Team Number :	apmcm2112772
Problem Chosen :	C

2021 APMCM summary sheet

The transformation of China's Saihanba Forest Farm from a desert to an oasis is an excellent example of the establishment of an ecological reserve. This article mainly uses the TOPSIS evaluation model based on the entropy weight method, the multiple regression model and the evaluation model based on BP Neural network to study the impact of the establishment of the Saihanba Ecological Reserve on the local ecological environment and the ability to resist wind and sand, and how to scientifically Planning to establish ecological protection zones.

For question 1, we need to establish a TOPSIS evaluation model of Saihanba′s impact on the environment. First, select the indicators and data that have an impact on the ecological environment in Saihanba area. For data that cannot be queried, we use cubic spline interpolation to fill in. In order to reduce the complexity of the model, we use the entropy method to reduce the dimensionality to obtain the index set. Finally, we can get the environmental score of Saihanba over the years and find that the ecological environment has improved significantly.

In response to question 2, we first use the Spearman correlation coefficient to test the ecological environment of Saihanba and Beijing′s ability to resist sand and dust, and find that the two have strong convergence, indicating that the indicator set in question 1 can be used to represent Saihanba. Next, establish a multiple regression model of the relationship between Saihanba′s ecological indicators and Beijing′s ability to resist sand and dust. It is found that the restoration of Saihanba Forest Farm has a significant effect on Beijing's ability to resist sand and dust.

Aiming at problem three, we first establish a BP neural network model between ecological evaluation indicators and three ecological capabilities based on the data of Saihanba. Then, the comprehensive ecological index is obtained through the weighted average of these three ecological capacity factors, and the scale of the ecological area that needs to be established in the place is judged through the negative index equation. Then, the carbon absorption capacity of the area was recalculated and compared with the carbon absorption capacity before construction. It was found that the carbon absorption capacity of each area had a significant upward trend.

In response to question 4, we chose Australia, which accounts for a large amount of desert area, to analyze and divide it into 7 regions. Based on the model established in question 3, we calculated the actual scale of ecological regions that can be built in each region of Australia and the carbon absorption after construction. The conclusion of the change trend of ability is similar to the third question.

Finally, we also wrote a non-technical report to describe the built model and put forward feasible plans and suggestions for the establishment of ecological reserves.

Keywords: Entropy method; TOPSIS; Spearman correlation coefficient; Multiple regression model; BP Neural network

Contents

1 Introduction

1.1 Background

We all know that to build a beautiful China, we must adhere to the concept that green waters and green mountains are golden mountains and silver mountains, and sustainable development. China's Saihanba Forest Farm is a good example. With the help of the Chinese government, the Saihanba Forest Farm has been transformed from a desert before its establishment into an oasis today. Compared with 1962, the forest coverage of Saihanba has increased from 11.4% to 80% in more than half a century. Above, Saihanba can absorb 747,000 tons of carbon for Beijing and Tianjin every year. It can also effectively resist sandstorms and conserve water. And it can be seen that the afforestation history of Saihanba has far-reaching developmental significance. Therefore, we can build models based on the historical data of Saihanba and reflect the mission of prospering ecology in more places.

1.2 Problem requirements

(1) Select appropriate indicators and collect relevant data, establish an evaluation model of Saihanba's impact on the ecological environment, and analyze the environmental conditions before and after Saihanba's restoration.

(2) Select appropriate indicators and collect relevant data to establish an evaluation model of Saihanba's impact on Beijing's Anti-dust ability, and evaluate Saihanba's influence on Beijing's Anti-dust ability.

(3) Collect data to establish a mathematical model, extend Saihanba's ecological model to the whole of China, determine where the ecological zone needs to be established, and determine the scale of the proposed ecological zone.

(4) Choose another country in the Asia-Pacific region to build a mathematical model and collect data, analyze which locations in this country need to establish an ecological zone, determine the scale of the proposed construction, and evaluate its impact on the absorption of greenhouse gases and the reduction of carbon emissions.

2 Problem analysis

2.1 Analysis of Problem One

In order to analyze the environmental conditions before and after the restoration of Saihanba, we can objectively assign an evaluation score to the two for comparison. After collecting relevant indicators and data, we can use interpolation to fill in missing data, and use entropy to find several important indicators. Then, we can use the TOPSIS method to evaluate the ecological environment of Saihanba through these weighted indicators, and finally get the annual ecological environment score of Saihanba from 1962 to 2020, and conduct a comparative analysis.

2.2 Analysis of Problem Two

Problem Two requires us to study the relationship between Saihanba's recovery and Beijing's Anti-sand dust ability. We need to find a new set of indicators to evaluate this pair of relationships. But before that, we can also try to test the correlation between Saihanba's ecological environment and Beijing's Anti-sand dust ability. If the correlation between the two is good, we can use five indicators for evaluating Saihanba's ecological environment to evaluate Beijing's Anti-sand dust ability. Next, we can roughly determine a multiple regression equation based on the line chart of Beijing's Anti-sand dust ability over time, and finally establish a mathematical model of the relationship between the recovery of Saihanba and Beijing's Anti-sand dust ability.

2.3 Analysis of Problem Three

In order to determine the scale of the establishment of the ecological zone, we can set a comprehensive ecological index for evaluation. The comprehensive ecological index can be determined by three ecological ability factors, such as the Anti-sand dust ability, Carbon absorption capacity and Water purification capacity. We can establish a model of the relationship between evaluation indicators and ecological capabilities through the BP Neural network, and then we can use the weighted average of the three ecological ability factors to determine the comprehensive ecological indicators. After obtaining the comprehensive ecological indicators, we can use the corresponding equations to judge whether the area needs to be established and the scale of the ecological area. Assuming that the ecological area is successfully completed in accordance with the construction plan, we can recalculate the Forest coverage of the area and bring it into the Neural network model to find the Carbon absorption capacity after the ecological area is constructed, and then combine it with the ecological area. The previous Carbon absorption capacity is compared, and a conclusion is drawn on the impact of the ecological zone on carbon neutrality.

2.4 Analysis of Problem Four

Problem Four is similar to Problem Three, we only need to find the relevant indicator data of a country in the Asia-Pacific region and bring it into the model in Problem Three to draw conclusions.

3 Model assumptions

(1) Assuming that the change in Beijing's Anti-sand dust ability is only related to the ecological changes in Saihanba.

(2) The change trend of Beijing's Anti-sand dust ability is logarithmically related to the evaluation index.

(3) It is assumed that the establishment of Saihanba Ecological Zone will have the same impact on the environment as other regions in the world.

(4) Assuming that when ecological zones are established in various administrative regions in China, the weights of ecological capabilities as evaluation factors are consistent with those of Saihanba.

(5) Suppose that in the initial stage of the establishment of the ecological zone, a larger scale of ecological zone needs to be built to significantly change the ecological environment; as the scale of the ecological zone gradually increases, the rate of improvement of the ecological environment also increases accordingly.

(6) It is assumed that all the ecologically-developed land planned by the government can be used to build an ecological zone.

4 Symbol description

Symbol	Description	Unit
V_{ij}	Normalized standard matrix	/
P_{ij}	Under the index i, the characteristic proportion of the j-th evaluation object	%
e_j	Entropy of the j-th indicator	/
d_j	The coefficient of variance for the index j	%
w_j	Entropy weight of the index j	%
Nh	Number of hidden layer neurons	PCS
N_i	Number of neurons in the input layer	PCS
No	Number of neurons in the output layer	PCS

5 Model building and solution of question 1

In order to analyze the environmental conditions before and after the restoration of Saihanba, we can objectively assign an evaluation score to the two for comparison. Among them, the TOPSIS method is a good objective evaluation method. TOPSIS is a systematic evaluation method suitable for multi-index analysis. We can collect a series of indicators representing the ecological environment. By calculating the weighted Euclidean distance between the relevant indicators and the positive and negative ideal solutions for each year, we can obtain the indicators and positive ideals for each year. The closeness of the solution is used as the basis for evaluation.

5.1 Data preprocessing

First, we need to select the representative factors that affect the ecological environment in the Saihanba area as the index set. By consulting relevant information, we selected 10 indicators of Forest coverage, Air quality index, Water quality index, Soil pollution index, Biological abundance index, Fixed carbon dioxide, Oxygen released, Forest tree species structure, Climate index, and Ecological resilience[1].

We tried to collect data on these indicators in the Saihanba area since 1962, but we found a problem. It is difficult for us to accurately find the historical data of the Saihanba area over a long period of time. In order to establish a more reasonable model, we consider Use the collected data to interpolate the unknown data.

We can use cubic spline interpolation algorithm to smoothly connect missing data. Let $S(x)$ be a cubic spline interpolation function, and its construction method has the following form:

$$S(x)=\begin{cases} S_0(x), & x\in[x_0,x_1], \\ S_1(x), & x\in[x_1,x_2], \\ \quad\vdots \\ S_{n-1}(x), & x\in[x_{n-1},x_n]; \end{cases} \quad S_i(x)\in C^3([x_i,x_{i+1}]) \tag{1}$$

At the same time meet the conditions:

$$\begin{cases} S(x_i)=f_i, \\ S_{i-1}(x_i)=S_i(x_i), \\ S'_{i-1}(x_i)=S'_i(x_i), \\ S''_{i-1}(x_i)=S''_i(x_i); \end{cases} \tag{2}$$

In summary, we can get complete and more accurate data. The complete data can be found in the appendix. Among them, the results of the interpolation algorithm for forest coverage are shown in the figure below:

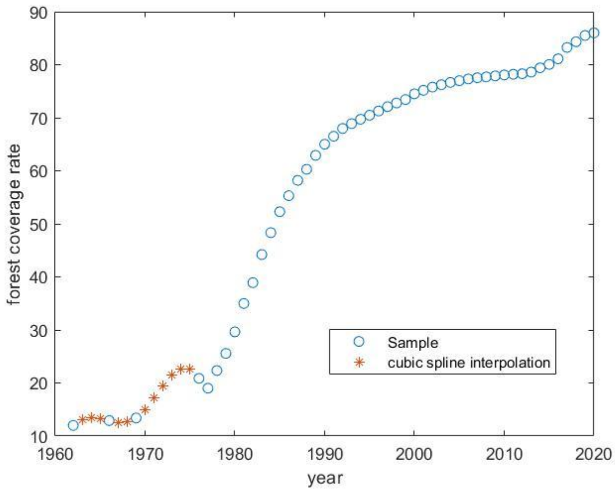


Figure 1:Forest coverage rate

5.2 Dimensionality reduction of indexes by entropy method

After obtaining the data of 10 indicators, we found that if all these indicators are used to evaluate the ecological environment of Saihanba, the complexity of the model will be slightly larger. In order to simplify the model, we can extract several of the most representative characteristic indicators from these 10 indicators as evaluation criteria.

Among them, the entropy method is a method that can objectively assign weights, which can reflect the effective information amount of the indicator. The larger the entropy value, the smaller the effective information amount of the indicator, and the smaller the weight of the indicator. Compared with the subjective weighting method, the entropy weight method is more explanatory for indicators. Therefore, we can use the entropy method to more accurately extract the most representative feature indicators among the above 10 indicators.

5.2.1 Establish and standardize the decision matrix

Firstly, We will select the years from 1962 to 2020 as the object set M_i, Multiple indicators are used as indicator set D_j. The specific data of different indicators in 59 years were selected horizontally to form the decision matrix.

Secondly, we need to divide the indicators into two categories according to their nature: benefit indicators and cost indicators. The so-called benefit index actually indicates that the larger the index is, the better; The cost index indicates that the smaller the index is, the better. Among these 10 indicators, we can divide them into the following categories:

- **Benefit indicators:**
  - **Forest coverage rate:** Forests play an important role in conserving water sources and conserving water and soil. It is easy to know that the higher the forest coverage rate, the more beneficial the impact on the terrestrial ecosystem with forests as the main body.
  - **Water quality index:** This index can reflect the quality of water with a dimensionless

value by integrating the test results of multiple water quality parameters, including ph, dissolved oxygen and other indicators, which is a reliable indicator to measure water quality.

❖ **Biological abundance index:** It reflects the abundance and poverty of species in the evaluated area. The higher the species abundance, the stronger the anti-disturbance ability of the ecosystem.

❖ **The amount of released oxygen:** This indicator can reflect the amount of oxygen released by the forest or vegetation in the evaluation area in a certain period of time. The higher the amount of oxygen released, the higher the ecological and environmental benefits.

❖ **Vegetation Index:** This index is a simple, effective and empirical measure of the vegetation status on the ground, reflecting the difference between the reflection of vegetation in the visible light and near-infrared bands and the soil background[2].

❖ **Climate index:** refers to a quantity that is composed of two or more climatic elements and represents a certain climatic characteristic. It includes aridity index, humidity index, monsoon index, and degree of land. It is mainly used for climate classification and zoning.

❖ **Ecological resilience:** also known as resilience, refers to the ability of the ecosystem to maintain the structure and pattern, that is, the ability of the system to restore its original function after being disturbed. The greater the ecological resilience, the faster the restoration of the ecosystem.

➢ **Cost indicators:**

❖ **Soil pollution index:** an index used to quantitatively describe the degree of soil pollution, the higher the degree of soil pollution, the greater the index.

❖ **Air Quality Index:** Reflects the concentration of pollutants (smoke, NO_2, SO_2, etc.) in the air. The greater the degree of air pollution, the higher the index.

❖ **Soil erosion rate:** refers to the ratio of land surface erosion and water and soil loss to the area of the evaluation area, and is a macro management indicator that reflects the overall situation of soil and water conservation.

For the benefit index, we use the following formula to normalize:

$$v_{ij} = \frac{\max\{x_{1j}, x_{2j}, \cdots, x_{nj}\} - x_{ij}}{\max\{x_{1j}, x_{2j}, \cdots, x_{nj}\} - \min\{x_{1j}, x_{2j}, \cdots, x_{nj}\}} \tag{3}$$

For cost indicators, we use the following formula to normalize:

$$v_{ij} = \frac{x_{ij} - \min\{x_{1j}, x_{2j}, \cdots, x_{nj}\}}{\max\{x_{1j}, x_{2j}, \cdots, x_{nj}\} - \min\{x_{1j}, x_{2j}, \cdots, x_{nj}\}} \tag{4}$$

According to the above formula, we can finally get the normalized standard matrix.

5.2.2 Calculate feature weights and establish entropy values

❖ **Calculate feature weights:**

In the standard matrix, for a certain index j, the larger the value of V_{ij} is, the more useful information the index can provide to the evaluated object. We calculate the feature weight P_{ij} by the following formula:

$$P_{ij}=\frac{v_{ij}}{\sum_{i=1}^{m} v_{ij}}(0\leqslant v_{ij}\leqslant 1, 0\leqslant p_{ij}\leqslant 1) \tag{5}$$

❖ **Calculate the entropy value:**

The entropy value e_j of the indicator is calculated by the following formula, the closer the entropy value is to 1, the more stable the column is.

$$e_j=-\frac{1}{\ln m}\sum_{i=1}^{m} p_{ij}\ln(p_{ij}) \tag{6}$$

❖ **Calculate the coefficient of difference:**

The greater the difference coefficient d_j is, the greater the amount of information provided by the index, and a higher index weight should be given.

$$d_j=1-e_j \tag{7}$$

❖ **Determine the index entropy weight:**

$$w_j=\frac{d_j}{\sum_{k=1}^{n} d_k}, j=1,2,\cdots,n \tag{8}$$

Through the above formula, the entropy weight can be obtained by dividing the difference coefficient calculated in the previous step by the sum of the difference coefficient of all indicators.

5.2.3 Solution of Entropy Method

We import the data into ***matlab*** to solve, sort the results of the entropy weight of many indicators, and extract the top 5 indicators with the most influence as our final indicator set according to the fault condition of the entropy value. The weights of the indicators are normalized to obtain the following weight matrix:

Table 1: Index weight matrix

Ranking	Indicators	Entropy weight
1	Forest coverage rate	0.3564
2	Air Quality Index	0.2529
3	Water quality index	0.1776
4	Soil pollution index	0.1399
5	Biological abundance index	0.0732

It can be seen that the Forest coverage rate is the most important for evaluating the quality of the environment, and its entropy weight reaches 0.35, followed by Air quality, Water quality, and Soil quality. Finally, it can be seen that the diversity of species has an impact on the evaluation of the environment. The quality of it also has a slight impact.

5.3 TOPSIS method to evaluate Saihanba's environmental impact

After obtaining the relevant indicators and data, we can use the TOPSIS method to calculate the annual environmental score of Saihanba since 1962. By comparing the environmental scores over the years, we can evaluate the impact of Saihanba on the environment.

5.3.1 Data set processing

We set the five indicators obtained in the previous step of dimensionality reduction as indicator set $D=(D_1,D_2,\cdots,D_n)$, and the situation of Saihanba from 1962 to 2020 as object set $M=(M_1,M_2,\cdots,M_m)$. The relevant data of each indicator from 1962 to 2020 is recorded as $x_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$.

Similarly, in order to eliminate the influence of dimensions, we can also normalize the data through benefit indicators and cost indicators. Through the two normalization formulas (3) (4), the standardized matrix is finally obtained.

Since the weight of each indicator is different, we can use the result obtained by the entropy weight method in 5.2.3 as the weight matrix W_j of indicator set $D=(D_1,D_2,\cdots,D_n)$. We multiply the weight matrix W with the standard matrix obtained in the previous step to obtain the weighted decision matrix:

$$\boldsymbol{R}=(r_{ij})_{m\times n}:\quad r_{ij}=w_j\cdot v_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n) \tag{9}$$

5.3.2 Establishment of TOPSIS model

(1) Calculate positive ideal solution and negative ideal solution

First, we need to use the weighted decision matrix obtained in the previous step to first calculate the positive and negative ideal solutions. The so-called positive ideal solution is a set consisting of the maximum value of the benefit index and the minimum value of the cost index. It can also be called the best solution. The negative ideal solution is the opposite:

$$S_j^+=\begin{cases}\max\limits_{1\leq i\leq m}\{r_{ij}\},j=1,2,\cdots,n;\ D_j\ \text{is the benefit indicator}\\ \min\limits_{1\leq i\leq m}\{r_{ij}\},j=1,2,\cdots,n;\ D_j\ \text{is the cost indicator}\end{cases} \tag{10}$$

The formula above is the positive ideal solution, and the formula below is the negative ideal solution.

$$S_j^- = \begin{cases} \min\limits_{1\leq i\leq m}\{r_{ij}\}, j=1,2,\cdots,n;\ D_j \text{ is the benefit indicator} \\ \max\limits_{1\leq i\leq m}\{r_{ij}\}, j=1,2,\cdots,n;\ D_j \text{ is the cost indicator} \end{cases} \tag{11}$$

(2) Calculate Euclidean distance

Next, we calculate the Euclidean distance between each object and the positive and negative ideal solutions by the following formula:

$$\begin{cases} Sd_i^+ = \sqrt{\sum_{j=1}^{n}(S_j^+ - r_{ij})^2}, i=1,2,\cdots,m \\ Sd_i^- = \sqrt{\sum_{j=1}^{n}(S_j^- - r_{ij})^2}, i=1,2,\cdots,m \end{cases} \tag{12}$$

(3) Calculate the relative closeness between object and the positive ideal solution

In order to explore the impact of changes in Saihanba on the ecological environment, we need to calculate the relative closeness η_i between the indicators of each year and the positive ideal solution, and evaluate the pros and cons of the ecological environment according to η_i excellent.

The expression for η_i is as follows:

$$\eta_i = \frac{Sd_i^-}{Sd_i^+ + Sd_i^-}, i=1,2,\cdots,m \tag{13}$$

5.3.3 Solution of TOPSIS model

Next, we import the sorted data, use $matlab$ programming to solve the established model, we can get the environmental score of Saihanba over the years, as shown in the following figure:

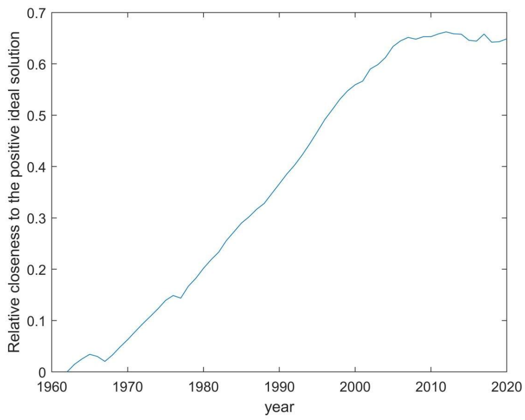


Figure 2: The relative closeness between the set of indicators and the positive ideal solution from 1962 to 2020

6 Establishment and solution of Problem Two Model

6.1 Model preprocessing

6.1.1 Handling of problems

Promblem 2 requires us to study the relationship between the restoration of Saihanba and the ability of Beijing to resist sand and dust, and establish a mathematical model that can quantitatively evaluate the impact of Saihanba on the ability of Beijing to resist sand and dust. At this time, we need to find a new set of indicators to evaluate this pair of relationships, but this is inefficient, and cumbersome data collection will increase the difficulty of building the model. In view of the fact that we have found five indicators of Saihanba's impact on the ecological environment in Promblem 1, we can try to test the correlation between Saihanba's ecological environment and Beijing's ability to resist sand and dust, if the correlation between the two is good, We can use the five indicators to evaluate the ecological environment of Saihanba to evaluate Beijing's ability to resist sand and dust.

Among them, the Spearman correlation coefficient is a method to study the convergence between two groups of variables. It does not require high sample size, and does not need to assume the normality of the population, and good results can be obtained. Therefore, here we use the Spearman rank correlation coefficient to test.

The definition of Spearman correlation coefficient is as follows:

$$\begin{cases} r_{SP} = 1 - \dfrac{6\sum\limits_{i=1}^{n} d_i^2}{n(n^2-1)} \\ t = r_{SP}\sqrt{\dfrac{n-2}{1-r_{SP}^2}} \sim t(n-2) \end{cases} \tag{14}$$

Where n is the sample size and d_i is the difference between the levels of the two sets of data. When $|t| \geqslant t_{1-\alpha/2}$, the null hypothesis is rejected: there is no correlation between the two groups of variables studied, that is, there is a significant convergent relationship between the two variables.

The test results are as follows:

Table 2: The test results

Spearman rank correlation ρ	p-value
0.9680	6.4030e-36

We can see that the value of ρ is as high as 0.9680, and p-value is almost 0, indicating that the ability to resist sand and dust storms has a significant convergent relationship with the impact of the first question on the ecological environment. We can use the five indicators in question one for subsequent modeling.

6.1.2 Data processing

In order to facilitate the establishment of the model, we need to quantify the two abstract concepts of Saihanba's annual restoration process and the ability to resist sand and dust: we use the five indicators obtained in Problem 1 (Forest coverage, Air quality index, Water quality index, Soil Pollution index and Biological abundance index) during the period of 1962-2020 to represent the annual recovery process of Saihanba; for the Anti-dust ability, we also based on the frequency and severity of sandstorms that occurred in Beijing each year from 1962 to 2020. Quantify the abstract index of Anti-sand dust ability into concrete data, and finally organize all the data into an excel table.

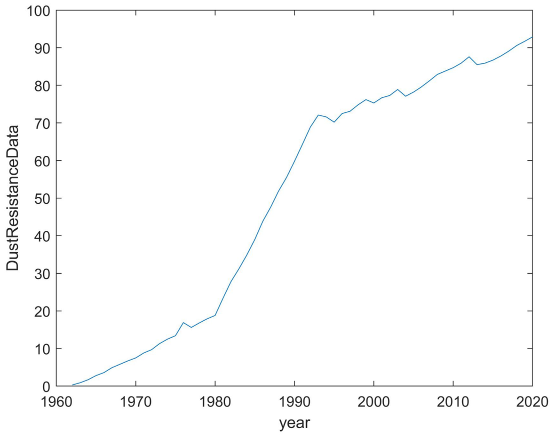


Figure 3:Variation of dust resistance in Beijing from 1962 to 2020

By observing the above figure, we found that the change trend of Beijing's Anti-dust ability index is similar to the logarithmic growth form, so we tried to use the logarithmic form of the function to establish a function that can reflect the recovery and Anti-sand dust of Saihanba Mathematical model of the relationship between abilities.

6.2 Establishment and solution of multiple regression model

We define the ability to resist sand and dust as y^D and the index set as $G(x^f, x^w, x^b, x^s, x^a)$. The variables x^f, x^w, x^b, x^s, x^a represent Forest coverage, Water quality index, Biological abundance index, Soil pollution index and Air quality index.

Assuming that the change in Beijing's ability to resist sand and dust is only related to the ecological changes in Saihanba, according to the obtained data, we take the five indicators of Forest coverage, Air quality index, Water quality index, Soil pollution index, and Biological abundance index as self Variables, take Beijing's ability to resist sand and dust as the dependent variable, and establish a multiple regression model that can reflect the relationship between them.

As mentioned above, the change trend of the ability to resist sand and dust is very similar to the logarithmic function. We try to build a model based on this, and the form is as follows:

$$y^D = \alpha_0 + \alpha_1 lnx^f + \alpha_2 lnx^a + \alpha_3 lnx^w + \alpha_4 lnx^s + \alpha_5 lnx^b \tag{15}$$

We can set:

$$\begin{aligned} Q &= \sum_{i=1}^{5} \left(y^D - \overline{y^D}\right)^2 \\ &= \sum_{i=1}^{5} \left(y^D - (\alpha_0 + \alpha_1 lnx^f + \alpha_2 lnx^a + \alpha_3 lnx^w + \alpha_4 lnx^s + \alpha_5 lnx^b)\right) \end{aligned} \tag{16}$$

Let Q be the smallest to get an equation with the best fit, and then let:

$$\frac{\partial Q}{\partial \alpha_i} = 0(i = 0, 1, 2, 3, 4, 5) \tag{17}$$

We can use matlab programming calculation to find the value of each coefficient:

Table 3: The value of each coefficient

parameter	α_0	α_1	α_2	α_3	α_4	α_5
estimated value	81.8011	22.027	-15.455	25.376	0.912	-30.199

Then bring in the value of each parameter to get the final mathematical model:

$$y^D = 81.8011 + 22.02761 lnx^f - 15.45556 lnx^a + 25.3765 lnx^w + 0.9127 lnx^b - 30.1999 lnx^s \tag{18}$$

Through the above model, it can be found that the coefficients before the variables lnx^f and lnx^w are as high as 22.0276 and 25.3765. Sand and dust capabilities have also been significantly improved. The coefficients of the variables lnx^a and lnx^s are -15.45556 and -30.1999, indicating that as the degree of Air pollution and Soil pollution in Saihanba decreases, Beijing's ability to resist sand and dust has increased significantly.

According to the nature of the logarithmic function, we can know that when the index set $G(x^f, x^w, x^b, x^s, x^a)$ just started to change, that is, when the Saihanba project was initially constructed, Beijing's Anti-dust ability was still relatively slow. However, in the middle of construction, the ability to resist sand and dust began to increase significantly, and the frequency and severity of sandstorms in Beijing were significantly reduced. Now, the Forest coverage rate is close to 86%, gradually reaching a state of saturation, and Beijing's Anti-dust ability has also begun to slow down and gradually maintained at a relatively good value.

6.3 Model checking

We use hypothesis testing methods for the obtained model, use the coefficient of determination test, F test and T test to test the model, and draw a fitting effect graph to observe the fitting effect of the model.

(1) Coefficient of determination test and F test

Table 4: Coefficient of F test

Source of Variance	Degree of freedom	sum of square	Mean square	F-value	p-value
Return	5.000000	62044.03	12408.81	654.22	0.000000
Residual	53.00000	1005.27	18.97		
Total	58.000000	63049.31			

We can see that the p-value of the F test is 0, rejecting the null hypothesis: the coefficients of all independent variables are 0, and the F test is passed. Moreover, the value of the determination coefficient is as high as 0.9826, indicating that the fitting effect of the model is still relatively good.

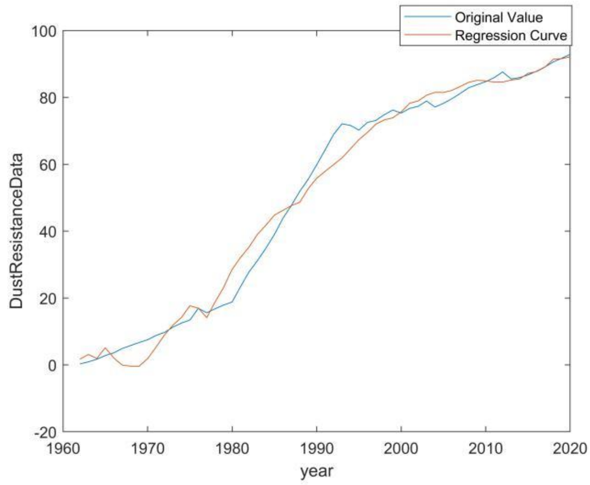
(2) T test

Table 5: Coefficient of T test

Variable	estimated value	Standard error	t-value	p-value
Constant term	81.8011	63.9919	1.2783	0.0267
$\boldsymbol{x}^{f}$	22.0276	2.9625	7.4356	0.0000
$\boldsymbol{x}^{a}$	-15.4556	14.3875	-1.0742	0.0876
$\boldsymbol{x}^{w}$	25.3765	3.2583	7.7883	0.0000
$\boldsymbol{x}^{b}$	0.9127	2.6677	0.3421	0.0336
$\boldsymbol{x}^{s}$	-30.1999	3.9228	-7.6986	0.0000

From the above table, we can see that 3 variables of the model: $\boldsymbol{x}^{f}, \boldsymbol{x}^{w}, \boldsymbol{x}^{s}$,the p-value of the t-test is all equal to 0, and the values of the remaining variables and constant terms are also less than 0.10. The more excellent ones have passed the t-test, indicating The Linear Relationship between the independent variable and the dependent variable of the model is significant, and the model fitting effect is better.

(3) Fitting effect diagram



Figuer 4:Model fitting effect diagram of "impact of Saihan dam on anti dust capacity of Beijing from 1962 to 2020"

By observing the regression curve and the original value, we can see that around 1970, the fitted value deviates from the original value trend. As mentioned above, our original value Anti-dust ability is based on the number and severity of sandstorms that occurred in Beijing that year. We suspect that the climate of Beijing during this time period reduced the frequency of sandstorms, but the actual Anti-dust ability is not as high as the original value.

Looking at the information, we found that Beijing experienced 114 consecutive days of drought-free rain from 1970 to 1971[3]. The surface was loose and the surface was loose. When strong winds passed, it was very easy to produce sandstorms. Fortunately, there were no strong winds in Beijing during this period, which reduced the occurrence of sandstorms. The frequency of resulting in the false high of the original value. Instead, the regression curve we fitted is more able to reflect Beijing's sudden drop in Anti-dust ability at that time.

From 1981 to 1996, the regression curve was slightly higher than the original value before 1988, and slightly lower after 1988, but the fitting effect is still good, and the unconsidered climate may be one of the reasons for the error[4]. After 1996, there was almost no deviation between the regression curve and the original value.

On the whole, the fitting effect of this model is still relatively good.

7 Establishment and solution of Problem Three Model

7.1 Model building ideas

In Problem 1, we have found the impact of Saihanba on the local ecological environment through relevant indicators. In order to extend Saihanba’s ecoloehensive Ecological indicators to judge these geographic locations. The comprehensive ecological index can be determined by three ecological capacity factors, such as the ability to resist sand and dust storms, as well as the carbon absorption capacity and the water purification capacity described in Problem 2.

The specific modeling steps are as follows:

- ✓ **Step1:** We can build a model of gical protection model to the whole country and determine which geographical locations need to build ecological zones, we can use a comprthe relationship between evaluation indicators and ecological capabilities based on the data of Saihanba over the years, and then we can use the weighted average of the three ecological ability factors to determine the comprehensive ecological indicators.
- ✓ **Step2:** Divide the geographical location into 34 administrative regions in China, collect the relevant evaluation indicators of these administrative regions, and obtain the comprehensive ecological indicators of each administrative region. Through comprehensive ecological indicators, we can judge whether it is necessary to build an ecological zone in this area, and if it needs to be built, go to the next step.
- ✓ **Step3:** Establish a relationship between the comprehensive ecological indicators and the scale of the ecological zone through an equation to determine the proportion of the ecological zone area that needs to be increased in the administrative area. Then, we compare it with the maximum ecological area that can be planned in the administrative area. If the area that can be planned is greater than the area that needs to be increased, the actual area that can be increased is the area that needs to be increased. Otherwise, the ecological area will be based on the largest area. The area can be planned for construction.
- ✓ **Step4:** Assuming that the ecological zone is successfully completed in accordance with the construction plan, we can recalculate the forest coverage of the area and bring it into the previously established model to solve the carbon absorption capacity of the area after the ecological zone is constructed, and compare it with The carbon absorption capacity of the ecological zone was compared before the construction, and the conclusion of the ecological zone's impact on carbon neutrality was reached.

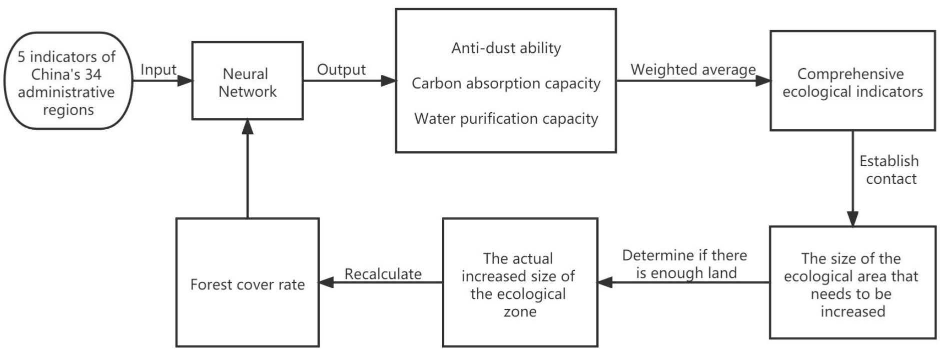


Figure 5:Step flow chart

7.2 The establishment of BP Neural Network

In Step 1, we need to establish a model of the relationship between evaluation indicators and ecological capabilities. At this time, the evaluation indicators can be the five factors in question 1: Forest coverage rate, Air quality index, Water quality index, Soil pollution index and Biological abundance Index; and the Ecological ability can be the ability to resist sand and dust storms, the ability to absorb carbon, and the ability to purify water. Therefore, we need to find out the mapping relationship between these two pairs of indicators, namely:

$$f: (x_1, x_2, x_3, x_4, x_5) \rightarrow (y_1, y_2, y_3) \tag{19}$$

This is a mapping relationship from 5 dimensions to 3 dimensions. We cannot simply use multiple regression to establish a suitable mathematical model. In the field of artificial intelligence, deep learning solves this problem well. Therefore, we decided to use BP Neural Network to achieve this mapping relationship.

7.2.1 Neural network design

Since the data we collect is non-linear, and in neural networks, if the data needs non-linear separation, there must be a hidden layer. Therefore, we can divide the architecture of the neural network into an input layer, a hidden layer, and an output layer. The input layer is mainly used to obtain input information; the hidden layer is mainly used for feature extraction, so that the neurons in the hidden layer react to a certain pattern; the output layer is used to receive the hidden layer and output the model results, which can be adjusted by adjusting the weights. Neuron stimulation forms a more superior response, in which the excitability of the output is the result.

In the neural network model we built, the input layer is forest coverage, air quality index, water quality index, soil pollution index and biological abundance index, and the output layer is the three indicators of sandstorm resistance, carbon absorption and water purification. For the hidden layer, the deeper the number of network layers, the stronger the ability to fit the function theoretically, and the better the effect, but in fact, the deeper the number of layers may also cause over-fitting problems and increase the difficulty of training, which makes the model difficult to converge[5]. Therefore, when determining the number of network layers, we only give priority to the case of one hidden layer. If the result of the model training is not ideal, we will appropriately add the number of hidden layers.

The number of neurons in the hidden layer is also a very particular variable. If too few neurons are used in the hidden layer, the model will underfit, and too many neurons may also lead to overfitting. Therefore, choosing an appropriate number of hidden layer neurons is crucial. However, we do not have an accurate way to determine this value. We can only rely on empirical formulas to determine the initial value of the number of neurons, and finally dynamically adjust the number of neurons according to the fitting results. In this model, the empirical model we propose is:

$$Nh = \frac{2}{3}N_i + No + \alpha \tag{20}$$

Among them, nh represents the number of neurons in the hidden layer, ni represents the number of neurons in the input layer, no represents the number of neurons in the output layer, and α can be any integer, generally starting from 0 and going to both ends. Since in this model, there are 5 neurons in the input layer and 3 neurons in the output layer, we can initially set the number of neurons in the hidden layer to 7.

In addition, when solving nonlinear problems, we need to select an appropriate activation function to introduce nonlinear features into the network. If we don't use the activation function, then the output signal will be just a simple linear function, and its ability to learn more complex function mapping from the data is even smaller. In addition, since the number of network layers in this model is not very deep, we can use the traditional tansig activation function.

Finally, the Neural network model we get is as follows:

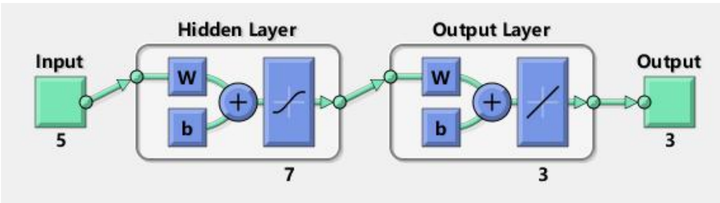


Figure 6:Neural network model

7.2.2 Neural network solution

We import the collected data into $matlab$, use 75% of the index data for training, 15% of the index data for verification, and 15% of the index data for testing. As the number of iterations increases, the model gradually tends to converge. When the Neural network is iterated to 52 times, the optimal value of the Neural network parameters can be found and output as the result.

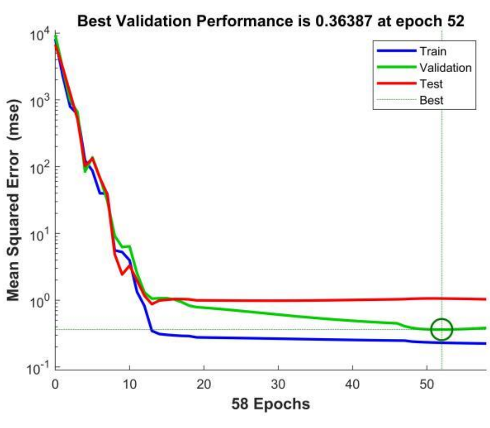


Figure 7:Index data training results

Next, we test the results of the model training, we can get the fitting effect of all the data sets, and find that the R value of the model is very close to 1, and we can think that the fitting effect of the neural network model is perfect.

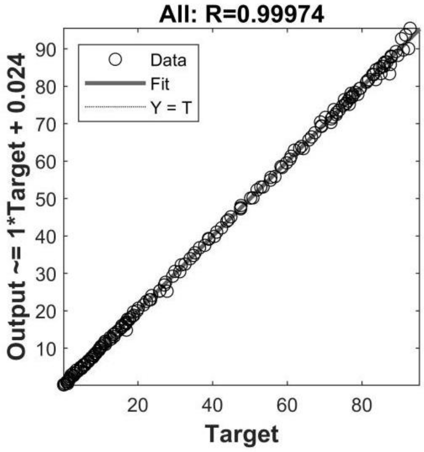


Figure 8:Fitting effect of all data sets

Therefore, we can bring the index data of 34 administrative regions into the neural network model for prediction, and we can get the anti-sand storm ability, carbon absorption ability and water purification ability of these administrative regions. Part of the data is shown in the following figure:

7.3 The location where the ecoregion was established

In order to explore whether these 34 administrative regions need to build ecological zones, we need to have an exact judgment basis, so we can set a comprehensive ecological index to make judgments.

In the previous section, we have calculated the three ecological abilities of 34 administrative regions, and the comprehensive Ecological index can be determined by the three ecological abilities. The easiest way is to take the average of the three. . However, since these three indicators have slightly different assessments of the local ecological environment, we must distinguish the importance of these three indicators.

We can assign a specific weight to various ecological abilities according to the establishment of existing ecological regions in China. By collecting relevant data, we found that the most important capacity for the establishment of the ecological zone is the capacity to absorb carbon, followed by Anti-dust ability and Water purification capacity. We can assign values to these indicators:

Table 6:Indicator assignment

	Anti-dust ability	Carbon absorption capacity	Water purification capacity

Weight	0.3	0.45	0.25

Therefore, the relationship between comprehensive ecological indicators and ecological capabilities is:

$$\beta = 0.3Y_1 + 0.45Y_2 + 0.25Y_3 \tag{21}$$

Through this relationship, we can find the comprehensive ecological indicators of 34 administrative regions, as shown in the following figure and table:

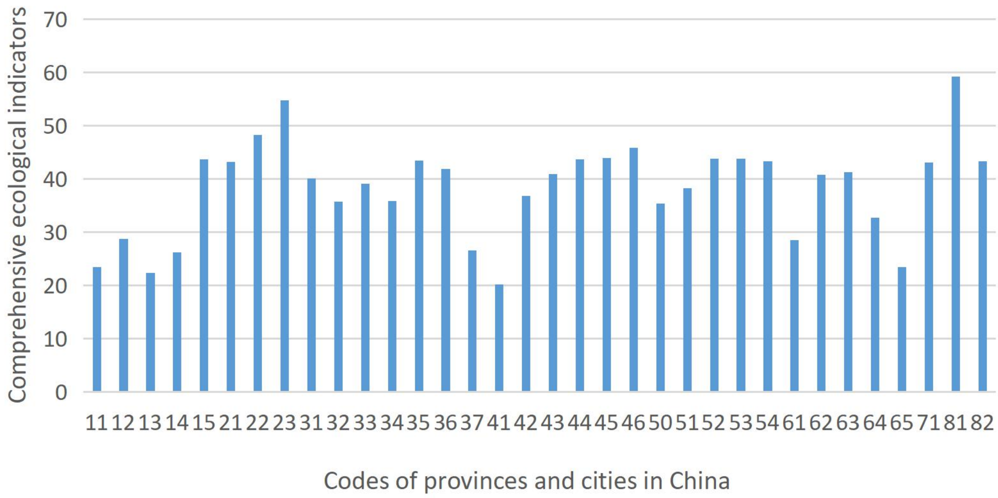


Figure 9:The comprehensive ecological indicators of 34 administrative regions

When the comprehensive ecological index is greater than 50, we can think that there is no need to build an ecological zone in the area, and when the comprehensive ecological index is less than 50, we think that the area needs to build a new ecological zone. We set the variable w as whether a new ecological area needs to be built in the area. A value of 1 indicates that it needs to be built, and a value of 0 indicates that it does not need to be built. Then the mathematical model can be expressed as:

$$w_i = \begin{cases} 1, & \beta_i \leq 50 \\ 0, & \beta_i > 50 \end{cases} \tag{22}$$

Therefore, it can be found from the above formula that there are 32 administrative regions in China that need to build ecological zones.

7.4 The scale of the establishment of the ecological zone

Next, we need to determine the scale of the ecological zone they need to build through the comprehensive ecological indicators of each administrative region. However, due to the inconsistency of the size of the various administrative regions, in order to facilitate the identification of the connection between the two, we can use the proportion of the ecological area that the administrative region needs to increase as a measurement standard.

Since in the initial stage of establishing an ecological zone from scratch, it is necessary to build a larger scale of ecological zone to have an impact on comprehensive ecological indicators. Let S_i be the proportion of ecological area that needs to be increased, and the function of S_i with respect to β_i is in the defined domain [0,50] Is continuous and derivable, that is, the constraint condition can be expressed as:

$$\frac{dS}{d\beta_0} > \frac{dS}{d\beta_1}(\beta_0 < \beta_1) \tag{23}$$

We can think that the equation of S_i with respect to β_i is a concave function in the domain of definition, and try to establish the connection between the two using an exponential equation. The preliminary model is as follows, where a and b are undetermined coefficients:

$$S_i = b(a^{-\beta_i}) \tag{24}$$

In order to determine the two undetermined coefficients a and b, we can perform the following analysis: When the comprehensive ecological index is 0, it means that the ecological environment of the area is extremely bad. By consulting relevant information, we can see that we will need at least 40% of the area to build an ecological zone in order to significantly improve the local ecological environment[6]. Only by building an ecological zone can the local ecological environment be significantly improved. At the same time, we believe that once the ecological zone is determined to be established, it is more cost-effective to build at least 5% of the area. Therefore, when the comprehensive ecological index is 80, we also believe that at least 5% of the area is needed to build an ecological zone.

That is, when $\beta_i = 0, S_i = 40$ and $\beta_i = 50, S_i = 5$. Substituting into (x) formula can be solved to get $a = 1.026, b = 40$, namely:

$$S_i = 40(1.042^{-\beta_i}) \tag{25}$$

In addition, we can also extend the domain of the model to [0,100], which is:

$$S_i = 40w_i(1.042^{-\beta_i}) \tag{26}$$

At this point, we can calculate the scale of the ecological zone that needs to be built in each administrative region, as shown in the following figure:

Table 7: The scale of the ecological zone that needs to be built in each administrative region

	Beijing	Tianjin	Hebei	Shanxi	Liaoning	Jilin	Heilongjiang
Anti-dust ability	45.39	45.85	35.68	45.34	50.65	60.95	66.54
Carbon absorption capacity	13.10	20.22	15.86	15.61	38.86	41.33	48.14

Water purification capacity	15.96	23.75	18.10	22.43	42.07	45.38	52.72

However, each administrative area is not completely consistent with regard to the free land planning. For example, areas such as Beijing, Shanghai, and Hong Kong have higher economic levels and denser populations. The remaining free land in these areas is less, and the proportion of land used for economic development and people's livelihood Larger, resulting in less land used to build ecological zones in these areas. However, through calculations, it is known that the scale of the ecological zone that needs to be built in these areas is relatively large, and the construction plan of the area cannot meet the optimization of the ecological environment of the area. At this time, we can only appropriately reduce the construction scale of the ecological zone.

We set the actual increaseable ecological area ratio of this place as Q_i and the planned ecological area ratio of this place as A_i, then the model can be updated as follows:

$$Q_i = \begin{cases} 40w_i\left(1.026^{-\beta_i}\right), & S_i < A_i \\ A_i, & S_i \geqslant A_i \end{cases} \tag{27}$$

In the end, we can get the actual scale of ecological zones that can be built in each administrative region, which can be represented by the following figure:

Figure 10:The percentage of increase in the actual ecological area that can be built in each administrative region

7.5 Assess changes in carbon neutrality

Assuming that the ecological zone is successfully completed in accordance with the construction plan, we can recalculate the forest coverage of the area and bring it back into the neural network model to calculate the carbon absorption capacity of the area after the construction of the ecological zone. We compare it with the carbon absorption capacity of the ecological zone before construction, as shown in the figure below:

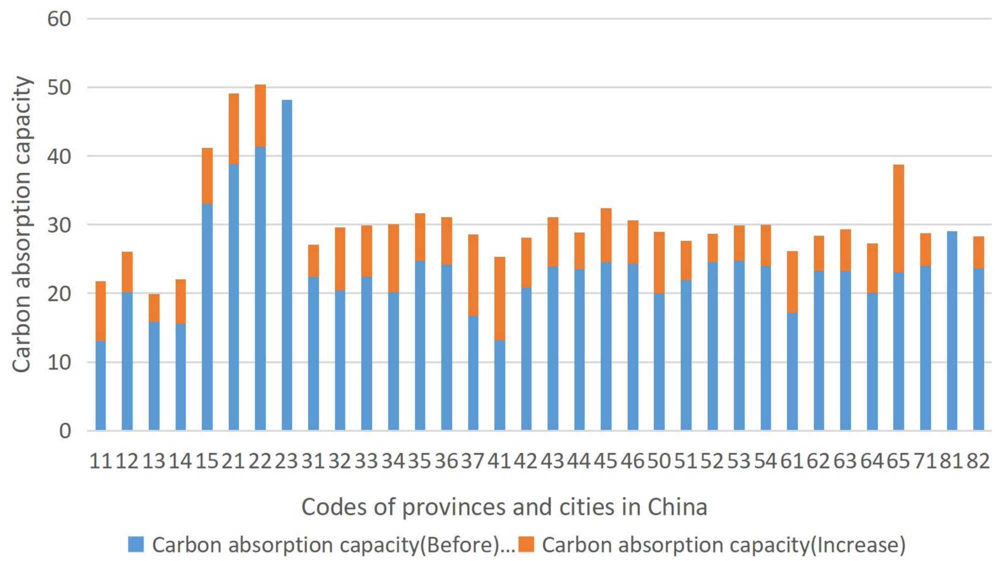


Figure 11:Comparison of carbon absorption capacity

It can be found that the carbon absorption capacity of each administrative region has a clear upward trend. We averaged the difference before and after the establishment of the ecological zone. As a result, the overall carbon absorption capacity increased by 7.17, which shows that the construction of the ecological zone will help China achieve carbon The neutralization goal has a better promoting effect.

8 Establishment and solution of Problem Four Model

Problem 4 requires us to analyze a country in the Asia-Pacific region, discuss which geographical locations of the country need to build an ecological zone and the scale of the ecological zone, and at the same time conduct an assessment of the carbon absorption capacity of the ecological zone. This is actually an extension of the model in question three, and we can use the model established in question three to solve it.

For countries in the Asia-Pacific region, we find that most parts of Australia are perennially dry, and a large area of land is desert. Therefore, the Australian government needs to build more ecological zones to enhance the territory's ability to resist sand and dust, so this can become Our ecoregion model is a more suitable research object.

We divide Australia into 7 geographical locations, namely Queensland, New South Wales, Victoria, Tasmania, Northern Territory, South Australia and Western Australia. By collecting the relevant index data of these seven geographic locations, and then bringing them into the neural network model for prediction, the anti-sand storm capacity, carbon absorption capacity and water purification capacity of these administrative regions can be obtained respectively.

In Australia, the important basis for establishing an ecological zone is its ability to resist sand and dust, we can slightly adjust the relationship (21) in the model to increase the weight of its ability to resist sand and dust storms, and obtain the following formula:

$$\beta = 0.45Y_1 + 0.35Y_2 + 0.2Y_3 \tag{28}$$

Then, we can calculate the comprehensive ecological indicators of each geographic location through the relational formula, as shown in the following table:

Table 8:Comprehensive ecological indicators for each geographic location

	Anti-dust ability	Carbon absorption capacity	Water purification capacity	Comprehensive ecological indicators
Queensland	67.97	24.10	33.30	45.68
New South Wales	72.38	25.19	35.14	48.41
Victoria	70.45	24.45	34.17	47.09
Tasmania	82.48	25.70	36.97	53.51
Northern Territory	53.45	25.62	31.79	39.38
South Australia	30.58	14.80	17.10	22.36
Western Australia	47.50	19.51	23.59	32.92

We can calculate the actual scale of the ecological area that can be built in each geographic location according to the size of the ecological area that can be planned for each geographic location in Australia and the formula (27), as shown in the following figure:

Figure 12:The percentage of increase in the actual ecological area that can be built in each administrative region

Assuming that the ecological zone is successfully completed in accordance with the construction plan, we can recalculate the forest coverage of the area and bring it back into the Neural Network model to calculate the carbon absorption capacity of the area after the construction of the ecological zone. We compare it with the carbon absorption capacity of the ecological zone before construction, as shown in the figure below:

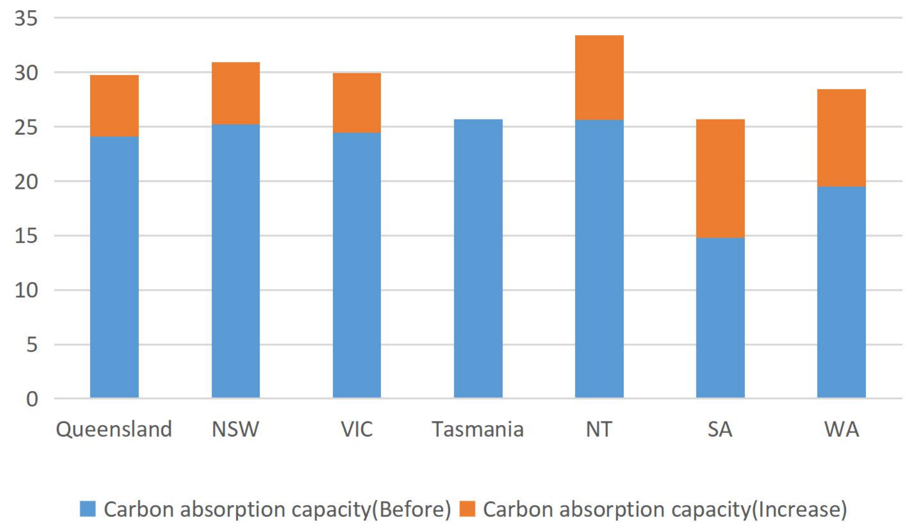


Figure 13:Comparison of carbon absorption capacity

9 Evaluation, Improvement and promotion of the Model

9.1 Evaluation of the Model

- **Advantage:**

(1) We have found the relationship between Saihanba's ecological environment and Beijing's ability to resist sand and dust through correlation research. There is no need to continue to collect complex data for research, which simplifies the difficulty of modeling.

(2) We use artificial intelligence algorithm neural network to analyze ecological capabilities. Compared with traditional regression models, the results we get will be more convincing.

(3) In the fourth question, Australia, a country with serious desertification, was selected as the research object, which has good representative significance and analytical value

- **Shortcoming:**

(1) In reality, there are many indicators for correctly evaluating an ecological region, and our model only uses a few more important indicators, which may not fully reflect the characteristics of the ecological region.

(2) When evaluating the ability to resist sand and dust, we judge by studying the frequency and severity of the occurrence of sand and dust storms. It is inevitable that there will be a bit of subjective evaluation, and the evaluation that may be obtained is not completely correct.

9.2 Improvement and promotion of the Model

(1) We can collect more indicators and data, and build a more complete and more complex mathematical model for more accurate evaluation of ecological indicators in various regions.

(2) We can apply the Saihanba ecological model to the whole country or even the whole world through the establishment of the Saihanba ecological model, and draw out the scale of the ecological zone that needs to be built according to the environment of each region.

10 References

[1] Study on the value of forest ecosystem service function in Saihanba Nature Reserve[D].Zhang Yunling.Hebei Normal University

[2] Design and implementation of forest resources data visualization system in Saihanba Forest Farm[D].Zhu Qianyu.Beijing Forestry University

[3] Dust weather in Beijing in recent 50 years and Its Control Countermeasures.Chen Guangting[J].Institute of environment and Engineering in cold and arid regions, Chinese Academy of Sciences

[4] History of severe dust storms and changes of surrounding ecological environment in Beijing[J].Chen Guangting.Institute of environment and Engineering in cold and arid regions, Chinese Academy of Sciences

[5] Liu lianglin, Wang Quanfeng, Lin Huangbin, parameter setting and application of BP neural network[J], infrastructure optimization, Vol. 28, No. 2, April 2007

[6] Quantitative discussion on the best protection benefit and forest coverage[D].Zhang Jian, Gong Yuanbo, Chen Linwu.College of forestry, Sichuan Agricultural University

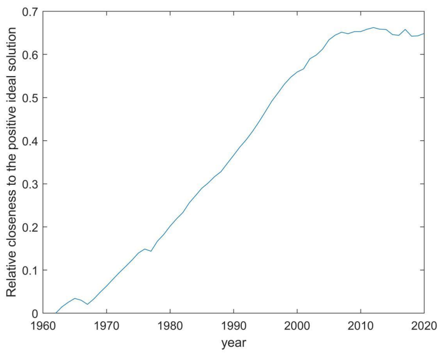
Report:

The construction of ecological civilization plays a very important role in the sustainable development of the world. China has always placed the construction of ecological civilization in a prominent position in the overall work. The transformation of Saihanba forest farm from desert to oasis in China is an excellent case of improving the local ecological environment by establishing an ecological reserve.

Figure:Comparison of Saihan dam in the past and present

Before **Now**

In order to study the effect of the establishment of Saihanba ecological area, we extract the indexes that can significantly represent the ecological environment of Saihanba by Entropy method: forest coverage rate, air quality index, water quality index, soil pollution index and biological abundance index. We can establish TOPSIS evaluation model of Saihan dam's impact on ecological environment through relevant data, and get the score of Saihan dam's ecological environment from 1962 to 2020. The higher the value, the better the local ecological environment.



We can see that when Saihanba forest farm has not been established, the score of the area is very low, indicating that the environment is very bad. However, in the following 60 years, Saihanba people jumped the ecological score to 0.65 through their hard-working hands, indicating that the local environment has improved greatly.

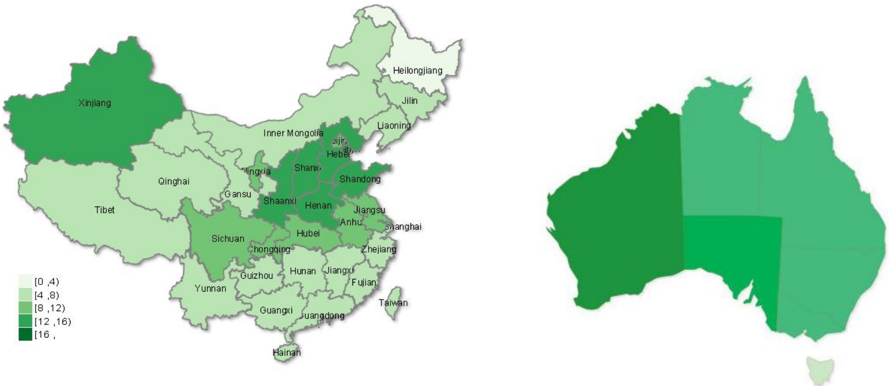
Then, in order to study the impact of Saihan dam on the anti dust ability of Beijing, we established a multiple regression model:

$$y^D = \alpha_0 + \alpha_1 lnx^f + \alpha_2 lnx^a + \alpha_3 lnx^w + \alpha_4 lnx^s + \alpha_5 lnx^b$$

According to the model, with the increase of Saihanba forest coverage and water quality index and the decrease of air pollution and soil pollution, the anti dust ability of Beijing shows an obvious upward trend. Therefore, we can think that the establishment of Saihanba ecological area has made a significant contribution to Beijing's anti dust ability.

Then, in order to extend the ecological protection model of Saihan dam to China, we calculate the three ecological indicators of anti sandstorm capacity, carbon absorption capacity and water purification capacity through BP neural network, and introduce comprehensive ecological indicators to judge whether it is necessary to establish an ecological area. Then, the proportion of ecological area to be increased in an administrative region is calculated according to the size of comprehensive ecological indicators, and the actual proportion of ecological area to be increased is obtained according to the idle land of each province.

At the same time, we can also extend the model to other countries in the Asia Pacific region, and can also calculate which geographical locations of these countries need to increase the ecological area. We take Australia, which is dry all the year round and has a wide desert area, as an example for analysis, as shown in the figure below:



The darker the color in the figure, the larger the proportion of ecological area actually needs to be increased in this area. If all regions are constructed according to this plan, we recalculate the forest coverage after the implementation of the scheme and bring it into the neural network model. It is found that the carbon absorption capacity of all regions has been significantly improved, which also shows that the construction of ecological areas is of great significance to the realization of carbon neutrality and, to a certain extent, also proves the popularization of our model.

To sum up, we suggest that each region can reasonably expand the ecological area according to these results without sacrificing our valuable ecological environment for rapid economic development, which is our bounden responsibility. The establishment of ecological area is to provide a more comfortable environment for our human habitat and add more beauty to our home.