ISSN: 2688-8653

Research on Air Temperature Prediction Model Based on BP Neural Network

Zhengling Wu

College of Arts and Sciences, Northeast Agricultural University, Harbin, Heilongjiang, 150030, China

Abstract

In recent years, the frequent occurrence of extreme climate caused by global warming has attracted people ' s attention. It is of great significance to master the global climate and temperature situation for human sustainable development. In this context, this paper focuses on the problem of global warming, and establishes a temperature prediction model based on BP neural network with the abnormal value of ocean temperature and carbon dioxide emissions as the influencing factors and the atmospheric temperature as the output. Combined with the strong nonlinear mapping ability of traditional BP neural network, the ARIMA(0,1,1) model is established to compare the error. The results show that the BP neural network model shows stronger fitting ability than the ARIMA model for the nonlinear temperature data with strong volatility. The error MAE, MAPE and RMSE were reduced by 1.5 %, 4.67 % and 4.33 % respectively, and the determination coefficient R-squared was increased by 0.0618. At the same time, BP neural network does not need professional meteorological knowledge, the model is simple and reliable, which provides data support for global warming prediction and has reference significance for practical work.

Keywords

Climate Warming; Sustainable Development; BP Neural Network Prediction; Error Analysis.

1. Introduction

With the intensification of human activities, a large number of burning coal, natural gas and other carbon-containing fuels lead to excessive emissions of greenhouse gases, resulting in the destruction of greenhouse gas balance in the atmosphere, resulting in continuous accumulation of greenhouse gases, making the global temperature rising, resulting in global warming.

Global warming directly leads to glacier melting, sea level rise, flooding a lot of land, changing the current world climate pattern ; at the same time, global warming has increased natural disasters and affected the current agricultural layout [1]. Mastering the global climate and temperature trend can predict the future climate, which is of great significance to the world ecological protection and human sustainable development. Many scholars have established models to predict atmospheric temperature, such as Hargreaves-Samani (HS) formula based on physical and plant transpiration, or statistical prediction methods, such as principal component analysis, grey prediction, time series analysis, LSTM, which have been applied in this field. In the above method, HS formula needs professional meteorological knowledge, grey prediction depends on the initial value and background value, and time series analysis for nonlinear data fitting effect is not good.

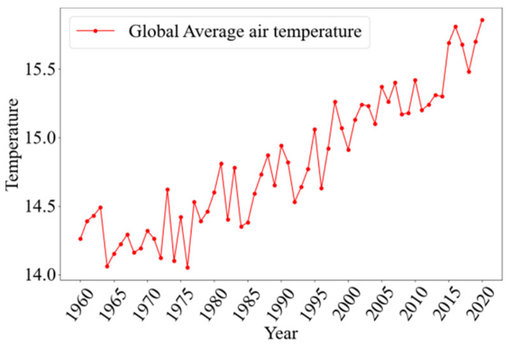
In view of the nonlinear characteristics of temperature fluctuation, this paper comprehensively considers the influence of ocean temperature anomaly and carbon dioxide emissions on atmospheric temperature. BP neural network is used to fit the model, and then the model

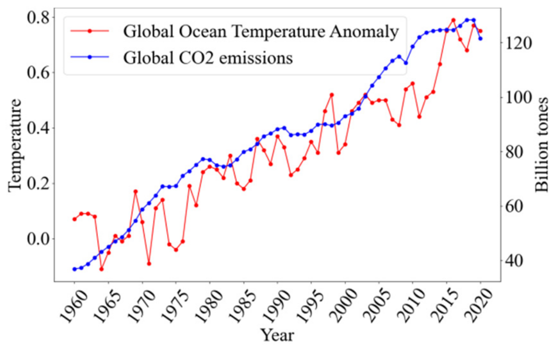
ISSN: 2688-8653

parameters and error estimation are tested. At the same time, ARIMA model is set to compare the goodness of fit.

2. Data Feature Description

In this paper, the global meteorological data are obtained from our World in data official website. The average global temperature, the abnormal value of global ocean temperature and the total global carbon dioxide emissions from 1960 to 2020 are screened out by Excel. Time Series Charts are drawn by Python:





(a) Global Average Air Temperature (b) Ocean Temperature and CO2 emissions

Figure 1. Time Series Charts

It can be seen from Figure 1 that the anomalous value of ocean temperature and carbon dioxide emissions have an obvious upward trend with the passage of time since 1960, which is consistent with the upward trend of atmospheric temperature fluctuation and is consistent with the research problem , indicating that atmospheric warming is serious [2]. Because the atmospheric temperature fluctuation is nonlinear, BP neural network is used to predict.

3. Material and Methods

3.1. Algorithmic Theory of BP Neural Networks

BP neural network is a typical multi-layer feedforward neural network composed of input layer, hidden layer and output layer. BP neural network calculation formula as (1) - (2) shows :

$$u_j = \sum_{i=1}^{n} w_{ij} x_i - o_j \tag{1}$$

$$p = f_s\left(\sum_{j=1}^{n} w_j o_j + \theta\right) \tag{2}$$

Equation (2) shows that the input data of neurons are subtracted from the threshold after weighted summation, x_j represents the input data, w_{ij} represents the weight of i-th input data to the j-th neuron, and o_j represents the threshold ; p denotes the local output, $f_s(\ )$ is the activation function [3], we use the Sigmoid function : $f_s(u) = \frac{1}{1+e^{-u}}$.

The conventional BP neural network includes data forward propagation and error back propagation. According to the error value, the connection weight w_{ij} and the threshold θ are continuously adjusted by gradient descent method, so that the error is continuously reduced to the expected value. The correction formula [4] of connection weight and threshold is shown in Equations (3)–(4):

$$w(k+1)=w(k)-\beta\frac{\partial E}{\partial w} \tag{3}$$

$$\theta(k+1)=\theta(k)-\beta\frac{\partial E}{\partial \theta} \tag{4}$$

In the formula, E represents the error between the output value of the neural network and the target value, and β represents the learning rate of the neural network.

3.2. Algorithm Steps

1) Initialize the connection weights and thresholds, and randomly select a dataset to provide to the network ;
2) Use connection weight and threshold to forward calculate the value of output layer and calculate error;
3) Adjusting Connection Weights and Thresholds of Connection Weights by Modified Formulas of Connection Weights and Thresholds;
4) Randomly select the next dataset and return to step 3 until all datasets are trained;
5) Reselect a data set randomly from all data sets and return to step 3 until the error of the network is set to a minimum or the number of learning reaches a predetermined value, the algorithm ends;

Output training results.

The algorithm flowchart is shown in Figure 2 :

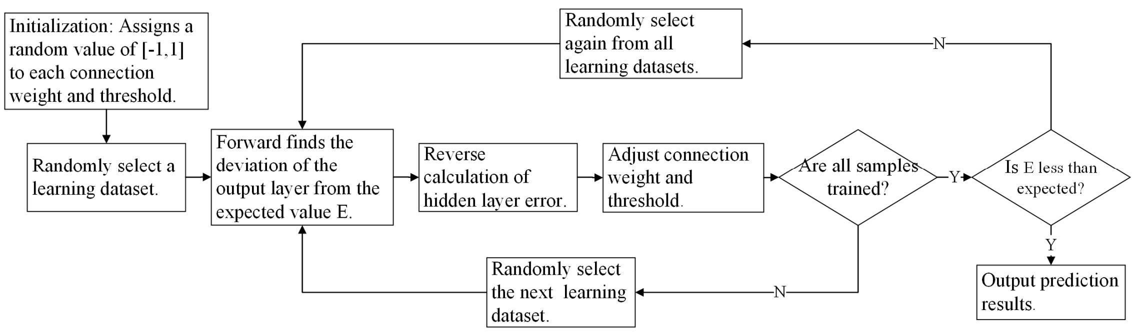


Figure 2. Algorithm flow chart

4. Experimental Simulation Analysis

4.1. Application of BP Neural Networks

In general, the application of BP neural network includes the following steps :

1)Determining the Structure of BP Neural Network

In this paper, the annual average temperature of the global ocean and the global annual carbon dioxide emissions are used as the input of the BP neural network, and the global annual average temperature is used as the output. Therefore, the number of nodes in the input layer is 2, and the number of nodes in the output layer is 1. The hidden layer can be multi-layer, this paper set the hidden layer is 1 layer. Increasing the number of hidden layer nodes can improve the model accuracy but reduce the computational efficiency, and may lead to overfitting problems.

Number of hidden layer nodes based on empirical formula [5]$h=\sqrt{m+n}+a$,where h is the number of hidden layer nodes, m and n are the number of input layer nodes and output layer nodes, respectively, and a is the adjustment constant between 1 and 10. Considering the model accuracy, overfitting problem and computational efficiency, the number of hidden layer nodes is set to 10.

2)Setting Parameters for Network Training

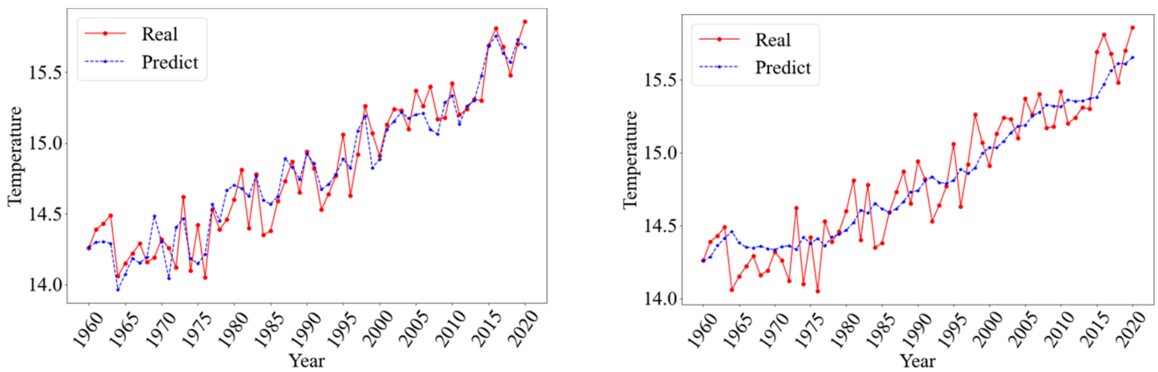
Before the training of BP neural network, it is necessary to set the maximum number of iterations of network training, the expected value of training target error and the activation function. In this paper, 1000 iterations are set, and the expected value of training target error is 1e-5. The activation function is selected as the Sigmoid function above.

4.2. Error Evaluation Method

The regression algorithm uses MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error) and RMSE (Root Mean Square Error) three kinds of error analysis methods and the determination coefficient R^2 to evaluate the accuracy of the prediction model, so as to verify the accuracy of the prediction model. The smaller the MAE, MAPE and RMSE is, the smaller the error is, the greater the goodness of fit is, and the closer R^2 is to 1, indicating that the regression model has better predictive effect on fitting.

4.3. Experimental Simulation Results

The processed data are used for training, and the trained data set is evaluated for error. At the same time, the ARIMA(0,1,1) model of global atmospheric temperature is established and compared with BP neural network. The fitting effect is shown in Figure 3. It can be seen that the goodness of fit of BP neural network for atmospheric temperature is stronger than that of ARIMA(0,1,1) model.



(c) Prediction results of BP neural networks (d) Prediction results of ARIMA(0,1,1)

Figure 3. Prediction results of each model

After the data are predicted by the model, the relative error of the model is calculated. The running results are shown in Table 1. By comparing the fitting effect of Figure 3 with the relative error of Table 1, it is found that the relative error of BP neural network is lower than that of ARIMA(0,1,1) model.

Table 2 is the prediction error of the model. By calculating the MAE, MAPE, RMSE and the determination coefficient R^2 of each prediction model, the MAE, MAPE, RMSE predicted by BP neural network are 0.0187, 0.1369 and 0.1094, respectively. Compared with ARIMA (0,1,1), they are reduced by 1.5 %, 4.67 % and 4.33 %, and the determination coefficient R^2 is 0.9227, which is increased by 0.0618.

Table 1. Relative error of each model

Year	Temperature	BP Model	Relative Error	ARIMA(0,1,1)	Relative Error
1960	14.26	14.263	0.0002	-	-
1961	14.39	14.300	0.0063	14.285	0.0073
1962	14.43	14.304	0.0088	14.364	0.0046
1963	14.49	14.292	0.0137	14.412	0.0054
1964	14.06	13.964	0.0068	14.460	0.0284
1965	14.15	14.074	0.0054	14.382	0.0164
1966	14.22	14.185	0.0025	14.352	0.0093
1967	14.29	14.154	0.0095	14.347	0.0040
1968	14.16	14.196	0.0025	14.359	0.0141
...	...	...	...	...	...
2019	15.7	15.730	0.0019	15.610	0.0057
2020	15.86	15.677	0.0115	15.653	0.0130

Table 2. Comparison of model prediction errors

Predict Model	MSE	RMSE	MAE	R-Squared
BP	0.0187	0.1369	0.1094	0.9227
ARIMA(0,1,1)	0.0337	0.1836	0.1527	0.8609

Based on the analysis of the training results, BP neural network has strong nonlinear mapping ability in analyzing the data with strong volatility of atmospheric temperature. Through the analysis of MAE, MAPE, RMSE and R^2 of BP neural network and ARIMA model, it can be concluded that the relative error of BP neural network is small, and the goodness of fit is higher, which can more accurately reflect the trend of atmospheric warming, and has stronger reference significance for practical work.

5. Conclusion

Aiming at the problem of global warming, this paper establishes an air temperature prediction model based on BP neural network with the abnormal value of ocean temperature and carbon dioxide emissions as the influencing factors and the atmospheric temperature as the output. Compared with the ARIMA (0,1,1) model, the model shows strong nonlinear mapping ability and can predict the global climate more accurately. Moreover, the BP neural network does not rely on professional meteorological knowledge, which has clear significance and strong reliability. It can explore and predict the current temperature data and has reference significance for practical work. At the same time, carbon dioxide emissions have a great impact on global climate, and carbon dioxide emissions are generally consistent with the trend of global warming. Therefore, reducing carbon dioxide emissions is also one of the important means to control global warming. Since the BP neural network may fall into local optimal solution, it may require multiple experiments to obtain the optimal solution ; at the same time, this paper only selects two influencing factors, namely the abnormal value of ocean temperature and carbon dioxide emissions. If multiple factors affecting global temperature are considered, it will have better results.

References

[1] Sun Jiejie, Jiao Wenxing, Wang Qian, Wang Tongli, Yang Hongqiang, Jin Jiaxin, Feng Huili, Guo Jiahuan, Feng Lei, Xu Xia, Wang Weifeng. Potential habitat and productivity loss of Populus deltoides industrial forest plantations due to global warming[J]. Forest Ecology and Management,2021,496.

[2] Wang Yuanhao. Global temperature forecast analysis based on ARIMA model and LSTM neural network [J]. Science and technology innovation, 2021(35):166-170.

[3] Liu Yanjun. Self-learning of neural network and its application [D]. University of Science and Technology of China, 2018.

[4] Wang Jianan, Wang Yuying, He Shulin, Shi Longmin, zhangyan Di, Sun Haiyang, liuyong. Soil moisture prediction model optimized by BP neural network based on improved genetic algorithm [J]. Application of Computer System, 2022,31 (02): 273-278. DOI: 10.15888/J.CNKI.CsA.

[5] Jiang Minghui, Wang Yalin, Zhao Xin speed skater, Huang Weiping. Application of k- Nearest Neighbor Discriminant Analysis in Personal Credit Evaluation [J]. Research on Quantitative Economy, Technology and Economy, 2004(02):143-147.