Global Temperature Prediction Study Based On AMIMA, Grey Prediction and Neural Network Models

Junyi Lv [1,*], Yuan Xie [1], Chenyu Wang [2]

[1] School of Public Health, Xuzhou Medical University, Xuzhou, Jiangsu, 221000

[2] School of Medical Information and Engineering, XuZhou Medical University, Xuzhou, Jiangsu, 221000

* Corresponding Author Email: 202005010319@stu.xzhmu.edu.cn

Abstract. Global warming has profoundly affected the survival and development of human beings and has become a worldwide problem, and the mitigation of global warming requires a more accurate prediction of global temperature development. In order to accurately predict global temperatures in future years and provide feasible solutions, this paper comprehensively compares the prediction results of three prediction models. Firstly, a time series model is chosen to map the past time series to describe the global temperature changes. In addition, this paper selects three models for analysis by comparing a time series ARIMA model, a grey prediction model and a neural network model, fits the model using data information on the global average temperature for 200 consecutive years, and trains and corrects the model to obtain the projected global temperature for the next 80 years. Considering the advantages and disadvantages of each model and combining the conclusions of previous studies, this paper concludes that the neural network model predicts the most accurate results. Meanwhile, the results indicate that in the future, if the continuous increase in global temperature is not controlled, global warming will continue to cause irreversible damage to the Earth's ecosystem, which in turn will have a negative impact on human health.

Keywords: ARIMA model, gray prediction model, neural network model, global temperature.

1. Introduction

Global warming has profoundly affected the survival and development of **mankind [**1-4], and global warming has become a focus of global concern.It has been noted that global land temperatures have shown a highly significant warming trend from 1981-2019,with an average annual increase of 0.835°C [5].

In order to accurately predict the global temperature change, Zhu found that the global temperature will continue to rise over time and the global warming problem is imminent by constructing a BP neural network model based on principal component analysis and grey prediction[6].Luo et al. predicted the global average surface temperature by constructing an LSTM model and predicted that the year of 2020 will have a higher probability of becoming one of the hottest years in history[7]. Wang used a combination of ARIMA multivariate autoregressive time series prediction, LSTM long and short-term memory network and other methods to build a mathematical model to predict the global temperature change in the next 20 years, and concluded that with the future CO2 emissions increasing year by year, the global land and ocean temperatures are on an upward trend and will reach a maximum value of 15.64 degrees Celsius in 2040 [8].Cheng et al. et al. used the dynamic data information of global average temperature to predict the average temperature change in the short-term future, and proved that the ARIMA model is suitable for predicting the global average temperature [9].

In this paper, the global temperature is predicted by establishing three prediction models, and the model is fitted by using the data information of the global average temperature for 200 consecutive years, and the model is trained and corrected by using the data, so as to obtain the global temperature prediction for the next 80 years.

2. Establishment of Global Temperature Prediction Models

In order to predict the future global temperature, we choose three models for analysis by comparison, which are time series ARIMA model, grey prediction model and neural network model. By combining the advantages and disadvantages of the three models, we analysed them and came up with the model with the best simulation results.

2.1. Establishment and solution of ARIMA model

2.1.1 Establishment of ARIMA model

Considering that the data are too complicated and numerous to be exact for each month, we simplified the calculation by averaging the data for 12 months of a year to obtain the annual average temperature using Matlab software. We took year as the time variable and drew a time series plot (Figure 1).

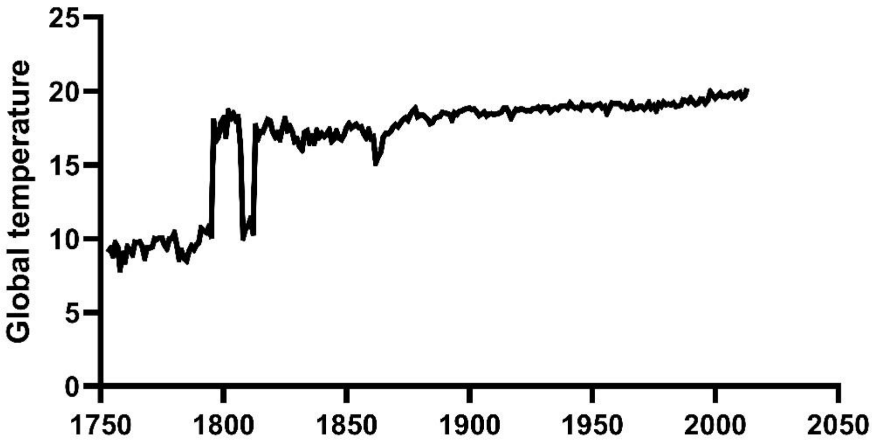

Figure 1: Global average annual temperature from 1753-2003

The time series plot can be used to describe the past global temperature levels, and the figure shows the data from 1753 to 2013. The line graph shows that the global temperature shows an overall zigzag increase, showing a large increase around 1803, and then gradually holds a high level of zigzag and smooth increase. The results of the model are shown in the Table1 below。

To determine which specific model to choose in the time series, we choose SPSS time series simulator to pick the optimal time series model, and finally we choose ARIMA (p,d,q) model.

Model principle:The time series may make a unit root process of order d, so we need to differentially process the data first and transform it into a smooth time series before modeling.

$$y_t^{'} = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^{'} + \varepsilon_t + \sum_{i=1}^{q} \beta_i \varepsilon_{t-i} \sum_{i=1}^{p} \alpha_i L_i \tag{1}$$

$$y_t^{'} = \triangle^d y_t = \left(1-L\right)^d y_t \tag{2}$$

$$\left(1-\sum_{i=1}^{q} \alpha_i L^i\right)\left(1-L\right)^d y_t = \alpha_0 + \left(1+\sum_{i=1}^{q} \beta_i L^i\right)\varepsilon_t \tag{3}$$

(Autoregressive Integrated Moving Average Model)

Table 1: Model calculation results

ARIMA Model Parameters				
	Estimate	Standard Error	t	Significance
Differences	1	None	None	None
Delay 1	0.133	0.057	2.323	0.021
Delay 5	0.287	0.058	4.956	0.000
Delay 12	0.253	0.058	4.326	0.000
Delay 17	-0.306	0.060	-5.106	0.000

Since the result is (p,1, q), we first differenced the data to the first order before analyzing it with ACF and PACF graphs.

2.1.2 Model checking

In order to test the fit of the model, we performed the Residual white noise test.The R-squared is 0.948, the closer to 1 the better the fit is, indicating that the fit is good and the results are more accurate. The results of the model fit are shown in the Table2 below.

Table 2: Residual white noise test

Model Statistics						
Models	Number of predicted variables	Stable R-squared	Statistics DF Significance			Number of outliers
Temperature-Model_1	0	0.182	11.624	14	**0.636**	0

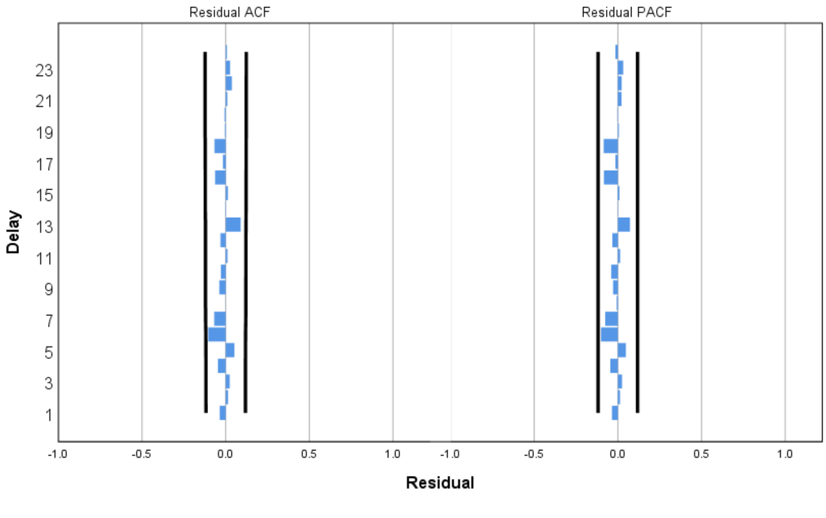

Figure 2: Residual ACF and Residual PACF

As can be seen from the ACF and PACF of the residuals in Figure 2 above, the autocorrelation coefficients and partial autocorrelation coefficients of all lag orders are not significantly different from 0. In addition, from the above table, it can be seen that the p-value obtained from the Q-test on the residuals is 0.636, i.e., we cannot reject the original hypothesis that the residuals are white noise series, so the Arima (0,1,17) model can identify the data well and prediction of global temperature for the next 80 years.

2.1.3 Prediction results and effects

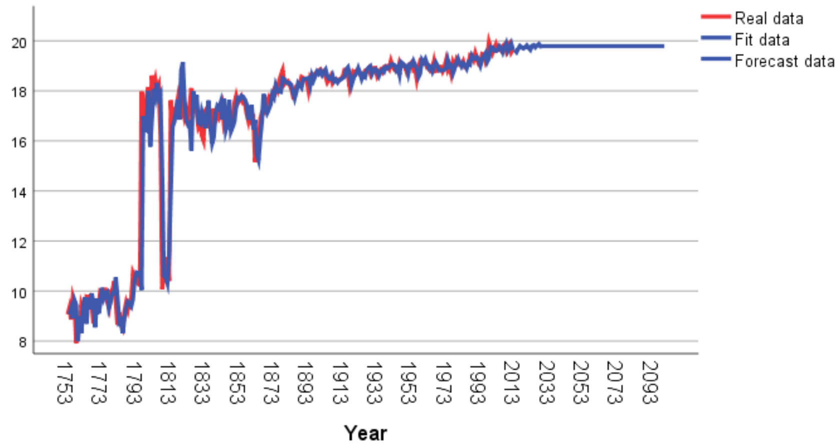

Figure 3: Prediction results by ARIMA

As can be seen from the Figure3, the time series plots of the real data and the fitted data almost overlap, which indicates that the ARIMA model fits the original data well; since 2013, due to the model limitations and the lack of data, the data predictions are in a smooth pattern, but never reach 20 degrees Celsius.

2.2. Gray predictive model

2.2.1 Establishment of Gray predictive model

Gray prediction model is to generate the original data to find the pattern of system changes, and generate a data series with strong regularity, and then build the corresponding differential equation model to predict the future development trend of things. In order to describe the past and predict the future global temperature level we use the gray prediction model, but due to the limitation of the gray prediction model itself, we only select the period of 2005-2013 for the past data.

Let the original time series of global temperature levels be:

$$x^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\right) \tag{4}$$

The original time series is summed once to get the new generated data series

$$x^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)\right) \tag{5}$$

The average of two adjacent values of the series after one accumulation, so that$z^{(1)}$for $x^{(1)}$ immediately adjacent to the average value of the generated series

$$z^{1} = \left(z^{(1)}(2), z^{(1)}(3), \ldots, z^{(1)}(n)\right) \tag{6}$$

The basic form of the GM (1, 1) model can be expressed as::

$$x^{(0)}(k) + az^{(1)}(k) = b \tag{7}$$

where b denotes the grey action of the equation and a denotes the data development coefficient Immediately after constructing the data matrix B and the data vector Y and parameters.

Thus, the GM (1, 1) model(4) can be expressed as

$$Y = Bu \tag{8}$$

We use the least squares method to obtain estimates of the parameters a,b as

$$u = \left(\frac{\alpha}{b}\right) = \left(B^T B\right)^{-1} B^T \tag{9}$$

The development coefficient obtained from the least squares fit is 0.0018503 and the amount of gray action is 19.6922, substituting the parameters into (1) equation gives:

$$x^{(0)}\left(k\right) = 0.0018z^{(1)}\left(k\right) + 19.7 \tag{10}$$

Using the above equation, future global temperatures can be predicted.

2.2.2 Results and conclusions

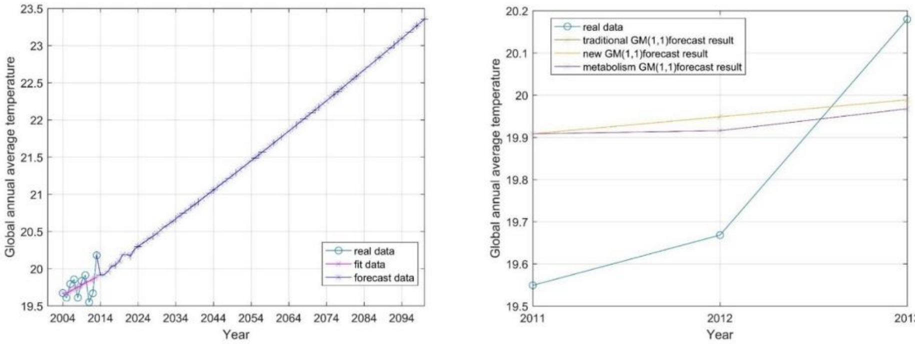

Figure 4: Experimental data predictions and Final prediction data effect

Matlab predictions were made using traditional GM (1,1), new information GM (1,1), and metabolic GM(1,1) respectively yielding error sums of squares, and we chose the one with the smallest error sum of squares for the subsequent predictions.

As shown in Figure 4 and Figure 5, the gray prediction model predicts a diagonal increase in global temperature in the future, but the predictions may be less accurate due to limitations in both the past data utilized by the gray prediction model and the years that can be predicted.

2.2.3 Model checking

Using the gray prediction model, residual tests and cascading deviation tests can be performed, and the corresponding residual and cascading deviation plots are made using, which are shown in the figure6 below.

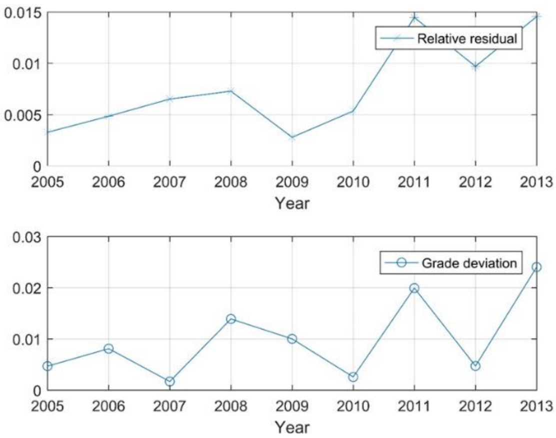

Figure 5: Relative residual and Grade deviation

The average relative residual was 0.0076426 and the average stepwise deviation was 0.0099513, both of which indicated that the model was a very good fit for the original data.

2.3. The structure of BP neural network

2.3.1 Establishment of BP neural network

BP neural network is a multi-layer network with error reverse propagation, which is composed of input layer nodes, hidden layer nodes and output layer nodes. This process has been reduced to an acceptable level of error to the network output, or to a predetermined number of learning times. The network structure is shown in Figure 6.

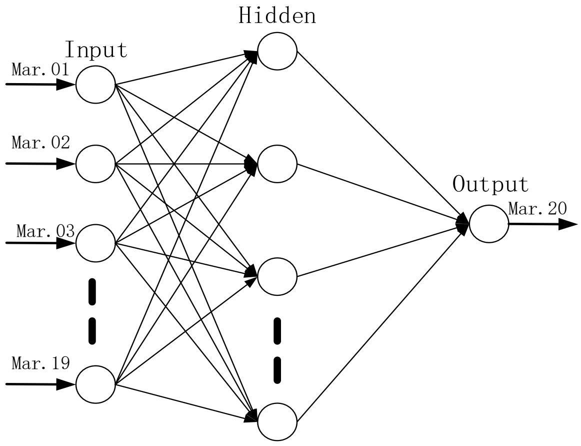

Figure 6. Neural network structure

The general model of artificial neural network consists of four basic elements, which are:

Let the input layer of BP network has n nodes, the hidden layer has q nodes, and the output layer has m nodes, and the weights between the input layer and the hidden layer are, and the weights between the hidden layer and the output layer are.The transfer function of the hidden layer is , and the transfer function of the output layer is , then the output of the nodes in the hidden layer is

$$z_k = f_1\left(\sum_{i=0}^{n} v_{ki} x_i\right) \qquad k = 1, 2, \ldots, q \tag{11}$$

The output of the output layer node is

$$y_j = f_2\left(\sum_{k=0}^{q} w_{jk} z_k\right) \qquad j = 1, 2, \ldots, m \tag{12}$$

The BP network completes the mapping of n-dimensional space vectors to m-dimensional space vectors.

According to the above principle, the computation process of artificial neural network can be summarized as the following steps:

(1) Initial value selection:$\omega(0)$

(2) Forward computation to find the output of all neurons:$a^{(k)}(t)$

(3) Calculation of the output layer:$\delta_j = (t_j - a_j) a_j (1 - a_j)$

(4) Calculate each hidden layer from back to front:$\delta_j = a_j (1 - a_j) \sum_i w_{ji} \delta_i$

(5) Calculate and save the correction amount of each weight:$\Delta w_{ij} = -\eta \delta_j a_i$

(6) Correction weights:$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}$

(7) If convergence is achieved, the process ends, if not, go to step (2)

2.3.2 Results and conclusions

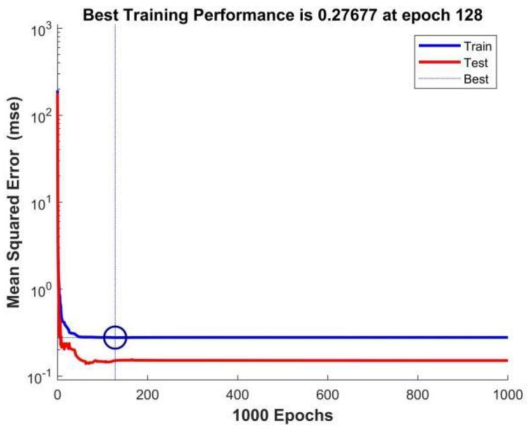

Figure 7: Performances

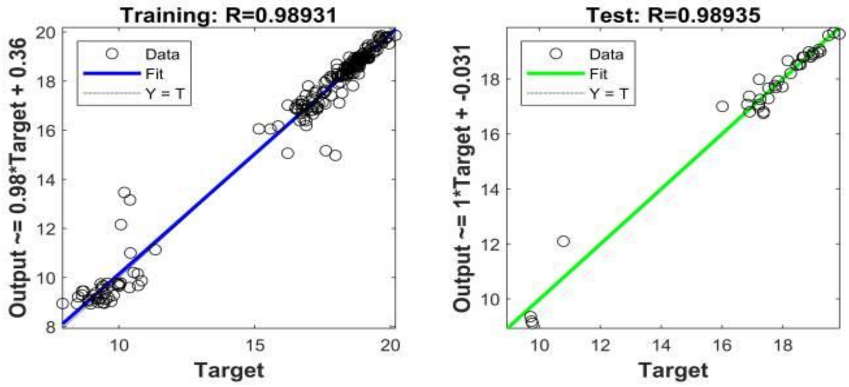

Figure 8: Regression

One epoch is equal to one training session using all the samples in the training set, and the parameters in the neural network are adjusted for each training session.

MSE: Mean Squared Error Mean Squared Error MSE = SSE/n

In general, the error in the validation data set may begin to increase. The Figure7 shows that the minimum MES is reached at 128 cycles.

As Figure 8 shows,the fitted values are regressed on tor decreases after more training phases, but as the network begins to overfit the training data, the true values, and the larger the square of R indicates a better fit, indicating that the neural network model fits better this time and can be used for subsequent prediction.

3. Discussion

Three models are used to predict the global temperature in 2050 and 2100 respectively.

The time series ARIMA model (Figure 9) was performed to predict the results showing that the temperature was steady since 2013, but never reached 20°C.

The gray projection model (Figure 10) shows that the global temperature has been increasing continuously since 1900 and has already reached 20 degrees Celsius in 2050 and 2100.

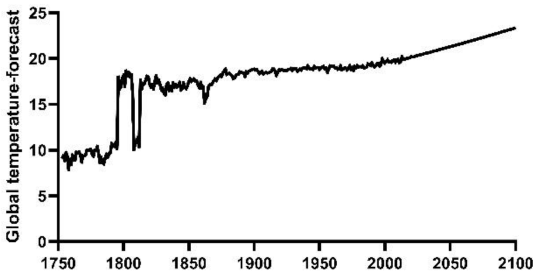

Figure 9: Global average annual temperature predicted by the ARIMA model.

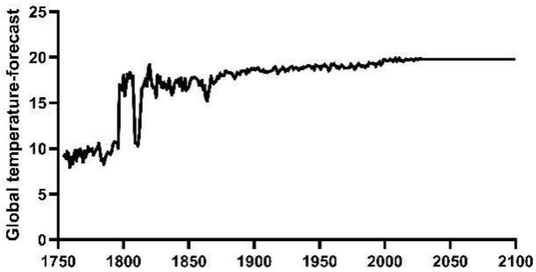

Figure 10: Global average annual temperature predicted by the grey forecast model.

The BP neural network model (Figure 11) predicts a small increase in temperature all the time, and finally slowly approaches flatness, with temperatures reaching 20 degrees Celsius in 2050 and 2100.This slow upward trend is consistent with Hou's conclusion from the BP neural network prediction model that future global average temperatures are on a slow upward trend, eventually levelling off [10].

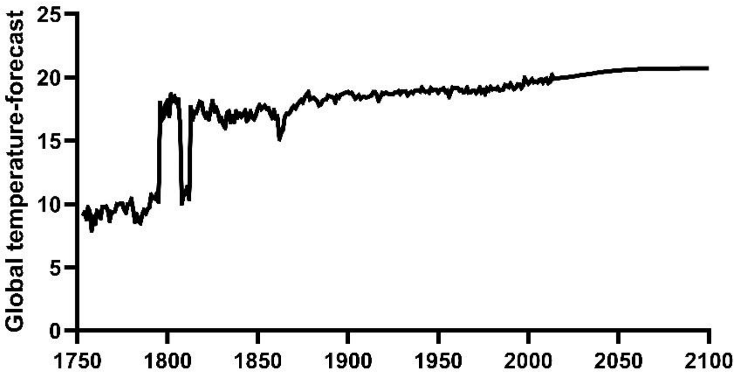

Figure 11: Global average temperature predicted by the Neural network model.

Among the three models, we believe that the neural network model is the most reliable, while the ARIMA model is less reliable due to the large number of unknown years until 2050 and 2100, and the data tend to be more consistent in the later years. The gray prediction model can only predict 16 years due to the limitation of the model itself on the prediction year, and the data in the later years lack reliability. The neural network model is a highly complex nonlinear dynamical learning system with high reliability.

4. Conclusion

In order to predict the future development of global temperature and provide feasible solutions, we use ARIMA model, grey prediction model, neural network model, the results show that in the future if we do not control the global temperature will be a continuous increase, global warming continues to cause irreversible damage to the earth's ecological environment, and thus will have a negative impact on human health.

To address the root causes of global warming, it is feasible to reduce heat emissions by converting heat energy into chemical energy, such as using photovoltaic power generation to electrolyse water, utilising wind and water energy to generate electricity, and replacing fuel cars with electric cars, all of which can address the issue of heat emissions from the root causes.

References

[1] Zheng Jingyun,Liu Yang,Wu Maowei et al. Characteristics of multiscale changes in temperature and regional differences during the medieval climate anomaly in China[J]. Journal of Geography,2019,74(07):1281-1291.

[2] Jin Kai,Wang Fei,Han Jianqiao et al. Impacts of climate change and human activities on NDVI changes of vegetation in China from 1982 to 2015[J]. Journal of Geography,2020,75(05):961-974.

[3] Climate Change: Researchers at National Institute of Amazonian Research Have Reported New Data on Climate Change (Extinction risks forced by climatic change and intraspecific variation in the thermal physiology of a tropical lizard) [J]. Global Warming Focus,2018.

[4] Neves T S,Serra M S,Manuela M. Climate change and economic growth: a heterogeneous panel data approach.[J]. Environmental science and pollution research international,2018,25(23).

[5] Shen Beibei,Song Shuaifeng,Zhang Lijuan et al.Characteristics of global temperature change from 1981-2019[J]. Journal of Geography,2021,76(11):2660-2672.

[6] Zhu Shengyao,Chen Jinjie. Research on global temperature prediction based on grey prediction and BP neural network[J]. Software Guide,2020,19(02):105-109.

[7] D.Y.Luo,F. Zheng,Q.L. Chen. Deep learning-based prediction of global mean surface temperature interannual signal time series[J]. Climate and Environment Research,2022,27(01):94-104.

[8] Wang Yuanhao. Analysis of global temperature prediction based on ARIMA model and LSTM neural network[J]. Science and Technology Innovation,2021(35):166-170.

[9] Cheng Yan,Hua Zhiqiang,Huang Yujie et al.Application of ARMA model in predicting global average temperature scenario[J]. Journal of Inner Mongolia University for Nationalities (Natural Science Edition),2023,38(02): 103-108.DOI: 10.14045/j.cnki.15-1220.2023.02.002.

[10] HOU Huiqing. Global climate change prediction model based on BP neural network[J]. Science, Technology and Innovation,2021, No.177(09): 10-11.DOI: 10.15913/j.cnki.kjycx.2021.09.005.