Proceedings of the 3rd International Conference on Green Energy, Environment and Sustainable Development (GEESD2022), X. Zhang et al. (Eds.)

doi:10.3233/ATDE220312

Research on Ecological Evaluation Model Based on Principal Component Analysis

Litao LIU[1], Lei LI and Haozhou MO
Pittsburgh Institute, Sichuan University, Chengdu 610207, Sichuan, China

Abstract. In recent years, China has adhered to the concept of sustainable development of respecting nature, protecting nature and living in harmony with nature. A typical example is China's long-term ecological management of Saihanba. Since 1962, after more than half a century of rescue, China has established the world's largest artificial forest, improving the ecological environment of Saihanba area, making Saihanba the lungs of surrounding cities (Beijing, Tianjin, etc.). We collect relevant environmental data from 2002 to 2020, established 11 indicators through principal component analysis, score the annual environmental quality of Saihanba with 99% confidence, and propose an environmental assessment model for Saihanba. In second part, we collect environmental data in Beijing in recent years, such as the frequency of sandstorms, particle concentration, etc., and aims to explore the benefits of Saihanba governance to surrounding cities. The results show that after the implementation of environmental management measures, the environmental quality of Saihanba has been improved, and it has had a positive impact on the surrounding cities, that is, improved the environmental quality of nearby cities such as the Beijing and Tianjin.

Keywords. PCA, ecological evaluation, Saihanba

1. Introduction

In recent years, China has adhered to the concept of sustainable development of respecting, protecting and living in harmony with nature [1]. A typical example is the ecological treatment of Saihanba in China for a long time. Saihanba, located in Chengde City, Hebei Province, China, was once a royal hunting ground in China in the Qing Dynasty, but due to a large number of reclaimed waste lands in the Qing Dynasty, the ecology of Saihanba area was degraded into plateau hills. Since 1962, after more than half a century of rescue, China has established the largest artificial forest in the world, improved the ecological environment of Saihanba area, made Saihanba become the lung of surrounding cities (Beijing, Tianjin, etc.), and provided a lot of clean resources to people. The successful transformation of Saihanba provides an alternative model for environmental governance in countries in the Asia-Pacific region [2-4]. Only when countries participate in global environmental governance actions can we solve the environmental problems facing the world at this stage, such as Greenhouse Effect. China has promised the world to achieve carbon peak in 2030 and car-bon neutralization by 2060, so it is imperative to vigorously promote national environmental governance and improve green vegetation coverage [5].

[1] Corresponding Author, Litao LIU, Pittsburgh Institute, Sichuan University, Chengdu 610207, Sichuan, China; Email: 1569589244@qq.com.

This work mainly analyzes the collected environmental data of Saihanba, and the data used are all from the official website of the national environmental department. Principal component analysis still plays an important role in environmental assessment in recent years. this year, many scholars use it to analyze environmental assessment problems and draw convincing conclusions [6-9]. Therefore, this paper also plans to use principal component analysis to evaluate the environmental quality since the implementation of environmental control measures in Saihanba. In addition, we analyzed the environmental quality changes in the surrounding cities of Saihanba in recent years, with the purpose of analyzing the positive impact on the environmental quality of the surrounding cities since the environmental management of Saihanba.

The key mathematical notations used in this paper are listed in Table 1.

Table 1. Notations used in this paper.

Symbol	Description	Unit	Symbol	Description	Unit
P1	Forest coverage	/	P9	Days of urban air quality up to standard (better than level II)	/
P2	Covered area (10,000)	mu	P10	Urban PM2.5 concentration	ug/m^3
P3	Forest accumulation (10,000)	m^3	P11	Surface water quality up to standard rate	/
P4	Water conservation capacity (100 million cubic meters)	m^3	Fscore	The function of the scoring in all the Part	/
P5	Carbon dioxide absorption (10,000)	ton	B1	Air quality index (AQI)	/
P6	Oxygen release (10,000)	ton	B2	Days of good air quality	/
P7	Tourist trips (10,000 people)	/	B3	Urban PM2.5 concentration	ug/m^3
P8	Tourism income (100 million)	yuan	B4	Urban PM10 concentration	ug/m^3

2. Eco-environmental Impact Evaluation Model of Saihanba

2.1. Data Description

Through the collection of environmental data of Saihanba, we collect relevant environmental indicators for some years from 2002 to 2020, which are Forest Coverage (%), Covered Area (10,000 mu), Forest Accumulation (10,000), Water Conservation (100 million), Carbon Dioxide Absorption (10,000 tons), Oxygen Release (10,000 tons), Tourist Trips (10,000 people), Tourism Income (100 million), Days of Urban Air Quality up to Standard (better than Grade II), Urban PM2.5 Concentration (microgram), Surface Water Quality Standard Rate (%). As the data sources are too messy, we have sorted out the data we need. We preprocess the data corresponding to each index to get the matrix diagram below, as shown in Figure 1.

From the matrix diagram obtained from the pretreatment, we can see that the overall trend of the selected positive indicators (such as carbon dioxide absorption, etc.) is increasing year by year, in which it is worth noting that the number of tourists, a positive index, dropped sharply in 2020, and our group believes that it is the impact of COVID-19. In the later analysis, we will ignore this untrusted value and choose the predicted value instead of this untrusted value to analyze the problem. Among them, the concentration of PM2.5 in urban area is a reverse index, and the overall trend decreasing year by year can also reflect the positive response of the restoration of Saihanba to the surrounding cities.

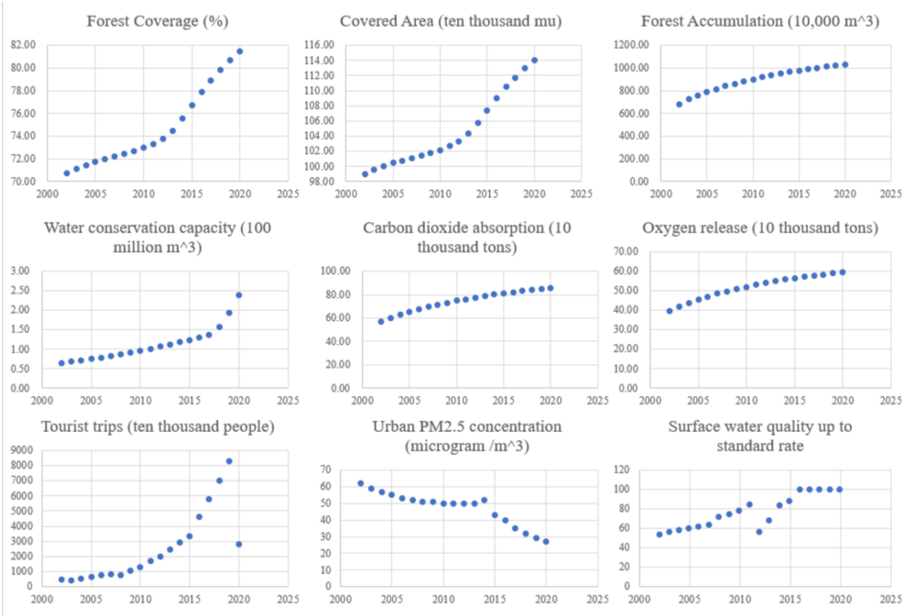

Figure 1. Data matrix diagram of Part 1 (Data sources include National urban air quality real-time release platform, NOAA Physical Sciences Laboratory (PSL), etc.)

2.2. The Combination of PCA and Data

Principal component analysis (PCA) can be traced back to the multivariate transformation analysis of non-random variables initiated by Pearson in 1901 and extended to random variables by Hotelling in 1933. PCA is a statistical process that allows users to summarize the content of information in large data tables through a set of smaller summary indices that are easier to visualize and analyze [3]. We plan to migrate and apply the existing PCA model to the establishment of this evaluation model, which mainly includes the following steps [5].

Through the analysis of the initial data, there are 11 available indicators P1-P11 (see Notations). At the same time, we deal with the negative indicators positively, for example, change the PM2.5 concentration to its inverse data. The following is the positive formula used.

$$x_{use} = \frac{1}{x_{initialvalue}} \tag{1}$$

$$y_{ij} = \frac{x_{ij} - \overline{x_j}}{s_j}, i = 2002 \ldots 2020, j = 1 \ldots 11 \tag{2}$$

$$r_{ij} = \frac{\sum_{k=1}^{n} (x_{ki} - \overline{x_i})(x_{kj} - \overline{x_j})}{\sqrt{\sum_{k=1}^{n} (x_{ki} - \overline{x_i})^2 \sum_{k=1}^{n} (x_{ki} - \overline{x_j})^2}}, (i, j = 1 \ldots 11) \tag{3}$$

Among them, r_{ij} is the correlation coefficient between the different Index; and then, we found that most of the indicators have a strong correlation, which means that it is very suitable to use principal component analysis to extract principal components. For example, the covariance between Tourist trips (P7) and Tourism income (P8) is 0.9959. We finally choose P11, P10, P9 and P8 as the environmental assessment index of Saihanba. The weights of the top four principal components are normalized, and the normalized mathematical expression is given below. Figure 2 shows the weight of each indicator.

$$R = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{4}$$

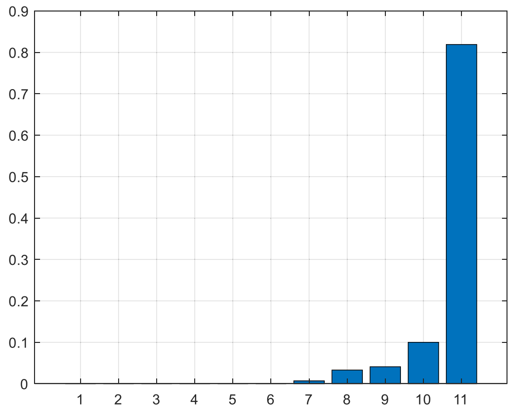

Figure 2. Weight of principal components.

Taking 99% degree of explanation as the standard, we select 4 principal components. The wight of principal components is 81.9%, 10.0%, 4.1% and 3.3%, respectively.

2.3. Results of Principal Component Analysis

Through the analysis of the data we can find, we find out four principal components, which are Surface water quality up to standard rate and Urban PM2.5 concentration, Days of urban air quality up to standard and Tourism income, with 99% as the standard of interpretation. By calculating the weight of the four principal components, we obtain the scoring function, as shown in Figure 3.

$$Fscore = 0.823 \times \mathrm{P11} + 0.102 \times \mathrm{P10} + 0.042 \times \mathrm{P9} + 0.03 \times \mathrm{P8} \tag{5}$$

Through the analysis of the ecological data of Saihanba from 2002 to 2020, we have realized the quantitative evaluation of the ecological environment of Saihanba. It is not difficult to find that with the continuous restoration of Saihanba, its green area, oxygen content and carbon dioxide absorption are increasing almost every year, and it has significantly improved the local water quality and urban air quality. at the same time, it has also promoted the development of tourism in Saihanba area in recent decades.

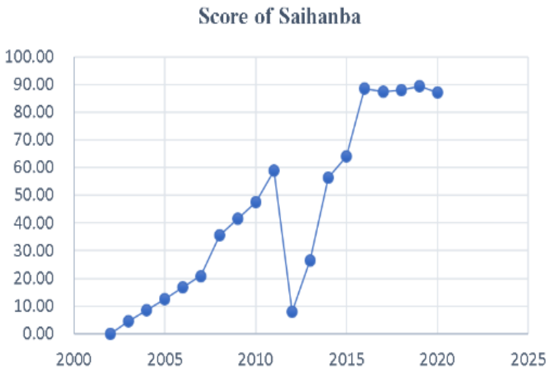

Figure 3. Ecological environment assessment curve of Saihanba from 2002 to 2020.

3. Environmental Quality Evaluation in Beijing in Recent Years

3.1. Data Description

The harm of sandstorm to human body is huge. For example, when particles enter the upper respiratory tract, these harmful substances stimulate and corrode the upper respiratory tract, causing damage to the defense system of the respiratory tract [10].

As shown in Figure 4, from the matrix diagram, it is not difficult to find that B1, B3 and B4 belong to negative index, and we need to deal with them positively in the subsequent analysis. In a whole view, the negative indicators show a weak trend, and we can find that the environment in Beijing is constantly improving. B2 is positive indicators, for B2 data map, its overall trend is increasing, which is a response to the improvement of Beijing's ecological environment.

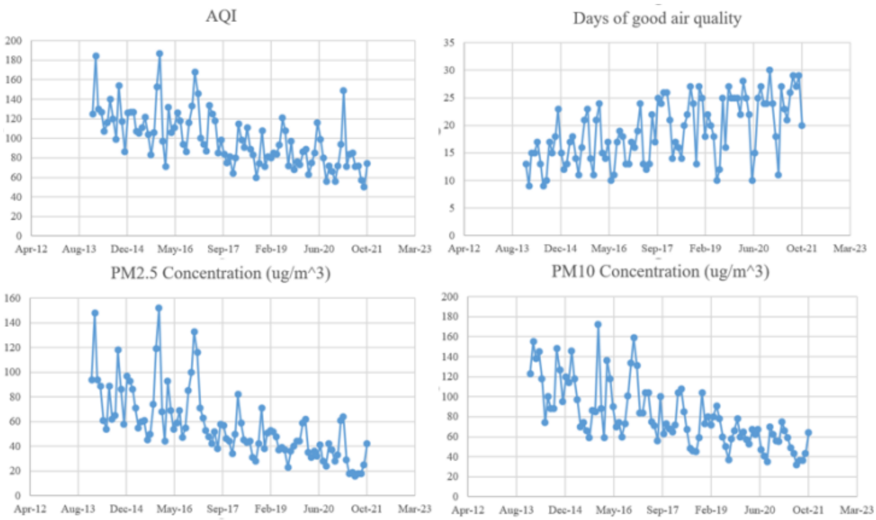

Figure 4. Matrix map of ecological data in Beijing.

3.2. The Combination of PCA and Data

We plan to continue to use the PCA to get the weight of each index (B1-B4), and then score Beijing's ecological environment on a monthly basis to get a score matrix from January 2014 to November 2021. We deal with the negative indicators positively using equation (1). And then, we standardize the data according to the first step.

$$y_{ij} = \frac{x_{ij} - \overline{x_j}}{s_j}, i = 1\ldots 95, j = 1,2,3,4 \tag{6}$$

Among them, s_j is the standard deviation, and y_{ij} is the corresponding value after standardization. The following figure shows all the standardized data (Partial). The correlation coefficient matrix is calculated:

$$r_{ij} = \frac{\sum_{k=1}^{n}(x_{kj} - \overline{x_i})(x_{kj} - \overline{x_j})}{\sqrt{\sum_{k=1}^{n}(x_{kj} - \overline{x_i})^2 \sum_{k=1}^{n}(x_{kj} - \overline{x_j})^2}}, (i, j = 1,2,3,4) \tag{7}$$

Among them, r_{ij} is the correlation coefficient between the different Index. Through observation, we found that most of the indicators have a strong correlation, which means that it is very suitable to use principal component analysis to extract principal components. For example, the covariance between AQI (B1) and Days of good air quality (B2) is 0.892.

Since all the indicators are used as the coefficients of the scoring function, there is no need for normalization, and the final weight is shown in Table 2.

Table 2. Final weight distribution.

The final scoring weight	
B4	79.5%
B3	15.2%
B2	3%
B1	2.3%

3.3. Results of Principal Component Analysis

Through the analysis of the limited data, we get the weight of four indicators and establish the following quantitative evaluation equation.

$$Fscore2 = 0.795 \times B4 + 0.152 \times B3 + 0.003 \times B2 + 0.0023 \times B1 \tag{8}$$

As shown in the analysis results in Figure 5, by comparing the evaluation curves of Saihan Dam and Beijing, consulting the strategic objectives of China's restoration of Saihan Dam, and analyzing the geographical location of Saihan Dam and Beijing, we can draw an obvious conclusion that the restoration of Saihan Dam has a positive impact on the ecological environment of the Chinese capital Beijing. The particulate matter data used is closely related to Beijing sandstorm, and this kind of targeted data can make us aware that the restoration of Saihan Dam is positive for the control of sandstorm in Beijing.

Figure 5. Monthly environmental assessment in Beijing from 2014 to 2021.

4. Conclusion

In the first part, we established a Saihanba environmental quality assessment model by principal component analysis, and we successfully quantified the improvement of the local environment since Saihanba was rehabilitated. In the second part, we established the Beijing environmental assessment model and observed that the governance of Saihanba had a positive impact on the surrounding cities and improved the environment of the surrounding cities. Therefore, the establishment of environmental governance zones is a positive response to sustainable development strategies and the long-term goal of creating well-being for future generations.

References

[1] Kostka G, Nahm J. Central–local relations: Recentralization and environmental governance in China. The China Quarterly, 2017, 231: 567-582.

[2] Wang M, Wang E, Zhou F, et al. Effect analysis of application of drip irrigation in sandstorm land protective forest. Journal of Fujian Forestry Science and Technology. 2009;2.

[3] Wold S, Esbensen K, Geladi P. Principal component analysis. Chemometrics and Intelligent Laboratory Systems. 1987;2(1-3):37-52.

[4] Wang SY, Liu JS, Yang CJ. Eco-environmental vulnerability evaluation in the Yellow River Basin, China. Pedosphere. 2008;18(2):171-82.

[5] Ringnér M. What is principal component analysis. Nature Biotechnology. 2008;26(3):303-4.

[6] Yue H, Liu Y, Li Y, et al. Eco-environmental quality assessment in China's 35 major cities based on remote sensing ecological index. IEEE Access, 2019, 7: 51295-51311.

[7] Touzani I, Machkor M, Boudouch O, et al. Environmental Impact Assessment of Taza City's Wastewater: Application through Principal Component Analysis. Scientifica, 2020, 2020.

[8] Xu S, Cui Y, Yang C, et al. The fuzzy comprehensive evaluation (FCE) and the principal component analysis (PCA) model simulation and its applications in water quality assessment of Nansi Lake Basin, China. Environmental Engineering Research, 2021, 26(2).

[9] Chen C, Ge Q, Tao Z, et al. Evaluation of Ecological Environment Quality in Chongqing Main City Area Based on Principal Component Analysis. Scientific Programming, 2022, 2022.

[10] Meo SA, Almutairi FJ, Abukhalaf AA, et al. Sandstorm and its effect on particulate matter PM 2.5, carbon monoxide, nitrogen dioxide, ozone pollutants and SARS-CoV-2 cases and deaths. Science of the Total Environment. 2021;795:148764.