Prediction of global temperature based on the BP neural networks

Zhonglin Huang [#], Junjie Wang [*,#], Zedan Gongbo [#], Qiao Li [#], Yutong He [#]

School of Network and Communication Engineering, Chengdu Technological University, Chengdu, 611730

* Corresponding Author Email: junjie20230116@163.com

[#]These authors contributed equally.

Abstract. The problem of global warming is getting worse as global temperature has risen year by year to an alarming level. This paper is first based on historical data on global temperatures, two prediction models were developed using Gray Prediction Method and BP Neural Network model, whose results show that the root mean square error test value of BP neural network is about 0.03 lower than that of the gray prediction method (0.1627).Thus, the BP neural network prediction model is proved effective, and then the model was analyzed by gray correlation theory, which found that carbon dioxide concentration, forest cover area and human population are affecting the global temperature, and the Gray correlation of the fisrt two factors is as high as 0.8, which makes the two the main factors. In the end the paper gave measures to contain global warming.

Keywords: global temperatures, Gray Prediction Method, BP Neural Network.

1. Introduction

At the time that the temperature around the world is increasing year by year, it is of great significance to accurately predict the global temperature and study the temperature changes to understand the future trend and trend of the climate anomaly [1]. Short-term annual average temperature forecast is an important component of showing the future global temperature change trend. The prediction factors directly depend on the carbon dioxide emissions over the years, the heat absorption of the earth and the change of ocean surface temperature and other factors. The climate prediction model [2-3] of BP neural network is adopted, and the prediction results are more accurate. In conclusion, this paper will adopt two models of gray prediction [4] and BP neural network prediction, and analyze the main factors affecting the temperature change through gray association method.

2. The basic funamental of BP Neural Network

2.1. Basic Principles of BP Neural Network

BP neural network has arbitrary complex mode classification ability and excellent multidimensional function mapping ability. Structurally speaking, BP network has an input layer, hidden layer and output layer [5].

The basic BP algorithm includes two processes: forward propagation of signal and back propagation of error. That is, the error output is calculated in the direction from input to output, while the adjustment weight and threshold value are calculated in the direction from output to input. The grid of the neural network learning mechanism is shown in Figure 1:

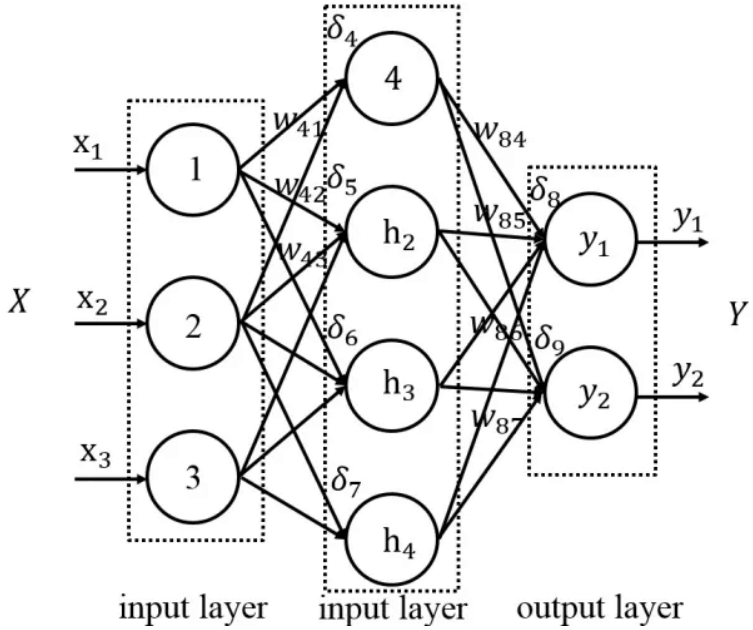


Figure 1. grid of the neural network learning mechanism

The algorithm flow chart of BP neural network is mainly composed of input layer, hidden layer, output layer, etc. See the following flow chart of the Figure 2.

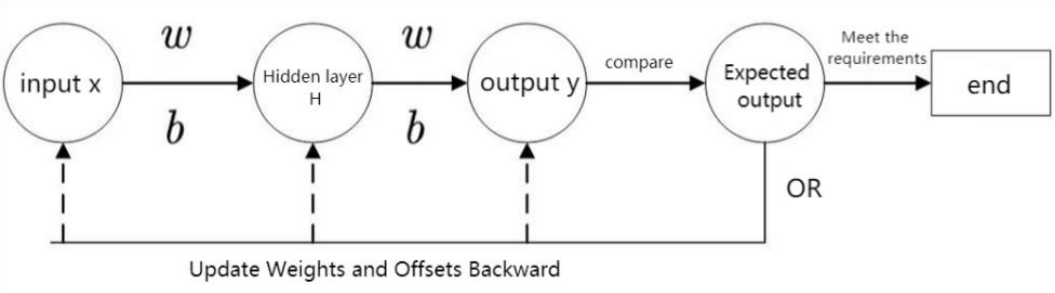


Figure 2. Algorithm flow chart

*Step*1**:** Conduct grid initialization

The random number in each connection weight is given to an interval (-1,1) for each connection weight, and then the error function is set, and then the maximum accuracy value and the maximum learning number are given. $e \varepsilon M$

*Step*2**:** randomly select the first input sample and the corresponding expected value output

$$\begin{aligned} x(k) &= (x_1(k), x_2(k), \cdots, x_n(k)) \\ d_0(k) &= (d_1(k), d_2(k), \cdots, d_q(k)) \end{aligned} \tag{1}$$

*Step*3**:** Calculate the input and output of each neurons in the hidden layer

$$hi_h(k) = \sum_{i=1}^{n} w_{ih} x_i(k) - b_n, h = 1, 2, \cdots, p \tag{2}$$

$$ho_h(k) = f(hi_h(k)), h = 1, 2, \cdots, p \tag{3}$$

$$yi_o(k) = \sum_{h=1}^{p} w_{ho} ho_h(k) - b_o, o = 1, 2, \cdots, q \tag{4}$$

$$yo_o(k) = f(yi_o(k)), o = 1, 2, \cdots, q \tag{5}$$

*Step*4 : Using the desired input and the actual output of the grid, calculate the partial derivative of the error function on each neurons in the output layer, and then establish weights with the input and output of each neurons in the hidden layer $\delta_0(k)$

*Step*5 **:** Calculate the global error

$$E = \frac{1}{2m}\sum_{k=1}^{m}\sum_{o=1}^{q}(d_o(k) - y_o(k))^2 \tag{6}$$

*Step*6 **:** Determine whether the grid error meets the requirements

3. Results

The data used in this paper are from a data set containing 239,177 temperature records worldwide and data found on the Berkeley Earth data page (http://berkeleyearth.org/data/).

3.1. Global temperature prediction model based on Gray Prediction Method

There requires the prediction of future global temperature levels by building two models, and from the known data analysis given in the attachment, we will build the gray prediction model [6] and the BP neural network prediction model.

(1) Model building process

*Step*1 : Raw data is set to $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$;

*Step*2 : Calculate the generated sequences. $x^{(1)}$

$x^{(0)}$ That is, to add up the generated sequence of the raw data to get a new generated sequence:

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)) \tag{7}$$

among $x^{(1)}(k) = \sum_{i=1}^{j} x^{(0)}(i), k = 1, 2, \cdots, n,$

*Step*3 $x^{(1)}$: Define the gray derivative as

$$d(k) = x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1) \tag{8}$$

$z^{(1)}$ $x^{(1)}$ To generate a sequence:

$$z^{(1)}(k) = \alpha x^{(1)}(k) + (1-\alpha)x^{(1)}(k-1), k = 2, 3, \cdots, n \tag{9}$$

To sum up, the gray prediction model is obtained:

$$x'^{(1)}(k+1) = (x^{(0)}(1) - \frac{b}{a})e^{-ak} + \frac{b}{a}, k = 1, 2, \cdots, n-1 \tag{10}$$

However, in using the gray prediction model, in order to ensure the feasibility of the GM (1,1) modeling method, it is necessary to test the data given in the known attachment. The inspection steps are as follows:

$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$ First, set the raw data as, and then calculate the level ratio of the series, where the following relationship is satisfied: $\lambda(k)$ $\lambda(k)$

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, k = 2, 3, \cdots, n \tag{11}$$

$X = (e^{\frac{-2}{n+1}}, e^{\frac{2}{n+1}})$ Now there is a tolerable interval limit for the level ratio. If the model can make gray prediction, the level ratio should fall within the interval range, otherwise the transformation will be processed. For this question, the data obtained by grading testing of the data are shown in Appendix II:

The sequence values and the rank ratios are shown through the Table1. All the level ratios of the translated sequence are located within the interval (0.98,1.02) [7], indicating that the translated sequence is suitable for constructing the gray prediction model.

Table 1. Values and rank ratio

developmental quotient a	The amount of action in gray b	The posterior difference is compared to the C value
0	40.166	0.356

The Table1 shows the development coefficient, the gray action amount, and the posterior difference ratio. The posterior difference ratio can verify the accuracy of the gray prediction, and the smaller the posterior difference ratio is, the higher the gray prediction accuracy is. Generally, if the posterior difference ratio C value is less than 0.35, the model accuracy is high, and the C value is less than 0.5 indicates that the model accuracy is qualified. As can be obtained from the above analysis, the posterior difference ratio value is 0.356, indicating that the model accuracy is qualified, and the conditions of the gray prediction model are met.

Fthe data in the attachment and compare the predicted and actual values to obtain the fitted data map, see the Figure 3. The smaller the relative error value, the better the effect. In general, less than 20% indicates a good fit. Through table analysis, the average relative error of the model was 0.659%, indicating that the model fits well.

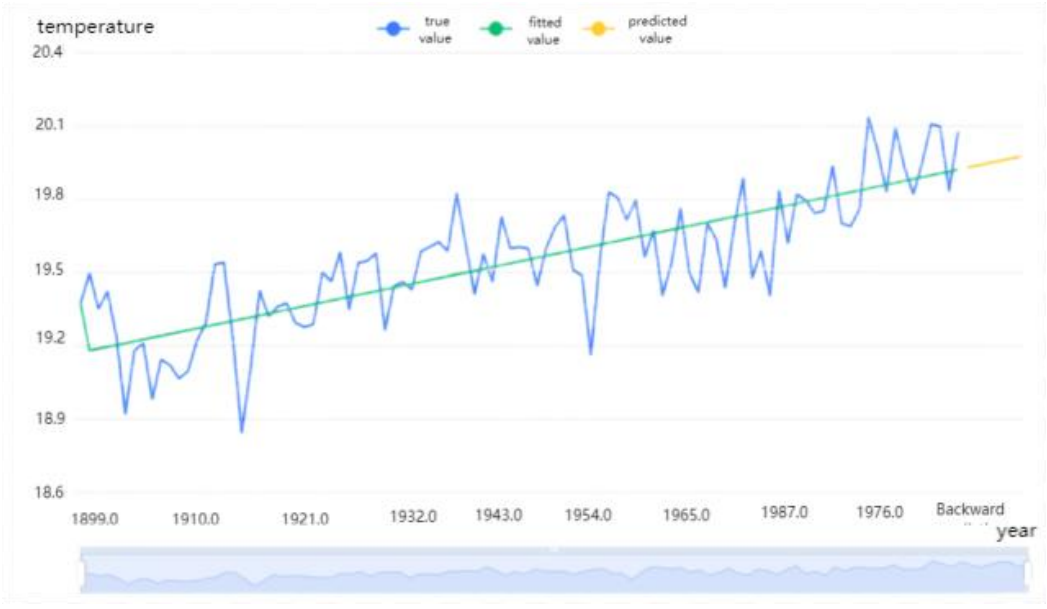


Figure 3. Fit prediction plots of the grey prediction model

(2) Gray prediction test

After the final model is obtained, the model is required to conduct the test step, test the predicted value, judge the reliability of the model, and conduct the residual test.

Residual difference test: calculate the relative residual difference

$$\varepsilon(k) = \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)}, k = 1, 2, \cdots, n \tag{12}$$

$|\varepsilon(k)| < 0.1$Provisions if all, considered to meet higher requirements, if all, considered to meet general requirements.$|\varepsilon(k)| < 0.2$

The residual statistics are calculated as shown in the Table2:

Table 2. Residual statistics[a]

	minima	maxima	mean	standard deviations	N
predicted value	19.176661	19.913324	19.544993	.2159088	99
residual	-.4729993	.3308230	0E-7	.1626745	99
normal expected value	-1.706	1.706	.000	1.000	99
Standard residuals	-2.893	2.023	.000	.995	99

The standard deviation of the found residue is 0.1627, which meets the compliance range and is considered to meet the general requirements.

(3) Model improvement

For the improvement of the GM (1,1) grey prediction model, by reviewing the literature[8]Get the essence of gray action, in the same sequence, if choose the same length sequence of the model, over time, under the premise of constant development coefficient, the gray effect will change, and this topic, in every 10 years for a group, to conduct the establishment of the model, just conform to the characteristics of the time-varying nature, so a way to improve the model, to reduce the error.

Now assume Model 1:

$$x_1^{(0)}(k) = Ae^{\lambda(k-1)}, k = 1, 2, \cdots, n \tag{13}$$

Model 2:

$$x_2^{(0)}(k) = Ae^{\lambda(k-1)}, k = n+1, n+2, \cdots, 2n \tag{14}$$

a b c d The development coefficient known as model 1 is gray action quantity; the development coefficient of model 2 is gray action quantity, and the following relation is obtained:

$$c = a, d = e^{\lambda t}b \tag{15}$$

Therefore, this problem directly uses the structure of model 1 to predict the data of model 2, which will cause relatively large errors, and

$$d = e^{\lambda t}b \approx b + \lambda bt \tag{16}$$

t Therefore, the analysis shows that the gray action amount is a function of time, and this problem regards it as an invariant to predict, which will inevitably cause a lot of error, so it is improved instead of building a new model.$b_1 + b_2 * n$ b

The modeling process is as follows:

$$x^{(0)}(k) + az^{(1)}(k) = b_1 + b_2 * n \tag{17}$$

This as an optimized gray differential equation.

Repeat the previous steps of solving the differential equation again, and first accumulate the raw data again:

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)) \tag{18}$$

$x^{(1)}(k)=\sum_{i=1}^{j}x^{(0)}(i),k=1,2,\cdots,n,$ The white differential equation corresponding to the model is then obtained again

$$\frac{dx^{(1)}}{dt}+ax^{(1)}=b_1+b_2t \tag{19}$$

a、b_1,b_2 When solving the value, the single linear regression estimation parameter matrix is again used to establish the differential equation model by using the discrete data series, which has obtained the GM (1,1) gray differential equation, and finally restored to the original data as follows:

$$x^{(0)}(k+1)=x^{(1)}(k+1)-x^{(1)}(k)=(1-e^{a})(x^{(0)}(1)-\frac{b_1}{a}-\frac{b_2}{a}+\frac{b_2}{a^2})e^{-ak}+\frac{b_2}{a},k=1,2,\cdots,n \tag{20}$$

After theoretical analysis, it is known that the improved model has better model accuracy.

3.2. Global temperature prediction model based on BP Neural Network

Through the output of the network, the learning error can be calculated. When the preset accuracy is reached, the forward transmission is over, otherwise, the next sample and the expected value are selected to repeat the above step. During backpropagation, the error signal is transmitted from back to forward, modifying the connection weights layer by layer. This process is called an iteration. The BP algorithm requires multiple iterations to make the error converge to minimal accuracy. Therefore, the learning time, number of iterations and final error accuracy of the network become important factors to measure the performance of BP neural networks.

According to the global average temperature given in the original setting data, according to the BP neural network model, the number of iterations is set to 1000 times, and then the number of hidden layers in the first layer is set to 100, and a group of grids are selected for the above step.

With MATLAB software programming, the neural network prediction value of global temperature Figure4 and the forecast diagram Figure5 are as follows:

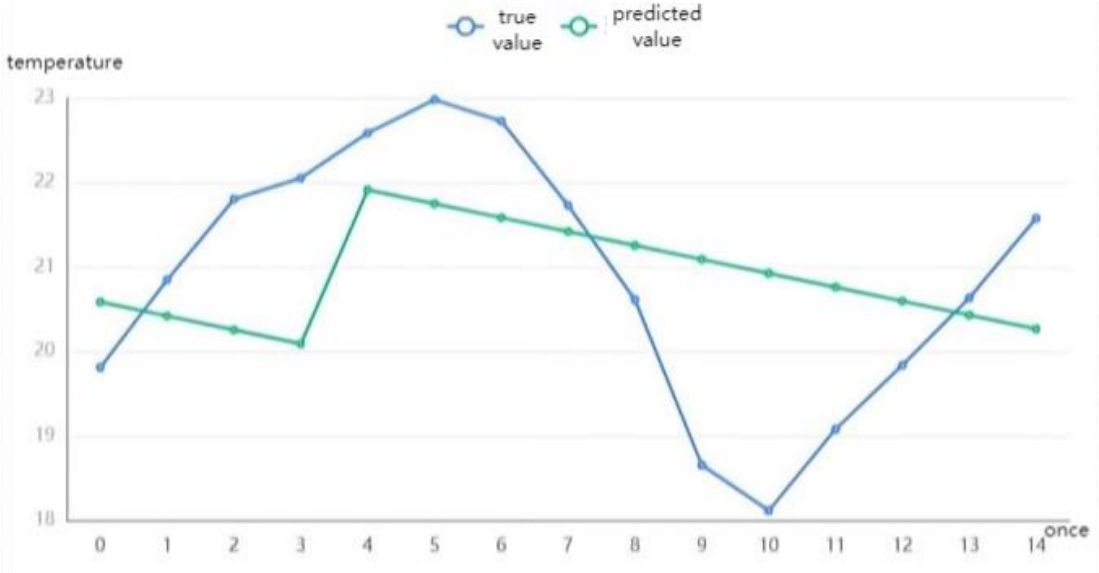


Figure 4. Neural network prediction plots in Fig

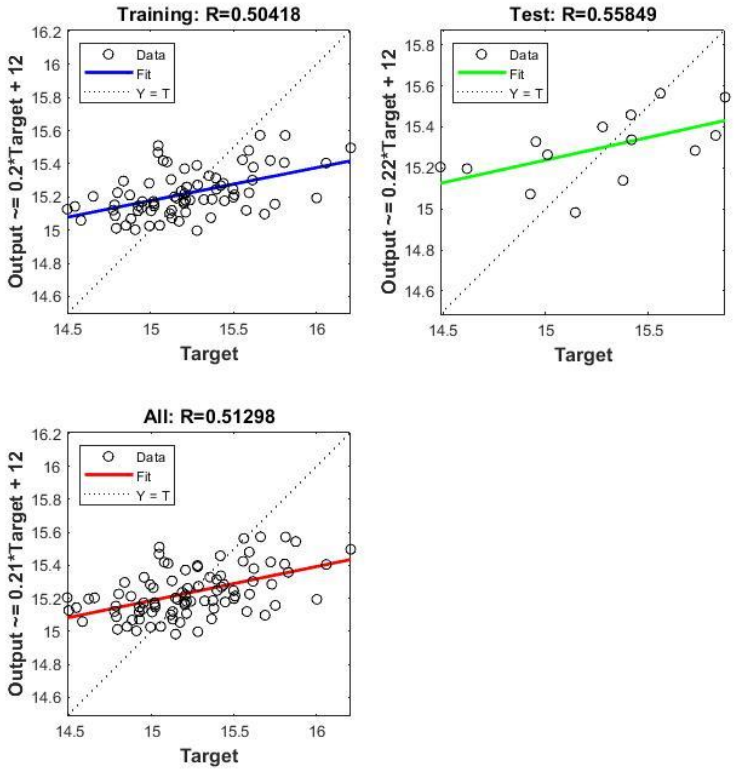


Figure 5. Test data fitting to Fig

For example, the above R fitting figure is predicted after the end of the training cycle, combining the prediction results with the training results to explain the fitting effect more intuitively. Overall, the model accuracy is high.

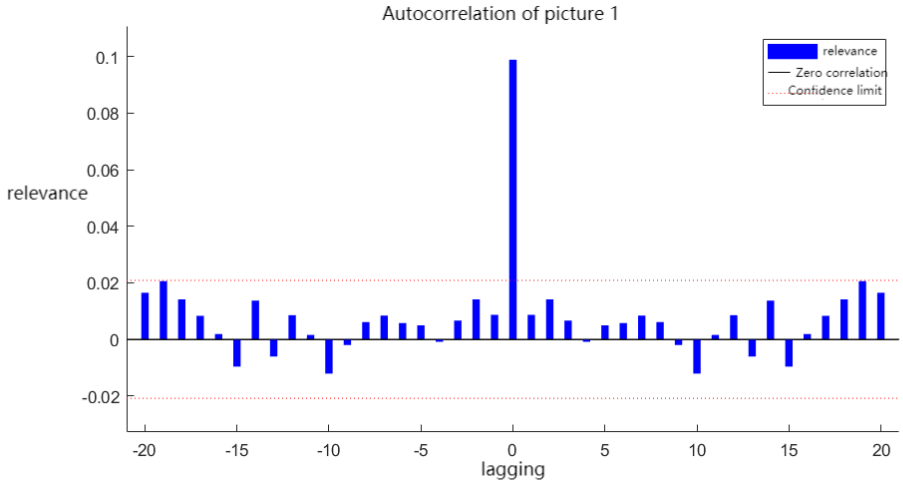


Figure 6. Neural network error autocorrelation plot

This figure shows that within the error range, the neural network prediction results fluctuate up and down in the zero correlation, except for the zero point is within the confidence limit, so the prediction results of the model have good error correlation and the prediction confidence is high.

The test results are shown as Table3:

Table 3. Model evaluation results

	MSE	RMSE	MAE	MAPE	R^2
training set	0.028	0.168	0.129	0.664	0.411
test set	0.043	0.206	0.168	0.843	-0.008

According to the analysis in the above table, MSE is the mean square error: the expected square of the difference between the predicted value and the actual value [9]. The lower the value, the higher the model accuracy, with 0.028 in the table, indicating the higher accuracy

The RMSE is the root mean square error: the square root of the MSE, the smaller the value, the higher the model accuracy, and the value is 0.168

MAE is the average absolute error: the average of the absolute error, which can reflect the actual situation of the predicted value error. The smaller the value, the higher the model accuracy, with a value of 0.129.

The MAPE is the average absolute percentage error: it is a percentage value. The smaller the values, the higher the model accuracy.

R^2 compares the predicted value, closer to only the mean, and the accuracy of the model is higher, and the median value is 0.411, which is a large gap from 1, so the accuracy of the model in the fit is not very high.

To sum up, the first three evaluation results are relatively small, indicating that the fitting accuracy of the model is high, while the last two relative values are high and the fit of the model is relatively low. The analysis reason may be that some discretization points differ too much from the actual temperature value, but overall, the model fit is good.

3.3. Analysis of influence factors of global temperature based on gray correlation theory

For subsystems or different factors in a system, the degree of correlation will be quantified by the correlation degree. The reference sequence is now set as, the temperature in this question is set as the main sequence, and the comparison sequence is established: $Y_0 = \{ y_0(i), i = 1, 2, \cdots, n \}$

$$X_{nm} = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} \tag{21}$$

n m Among them: is the number of years, is the number of elements, the population number, forest coverage area and carbon dioxide concentration are taken as three subsequences, as the comparison sequences.

Due to the different dimensions in the data composition, the subsequent solution will affect, so the data is normalized first. Here, the mean value is used to remove the dimensions. There are the following formulas:

$$\hat{y}_k = \frac{y_k}{\overline{y_i}}, \hat{x}_{kj} = \frac{x_{kj}}{\overline{x_j}} \tag{22}$$

among, $\overline{y_i} = \frac{1}{n}\sum_{k=1}^{n} y_k, \overline{x_j} = \frac{1}{n}\sum_{k=1}^{n} x_{kj}, (j = 1, 2, \cdots, m)$

The gray correlation coefficient is next calculated as $\sigma_j(k)$

The index correlation size for all factors is calculated, and the results are shown in the following Table4:

Table 4. Table of the gray association coefficient

Correlation coefficient results			
year	Carbon dioxide concentration	Global forest area	population size
1959.0	0.674	0.722	0.336
1960.0	0.690	0.736	0.342
1961.0	0.685	0.752	0.344
1962.0	0.717	0.725	0.356
1963.0	0.708	0.741	0.360
1964.0	0.745	0.718	0.376
1965.0	0.731	0.734	0.380
1966.0	0.713	0.764	0.382
1967.0	0.752	0.735	0.401
1968.0	0.770	0.727	0.414
1969.0	0.743	0.765	0.414
1970.0	0.760	0.756	0.428
1971.0	0.793	0.732	0.451
1972.0	0.766	0.766	0.452
1973.0	0.756	0.794	0.457

According to the Table4, gray correlation analysis for three evaluation items (CO 2 concentration, global forest area, population number) and 63 data, and take the global average temperature as the parent sequence, study three evaluation items (CO 2 concentration, global forest area, population number and global average temperature, and provide analytical reference based on correlation, using gray correlation analysis, resolution coefficient is 0.5, correlation coefficient is calculated by correlation coefficient calculation formula, and correlation coefficient value, and then the correlation value is calculated for evaluation judgment.The corresponding correlation coefficient plots is shown in Figure 7.

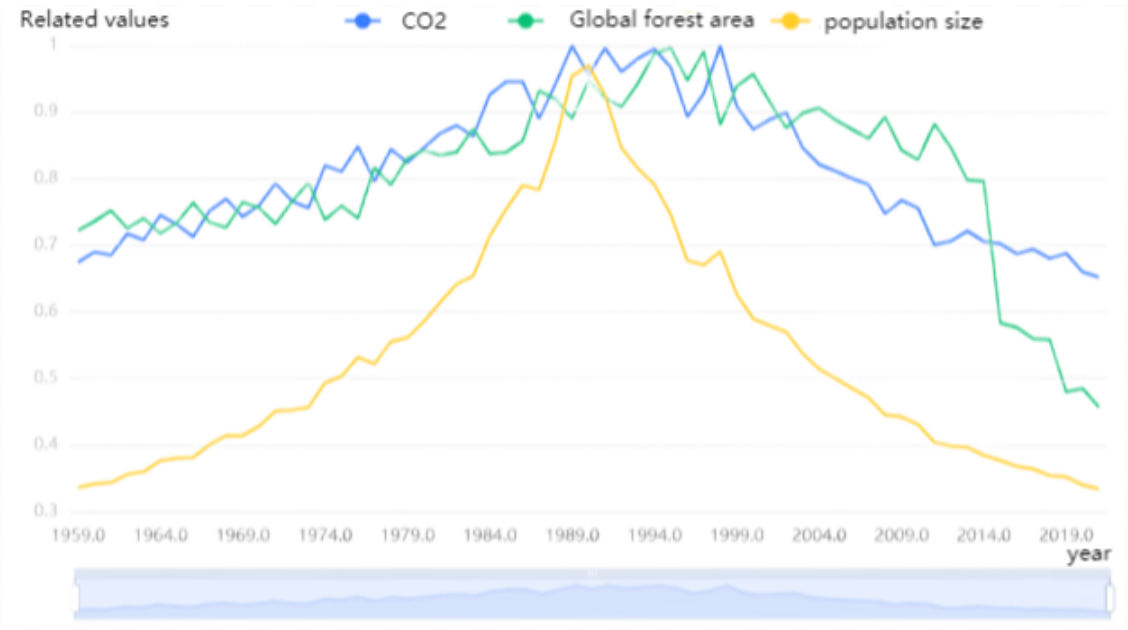


Figure 7. Correlation coefficient diagram

The correlation coefficient represents the degree of association value between the carbon dioxide concentration of the subsequence,the global forest area,and the corresponding dimension of the population number pair and the parent sequence,and the larger the number,the stronger the association is

Table 5. Gray correlation degree result Table

Correlation results		
Evaluation item	Correlation degree	ranking
Carbon dioxide concentration	0.813	1
Global forest area	0.805	2
population size	0.537	3

Based with the Table5, the correlation value is used to rank the three evaluation objects. The correlation value is between 0 and 1, the greater the value, the stronger the correlation between it and the parent sequence, the higher the evaluation [10]. As can be seen from the above table, for the three evaluation items, the highest carbon dioxide concentration was highest (correlation: 0.813), followed by the global forest area (correlation: 0.805).

4. Conclusions

The gray prediction method is used for modeling, through the level test, found that the gray prediction model conditions, and finally the residual test, the calculated value of 0.1627, all less than 0.2, is considered to meet the general requirements. The model is also improved, and the error is reduced.

The BP neural network model was also established to take the annual average temperature as the input layer, and then obtain the weight between the layers. After the cycle training, the future temperature can be predicted when the error meets the accuracy, and the root mean square error test was tested. The test value is about 0.03, which was found to be less than 5%, so this model passed the test index.

We use there prediction model to generalize to the future, and to obtain the correlation size of three factors and global temperature warming.

It is found that the decrease of carbon dioxide concentration suppresses the increase of global air temperature, and the forest coverage area has a direct relationship to the change of carbon dioxide concentration, and the change of global average air temperature is indirectly related, so the following measures can reduce the carbon dioxide concentration:

1. Adopt clean energy sources

The first approach is to use clean fossil fuels with low carbon content. However, using natural gas and shale gas to replace coal also has many problems. Firstly, using fossil energy and depletion; while the proportion of carbon content has decreased, but there is still half of the emissions. In addition, non-fossil energy, such as nuclear, hydropower, wind and solar, can be developed. But in terms of total power supply, the proportion of renewable energy is still very small. Globally, new renewable energy sources, namely wind and solar energy, still account for less than 5% of the world's major energy supply. We still need to analyze the relevant energy sources later. The use of low-carbon energy and renewable energy is obviously the way out, but the energy structure involves too many factors, which still needs to be continuously explored.

2. Afforestation in the desert

In the limited soil area, it is difficult to plant trees in large areas, so the desert and soil desertification has become the first choice, but it is a very long process to turn the desert into an oasis. It still takes a lot of hard work to grow. Although planting trees in the desert may seem like a good decision to solve the greenhouse effect, the data shows that the slower the surviving trees grow after 50 years, the lower their ability to absorb carbon dioxide. Especially for those dead trees, it will only increase the burden on the earth. In particular, different trees have different density at their ages, and they also vary in their ability to absorb and release carbon dioxide. So how to afforestation in the desert is still a problem worth further study.

References

[1] Wei Fengying, Cao Hongxing. Air temperature mutation analysis and its trend prediction study in China, Northern Hemisphere and Global [J]. Atmospheric Science, 1995 (02): 140-148.

[2] Hou Huiqing. Global climate change prediction model based on the BP neural network [J]. Technology and Innovation, 2021(09): 10-11.DOI: 10.15913/j.cnki.kjycx. 2021.09.005.

[3] Song Tianyang. Research on global average annual air temperature prediction based on genetic neural network [D]. North China Electric Power University, 2020.

[4] A gray prediction model based on the BP neural network [D]. Yan Jianbo. Xi'an University of Technology, 2009

[5] Xu Hong. Prediction of Employment Based on BP Neural Network [J]. Hunan Technology and Business University, 2020, (15).

[6] He Jie, Hu Da, Huang Xiaolin. Deformation prediction of tunnel surrounding rock based on an improved grey theory model [J]. Hunan University of Technology, 2017,42 (05)

[7] Gao Zhichao. Implementation carbon neutralization prediction analysis based on carbon emission data [J]. Electronic Technology and Software Engineering, 2022 (22): 201-204.

[8] Li Junfeng, Dai Wenzhan. Research on GM (1,1) improved model and the application of power generation modeling in Shanghai [J]. System Engineering Theory and Practice, 2005, (3): 140-144.

[9] Wang Chao. Short-term air temperature prediction study based on a deep neural network [D]. Anhui University, 2021.DOI: 10.26917/d.cnki. ganhu. 2021.002046.

[10] Zhang Yu, Zhao Wenying, Liao Fei, Cao Shuo, Dong Ying. Comprehensive evaluation of medical development level in provinces in the Yangtze River Basin based on principal component-grey correlation analysis [J]. Scientific and technological Innovation and application, 2022,12(30): 29-32.DOI:10.19981/j.CN23-1581/G3.2022.30.007.