu Bingtao, Chen Yudong. Screening of nitrogen and phosphorus monitoring indicators for rice fields in the lower reaches of the Yangtze River based on cluster analysis and principal component analysis[J]. Journal of Ecology and Rural Environment, 2020, 36(11): 1495-1504.

[8] Li Pengfei, Zhang Xingchang, Hao Mingde, et al. Evaluation of soil quality for reclamation in Loess Plateau mining areas based on minimal dataset[J]. Transactions of the Chinese Society of Agricultural Engineering, 2019, 35(16): 265-273.

[9] Li Xia, Zhu Wanze, Shu Shumiao, et al. Principal component analysis based soil quality evaluation of grassland in the dry warm valley of the middle reaches of the Dadu River[J]. Acta Ecologica Sinica, 2021, 41(10): 3891-3900.

[10] Yuan Ye, Wang Hao, Luo Yang, et al. Evaluating the impact of tillage methods on soil quality in black soil farmland based on the minimum dataset[J]. Journal of Maize Sciences, 2023, 31(4): 148-157.

[11] Bao Lingfeng, Yang Mingying, Yin Xingsheng, et al. Evaluation of soil quality and obstacle diagnosis for tobacco planting in Baoshan City based on minimal dataset[J]. Southwest China Journal of Agricultural Sciences, 2023, 36(3): 612-622.

[12] Yu P J, Han D L, Liu S W, et al. Soil quality assessment under different land uses in an alpine grassland[J]. CATENA, 2018, 171: 105202, doi: 10.1016/j.catena.2018.02.015.

[13] Zhang J H, Ding Q D, Wang Y J, et al. Soil quality assessment and constraint diagnosis of salinized farmland in the Yellow River irrigation area in northwestern China[J]. Geoderma Regional, 2023, 34: e00684, doi: 10.1016/j.geodrs.2023.e00684.

[14] Raiesi F, Pejman M. Assessment of post wildfire soil quality and its recovery in semi-arid upland rangelands in central Iran through selecting the minimum data set and quantitative soil quality index[J]. CATENA, 2021, 201: 105202, doi: 10.1016/j.catena.2021.105202.

[15] Cong P, Wang J, Li Y Y, et al. Changes in soil organic carbon and microbial community under varying straw incorporation strategies[J]. Soil and Tillage Research, 2020, 204: 104735, doi: 10.1016/j.still.2020.104735.

[16] Jiang Congze, Shou Na, Gao Wei, et al. Comprehensive evaluation of soil quality of different land use types in the northeast margin of the Qinghai Tibet Plateau[J]. Chinese Journal of Applied Ecology, 2022, 33(12): 3279-3286.

[17] Mei Nan, Gu Yan, Li Dezhong, et al. Evaluation of soil quality of black soil cultivated layer in Jilin Province based on minimal dataset[J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(12): 91-98.

[18] Chen Fangzheng, Ren Jian, Liu Sihan, et al. Comprehensive evaluation of soil fertility of cultivated land in the southern Dongting Lake Basin based on minimal dataset[J]. Chinese Journal of Soil Science, 2021, 52(6): 1348-1359.

[19] Shi Dongmei, Jiang Na, Jiang Guangyi, et al. Factors and sensitivity analysis on the quality of cultivated layer in purple soil slope farmland[J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(3): 135-143.

[20] Guo J H, Liu X J, Zhang Y, et al. Significant acidification in major Chinese croplands[J]. Science, 2010, 327(5968): 1008-1010.

[21] Zhou Wenyu, Yang Xiaohu, Yang Haichang, et al. Evaluation of soil quality in typical oasis farmland based on minimal dataset[J]. Journal of Irrigation and Drainage, 2023, 42(7): 117-122.

[22] Liu Likun, Zhao Guangju, Mu Xingmin, et al. Soil quality evaluation along the Qinghai Tibet Highway based on minimum dataset[J]. Research of Soil and Water Conservation, 2022, 29(2): 125-131.

[23] Liu Xiangjun, Qiao Guanyu, Guo Fenghao, et al. Evaluation and obstacle analysis of topsoil quality in the Huanghuaihai dry farming area based on minimum dataset[J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(12): 104-113.

[24] Yu Zhongyang. Study on soil organic carbon change and carbon nitrogen coupling relationship in facility agriculture[D]. Xining: Qinghai University, 2023.

[25] Cao Gang, Bi Shuhai, Zhao Mingxin, et al. Comprehensive evaluation of soil environmental factors under different coverage conditions in pear orchards in arid areas[J]. Arid Land Geography, 2022, 45(3): 890-900.

[26] Song Hongbing, Guo Weiling. Comprehensive evaluation of soil quality in Jianghuai hilly area based on minimum dataset[J]. Agriculture and Technology, 2023, 43(12): 93-98.

[27] Han T F, Liu K L, Huang J, et al. Temporal and spatial characteristics of paddy soil potassium in China and its response to organic amendments: A systematic analysis[J]. Soil and Tillage Research, 2024, 235: 105894, doi: 10.1016/j.still.2023.105894.

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