

A study on global warming based on time series prediction model

Yao Zhang^{1, *}

¹ College of Physics and Information Engineering, Fuzhou University, Fuzhou, China

* Corresponding Author Email: 052003115@fzu.edu.cn

Abstract. Global warming is a natural phenomenon that has been hotly discussed in recent years. The reason for this is that since the industrial revolution, human activities have become more frequent and the development of carbon resources has become more intense, leading to an increase in the concentration of greenhouse gases in the air every year [4]. In this paper, we preprocess the given data and then find the missing data ourselves to determine whether the temperature increase in March 2022 leads to a greater temperature increase than in the past decade. Then, we build a time series prediction model and an exponential smoothing model to describe the past and predict the future temperatures. With the help of these two models, we give specific temperatures for both models in 2050 and 2100 and predict the time when the global temperature reaches 20° C. Finally, the prediction accuracy of the two models is evaluated and compared.

Keywords: Global warming; ARIMA; exponential smoothing method; greenhouse effect.

1. Introduction

Global warming must be one of the most concerning issues in the international community today. The temperature trend is gradually increasing. This has led to the dissolution and melting of many primitive glaciers or frozen soils. As the recent news: "Scientists find ancient viruses still partially active in thawing permafrost" shows that in addition to the impact of global warming on the environment and biodiversity, it also has a relatively direct impact on humans. After all, following the outbreak of the New Crown pneumonia epidemic in recent years, we believe that no one wants to go through similar or worse times [1].

Based on the data tables given, we need to first supplement and pre-process the data. Pre-processing usually includes a list of groups, conversion of data types, etc. After that, we need to make a relatively rigorous judgment: whether the rise in global temperature in March 2022 has led to a greater rise than in the past decade. Then two models need to be built to predict the global temperature trend after 2022 and give the approximate year when the global temperature reaches 20°C. We will then analyze the fit of the two models we have built.

2. Analytical Observability

2.1. Observability analysis

First of all, we can make a visual analysis of all the data obtained to roughly grasp the trend of global temperature changes.

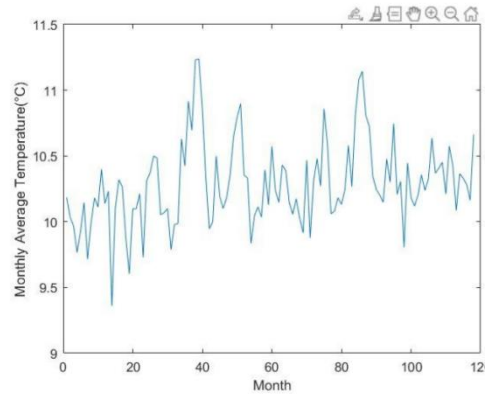


Fig. 1 Observability trend chart of actual temperature

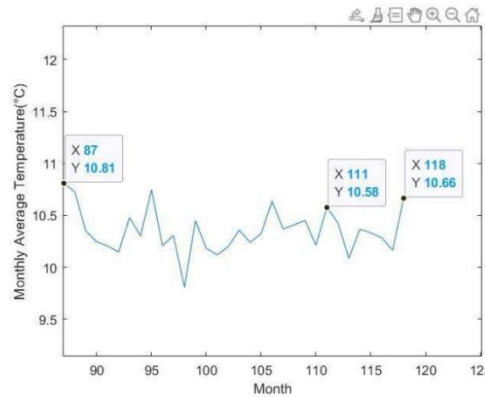


Fig. 2 Enlarge the last visual trend chart

The above is the result diagram of visual analysis. The first one is from ten years ago to October 2022. The latter one is a trend chart obtained by terminal enlargement of the previous one. (X=111, Y=10.58) is the coordinates of March 2022.

From these two figures, we can intuitively see that there is no significant upward trend after March 2022.

2.2. Difference analysis

In order to verify the conclusions of the observability analysis, we still need to make specific calculations. Here we divide all data in the first ten years of March 2022 and all data after March 2022 into two groups, and use independent sample T test to judge the difference between the two groups of data [5].

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s\bar{X}_1 - \bar{X}_2} = \frac{\bar{X}_1 - \bar{X}_2}{s\bar{X}_1 - \bar{X}_2}, v = n_1 + n_2 - 2 \quad (1)$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_c^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (2)$$

$$= \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sum X_1^2 - (\sum X_1)^2/n_1 + \sum X_2^2 - (\sum X_2)^2/n_2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (3)$$

$$= \frac{\bar{x}_1 - \bar{X}_2}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}, v = n_1 + n_2 - 2 \quad (4)$$

Independent sample t-test is used to analyze the difference between two groups of direct quantitative data, which is a method of difference test. It has many applicable conditions. It has many applicable conditions. The most important is that the observed variables in each group are normal (or nearly normal) distribution & the variance of the two groups of observed variables is equal. Therefore, we need to do a normality test for the two groups of data here.

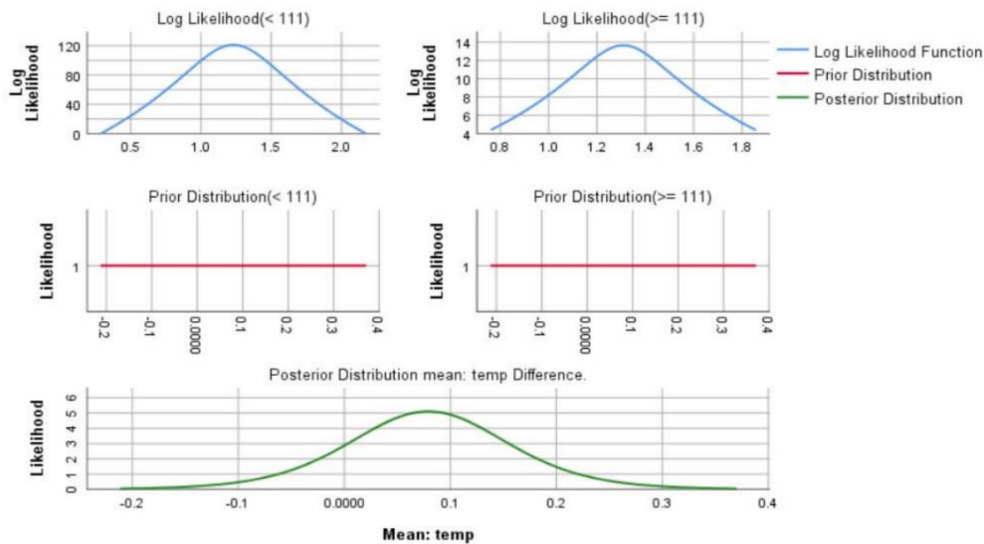


Fig. 3 Results of normality test

From the above images, we can know that these two groups of data are approximately normal distribution. Then we use independent sample t test to get the following images:

➔ **Bayesian Independent**

Group Statistics				
temp	timeorder	N	Mean	Std. Deviation
< 111		110	1.22977	.333667
>= 111		8	1.30900	.194025

Posterior Distribution Characterization for Independent Sample Mean^a

temp	Posterior			95% Credible Interval	
	Mode	Mean	Variance	Lower Bound	Upper Bound
	.07923	.07923	.008	-.09509	.25354

a. Prior for Variance: Diffuse. Prior for Mean: Diffuse.

➔ **T-Test**

Group Statistics				
temp	timeorder	N	Mean	Std. Deviation
>= 111.00		8	1.30900	.194025
< 111.00		110	1.22977	.333667

Independent Samples Test									
Levene's Test for Equality of Variances					t-test for Equality of Means				
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
Equal variances assumed	1.842	.177	.662	116	.509	.079227	.119719	-.157891	.316345
Equal variances not assumed			1.048	10.304	.319	.079227	.075616	-.088584	.247039

Fig. 4 Independent t-test results

Till now we can conclude that there is no significant difference between the two groups of data. The results are the same as those obtained from our previous visual analysis.

3. Time series forecasting models

3.1. Global Temperature Level Projections

To describe the past, we can directly use all the data in the past to make a distribution map. This step can also be called visual analysis. For subsequent prediction, we use ARIMA method and exponential smoothing method here. ARIMA model is a linear model proposed by Box and Jenkins for time series analysis and prediction. In time series analysis, three parameters need to be set, namely autoregressive order (p), differential order (d) and moving average order (q) [2]. The general form of ARIMA (p, d, q) is as follows:

$$\begin{cases} \phi(B)\nabla^d y_i = \theta(B)\varepsilon_i \\ E(\varepsilon_i) = 0 \\ E(x_i \varepsilon_s) = 0 \end{cases} \quad \begin{cases} \text{var}(\varepsilon_i) = \sigma_\varepsilon^2, E(\varepsilon_i \varepsilon_s) = 0, s \neq i \\ s < t \end{cases} \quad (5)$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (6)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (7)$$

Exponential smoothing method is a time series forecasting method which is based on the actual number and forecast number of an indicator in the current period and introduces a simplified weighting factor, namely smoothing coefficient, to obtain the average. That is to say, a special weighted average method is used to give greater weight to the historical data close to the prediction period, and the weight decreases from near to far according to the exponential law [3].

The ARIMA prediction method needs to determine the normality of the data itself, and here ACF and PACF are used as the test basis. The difference also needs to be checked.

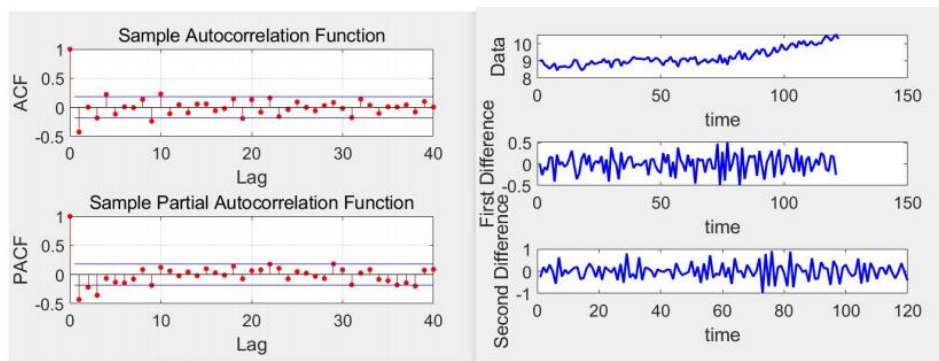


Fig. 5 Normality test and difference smoothing test

Through calculation, we can know that the parameters of ARIMA here are 1, 1 and 0 respectively. Therefore, the following images are drawn:

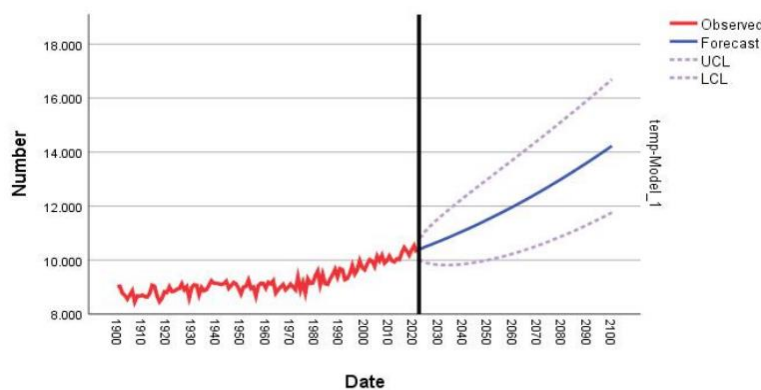


Fig. 6 The figure of ARIMA(1.1.0)

The second prediction method we adopt is Brown exponential smoothing method in exponential smoothing. The use conditions of exponential smoothing method are much simpler than ARIMA

prediction method. However, in terms of long-term prediction, the prediction accuracy of exponential smoothing is much worse than ARIMA.

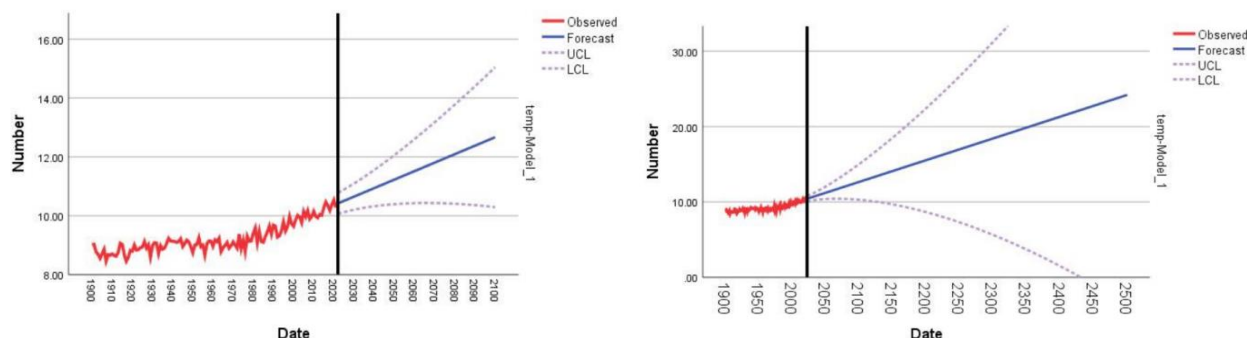


Fig. 7 Normality test and difference smoothing test

Based on the two models above, we can output the specific temperature for each specific year for both models. We can output specific temperature data for each specific year for both models. the specific temperature data for 2050 for ARIMA, Brown and 2100 and before and after for ARIMA and Brown exponential smoothing methods are as follows:

Table 1. ARIMA (1.1.0) model in 2050 and 2100

2047	2048	2049	2050	2051	2052	2053
11.380	11.424	11.468	11.513	11.558	11.604	11.649
12.809	12.880	12.950	13.020	13.090	13.161	13.231
9.950	9.968	9.986	10.006	10.026	10.046	10.068

2097	2098	2099	2100	2101	2102	2103
14.041	14.104	14.167	14.230	14.294	14.359	14.424
16.460	16.538	16.617	16.696	16.776	16.855	16.935
11.622	11.669	11.717	11.765	11.813	11.862	11.912

Table 2. Brown index smoothing model in 2050 and 2100

2047	2048	2049	2050	2051	2052	2053
11.14	11.17	11.20	11.23	11.26	11.29	11.32
11.91	11.96	12.01	12.06	12.11	12.17	12.22
10.38	10.38	10.39	10.40	10.40	10.40	10.41

2097	2098	2099	2100	2101	2102	2103
12.58	12.61	12.64	12.67	12.70	12.73	12.76
14.85	14.92	14.98	15.05	15.11	15.18	15.24
10.32	10.31	10.30	10.29	10.29	10.28	10.27

We find that neither of the two models can reach 20°C in 2050 and 2100. When the temperature reaches 20°C, the specific corresponding years of the two models are as follows:

Table 3. ARIMA (1.1.0) model reaching 20°C for specific years

Model	2172	2173	2174	2175	2176	2177	2178
temp-Model_1							
Forecast	19.803	19.894	19.986	20.078	20.170	20.263	20.356
UCL	23.203	23.306	23.408	23.511	23.615	23.719	23.823
LCL	16.403	16.483	16.563	16.644	16.725	16.807	16.889

Table 4. Specific years in which the Brownian smoothing model reached 20°C

2352	19.93
2353	19.96
2354	19.99
2355	20.02
2356	20.05
2357	20.08
2358	20.11
2359	20.14

3.2. Model Evaluation

For the above two models, we compare the past part of the model with the actual past part of the database, and analyze the fitting accuracy of the model. The results are as follows:

Table 5. Fitting accuracy of ARIMA (1.1.0) model

Model Summary

Model Fit												
							Percentile					
Fit Statistic	Mean	SE	Minimum	Maximum	5	10	25	50	75	90	95	
Stationary R-squared	.190	.	.190	.190	.190	.190	.190	.190	.190	.190	.190	
R-squared	.839	.	.839	.839	.839	.839	.839	.839	.839	.839	.839	
RMSE	.200	.	.200	.200	.200	.200	.200	.200	.200	.200	.200	
MAPE	1.768	.	1.768	1.768	1.768	1.768	1.768	1.768	1.768	1.768	1.768	
MaxAPE	4.950	.	4.950	4.950	4.950	4.950	4.950	4.950	4.950	4.950	4.950	
MAE	.163	.	.163	.163	.163	.163	.163	.163	.163	.163	.163	
MaxAE	.450	.	.450	.450	.450	.450	.450	.450	.450	.450	.450	
Normalized BIC	-3.103	.	-3.103	-3.103	-3.103	-3.103	-3.103	-3.103	-3.103	-3.103	-3.103	

Model Statistics

Model	Number of Predictors	Model Fit statistics Stationary R-squared	Ljung-Box Q(18)			Number of Outliers
			Statistics	DF	Sig.	
temp-Model_1	1	.190	32.646	17	.012	0

Table 6. Fitting Accuracy of Brownian Exponential Smoothing Model

Model Summary

Model Fit												
								Percentile				
Fit Statistic	Mean	SE	Minimum	Maximum	5	10	25	50	75	90	95	
Stationary R-squared	.767	.	.767	.767	.767	.767	.767	.767	.767	.767	.767	
R-squared	.866	.	.866	.866	.866	.866	.866	.866	.866	.866	.866	
RMSE	.180	.	.180	.180	.180	.180	.180	.180	.180	.180	.180	
MAPE	1.588	.	1.588	1.588	1.588	1.588	1.588	1.588	1.588	1.588	1.588	
MaxAPE	4.443	.	4.443	4.443	4.443	4.443	4.443	4.443	4.443	4.443	4.443	
MAE	.146	.	.146	.146	.146	.146	.146	.146	.146	.146	.146	
MaxAE	.413	.	.413	.413	.413	.413	.413	.413	.413	.413	.413	
Normalized BIC	-3.389	.	-3.389	-3.389	-3.389	-3.389	-3.389	-3.389	-3.389	-3.389	-3.389	

Model Statistics

Model	Number of Predictors	Model Fit statistics Stationary R-squared	Ljung-Box Q(18)			Number of Outliers
			Statistics	DF	Sig.	
temp-Model_1	0	.767	14.385	17	.640	0

By comparing the two, we can know that ARIMA (1.1.0) has a better fitting effect in this database. It should also prove that the aforementioned Brownian exponential smoothing model is not suitable for long-term high data volume prediction.

4. Summary

In the current situation, the global temperature rise trend is still gradually increasing and the situation is becoming more and more pessimistic. Therefore, relevant measures must be taken to slow down the economic growth rate. Our team first predicted the general trend of global temperature change and then analyzed the factors that will lead to the global temperature increase. In analyzing the causes of global temperature change, our team found that the continued production of carbon dioxide is the main reason for the continued rise in temperature. Although factors such as population and GDP are also associated with global temperature increases, they are not the main cause. Therefore, if we take too many relevant measures in these areas, we may not get effective results. Therefore, the most effective way to curb global warming is through measures to reduce CO₂ emissions.

In this paper, the time series forecasting method and exponential smoothing method are used in predicting the future global temperature level. The time series model can perform a smoothness test on the time series, so as to determine whether the time series data are smooth or not. In addition, the time series model can also perform seasonal decomposition so as to determine the periodicity of the data. In addition, the time series model can also perform a unit root test so as to determine the smoothness of the data. The exponential smoothing model can perform a smoothness test on the time series data, so as to determine whether the time series data are smooth or not. In addition, the exponential smoothing model can also perform seasonal decomposition so as to determine the periodicity of the data. In addition, the exponential smoothing model can also perform unit root tests and thus determine the smoothness of the data.

However, the exponential smoothing model assigns less weight to the forward data and more weight to the near-term data, so it can only make short-term forecasts. In addition, the prediction results of the exponential smoothing model are affected by the exponential smoothing value, so some adjustment is required to obtain more accurate prediction results.

References

- [1] Reddy D M. ARIMA time series modeling of the COVID-19 outbreak in India [J]. IJMH, 2020, 4: 41-45.
- [2] Miller J W. ARIMA time series model of full truckload transport prices[J]. Forecasting, 2018, 1(1): 121-134.
- [3] Zhang Y, Choo W C, Ho J S, et al. Single or Combine? Travel demand volatility forecasting with exponentially weighted and smoothed transition combination methods [J]. Computation, 2022, 10(8): 137.
- [4] Hao Yu. Coping with global warming! Urgent! [J]. Nature Discovery, 2022(7):30-34.
- [5] Zhu D.B., Wu X.J., Xiang Y., et al. Application of t-test and HorRat equation in method validation[J]. Shanghai Chemical Industry, 2022, 47(2): 44-46.