ISSN:2790-1688 DOI: 10.56028/aetr.4.1.494.2023

Carbon Dioxide Concentration and Sea-land Temperature Changes——Based on Mathematical Model Prediction

Zixin Zhou[1, a,*], Peiru Tian[1,b], Yaofu Zhang[2,c], Dongsheng Shi[2,c]

[1]Ningbo foreign language school, Ningbo, 315100, China

[2]Ningbo Xiaoshi High School, Ningbo, 315012, China

[3]Nanjing Foreign Language School, 210000, China

[4]Shenzhen International Foundation Collage, ShenZhen, 518000, China

[a] 195648937@qq.com, [b] 1130759942@qq.co m, [c] 2982339367@qq.com,[d] 46821471@qq.com

Abstract. With a large amount of carbon dioxide emissions from factories, there will be global warming, and the rising sea level and melting glaciers at the poles will pose a threat to the environment on which human beings depend. We build a function to fit the past development pattern and future development trend of CO2/land-ocean temperature based on the given data of CO2 concentration/land-ocean temperature from 1959 to 2021s. We will also study the correlation between temperature and carbon dioxide, Also, based on the error distribution, we discussed the reliability of the prediction method based on CO2 concentration and gave a method to judge the reliability of the prediction results.

Keywords: CO2 Concentration, Land-Ocean Temperature, Prediction Model, Correlation.

1. Introduction

With the development of industry, although people's living standards and conditions have been constantly improved, environmental problems have become more and more serious. The emissions from factories are increasing the concentration of carbon dioxide and causing global warming [1,2], which will pose a great threat to the environment on which human beings depend. Global warming will lead to sea level rise and redistribution of precipitation, as well as disrupt the biological and food chains. More seriously, it will increase the frequency and intensity of extreme weather (such as droughts, floods, thunderstorms, hail, storms, hot weather, and dust storms) in some regions [3-5]. In recent years, CO2 concentration has still been rising rapidly, which will lead to more severe global warming. Therefore, we play an important role in studying the future trend of CO2 concentration.

Our research work is based on historical data on CO2 concentration and land-ocean temperature, and we establish the trend of CO2 concentration/land-ocean temperature by the fitting method. the first task we need to deal with is to calculate the 10-year average growth rate of CO2. then we build multiple mathematical models to represent the change in CO2 levels over time. Linear and nonlinear regression methods were used to establish the corresponding mathematical model, and the regression equation between CO2 content and time t was established. After plugging in the data to calculate, we can calculate the predicted CO2 concentration in future.

Also, we build a prediction model for land-ocean temperature change and make an analysis of the relationship between carbon dioxide concentration and temperature change.

2. Assumptions and Definition of Variables

Our research work is based on historical data on CO2 concentration and landocean temperature, and we establish the trend of CO2 concentration/land-ocean temperature by the fitting method. Therefore, the data we use should be correct enough. For this, we make some reasonable assumptions：

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.494.2023

Assumption 1: In this paper, it is assumed that in the short term, no major changes in global industrial development, human activities, etc.

Justification 1: At present, the changes in CO2 concentration and land-ocean temperature are mainly due to human production activities. When we study the changes in CO2 concentration temperature, we do not consider extreme contingencies and huge changes in the policy environment that may lead to shifts in industrial development and changes in the development patterns of human production activities, such as volcanic eruptions, massive forest fires, and huge restructuring of industrial production on a global scale due to huge changes in environmental policies.

The variables in the model are explained in Table 1.

Table 1 Description.

Symbols	**Instructions**
t	Year
x(t)	t (Year) Atmospheric CO2 Concentration
y(t)	t (Year) Land-Ocean Temperature
A	Coefficient Matrix
ε	Random Error
L	Residual Sum of Squares
$x^{(0)}(t)$	(t+1958) (Year) Atmospheric CO2 Concentration
$x^{(1)}(t)$	Generates the Cumulative Sequence
R^2	Coefficient of Fit

3. Models and Predicted Results

3.1 Analysis of Average Increase of CO2 in Each 10 Years

We determined the pattern of the relationship between year and atmospheric CO2 concentration using a scatter plot based on the average value of CO2 (expressed as a mole fraction in dry air) for March each year at Mauna Loa Observatory in Hawaii, USA. The horizontal axis of the coordinates represents the year, using variable t, and the vertical axis represents the CO2 concentration, using variable x. The data from 1959 to 2021 is shown in Figure 1.

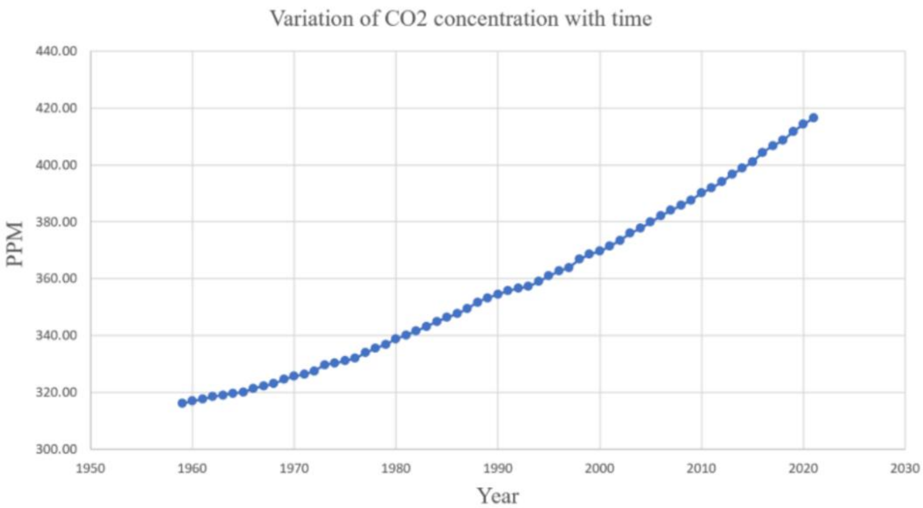

Figure 1 Data of CO2 Concentration from 1959 to 2021.

It can be seen from the figure that the CO2 concentration is correlated with the year. Linear regressions were performed every 10 years using data from 1959-1968, 1960-1969,…, and 1995-2004, respectively, to obtain x=r*t + a, with r representing the average growth rate of atmospheric CO2 concentration every 10-year period, as shown in Table 3.1. Comparing the data in

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.494.2023
the table, it can be concluded that the growth rate of atmospheric carbon dioxide concentration is the largest in the 10 years from 1995 to 2004. Therefore, we agree that the March 2004 increase in CO2 resulted in a larger increase than observed over any previous 10- year period.

Table 2 Average Growth Rate of CO2 Concentration Every 10 Years.

Years	Growth Rate （CO2）	Years	Growth Rate （CO2）	Years	Growth Rate （CO2）
1959-1968	0.75	1972-1981	1.36	1985-1994	1.39
1960-1969	0.81	1973-1982	1.39	1986-1995	1.37
1961-1970	0.88	1974-1983	1.49	1987-1996	1.36
1962-1971	0.92	1975-1984	1.55	1988-1997	1.33
1963-1972	0.99	1976-1985	1.58	1989-1998	1.45
1964-1973	1.09	1977-1986	1.55	1990-1999	1.58
1965-1974	1.14	1978-1987	1.54	1991-2000	1.67
1966-1975	1.13	1979-1988	1.58	1992-2001	1.74
1967-1976	1.13	1980-1989	1.61	1993-2002	1.80
1968-1977	1.15	1981-1990	1.63	1994-2003	1.83
1969-1978	1.17	1982-1991	1.61	1995-2004	1.85
1970-1979	1.23	1983-1992	1.55	-	-
1971-1980	1.31	1984-1983	1.45	-	-

3.2 CO2 Concentration Prediction

We use the year as the independent variable t and CO2 concentration as the dependent variable, and build regression models as shown in equations (1~4), respectively. The error term ε is a random variable with expectation 0.

Linear: $y(t)=a*t+b+\varepsilon$ (1)

Quadratic: $y(t)=a*t\ 2+b*t\ +c+\varepsilon$ (2)

Cubic: $y(t)=a*t\ 3+b*t\ 2+c*t+d+\varepsilon$ (3)

Logarithm: $y(t)=a*\ln(t)+b+\varepsilon$ (4)

We use the least squares method to solve each regression model parameter, as shown in Table 3.

Table 3 Regression Model Parameter.

Functions	Parameters			
	constant	b1	b2	b3
Linear	-2824.9	1.599	---	---
Quadratic	48771	-50.276	0.013	---
Cubic	-705.184	0.000	0.000	1.35E-07
Logarithm	-23794.113	3179.58	---	---

Substituting the parameters in Table 2 into equations (1~4). Therefore, the corresponding equations of the regression model are obtained as equation (5~8), respectively.

Linear: $y(t)= 1.599*t -2824.904$ (5)

Quadratic: $y(t)= =0.013*t\ 2 -50.276*t +48771$ (6)

Cubic: $y(t)= 0.0000001348*t\ 3+0.000*t\ 2+0.000*t-705.184$ (7)

Logarithm: $y(t)= 3179.58\ln(t)-23794.113$ (8)

Here we compare and analyze each model, and draw a graph of the sample data and model predictions. The graphs of linear regression, quadratic polynomial, and cubic polynomial regression

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.494.2023
models are shown in Figure 2. The x-axis represents the time, and the y-axis represents CO2 concentration. Using the Regression Models we build, Figure 2(a) is the Linear Model, Figure 2(b) is the Quadratic polynomial model, Figure 2(c) is the Cubic polynomial model, and Figure 2(d) is the logarithm model. The graphs show the sample point data in green and red. The predicted values are shown in red.

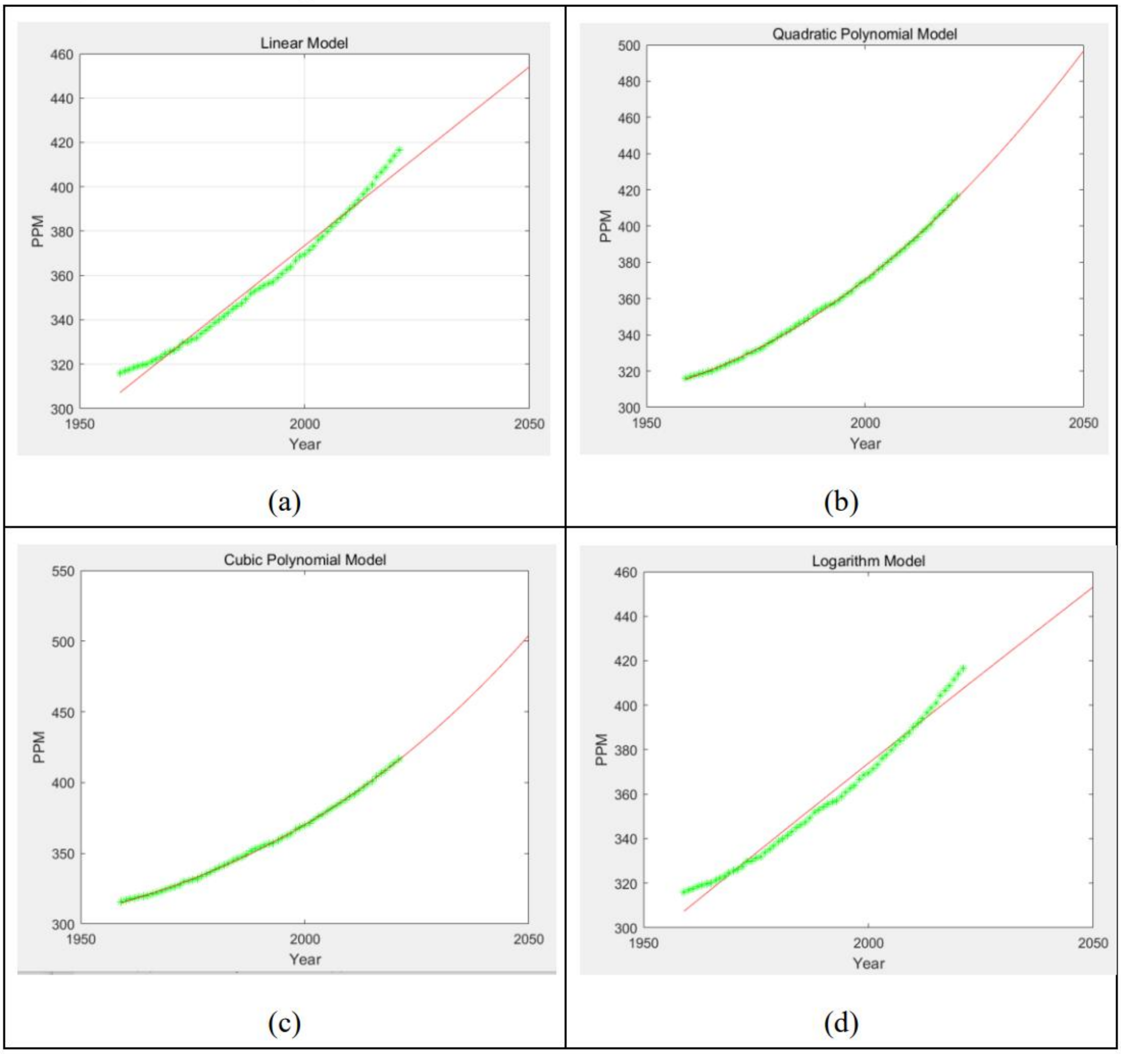

Fig.2 Sample Data and Model Prediction Comparison (Years_PPM).

The predicted CO2 concentration curve is shown in Figure 3. The x-axis is set to 1959- 2100, indicating that it starts from 1959 and is predicted to 2100, and the y-axis is set to the PPM value. The graph shows sampling point data in green, linear regression predicted values in red, quadratic polynomial regression predicted values in blue, cubic polynomial regression predicted values in yellow, and logistic regression predicted values in black.

ISSN:2790-1688
DOI: 10.56028/aetr.4.1.494.2023

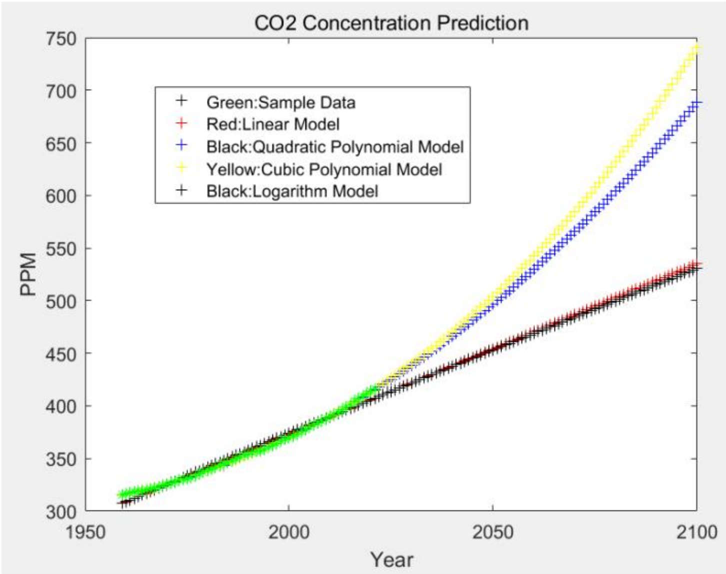

Fig.3 CO2 Concentration Prediction.

The values of CO2 concentration predicted to be reached in 2050 were solved by substituting t=2050 into the four regression models we built for the equation. The equation is also solved by substituting y(t)=685 to obtain the year corresponding to the CO2 concentration reaching 685, as shown in Table 4.

Table 4 Prediction of the Year Reaching the Specified CO2 Concentration.

Functions	Prediction Result		Year of Reaching 685 ppm
	2050	2100	
Linear	454.18	534.88	2194
Quadratic	496.81	688.33	2100
Cubic	497.94	696.22	2098
Logarithm	466.29	583.70	2136

3.3 Selection of CO2 prediction model

We calculated the correlation coefficient R^2 and F-test for the established linear, quadratic polynomial, cubic polynomial, and logarithm regression models, as shown in Table 5.

Table 5 R2 and F-test of different CO2 prediction models.

Models	R2	F-test
Linear	0.981	3242.148
Quadratic	0.982	3446.278
Cubic	0.983	3669.816
Logarithm	0.980	3055.278

Analyzing the R^2 and F-test values corresponding to various different regression models in Table 3.4, we can see that the values of R2 range from 0.980 to 0.983, which are close to 1, all reflecting a strong correlation between time and CO2 concentration. The F-test results are reasonable and less sensitive. From Figure 2 and Figure 3 we can see that the sample data are compared with the model prediction data, although the quadratic and cubic polynomial fits are high and the residuals are smaller. The rate of increase of CO2 concentration with time is too fast. According to the given data set, the concentration of CO2 in the atmosphere reached 377.7 ppm in March of 2004, resulting in

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.494.2023

the largest 10-year average increase up to that time. but the rate of increase tends to stabilize in the later 10-year period. Therefore, we believe that the Quadratic regression model is more accurate.

3.4 Land-Ocean Temperatures Changes Prediction

Based on the Data Source Credit from the National Aeronautics and Space Administration Goddard Institute for Space Studies, We use the year as the independent variable t and Land-Ocean Temperatures Changes as the dependent variable, and build a regression model respectively, the same as equations (1~4). Like the method for solving the atmospheric CO2 concentration parameters, we here also use the least squares method to solve each regression model parameter, we get the following results.

Linear: $y(t)= 0.017*t -33.117$ (9)

Quadratic: $y(t)= 0.000004228*t^2+0.000*t-16.395$ (10)

Cubic: $y(t) =0.000000001417*t^3+0.000*t^2+0.000*t-10.821$ (11)

Logarithm: $y(t) =33.445*\ln(t)-253.694$ (12)

we compare each model, draw a graph of the sample data and model predictions, calculate the mean square error, correlation coefficient R2 and F-test, and find the most suitable model.

Table 6 MSE, R2, and F-test of different Temperature Change prediction models.

Models	MSE	R2	F-test
Linear	0.0113	0.89	537.971
Quadratic	0.0085	0.92	544.607
Cubic	0.0082	0.92	551.201
Logarithm	0.0116	0.89	531.301

According to the value of MSE, R2, and F-test corresponding to the various regression models in Table 5, we can conclude that the quadratic polynomial regression model is the most appropriate, with a smaller MSE value for the mean square error, a larger R2 for the correlation coefficient, a reasonable and not too sensitive F-test result.

3.5 Correlation Analysis Between Co2 Concentration and Land-Ocean Temperatures

3.5.1 Correlation Analysis between CO2 Concentration and Land-Ocean Temperatures based on Regression Models

In order to better analyze the relationship between CO2 concentration and land-ocean temperatures, a scatter plot of atmospheric CO2 concentration versus global annual average surface air temperature variation (in degrees Celsius) is first drawn. it can be seen that both temperature change and carbon dioxide concentration.

ISSN:2790-1688 | DOI: 10.56028/aetr.4.1.494.2023

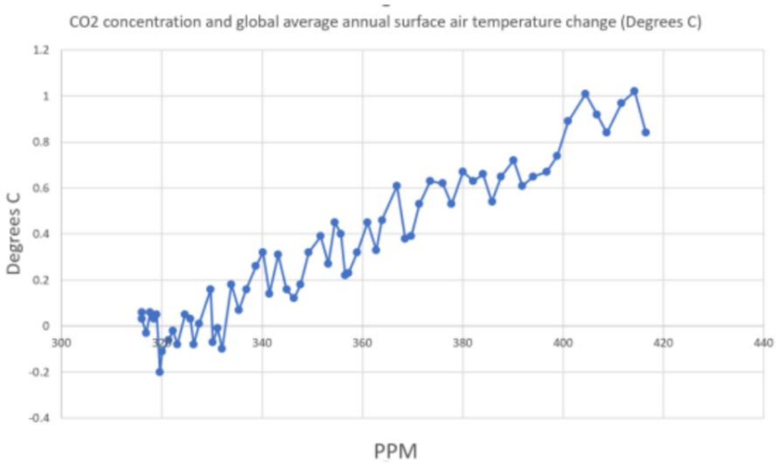

Fig.4 Data of CO2 Concentration and the Average Land-Ocean Temperature Change (°C)

We developed linear, quadratic polynomial, cubic polynomial and logarithmic regression models with CO2 concentration as the independent variable and Land-Ocean Temperatures Changes as the dependent variable, respectively. The corresponding equations of the regression model are obtained as equation (13~16), respectively. The correlation coefficient R2 and F-test of the various regression models we developed were calculated as shown in Table 6.

Linear: y(t) =0.010*x-3.393 (13)

Quadratic: y(t) =0.00000295*x 2+0.008*x-3.008 (14)

Cubic: y(t) =0.000* x 3+0.00000295*x 2+0.008*x-3.008 (15)

Logarithm: y(t) =3.776*ln(x)-21.835 (16)

Table 7 R2, and F-test of different prediction models.

Models	R2	F-test
Linear	0.924	743.036
Quadratic	0.924	365.700
Cubic	0.924	365.700
Logarithm	0.922	724.092

By comparing the values of R2 and F-test corresponding to various different regression models, we conclude that there is a relationship between CO2 concentration and land and ocean temperature changes. It is a strong positive correlation, which is verified by all four regression models.

3.5.2 Correlation Analysis between CO2 Concentration and Land-Ocean Temperatures based on Neural Network Model

This subsection sets up the sample set, data set, and test set based on the neural network model. We use 75% of the sample data as a Training set, 15% of the sample data as a Test set, and 15% of the sample data as a Validation set. After training, the curve of CO2 concentration and land-ocean temperature variation was fitted based on the BP neural network model [6,7] as shown in Figure 5.

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.494.2023

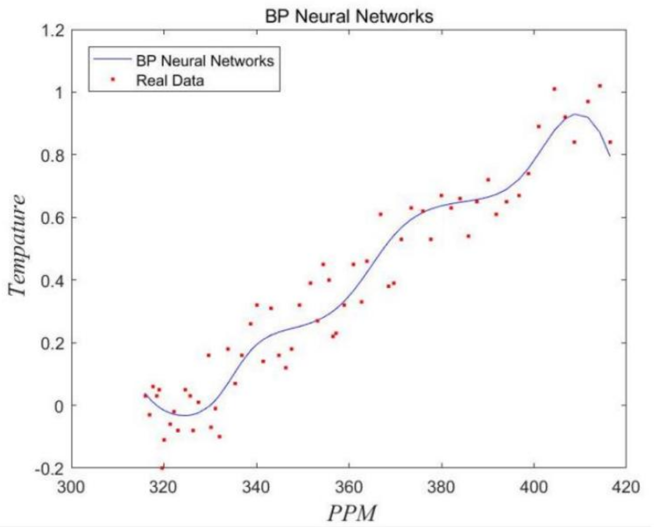

Fig.5 PPM and Land-Ocean Temperature Changes based on Neural Network Model

The best verification performance is 0.0070502 at epoch 3. The Mean Squared Error (MSE) equals 0.0070502 at that time, as shown in figure 7(a). Error histogram with 20 bins as shown in figure 7(b) represents the error classification in neural network training. In the figure, blue represents the training set, green represents the verification set, and red represents the test set, distributed on both sides of Zero Error. The correlation coefficient R2 of the training set is 0.97284, the R 2 of the Test set is 0.93483, the R 2 of the verification set is 0.96435, and all correlation coefficient R2 of CO2 concentration and land-ocean temperatures changes based on Neural network model is 0.96611. These values of R2 are all close to 1, indicating high fitting. The sentence must end with a period.

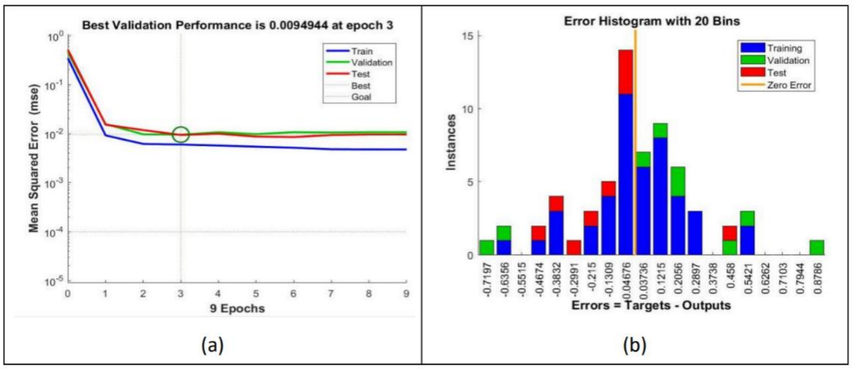

Fig.7 Model Training Based on BP Neural Network

3.6 Model Reliable Analysis of Correlation Between Co2 Concentration and Land-Ocean Temperature Prediction

It can be seen from the above analysis results of the land-ocean temperature prediction method based on CO2 concentration. Although the prediction results of the BP neural network model have a good fitting effect on historical data. However, from the changing trend of the prediction results based on the BP neural network, we can see that the BP neural network has the problem of overfitting, as shown in Fig. 4.6. Therefore, we choose a linear regression model to predict future changes in land ocean temperature based on CO2 concentration. For the prediction of a single point, the prediction error follows a normal distribution. It is assumed that the distribution of prediction error is as follows:

$$e = y - \hat{y} \sim N(0, \sigma^2)$$

Based on the distribution of errors, we can see that the prediction result $\hat{y}$ falls near the true value y with a certain probability. Therefore, at a certain confidence level, such as 5%, the

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.494.2023
confidence interval of the predicted value $\hat{y}$ is the 95% probability interval of the error distribution with the real value y as the center. Therefore, the confidence interval of the predicted value is [$\hat{y} - 2 * \sigma$, $\hat{y} + 2 * \sigma$]. In addition, it should be noted that when predicting the land-ocean temperature at a certain time in the future based on the CO2 concentration, the CO2 concentration value is also based on the prediction model, and there are certain errors in the prediction results. Therefore, we also need to consider the prediction error of CO2 concentration. Then the corresponding confidence interval is [$\hat{y} - 2 * \sigma - 0.01 * \sigma'$, $\hat{y} + 2 * \sigma + 0.01 * \sigma'$].

All paragraphs of section, subsection and sub subsection should have a first line indent of 0.5-centimeter. All paragraphs should be justified on both sides.

4. Conclusion

In recent years, with the continuous increase of CO2 concentration, the problem of global warming is also intensifying, which has led to a series of serious consequences. Therefore, the prediction of CO2 concentration and its relationship with temperature change plays an important role in the study of global warming. In this paper, the prediction methods for the future change trend of CO2 concentration and temperature are studied, and a variety of effective prediction models are established, which can predict the past and future CO2 concentration and land-ocean temperature. The advantages of our research mainly include the following:(1) We define a simple and effective average increase calculation method. By analyzing the change of average increase, we can more accurately understand the changing trend of CO2 and more clearly judge the future change of CO2 concentration. (2) We used a variety of linear and nonlinear regression functions and BP neural network methods to establish prediction models for CO2 concentration and land-ocean temperature respectively. And through the evaluation of the prediction model, the most accurate prediction model can be selected. (3) By establishing a prediction model of land-ocean temperature based on CO2 concentration and combining it with a correlation test, the correlation degree between them is analyzed, and the impact of CO2 concentration change on land-ocean temperature change is verified.

The main disadvantage of the model and research content in this paper is that our model is based on time variables and mainly establishes the prediction function of land-ocean temperature/CO2 concentration. However, the change of time is fixed, which is not the reason for the change in land-ocean temperature. When the internal factors, such as the factors related to CO2 emissions and absorption, change greatly, the time-based prediction model will become invalid.

References

[1] Richardson, T. K. (2010). Estimates of global sources and sinks of carbon from land cover and land use changes. AGU Fall Meeting Abstracts.

[2] Jenkinson, D. S., Adams, D. E. & Wild, A. (1991). Model estimates of co2 emissions from soil in response to global warming. Nature, 351(6324), 304-306.

[3] Raddatz, R. L. (2003). Agriculture and tornadoes on the Canadian prairies: potential impact of increasing atmospheric co2 on summer severe weather. Natural Hazards, 29(2), 113-122.

[4] Stewart, R. E. (2010). Drought, Climate Change and the Canadian Prairies. 6th Alexander von Humboldt International Conference on Climate Change, Natural Hazards, and Societies, held March 15-19 in Merida, Mexico. http://meetings.copernicus.org/avh6, id.AvH6-28.

[5] Rogers, H. (2011). Impact of increasing atmospheric co2 on carbon dynamics under different tillage practices. Chinese Journal of Clinical Oncology & Rehabilitation, (1), 75-84.

[6] Zhang, Y. & Wu, L. (2009). Stock market prediction of s&p 500 via combination of improved bco approach and bp neural network. Expert Systems with Applications, 36(5), 8849-8854.

[7] Wei-tao, et al. (2019). "Study on Strawberry CO2 Gas Fertilizer in Greenhouse Based on BP Neural Network."