Impacts of fertilization rate on soil quality in Taihu Lake watershed

Xu Wenping, Gao Hongju*

(*College of Information and Electrical Engineering, China Agriculture University, 100083 China*)

Abstract: The application of chemical fertilizers affects the soil quality greatly. In this paper, data were collected from Taihu Lake watershed area includes soil nutrients and heavy metals in 1477 topsoil samples (PH, OM, N, P, K, AVP, AVK, CEC, Cd, Cr, Cu, Pb, Zn, AVZn, Hg, As and Se), and the grain yield per unit area and fertilization rate per unit area in successive 35 years. Effects of fertilizers on soil quality were investigated using correlation matrix, statistical analysis, PCA (Principal Components Analysis), FCM (Fuzzy c-means clustering analysis) and time series analysis. The fertilization rate of Hangzhou, Zhenjiang etc. was higher through time series analysis. Compared to 1982, the fertilization rate per unit area increased in 2000. However, the partial fertilizer productivity decreased gradually after 1995. Correlation coefficients of PH and P, OM and N are 0.62, 0.90 respectively, which imply that soil acidification and the increase of OM are partly caused by the heavy use of fertilizers. Nevertheless, the average content of OM was increased by only 4.977 g/kg with the continuous application of fertilizers from 1982 to 2000, showing that the fertilizer unabsorbed cannot retain in the topsoil. FCM divided samples into six clusters depending primarily on soil quality attributes, and values of pH are showing the tendency of soils towards acidification for all clusters. Soil quality of clusters which were mainly distributed in Hangzhou, Changzhou and Zhenjiang was bad. Therefore, the heavy use of fertilizers for long term has a negative impact on soil quality and the partial fertilizer productivity.

Keywords: fertilization rate per unit area, soil quality in Taihu Lake, PCA, FCM, time series analysis

Citation: Xu, W. P., and H. J. Gao. 2017. Impacts of fertilization rate on soil quality in Taihu Lake watershed. International Agricultural Engineering Journal, 26(3): 375–384.

1 Introduction

Taihu Lake watershed is an area with integrated economic development in both agriculture and industry. Thus, there is a direct impact of soil quality on the development of local agricultural economy. In recent decades, it has been well documented that ecosystem of Taihu Lake watershed was destroyed (Liu et al., 2017; Zhao et al., 2017; Li et al., 2017), and some thoughts and researches on ecological balance were induced by the explosion of incident of cyanobacteria in Taihu Lake. Fertilization of farmland crops in Taihu Lake watershed was much higher than the crop demand actually, as described in previous study (Xia et al., 2003). Wang et al. (2016) reported that excessive use of chemical fertilizer and massive emissions of livestock excreta were the main reasons for aggravating environmental pollution. They proposed suggestions to relieve the environmental load of nitrogen and promote scientific fertilization. It was also demonstrated that runoff carrying N and P nutrients from chemical fertilizer inputs in agricultural areas is the major contributor to eutrophication in the lake basin (Zhang et al., 2010; Chen et al., 2010; Li et al., 2000). Accordingly, excessive use of fertilizer becomes a common phenomenon, hence, results in water pollution, soil acidification and lower nitrogen use efficiency as researchers have shown in their reviews (Antonangelo et al., 2017; Hu et al., 2017; Liu et al., 2016).

Statistical analysis and correlation analysis were used basically to reveal impacts of fertilization rate on soil

Received date: 2017-07-09 **Accepted date:** 2017-08-27
* **Corresponding author: Gao Hongju,** Associate Professor, College of Information and Electrical Engineering, China Agriculture University, 100083 China. Email: hjgao@cau.edu.cn. Tel: +86-10-62738830.

quality (Yang et al., 2014; Zhang et al., 2010). It was found that more attention has been spent to the sustainable agro-ecosystem and researches are dependent on its practical significance (Stavi et al., 2011; Xia et al., 2010). Fuzzy clustering method was applied into the division of soil management zones, and then the soil can be classified according to the characteristics of soil properties and managed conveniently (Rahul et al., 2015; N. Davatgar et al., 2012; Guo et al., 2013). There are many different algorithms for clustering, fuzzy clustering and fuzzy c-means clustering are applied generally. Two types of clustering methods were used to analyze soil management zones and compare the performance of two methods in the classification management (Hot et al., 2015; Ghosh et al., 2013). There are fewer attributes integrated when studying the soil quality. Followed by it, Bansod et al. (2013) was able to perform fuzzy c-means clustering based on principal component analysis to reduce the number of soil analysis needed. Time series analysis was conducted that account better to explain the changes in attributes (Lopez-Garcia et al., 2017). In conclusion, the degradation of soil fertility is the result of long-term fertilization, and the sustainable development of soil aroused the attention of many researchers. Although, much effort of predecessors contributes to revealing the relationships with fertilization and soil fertility, while not combines the time series for fertilization rate with the soil quality and partial fertilizer productivity together.

The primary goal in this paper is to study the effects of fertilization rate on soil quality by combining these methods, such as statistical analysis, correlation analysis, principal component analysis, fuzzy c-means clustering and time series analysis. Time series analysis was employed to analyze the correlation between the fertilization rate and the partial fertilizer productivity. Effects of fertilization rate on soil quality were analyzed. The rest of paper is as follows: there is a brief introduction to the study area of Taihu Lake. The data collating is presented. Methodologies used in this paper are introduced. The results and discussion about the analysis of each part are illustrated, and conclusions are given at last.

2 Study area and data

2.1 Study area

Taihu Lake watershed (Figure 1) is one of the China's five major freshwater lakes, which is located between latitude 30°55′40″-31°32′58″ and longitude 119°5′32″-120°36′10″, covered an area of 39600 square kilometers. It is dominated by irrigated agriculture, mostly occupied by permanent plain. Taihu Lake watershed belongs to the water network in southern plain with a dense population, rapid development and developed economy.

With the sufficient sunlight resources, it is favorable to develop agriculture in Taihu Lake watershed, and thus its main crops here are rice, wheat and other cash crops.

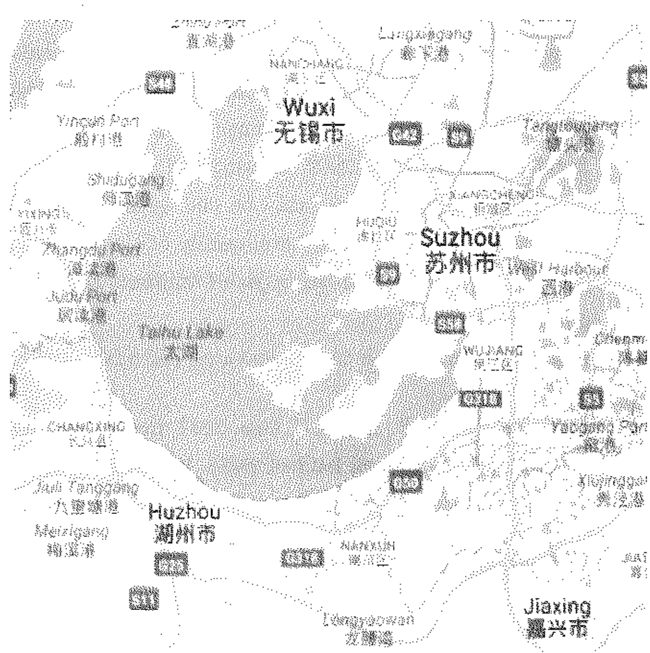

Figure 1 Boundary map of the study area

2.2 Data set introduction

There are two parts of data mainly involved in the research, they are soil nutrient and heavy metal, fertilization rate and grain yield respectively.

A total of 1477 samples with soil nutrients (N, P, K, PH, CEC, OM, AVP, AVK) and heavy metals (Cd, Cr, Cu, Pb, Zn, AVZn, Hg, As, Se) were collected from Nanjing soil database. Changes of nutrients and heavy metals in soil can be regarded as a criterion for assessing the soil quality.

The data of grain yield and fertilization rate were reorganized using equations which were given as follows before analysis. The grain yield per unit area, the fertilization rate per unit area and the partial fertilizer productivity were abbreviated as GY, FR and PFP here respectively.

GY (T/Ha)=grain yield/crop area.

FR (T/Ha)=fertilization rate/cultivated area.

PFP (kg/kg)=grain yield/fertilization rate.

Data were collated into the data of fertilization rate per unit area, grain yield per unit area and partial fertilizer productivity with different cities in successive 35 years. These data were prepared for time series analysis.

The detailed sampling sites around Taihu Lake are illustrated in Figure 2. These points are mainly distributed in Wuxi, Changzhou, Suzhou, Zhenjiang, Hangzhou, Jiaxing and Huzhou.

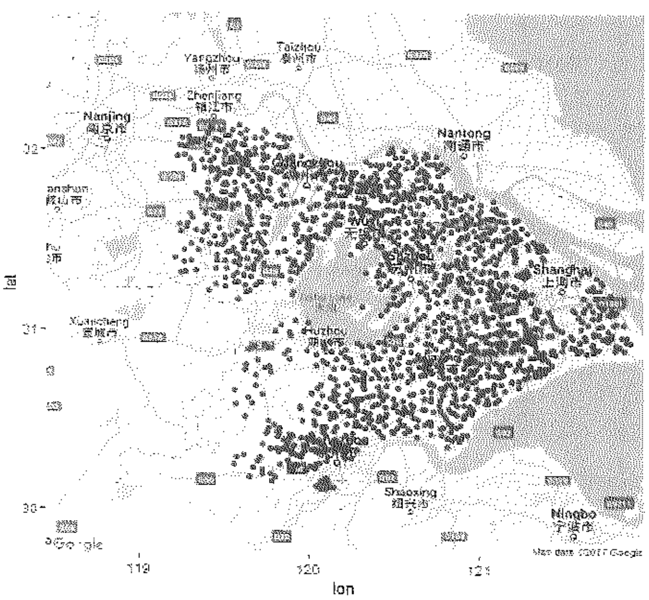

Figure 2 Distribution of sampling sites for analyzing the soil quality

3 Methodologies

3.1 Data preprocessing

3.1.1 Normalization

Because of differences in dimensions, the data should be standardized before analysis. The processed data were conformed to the standard normal distribution and the equation is given as follows:

$$X = \frac{x - \mu}{\sigma} \tag{1}$$

where, the μ means mean, and the σ stands for the standard deviation.

3.1.2 Interpolation

Data were collected manually and then there were some vacant values inevitably which cannot be deleted directly due to the less sample points. Therefore, these missing values were interpolated.

3.2 Multivariate statistical techniques

Multivariate analysis of the soil quality was performed through correlation analysis, descriptive statistical analysis and principal component analysis. Mathematical and statistical computations were performed using Microsoft Office Excel and R (i386 3.3.3).

3.2.1 Correlation analysis

Correlation analysis is a measurement of relationship between two or more variables. To verify possible relations between variables of soil, Pearson's coefficient of linear correlation was calculated between different variables (Jose et al., 2008).

3.2.2 Statistical analysis

Statistical analysis is usually categorized into three parts to analyze the data, such as centralized trend analysis, depending on mean, median and mode to analyze the data mainly. The distribution of these properties was tested using values of skewness and kurtosis. The size and concentration of the data can be observed systematically by statistical analysis.

Data of soil quality were analyzed using this method to obtain characteristics of size and concentration.

3.2.3 Principal component analysis

Principal Component Analysis was abbreviated as PCA in general. PCA is a multivariate method mainly used for reducing data dimensions. The contribution of PCs (principle components) to the preservation of original information is measured by its variance. The complexity of multi-index analysis will be declined greatly and the reliability of system operation will be improved.

The possible origins of nutrients and heavy metals in samples were investigated by PCA. The PCs transformed from the original variables were extracted according to eigenvalues >1.0. Due to different dimensions of each variable, first of all, normalization was done prior to PCA. The data of soil quality were analyzed using PCA to reduce dimensions and prepare for the clustering.

3.3 Fuzzy c-means clustering analysis

Fuzzy c-means was abbreviated as FCM in general. FCM based on PCA was used to classify the selected PCs and aim at identifying natural clusters.

FCM is completed by R software, and the only shortcoming is that the number of clusters as a parameter which is set in advance. Selecting the appropriate number of clusters plays a key role in the result of clustering analysis. Thus, one critical issue is to find the best NC (number of clusters) based on some indices, which are PE (the partition entropy index), PC (the partition coefficient index), MPC (the modified partition coefficient index), SIL (the silhouette index) and XB (Xie and Beni index) respectively. NC is an optimal value only when XB and PE are minimum values and SIL, PC and MPC are maximum values respectively. The data obtained by PCA were clustered to divide the soil into several parts relying on the characteristics of soil quality, and then analyze each part of soil.

3.4 Time series analysis

Time series analysis is a statistical method of dynamic data processing, and focuses on studying changes of data sequences. The time series diagrams were plotted for analyzing changes of fertilization rate per unit area and partial fertilizer productivity following years.

4 Results and discussion

4.1 Fertilizer based on spatial and temporal analysis

Comparisons of fertilization rates in different cities from 1980 to 2014 were illustrated in Figure 3. It is simply appeared that in Hangzhou, Zhenjiang and Changzhou, the fertilization rates per unit area are higher than those of other cities.

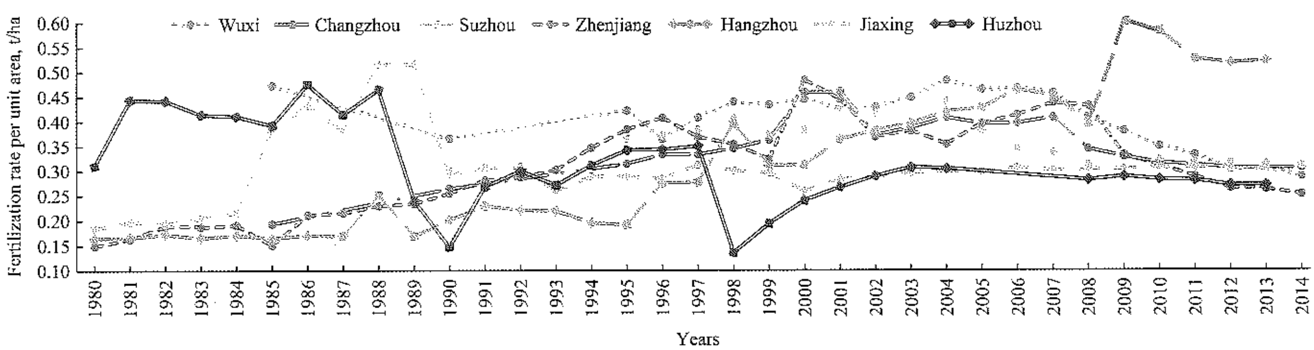

Figure 3 Comparisons of fertilization rates per unit area in different cities

Taking Jiangyin as an example, the time series of fertilization rate per unit area and partial fertilizer productivity were shown in Figures 4 and 5, respectively. Compared to 1982, the fertilization rate per unit area is largely increased in 2000. Furthermore, the trend of partial fertilizer productivity is steeper from 1991 to 1995 corresponding to a remarkably increase, and then began to decline as the fertilization rate increased progressively after 1995.

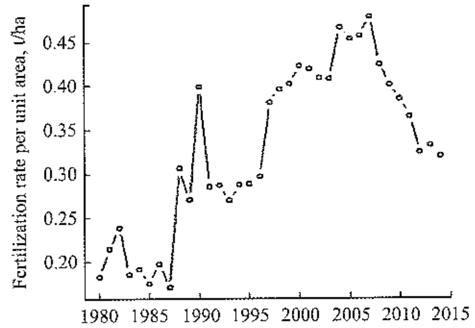

Figure 4 Time series of fertilization rate per unit area

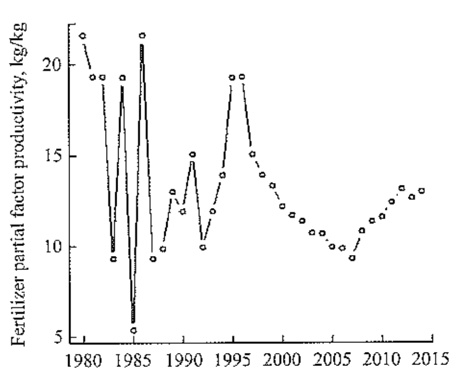

Figure 5 Time series of partial fertilizer productivity

It has been fully proved that the excessive fertilizer was applied in Taihu Lake area, and some parts of unabsorbed fertilizer were washed into the deeper horizons of soil. As observed above, the partial fertilizer productivity decreased after 1995, and the absorbed rate of crops also decreased. In short, there is a negative impact of fertilization rate on agricultural production.

4.2 Results for soil quality

4.2.1 Correlation analysis

For better visualization, correlation coefficient matrices were shown in Figures 6, 7 and 8 respectively. Digits in correlation coefficient matrix represent Pearson correlation coefficient. In addition, different colors in matrix are corresponding to the sizes of coefficient, and the color changes from deep to shallow representing the correlation changes from strong to weak.

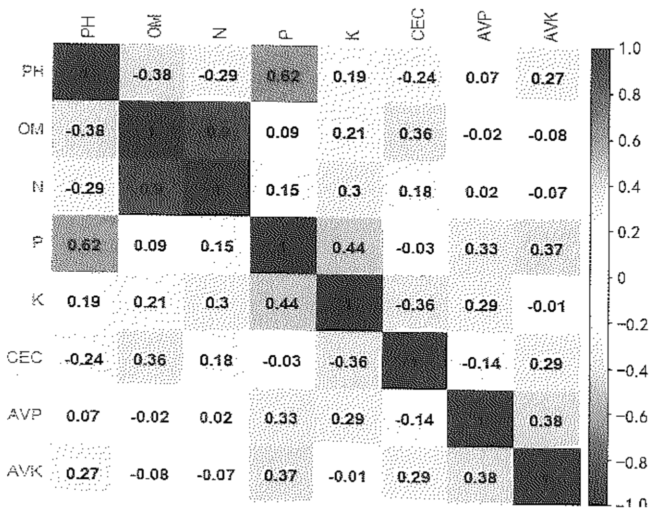

Figure 6 Correlation coefficient matrix of soil nutrients

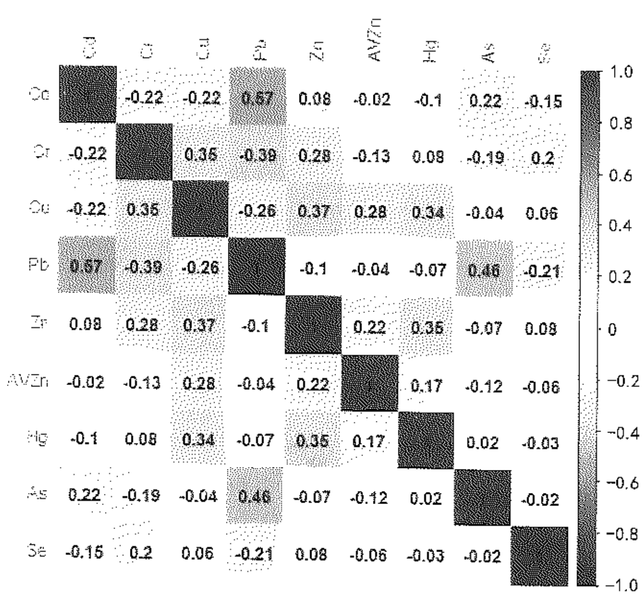

Figure 7 Correlation coefficient matrix of heavy metals

It can be seen that correlation coefficients between PH and P, OM and N are 0.62, 0.9 respectively, with strong correlations. The correlation between other properties is weak, particularly the correlation between AVK and K. A possible explanation is that the increase in OM and the soil acidification are probably caused by the application of fertilizers extensively. Furthermore, the correlation between AVK and K is –0.01, with relatively weak negative correlation, which implies that content of AVK in soil is restricted by the soil fertility instead of the fertilization rate.

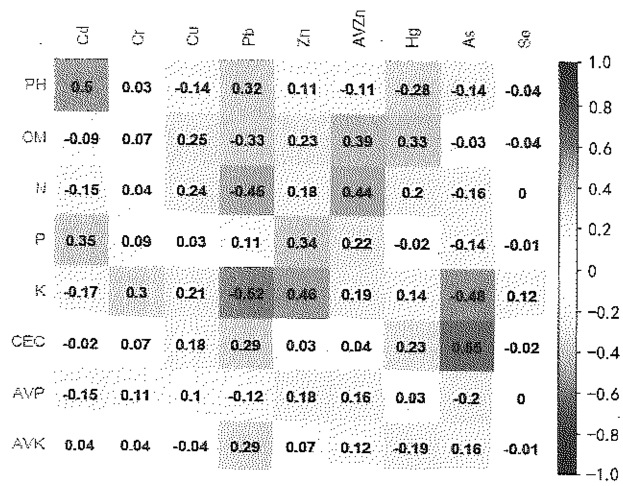

Figure 8 Correlation coefficient matrix between nutrients and heavy metals

As illustrated in Figure 7, correlations between variables in heavy metals are weak generally, indicating the diversity of resources of heavy metals. But correlation coefficients between Pb and As, Pb and Cd are 0.46, 0.57 respectively, with a relatively strong positive correlation. It can be concluded that resource of Pb is similar with As and Cd with a great possibility.

A common belief regarding CEC and As is that their contents in soil can reflect the status of soil fertility. As observed above, there is a strong positive correlation between As and CEC with the correlation coefficient of 0.55. The correlation between Cd and PH is also strong, with a coefficient of 0.5. Accordingly, the content of As affects the soil fertility with the accumulation of As in soil, and the content of Cd tends to favor the soil acidification.

Scatter plots of two variables were shown in Figures 9 and 10. It can be seen that, the regional distribution of some heavy metals presents a trend of polarization. Most of sampling sites have a relatively high content of Pb, with an uneven distribution, while the regional distribution of Cd is more uniform. Soil pollution caused by the use of fertilizers is non-point source pollution, while the soil pollution caused by heavy metals is point source pollution. This means that fertilizer is not the only factor that causes the heavy metal pollution.

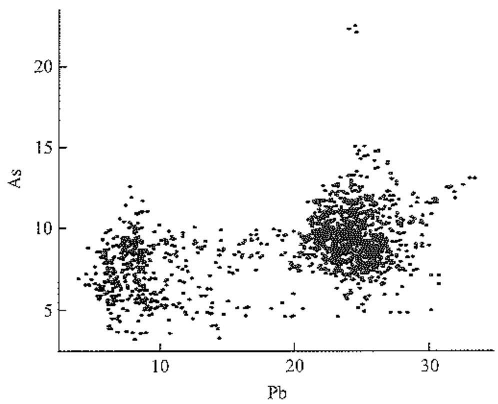

Figure 9 Scatter plot with Pb and As

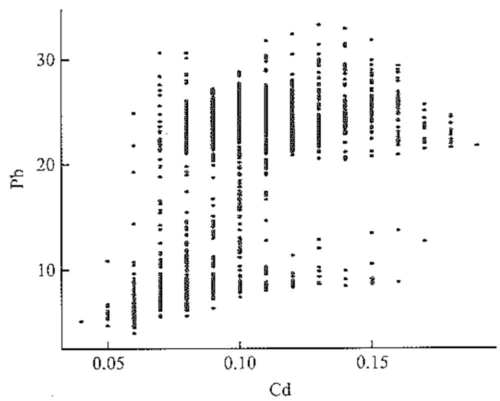

Figure 10 Scatter plot with Pb and Cd

4.2.2 Statistical analysis

The most frequently used parameters (Mean, Min, Max, Var, CV), which can represent the size and concentration of data, were calculated. Where, CV refers to the coefficient of variation. It is well known that OM is a crucial component to maintain the soil fertility and provide nutrients for crops. In brief, the content of OM is regarded as an important factor to assess the soil fertility. Results of statistical analysis were summarized in Tables 1 and 2, and contents of CEC in 1982 are vacant. For analyzing intuitively, Figures 11 and 12 were shown. Compared to 1982, PH has a downward trend in 2000 facing the risk of acidification. As can be seen from Table 1, compared to 1982, OM and N have an upward trend, and this is a general phenomenon with the continuous fertilization year after year. But the average content of OM increases only by 4.977 (g/kg), so that ultimately the small amount of fertilizers was trapped in soil and most of them had been washed into the deep soil. Although acknowledged, fertilization exactly compensates nutrient requirements. Nonetheless, there is a tendency of soil towards acidification likely due to the long-term fertilization.

Table 1 Statistical analysis of nutrients in 1982 and 2000

Nutrients	Year	Mean	Min	Max	Var	CV
PH	1982	6.74	5.44	8.13	0.305	0.082
	2000	6.215	4.96	8.39	0.543	0.119
OM, g/kg	1982	26.242	13.88	45.21	38.447	0.236
	2000	31.219	16.35	42.5	28.318	0.17
N, g/kg	1982	1.555	0.8	2.66	0.107	0.211
	2000	1.769	0.82	2.83	0.132	0.205
P, g/kg	1982	0.704	0.17	1.63	0.074	0.387
	2000	0.654	0.33	1.35	0.024	0.235
AVP, mg/kg	1982	11.382	3.53	88.96	47.451	0.605
	2000	10.677	4.07	31.21	21.779	0.437
AVK, mg/kg	1982	90.134	34.29	178.57	552.823	0.261
	2000	85.701	45.69	152.66	287.892	0.198
CEC, cmol/kg	1982					
	2000	16.271	7.85	23.77	7.974	0.174

Table 2 Statistical analysis of heavy metals in 2000

Heavy metals, g/kg	Mean	Min	Max	var	CV
Cd	0.102	0.04	0.19	0.0007	0.268
Cr	64.028	29.59	113.13	111.874	0.165
Cu	26.307	16.35	93.99	34.745	0.224
Pb	19.602	3.98	33.45	57.546	0.387
Zn	102.993	51.26	200.09	380.748	0.189
AVZn	1.905	0.76	11.31	1.073	0.544
Hg	0.207	0.05	1.2	0.017	0.638
As	8.885	3.19	22.58	4.36	0.235
Se	0.891	0.16	57.26	30.18	6.163

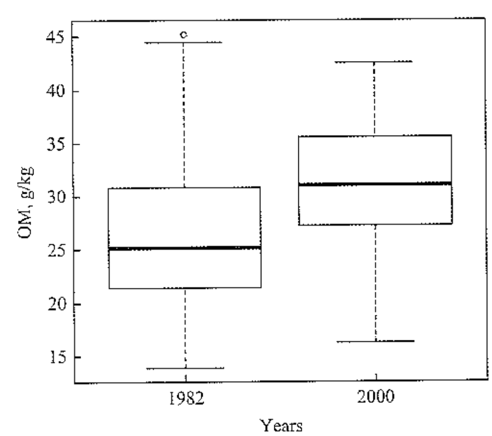

Figure 11 Comparison of OM between 1982 and 2000

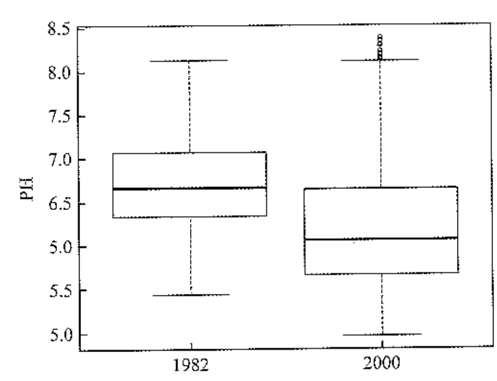

Figure 12 Comparison of PH between 1982 and 2000

Table 2 shows that variances of Cr, Cu and Zn are significantly different from variances of other heavy metals. Variances of Cr, Cu and Zn are much bigger, as this means that distributions of Cr, Cu and Zn are uneven especially. The variance of Cd is lower, and then its distribution is even relatively. The mean of heavy metals varies greatly. For instance, the average content of Cd is 0.102 g/kg; whereas Zn is 102.993 g/kg. It is showed that heavy metals in different areas are uneven and their resources are varied from each other. These results are consistent with the results of correlation coefficient analysis.

In soil science studies, it is a common sense that CV represents the spatial variability of soil properties. CV is usually used for comparing the discrepancy of different dimensions or different scales of measurement. CV of Se has reached 6.163, showing a strong variability, and CV of AVZn and Hg are 0.544, 0.638 respectively, changing obviously. Accordingly, there is an obviously different from each other in different regions with respect to soil characteristics, though they are all around Taihu Lake area.

4.2.3 Results for PCA

PCA was performed to aggregate the variability in these properties and was prepared for FCM. More specifically, PCA was adopted to characterize the type of soil around Taihu Lake, while FCM was used for grouping the soil.

Each PC (Principal Component) is a linear expression of all original variables, and coefficients of the expression are corresponding to eigenvectors. It is possible to infer the variable of original data mainly expressed by PC according to the coefficients of expression which has the larger coefficient.

Six PCs were extracted as listed in Table 3, and the cliff of this analysis was presented in Figure 13. Where, Comp.1 denotes the first PC. As shown in Table 3, the first PC and the sixth PC are accounting for 46.60% and 2.82% respectively, with a large difference. These six PCs are accounting for 93.58% of original information.

4.2.4 Fuzzy c-means clustering

The data, which are organized into PC values after PCA, are used for FCM. FCM based on PCA is completed by R software to analyze characteristics of soil.

Table 3 Information about PCs

PCs	Eigenvalue	Component loading, %	Cumulative loading, %
Comp.1	5.09	46.60	46.60
Comp.2	3.88	24.01	70.60
Comp.3	2.86	9.53	80.14
Comp.4	1.99	5.91	86.05
Comp.5	1.78	4.72	90.76
Comp.6	1.20	2.82	93.58

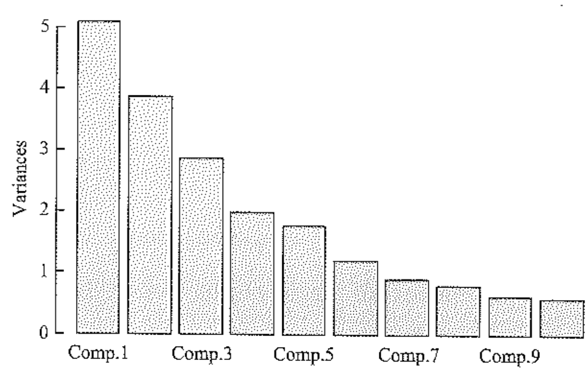

Figure 13 Cliff of the PCA

The first thing to do is selecting the optimal NC (number of clusters). Additionally, different indices followed by different clusters were listed in Table 4, and the line chart was also given in Figure 14 to find the optimal value more intuitively. When NC is six, XB is a minimum value and MPC and SIL are maximum values. Selecting six as the optimal value with comprehensive consideration, despite PE is not a minimum value and PC is not a maximum value. At last, the spatial distribution of each cluster was shown in Figure 15, and different colors represent different clusters.

Table 4 Different indicators followed by different NC

NC	PE	PC	MPC	SIL	XB
2	0.5964639	0.5902092	0.1804183	0.3541781	0.9211374
3	0.9408203	0.4436973	0.1655459	0.393596	0.9293091
4	1.188661	0.3617968	0.1490624	0.3424196	0.876639
5	1.324459	0.3399689	0.1749611	0.3819176	0.5167741
6	1.447187	0.3180616	0.1816739	0.3837727	0.4628576
7	1.55341	0.2983572	0.1814168	0.3641149	0.610551
8	1.690754	0.2656551	0.1607487	0.3232503	0.5451866

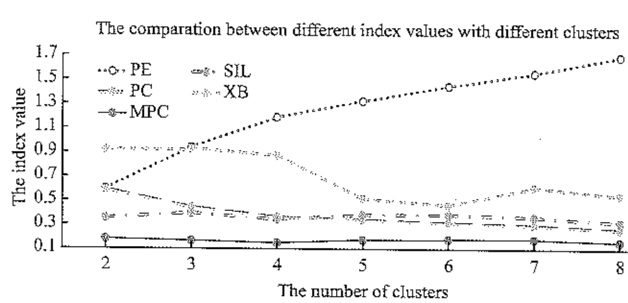

Figure 14 Indicators with different clusters

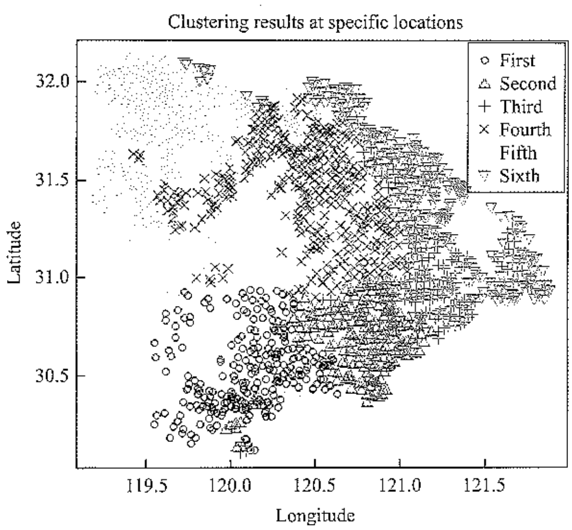

Figure 15 Distribution of each cluster

Attribute values of each cluster were detailed in Tables 5, 6 and 7 respectively. The soil status of each cluster was evaluated in terms of Table 8, which was performed according to the unified system of whole country. Results were given in Table 9. Accordingly, the order of soil quality of six types is 3, 6, 2, 4, 5, and 1, from good to bad. The soil quality of clusters which are mainly distributed in Hangzhou, Changzhou and Zhenjiang is poor. A further scrutiny of Figure 3 shows that the soil quality is poor in areas where more fertilizer has been applied.

As demonstrated above, heavy metals have a strong spatial variability, and their recourses are diverse. The soil acidification and the decrease in partial fertilizer productivity are caused by excessive fertilization around Taihu Lake. In conclusion, there is a serious impact of excessive fertilization on grain yield in the long run.

Table 5 Nutrients for six clusters in 1982

Cluster	PH	OM	N	P	AVP	AVK
first	6.3032	28.3946	1.6052	0.5214	0.0184	0.0718
second	6.5563	31.5525	1.8191	0.7102	0.0104	0.1090
third	6.9904	31.5777	1.8409	0.7778	0.0160	0.1147
fourth	6.5181	27.0051	1.6234	0.7254	0.0079	0.0811
fifth	6.6064	19.7762	1.2597	0.5078	0.0077	0.0836
sixth	7.6555	22.8296	1.3518	1.0557	0.0113	0.0975

Table 6 Nutrients for six clusters in 2000

Class	PH	OM	N	P	K	CEC	AVP	AVK
first	5.8220	33.1322	1.9700	0.5635	19.9481	13.0229	0.0103	0.0680
second	5.7698	35.3835	2.0588	0.6808	19.0957	17.3801	0.0137	0.0870
third	6.6525	34.6310	2.0912	0.8493	18.3840	17.0112	0.0142	0.1080
fourth	5.7865	32.7103	1.7507	0.5987	14.7163	19.1256	0.0084	0.0845
fifth	6.1697	26.1439	1.4273	0.5198	13.7752	15.7302	0.0101	0.0858
sixth	7.3551	27.5836	1.5466	0.8345	18.2631	14.8704	0.0100	0.0894

Table 7 Heavy metals for six clusters in 2000

class	Cd	Cr	Cu	Pb	Zn	AVZn	Hg	As	Se
first	0.0802	62.9400	25.9454	9.9087	104.4396	2.0796	0.2388	6.5982	1.0902
second	0.0776	80.0936	33.5320	8.8556	118.3215	2.0113	0.2907	8.0972	4.1014
third	0.1085	58.6890	25.5779	22.0846	107.1655	3.2326	0.1637	8.4638	0.2355
fourth	0.1037	58.7296	25.9502	24.1195	99.1198	1.7656	0.2547	11.1069	0.2436
fifth	0.1032	62.7426	24.0809	24.3359	86.3377	1.3323	0.1289	9.2283	0.2043
sixth	0.1337	64.5314	24.4299	24.7260	111.9929	1.6265	0.1619	8.6579	0.2259

Table 8 Classification

Classification	OM, g/kg	N, g/kg	AVP, g/kg	AVK, g/kg	AVN, g/kg	PH
1, Extremely high	>40	>2	>0.04	>0.2	>0.15	<4.5
2, Very high	30-40	1.5-2	0.02-0.04	0.15-0.2	0.12-0.15	4.51-5.50
3, High	20-30	1-1.5	0.01-0.02	0.1-0.15	0.09-0.12	5.51-6.50
4, Middle	10-20	0.75-1	0.005-0.01	0.05-0.1	0.06-0.09	6.51-7.50
5, Low	6-10	0.5-0.75	0.003-0.005	0.03-0.05	0.03-0.06	7.51-8.50
6, Very low	<6	<0.5	<0.003	<0.03	<0.03	>8.50

Table 9 Classification of each cluster

Category	PH	OM	N	AVP	AVK
1	Slightly acid	Very high	Very high	High	High
2	Slightly acid	Very high	Extremely high	High	Middle
3	Neutral	Very high	Extremely high	High	High
4	Slightly acid	Very high	Very high	Middle	Middle
5	Slightly acid	High	High	High	Middle
6	Neutral	High	Very high	High	Middle

5 Conclusions

The application of fertilizers for long term has caused serious impacts on soil quality. The fertilization rate around Taihu Lake watershed is considered as a clue, impacts of fertilization rate on soil quality were analyzed through correlation analysis, statistical analysis, fuzzy c-means clustering, time series analysis, etc.

Through time series analysis, it is reported that excessive fertilization rate leads to the decrease in partial fertilizer productivity. This has been a common phenomenon and affects the absorbed rate for crops.

There is a noticeable trend of soil acidification through statistical analysis. The average content of OM increased inconspicuously with the continuous application of fertilizers, showing that the part of fertilizer can remain in the topsoil and most of them can be washed into the deep soil substantially. Correlation coefficients between PH and P, OM and N are 0.62 and 0.9 respectively with the strong positive correlation. It is revealed that soil acidification is caused by the application of fertilizers. After that, FCM was employed to divide the soil into six types, and the lager the fertilization rate is, the worse the soil quality would be.

In conclusion, excessive use of fertilizers for long term causes soil acidification and the decrease in soil fertility. If it is failed to take real-time measures to control the fertilization rate, ecosystem in Taihu Lake will be caught in a vicious circle with respect to the fertilization rate and soil quality. As such, it is imperative to fertilize moderately, which can protect the soil against natural tendency to acidification, guide the agro-production and also maintain the agricultural sustainable development.

Acknowledgments

We acknowledge the sharing of the soil data from China Soil Database, and are grateful for the support from National Natural Science Foundation of China through project 31371531.

[References]

[1] Antonangelo, J. A., J. F. Neto, C. A. C. Crusciol, and L. R. F. Alleoni. 2017. Lime and calcium-magnesium silicate in the ionic speciation of an Oxisol. *Scientia Agricola*, 74(4): 317–333.

[2] Bansod, B. S., and O. P. Pandey. 2013. An application of PCA and fuzzy C-means to delineate management zones and variability analysis of soil. *Eurasian Soil Science,* 46(5): 556–564.

[3] Chen, D. J., J. Lu, and Y. N. Shen. 2010. Artificial neural network modelling of concentrations of nitrogen, phosphorus and dissolved oxygen in a non-point source polluted river in Zhejiang Province, Southeast China. *Hydrological Processes*, 24(3): 290–299.

[4] Davatgar, N., M. R. Neishabouri, and A. R. Sepaskhah. 2012. Delineation of site specific nutrient management zones for a paddy cultivated area based on soil fertility using fuzzy clustering. *Geoderma*, 173-174: 111–118.

[5] Elma, H., and V. P. Bugarin. 2015. Soil data clustering by using K-means and fuzzy K-means algorithm. *23rd Telecommunications forum TELFOR 2015*, 2015, *Belgrade, Serbia. 24-26 Nov.* 890–893.

[6] Ghosh, S., and S. K. Dubey. 2013. Comparative analysis of K-means and fuzzy C-means algorithms. *International Journal of Advanced Computer Science and Application*, 4(4): 35–39.

[7] Guo, Y., Y. F. Tian, H. H. Wu, and Z. Shi. 2013. Study on soil partitioning of farmland based on multi-source data and fuzzy k-means. *Acta Pedologica Sinica*, 50(3): 441–447.

[8] Hu, M., Y. S. Xiang, Z. Zhang, R. H. Cong, F. Y. Huang, J. Q. Zhang, L. L. Shang, and J. W. Lu. 2017. Variation characteristics of farmland soil pH in the past 30 years of Enshi Autonomous Prefecture. *Chinese Journal of Applied*

Ecology, 28(4): 1289–1297.

[9] Janssen, A.B.G., V. C. L. de Jager, J. H. Janse, X. Z. Kong, S. Liu, Q. H. Ye, and W. M. Mooij. 2017. Spatial identification of critical nutrient loads of large shallow lakes: Implications for Lake Taihu (China). *Water Research*, 119: 276–287.

[10] Li, H.B., P. Xing, and Q. L. Wu. 2017. Genus-specific relationships between the phytoplankton and bacterioplankton communities in Lake Taihu, China. Hydrobiologia, 795(1): 1–14.

[11] Lopez-Garcia, P., M. D. Gelado-Caballero, C. Collado-Sanchez, and J. J. Hernandez-Brito. 2017. Solubility of aerosol trace elements: Sources and deposition fluxes in the Canary Region. *Atmospheric Environment*, 148: 167–174.

[12] Liu, X., S. S. Xu, J. W. Zhang, Y. F. Ding, G. H. Li, S. H. Wang, Z. H. Liu, S. Tang, C. Q. Ding, and L. Chen. 2016. Effect of continuous reduction of nitrogen application to a rice-wheat rotation system in the middle-lower Yangtze River region (2013-2015). *Field Crops Research*, 196: 348–356.

[13] Molin, J. P., and C. N. de Castro. 2008. Establishing management zones using soil electrical conductivity and other soil properties by the fuzzy clustering technique. *Sci.Agric.* (Piracicaba, Braz), 65(6): 567–573.

[14] Stavi, I., R. Lal, and L. B. Owens. 2011. On-farm effects of no-till versus occasional tillage on soil quality and crop yields in eastern Ohio. *Agronomy for Sustainable Development*, 31(3): 475–482

[15] Taoufik, G., I. Khouni, and A. Ghrabi. 2017. Assessment of physico-chemical and microbiological surface water quality using multivariate statistical techniques: a case study of the Wadi EI-Bey River, Tunisia. *Arabian Journal of Geosciences*, 10(7): 181.

[16] Tripathi, R., A.K. Nayak, M. Shahid, B. Lal, P. Gautama, R. Raja, S. Mohanty, A. Kumar, B. B. Panda, and R. N. Sahoo. 2015. Delineation of soil management zones for a rice cultivated area in eastern India using fuzzy clustering. *Catena*, 133: 128–136.

[17] Wang, D., Y. H. Wang, H. Yang, Z. C. Cai, and X. Y. Yan. 2016. Nitrogen flow features of crop-livestock breeding system in Taihu Lake basin. *Research of Environmental Sciences*, 29(3): 457–464.

[18] Xia, L. Z., and L. Z. Yang. 2003. Research and control of non - point source pollution in Taihu Lake watershed. *Resources and Environment in the Yangtze Basin*, 12(1): 45–49.

[19] Xia, Q. F. 2010. Effects of Chemical Fertilizer Application on Soil Ecosystem and Counter measures. *Journal of Guangdong Agricultural Sciences*, (2010) 9.

[20] Yang., Y. F., T. Shao, M. Y. Lv, S.Y. Sun, J. Y. Zhang, and P. J. Zhang. 2014. Characteristics of soil nutrient and heavy metal pollution in wetlands under different land use patterns in Longwo Lake of Wuhu City, Anhui Province, China. *Chinese Journal of Ecology*, 33(5): 1312–1318.

[21] Zhang, B. Y., T. L. Chen, and B. Wang. 2010. Effects of Long-term uses of chemical fertilizers on soil quality. *Chinese Agricultural Science Bulletin*, 26(11): 182–187.

[22] Zhang, Q. L., Y. X. Chen, G. Jilani, I. H. Shamsi, and Q. G. Yu. 2010. Model AVSWAT apropos of simulating non-point source pollution in Taihu Lake basin. *Journal of Hazardous Materials*, 174(1-3): 824–830.

[23] Zhao, Z. H., Y. Jiang, Q. Y. Li, Y. J. Cai, H. B. Yin, L. Zhang, and J. Zhang. 2017. Spatial correlation analysis of polycyclic aromatic hydrocarbons (PAHs) and organochlorine pesticides (OCPs) in sediments between Taihu Lake and its tributary rivers. *Ecotoxicology and Environmental Safety*, 142: 117–128.