ISSN 2616-5872 Vol.4, Issue 1: 42-46, DOI: 10.25236/AJEE.2022.040109

Research on the Ecological Environment Evaluation of Saihanba Based on Entropy Weight Method

Kaibin Xu

College of Mechanical and Electrical Engineering, Sichuan Agricultural University, Ya'an, Sichuan, 625000, China

Abstract: *This paper mainly studies the ecological evaluation of Saihanba. First, this paper calculate the Pearson correlation coefficient of the collected Saihanba data indicators, and after obtaining their significant correlations, because too many indicators are not conducive to the calculation, perform cluster analysis on the data, due to the K-means clustering method The subjectivity is too strong, use systematic clustering method to cluster the target, and use the elbow cri-terion to obtain the K value. Finally import the classified data indicators into MATLAB, and use the entropy weight-TOPSIS model to score and quantitatively evaluate Saihanba's environmental rating.*

Keywords: *Ecological evaluation, Pearson correlation coefficient, K-means, Entropy weight-TOPSIS model*

1. Introduction

Saihanba Forest Farm was originally a vast plateau outside the Great Wall with white sands and sparse trees[1]. The species is scarce and the ecosystem is fragile. Moreover, its harsh ecological environment has severely affected the weather conditions in Beijing, which is not conducive to the sustainable development of our country[2].

Since 1962, more and more young people have planted their dreams of restoring forests and repainting a piece of greenery in Saihanba. Today, the ecology of Saihanba has been restored and the water quality has been steadily improved[3]. It plays an important role in maintaining ecological balance and stability, and has become an important ecological barrier to guard Beijing and Tianjin. With the new goal of ecological restoration, the people of Saihanba[4] have created three major projects: afforestation, natural improvement of artificial forests, and near-planning and cultivation of natural forests, striving to bring artificial forests closer to natural forests.

2. Model preparation

2.1. Pearson correlation analysis

In statistics, the Pearson product-moment correlation c-oefficient is usually repre-sented by r or p, and is used to measure the relationship between two variables X and Y, the value range is between [-1, +1]. The Pearson correlation coefficient is widely used in academic research to measure the strength of the linear correlation between two varia-bles.

First, perform descriptive statistics on Saihanba’s forest coverage, coverage area, forest accumulation, water conservation, carbon dioxide absorption, and oxygen release from 1962 to 2021. Then make a scatter diagram to observe whether there is a linear relationship between each group of variables.

It can be seen from the graph that there is a strong correlation between most of the variables. In the next step, we will calculate the Pearson correlation coefficient and use the corrcoef function of MATLAB to find the correlation coefficient matrix between the variables as shown in the following figure 1:

From the data comparison table 1 in the above table, it can be seen that there is a significant correlation between the various data. And the number of indicators is large, and the year is long, the next step is to choose a clustering algorithm to classify the data and indicators, which is convenient for later scoring statistics.

ISSN 2616-5872 Vol.4, Issue 1: 42-46, DOI: 10.25236/AJEE.2022.040109

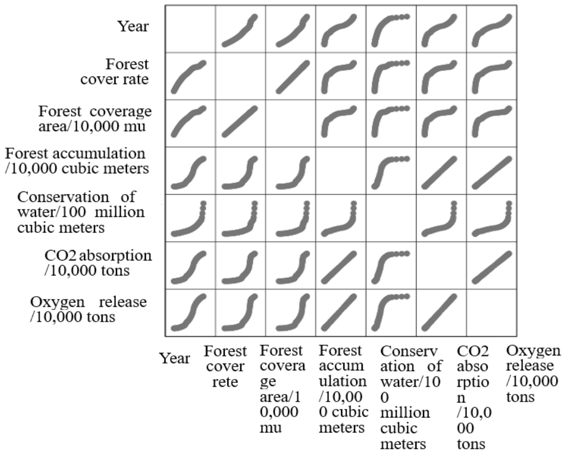

Figure 1: Scatter plot

Table 1: Correlation coefficient matrix

1	0.974741523	0.974749771	0.961694523	0.853825204	0.961693173	0.96169622
0.97474152	1	0.999999982	0.884517674	0.760059597	0.884518964	0.884520571
0.97474977	0.999999982	1	0.884533241	0.760089795	0.884534527	0.884536139
0.96169452	0.884517674	0.884533241	1	0.868979022	0.999999996	0.99999999
0.8538252	0.760059597	0.760089795	0.868979022	1	0.868965845	0.868976832
0.96169317	0.884518964	0.884534527	0.999999996	0.868965845	1	0.999999988
0.96169622	0.884520571	0.884536139	0.99999999	0.868976832	0.999999988	1

2.2. Cluster analysis

Cluster analysis refers to the analysis process of grouping a collection of physical or abstract objects into multiple classes composed of similar objects. The K-means clus-tering analysis method in dynamic analysis is used in this paper[5]. Since the dimensions of each index are different, first eliminate the dimensions between each index. The results obtained are shown in the Table 2:

Table 2: Descriptive Statistics Table

	N	Min	Max	Mean	Standard Deviation
Forest cover rate	60	13.5714285714286	82.2142857142857	55.9858998621786	19.23176289670105
Forest coverage area/ 10,000 mu	60	19.0000000000000	115.1000000000000	78.3802598070501	26.92446805538147
Forest accumulation/ 10,000 m^3	60	33.0000000000000	1036.800000000000	424.961851804728	372.0385901213838
Conservation of water/ 10^7 m^3	60	0.090000000000000	2.840000000000000	0.59180045275135	0.565426462091106
CO2 absorption/ 10,000 tons	60	2.73822337962963	86.03000000000000	35.2618326685578	30.87044744226722
Oxygen release/ 10,000 tons	60	1.90462962962963	59.84000000000000	24.5271192245321	21.47259763972184
valid cases	60				

After that, perform K-means clustering analysis on the data with dimensions removed, and select the number of clusters to be 2 and 3 respectively for clustering.

When the clustering result is set to 2, the results obtained are based on 1987. The six data indicators after 1987 have been significantly improved compared with those before 1987. The transformed

ISSN 2616-5872 Vol.4, Issue 1: 42-46, DOI: 10.25236/AJEE.2022.040109

Saihanba has rich natural resources, improved forest quality, improved water quality, and enhanced ecosystem regulation capabilities.

When the set result is 3, the result obtained distributes the 60-year data at two time points in 1983 and 2001. During 1964-1983, the environment of Saihanba showed a slow growth trend, and in 1983-2001 During the year, the environmental condition of Saihanba was in a transitional period. From 2001 to 2021, the environmental situation of Saihanba was at a rapid growth rate. Optimization, ecological restoration work has made remarkable achievements, and the improvement rate of Saihanba's ecological environment has shown a rapid growth.

Since the K-means clustering algorithm requires the operator to set the K value, and the system clustering algorithm will be used. Use the elbow criterion to analyze the value of K. The coefficients are stored in the table of data:

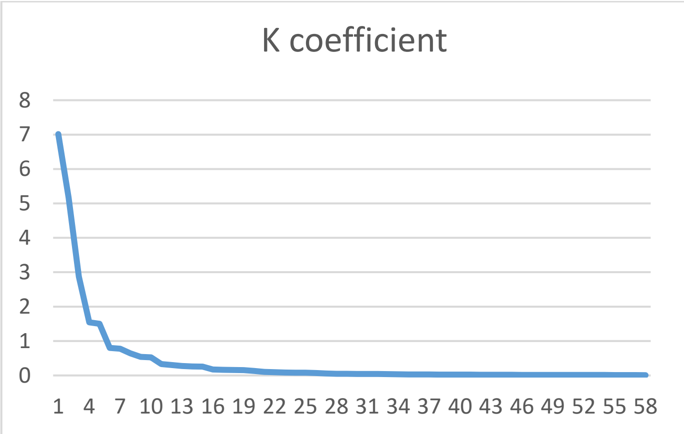

Figure 2: Line graph of aggregation coefficient

According to the polyline graph of the aggregation coefficient, it can be seen that when the K value is from 1 to 3, the degree of distortion changes the most. After more than 5, the degree of distortion changes significantly. Therefore, the elbow is K=3, so the number of categories can be set to 3.

The results obtained divide 1962-2000 into the first category. During the 37 years, the environmental condition of Saihanba has shown a low growth trend, and the period from 2001 to 2019 is divided into the second category. Therefore, in these 17 years, the environmen-tal condition of Saihanba has shown a transitional trend. In the period of 2020-2021, The environmental conditions of Saihanba show a high growth trend.

3. Establishment of model

The essence of entropy weight is that when evaluating the thing, it should be given a smaller weight, otherwise, it should be given a larger weight[6].

For information entropy: suppose that x represents a certain situation that event X may occur, and p(x) represents the probability of occurrence of this situation. We can define

$$I(x) = -ln(p(x)), \text{ because } 0 \leq p(x) \leq 1, \text{ so } I(x) \geq 0. \quad (1)$$

The possible situations if event X are $x_1, x_2, \ldots x_n$. Then we can define the information entropy of event X as:

$$H(x) = \sum_{i=1}^{n}[p(x_i)I(x_i)] = -\sum_{i=1}^{n}[p(x_i)\, ln(p(x_i))] \quad (2)$$

It can be seen from the above formula that the essence of information entropy is the expected value of the amount of information.

The basic steps of the entropy method are as follows:

Step 1: Build the original matrix

First, according to the data of the past 60 years in Saihanba, the original matrix is constructed:

ISSN 2616-5872 Vol.4, Issue 1: 42-46, DOI: 10.25236/AJEE.2022.040109

$$X = (x_{ij})_{n\times m} \quad (i = 1,2,3,\dots,n, j = 1,2,3,\dots,m) \tag{3}$$

Step 2: Normalized processing data

Normalize the selected indicators to initially eliminate the differences caused by dimensions:

For positive indicators: $Y_{ij} = \frac{X_{ij}-min(X_{ij})}{max(X_{ij})-min(X_{ij})}$

For negative indicators: $Y_{ij} = \frac{max(X_{ij})-X_{ij}}{max(X_{ij})-min(X_{ij})}$

Step 3: Calculate the entropy

$$e_j = -\frac{1}{ln\,n}\sum_{i=1}^{n\Sigma_{ij}} p_{ij} ln \tag{4}$$

Step 4: Calculate the index weight

$$w_i = \frac{1-e_j}{m-\sum e_j}(j = 1,2,3 \dots, m) \tag{5}$$

Step 5: Calculate single index and comprehensive score

The TOPSIS method is more efficient in comprehensive evaluation. Therefore, the entropy weight method and the TOPSIS method are used to process the collected raw data of each province, calculate the weight and convert it into the corresponding score, and then obtain its comprehensive development index Sj, Based on this, comprehensive evaluation and analysis of the development of various regions are carried out.

$$S_{ij} = w_i \times X_{ij} \qquad S_i = \sum_j^n S_{ij} \tag{6}$$

4. Model results

Import the data into MATLAB, first standardize the data, and then use the entropy method to determine the weight of each indicator, and get the following Table 3:

Table 3: Index weight table

Forest cover rate	Forest coverage area/ 10^4 mu	Forest accumulation/ 10^4 m^3	Conservation of water/ $10^7 m^3$	CO2 absorption/ 10^4 tons	Oxygen release/ 10^4 tons
0.0387	0.0387	0.2363	0.2138	0.2363	0.2363

After that, bring the data divided into three categories in the system cluster into the program to calculate the score and get the score:

Table 4: Score Table

1962-2000	2001-2019	2020-2021
0.0072	0.032	0.053

It can be seen from the score that during 1962-2000, although the ecological environment of Saihanba has been improved, the overall ecological environment is not optimistic. Data show that the forest coverage rate has increased from 13.57% to 69.44%, and the forest coverage area has increased from 190,000 mu to 972,200 mu. In 2000, the forest stock volume reached 5,445,500 cubic meters, the water conservation volume reached 58 million cubic meters, and the carbon dioxide absorption reached 451,800 tons and oxygen release amounted to 314,300 tons. Compared with the data in 1962, these ecological environmental indicators have been improved to varying degrees, but there is still a lot of room for improvement, and the ecosystem is still in a relatively fragile state. Ecological benefits need to be improved, and the regulation capacity of the ecosystem needs to be improved.

During 2001-2019, the ecological environment of Saihanba has been restored to a certain extent, the ecological structure layout has been optimized, and the forest coverage rate has increased to 80.66%. The area has reached 1,129,300 mu, and these two indicators have been steadily improved on the basis of 2000. In addition, the ecological benefits of Saihanba have also been improved. The forest volume has doubled compared with 2000, reaching 10.207 million cubic meters, and the amount of water conserved has tripled compared with 2000, reaching 194 million cubic meters. The amount of carbon dioxide absorption has increased. In 2000, it nearly doubled to 846,900 tons, and the amount of oxygen released nearly doubled from 2000 to 589,100 tons. It can be seen that from 2001 to 2019, various ecological environment indicators have been greatly improved, the rate of improvement of the ecological environment is high, and the rate of restoration of the ecological environment of Saihanba has shown a rapid growth trend.

During 2020-2021, Saihanba has changed from barren sandy land to a green pearl, becoming a green miracle of ecological environment management. The forest coverage rate has reached a new high, from 13.57% to 82.21%, an increase of 68.64%; the forest coverage area has attracted worldwide attention, increasing from 190,000 mu to 1.151 million mu, an increase of 961,000 mu in total; Saihanba is rich in natural resources There are 261 species of terrestrial wild vertebrates, 32 species of fish, 660 species of insects, 179 species of macrofungi, and 625 species of plants, which together constitute the rich ecosystem of Saihanba; the ecosystem’s ability to regulate continues The ecological benefits are very optimistic. The forest stock volume ranks in the forefront, as high as 10.368 million cubic meters, which is about 31.42 times the data in 1962; the water quality has been improved, and the amount of conserved water has been continuously increased. It has reached 284 million cubic meters. Compared with 1962, it has increased by about 30.5 times; the carbon dioxide absorption is as high as 860,300 tons, which is about 31.4 times of the 1962 data; the oxygen release has increased from 19,000 tons to 598,400 tons, an increase of nearly 30.5 times. It can be seen that from 1962 to 2021, the improvement of Saihanba's ecological environment has achieved remarkable results, all indicators of the ecological environment have been improved by leaps and bounds, and the ecological benefits have been greatly improved. At present, Saihanba has become an important ecological barrier to guard Beijing and Tianjin, playing a major role in resisting wind and sand, improving the climate and conserving water sources.

5. Conclusion

Through ecological evaluation analysis,the improvement of Saihanba's ecological environment has achieved remarkable results, all indicators of the ecological environment have been improved by leaps and bounds, and the ecological benefits have been greatly improved.

The cluster analysis greatly improves the operating efficiency of the model and facilitates subsequent scoring and evaluation, and the computer is used to assign weights to various indicators to reduce subjectivity. Cluster analysis uses systematic clustering and elbow criteria to determine the K value, which makes the value more accurate and reliable.

References

[1] Xiaoxia Wang. Research on the Cooperative Legislation Mechanism of Ecological Protection Regions in the Upper Yellow River [D]. Northwest Normal University, 2021.

[2] Ziqiang Wang, Jinliang Ke, Bangkai Ke. Eight major environmental problems in the Asia-Pacific region [J]. World Environment, 1997(04): 41-43.

[3] Wang Tao, Yuanhui Deng, Zhao Yu, Yating Dong, Boran Li, Xumin Li. Analysis of the main impact indicators of the innovation and entrepreneurship ecological environment of the regional double innovation demonstration base [J]. Today Keyuan, 2021(07): 13-25.

[4] Changfu Yu. Research on the Index System of Environmental Impact Assessment of Farmland Water Conservancy Projects [J]. Environmental Science and Management, 2020, 45(08): 178-182.

[5] Changming Li, Hongchen Zhang, Wang Chao, Xiaoguang Li, Lu Yang, Chaoyue Qian. An efficient yin and yang k-Means clustering algorithm [J]. Journal of Jilin University (Science Edition), 2021, 59(06): 1455 1460. DOI: 10.13413/j.cnki.jdxblxb.2020406.

[6] Hanshan Si. Research on the comprehensive evaluation of the development of 4 provinces in the middle reaches of the Yellow River based on the entropy weight-TOPSIS model [J]. Science and Technology and Industry, 2021, 21(11): 203-207.