# Ecoregion Model Based on Fitting and Clustering

Jiachen Yang [a], Lu Yang [b] Shilei Shi [c]

School of Economics, Shandong University of Technology, Zibo, China, 255049

[a]2917756199@qq.com, [b]lu_j1974@126.com, [c]shishilei1234@126.com

**Abstract.** With the gradual restoration of Saihanba forest, the positive ecological impact of Saihanba on Beijing, Tianjin and other surrounding areas has become more and more obvious, such as water supply and sandstorm conditions have been improved to a considerable extent. In this paper, the environmental impact assessment model of Saihanba is established, the fitting equation of forest accumulation and growth is constructed, and nine three-level indicators are selected from three aspects of afforestation, water saving and atmospheric regulation, and are calculated by entropy weight method and TOPSIS method. Comprehensive score of ecological environment. At the same time, this paper builds an index system of Beijing's ability to resist sand and dust storms, and uses Spearman's correlation coefficient to analyze the correlation between Saihanba Forest Farm and Beijing's ability to resist sand and dust storms, and conduct a comprehensive evaluation of Beijing's dust-proof ability. The results show that with the restoration of Saihanba, Beijing's sandstorm defense capability has been improved. Secondly, this paper constructs a cluster analysis model based on carbon emission absorption rate, and uses regression fitting method to predict the carbon emission absorption rate of ecological regions in the next few years. The results show that by 2060, the establishment of ecological regions can achieve carbon neutrality in China. and target. And based on this model, the major cities in Vietnam are determined to determine the geographical location of the eco-regions, and the eco-protected areas are calculated according to the percentage of air quality improvement in each city.

**Keywords:** Forest Ecoenvironment Evaluation, Clustering analysis, fitting analysis, entropy weight method.

## 1. Introduction

With the implementation of the sustainable development strategy and the introduction of the concept that "clear waters and lush mountains are invaluable assets", China's determination to protect the ecological environment has become stronger and stronger. In recent years, the state has introduced a series of policies to coordinate mechanisms and build an ecological civilization system, so as to promote the economic and social transformation to green growth in all respects and build China into a modern country that is prosperous, strong, culturally advanced, harmonious and beautiful.

Since the establishment of Saihanba Mechanical Forest Farm in Hebei Province in 1962, the largest artificial forest farm in the world has been built after three generations of hard work and continuous struggle. The original forest area in Saihanba has increased from 240,000 mu to 1.151,000 mu, and the forest coverage rate has increased from 11.4%to 82%. The forest volume has increased from 330,000 cubic meters to 10,368,000 cubic meters now. As a result, Saihanba can provide 1.37*108m3 of clean water, 7.47*106t of carbon sequestration and 5.45*106t of oxygen for The Beijing Tianjin.

Therefore, this paper establishes the Saihanba ecological environment impact assessment model, evaluates the Saihanba area, and quantitatively evaluates the environmental impact after the restoration of Saihanba. At the same time, a mathematical model of the influence of Saihanba on the sand-dust storm capability of Beijing was established, and the impact of Saihanba on the sand-dust storm resistance capability of Beijing was quantitatively evaluated. Under the analysis of this research, determine which geographical locations in China need to be built as ecological reserves and their scale or number. And select a country in the Asia-Pacific region, discuss the geographical location and the number or scale of the establishment of ecological reserves in the country, and then evaluate its impact on greenhouse gas absorption and carbon reduction.

# 2. Model establishment and analysis

## 2.1. Evaluation model of Saihanba's impact on environment

### 2.1.1. The equation of forest accumulation growth

By observing the data of forest stock change with years in Saihanba, it is obvious that abnormal values can be found, and we eliminate the unreasonable data. Then, according to the data, the fitting curve of forest stock with the change of year is made as shown in Figure 1below.

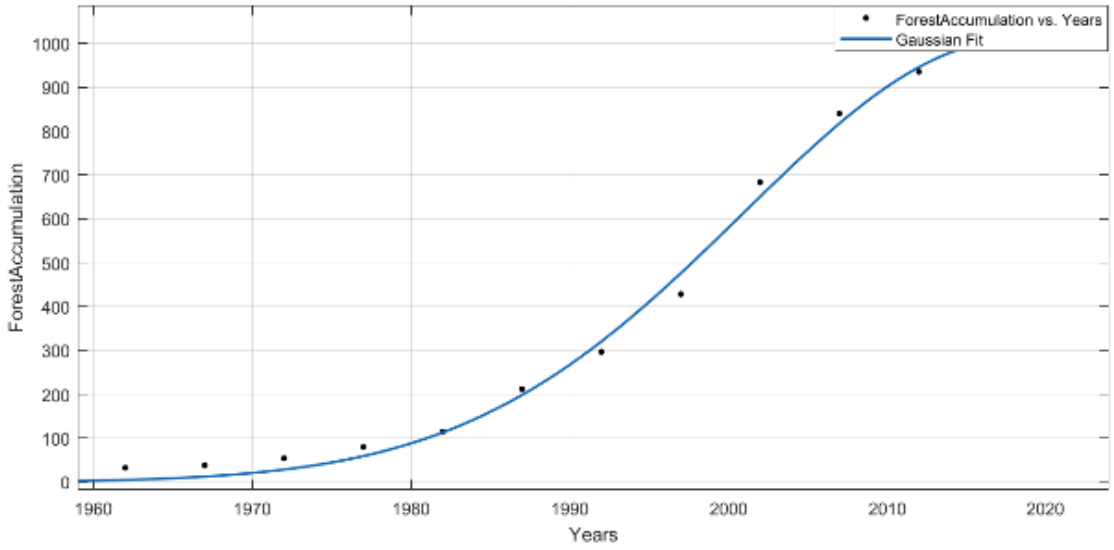



**Figure 1.** Fitting curve

It can be roughly observed from the image that with the migration of years, the forest stock also increases, and the increasing speed increases first and then decreases. We can obtain the following fitting equation:

**Table 1.** The fitting equation and its $R^2$

| Fitting equation | $R^2$ |
|---|---|
| $y = a_1 e^{-\left(\frac{x-b_1}{c_1}\right)^2} + a_2 e^{-\left(\frac{x-b_2}{c_2}\right)^2}$ | 0.9962 |

Among them:

$a_1 = 5.917e+15\,(-9.468e+31, 9.468e+31)$; $b_1 = 2037\,(-7.157e+15, 7.157e+15)$;

$c_1 = 2.836\,(-5.577e+14, 5.557e+14)$; $a_2 = 1009\,(808.2, 1210)$;

$b_2 = 2018\,(2010, 2027)$; $c_2 = 24.52\,(18.49, 30.54)$

### 2.1.2. Ecological Environmental Impact Assessment System

In order to analyze and establish the evaluation model of Saihanba's impact on ecological environment [4], we need to establish an ecological environmental impact evaluation system. Afforestation factors [1], water conservation factors and air regulation factors were selected as secondary indicators. Forest cover rate, cover area, forest accumulation [2], conservation of water, average annual precipitation [3], annual runoff, fixed amount of carbon dioxide, oxygen release and gale days were taken as three indexes [4].

| Primary indicators | Secondary indicators | Third level ndicators |
|---|---|---|
| The impact on the environment after the restoration of Saihanba | Afforestation elements | Forest cover rate |
| | | Covered area |
| | | Forest accumulation |
| | Water conservation factors | Conservation of water |
| | | Annual runoff |
| | | Average annual precipitation |
| | Atmospheric regulating element | CO2 absorption |
| | | O2 release |
| | | Windy days |

**Figure 2.** Construction of ecological environment index system

### 2.1.3. Topsis optimal solution distance method

TOPSIS algorithm is a commonly used multi-objective and finite scheme comprehensive evaluation method. This method makes full use of the original data information to calculate the comprehensive evaluation value of each scheme, and ranks each scheme according to the value, so as to accurately reflect the gap between each evaluation scheme. The entropy weight rule is an objective weight assignment method. The smaller the variation degree of the index is, the smaller the amount of information is reflected, and the lower the corresponding weight is.

### 2.1.4. Evaluation model of Saihanba's impact on environment

In this paper, the nine three-level indicators obtained above are imported, and the weight of the indicator is obtained through the information reflected in the indicator. Therefore, the Saihanba ecological environment impact assessment model can be established.

$$Greening = 0.03X_1 + 0.03X_2 + 0.16X_3 \tag{1}$$

$$Watering = 0.17X_4 + 0.16X_5 + 0.16X_6 \tag{2}$$

$$Atmousphere = 0.13X_7 + 0.15X_8 + 0.01X_9 \tag{3}$$

$$Score = 0.22Greening + 0.49Watering + 0.29Atmosphere \tag{4}$$

The data of each level 3 index every five years from 1966 to 2020 reflect the corresponding level 2 index, and the level 2 index ultimately plays a role in the eco-environmental impact score. The ecological environmental impact scores over the years after normalization are shown in the table below.

**Table 2.** Ecological environment score of each year

| Year | Score | Ranking |
|---|---|---|
| 1962 | 0.0001 | 13 |
| 1967 | 0.0064 | 12 |
| 1972 | 0.0133 | 11 |
| 1977 | 0.0209 | 10 |
| 1982 | 0.0292 | 9 |
| 1987 | 0.0417 | 8 |
| 1992 | 0.0549 | 7 |
| 1997 | 0.0731 | 6 |
| 2002 | 0.1070 | 5 |
| 2007 | 0.1281 | 4 |
| 2012 | 0.1435 | 3 |
| 2017 | 0.1591 | 2 |
| 2020 | 0.2227 | 1 |

As can be seen from the table above, with the increase of years, the score of Saihanba's impact model on ecological environment keeps rising [5], and the specific change trend is shown in the table below:

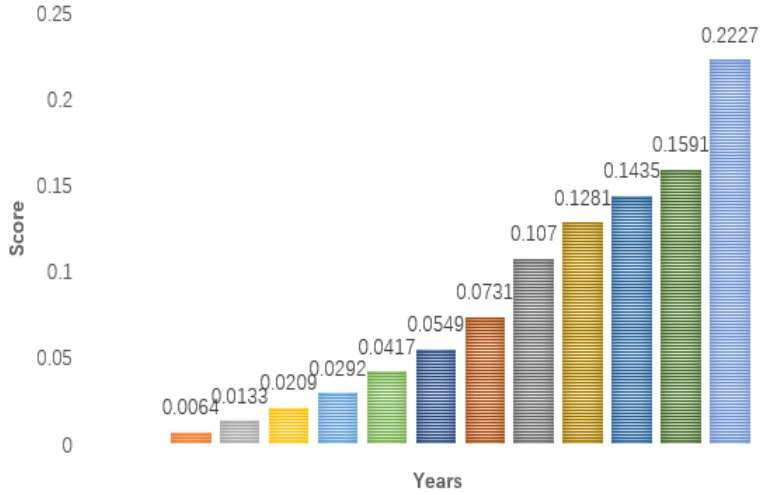



**Figure 3.** Bar chart of scores in each year

Therefore, the mean of model score before 2000 is 0.029938, and the mean of model score after 2000 is 0.15208. After the quotient of model score before and after 2000, the average score after 2000 is 5.079916 times, that is, before and after the restoration of Saihanba, its ecological environment has been significantly improved. The quality of the ecological environment is five times higher than before.

### 2.2. Sandstorm resistance model in Beijing based on Spearman correlation coefficient

#### 2.2.1. Index selection and relationship establishment

This paper first studies the relationship between the restoration of Saihanba Forest Farm and Beijing's resistance to sandstorms, and then studies the impact of Saihanba Forestry restoration on Beijing's resistance to sandstorms. Therefore, the forest area and forest coverage rate were selected as the indicators reflecting forest restoration, and the air quality index, average wind speed, number of days with strong winds, and average PM10 concentration were selected as indicators reflecting the ability of Beijing to resist sand and dust storms.

#### 2.2.2. The Spearman correlation coefficient

In order to measure the linear relationship between two variables, Pearson correlation coefficient analysis or Spearman correlation coefficient is generally used for analysis. Pearson's correlation

coefficient can be used to describe the correlation between variables only when the two variables are continuous and linear, and Pearson's correlation coefficient can be hypothesis tested only when the two variables are normally distributed. In contrast, Spearman correlation coefficient does not use restrictions. This data was analyzed by Spearman coefficient because both linear relationship and normal distribution test were passed.

Definition: X and Y are two groups of data, and their Spearman correlation coefficient is

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \tag{5}$$

Among them, $d_i$ is the Poor grades betwwen $X_i \& Y_i$.

From the above calculations, it can be concluded that there is a significant negative correlation between the restoration of the Saihanba forest farm and the indicators reflecting the ability of Beijing to resist sand and dust storms. In other words, with the continuous recovery of Saihanba Forest Farm, the frequency of sandstorms in Beijing is getting lower and lower, and Beijing's ability to resist sandstorms is getting stronger and stronger.

### 2.2.3. Scoring of anti-sand dust ability in Beijing based on entropy weight method

Referring to the above-mentioned evaluation model of the impact on the ecological environment before and after the restoration of Saihanba, this paper also uses the entropy weight method to score the sand and dust resistance of Beijing, and obtains the quantitative index of Beijing's sand and dust resistance around 2000, and obtained the following weights and scores through entropy weight method.

**Table 3.** Index weight

| Index | Weight | Index attribute |
|---|---|---|
| AQI | 0.1975 | Negative index |
| Average Wind Speed | 0.1412 | Negative index |
| Wind Days of a Year | 0.1904 | Negative index |
| Mean Concentration of PM10 | 0.4709 | Negative index |

It is easy to observe that the annual average of Beijing's sandstorm resistance ability before 2000 is 0.0393, and that of the model after 2000 is 0.17644. The average score after 2000 is 4.489567 times of that before 2000. That is to say, before and after Saihanba is restored, Beijing's ability to defend against sandstorms has been significantly improved, reaching nearly 4.5%. The specific change trend is shown in the figure below.

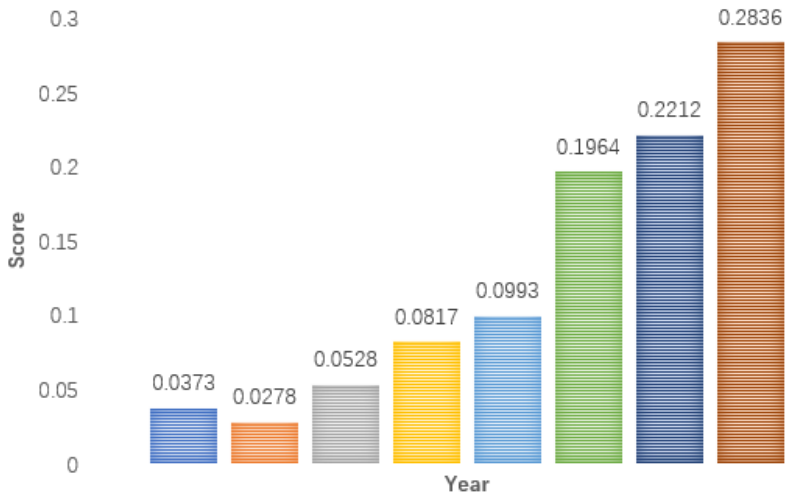



**Figure 4.** Bar chart of scores in each year

### 2.3. Cluster Analysis Model Based on Carbon Emission Absorption Rate

This paper, based on the calculation of carbon emissions in each province, determines which geographic locations need to establish ecoregions, and establishes ecoregions in areas with high carbon emissions to achieve carbon neutrality. According to the data, the carbon emission coefficients of coal, oil, natural gas and other fuels in each province are shown in the following table:

**Table 4.** Carbon Emission Coefficient

| Type of fuel | Coal | Oil | Natural gas |
|---|---|---|---|
| Carbon emission coefficient | 0.7329 | 0.5574 | 0.4226 |

Then, by querying the annual coal, oil, natural gas and other fuel consumption of each province in China and weighting them according to the following formula, the carbon emission formula of each province can be obtained:

$$W = \sum \alpha_i \cdot q_i \tag{6}$$

Among them, $W$ is Carbon emissions, $\alpha_i$ is Energy carbon emission coefficient, $q_i$.is Coal, oil, natural gas and other fuels used.

It is known that each hectare of forest can absorb 1 ton of carbon per day, the formula of carbon absorption of each province can be obtained:

$$Q = \beta \cdot S \tag{7}$$

There Q is carbon absorption, S is the amount of carbon dioxide converted by forest per unit area, and S is the forest area in the province.

Carbon sequestration - emission conversion rate as shown below

$$H = \begin{cases} \frac{Q}{W} & , \ Q < W \\ 1 & , \ Q \geq W \end{cases} \tag{8}$$

#### 2.3.1. The transformation ratio of each province was analyzed by cluster analysis.

This paper builds a cluster analysis-based model to identify which provinces have greater demand for ecoregions based on the transitions in each province.

Cluster analy[sis] [6], also known as group analysis, is a statistical analysis method to study (sample or index) classification problems, and is also an important algorithm of data mining. Cluster analysis is based on similarity and there are more similarities between patterns in a cluster than between patterns not in the same cluster.

The corresponding conversion rate H of each province obtained from equation (8) was analyzed by clustering with SPSS, and the results are as follows

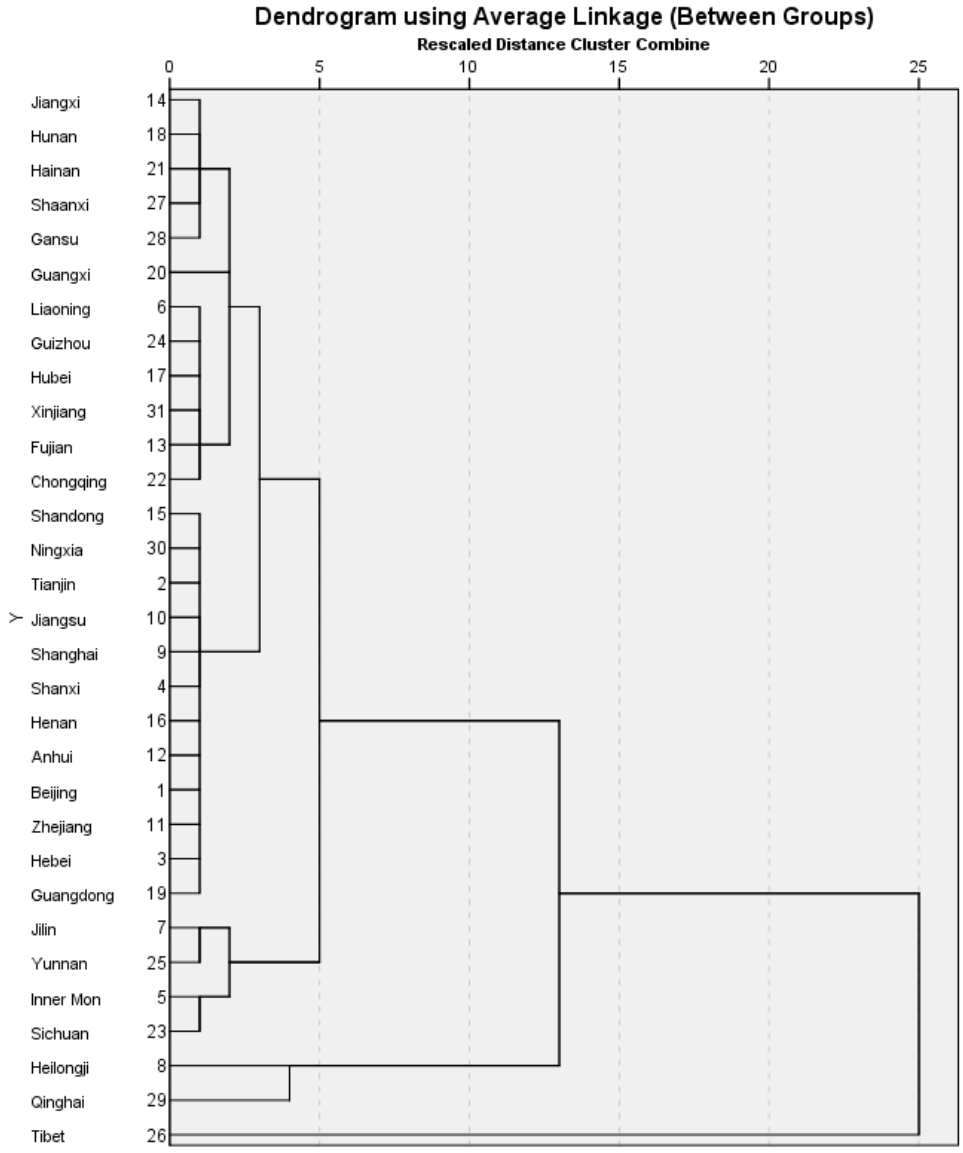



**Figure 5**. Cluster analysis results

Through the clustering model, we found that Shandong and Ningxia, Tianjin, Jiangsu, Shanghai, Shanxi, Henan, Anhui, Beijing, Hebei, Hebei, Guangdong and other 12 provinces and cities of carbon absorption [7] lower emissions conversion, there is a great need to establish ecological zones.



**Figure 6.** Location of proposed ecological zone

According to the above analysis of the established model, three ecological protected areas are planned to be built: one is a large ecological protected area in East China, covering Zhejiang, Shanghai, Jiangsu, Anhui, Henan and Hebei. The other is a medium-sized ecological reserve in northern China, including Ningxia, Tianjin, Shanxi, Beijing and Hebei. Finally, there are small-scale protected areas in southern China, including Guangdong.

### 2.3.2. Assess the impact of establishing ecological zones on carbon neutrality in China

China's policy goal is to achieve "carbon peak" in 2035 and "carbon neutrality" in 2060 [8-10]. This paper uses a clustering model, and the carbon sequestration carbon conversion rates of 12 provinces and cities, including Shandong, Ningxia, Tianjin, Jiangsu, Shanghai, Shanxi, Henan, Anhui, Beijing, Zhejiang, Hebei, and Guangdong, have high similarity, so , this paper will take Shandong as the representative, refer to the volume of the ecological zone initially established in Saihanba Forest Farm, under the condition of establishing an ecological zone of 1.5 million mu and no ecological zone [11], to simulate Shandong, and predict the carbon dioxide conversion in the next few years. The fitted curve is shown in the figure below.

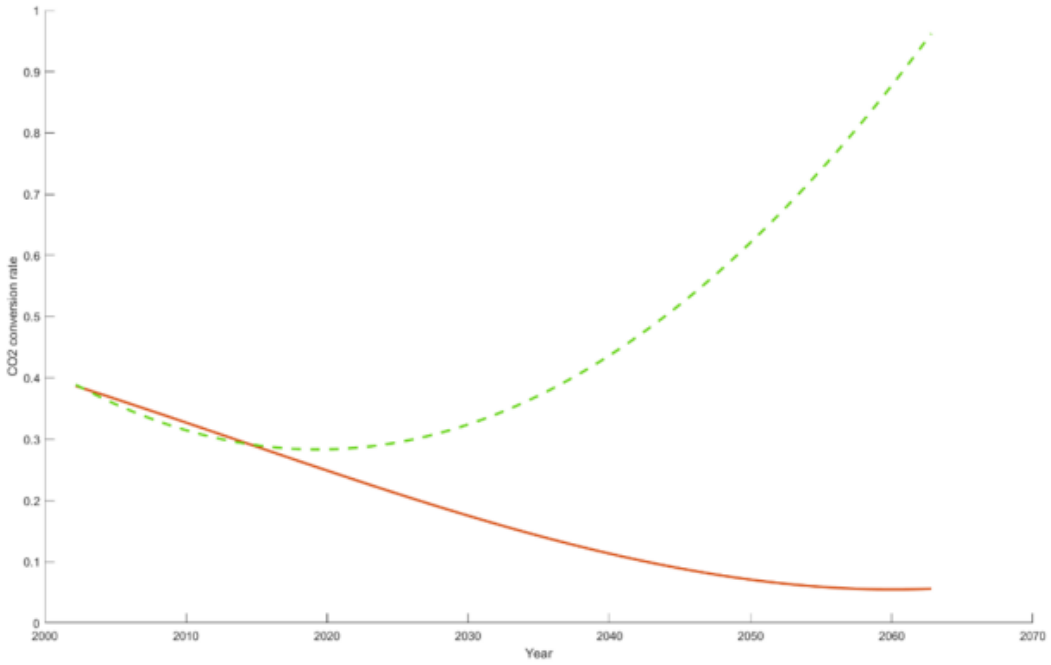



**Figure 7.** Fitting curve

From the figure, we know that the establishment of ecological protection zone in Shandong will certainly promote the goal of carbon neutrality in 2060, and the "carbon peak" can also be achieved in advance. By magnifying the above prediction to 12 provinces, we can obtain the predicted value of our ecological zone construction, and thus we can predict that by 2060, China's carbon neutrality target will be basically achieved.

## 3. Model promotion

This paper extends the above model and selects Vietnam as the research target by looking at the ranking of countries with the worst air quality provided by WHO on the IQAir website and combining the actual situation of countries in the Asia-Pacific region.

### 3.1. Establish model for finding cities need protect

To build ecological protection zones in certain cities in Vietnam, it is necessary to calculate the indicators of their respective large cities [12]. In cities with high AQI values, large enough urban

areas and low forest coverage, ecological protection zones should be established in order to reduce carbon.

Let the air quality index of each city be qi, sort all the qi values of the eleven large cities collected from small to large and take their quantile.

$$M_{\phi}=\begin{cases} x_{[np+1]} & np\notin\{k|0,1,2,...\} \\ \frac{1}{2}(x_{[np]}+x_{[np+1]}) & np\in\{k|0,1,2,...\} \end{cases} \tag{9}$$

Let p = 0.75, then $M_{0.75}$=76

When $q_i$>=$M_{0.75}$, We believe that the air quality in this city is poor, and it is necessary to establish an ecological protection zone in this city for ecological protection.

### 3.2. Statistics of the collected air quality indexes of major cities

#### 3.2.1. Solve the percentage of air quality improvement in each city

$$W_i=\frac{(q_i-M_{0.75})}{q_i}\times 100\% \tag{10}$$

$W_i$ Indicates the percentage improvement of air quality in each city [13].

#### 3.2.2. Calculation of the area of ecological protection zone

According to relevant data, there is a strong correlation between air quality and urban forest coverage. It can be assumed that the percentage of air quality improvement and the increase in forest coverage. Let the original forest coverage rate of each city be hi and the land area be $S_i$

Then the area of the ecological protection zone is

$$\Delta S_i=S_i\cdot h_i\cdot(1+w_i)-S_i\cdot h_i=S_i\cdot h_i\cdot w_i \tag{11}$$

#### 3.2.3. The analysis about carbon after setup of ecological reserve

$$\Delta S_i^{'}=\begin{cases} \frac{t}{t_2}\Delta S_i & 0<t\le t_2 \\ \Delta S_i & t>t_2 \end{cases} \tag{12}$$

Querying data shows that the carbon absorption index per hectare of forest is μ, then the carbon dioxide absorption Q of the newly-added ecological protection zone in the city.

Therefore, we can predict the amount of carbon dioxide absorption in the future for many years, and the fitting curve is shown in the figure below.

$$Q=u\Delta S_i^{'}=\begin{cases} \frac{ut}{t_2}\Delta S_i & 0<t\le t_2 \\ u\Delta S_i & t>t_2 \end{cases} \tag{13}$$

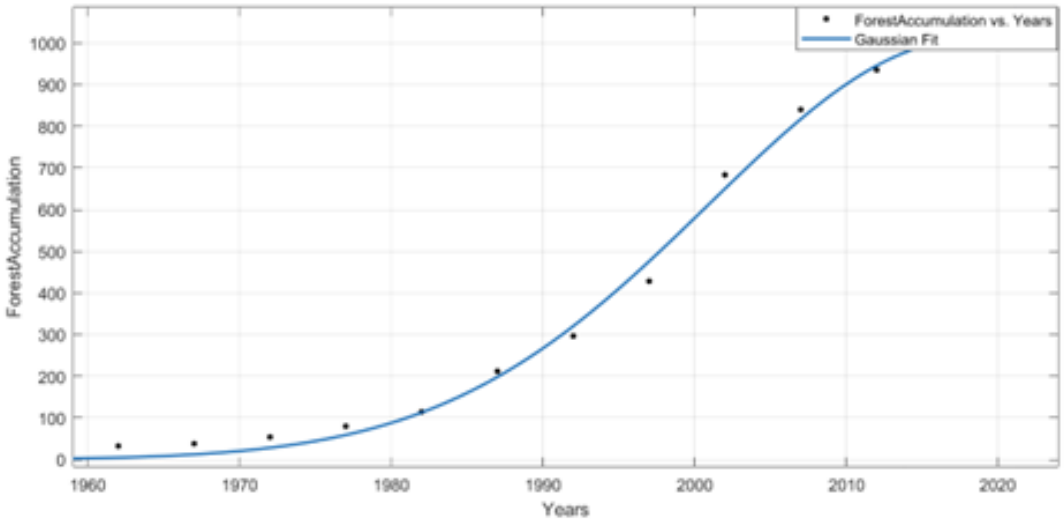



**Figure 8.** The fitting results

From this, we can predict that by 2035, the contribution of Vietnam's ecological reserves to CO2 will stabilize. In other words, the establishment of ecological zones can promote the absorption of greenhouse gases to reduce carbon emissions.

## 4. Evaluation of the model

Under the condition that multiple factors jointly affect the ecological environment, this paper analyzes multiple indexes, and divides several three-level indexes under each dual index, which makes the logic more rigorous and the conclusion more convincing. In terms of data sources, this paper selects data from several economic databases and references to ensure reliable and abundant data sources. And data visualization technology, with the help of graphic visualization method to make the data display more intuitive, efficient. This paper not only carries on the quantitative analysis to the problem, but also combines the reasonable qualitative analysis to make the main logic of the paper more perfect. In addition, in order to improve the accuracy of the problem, the algorithm is compared and analyzed, so as to improve the overall accuracy of the problem

## References

[1] Yang Huijuan, Fan Dongdong, Yu Xiaohong, Evaluation of visual landscape quality of Saihanba Forest based on GIS, Journal of Northwest Forestry College, 2020, 35 (5), 225 – 232.

[2] Wang Dong, Feng Kaibin, Yin Xin, Zhang Hongwei, Li Shuang, Analysis and evaluation of forest resource quality in Saihanba Mechanical Forest Farm, Hebei Forestry Science and Technology, 2020 (3), 1.

[3] Ren Yanlin, Wavelet analysis of precipitation variation in Saihanba Region from 1965 to 2011, Journal of Peking University (Natural Science), 2012(11), 48, 6.

[4] T V Bogachev; Bogachev T V; Alekseychik T V; Stryukov M B;Yakovlev A V, Construction of an economic and environmental assessment of the state of regional forestry in the Russian Federation based on the methods of the fuzzy set theory, IOP Conference Series: Earth and Environmental Science, 2020 (11),

[5] Chang Weiqiang, Analysis on dynamic change of forest resources in Saihanba Mechanical Forest Farm, Forest resource management, 2018 (12), 6.

[6] Jiang Dahai, Preliminary study on quantitative assessment and risk management of Sandstorm disaster, Lanzhou University, 2013 (05).

[7] Gui Hailin, Qin He, Zhao Peitao, Wang Fei, Tang Zhijun, Wang Jikang, Zhu Yuanyuan, Chu Yangxi, Boundary layer characteristics and source analysis of a dust weather in Beijing in spring 2018, Journal of Meteorology and Environment, 2021 (8), 37, 4.

[8] Deng Xu, Xie Jun, Teng Fei, what is carbon neutral?, Advances in climate change research, 2021 (01), 17, 1.

[9] Si Shoukui, Mathematical Modeling Algorithms and Procedures, Naval Aeronautical and Engineering College, 2007 (09).

[10] Wang Can, Zhang Yaxin, The realization path and policy system of carbon neutral vision, Environmental Management in China, 2020 (12), 58 - 64.

[11] Zhang Yaxin, Lu Huilin, Wang Can, Analysis of international trends in carbon neutrality initiatives, Advances in climate change research, 2021 (01), 17, 1.

[12] Li Jun, Vietnam is facing an ecological crisis, Indochina, 1986 (02); 13 - 15.

[13] Jiao Nianzhi, Research and development of Marine "negative emissions" technologies to support China's demand for "carbon neutrality", Proceedings of the Chinese Academy of Sciences, 2021 (01), 179 - 187.