The Relation Between the CO_2 Concentration Levels and the Temperature

Yuyuan Lin, Zidong Zeng, Lang Chen, Zifeng Deng

Guangdong Experimental High School AP Curriculum, Guangzhou 510000, Guangdong, China

Abstract: Carbon dioxide can change the heat balance of the atmosphere. To study the relationship between CO_2 and temperature change, we use the given CO_2 concentrations to implement linear regression model, gray time series forecasting model, back-propagation model, and auto-regressive and moving average model to establish the growth function of CO_2 concentration. Errors are evaluated to choose the most suitable model. then we use the most suitable model in step one to further predict future land-ocean temperature. Finally, we use grey relational analysis to analyze the relationship between CO_2 concentrations and land-ocean temperatures since 1959.

Keywords: Global warming; prediction and fitting.

1. Introduction

Carbon dioxide can change the heat balance of the atmosphere. It can absorb infrared radiation from the Earth, causing an increase in near-surface atmospheric temperature. The warming of the near-surface atmosphere causes enhanced evaporation from the ground, resulting in an increase in atmospheric water vapor. The increase of water vapor in the atmosphere will further enhance the absorption of infrared radiation from the Earth by the near-surface atmosphere. As a result of this interaction, the increase of carbon dioxide in the atmosphere will change the original atmospheric heat balance and cause global warming.

To study the relationship between CO2 and temperature change, we use the given CO2 concentrations to implement linear regression model, gray time series forecasting model, back-propagation model, and auto-regressive and moving average model to establish the growth function of CO2 concentration. Errors are evaluated to choose the most suitable model. then we use the most suitable model in step one to further predict future land-ocean temperature. Finally, we use grey relational analysis to analyze the relationship between CO2 concentrations and land-ocean temperatures since 1959.

2. Assumptions of the Models and Notation

Assumption 1: We assume that there will be no future extreme natural conditions that would cause dramatic changes in carbon dioxide.

Justification 1: Extreme situations which dramatically increase the concentration of CO2, including volcanic eruption, or mountain fire of which the former has tiny probability of occurrence and the later has similar trend of increasing occurrence from 1959-2021 to the more distant future. Therefore, it is reasonable to ignore these extreme factors when establishing a model in this paper.

Assumption 2: We assume that recent climate patterns will continue in the future.

Justification 2: We anticipate that it will take a while for these initiatives to progress from research to mature operation and considerably benefit the ecological environment, even though changes in government regulations and technical advancement may lower future greenhouse gas emissions. There won't be any significant shifts in climatic patterns at least between now and 2100.

Assumption 3: We assume that the concentration of carbon dioxide is the only factor affecting global warming.

Justification 3: Apart from carbon dioxide; methane, nitrous oxide, hydrochlorofluorocarbons (HCFCs), hydrofluorocarbons (HFCs) and ozone are also common greenhouse gases that contribute to the overall warming effect. However, according to observations by the NOAA Global Monitoring Lab, in 2021 carbon dioxide alone was responsible for about two-thirds of the total heating influence of all human-produced greenhouse gases. [3] Therefore, for the sake of simplicity, we only take the effect of CO2 into consideration in terms of temperature change.

Assumption 4: We assume that the rising temperature will not decrease the solubility of CO2 in ocean to the extent which largely affects carbon dioxide level in the atmosphere.

Justification 4: The solubility of CO_2 in water has been the subject of much investigation since its discovery in 1803. A model based on Henry's rule could be able to predict how much CO_2 will be absorbed by water as a function of temperature and pressure. The Henry's constant (MPa/mol percentage), according to a critical examination of 91 publications in prestigious journals from around the world, grows linearly as temperature rises from 0 to 150 °C. [4] However, since the standard deviation of given temperature data calculated by excel is 0.335466113, a rather small value, it is reasonable to ignore the effect of temperature change on CO_2 solubility.

Table 1 Notation

t	discrete time variable for years
C(t)	history data of CO2 concentration level
a	for cubic exponential smoothing model
p	prediction for auto regressive integrated
e	white noise

3. Model Theory, Implementation and Result

3.1 CO2 Concentration Changed over Time Elapsed

3.1.1 Data Processing

We use two methods to detect the outliers which may affect the model and prediction. Firstly, we draw a best fitting curve based on scattered data using excel. We choose to use polynomial function for curving fitting, and the R-squared values=0.9181 indicates the data are to the fitted regression line. Since no value deviates from the trend line in particular, we assume that there is no outlier.

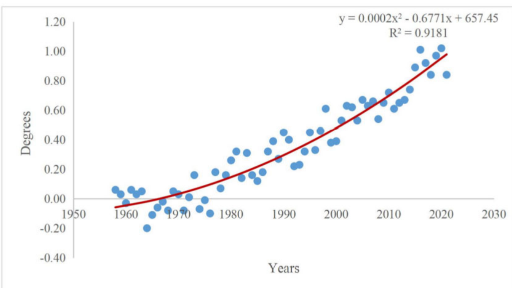


Figure 1 Fitted Regression Line of co2 Concentrations from 1959 to 2021

3.1.2 Polynomial Fitting

Let the polynomial be $y(w,x) = w_0 + w_1x + w_2x^2 + \cdots w_Mx^M = \sum_{j=0}^{M} x_jx^j$. and we transform the polynomial into a mean-square error function:

$$E(w) = 1/2\sum_{n=0}^{N}\{y(x_n,w)-t_n\}^2$$

We need to fit a curve such that the value of E takes a minimal value, where E is the mean-square error function. When the number of parameters is too large, because the noise in the dataset is given excessive importance, the fitting results will tend to include the noise in the model, and over-fitting may occur, i.e., the model and parameters fit well for a given dataset, but cannot give suitable estimates of the dependent variable for the new independent variable. Therefore, we can use the regularization method to add a penalty term to the error function to obtain:

$$E(w) = 1/2 \sum_{n=0}^{N}\{y(x_n,w)-t_n\}^2 + \lambda/2||x||^2$$

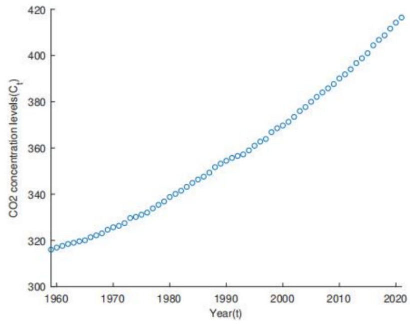


Figure 2 Measured and Predicted Carbon Dioxide Concentrations from 1959 to 2100, Generated by Polynomial Fitting

3.1.3 Grey GM (1,1) Model

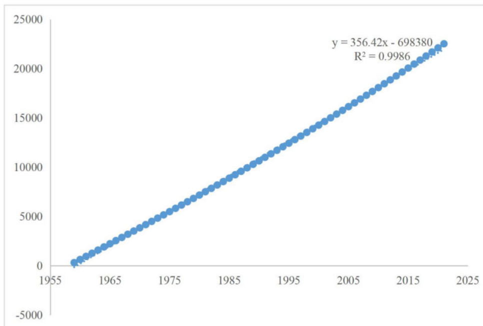
First of all, in order to confirm the feasibility of the data using the GM (1,1) model, we should calculate $\lambda(k) = \frac{x_0(k-1)}{x_0(k)}$, k=2, 3…, n_0. if λ(k)is in the interval ($e^{\frac{-2}{n+1}}$, $e^{\frac{2}{n+2}}$), the GM (1,1) model can be used.
Secondly, we let the original sequences for the 63 years 1959-2021 be respectively $x_1^{(0)}$, $x_2^{(0)}$, $x_3^{(0)}$ … $x_{63}^{(0)}$. Then we make an accumulation to generate a new sequence $x_1^{(1)}$, $x_2^{(1)}$, $x_3^{(1)}$ … $x_{63}^{(1)}$, where there are

$$x_k^{(1)} = \sum_{i=1}^{k} x_i^{(0)}, k = 1,2,3 \dots 63$$

Afterwards, the mean series is generated:

$$z_k^{(1)} = \alpha x_k^{(1)} + (1-\alpha)x_{k-1}^{(1)}$$

where 0 ≤ α ≤ 1, are the weights. Usually α = 0.5. Since the new sequence generated looks like a straight line, we can approximate this new sequence with an expression for a straight line.



 https://doi.org/10.1051/e3sconf/202337503014

Figure 3 Function of $x_k^{(1)} = \sum_{i=1}^{k} x_i^{(0)}$

Therefore, we construct a first-order ordinary differential equation to solve for the functional expression of the fitted curve. From this figure, the gray differential equation is established as

$$x_k^{(0)} + az_k^{(1)} = b$$

The corresponding GM (1, 1) whitening differential equation is

$$\frac{dx^{(1)}}{d_t} + ax_t^{(1)} = b$$

Shifting the term of the grey differential equation gives.

$$-az_k^{(1)} + b = x_k^{(0)}$$

a,b are parameters to be determined, the above equation can be written in the form of a matrix in the form of Xβ = Y

Using least squares method, we can determine the parameter matrix β:

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

The resulting estimates of the parameters a,b are brought into the whitening equation to obtain the sequence $x_k^{(1)}$:

$$\widehat{x_k^{(1)}} = (x_1^{(0)} - b/a)\, e^{-\alpha(k-1)}(1 - e^{a}), k = 1,2,3 \ldots 63$$

when k=1,2,3...63, we can get the fitted value; when k > 63, we can get the predicted value. We realize this model by using MATLAB. Carbon dioxide concentrations predicted from year 2022 to year 2100 are as follows:

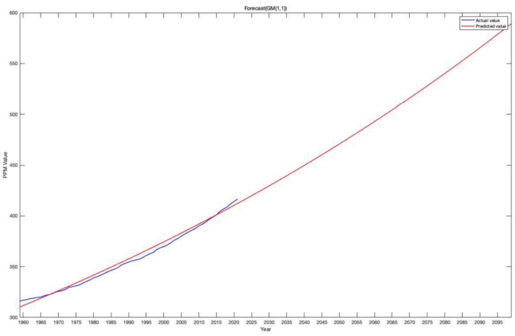


Figure 4 Measured and Predicted Carbon Dioxide Concentrations from 1959 to 2100, Generated Using Grey GM (1,1) model

3.1.4 Autoregressive Integrated Moving Average Model

From the fitted curves of the sample time series of carbon dioxide concentrations and years, it is clear that carbon dioxide concentration cannot continue "inertially" in its current state for some time in the future. Therefore, the series needs to be differenced.

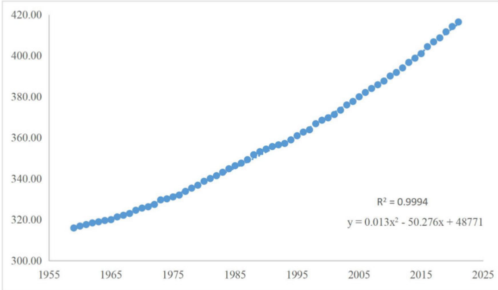


Figure 5 Fitted Curves of the Time Series of Carbon Dioxide Concentrations and Years

The model was initially determined as ARIMA(p,1,q) due to the first-order differencing of the original data. Based on the characteristics of the differenced PPM data, the order of the model can be determined based on the autocorrelation and then autocorrelation plots (ACF & PACF plots). The autocorrelation and partial autocorrelation plots are plotted with the help of the functions acf() and pacf() in MATLAB as follows:

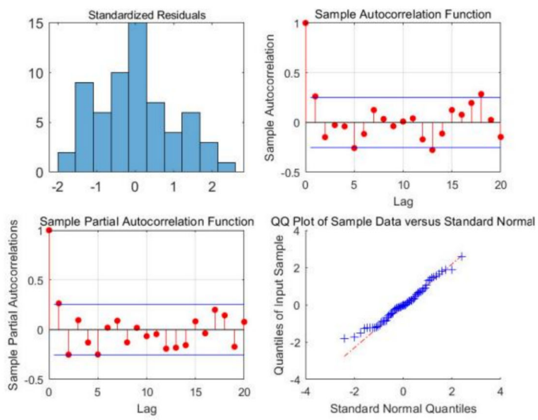


Figure 6 The autocorrelation and partial autocorrelation plots

According to the autocorrelation and partial autocorrelation plots, judging the stage of the model generally requires subjective guesses. From the information in the graphs, we can conclude that p=2 and q=1 in the ARIMA model.

According to the graph of residual test results, standardized residuals is to see if the residuals are close to normal distribution. According to the QQ plot below, the blue points are near the red line, so the residuals are close to the normal distribution. Therefore, the model ARIMA (2,1,1).Based on the ARIMA (2, 1, 1) model, the carbon dioxide concentration values can be predicted for 2022-2100. The prediction can be done by MATLAB function, results are shown in the following figure.

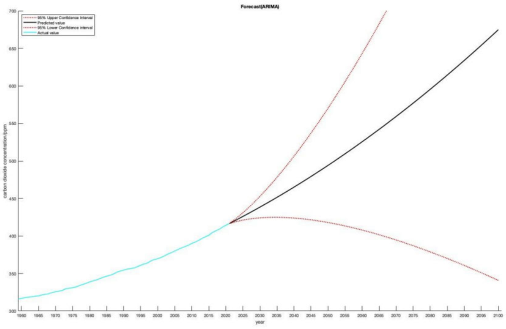


Figure 7 Measured and Predicted Carbon Dioxide Concentrations from 1959 to 2100, Generated by Autoregressive Integrated Moving Average Model

3.2 Comparison

We use three models in total to describe the changes in carbon dioxide concentrations over time and predict future carbon dioxide concentrations from 2022 to 2100.
R2 is a measure of the goodness of fit of a model. We compare the R2 value for each model, it shows that the R2 values of polynomial fitting, grey GM (1,1) model, and autoregressive integrated moving average model are 0.8912,0.9867, and 0.9994, respectively, which proves ARIMA the most accurate model.

3.3 Model Future Land-Ocean Temperatures Changes

To determine future land-ocean temperatures changes, we use autoregressive integrated moving average model again due to its relative accuracy. Since we have discussed the principle of the model in the previous part, we will emphasize on the processes for predicting future land-ocean temperatures changes, which we have not covered before.
First of all, since the temperature for 1951-1957 were not found, we first used fitting to graph the data for 1958-1980 and expressed them as a function, The function gives y = 11.688ln(x) -88,639, where x represents the year and y represents the temperature value. the year is very unstable and irregular in the data co-fitted to the temperature. Therefore, a 1st order difference treatment is required for this series. In this case, the time series is weakly stationary. On the basis of the first-order difference of the given data, it was first determined that the model should be ARIMA (p, 1, q), and then the p and q values were determined based on the autocorrelation and then autocorrelation plots (ACF & PACF), respectively. Since the p, q values are subjective, we still set p=2 and q=1 here.

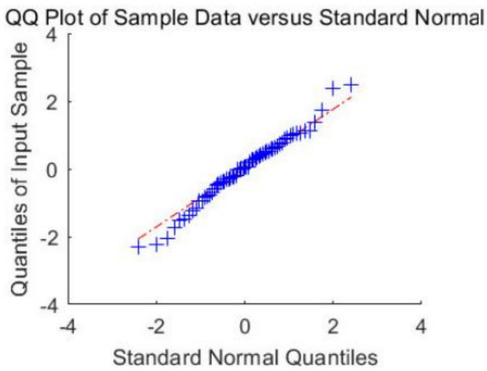


Figure 8 QQ Chart (Sample data-Standard Normal Distribution)

In this figure, almost all the blue points are very close to the red line, which means that the residuals are close to the normal distribution. Therefore it is feasible to use ARIMA (2,1,1) to predict future temperature.

4. Relationship Between CO2 Concentrations and Land-Ocean Temperatures Since 1959

The first method of relationship analysis in this paper is Grey relational analysis. Using excel, we can obtain matrix of the CO2 concentration and the temperature. after nondimensionalization. As calculated by MATLAB, the value of gray correlation is 0.9298, which shows high correlation between carbon dioxide concentration and temperature.
We also use linear regression model to further confirm the correlation between carbon dioxide concentration and temperature. By using excel, we can find that carbon dioxide concentration and temperature shows a linear correlation where the goodness of fit R2=0.9216, approximately the same as the value of gray correlation, which proves that either of the method is feasible.

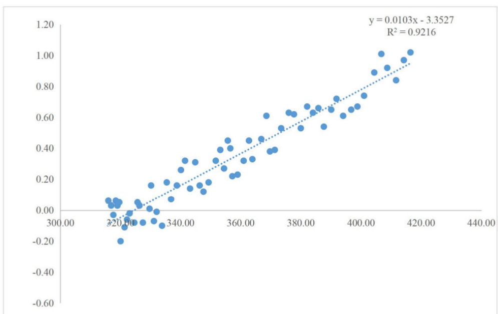


Figure 9 Correlation between Carbon Dioxide Concentrations and Temperatures

5. Conclusion and Possible Recommendations for the Future

We obtain the prediction data of CO2 concentrations and land-ocean temperatures by ARIMA model and compute the Spearman rank correlation coefficient of our two time series as 0.9989. Thus we can conclude that our model is reliable into the future by the probability which is close to 1. Notably, we have not used the history data, which are given in stable conditions and environment. Therefore, if the climate conditions and environment change acutely, our model's ability to predict future CO2 concentration levels and land-ocean temperatures may be worse, i.e., when extreme weather appears our model requires more improvement.

References

1. ZHANG Zhenghai, ZHANG Yayu, LEI Huixia, SHAO Cai, WU Chen, QIAN Jiaqi. Effect of light intensity-temperature-humidity and carbon dioxide on photosynthetic rate of ginseng[J]. Specialty Research,2020,42(02):41-46.
2. Xiao Bo. A coupled calculation method of temperature-pressure field for supercritical carbon dioxide dry fracturing[J]. China Science and Technology Paper,2020,15(01):119-124.
3. Li Yaohua. Effects of elevated temperature and carbon dioxide concentration on the biological characteristics of Aphis minor wasps[D]. Northwest Agriculture and Forestry University, 2018.
4. Zhang YJ. Experimental study on the effects of temperature and carbon dioxide on human comfort in a small space[D]. Chongqing University,2018.
5. Liang W B. Analysis of high exhaust temperature of CP-11 type carbon dioxide high-pressure compressor[J]. Light Industry Science and Technology,2017,33(04):50-52+60.
6. Feng Mengmeng,Tan Yingxin, Zhang Xiqing, Yang Yajun. Experimental study on the effect of initial temperature on the effect of carbon dioxide explosion suppression[J]. Fire Science and Technology,2015,34(11):1423-1425.
7. Wu Xiaohong, Xiao Min,Yang Wenyuan, Ren Bin, Ma Jingang. Experimental design of using temperature sensor to investigate the relationship between carbon dioxide and greenhouse effect[J]. Reference for teaching secondary school chemistry, 2015, (11):57-58.
8. Coco Zhang, Qi Y, Fu X, Miao B, Bai Xuejiao. Carbon dioxide temperature measurement based on TDLAS primary harmonics[J]. Optoelectronics, 2014, 34(04):300-303+307+312.
9. Translated with www.DeepL.com/Translator (free version) Climate Change, N. G. (2022, September 1). Carbon Dioxide Concentration | NASA Global Climate Change. Climate Change: Vital Signs of the Planet. Retrieved November 15, 2022, from https://climate.nasa.gov/vital-signs/carbon-dioxide
10. United Nations. (n.d.). The ocean – the world' s greatest ally against climate change. https://www.un.org/en/climatechange/science/climate-issues/ocea