Threat assessment of Tianjin Qilihai wetland from human activities based on principal component analysis

Qiang Chu1 Baocui Liang1 Shitao Peng1 Tianli Zheng1 Guoguang Li2 Kailei Zhang1 Liankun Liu1 and Jianna Jia1*

1Key Laboratory of Environmental Protection Technology on Water Transport, Ministry of Transport, Tianjin research institute for water transport engineering, M.O.T., Tianjin, 300456, China

2Environmental Criteria, Ministry of Education, and Tianjin Key Laboratory of Environmental Remediation and Pollution Control, College of Environmental Science and Engineering, Nankai University, Tianjin, 300071, China

*Corresponding author’s e-mail: pengshitao@tiwte.ac.cn

Abstract. Tianjin Qilihai wetland is an important wetland nature reserve in north China, which has important ecological value. Over the last few decades, the Qilihai wetland has been developed by human on a large scale, and the natural ecosystem has been seriously disturbed by human activities. In this study, principal component analysis was conducted on 7 indicators that may have potential impacts on the Qilihai wetland from 2013 to 2017 to assess the threat of human activities to the wetland area. The results show that from 2013 to 2017, the impact of human activities on the Qilihai wetland is increasing year by year, among which the industrial development, transportation and the increase of human activities around the wetland may be the main threat for the wetland.

1. Introduction

Wetland ecosystem is one of the most important ecosystems on earth [1], which plays an extremely important role in water quality purification and global carbon fixation. Tianjin Qilihai wetland is an important wetland nature reserve in China, located in the southwest of Ninghe County, Tianjin, China. It is one of the many ancient lagoon wetlands left in the Tianjin Plain by the process of regression since the Holocene [2]. Qilihai wetland has rich natural resources and ecological value. In the wetland, the geographical components of plants are diverse, and the Pan-tropical species, North temperate species and Old World temperate species are dominant. There are about 69 families and 290 species of plants, and herbaceous plants are dominant. Wetland area is rich in wildlife species, among them because the wetland is located in the East Asian - Australasian bird migration area, is particularly rich in wetland bird resources. The survey in the area found about 184 bird species, belonging to 16 orders, 39 families, 88 genera, including national class I protect 3 species, class II protect 25 species [3]. In addition to animal and plant resources, as one of the three famous ancient coastal relics in the world, Tianjin Qilihai wetland also has three different kinds of geological landscapes in the world, including ancient coastal relics, ancient lagoon relics and paleontological relics.

Due to the rich natural resources, the Qilihai wetland has been developed and utilized on a large scale in the past few decades. It is seriously disturbed by human activities and the natural ecosystem is degraded. The major problems of natural wetlands include shrinking area, simplification of landscape...
types, destruction of plant and animal community structure and deterioration of water quality. Human activities are the main factors influencing the wetland ecosystem. Between 1982 and 2005, land reclamation and aquaculture are widespread in the area of the wetland, the wetland in large areas of natural waters and farmland is converted, the breeding significantly reduces natural wetland, artificial wetlands are on the increase, making wetland biodiversity decreased significantly at the same time also reduces the richness of wetland landscape. Based on the pressure-State-response (PSR) model, the ecological vulnerability analysis of the Qilihai wetland shows that most of the wetland areas in Qilihai belong to the moderate vulnerable areas [4]. Therefore, the ecological protection and restoration work of the Qilihai wetland is also gradually advancing.

As for the human threats on the Qilihai wetland, it is an important part of the research work to analyze the main threat factors and evaluate the annual changes of the threats. In the study of human threat analysis and assessment, principal component analysis, constructed by Pearson and further developed by Hotellin, is widely used in various natural and social disciplines [5]. Principal component analysis transforms the original data into a new coordinate system by means of linear transformation [6], and makes comparative analysis of different principal components by means of variance. In this study, principal component analysis was used to analyze the human activity factors that may pose potential threats to the Qilihai wetland, so as to grasp the influence characteristics of human activities on the Qilihai wetland in different years.

2. Methods

2.1 Human threat factor selection

As Tianjin Qilihai wetland is located in Ninghe District of Tianjin, the threat of human activities mainly come from Ninghe region. Therefore, in this study we use national economic and social development statistics dataset of Ninghe District, Tianjin in the five years from 2013 to 2017.

The Qilihai wetland reserve contains core area, buffer area and experimental area. The core areas are mainly reed fields and water areas, and human activities are prohibited. The buffer area and the experimental area are distributed with rural land, cultivated land and land for agricultural facilities. In addition, a certain number of artificial fish ponds are distributed in the experimental area, and highways are distributed around the wetland.

Therefore, seven indicators with the greatest impact on wetland are selected, namely, annual gross product, agricultural gross product, freshwater aquaculture area, crop planting area, total industrial output value above designated scale, total highway mileage and population.

2.2 Principal component analysis

For the selected factors, principal component analysis is used to analyze human activity threat. The principle of this method is to replace the original indexes with a new set of unrelated comprehensive indexes with the idea of dimension reduction. The mathematical method is to make linear combination of the original N indexes to form a new comprehensive index.

The mathematical model of principal component analysis is as follows:

\[
F_1 = a_{11}Z_{X1} + a_{21}Z_{X2} + \ldots + a_{p1}Z_{Xn} \quad (1)
\]

\[
F_2 = a_{12}Z_{X1} + a_{22}Z_{X2} + \ldots + a_{p2}Z_{Xn} \quad (2)
\]

\[
\ldots
\]

\[
F_m = a_{1m}Z_{X1} + a_{2m}Z_{X2} + \ldots + a_{pm}Z_{Xn} \quad (3)
\]

Where \(a_{ij}, a_{2i}, \ldots, a_{pi} \quad (i=1, \ldots, m)\) is the eigenvectors corresponding to the eigenvalue of covariance matrix \(\Sigma\) of \(X, Z_{X1}, Z_{X2}, \ldots, Z_{Xn}\) is the value formed after the original variable has been standardized.

After standardizing the original variables, m principal components can be formed through the above linear combination method. Then the variance of different principal components is compared. The larger variance is, the more information it contains. After linear transformation, the variance of \(F_1\) should be the largest of all linear combinations, so \(F_1\) is called the first principal component. If the first principal component is insufficient to represent the information of the original N indicators, \(F_2\) is selected, that
is, the first and second linear combination is considered at the same time. In order to effectively reflect the original information, the existing information of F1 no longer appears in F2, that is, Cov (F1, F2) = 0, and F2 is the second principal component, and also the third, the fourth, etc.

When extracting principal components, the eigenvalue can be regarded as an indicator of the influence force of principal components to some extent. If the eigenvalue is less than 1, it indicates that the explanatory power of the principal component is less than the average explanatory power by directly using a original variable. Therefore, the extraction principle of the number of principal components is to adopt the first m principal components whose corresponding eigenvalue is greater than 1.

3. Results and discussions

Set 7 human activity threat factors as original variables from X1 to X7. X1 is the annual gross product, X2 is the agricultural gross product, X3 is the freshwater farming area, X4 is the crop planting area, X5 is the total industrial output value above designated scale, X6 is the total highway mileage, and X7 is the population. The statistical values of the seven variables from 2013 to 2017 are shown in Table 1.

The data is standardized and the correlation coefficient matrix table (Table 2) and variance decomposition principal component extraction table (Table 3) are obtained. After principal component analysis, according to the results of principal component extraction and extraction principle, the first two principal components of F value is greater than 1, the first eigenvalue cumulative percentage is greater than 95.34%. So the first two principal component can reflect most of the information, and we get their initial factor loading matrix (Table 4). The initial factor loading matrix shows that the industrial output value above designated scale, total highway mileage and population are the three factors with the highest load on the first principal component, and the first principal component mainly reflects the information of these indicators. Annual gross product, agricultural gross domestic product and freshwater aquaculture area are the three factors with high load on the second principal component, and the second principal component mainly reflects the information of these three indicators.

The eigenvectors A1 of the first principal component and A2 of the second principal component were obtained by using the initial eigenvalues and initial factor loading matrix. Using the eigenvectors of the two principal components and the standardized value of the 7 original variables (Table 5), the expressions of the two principal components are obtained as follows:

\[ F_1 = 0.30ZX_1 - 0.27ZX_2 + 0.30ZX_3 + 0.41ZX_4 - 0.43ZX_5 - 0.46ZX_6 + 0.42ZX_7 \]  (4)
\[ F_2 = 0.50ZX_1 + 0.53ZX_2 + 0.47ZX_3 - 0.34ZX_4 + 0.27ZX_5 + 0.00ZX_6 + 0.25ZX_7 \]  (5)

After calculating the ratio of the eigenvalue of each principal component to the sum of the eigenvalue of all principal components extracted, then taking this ratio as the weight coefficient of the corresponding principal component to obtain the comprehensive value of the principal component:

\[ F = \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2} \]  (6)
\[ \lambda_1 = 4.225, \ \lambda_2 = 2.449, \ \text{and} \ \lambda_1 + \lambda_2 = 6.674, \ \text{then the comprehensive value of the principal component can be obtained:} \]
\[ F = \frac{4.225 \times F_1 + 2.449 \times F_2}{6.674} \]  (7)

According to the principal component analysis results from 2013 to 2017, the sorted comprehensive value of the principal component can be obtained (Table 6).
Table 2. The correlation coefficient matrix

|       | X1     | X2     | X3     | X4     | X5     | X6     | X7     |
|-------|--------|--------|--------|--------|--------|--------|--------|
| X1    | 1.000  | 0.297  | 0.928  | 0.097  | -0.241 | -0.553 | 0.877  |
| X2    | 0.297  | 1.000  | 0.252  | -0.917 | 0.840  | 0.537  | -0.156 |
| X3    | 0.928  | 0.252  | 1.000  | 0.142  | -0.195 | -0.677 | 0.751  |
| X4    | 0.097  | -0.917 | 0.142  | 1.000  | -0.952 | -0.813 | 0.504  |
| X5    | -0.241 | 0.840  | -0.195 | -0.952 | 1.000  | 0.774  | -0.652 |
| X6    | -0.553 | 0.537  | -0.677 | -0.813 | 0.774  | 1.000  | -0.729 |
| X7    | 0.877  | -0.156 | 0.751  | 0.504  | -0.652 | -0.729 | 1.000  |

Table 3. Variance decomposition principal component extraction

| Principal component | Total No. | Percentage of variation (%) | Cumulative percentage (%) | Total No. | Percentage of variation (%) | Cumulative percentage (%) |
|---------------------|-----------|-----------------------------|---------------------------|-----------|-----------------------------|---------------------------|
| 1                   | 4.225     | 60.358                      | 60.358                    | 4.225     | 60.358                      | 60.358                    |
| 2                   | 2.449     | 34.986                      | 95.344                    | 2.449     | 34.986                      | 95.344                    |
| 3                   | 0.311     | 4.437                       | 99.782                    |           |                             |                           |
| 4                   | 0.015     | 0.218                       | 100.000                   |           |                             |                           |
| 5                   | 0.000     | 0.000                       | 100.000                   |           |                             |                           |
| 6                   | 0.000     | 0.000                       | 100.000                   |           |                             |                           |
| 7                   | 0.000     | 0.000                       | 100.000                   |           |                             |                           |

Table 4. Initial factor loading matrix

| Principal component | Principal component 1 | Principal component 2 |
|---------------------|-----------------------|-----------------------|
| X1                  | 0.617                 | 0.777                 |
| X2                  | -0.563                | 0.826                 |
| X3                  | 0.625                 | 0.739                 |
| X4                  | 0.841                 | -0.535                |
| X5                  | -0.882                | 0.422                 |
| X6                  | -0.950                | -0.008                |
| X7                  | 0.865                 | 0.392                 |

Table 5. Standardized value of the 7 original variables

| ZX1   | ZX2   | ZX3   | ZX4   | ZX5   | ZX6   | ZX7   |
|-------|-------|-------|-------|-------|-------|-------|
| -1.661| -0.451| -1.789| -0.255| 0.348 | 1.211 | -1.344|
| -0.078| 0.218 | 0.447 | -0.210| 0.348 | -0.282| -0.501|
| 0.271 | 0.545 | 0.447 | -0.322| 0.648 | -0.080| -0.025|
| 0.940 | 1.148 | 0.447 | -0.925| 0.430 | 0.597 | 0.622 |
| 0.528 | -1.460| 0.447 | 1.712 | -1.775| -1.446| 1.248 |

Table 6. Sorted comprehensive values of the principal component

| Year | F1 value | F1 ranking | F2 value | F2 ranking | F value | F difference | Comprehensive ranking |
|------|----------|------------|----------|------------|---------|--------------|-----------------------|
|      |          |            |          |            |         |              |                       |
2013  -2.298  5  -2.068  5  -2.214  -  5
2014  -0.263  2  0.329  3  -0.046  2.168  4
2015  -0.315  3  0.912  2  0.135  0.181  3
2016  -0.474  4  1.869  1  0.386  0.251  2
2017  3.351  1  -1.042  4  1.739  1.353  1

According to the sorted comprehensive values of the principal component in years 2013 to 2017, F2017 > F2016 > F2015 > F2014 > F2013. As the selected threat factors of human activities are all data types that can have potential negative impact on the ecological environment of Qilihai wetland, the statistical results of principal component analysis show that the threat of human activities is on the rise from 2013 to 2017. This is consistent with the economic development of surrounding areas in recent years and the continuous human impact on the ecological environment of Qilihai wetland. In addition, the difference between F values in different years is the largest in 2013-2014, followed by the difference in 2016-2017. Such results indicate that the Qilihai wetland was the most seriously threatened by human activities during 2013-2014, followed by 2016-2017. According to the results of initial factor loading matrix, the industrial output value above designated scale, total highway mileage and population are the most important indicators affecting the wetland, indicating that industrial development, transportation and the increase of human activities around the wetland within the region may be important factors affecting the ecological environment of the wetland.

Acknowledgments
Authors wishing to acknowledge the funding of the National Water Pollution Control and Treatment Science and Technology Major Project (2017ZX07107-004), the Fundamental Research Funds for the Central Public Welfare Research Institutes (TKS20200405 and TKS190301), and Science and Technology Project of Tianjin Municipal Bureau of Planning and Natural Resources (Study on ecological environment effect of restoration project of Tianjin Ancient Coast and Wetland Reserve).

References
[1] Lewis, W. M., Ebrary, I. (1995). Wetlands: Characteristics and Boundaries. Wetlands Characters & Boundaries, pp. 901-902.
[2] Qin L. (2012). Study on environmental Evolution of Qilihai Ancient Lagoon wetland in Tianjin. Wetland Science, 10 (2): 181-187.
[3] Peng, S.T., Zhao, Y.D., Zhang, G.Y., Jing L., Zhan, S.F., Yu, H. (2009). Biodiversity status and Protection countermeasures of Qilihai wetland. Watercourse and Port 02: 135-138.
[4] Lin J.T. (2014). Research on wetland restoration of Qilihai Based on scientific protection of landscape resources in ancient Geological time. Tianjin University.
[5] Baldi, P., Hornik, K. (1989). Neural networks and principal component analysis: Learning from examples without local minima. Neural Networks, 2 (1): 53-58.
[6] Li C.F., Yang Z.B. (2007). Eigenvector calculation in SPSS principal Component analysis. Statistical Education, 000 (003): 10-11.