Team Number:	apmcm2103506
Problem Chosen:	C

Ecological Protection Construction and Environmental Impact based on the Successful Experience of Saihanba

Green water and green mountains are golden mountains and silver mountains, just as General Secretary Xi said. With the help of the Chinese government, China's Saihanba Forest Farm has recovered from the desert and has now become an eco-friendly green farm. China's contributions and achievements in Saihanba are notable. We hope to build a more generalized ecological evaluation model based on Saihanba's ecological restoration system to help more regions in China and abroad to build a better ecological zone.

Question one requires a comparative analysis of the environmental conditions before and after the restoration of Saihanba. Based on more than 10 data sets, we extracted **a total of 69 indicators**. Due to the inconsistency of the completeness of the indicators in different time periods, we divided them into 4 time periods **(1962-1982, 1983-2004, 2005-2013, 2014-2021)** According to statistical analysis, the number of indicators selected in the 4 time periods is inconsistent. We draw the **time series diagram** of the whole time period for the indicators with higher integrity, conduct a rough analysis of the data, and use the gray correlation analysis to obtain the importance indicators of different time periods. Several time periods indicate that **Forest_Stock** is the most important indicator. This will serve as our second question of Saihanba benchmark data.

Question two needs to evaluate the role of Saihanba in the fight against sandstorms in Beijing. Based on Beijing's sandstorm disaster data set, we conducted **one-way ANOVA** to determine whether there are significant differences in the number of sandstorms and sandstorm wind speeds in different periods of time. Finally, we obtained a test value of less than 0.05, which confirmed that Saihanba played a more significant role in resisting sandstorms in Beijing.

Question three requires ecological zone planning and area calculation for other areas with poor ecological environment in the country. We have obtained data on air quality, carbon emissions and the degree of desertification in the corresponding provinces in some cities across the country. Using the **K-means** algorithm, **15 ecologically worst cities in Xinjiang and Nei Menggu are extracted** when the best K is 3, and ecological regions are constructed based on these cities. We assign different weights to the obtained data, find the scoring function for each city, and use the Saihanba Ecological District in Chengde City as the benchmark, focusing on the **degree of desertification** and **regional area limitations**, and calculate the demand for these 15 worst cities. The size of the constructed ecological zone, and **the conclusion that if the ecological zone is maintained for 10 years, the carbon absorption level of the city can be increased by about 11%**.

Question four requires the operation of question three in other parts of the Asia-Pacific region. We select Australia as the goal, hoping to improve Australia's desertified cities, consult the map and select Australia's city near the desert, **Albury**, and conduct a comprehensive analysis of **New South Wales**.

Keywords: Grey Relational Analysis Comprehensive Evaluation of Rank Sum Ratio K-means Analysis of Variance Time Series

Contents

1. Introduction

1.1 Problem Restatement

For question 1, we first need to find some indicators related to resistance of sandstorm ,environmental protection, and maintenance of ecological balance and stability. The selection of data is particularly critical. We have extracted a total of **69 indicators** and selected different time spans (Due to large missing data, the complete data from 1962 to 2021 is not available), and use different indicators to construct different scores and use **gray relational analysis** to extract important indicators, so as to conduct a comparative analysis before and after the restoration of Saihanba.

For question 2, we explored the correlation between the number of sandstorms and wind speed in Beijing and the afforestation indicators of Saihanba, combined with **the important indicators extracted from the first question** (such as Forest_Stock), and used variance analysis and family series to explore the relationship between time Regarding the importance of the changes in the results, the contribution of Saihanba to alleviating the sandstorm in Beijing is derived.

For question 3, we need to extend Saihanba's ecological protection model to the whole country. From the desert data set, we use **K-means** to identify areas that urgently need to adjust the ecological environment, **precising to prefecture-level cities**. According to the relevant indicators in the first question, set the weight to construct a focus on score. **Taking Saihanba in Chengde City as a benchmark**, we then estimate the area of the ecological zone that each city needs. For the sake of simplicity, we only calculated **the values of the top 15 regions** that most need to adjust the ecological model.

For question 4, we selected the **Australian region** where the desertification is more serious in the Asia-Pacific region, and used the method in the third question to do the same to carry out extended research.

1.2 Background Information

As General Secretary Xi Jinping said, green water and green mountains are golden mountains and silver mountains. "Ecological prosperity leads to the prosperity of civilization, and ecological decline leads to the decline of civilization." From the perspective of the development of human history, Xi Jinping has repeatedly expounded on the relationship between man and nature, the rise and fall of civilization and the destiny of the nation, environmental quality and the well-being of the people. Profoundly, the epistemology of how to deal with the relationship between human production and the natural environment has been developed to a new height, reflecting the historical responsibility for ecological issues and the overall development view.

With the help of the Chinese government and the efforts of the people through generations, Saihanba has created a miracle of turning from a desert into a sea of forests. Based on the successful experience of Saihanba, we attempt to promote the ecological protection model of Saihanba nationwide, and provides an example for the ecological protection model of the Asia-Pacific region. We first use grey relational analysis method and rank sum ratio comprehensive evaluation method to establish some ecological environment evaluation models. Secondly, through canonical correlation analysis and one-way analysis of variance, we prove that Saihanba Forest Farm plays an important role in resisting sandstorms in Beijing. Based on the successful experience of Saihanba, we collect relevant indicators of various cities in China, and analyse the geographical locations of ecological protection areas that need to be built through cluster analysis, as well as their specific scale and quantity. Finally, we chose Australia, an Asia-Pacific country, to demonstrate its demand for ecological reserves through a model, including geographic location, quantity, and scale, and evaluated these contributions to absorbing greenhouse gases and achieving carbon neutrality.

2. Assumptions and Notations

2.1 Assumptions

To simplify our modelling, we make the following assumptions:

Assumption 1 *The data we found in official websites, government yearbooks and published papers are all true and reliable. Also, our model building and solving are based on these data.*

Assumption 2 *The level of scientific and technological development from 1962 to the present is stable, and the changes in the data we use over time have nothing to do with its detection methods.*

Assumption 3 *Despite the differences in topography and climate, Saihanba's ecological protection model has a benchmarking role in the country and is of significance for promotion in the Asia-Pacific region and even the world.*

2.2 Notations

In this work, we use the nomenclature in Table 1 in the model construction. Other none-frequent-used symbols will be introduced once they are used.

Table 1 Notations used in this literature

Symbol	Meaning	Unit
$\xi_i(k)$	Ray correlation coefficient	–
f_i^*	Gray weighted relevance degree r_i ranking index value	–
A	Decision matrix	–
B	Normalized decision matrix	–
C	Weighted norm matrix	–
RSR	Rank sum ratio	–
$WRSR$	Weighted rank sum ratio	–

3. Data Search and Exploration

GHG emissions for Beijing, Chengde, Weichang County and other urban areas Emission data were obtained from (emission inventories were compiled based on the latest revision of energy data (2015) from the China Bureau of Statistics), CEADS and global real-time carbon data [2]. Missing values for intermediate years were obtained using an interpolation algorithm.

World national GHG emissions from the World Development Indicators. Agriculture, forestry, animal husbandry and fishery production, total area, forest cover, plantation forest area, forest stand area, plantation forest data, agricultural products value from Chengde City, Hebei Province, China, annual statistical yearbook [1], carbon dioxide absorption and oxygen release, tourism visits and income, surface water quality and quantity, climate indicators of Weichang.

County are collected from China Forestry and Grassland Science Data Center [2] and Saihanba Mechanical Forestry Field. Weichang County air pollution indicators from China Real-Time Air Pollution Index [5], sandstorm data from the paper (Research on the spatial and temporal distribution pattern of sand and dust disasters in Chengde City in the past 50 years - Wanfang data) [3], precipitation levels from historical weather data of Weichang County (temperature, humidity and other indicators) [7]. Beijing dust storm data from (Gale and Sand-dust Storm disaster data of 1954-2017 in Beijing) [4]. Saihanba soil data and other meteorological data from the National Forestry and Grassland Data Science Center [2] [3].

Historical monthly air quality statistics for Beijing, Chengde and other related cities from Air pollution in the world [7]. Historical weather data (temperature, humidity and other indicators) for Weichang County from rp5.ru [5]. China sanding data from the Western China Environmental and Ecological Science Data Center. Indicators used in the article are referenced from (Saihanba Forest Ecological Protection and Restoration - Wanfang Data) [6].

Australian Desertification Data from NSW GOVERNMENT(`https://data.nsw.gov.au/`), Austrilia Bureau of Meteorology (`http://www.bom.gov.au/`), weather data from Air pollution in the world, UNEP, World Meteorological Organization, air Pollution Index from Air pollution in the world and IQAir [7].

4. Model Construction

4.1 Saihanba Ecological Environment Evaluation Model

4.1.1 Data Processing and Analysis

We found a total of **69 relevant indicators**, which were divided in detail from 10 levels. The ten first-level indicators are ***Growth and Decline of Forest Resources, Ecological Purification Capacity, Travel, Growth of Forest Resources, Paddock Air Pollution Index, Wei Chang Paddock Air Pollution Index, Windy Weather above Level 4 in Saihanba, Man-made Forestry, Agricultural Output Value, Wei Chang Sand Storm. The secondary indicators are shown in Figure 1***. Among them, we use Weichang's weather data (pollution, temperature, and humidity, etc.) as the weather data

Growth and Decline of Forest Resources	Total_Area(10,000 hectares) Stand_Area(10,000 hectares) Standing_Wood_Total_Accumulation(10,000 cubic meters) Artificial_Forest_Area(10,000 hectares) Forest_Coverage_Rate(%) Cover_Area(10,000 hectares) Forest_Stock(10,000 cubic meters)
Ecological Purification Capacity	Carbon_Dioxide_Absorption(10,000 tons) Oxygen_Release(10,000 tons)
Travel	Tourist_Number(10,000 person) Tourism_Revenue(100 million yuan)
Growth of Forest Resources	Water_Conservation(100 million cubic meters) Surface_Water_Quality_Standard_Rate(%)
Paddock Air Pollution Index	PM2.5 PM10 O3 NO2 SO2 CO Urban_Air_Quality_Standard_Days/Better_Than_Grade_ii(days) Urban_PM2.5_Concentration(per microgram per cubic meter) AQI
Wei Chang Paddock Air Pollution Index	T(air temperature 2m above ground level (Celsius)) P(Atmospheric pressure at mean sea level (millimeters of mercury)) U(Relative humidity at ground height of 2m (%)) FF(average wind speed at ground height of 10-12 m (m/s) within 10 minutes prior to observation) FF3(maximum gust (m/s) at ground height of 10-12 m between observations) Tn(the lowest temperature (Celsius) within the past period of time (not exceeding 12h)) Tx(highest temperature in the past period of time (not exceeding 12h) (Celsius)) VV Td(horizontal visibility (km)) N (the number of all cloud layers C1 observed, the number of cloud layers Cm observed when there is no cloud layer C1) WW W1

Figure 1a Indicators for Saihanba analysis.

of Saihanba. By reading the literature, we roughly divide the time axis into 4 periods: **1962 to 1982, 1983 to 2004, 2005 to 2013, 2014 to 2021**. The first reason for this division is due to the large lack of data in different years, and secondly, we understand the period from 1983 to 2000 was a period of the rapid increase in the stock volume of the forest farm, which fully doubled. We believe that this division can show the different development history of the Saihanba Forest Farm.

Regarding to the data situation, we made the following **4 descriptive statistics** based on these 4 periods and listed the approximate distribution and **missing data**. We use **Random Fores** for all missing values, filling in the data by filling in the prediction results of the model. Figure 2 shows the missing data information from 1962 to 1982. We focused on exploring the timing of some data and drew a series of timing diagrams.

For the processing of some indicators, we need to make relevant explanations.

RRR((Mean_Annual_Precipitation)(mm)) is the sum of daily precipitation in a year, which may be different from actual precipitation.

	W2
	Tn (the lowest temperature (degrees Celsius) in the past period of time (not more than 12h))
	Tx (the highest temperature (degrees Celsius) in the past period of time (not more than 12h))
	Cl
	Nh
	H
	Cm
	Ch
	VV
	Td (horizontal visibility (km))
	RRR((Mean_Annual_Precipitation)(mm))
	tR (dew point temperature (degrees Celsius) at a height of 2 meters from the ground)
	Precipitation_Grade
Windy Weather above Level 4 in Saihanba	Days_With_Force_Four_Gales
	Days_With_Force_Five_Gales
	Days_With_Force_Six_To_Eight_Gales
Man-made Forestry	Artificial_Afforestation_Area
	That_Year_New_Seal_Mountain (sand) Afforestation_Area
	Artificial_Replacement_Area
	Area_Of_Forest_Tending
	Closed_Forest_Area
	Surrival_Rate_Of_Afforestation
	Vegetable_Planting_Area
Agricultural Output Value	Seed_Yield
	Total_Output
	Total_Output_of_Aquatic_Products
	Gross_Output_Value (agriculture, forestry, animal husbandry and fishery)
	Value_Of_Agricultural_Production
	Value_Of_Forestry_Production
	Value_of_Farming_Production
	Value_of_Fishing_Production
	Commodity_Output_Value (Agricultural, forestry, animal husbandry and fishery)
Wei Chang Sand Storm	Floating_Dust
	Blowing_Sand
	Sand_Storm

Figure 1b Indicators for Saihanba analysis (Continued).

Air pollution indicators (PM2.5, PM10, AQI, etc.) take the annual average, annual minimum, and annual maximum.

Wei Chang Sand Storm indicators have only truncated data, that is, point data with an interval of 3 years. We use linear interpolation to turn it into continuous data.

The wind speed class data is taken from the annual minimum and annual maximum.

First, perform a rough analysis of the data. From the time series chart below, it can be seen that between 1962 and 2021, ***the forest coverage of the forest farm*** (reflected by Forest_Covery, Cover_Area, Forest_Stock) expanded with the growth of the year. ***The absorption of carbon dioxide, and the release of oxygen*** increased with the growth of the year. It shows that Saihanba has played a good role in controlling carbon emissions. ***The amount of water conservation*** (the amount of water in a forest is an important indicator for measuring the ecology of a region.) has also increased year by year, and the growth rate has accelerated after 2015. Shows the good features of the Saihanba forest system.

In addition, **for the data from 2014 to 2021**, the indicator dimension is relatively high (47 dimensions), and the sample size (8) is small, which makes the statistical

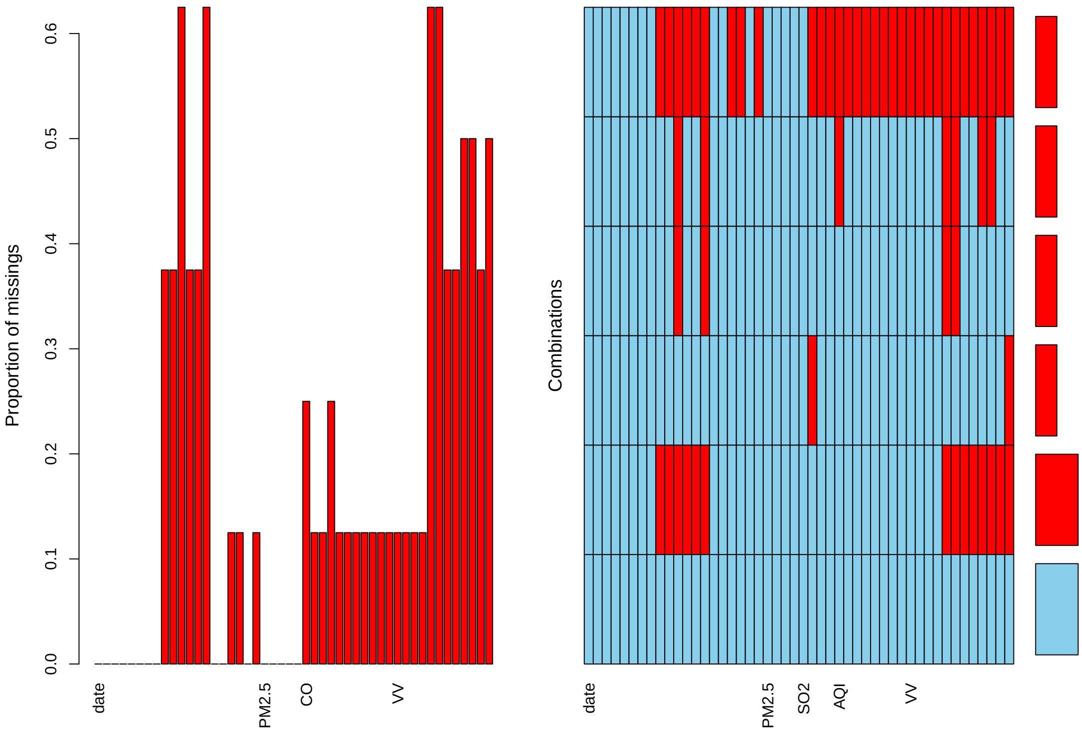

Figure 2 Missing Valus from 2014 to 2021.

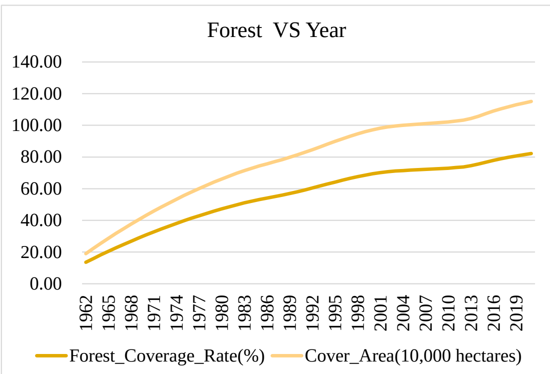

Figure 3 Forest Indicaters VS Year.

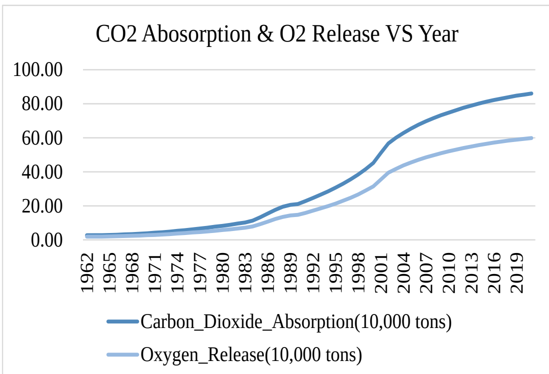

Figure 4 Carbon Dioxide Abosorption and oxygen Release VS Year.

work particularly complicated. We hope to reduce the dimension as much as possible. After many experiments, we finally reduce irrelevant variables or indicators that have less impact on the results, eventually reduced to **23 effective indicators** (descriptive statistical results in the appendix for details) as shown in Figure 7.

As can be seen from the correlation matrix of the 2014-2021 in Figure 8, there is a strong correlation between the 47-dimensional features. We need to use **PCA** to reduce the data dimension. We select 95% of the interpretation and try to extract the main components, but due to the main components, the interpretability is poor. **We only select the dimensions that have a very low proportion in each main component to discard, so as to obtain 23 effective variables**.

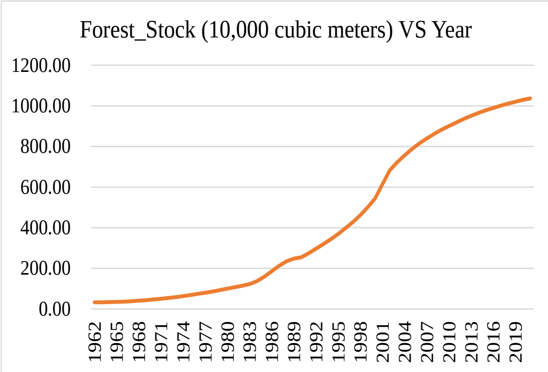

Figure 5 Forest Stock VS Year.

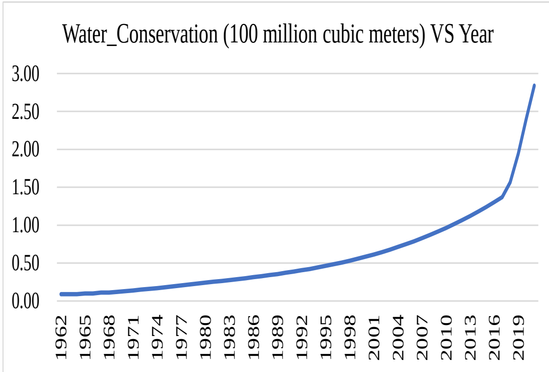

Figure 6 Water Conservation VS Year.

4.1.2 Model Establishment and Solving

After the rough analysis of the data, we began to analyze each stage in detail.

We built a gray correlation model, TOPSIS to extract the main indicators of the four periods, and used the rank sum ratio comprehensive evaluation model to find the year ranking.

First of all, when we transform the indicators, we need to unify the types of evaluation model indicators. We unify the indicators as very large indicators. We hope that the higher the value, the better. Therefore, we discard the interval data of temperature and convert the very small indicators to extremes. Large indicators use the following methods:

$$X_i' = \max\left(X_i\right) - X_i \tag{1}$$

The specific steps of **gray correlation analysis** are as follows

1. Determine the comparison object (evaluation object) and the reference sequence (evaluation standard).
   There are m evaluation objects, n evaluation indicators, the reference number sequence is $x_0 = \{x_0\left(k\right)|k = 1, 2, ..., n\}$, and the comparison sequence is $x_i = \{x_i\left(k\right)|k = 1, 2, ..., n\}, i = 1, 2, ..., m$.
2. Determine the weight corresponding to each indicator value.
   The analytic hierarchy process can be used to determine the weight $w = [w_1, ..., w_n]$ corresponding to each indicator, where $w_k, k = 1, 2, ..., n$ is the weight corresponding to the k-th evaluation indicator.
3. Calculate the gray correlation coefficient.

$$\xi_i\left(k\right) = \frac{\underset{s}{\min}\underset{t}{\min}\left|x_0\left(t\right) - x_s\left(t\right)\right| + \rho\underset{s}{\max}\underset{t}{\max}\left|x_0\left(t\right) - x_s\left(t\right)\right|}{\left|x_0\left(k\right) - x_i\left(k\right)\right| + \rho\underset{s}{\max}\underset{t}{\max}\left|x_0\left(t\right) - x_s\left(t\right)\right|} \tag{2}$$

   In order to compare the sequence x_i, for the reference sequence x_0, the correlation coefficient on the k-th index, where $p \in [0, 1]$ is the resolution coefficient. In the formula, $\underset{s}{\min}\underset{t}{\min}\left|x_0\left(t\right) - x_s\left(t\right)\right|$ and $\underset{s}{\max}\underset{t}{\max}\left|x_0\left(t\right) - x_s\left(t\right)\right|$ are called two levels, the minimum difference and the two maximum difference. Generally speaking, the larger the resolution coefficient ρ, the larger the resolution; the smaller the ρ, the

	Total_Area(10,000 hectares)	Stand_Area(10,000 hectares)	Standing_Wood_Total_Accumulation(10,000 cubic meters)	Artificial_Forest_Area(10,000 hectares)	Forest_Coverage_Rate(%)
count	22	22	22	22	22
mean	9.5691	5.7791	440.8264	4.4882	62.0918
std	0.2072	0.6793	195.8432	0.9136	6.8204
min	9.3900	4.7000	111.8000	2.9000	51.0400
25%	9.4825	5.3200	265.0400	3.6550	56.2500
50%	9.5000	5.7300	480.2100	4.7600	62.3150
75%	9.5100	6.1500	610.4100	5.3300	68.3450
max	10.2000	7.0400	714.1800	5.5400	71.4700

	Cover_Area(10,000 hectares)	Forest_Stock (10,000 cubic meters)	Carbon_Dioxide_Absorption(10,000 tons)	Oxygen_Release(10,000 tons)	Water_Conservation(100 million cubic meters)
count	22	22	22	22	22
mean	86.9282	376.0727	31.2064	21.7050	0.4541
std	9.5489	192.1380	15.9436	11.0896	0.1321
min	71.4600	123.0000	10.2100	7.1000	0.2800
25%	78.7500	238.0825	19.7575	13.7400	0.3450
50%	87.2350	332.1500	27.5600	19.1700	0.4300
75%	95.6775	491.6925	40.8025	28.3750	0.5525
max	100.0500	757.2700	62.8400	43.7100	0.7100

	Days_With_Force_Six_To_Eight_Gales	Floating_Dust	Blowing_Sand	Sand_Storm	Mean_Annual_Precipitation
count	22	22	22	22	22
mean	56.7727	5.9636	9.5136	1.1864	480.9109
std	14.9282	1.0706	4.3858	0.9682	33.2152
min	42.0000	4.0000	1.9000	0.0000	440.8000
25%	49.0000	5.0500	5.5500	0.3250	440.8000
50%	49.0000	6.2500	11.1000	1.1000	480.9100
75%	64.0000	6.7750	12.4750	1.9500	513.0000
max	95.0000	8.0000	15.0000	3.0000	513.0000

Figure 7 Descriptive statistics from 1983 to 2004.

smaller the resolution.

4. Calculate the gray-weighted relevance degree.
   The calculation formula of the gray-weighted relevance degree is

$$r_i = \sum_{k=1}^{n} w_i \xi_i (k) \tag{3}$$

   where r_i is the gray weighted relevance degree of the i-th evaluation object to the ideal object.

5. Evaluation analysis
   According to the size of the gray weighted relevance degree, the evaluation objects are sorted, and the relevance order of the evaluation objects can be established. The greater the relevance degree, the better the evaluation result.

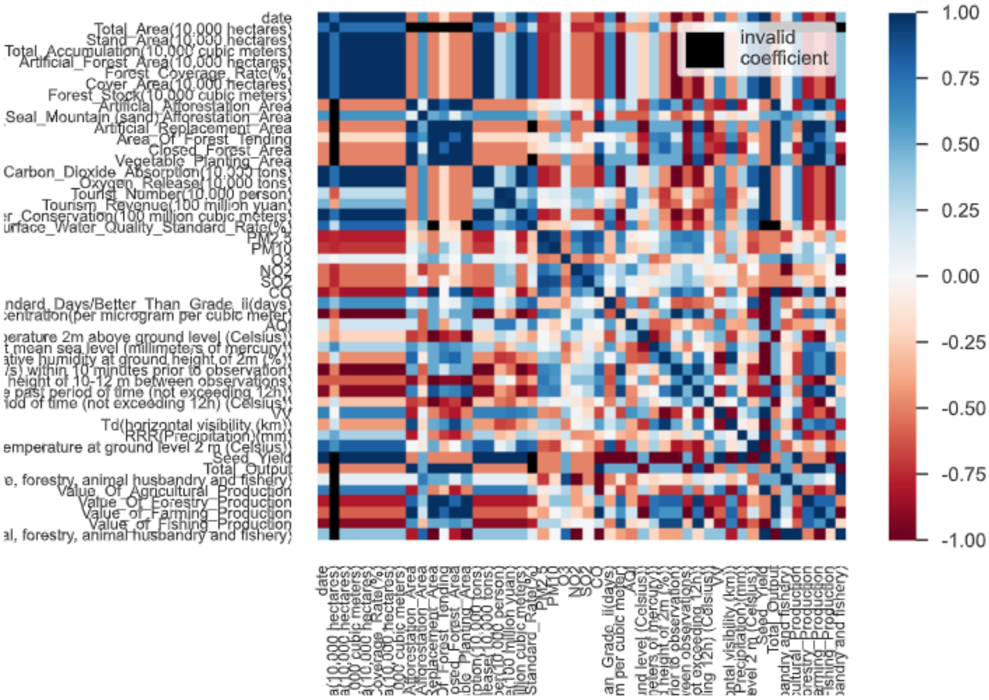

Figure 8 Correlation matrix.

The specific algorithm of the **TOPSIS** method is as follows

1. Use the vector programming method to obtain the normalized decision matrix.
   Suppose the decision matrix $A = \left(a_{ij}\right)_{m\times n}$ of the multi-attribute decision-making problem, and the normalized decision matrix $B = \left(b_{ij}\right)_{m\times n}$, where

$$b_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^{m} a_{ij}^2}} \tag{4}$$

2. Construct a weighted norm matrix $C = \left(c_{ij}\right)_{m\times n}$.
   Let the weight vector of each attribute given by the decision maker as $w = [w_1, w_2, \ldots, w_n]^T$. Then,

$$c_{ij} = w_j \cdot b_{ij},\ i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n \tag{5}$$

3. Determine the positive ideal solution C^* and the negative ideal solution C^0.
   Suppose the value of the jth attribute of the positive ideal solution C^* is c_j^*, and the value of the jth attribute of the negative ideal solution C^0 is c_j^0, then
   Positive ideal solution:

$$c_j^* = \begin{cases} \max\limits_{i} c_{ij},\ j \text{ is benefit attribute} \\ \min\limits_{i} c_{ij},\ j \text{ is cost attribute} \end{cases} \quad j = 1, 2, \ldots, n \tag{6}$$

Negative ideal solution:

$$c_j^0 = \begin{cases} \min\limits_i c_{ij}, \ j \text{ is benefit attribute} \\ \max\limits_i c_{ij}, \ j \text{ is cost attribute} \end{cases} \quad j = 1, 2, ..., n \tag{7}$$

4. Calculate the distance from each scheme to the positive ideal solution and the negative ideal solution.
   The distance from alternative d_i to the positive ideal solution is

$$s_i^* = \sqrt{\sum_{j=1}^{n} \left(c_{ij} - c_j^*\right)^2}, \ i = 1, 2, ..., m \tag{8}$$

   The distance from alternative d_i to the negative ideal solution is

$$s_i^0 = \sqrt{\sum_{j=1}^{n} \left(c_{ij} - c_j^0\right)^2}, \ i = 1, 2, ..., m \tag{9}$$

5. Calculate the ranking index value of each plan (i.e. comprehensive evaluation index).

$$f_i^* = \frac{s_i^0}{s_i^0 + s_i^*}, \ i = 1, 2, ..., m \tag{10}$$

6. Arrange the pros and cons of the schemes in descending order of f_i^*.

The steps of the **rank sum ratio comprehensive evaluation method** are as follows

1. Rank
   Arrange the m evaluation indexes of n evaluation objects into an original data table with n rows and m columns. Compile the rank of each evaluation object for each index, among which the benefit index is compiled from small to large, the cost index is compiled from large to small, and those with the same index data are compiled with the average rank. The resulting rank matrix is denoted as $R = \left(R_{ij}\right)_{n\times m}$.
2. Calculate the rank sum ratio (RSR).
   According to Equation 11 to calculate the rank sum ratio.

$$RSR_i = \frac{1}{mn} \sum_{j=1}^{m} R_{ij}, \ i = 1, ..., n \tag{11}$$

   When the weight of each evaluation index is different, calculate the weighted rank sum ratio ($WRSR$), the calculation formula is

$$WRSR_i = \frac{1}{n} \sum_{j=1}^{m} w_j R_{ij}, \ i = 1, ..., n \tag{12}$$

   where w_j is the weight of the j-th evaluation index and $\sum_{j=1}^{m} w_j = 1$.
3. Calculate the probability unit.
   Compile the RSR (or $WRSR$) frequency distribution table from small to large, list each group of frequencies f_i, calculate the cumulative frequency of each group cf_i,

calculate the cumulative frequency $p_i = cf_i/n$, convert p_i to the probability unit Probit_i. Probit_i is the p_i quantile of the standard normal distribution plus 5.

4. Calculate the linear regression equation. Take the probability unit Probit_i corresponding to the cumulative frequency; as the independent variable, use RSR_i, (or $WRSR_i$) as the independent variable, and calculate the linear regression equation, that is, $RSR(WRSR) = a + b \times \text{Probit}$.
5. Sort by bins.
   According to the estimated value of RSR ($WRSR$) calculated by the regression equation, the evaluation objects are sorted by bins.

4.2 Beijing Sandstorm Resistance Model

4.2.1 Data Processing

We use the Beijing sandstorm data set (including the number of sandstorms per year, sandstorm wind speed and other indicators), combined with the important indicator Forest_Stock (10,000 cubic meters) analyzed in the first question for follow-up analysis. For wind speed, it is not meaningful to get the average value. We choose the annual maximum wind speed as the evaluation index. The wet temperature index also selects the annual minimum and annual average as the evaluation index.

The data span of sandstorms in Beijing is 1954-2021. Therefore, we discarded the data from 1954 to 1961 and started statistical analysis from 1962, the starting year of Saihanba.

4.2.2 Model Establishment and Solving

First, we draw the relevant timing diagram as shown in Figure 9. It can be seen that the maximum wind speed and the minimum wind speed caused by sandstorms have great volatility, but the overall trend shows a downward trend. With the obvious increase of Forest_Stock, the above 2 values have a significant decrease. The afforestation work in Saihanba has a positive effect on resisting sandstorms in Beijing. However, in Figure 9 and 10, the number of sandstorms in Beijing is on the rise, and the data is very unstable. We believe that the fluctuation of the data may be caused by the long time and missing data.

Next, we performed an analysis of variance. We roughly divide the time into 1962-1990 and 1991-2021, and explore whether there are significant differences in Beijing's sandstorm index in different time periods. The final analysis shows that there are significant differences in Beijing's humidity and the maximum wind speed in the year when sandstorms are encountered. The P-value is 0.002397, less than 0.05, there is a significant difference. The decrease in wind speed found in the time series further confirms the positive effect of Saihanba in resisting sandstorms in Beijing.

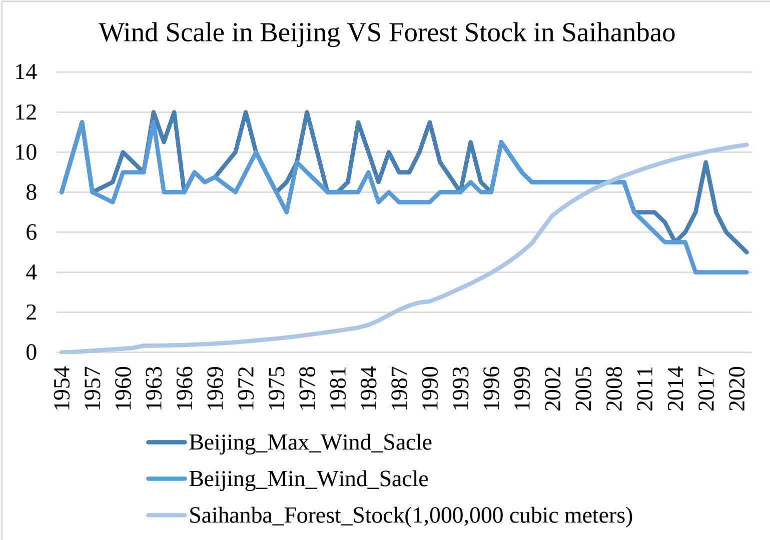

Figure 9 Wind Scale in Beijing VS Forest Stock in Saihanbao.

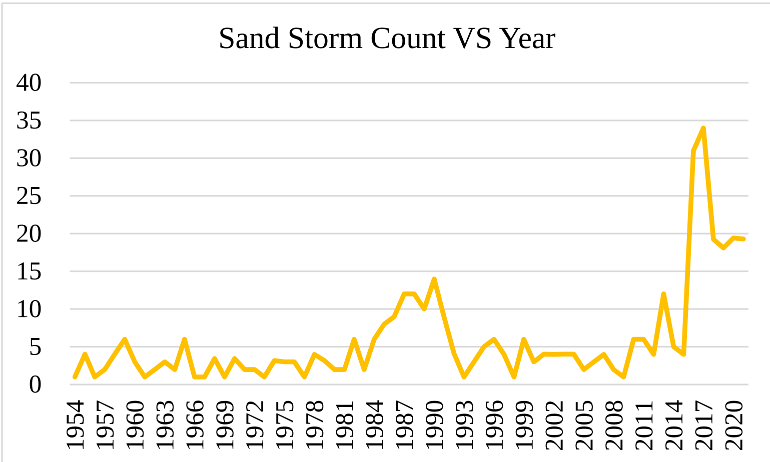

Figure 10 Sand Storm Count VS Year.

4.3 The Extension Model of Saihanba Model

4.3.1 Data Processing

First, we collected a national desert data set, and obtained the area and proportion of deserts and sandy land in the provinces and regions of the country. Part of the desert data is shown in the Figure 12.

Due to the limited collection of indicators in cities across the country, we only collected the following indicators as input: **AIR**(CO2_Release, AVG_PM25, AVG_PM10, AVG_SO2, AVG_NO2, AVG_CO, AVG_O3) **DESERT**(Flow Area, Semi-flow Area, Semi-fix Area, Fix Area)

All air indicators are averaged annually. Since there is no desertification data for each county-level city, we make the desertification data for each city equal to the desertification data for each province.

Source	SS	df	MS	F	P-value	F crit
Inter group	244.9292	67	3.655659	2.004197	0.002397	1.496169
Intra group	124.0321	68	1.824002			
Total	368.9613	135				

Figure 11 Oneway-ANOVA Result.

Province	Flow Area	Proportion1	Semi-flow Area	Proportion2	Semi-fixed Area
Xin Jiang	31455290	60.45	14095382	27.09	5087454
Nei Menggu	12833773	25.93	9491884	19.18	14913818
Qing Hai	4226511	36.26	5329469	45.72	1872936

Province	Proportion3	Fixed Area	Proportion4	Total Area	National Proportion
Xin Jiang	9.78	1397106	2.68	52035232	41.8862
Nei Menggu	30.13	12252173	24.76	49491648	39.8387
Qing Hai	16.07	227362	1.95	11656278	9.3828

Figure 12 Desert data example.

4.3.2 Model Establishment and Solving

We perform K-means clustering on all 292 cities, hoping to get the potential classification. First, we use the elbow diagram, as shown in Figure 13 to get the best K value:

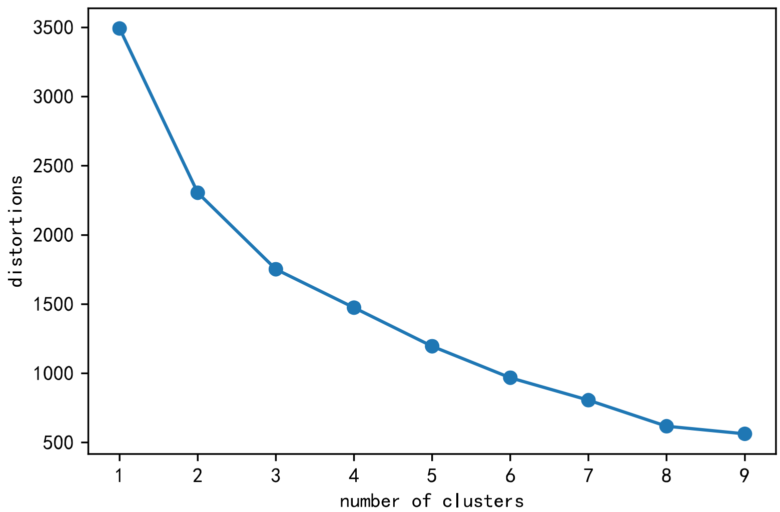

Figure 13 Elbow figure for K-means clustering.

Since we only want to be able to distinguish the most obvious data, that is, the regional cities that need to build protected areas most, there is no need to distinguish them in detail. There is an obvious turning point at $K = 3$ in the figure. We take $K = 3$ for cluster analysis, and the results are shown in Figure 14.

The three categories are: 91, 25, and 176 cities, and the main provinces are marked in Figure 15.

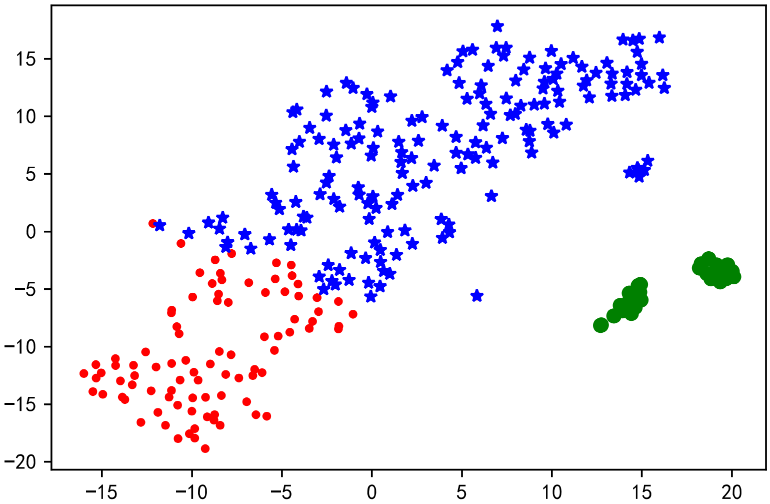

Figure 14 Result of K-means clustering.

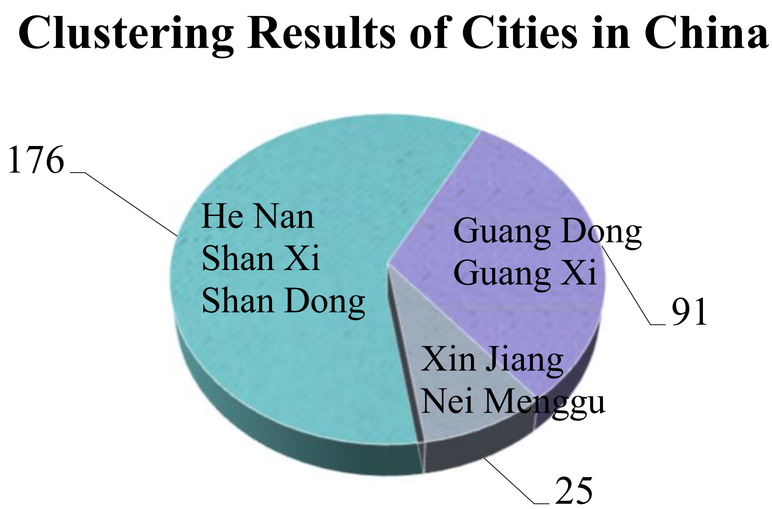

Figure 15 Specific provinces of K-means clustering.

Among the 25 cities in the cluster, the only major cities are Xinjiang and Inner Mongolia, which coincides with the top two with severe desertification. After the index is maximized, we use the rank sum ratio comprehensive evaluation to rank the 25 cities. The 15 cities with the lowest scores are: Turpan, Urumqi, Hotan area, Baotou, Changji Hui Autonomous Prefecture, Kashgar area, Hohhot , Wujiaqu, Aksu area, Wuhai, Shihezi, Yili Kazakh Autonomous Prefecture, Chifeng, Hami area, Bayannaoer.

We visualize the results of the comprehensive evaluation using the rank sum ratio as shown in Figure 16. The deeper the yellow, the lower the score, and the more it is necessary to build an ecological zone. The green represents a model city with a good natural environment.

We hope to explore the more detailed classification of the classes (91 cities) in which Chengde (Saihanba) is located. We further K-means clustering on it, and divided into 8 classes. The visualized map is as shown in figure 17.

Now, we are going to calculate the area of ecological regions for these 15 cities. Among the indicators we have, the indicators that affect the area of the ecological zone are: desertification indicators, air condition indicators (large particles), carbon emission

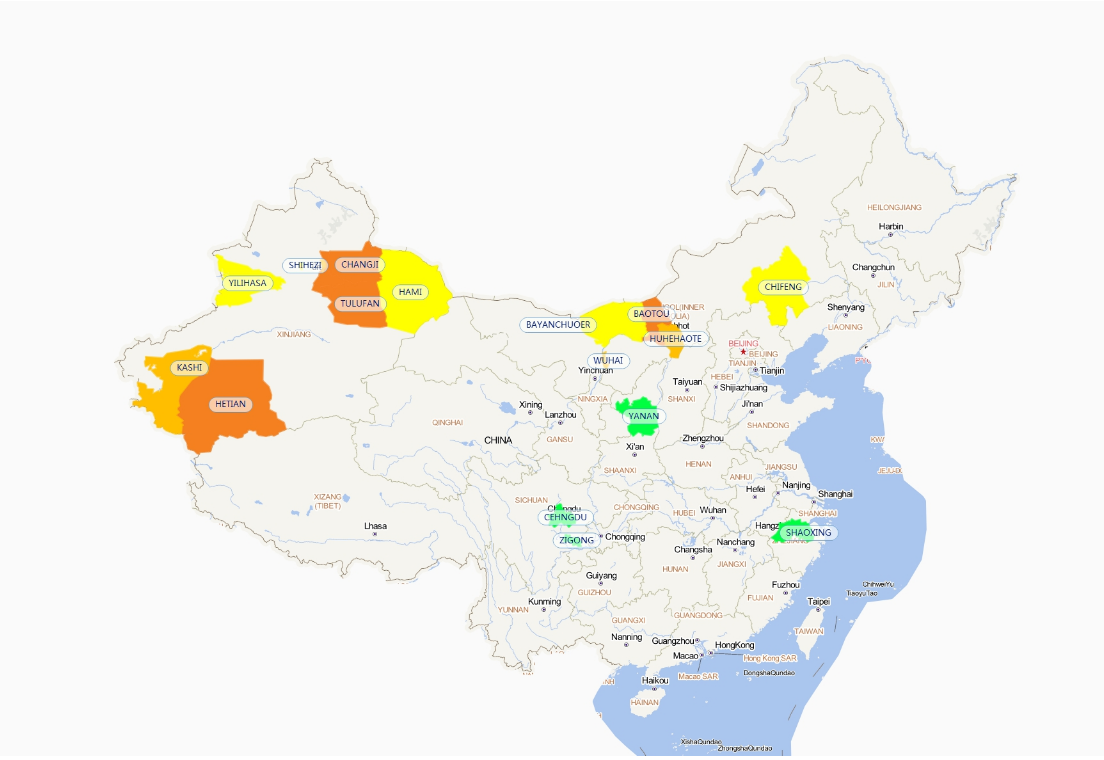

Figure 16 Visual results of comprehensive evaluation using the rank sum ratio.

indicators, and geographic area indicators. After repeated tests, we assign weights to them respectively: Desert-0.1, Air-0.05, Carbon-0.05, Area-0.8. Geographical area occupies much weight, because geographic area plays a very restrictive role.

For the area of the ecological zone of each city, we obtained a set of scores:

$$Ecology_{Area_i} = f\left(score_{city_i}\right) \cdot Saihan_{Area} / f\left(score_{city_{chengde}}\right) \tag{13}$$

The result is shown in Figure 18.

However, The data we use is in 2000. On the one hand, it is difficult for us to find the correlation coefficient in recent years. On the other hand, the data analysis in 2000 can test our results from another aspect. Judging from the results we have obtained, most cities conform to our judgment, that is, the need to establish ecological protection areas. For example, Chifeng has established 7 national protection areas and 9 provincial protection areas, and Bayannaoer has established 2 national protection areas and 4 provincial protection areas.

Regarding the improvement of carbon emissions, if these urban areas can build ecological areas according to the construction level of Saihanba in Chengde and maintain the construction for 10 years, the carbon dioxide absorption capacity can be increased by 10.84% (obtained from the 2012-2021 carbon dioxide absorption data of Saihanba).

Figure 17 Visual map of K-means clustering.

City	City Area (Square km)	Score	Ecological Area (Square km)
Chengde	**39519**	**1.736316**	**706**
Turpan	69713	2.15573	876.537262
Urumqi	13800	1.99232	810.093322
Hotan area	248100	2.724875	1107.955987
Baotou	27691	2.099285	853.586247
Changji Hui Autonomous Prefecture	73660.41	2.138409	869.494241
Kashgar area	162000	2.483331	1009.74219
Hohhot	17200	2.059538	837.424637
Wujiaqu	711	1.911733	777.325749
Aksu area	132500	2.31613	941.757147
Wuhai	1754	2.011131	817.74176
Shihezi	6007	1.91255	777.658297
Yili Kazakh Autonomous Prefecture	268593	2.695824	1096.143597
Chifeng	90021	2.202231	895.44468
Hami area	142100	2.260177	919.006062
Bayannaoer	65000	2.102817	855.022131

Figure 18 Scale Score.

4.4 The Development Model of the Asia-Pacific Region Based on Australia

4.4.1 Data Processing

We found that the desertification in Australia is particularly serious in the Asia-Pacific region, so we focus on analyzing Australia's ecological construction. By consulting news and Australian data, we found that Abury is located in New South Wales, the city with the worst air index, and New South Wales is close to a large desert, so we focus on analyzing New South Wales. We perform descriptive statistics on the acquired data set, and the results are shown in Figure 19

	CO2.Emissions	CO2_ALL	AVG_PM25	AVG_PM10	AVG_O3
count	51	51	51	51	51
mean	111.8455	371.4545	31.3674	15.1433	23.9123
std	31.3438	85.8733	0.9535	1.2660	0.3936
min	55.0000	216.4400	29.6025	12.9607	21.9000
25%	81.0000	286.3650	30.9077	14.3253	23.8346
50%	114.0000	376.3300	31.3552	14.9723	23.9464
75%	143.4150	458.0550	31.7879	15.6292	23.9626
max	150.0000	476.0100	35.4056	20.6556	25.4902

	AVG_NO2	Global.Solar.Exposure	Precipitation	Maximum.Temperature
count	51	51	51	51
mean	354.2756	206.0893	353.2081	162.9890
std	4.2737	6.5973	325.6484	122.6969
min	331.0000	186.0000	0.0000	0.0000
25%	354.0426	202.0170	32.2864	25.8020
50%	354.5000	205.9460	352.6000	252.4000
75%	356.0165	208.9340	626.3000	269.3500
max	364.0000	224.1000	916.4000	286.0000

Figure 19 Descriptive statistics.

4.4.2 Model Establishment and Solving

Due to the small amount of Australian data obtained, direct operation is not possible. We re-fitted the carbon dioxide absorption data of the original Saihanba (CO2 absorption growth rate-time), and the fitting results obtained are as shown in Figure 20 and 21.

For Saihanba, we found that the growth rate of carbon dioxide absorption is about a power exponential decline trend, and the amount of carbon dioxide absorption shows a logarithmic increase year by year. After the ecological zone is perfected, the growth rate of carbon dioxide absorption will decrease year by year, gradually reaching a stable carbon dioxide absorption rate.

We expect that Australia's capacity to absorb carbon emissions will have the same curve. We draw the same trend curve for Australia's CO2 absorption, which is a logarithmic upward trend. Let us make Australia's C02 absorptive capacity function as shown in Equation 14.

$$y = a \ln (x) + b \tag{14}$$

where y is the amount of CO_2 absorbed, x is the annual carbon dioxide emissions, a and b are the parameters to be estimated.

The least square method is used to optimize the modification to minimize the loss, so as to solve a and b.

For Australia's CO2 absorption growth rate, we also hope that the fitted rate of

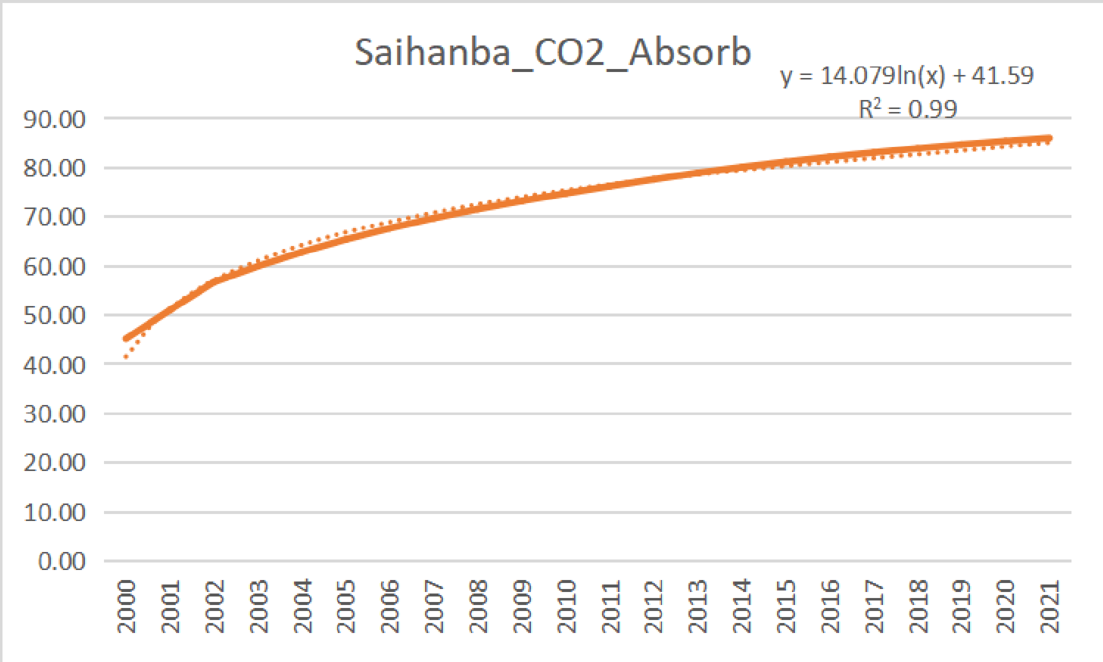

Figure 20 CO_2 **Absorb in Saihanba.**

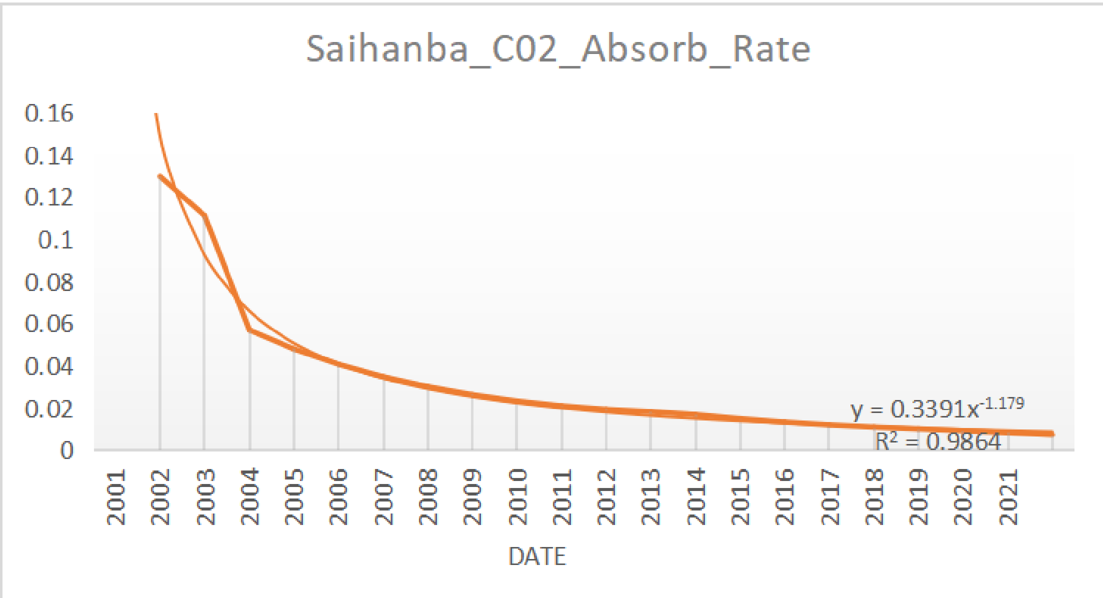

Figure 21 CO_2 **Absorb Rate in Saihanba.**

change of y conforms to the exponential decline as shown in Equation 15.

$$\frac{y' - y}{y} = p^* y^q \tag{15}$$

where y' is the CO_2 absorption of the year, y is the CO_2 absorption of the previous year, p and q are estimated parameters ($q < 0$).

Finally, according to the fitting results, it is determined that the construction and development of the Saihanba Ecological Zone can have a good absorption effect on the carbon emissions of New South Wales, Australia and reduce carbon emissions.

5. Non-technical Report

To: Asia-Pacific Mathematical Contest in Modeling Organizing Committee (APMCM)
From: Team # apmcm210350
Subject: Report to APMCM on Construction of Ecological Reserve
Date: November 28, 2021

Saihanba's impact on the ecological environment is significant. We found a lot of data from the official website, yearbook and papers, and conducted a series of data analysis based on this.

The shocking transformation of Saihanba from a desert to a forest farm has improved the harsh ecological environment of Saihanba itself. We have established three evaluation models, and the unanimous results are that there are significant differences in the environmental conditions before and after the restoration of Saihanba. It can be seen that the establishment of ecological reserves is vital to the change and protection of the ecological environment, and it is also of practical value.

In addition to its extraordinary value, Saihanba Forest Farm has also made significant contributions to Beijing's prevention of sand and dust storms, and even laid the cornerstone for China to achieve carbon neutrality. Through the analysis of variance, we found that the restoration of the Saihanba Forest Farm had a significant effect on the sandstorm in Beijing. This shows that the ecological model of Saihanba is worthy of being extended to the whole country.

In addition, there are many desert areas in the Asia-Pacific region, and Saihanba has set an example for these areas. Taking Australia as an example, we analyzed data such as carbon dioxide and rainfall. We believe that the successful experience of Saihanba can provide unlimited possibilities for Australia to achieve carbon neutrality and improve the ecological environment.

For the construction of ecological protection areas, we have some suggestions and plans as follows:

First, actively study the spirit of relevant policies and documents of the Party Central Committee. China has always emphasized the importance and necessity of ecological protection. General Secretary Xi Jinping is particularly focused on ecological issues and is particularly concerned about my country's goal of achieving carbon neutrality. Therefore, it is necessary to learn the relevant spirit, especially to understand the meaning of ecological protection from the source.

Second, comprehensively consider the regional characteristics and balance the layout of ecological forest land, economic development site and industrial land. Ecological protection is not blindly planting trees. It is true that we support the practice of afforestation to improve the environment. However, urban development requires economic development construction sites and industrial land, and these necessary land cannot be requisitioned as ecological forest land at will. Balanced land use layout needs to combine local characteristics and development needs, and formulate plans in accordance with individual characteristics, and cannot be generalized.

Third, based on local characteristics and actual conditions, we will learn from the successful experience of Saihanba and strive to promote the construction of ecological protection areas. The successful experience of Saihanba is worthy of being promoted, because the contribution of Saihanba Forest Farm to the ecological environment is obvious to all. Each city has its own characteristics and actual conditions. Based on these individual conditions, each city should work hard to speed up the construction of ecological reserves and contribute to China's carbon neutrality.

6. Conclusion

The establishment of ecological protection zones is of great significance to the ecological environment. As far as Saihanba is concerned, changing from a desert to a forest farm has greatly improved the ecological environment of Saihanba itself. At the same time, its changes have had a positive impact on the prevention and control of sandstorms in Beijing. Therefore, the successful experience of Saihanba deserves to be promoted nationwide, and it is also a benchmark for the Asia-Pacific region and even the world.

References

[1] & Chengde Investigation Team of the National Bureau of Statistics Chengde City Statistics Bureau. Chengde yearbook. https://www.aqistudy.cn/historydata/monthdata.php?city=%E6%89%BF%E5%BE%B7. November 28, 2021.

[2] Saihanba Data. National forestry and grassland science data sharing service platform. http://www.forestdata.cn/search-data.html?id=0&key=%25E5%25A1%259E%25E7%25BD%2595%25E5%259D%259D. November 28, 2021.

[3] Y Li. Research on the temporal and spatial distribution pattern of sand and dust disasters in chengde city in the past 50 years. *MSc Thesis*, 2011.

[4] n.d. Gale and sand-dust storm disaster data of 1954-2017 in beijing disaster risk reduction knowledge service. https://www.aqistudy.cn/historydata/monthdata.php?city=%E6%89%BF%E5%BE%B7. November 28, 2021.

[5] n.d. Weather in weichang township. http://rp5.ru/archive.php?wmo_id=54311&lang=cn. November 28, 2021.

[6] n.d. Western china environmental and ecological science data center. http://westdc.westgis.ac.cn. November 28, 2021.

[7] Chengde Yearbook. Chengde city air quality monthly statistics and historical data. https://www.aqistudy.cn/historydata/monthdata.php?city=%E6%89%BF%E5%BE%B7. November 28, 2021.

Appendix

Input Python Code for Grey Association Analysis:

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
data = pd.read_csv("new2014-2021.csv")

# Determine the analysis sequence
Son_sequence = data.loc[:,['Standing_Wood_Total_Accumulation.10.000.cubic.meters.',
        'Artificial_Forest_Area.10.000.hectares.', 'Forest_Coverage_Rate...',
        'Cover_Area.10.000.hectares.', 'Forest_Stock.10.000.cubic.meters.',
        'Artificial_Afforestation_Area', 'Area_Of_Forest_Tending','RRR.Precipitation..mm.']]
Mother_sequence = data.loc[:,['Carbon_Dioxide_Absorption.10.000.tons.', 'Oxygen_Release.10.000.tons.',
        'Tourist_Number.10.000.person.', 'Tourism_Revenue.100.million.yuan.',
        'Water_Conservation.100.million.cubic.meters.',
        'Surface_Water_Quality_Standard_Rate...', 'PM2.5', 'PM10', 'O3', 'NO2',
        'CO', 'Urban_Air_Quality_Standard_Days.Better_Than_Grade_ii.days.',
        'Urban_PM2.5_Concentration.per.microgram.per.cubic.meter.', 'AQI',
        'U.Relative.humidity.at.ground.height.of.2m.....']]

#Normalization
name = Mother_sequence.columns
num = Mother_sequence.shape[1]
for i in range(0,num):
    data = Mother_sequence[name[i]]
    Mother_sequence[name[i]] = [(float(x) - min(data))/(max(data) - min(data)) for x in data]

# get the final score
row_num = Mother_sequence.shape[0]
result = {}
for i in range(0,row_num):
    total = sum(Mother_sequence.iloc[i])
    result[i] = total/num

col_name=Son_sequence.columns.tolist()
col_name.insert(0,'score')
Son_sequence = Son_sequence.reindex(columns=col_name)
df = pd.DataFrame([result])
Son_sequence['score'] = df.T
Son_sequence.head()

# Grey Association
gray = Son_sequence

#scale the data
```

```
gray=(gray - gray.min()) / (gray.max() - gray.min())

std=gray.iloc[:,0]
ce=gray.iloc[:,1:]
n=ce.shape[0]
m=ce.shape[1]

a=zeros([m,n])
for i in range(m):
    for j in range(n):
        a[i,j]=abs(ce.iloc[j,i]-std[j])
c=amax(a)
d=amin(a)

result=zeros([m,n])
for i in range(m):
    for j in range(n):
        result[i,j]=(d+0.5*c)/(a[i,j]+0.5*c)
#Average, get the gray correlation value
result2=zeros(m)
for i in range(m):
        result2[i]=mean(result[i,:])
RT=pd.DataFrame(result2)
```

Input Python Code for Cities Clustering using K-means:

```
#K_Means Algrithum
from sklearn.manifold import TSNE
import pandas as pd
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt
#elbow
#parameter initialization
input_file = r'E:\qq-doc\FileRecv\q2.csv'
output_file = 'res.csv' #result csv
num_k = 3 #number of categories
iteration = 100 #Maximum number of cycles of clustering
data = pd.read_csv(input_file,index_col=0,encoding='ISO-8859-1') #ead
    data
data_zs = 1.0*(data - data.mean())/data.std() #Data normalization
d=[]
for i in range(1,10):
    estimator = KMeans(n_clusters=i,max_iter=iteration)  #
        Constructing the clusters
    estimator.fit(data_zs)
    d.append(estimator.inertia_)
plt.plot(range(1,10),d,marker='o')
plt.xlabel('number of clusters')
plt.ylabel('distortions')
plt.savefig('elbow.svg')
plt.show()
model = KMeans(n_clusters = 3, max_iter = iteration) #According to
    the elbow rule, there are 3 categories
model.fit(data_zs) #fit data

#Print Results
r1 = pd.Series(model.labels_).value_counts() #Count the number of
    each category
```

```
r2 = pd.DataFrame(model.cluster_centers_) #Identify cluster centers
r = pd.concat([r2, r1], axis = 1) #Get the number under the category
    corresponding to the clustering center
r.columns = list(data.columns) + [u'categories']

r = pd.concat([data, pd.Series(model.labels_, index = data.index)],
    axis = 1)
r.columns = list(data.columns) + [u'clustering']
r.to_csv(outputfile) #save result
#Data dimensionality reduction with TSNE and presentation of
    clustering results
tsne = TSNE()
tsne.fit_transform(data_zs) #Data Dimensionality Reduction
tsne = pd.DataFrame(tsne.embedding_, index = data_zs.index) #Convert
    data format

import matplotlib.pyplot as plt
plt.rcParams['font.sans-serif'] = ['SimHei']
plt.rcParams['axes.unicode_minus'] = False

d = tsne[r[u'clustering'] == 0]
plt.plot(d[0], d[1], 'r.')
d = tsne[r[u'clustering'] == 1]
plt.plot(d[0], d[1], 'go')
d = tsne[r[u'clustering'] == 2]
plt.plot(d[0], d[1], 'b*')
plt.savefig('maomao.svg')
plt.show()
```

Input Python Code for the Extension Model:

```
import pandas as pd
data = pd.read_csv("Q3_TEST.csv")
data = data.drop("Province",axis = 1)
weights = [0.4,0.2,0.2,0.2]
num = data.shape[1]
name = data.columns
ecology = data
#normalization
for i in range(0,num):
    data = ecology[name[i]]
    ecology[name[i]] = [((float(x) - min(data))/(max(data) - min(data
        ))+1) for x in data]
ecology['score'] = 0
sample_num = data.shape[0]
re = {}
for i in range(0,sample_num):
    re[i] = 0.05*ecology['CO2_Release'][i]+ 0.05*(ecology['AVG_PM25'
        ][i]+ecology['AVG_PM10'][i]+ecology['AVG_SO2'][i]+ecology['
        AVG_NO2'][i]+ecology['AVG_CO'][i]+ecology['AVG_O3'][i])+ 0.8*
        ecology['Area( square kilometers)'][i]+ 0.1*(ecology['Flow
        Area'][i]+ecology['Semi-flow Area'][i]+ecology['Semi-fix Area'
        ][i]+ecology['Fix Area'][i])
df = pd.DataFrame([re])
ecology['score'] = df.T
data = pd.read_csv("Q3_TEST.csv")
data = data.drop("Province",axis = 1)
saihanba = 706
```

```
rate = ecology['score'][0]
city = data['Fix Area'][0]
result = {}
for i in range(1,sample_num):
    result[i] = saihanba * ecology['score'][i] / rate
print(result)
ecology['ecology_area'] = 0
ecology['ecology_area'][0] = saihanba
df = pd.DataFrame([result])
df = df.T
```

	Total_Area(10,000 hectares)	Stand_Area(10,000 hectares)	Standing_Wood_Total_Accumulation(10,000 cubic meters)	Artificial_Forest_Area(10,000 hectares)
count	7	7	7	7
mean	10.24	4.014285714	76.65714286	1.7
std	0.021602469	0.371880677	18.9777692	0.64807407
min	10.21	3.5	50.3	0.8
25%	10.225	3.755	63.48	1.25
50%	10.24	4.01	76.66	1.7
75%	10.255	4.275	89.835	2.15
max	10.27	4.53	103.01	2.6

	Forest_Stock(10,000 cubic meters)	Carbon_Dioxide_Absorption(10,000 tons)	Oxygen_Release(10,000 tons)	Water_Conservation(100 million cubic meters)
count	21	21	21	21
mean	61.70714286	5.120952381	3.561428571	0.159047619
std	26.49582592	2.19823089	1.529664296	0.057784493
min	33	2.74	1.9	0.09
25%	38.53	3.2	2.22	0.11
50%	54.57	4.53	3.15	0.15
75%	80.35	6.67	4.64	0.21
max	115.06	9.55	6.64	0.26

	Sand_Storm	Precipitation_Grade	Surrival_Rate_Of_Afforestation	Mean_Annual_Precipitation
count	21	18	18	21
mean	1.157142857	2.388888889	72.34777778	427.7190476
std	1.159556565	0.849836586	24.92648191	13.79469532
min	0	1	28.28	413.6
25%	0	2	69.3625	413.6
50%	0.9	3	80.33	440.5
75%	2.4	3	89.85	440.5
max	3	3	98.84	440.8

	Forest_Coverage Rate(%)	Cover_Area(10,000 hectares)	Floating_Dust	Blowing_Sand
count	21	21	21	21
mean	33.53857143	46.9547619	11.37142857	23.11428571
std	11.28637731	15.80006475	5.087252978	5.040464831
min	13.57	19	6	13.6
25%	24.83	34.76	6.5	19.6
50%	34.6	48.43	10.2	23.8
75%	42.92	60.09	15.8	26.8
max	49.86	69.8	20	31

Figure 22 Descriptive statistics from 1962 to 1982.

	Total_Area(10,000 hectares)	Stand_Area(10,000 hectares)	Standing_Wood_Total_Accumulation(10,000 cubic meters)	Artificial_Forest_Area(10,000 hectares)
count	9	9	9	9
mean	9.422222222	7.206666667	856.2133333	5.706666667
std	0.058476016	0.074161985	74.88426737	0.074161985
min	9.33	7.08	743.66	5.58
25%	9.38	7.16	802.62	5.66
50%	9.43	7.22	858.36	5.72
75%	9.47	7.26	910.88	5.76
max	9.49	7.3	963.4	5.8

Figure 23 Descriptive statistics from 2005 to 2013.

	Total_Area(10,000 hectares)	Stand_Area(10,000 hectares)	Standing_Wood_Total_Accumulation(10,000 cubic meters)	Artificial_Forest_Area(10,000 hectares)	Forest_Coverage_Rate(%)
count	8	8	8	8	8
mean	9.3325	7.48	1007.14	5.89	79.14875
std	0.0046291	0.09797959	23.8090403	0.048989795	2.337400494
min	9.33	7.34	973.12	5.82	75.52
25%	9.33	7.41	990.13	5.855	77.5925
50%	9.33	7.48	1007.14	5.89	79.375
75%	9.3325	7.55	1024.15	5.925	80.86
max	9.34	7.62	1041.16	5.96	82.21

Figure 24 Descriptive statistics from 2014 to 2021.

AQI	Air Pollution Level	Health Implications	Cautionary Statement (for PM2.5)
0 - 50	Good	Air quality is considered satisfactory, and air pollution poses little or no risk	None
51 -100	Moderate	Air quality is acceptable; however, for some pollutants there may be a moderate health concern for a very small number of people who are unusually sensitive to air pollution.	Active children and adults, and people with respiratory disease, such as asthma, should limit prolonged outdoor exertion.
101-150	Unhealthy for Sensitive Groups	Members of sensitive groups may experience health effects. The general public is not likely to be affected.	Active children and adults, and people with respiratory disease, such as asthma, should limit prolonged outdoor exertion.
151-200	Unhealthy	Everyone may begin to experience health effects; members of sensitive groups may experience more serious health effects	Active children and adults, and people with respiratory disease, such as asthma, should avoid prolonged outdoor exertion; everyone else, especially children, should limit prolonged outdoor exertion
201-300	Very Unhealthy	Health warnings of emergency conditions. The entire population is more likely to be affected.	Active children and adults, and people with respiratory disease, such as asthma, should avoid all outdoor exertion; everyone else, especially children, should limit outdoor exertion.
300+	Hazardous	Health alert: everyone may experience more serious health effects	Everyone should avoid all outdoor exertion

Figure 25 AQI Explanation.