Scholars Journal of Engineering and Technology

OPEN ACCESS

Abbreviated Key Title: Sch J Eng Tech
ISSN 2347-9523 (Print) | ISSN 2321-435X (Online)
Journal homepage: https://saspublishers.com

Comprehensive Evaluation of Soil Fertility Based on Fuzzy Dynamic Clustering Algorithm

Sun Gongbing[1], Yuan Yuping[2*]

[1]School of Food and Pharmaceutical Engineering, Wuzhou university, Guangxi Wuzhou 543002, China
[2]School of Data Science & Software Engineering, Wuzhou University, Guangxi Wuzhou 543002, China

DOI: https://doi.org/10.36347/sjet.2024.v12i12.002 | **Received:** 19.05.2024 | **Accepted:** 26.06.2024 | **Published:** 07.12.2024

***Corresponding author:** Yuan Yuping
School of Data Science & Software Engineering, Wuzhou University, Guangxi Wuzhou 543002, China

Abstract **Original Research Article**

The distribution of soil nutrients in cultivated land was clarified and the soil fertility was systematically evaluated and analyzed. First, five important indicators that have a greater impact on soil comprehensive fertility were selected, and the index data were standardized. According to the different impacts of each indicator on soil fertility, the average value of partial correlation coefficient between each indicator and other indicators was calculated, and a certain weight was given to each indicator. The inverse absolute value method was used to construct a fuzzy similarity matrix, and the fuzzy similarity matrix was transformed into a fuzzy equivalent matrix to obtain dynamic clustering results. Finally, statistics were used to determine the optimal threshold value as 0.993 and the relative optimal soil classification number as 7 categories. The experimental results are in agreement with the actual situation, which proves the effectiveness of the fuzzy dynamic clustering algorithm.

Keywords: Soil fertility; Fuzzy dynamic clustering; Fuzzy similarity matrix; Comprehensive evaluation.

INTRODUCTION

Soil is closely related to human society. In precision agriculture, global Positioning system (GPS), remote sensing system (RS), geographic information system (GIS) and sensor technology can obtain a large amount of soil nutrient data. The comprehensive fertility of soil is high, which is conducive to the growth and maturity of crops, the increase of cash crops and the increase of farmers' income. In the previous evaluation process of soil comprehensive fertility, there were no quantitative indicators with objective standards, and it was usually said that the soil was fertile or poor. So far, scientific research has not established a clear and unified standard for the classification of soil comprehensive fertility [1]. In particular, China has a vast territory and complex and diverse natural environment, so it is very difficult to introduce a unified evaluation system for soil comprehensive fertility in different places [2]. At present, most comprehensive fertility evaluation of soil is mainly focused on a certain land or soil with similar properties.

In recent years, domestic soil researchers have divided cultivated land into different grades according to the fertility level, and then determined the application level. For example, Wang Hao *et al.*, [3] conducted field plot experiments to investigate the effects of partial replacement of fertilizer by organic fertilizer on rice yield, nutrient utilization, nitrogen and phosphorus fertilizer efficiency and soil fertility at the same phosphorus application level. Hu Litao *et al.*, [4] used fuzzy mathematics method and principal component analysis method to comprehensively evaluate the soil pH value, organic matter, alkali-hydrolytic nitrogen, available phosphorus and available potassium contents in Taipingba Township, a major tobacco planting area in Fengdu. Ji Fenghui [5] and Zhang Ying [6] used hierarchical analysis, comprehensive index method and fuzzy mathematics theory to determine the weight and membership of each evaluation factor, calculate the comprehensive land fertility index, and divide the cultivated land in Tongliao City into five grades according to the accumulation curve method, and analyzed the cultivated land grade and nutrient status.

In this paper, 857 farm cultivated land was selected as the research object, the main factors affecting soil fertility were analyzed, the data of factors affecting soil comprehensive fertility were standardized, weighted, fuzzy similarity matrix was constructed, and fuzzy dynamic clustering algorithm was adopted to study the classification of soil comprehensive fertility.

1. Basic Steps of Fuzzy Dynamic Clustering Algorithm

1.1 Steps of Analytic Hierarchy Process to Determine the Weight Coefficient

Step 1: Construct pairwise judgment matrix A;

Step 2: Normalize the initial vector $w^{(0)}$ with any n dimension;

Step 3: Calculation $\overline{w}^{(k+1)} = Aw^{(k)}, k = 1, 2, \cdots,$ and normalization $\overline{w}^{(k+1)}$;

Step 4: Take any value $\varepsilon > 0$, when $|w_i^{(k+1)} - w_i^{(k)}| < \varepsilon, i = 1, 2, \cdots, n$, it is true, $\overline{w}^{(k+1)}$ is the desired feature vector; Otherwise, return Step 2;

Step 5: Calculate the maximum eigenvalue λ and calculate the consistency index $CI = \frac{\lambda - n}{n-1}$;

Step 6: If $CR = \frac{CI}{RI} < 0.1$, through the consistency test, the weight vector is obtained; Otherwise, reconstruct the judgment matrix and return to Step 1.

1.2 Steps of Fuzzy Dynamic Clustering Algorithm

Step 1: Data standardization, that is $x'_{ij} = \frac{x_{ij} - \min\{x_{ij}\}}{\max\{x_{ij}\} - \min\{x_{ij}\}}$;

Step 2: Calculate the weight, that is: $Y = \begin{bmatrix} x'_{11} & x'_{12} & \cdots & x'_{1m} \\ x'_{21} & x'_{22} & \cdots & x'_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ x'_{n1} & x'_{n2} & \cdots & x'_{nm} \end{bmatrix} \cdot \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_m \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1m} \\ y_{21} & y_{22} & \cdots & y_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ y_{n1} & y_{n2} & \cdots & y_{nm} \end{bmatrix}$;

Step 3: Calculate the similarity r_{ij} between fuzzy set i and fuzzy set j ., $r_{ij} = \frac{\sum_{k=1}^{m} y_{ik} \cdot y_{jk}}{\sqrt{\sum_{k=1}^{m} y_{ik}^2} \cdot \sqrt{\sum_{k=1}^{m} y_{jk}^2}}$;

The fuzzy similarity matrix $R = (r_{ij})_{n \times n}$ is obtained.

2. Comprehensive Evaluation of Soil Fertility by Fuzzy Dynamic Clustering Algorithm

2.1 Standardization of soil fertility index factor data

PH value, organic matter, nitrogen, phosphorus and potassium were selected as the indexes affecting soil fertility. The membership function of the factors affecting soil comprehensive fertility was established, and the membership value of the factors affecting soil comprehensive fertility was calculated. Within a certain range, the effect curves of all factors affecting soil comprehensive fertility on crops are S-shaped, and the membership function of the S-shaped curve is transformed into the corresponding membership function of the broken line.

$$f(X) = \begin{cases} 1.0 & x \geq x_2 \\ 0.9(x - x_1)/(x_2 - x_1) + 0.1 & x_1 \leq x \leq x_2 \\ 0.1 & x \leq x_1 \end{cases}$$

According to literature [7] and the actual situation of acidic soil, the corresponding values of turning points in the curve are determined as shown in Table 1.

Table 1: Turning point values of membership function curves of soil fertility factors

Item	PH value	Organic matter（g/kg）	alkali-hydrolyzed nitrogen （mg/kg）	Available Phosphorus （mg/kg）	Available potassium（mg/kg）
X1	4.5	10	60	5	50
X2	6.0	30	120	20	150

The membership value of each factor affecting soil comprehensive fertility ranges from 0.1 to 1, which reflects the membership degree. 0.1 indicates that a factor affecting soil comprehensive fertility is seriously lacking in the soil, which is very unfavorable to the growth of crops. The minimum value here is 0.1 in order to meet the reality of production and life. See Table 2.

Table 2: Standardized index data of five indicators of 857 farm

Soil numbering	PH value	Organic matter	Nitrogen	Phosphorus	Potassium
1	0.886	0.857	1	1	0.478
2	0.634	1	1	1	1
3	0.634	0.975	1	1	0.828
4	0.754	1	1	1	1
5	0.58	1	1	1	0.633
...	...	...	...	...	...
933	0.886	1	1	1	0.820
934	0.784	1	1	1	1
935	0.694	1	1	1	1
936	0.700	1	1	1	1

2.2 Weight determination of factors affecting soil comprehensive fertility

The factors that affect the overall fertility of soil have different contributions to the overall fertility of soil. The actual contribution rate of each factor to soil fertility can only be obtained by studying the correlation between the two factors separately under the premise of excluding or fixing other factors affecting soil fertility. The average value of partial correlation coefficient () between each factor and other factors was calculated, and the proportion of the average value () to the sum of the average value of partial correlation coefficient () of all fertility indicators was the weight () of the single factor affecting soil comprehensive fertility.

$$R=\begin{bmatrix} 1.0000 & -0.0073 & -0.0493 & -0.1228 & 0.0594 \\ -0.0073 & 1.0000 & 0.1996 & 0.0581 & 0.1301 \\ -0.0493 & 0.1996 & 1.0000 & 0.0083 & 0.0320 \\ -0.1228 & 0.0581 & 0.0581 & 1.0000 & -0.0009 \\ 0.0594 & 0.1301 & 0.0320 & -0.0009 & 1.0000 \end{bmatrix}$$

$$R^{-1}=\begin{bmatrix} 1.0216 & -0.0021 & 0.0518 & 0.1251 & -0.0620 \\ -0.0021 & 1.0620 & -0.2073 & -0.0603 & -0.1314 \\ 0.0518 & -0.2073 & 1.0442 & 0.0097 & -0.0095 \\ 0.1251 & -0.0603 & 0.0097 & 1.0188 & 0.0011 \\ -0.0620 & -0.1314 & -0.0095 & 0.0011 & 1.0211 \end{bmatrix}$$

The average partial correlation coefficients between each factor and other factors were 0.0275, 0.2474, 0.3027, 0.0550 and 0.2474, respectively. Therefore, the contribution rates of single fertility indexes in the characterization of soil fertility were 0.0313, 0.2811, 0.3440, 0.0625 and 0.2811, respectively. The weight reweighting of each evaluation factor is shown in Table 3.

Table 3: Weights of each evaluation factor

Item	PH value	Organic matter	Nitrogen	Phosphorus	Potassium
Weight	0.0313	0.2811	0.344	0.0625	0.2811

2.3 Establishment of fuzzy similarity matrix

In this case, different indicators are treated differently, and the weights of different indicators are assigned to the indicators to achieve the accuracy of data results. Weighting process of 857 farm soil important index data:

$$Y=\begin{bmatrix} 0.886 & 0.857 & ... & 0.479 \\ 0.634 & 1 & ... & 1 \\ ... & ... & ... & ... \\ 0.7 & 1 & ... & 1 \end{bmatrix} \bullet \begin{bmatrix} 0.0313 & 0 & ... & 0 \\ 0 & 0.281 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & 0 & 0.281 \end{bmatrix} = \begin{bmatrix} 0.0277 & 0.241 & ... & 0.135 \\ 0.0198 & 0.281 & ... & 0.281 \\ ... & ... & ... & ... \\ 0.0219 & 0.281 & ... & 0.281 \end{bmatrix}$$

The result of clustering depends on fuzzy similar matrix, and the construction method of fuzzy similar matrix is not unique, which is easy to cause the instability of clustering result. In this paper, the reciprocal absolute value method is used to construct fuzzy similarity matrix. The approximation degree r_{ij} of fuzzy set i and fuzzy set j is established.

$$r_{ij}=\begin{cases} 1 & i=j \\ \dfrac{c}{\sum_{k=1}^{m}|y_{ik}-y_{jk}|} & i\neq j \end{cases}$$

To ensure that the elements of the matrix are between [0,1]. Set c to 0.1814 . is set. The fuzzy similarity matrix $R=(r_{ij})_{n\times n}$1 of 857 farm soil data was obtained.

$$R=\begin{pmatrix} 0.154 & 0.253 & \cdots & 0.146 \\ 0.253 & 0.734 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 0.146 & 1 & \cdots & 0.462 \end{pmatrix}$$

The fuzzy similar matrix is also changed into fuzzy equivalent matrix $t(R)$ by the quadratic self-synthesis method:

$$t(R)=\begin{pmatrix} 1 & 0.97 & \cdots & 0.98 \\ 0.97 & 1 & \cdots & 0.97 \\ \cdots & \cdots & \cdots & \cdots \\ 0.98 & 0.97 & \cdots & 1 \end{pmatrix}$$

The dynamic clustering result is obtained by arranging the values of different elements in fuzzy equivalent matrix $t(R)$ from the largest to the smallest and λ in order.

2.4 Definite classification number

Using the F statistic to determine the optimal threshold λ, calculate the center vector x in the original sample data matrix D:

$$\overline{x_k}=\left(\frac{1}{n}\sum_{i=1}^{n}x_{ik}\right)(k=1,2\cdots,m)$$

Let the number of categories corresponding to the value of λ be r, the number of samples of class j be n_j, the number of samples of class j be recorded as $x_1^{(j)},x_2^{(j)},\cdots,x_{n_j}^{(j)}$, and the cluster center vector of class j be $\overline{x}^{(j)}=\left(\overline{x}_1^{(j)},\overline{x}_2^{(j)}\cdots\overline{x}_n^{(j)}\right)$. Among them,

$$\overline{x}_k^{(j)}=\frac{1}{n}\sum_{i=1}^{n_j}x_{ik}^{(j)}(k=1,2,\cdots,m)$$

As the F statistic, we get:

$$F=\frac{\sum_{j=1}^{r}n_j\left\|\overline{x}^{(j)}-\overline{x}\right\|/(r-1)}{\sum_{j=1}^{r}\sum_{i=1}^{n_j}\left\|x_i^{(j)}-x^{-(j)}\right\|/(n-r)}$$

The denominator represents the distance between samples within the class, and the numerator represents the distance between classes. Therefore, the F-value is proportional to the classification effect. The larger the F-value is, the larger the distance between the classes formed by the cluster is, and the better the classification effect of fuzzy dynamic clustering is. If the

F-value is smaller, the fuzzy dynamic clustering classification effect is worse. According to the principle of mathematical statistics analysis of variance, if

$F > F_\alpha(r-1, n-r), (\alpha = 0.05)$

The difference between the formed classes and classes is significant. If there is more than one value of $F > F_\alpha(r-1, n-r)$, you need to further compare the size of $(F - F_\alpha)$ to find a satisfactory F-value from the larger one. The value of the F statistic is calculated separately when the value of λ is different. When the value of $F > F_{0.05}(r-1.n-r)$ is calculated, the value of $F - F_{0.05}(r-1.n-r)$ is obtained, and the larger value of λ is 0.993. When λ=0.993, the truncated matrix $R_{0.993}$ is:

$$R_{0.993} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

2.5 RESULTS AND ANALYSIS

Based on the fuzzy dynamic analysis results of soil PH value, organic matter, nitrogen, phosphorus and potassium of 857 farm in autumn 2018, five factors affecting soil comprehensive fertility data were analyzed, as shown in Table 4.

Table 4: Comparison of clustering results

λ	Class number	F-value
1	134	0.354
0.999	112	0.672
0.998	90	1.128
0.997	73	2.114
0.996	49	3.263
0.995	24	3.971
0.994	13	4.576
0.993	7	4.684
0.992	5	3.914
0.991	3	2.547
0.990	2	1.260
0.899	1	0.421

In the table 5, λ represents the degree of similarity between the data, the number of classes corresponds to the number of classes whose similarity is λ, and the value of F is the value of F distribution calculated using probability statistics. As can be seen from Table 5, when λ=0.993, the F value is the largest and the classification is better.

3. CONCLUSION

According to the values of five important soil indicators, combined with the classification results, the clustering results of the application of Autonavi Open platform are shown in Figure 1.

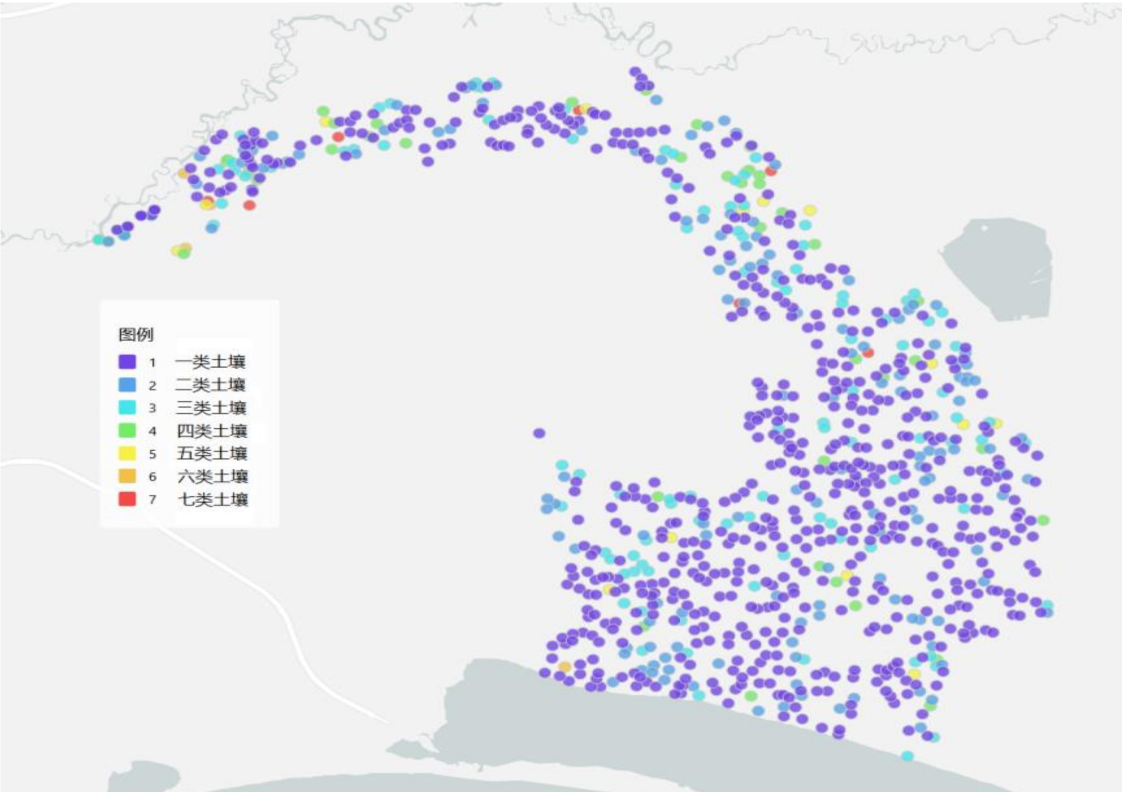

Figure 1. Graph of clustering results using Gaode Open Platform

The soil marked 1 in Figure 1 is the soil with the best comprehensive soil fertility in farm 857, and so on, the soil marked 7 is the soil with the worst comprehensive soil fertility in the clustering results. 64.74% of 857 farm soils were classified as one type of soil in the clustering results. 15.28% of 857 farm soils were classified as category 1 and category 2 soil in the clustering results. 11.43% of 857 farm soils were classified as one type and three types of soil in the clustering results. 5.02% of 857 farm soils were classified as one type and four types of soil in the clustering results. 1.92% of 857 farm soils were classified as one type and five types of soil in the clustering results. 0.75% of 857 farm soils were classified as one type, six types of soil in the clustering results. 0.85% of 857 farm soils were classified as one type in the clustering results, which were classified as seven types of soils.

In order to better reflect the level of soil comprehensive fertility, we must grasp the most important factors affecting soil comprehensive fertility, and then make an effective analysis. The multi-index comprehensive evaluation model has wide application value in soil comprehensive fertility evaluation, but it should be more strict in index selection, function threshold determination, weight determination of influencing factors and so on, so as to reduce the influence of subjective factors.

Acknowledgments

This work was financially supported by the wuzhou university talent introduction research start-up fund "Twin Support Vector Machine and its Application in Hydrological Prediction" (Project No. WZUQDJJ17143)

REFERENCES

1. Huang, Y. & Yang, Z. F. (2008). Research status and prospect of land quality evaluation in China. *Geological Bulletin of China*, 27(2): 207-211.
2. Wu, Chunlin., Wang, Ruiting. & Ding, Kun. (2018). Research status and progress of geochemical investigation and evaluation of soil quality in China [J]. *Northwest Geology*, 51(3): 240-252.
3. Wang, Hao., Li, Yanjun. & Sun, Xiaogen. (2019). Effects of organic fertilizer substitution on soil fertility, nitrogen and phosphorus utilization and rice yield in Chaohu Lake Basin [J]. *Fertilizer and Health*, 51 (02) :45-51. (in Chinese)
4. Hu, Litao., Li, Hang. & Su, Mengdi. (2024). Comprehensive evaluation of soil fertility for tobacco planting in Fengdu County, Chongqing [J], *Zhejiang Agricultural Sciences*, 03:1-6.
5. JI, Fenghui., Hou, Mihong. & GE, Haifeng. (2017). Soil fertility evaluation and nutrient analysis of cultivated land in Tongliao City [J], *Journal of Inner Mongolia University for Nationalities (Natural Science Edition)*, 32(05) : 429-436.
6. Zhang Ying, Z. Y., Zhao GengXing, Z. G., Wang ZhuoRan, W. Z., Chen XiaoNa, C. X., & Su BaoWei, S. B. (2018). Fertility evaluation and characteristics analysis of brown soil cultivated land in Shandong Province.
7. Li, Yinghui., Jiang, Xiaosan. & Wang, Zhenhua. (2019). Comprehensive evaluation of soil quality of agricultural land based on soil fertility and heavy metal pollution risk: A case study of Boxing County, Shandong Province [J]. *Chinese Journal of Soil Science*, 52(5): 1052-1062.