DOI: 10.23977/erej.2022.060114
ISSN 2616-3756 Vol. 6 Num. 1

Study on Influencing Factor Index of Saihanba Forest Farm Based on Genetic Algorithm

Hailong Sun

Huazhong Agricultural University, Wuhan, Hubei, 430000, China

Keywords: Saihanba Forest Farm, Index system, Genetic algorithm

Abstract: For Saihanba forest farm, 19 indexes are selected and divided into three levels: pressure, state and response. The data dimension is deduced by curve projection pursuit model, and the environment change is analyzed by accelerated genetic algorithm. Then input the data into the model, and finally get the ecological environment score 1. 3945 in 1980 and 1 in 2001. 6475 2020. This shows that the environmental conditions have been significantly improved after the establishment of Saihanba.

1. Introduction

Saihanba forest area is located at 42°02 ~ 42°36 'N, 116°51' ~ 117°39 'E, located in the north of Weichang County, Hebei Province, for the upper reaches of the Luan River and Liao River.

In the Qing Dynasty, due to the corruption of local officials and the decadence of finance, the Qing government, which was troubled at home and abroad, opened the enclosure for reclamation in 1863, and the forest vegetation was destroyed in 1962, Saihanba Mechanical Forest Farm was formally established. Since 1962, 369 young people with an average age under 24 have come to this wasteland filled with yellow sand. Now it is the largest forest park in northern China.

2. Indicators Selection

Firstly, this paper analyzed the main research results of ecological environment assessment in many countries at home and abroad, and then combined with the characteristics of ecological environment in Saihanba constructed the environmental assessment index system of Saihanba [1], which includes Pressure, Statue and Response.

According to “Technical Criteria for Ecosystem Status Evaluation” the evaluation indicators of Saihanba’s impact on ecological environment are selected [2]. Firstly, by combing the relevant research results, 32 indicators with high frequency are obtained by using the frequency analysis method. Combined with the relevant characteristics of the climate in Saihanba area, the selected indicators are adjusted for the first time, and some indicators with little reference significance to this kind of climate are eliminated.

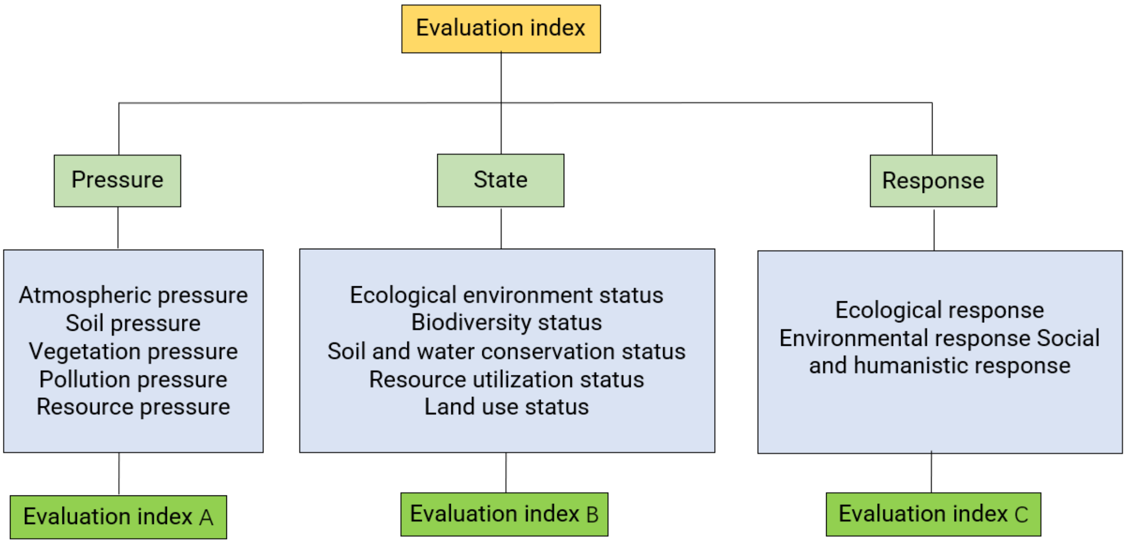

Figure 1: Evaluation Index

3. Construction curve projection pursuit dynamic clustering model

Projection pursuit (PP) is an exploratory data analysis method directly driven by sample data.

3.1 Curve projection

Hyperbolic tangent function equation:

$$y = \frac{e^x - 1}{e^x + 1} \tag{1}$$

Let $\vec{a}$ be the unit vector of m-dimensional projection direction, and its components $a_1, a_2, \ldots, a_m$ are the weights of M evaluation indexes respectively, and have:$\sum_{j=1}^{m} a_j^2 = 1$ $(-1 \le a_j \le 1)$. Then the projected eigenvalue of the curve can be expressed as:

$$z_i = \frac{e^{\sum_{j=1}^{m} a_j x_{ij}} - 1}{e^{\sum_{j=1}^{m} a_j x_{ij}} + 1} (i = 1,2,3 \ldots n) \tag{2}$$

3.2 Construction of projection dynamic clustering index

As mentioned above, the dynamic clustering method is used to construct a new projection index [3-5].

$$\varphi_t = \{z_i | d(A_t - z_i) \le d(A_h - z_i)\} \begin{cases} h = 1,2,3,\ldots,p \\ t \ne h \end{cases} \tag{3}$$

Here, $d(A_t - z_i) = |z_i - A_t|, d(A_h - z_i) = |z_i - A_h|, A_t, A_h$ denote the t-th and H-th order nuclei respectively. Here we assume that the nuclei are monotonically decreasing functions. The aggregation degree of samples in the class is expressed by the length and in the class, that is:

$$d(\vec{a}) = \sum_{i=1}^{p} [\max(\varphi_t) - \min(\varphi_t)] \tag{4}$$

The smaller $d(\vec{a})$ here means the higher the aggregation degree of samples in the class. The dispersion degree between samples can be expressed by the dispersion degree between samples, that is: $s(\vec{a})$=$\sum_{i,k\epsilon n,i\neq k}|\varphi_i-\varphi_k|$, and the larger the $d(\vec{a})$, the higher the dispersion between samples. The basic requirement is to gather the samples as much as possible and disperse the samples between classes as much as possible. The evaluation index of curve projection pursuit dynamic clustering is defined as:

$$Q(\vec{a})=s(\vec{a})-d(\vec{a}) \tag{5}$$

Therefore, the quantitative model of Saihanba impact on ecological environment can be described as the following nonlinear optimization problems:

$$\begin{cases} maxQ(\vec{a}) \\ s.t\sum_{j=1}^{m}a_j^2=1 \\ -1\leq a_j\leq 1 \end{cases} \tag{6}$$

4. Solution based on accelerated genetic algorithm

(1) Encoding: The real coding adopted by RAGA utilizes the following linear transformation:

$$a_j=a_{jmin}+y_i\left(a_{jmax}-a_{jmin}\right)(j=1\dots n) \tag{7}$$

Where a_{jmax} and a_{jmin} are the maximum and minimum values of the projection direction components. The transformation can be regarded as encoding the projection direction a=($a_1,a_2,\dots,a_n$) as y=($y_1,y_2,\dots,y_n$).

(2) Initialize population: Let the number of individuals of the initialization population be M, randomly generate M groups of primary projection directions $\{a_{j,i}^{(0)}\}$ in $[a_{jmin},a_{jmax}]$, and bring them into the coding rules in step 1 to obtain the primary coding sequence $\{y_{j,i}^{(0)}\}$. The sequence is successively brought into the objective function Q (a) to obtain the objective function value $\{f_i^{(0)}\}$ of each individual in the population. The individuals are sorted in ascending order according to the value of the objective function, and the top N(N< M) individuals are selected as excellent individuals.

(3) Fitness evaluation: The smaller the objective function value, the higher the fitness of the individual. Since the objective function value may be 0, an adaptive factor of 0.001 is added. Therefore, the fitness function is defined as:

$$S_i=\frac{1}{(f_i^{(0)})^2+0.001} \tag{8}$$

(4) Produce the first generation group: The probability that the i individual is selected is $E=\frac{S_i}{\sum_{i=1}^{m}S_i}$. defining $\{p|p_i=\sum_{k=1}^{i}E_i\}$ and p will divide the [0,1] interval into m sub interval $\{[0,p_1),(p_1,p_2]\dots(p_M,1]\}$. Then (M-N) random numbers $Rand\in R^{M-N}$are randomly generated, and the individual selection method is:

$$y_{j,k}^{(0)}=y_{j,i}^{(0)}R^k\in(p_{i-1}-p_i) \tag{9}$$

(5) Second progeny population was produced by hybridization: According to the selection probability E_i obtained in step 4, a pair of individuals $y_{j,k1}$ and $y_{j,k2}$ are randomly selected for hybridization. Due to the limitation of coding method, this paper adopts the following random linear combination to obtain offspring individual $y_{j,i}^{(2)}$:

$$\begin{cases} y_{j,i}^{(2)}=r_1\times y_{j,i}^{(0)}+(1-r_1)\times y_{j,i}^{(0)}r_3<0.5 \\ y_{j,i}^{(2)}=r_2\times y_{j,i}^{(0)}+(1-r_2)\times y_{j,i}^{(0)}r_3\geq 0.5 \end{cases} \tag{10}$$

Where r_1, r_2, r_3 is a random number in the interval [0,1].

(6) Perform mutation operation to generate the third generation population: The probability of n random numbers $pm_i = 1 - E_i$ is used to replace the $y_{j,i}$ individual for variation:

$$\begin{cases} y_{j,i}^{(3)} = r_4 r_5 < pm_i \\ y_{j,i}^{(3)} = y_{j,i}^{(0)} r_5 \geq pm_i \end{cases} \tag{11}$$

5. Evaluation criteria and results

For the evaluation index, refer to the classification standard of ecological environment status in the technical criterion for economic system status evaluation, and divide the ecological environment status into excellent, good, average and poor. The specific index is as follows. In this paper, 10 samples are randomly generated within the level range of each level standard, that is, 5 level standards are extended to 50 sample sequences. According to the generated sample data at all level

Table 1: Classification criteria

Level	Index
I	EI≥75
II	55≤EI<75
III	35≤EI<55
IV	20≤EI<35
V	EI<20

Table 2: Scope of categories

Level	New Index
I	1.723514173
II	1.623412564
III	1.498222315
IV	1.385215423
V	1.244587456

We have collected these 19 indicators of Chengde city where Saihanba is located in the past 40 years from China Statistical Yearbook and major databases. The specific data are put in the appendix. The above data are normalized according to the above normalization principle and put into the curve projection pursuit dynamic programming model for calculation.

It can be seen from the score that the ecological environment of Chengde has gradually improved in recent 40 years. Draw the scores of Chengde in the above years into a scatter diagram, as shown in the figure below:

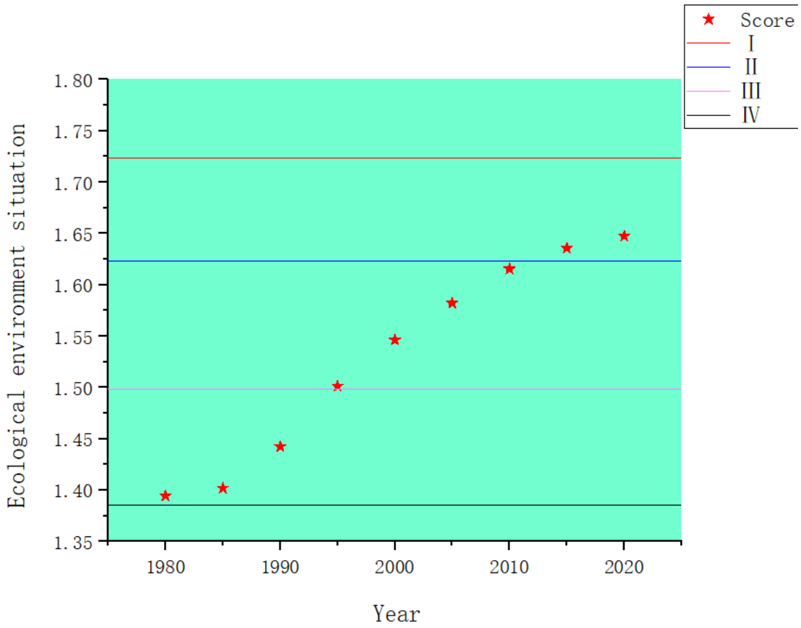

Figure 2: Scattered points of Chengde in various years

6. Conclusion

For Saihanba forest farm, this paper establishes a multi-level index system and analyzes it by using the algorithm. The final result is basically similar to the previous hypothesis. After the completion of Saihan dam, the impact on the ecological environment is relatively small. With the passage of time, the improvement effect on the environment is becoming more and more obvious. Finally, it will affect regional stability and make the ecological environment of Chengde city at a high level under the current standards.

References

[1] Zhang Li, Ren Guoyu. Evolution of sandstorm frequency and analysis of its climatic causes in northern China [J]. Journal of meteorology, 2003 (06): 744-750.

[2] Wang Shigong, Dong Guangrong, Chen Huizhong, Li Xiliang, Jin Jiong. Progress in dust storm research [J]. China desert, 2000 (04): 5-12.

[3] Wang Jiong, Xu Yueming, Guo Qing. Land use zoning and suggestions of Baoding City Based on cluster analysis [J]. China Agricultural Resources and zoning, 2011,32 (02): 63-67.

[4] Ning Youliang, Wu Shuirong, Li Zhiyong, Liu Shirong, Chen Shaozhi. Impact of climate change on China's forest ecosystem based on regional vegetation type assessment [J]. Journal of ecology, 2013,32 (08): 1967-1972. Doi: 10.13292/j.1000-4890.2013.0401.

[5] Xiao Peng. Study on impact assessment of land remediation on natural environment [D]. Yunnan University of Finance and economics, 2013.