

## Analysis of Global Warming and its Influencing Factors

Fan Zhang, Jiayu Chi, Mengting Song

School of Business Administration, Anhui University of Finance and Economics, Bengbu, China

### Abstract

In recent years, the earth is burning, whether global warming and its influencing factors, based on time series and gray prediction model, put forward the hypothesis and conclusion, provide the data of various factors, to evaluate the past temperature level, to predict the future global temperature level and its trend and analyze the influencing factors. For problem 1, first, based on the analysis of the global average air temperature data from March 2012-February 2022 and March 2022-October 2022, it is unscientific that the global temperature rise in March 2022 led to a larger rise than in the past 10 years. Second, the ARIMA model and the gray prediction model are used to describe the past global temperature levels and predict the future global temperature levels, respectively. The ARIMA model and the grey prediction model show that the global average annual temperature change has been stable in the past, while the prediction that the average temperature of 20.00°C in 2050 or 2100 is inaccurate. The grey prediction model shows that the global temperature levels will continue to rise in the future, and shows that the average temperature in 2,289 observations is expected to reach 20.00°C. According to the forecast, the global average temperature will not, as long as natural resources are properly used, reach 20.00°C. For problem 2, We first divide the Earth into the southern and northern hemisphere, and second into tropical and southern temperate and northern temperate zones. The linear relationship between global average temperature and the two conditions is discussed separately. According to the results, it significantly shows a strong correlation between the global average temperature and time. If unchecked, the global average temperature will grow over time. Second, using the principal component analysis method shows that may affect the global average temperature in the Middle East energy consumption, global hydropower consumption, world oil consumption, global niobium minerals, the world gas proven reserves, global crude oil production, global cadmium production, and to the results are tested, concluded that the principal component analysis method to influence the main indicators of the global average temperature method is more accurate. Therefore, the suggestions for global temperature change are made: to strive to reduce methane emissions; to further strengthen the strategic deployment and practical actions for energy revolution and low-carbon development, while increasing the intensity of energy low-carbon transformation. For problem 3, We propose the feasibility of this model evaluation and prediction results on slowing down or curbing global warming trends. One is direct emissions reduction, with countries signing contracts to force their carbon emissions and unite against global warming. Reduce heat emissions should be put on everyone's daily living habits, with a limited use of daily necessities, to achieve limited production, limited processing, so as to curb industrial flooding, reduce heat emissions. Second, the indirect emission reduction. With the development of industry, carbon dioxide emissions are constantly increasing, leading to a series of problems such as global warming and sea level rise. Afforestation can absorb the carbon dioxide contained in the atmosphere and reduce the greenhouse gas content in the atmosphere. Therefore, we must realize the importance of afforestation, take appropriate measures to improve the effect of afforestation, and better alleviate global warming.

## Keywords

**Prediction of Time Series; Gray Prediction Model; Global Warming; Correlation Factors.**

### 1. Introduction

Since the beginning of this year, there have been a large number of alarming temperature reports worldwide. Global warming is a phenomenon related to nature. This is due to the continuous accumulation of the greenhouse effect, resulting in an imbalance in the energy absorbed and released by the Earth's atmospheric system, and the continuous accumulation of energy in the Earth's atmospheric system, leading to temperature rise and global warming. Changes in the climate system arise from natural or internal operations and responses by external forces, both anthropogenic and non-human, such as solar activity, volcanic activity and greenhouse gases. Almost all climatologists have heard that the Earth has warmed in recent years. The causes of recent climate change are still the subject of active research, and the scientific consensus is that anthropogenic greenhouse gases are the main cause of global warming, which is not disputed in academia.

#### 1.1. Problem Background

Considering the analysis of the background information and the constraints of the problem statement, the following problems need to be specifically solved:

To address current reports and make reasonable projections of future global temperature levels, compare the rise in global temperatures in March 2022 with the rise over the past decade, based on the datasets provided, and analyze their causes; And establish two or more mathematical models to describe past temperature levels and make reasonable predictions of future temperature levels.

Using the established mathematical model, the global temperature in 2050 and 2100 is predicted respectively, and the average temperature of the global observation points at these two time points is observed to reach 20.00 degrees Celsius, if not, the average temperature in the observation model reaches 20.00 degrees Celsius, and the accuracy of the model is predicted.

Using known results and data, build a mathematical model to analyze the relationship between global temperature, time and location, if any, and interpret this relationship or prove that there is no relationship between them.

Collect data to analyze factors from natural disasters (e.g., volcanic eruptions, forest fires, COVID-19, etc.) and indicate their impact on global temperatures.

Analyze and derive the main causes of global temperature changes, and propose effective measures to curb or slow down global warming.

#### 1.2. Our Work

Two models, time series and grey prediction, are established to describe the past global temperature level and make reasonable predictions of future global temperature levels.

The prediction of the model is used to observe the time point when the average temperature of the global observation point reaches 20.00 degrees Celsius.

Analyze the known data collected and discuss its impact on global temperature to identify the main causes and propose effective containment or mitigation measures.

Prepare a one-page non-technical article for the APMCM Organizing Committee.

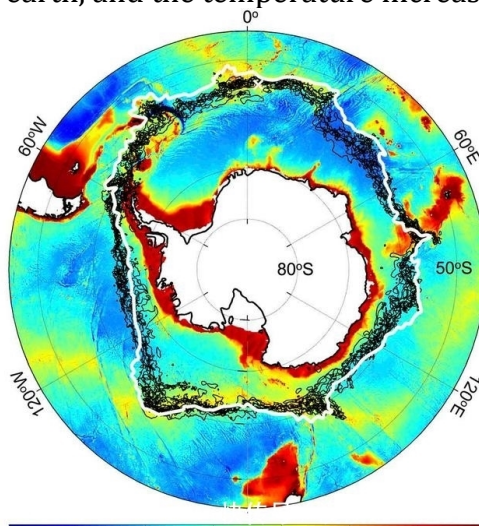
### 2. Assumption and Justifications

(1). Take the global temperature studied as a whole, as a system

- (2). Suppose that the data collected from NOAA is true and effective within the allowable error allowable range, and can reflect the situation of temperature change;
- (3). Suppose that global temperatures will not be affected by major natural disasters or extreme weather in recent years;
- (4). The influence of various micro influencing factors on temperature. The lag period is one year.
- (5). The same prediction assumptions and assumption measures are used for the prediction model

The impact of global climate change is entering "unknown areas of destruction" as countries fail to set adequate targets for reducing greenhouse gas emissions has warned. UN Secretary-General Antonio Guterres (Antonio Guterres) cited this year's floods in Pakistan, heat waves in Europe and record drought in parts of the United States and China as the result of a failure to combat climate change and contain fossil fuel production.

"It is not natural that the scale of these disasters is so large. They are the cost of a human addiction to fossil fuels," Guterres said in a statement. "This year's 'United in Science' report shows that climate impacts are moving into unknown territory. However, we are still consuming fossil fuels rapidly every year, even when the environment deteriorates rapidly." The picture below is the plan of the earth, and the temperature increasing as the eyes deepens.



**Figure 1.** Earth plan

Taking China as an example, the average temperature of the past five years has been collected, as shown in Table 1. Since 2017, the annual average temperature in China has been above 10°C, and the prevention and control measures have not achieved significant results, and the global warming is still a serious problem.

**Table 1.** Average temperature in China from 2017 to 2020

particular year	2017	2018	2019	2020	2021
temperature	10.6°C	10.1°C	10.34°C	10.25°C	10.5°C

### 3. Solution to Problem 1

#### 3.1. Step 1: Analyze Global Temperature Changes

Since the global average air temperature from March 2012 – February 2022 and March 2022 – October 2022 can be viewed as two independent samples and can be approximately considered to obey a normal distribution, see Figure 1.

In hypothesis testing, before using the independent sample t-test, three assumptions must be met: first, the observed value of each sample must be independent; second, the two populations of the sample from must be normally distributed; and third, the two populations of the sample from must have the same variance. In this data, the first two prerequisites for using the independent samples were met, and the analysis using SPSS yielded the independent sample t-test results, as shown in Table 1. In Levene's Test for Equality of Variances (see Figure 2), sig=0.171 > 0.05, indicates the variance of the two samples is homogeneous, indicating that the third prerequisite for the application of the independent sample t-test is also satisfied. Because the variance was homogeneous, available from the first row of Table 2, the significance level was sig=0.240 > 0.05, indicating that the test was not significant, with no significant difference in global mean temperatures between March 2012 – February 2022 and March 2022 – October 2022. The difference in global mean temperature may be caused by accidental random factors. Our generally assumed belief that the rise in global temperatures in March 2022 led to a larger rise than in the past 10 years is unscientific.

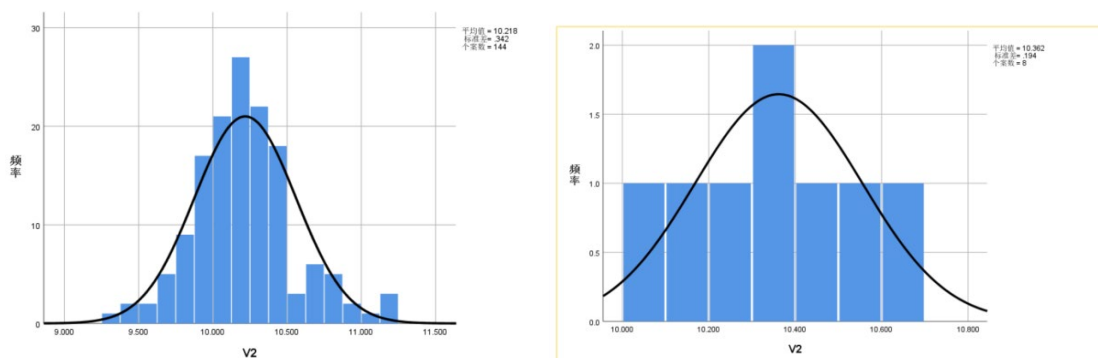


Figure 2. Two-sample normal distribution map

Table 2. Results of Homogeneity Test of Variance

	Levin Statistics	Df1	Df2	Sig.
Based on average	1.890	1	150	0.171
Based on median	1,832	1	150	0.178
Based on median and with adjusted degrees of freedom	1.832	1	146.202	0.178
Based on the average after pruning	1.857	1	150	0.175

Table 3. Table of the t-test results for independent samples

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

									Lower	Upper
total points	Equal variances assumed	1.890	0.171	-1.181	150	0.240	-0.144	0.122	-0.386	0.097
	Equal variances not assumed			-1.942	9.609	0.082	-0.144	0.074	-0.311	0.022

### 3.2. Step 2: Build Models to Describe the Past and Predict the Future

#### 3.2.1. ARIMA Model Introduction

(1) Diagnostic test

Diagnostic test of the ARIMA (2,1,0) model

Model significance test

For ARIMA (2,1,0), whether the residual is white noise sequence, according to the test results  $P > 0.05$ , the model residual sequence is white noise sequence, and the model is significant.

Coefficient significance test

**Table 4.** Coefficient significance test

Inspection parameters	t Statistic	P value	conclusion
mean $\mu$	1.1	0.027	Significantly non-zero
$\varphi_1$	-5.59	<0.0001	Significantly non-zero
$\varphi_2$	-2.45	0.016	Significantly non-zero

Based on the test results, the model parameters are significantly valid.

In summary, the model fit was performed well.

(2) Diagnostic test of the ARIMA (2,1,3) model

Model significance test

ARIMA (2,1,3) tests whether the residual is white noise sequence. According to the test result  $P > 0.05$ , the model residual sequence is white noise sequence and the model is significant.

Coefficient significance test

**Table 5.** Coefficient significance test

Inspection parameters	t Statistic	P value	conclusion
mean $\mu$	1.71	0.008	Significantly non-zero
$\varphi_1$	3.71	0.0003	Significantly non-zero
$\varphi_2$	-1.09	0.027	Significantly non-zero
$\theta_1$	-5.81	<0.0001	Significantly non-zero
$\theta_2$	1.8	0.043	Significantly non-zero
$\theta_3$	-0.432	0.016	Significantly non-zero

Based on the test results, the model parameters are significantly valid.

Model optimization

Using the optimal criterion function fixed order method, making the criterion function of the minimum model is the relatively optimal model, this paper uses AIC, AICc, BIC, where:

$$AIC(M) = -2 \ln(\text{maximum likelihood function}) + 2M$$

$$AICc = AIC + \frac{2M(M + 1)}{n - M - 1}$$

$$SBC(M) = -2 \ln(\text{maximum likelihood function}) + M \ln n$$

According to the results:

**Table 6.** Model optimization test

Model	AIC	AICc	SBC
ARIMA (2, 1, 0)	-47.5	-47.2	-36.3
ARIMA (2, 1, 3)	-67.4	-66.4	-47.8

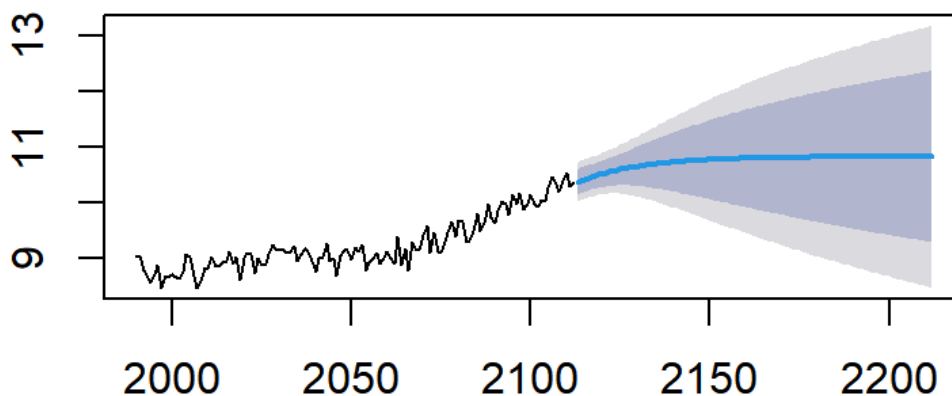
According to AIC, ARIMA (2,1,3) is optimal;

According to the AICc judgment, ARIMA (2,1,3) is optimal;

Judging by the SBC, ARIMA (2,1,3) is optimal.

In conclusion, we chose the ARIMA (2,1,3) model.

(3) Forecast



**Figure 3.** Future global trends

**Table 7.** The future global temperature levels

particular year	2023	2024	2025	2026	.....	2150	2151	2152	2153
Global average temperature	10.4	10.4	10.5	10.5	.....	10.7	10.7	10.7	10.7

According to the chart, it can be seen that the global average temperature will remain at 10.4-10.7 for a long time in the future. Meanwhile, the prediction that the average global observation temperature of 20.00 C in 2050 or 2100 is inaccurate. According to the test results, as long as the rational use of natural resources, the global average temperature will not reach 20.00 C to protect the environment.

(4) Advantages of the ARIMA model prediction

The modeling process of ARIMA model is relatively simple, clear and easy to understand. It has high accuracy for short-term prediction of global average temperature. Global average temperature prediction based on ARIMA model can provide valuable decision-making basis for decision makers.

### 3.2.2. The Grey Prediction Model

#### (1) Introduction to the grey prediction model

Grey prediction modeling is based on the grey module concept. Gray system theory believes that all random quantities are gray quantities and gray processes that change within a certain range and over a certain period of time. For the processing of gray quantity, it is not to seek its statistical law and probability distribution, but to find out the law from the irregular original data, that is, after processing the data through a certain way, make it becomes a more regular time series data, in the model. Its geometric meaning refers to the general term of the continuous curve given in the time and data 2 D plane of the generated sequence data and its bottom (i. e., the abscissa). Modules composed of known data columns, called white modules, while extrapolation by white modules to future modules, namely modules composed of predicted values, are called grey modules.

#### (2) Model Building

In the first step, the hierarchical ratio is tested and the feasibility of the modeling is analyzed. Set the time series to:

$$X_{(0)} = \{x_{(0)}^{(1)}, x_{(0)}^{(2)}, x_{(0)}^{(3)}, \dots, x_{(0)}^{(N)}\} \quad (1)$$

Where n represents having n observations. Then the level  $\sigma^{(k)}$  ratio is calculated and judged, only if all the model falls into the calculation range. The calculation and judgment formula of the grade ratio are respectively:

$$\sigma(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, \quad \sigma(k) \in \left( e^{-\frac{2}{n+1}}, e^{\frac{2}{n+1}} \right) \quad (2)$$

Through data processing, the original series  $X^{(0)}$  is accumulated once, and AGO (cumulative generation) is:

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

Therefore, the corresponding differential equation of the GM (1,1) model can be obtained as follows:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = \mu \quad (4)$$

Among them,  $a$  refers to developmental quotient,  $\mu$  refers to the amount of gray action,  $a$  and  $\mu$  Is the main parameter of the GM (1,1) grey prediction model, it be used to judge the prediction period, 1) When-  $a < 0.3$ , the model can make medium and long-term prediction; 2) when the short-term prediction is  $0.3 < -a < 0.5$ ; 3) when  $0.5 < -a < 1.0$ , the model needs to be corrected; 4) when-  $a > 1.0$ , the GM (1,1) model cannot be used for prediction.

In the second step, construct the data matrix B and the data vector Y, which are respectively:

$$B = \begin{pmatrix} -\frac{1}{2}|x^{(1)}(1) + x^{(1)}(2)| & 1 \\ -\frac{1}{2}|x^{(1)}(2) + x^{(1)}(3)| & 1 \\ \vdots & \vdots \\ -\frac{1}{2}|x^{(1)}(n-1) + x^{(1)}(n)| & 1 \end{pmatrix}$$

$$Y = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{pmatrix}$$

Set  $\hat{a}$  as the number of parameters to be estimated

$$\hat{a} = \begin{pmatrix} a \\ \mu \end{pmatrix}$$

It is obtained by using the least-squares method  $\hat{a} = (B^T B)^{-1} \times B^T \times Y$  Build the model and solve the generated values and the reduced values

According to the formula, the prediction model can be obtained:

$$\hat{X}^{(1)}(k+1) = \left( x^{(0)}(1) - \frac{\mu}{a} \right) e^{-ak} + \frac{\mu}{a} \quad (k = 0, 1, 2, \dots, n) \tag{5}$$

The reduction value can be obtained by the cumulative subtraction as:

$$\hat{\mathbf{x}}^{(0)} = \left( \hat{\mathbf{x}}^{(0)}(1), \hat{\mathbf{x}}^{(0)}(2), \dots, \hat{\mathbf{x}}^{(0)}(n-1) \right) \tag{6}$$

In the third step, the model is tested. The specific formula is:

The average relative error was at  $\varepsilon(\text{avg}) = \frac{1}{n-1} \sum_{k=2}^n |\varepsilon(k)|$  show;

The precision value to  $P^0 = (1 - \varepsilon(\text{avg})) \times 100\%$  represent:

The posterior difference ratio  $C = \frac{S_2}{S_1}$ , Where,  $S_1$  represents the variance of the residuals,  $S_2$  represents the variance of the residuals, The formula is:

$$S_1 = \sqrt{\frac{1}{N-1} \sum_{k=2}^N |E(k) - E|^2} \tag{7}$$

$$S_2 = \sqrt{\frac{1}{N} \sum_{k=1}^N |x^{(0)}(k) - \bar{X}|^2} \tag{8}$$

### (3)Fruit

We selected the annual average temperature data from 1973 to 2022 for import into the SPSS software, and first conducted the level ratio test to judge the suitability of the data sequence for model construction. The level ratio is the data of the last period / the current period.result display:.

**Table 8.** GM (1,1) Model-level ratio table

number	Original value	Grade ratio	Original value translation conversion	Ratio after conversion
1973	9.37	-	20.37	-
1974	8.88	1.055	19.88	1.025
1975	9.17	0.968	20.17	0.986
1976	8.78	1.044	19.78	1.02
1977	9.3	0.944	20.3	0.974
1978	9.13	1.019	20.13	1.008
...	...	...	...	...
2020	10.54	0.985	21.54	0.993
2021	10.28	1.025	21.28	1.012
2022	10.36	0.992	21.36	0.996

The raw data did not pass the level ratio test, so the translation transformation, that is, the translation conversion value of 11.00 is added to the original value, and the final translation conversion value is within the standard range interval [0.962,1.040], which means that this data is suitable for GM (1,1) model construction. The model building results are shown in the following table:

**Table 9.** Model building results

Coefficient of development	Grey action	Posterior error ratio <sup>C</sup>	Small error probability <sup>P</sup>
-0.001	19.994	0.1363	0.960

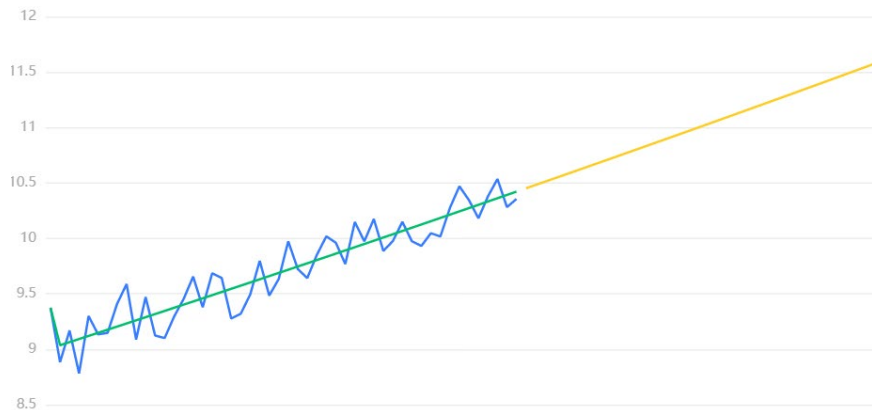
Where the posterior difference ratio <sup>C</sup> For the model accuracy level test, the smaller the value, the better. The general value is less than 0.35, and the model accuracy level is good. The value means that the model accuracy is qualified, and the value is less than 0.65 means that the model accuracy is basically qualified. If the value is greater than 0.65, the model accuracy level is unqualified; the small error probability <sup>P</sup> generally less than 0.7 means that the model is unqualified, less than 0.8 means that the model is barely qualified, less than 0.95 means that the model is qualified, and more than 0.95 means that the model accuracy is very good. The posterior difference ratio <sup>C</sup> 0.136 <=0.35, meaning that the model accuracy level is very good. In addition, the small error probability <sup>P</sup> 0.960 <1.0 means that the model accuracy is very good.

**Table 10.** The model fitting results

Index entry	Original value	Estimate	residual	relative error
1973	9.375	9.375	0	0
1974	8.884	9.036	-0.153	1.718
1975	9.169	9.064	0.105	1.141
1976	8.781	9.092	-0.311	3.541
1977	9.3	9.12	0.18	1.93
1978	9.135	9.148	-0.014	0.151
...	...	...	...	...
2019	10.379	10.333	0.046	0.445
2020	10.536	10.363	0.173	1.647
2021	10.283	10.393	-0.11	1.069
2022	10.356	10.422	-0.067	0.643

The above table shows the table of fitting results for the grey prediction model. The smaller the relative error value, the better. Generally, less than 20% indicates a good fit. The average model relative error is 1.446%, which means that the model fits well.

From the prediction results, the global temperature levels will continue to rise in the future, as shown in the chart below:



**Figure 4.** Prediction results

**Table 11.** Prediction results

Forecast order	Estimate
2023	10.452
2024	10.482
2025	10.512
2026	10.542
...	...
2060	11.588
2061	11.619
2062	11.651

The predicted guess from the prediction results shows that the average temperature of 20.00°C in 2050 or 2100 is incorrect, and according to the predicted trend, the predicted average temperature of 2289 observations will reach 20.00°C.

## 4. Solution to Problem 2

### 4.1 Temperature and Time (Correlation Analysis)

Step 1: It will be discussed in two modules

- (1) Divide the earth into the southern hemisphere and the northern hemisphere;
- (2) divided into tropical, southern temperate zone and northern temperate zone.

Step 2: The linear relationship between the global average air temperature and the above two conditions is discussed respectively:

For the Southern and Northern hemisphere:

One-way ANOVA was first performed and is presented in Table 13:

**Table 12.** ANOVA Table in the South and Northern Hemisphere

	Sum of squares	freedom	mean square	F	Significance
regression	94.806	2	47.403	272.830	0.000
residual	235.078	1353	0.174		
total	329.884	1355			

Subsequently, we performed a linear regression fit:

**Table 13.** The coefficient of the south and southern hemisphere

	Non standardized coefficient		Standardization coefficient	t	Significance	Collinearity statistics	
	B	Standard error	Beta			tolerance	VIF
constant	0.098	0.390		0.251	0.802		
Southern hemisphere average temperature	0.352	0.015	1.228	23.359	0.000	0.191	5.246
Northern Hemisphere average temperature	0.096	0.005	1.102	20.975	0.000	0.191	5.246

According to the table:  $y = 0.352x_1 + 0.096x_2 + 0.098$

For regional zones:

One-way ANOVA was performed first, as described in Table 15:

**Table 14.** Regional ANOVA

	Sum of squares	freedom	mean square	F	Significance
regression	93.504	3	31.168	178.268	0.000
residual	236.380	1352	0.175		
total	329.884	1355			

Subsequently, we performed a linear regression fit:

According to the table:  $y = -0.061x_1 + 0.357x_2 + 0.137x_3 + 3.402$

To sum up, we know that the global average temperature can be better represented by the above two conditions.

Step 3: Enter the time and discuss with the above two situations respectively:

(1)Time and the average temperature in the southern hemisphere and the northern hemisphere

**Table 15.** Table of regional zone coefficients

	Non standardized coefficient		Standardization coefficient	t	Significance	Collinearity statistics	
	B	Standard error	Beta			B	Standard error
constant	3.402	.321		10.604	.000		
Tropic average temperature	-.061	.016	-.138	-3.724	.000	.386	2.590
Southern hemisphere average temperature	.357	.018	2.225	20.147	.000	.043	23.019
Northern Hemisphere average temperature	.137	.008	2.216	17.882	.000	.035	28.963

**Table 16.** Correlation between the southern hemisphere and time

	Southern hemisphere	Northern Hemisphere
relevance	0.179	0.058
P value	<0.001	<0.001

According to the results, the mean air temperature in the southern and northern hemisphere is weakly correlated with the time, but its sig. <0.001 is significant, which indicates that the correlation is true.

**Table 17.** Correlation between tropical,...and time

	tropic	Southern hemisphere	Northern Hemisphere
relevance	0.254	0.093	0.046
P value	<0.001	0.001	<0.001

According to the results, the tropical correlation with time is stronger relative to the weak correlation term between the average air temperature and time in the tropical and southern temperate zones, and secondly, sig. <0.001, which is significant, indicating that the correlation is true.

In conclusion, the global mean temperature is discussed separately, and the correlation, although weak, exists.

(3)Time and the global average temperature:

**Table 18.** Correlation and time

	Global average temperature
relevance	0.061
P value	<0.001

According to the results, there is a strong correlation between global average temperature and time, second, sig. <0.001, significant, indicating that the correlation is true, and time and global

average temperature interact. If not controlled, the global average temperature will increase with time.

Step 4: Temperature and geographical location

(1) First, analyze the correlation between global temperature and geographical location in the southern hemisphere and the northern hemisphere, and verify the conclusion in the tropical and southern temperate and northern temperate zones divided according to latitude.

(2) Field difference analysis

**Table 19.** Field difference analysis

	Clustering category		F	P
	category 2(n=798)	category 1(n=570)		
Southern hemisphere average temperature	19.511±1.208	22.332±0.654	2570.084	0.0001
Northern Hemisphere average temperature	22.957±2.494	12.522±2.316	6174.777	0.0001

The above table shows the results of the quantitative field difference analysis, including results of the results of mean ± standard deviation, F test results, and significance P-value.

- The P-value for each analysis item was significant (P <0.05).
- If the is significant, the null hypothesis is rejected, indicating that there is a significant difference between the two groups of data. The difference can be analyzed according to the standard deviation of the mean ±, while it indicates that the data does not show a difference.

The results of the ANOVA show that: if the mean temperature of the southern hemisphere is 0.0001, the null hypothesis indicates that the mean temperature of the southern hemisphere is significantly different; the P value is 0.0001 and the null hypothesis is rejected, indicating that the mean temperature of the northern hemisphere varies significantly between the categories divided in the cluster analysis;

**Table 20.** Cluster summary

Clustering category	frequency	percentage %
Clustering category _1	570	41.667
Clustering category _2	798	58.333
total	1368	100.0

**Table 21.** Cluster the center point coordinates

Clustering category	Central value_ Average temperature in the southern hemisphere	Central value_ Northern Hemisphere average temperature
1	22.33	12.52
2	19.51	22.95

The results of cluster analysis showed that the cluster results were divided into 2 categories, with 570 frequency of cluster \_1,41.667%, and cluster \_2,798 frequency and 58.333%.

It can be concluded from the table above, that the average temperature of the region varies depending on the latitude, and so there is a correlation between the global temperature and the location.

We then verify this conclusion by analyzing the mean temperature clustering in three tropical, south temperate, and north temperate regions

**Table 22.** Field difference analysis

	Clustering category		F	P
	category 2(n=797)	category 1(n=571)		
Mean tropical temperature	25.215±0.747	23.747±0.994	972.714	0.0001
Average temperature of southern temperate zone	12.385±1.751	17.798±1.166	4140.823	0.0001
Average temperature of north temperate zone	20.857±3.867	6.33±2.962	5673.822	0.0001

For the variable tropical mean temperature, The significance P-value was 0.0001, Presence of significance at the level, Refreject the null hypothesis, It shows that the variable tropical mean temperature varies significantly between the categories divided by the cluster analysis; For the variable south temperate mean temperature, The significance P-value was 0.0001, Presence of significance at the level, Refreject the null hypothesis, It shows that the southern temperate mean temperature varies significantly between the categories divided by the cluster analysis; For the variable north temperate mean temperature, The significance P-value was 0.0001, Presence of significance at the level, Refreject the null hypothesis, It shows that the variable north temperate mean temperature varies significantly between the categories divided by the cluster analysis;

**Table 23.** Clustering summary

Clustering category	frequency	percentage %
Clustering category _1	571	41.74
Clustering category _2	797	58.26
total	1368	100.0

The results of cluster analysis showed that the cluster results were divided into 2 categories, with cluster category 1 frequency 571,41.74%, and cluster category 2 frequency 797 and 58.26%.

**Table 24.** Cluster the center point coordinates

Clustering category	Central value_ Mean tropical temperature	Central value_ Average temperature of southern temperate zone	Central value_ Average temperature of north temperate zone
1	23.74	17.79	6.33
2	25.21	12.38	20.85

From the table above, the average temperature varies between latitudes, so there is a correlation between global temperature and location.

Based on the data search, we collected 11 items of global temperature change, as shown in Table 26:

**Table 25.** Descriptive analysis of the relevant data

	minimum value	Maximum	mean value		Standard deviation	variance
			Statistics	Standard error		
CO2 emissions	358.96	396.74	377.1200	2.60567	11.65293	135.791
Global crude oil production	60520.30	72858.50	67425.41	940.96	4208.08	17707947.43
World oil consumption	3239.50	4185.10	3740.0760	66.34262	296.69323	88026.872
World proven natural gas reserves	136.72	208.44	167.9190	4.68249	20.94074	438.514
Global hydropower consumption	534.10	855.80	661.3400	21.61820	96.67953	9346.931
Energy consumption in the Middle East	331.71	785.30	532.6580	33.67204	150.58595	22676.127
Global cadmium production	15800	22200	19365.00	345.899	1546.907	2392921.053
Global output of niobium ore	15200	63400	38170.00	3959.779	17708.669	313596947.368
Global germanium production	44000	155000	86300.00	8126.014	36340.640	1320642105.263
Global titanium dioxide production capacity	3900000	7200000	5199000.00	228948.282	1023887.83	1048346315789.47
Global sponge titanium production	48000	209000	113665.00	12399.007	55450.047	3074707657.895

Global average air temperature is partly related to these factors listed in Table 1, but not every indicator has a significant impact on the global average air temperature. Too many indices will undoubtedly increase the difficulty and complexity of analyzing the analysis problem, and these multiple indicators may have some correlation. At this point, the principal component analysis method can be used to determine the main factors or indicators.

Principal Component Analysis (Principal Component Analysis) is a method often used when dealing with multiple variables. PCA can transform related index variables into unrelated index variables, convert more index variables into fewer index variables, and convert single meaning index variables into comprehensive index variables. Turning the related variables into the unrelated new variables with large variance can comprehensively reflect the main information contained by the original multiple variables; so the main influencing factors can be identified.

The principal component analysis and calculation steps are as follows:

Calculate the correlation coefficient matrix

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ M & M & M & M \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix} \tag{9}$$

$$r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}} \tag{10}$$

Because R is a real symmetric matrix, you just need to calculate its upper or lower triangle elements. Standardized processing of the raw data.

(2) Calculate the eigenvalues and eigenvectors

The eigenvalue is first found to solve the eigenequation  $\lambda_i$  (i=1, 2, ..., p), And make it in order of size,  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ ; The corresponding eigenvalues are then found separately  $\lambda_i$   $e_i$  (i=1, 2, ..., p). The eigenvalues are calculated from the correlation coefficient matrix and the contribution and cumulative contribution rates of each principal component.

**Table 26.** Eigenvalues and principal component contribution rate

component	total	Percent Variance	accumulate
1	8.649	86.487	86.487
2	0.696	6.965	93.452
3	0.241	2.409	95.861
4	0.176	1.755	97.616
5	0.125	1.250	98.866
6	0.061	0.612	99.478
7	0.036	0.356	99.835
8	0.011	0.113	99.948
9	0.005	0.046	99.994
10	0.001	0.006	100.000

(3) Calculate the principal component load:

$$p(z_k, x_i) = \sqrt{\gamma_k} e_{ki} \quad (i, k = 1, 2 \cdots p) \tag{11}$$

The principal component score can thus be further calculated:

$$z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1m} \\ z_{21} & z_{22} & \cdots & z_{2m} \\ M & M & M & M \\ z_{n1} & z_{n2} & \cdots & z_{nm} \end{bmatrix} \tag{12}$$

For eigenvalues  $\lambda_1 = 8.649$ ,  $\lambda_2 = 0.696$  The eigenvectors are found separately  $e_1$ ,  $e_2$ , Each variable was also calculated  $x_1, x_2, \dots, x_{17}$  Load on each principal component yields the principal component load matrix. All of the above steps are to apply the SPSS statistical analysis software for the principal component analysis, to find the feature root  $\lambda$ , the feature vector  $\mu$ ,  $E$  of the cumulative variance contribution rate and the principal component load matrix of the correlation matrix.

**Table 27.** Principal component load matrix

	component	
	1	2
Energy consumption in the Middle East	0.994	-0.038
Global hydropower consumption	0.976	0.123
World oil consumption	0.873	-0.131
Global sponge titanium production	0.363	-0.041
Global titanium dioxide production capacity	0.561	-0.069
Global germanium production	0.240	0.166
Global output of niobium ore	0.933	-0.133
World proven natural gas reserves	0.931	-0.268
Global crude oil production	0.915	-0.128
Global cadmium production	0.671	0.723

Table 27 shows that the first principal component of the No  $z_1$   $x_1, x_2, x_3, x_7, x_8, x_9$  Table 28 shows that the first principal component of the No  $z_2$   $x_{10}$  There is a large positive correlation.

The above analysis results show that using two principal components can replace these 10 indicators according to the contribution rate of the principal components, which greatly reduces the computational complexity. It can be seen according to the principal component load matrix  $x_1, x_2, x_3, x_7, x_8, x_9$  For the first principal component  $z_1$  exert a tremendous influence; enormous implications,  $x_{10}$  To the second principal component  $z_2$  exert a tremendous influence; enormous implications, So the indicators affecting these two principal components. The main indicators affecting the global average temperature are as follows: energy consumption in the Middle East, global hydropower consumption, global hydropower consumption, world oil consumption, global niobium minerals, world proven natural gas reserves, global crude oil production, and global cadmium production.

**Table 28.** Correlation coefficient with the global mean air temperature

	relevance	P value
Energy consumption in the Middle East	0.815	0.000
Global hydropower consumption	0.897	0.000
World oil consumption	0.832	0.000
Global sponge titanium production	0.708	0.000
Global titanium dioxide production capacity	0.611	0.000
Global germanium production	0.681	0.000
Global output of niobium ore	0.569	0.000
World proven natural gas reserves	0.874	0.000
Global crude oil production	0.915	0.000
Global cadmium production	0.871	0.000

Model evaluation: As can be seen from the above table, the correlation coefficient with the global average air temperature is greater than 0.8  $x_1, x_2, x_3, x_8, x_9, x_{10}$ . The data obtained by PCA is only one difference  $x_7$ , Other data indicators are consistent, indicating that the principal component analysis method to obtain the main indicators affecting the global average temperature is relatively accurate.

### 3. Main causes and countermeasures of global temperature changes

By analyzing the 10 indicators by the principal component analysis method, we conclude that the Middle East energy consumption, global hydropower consumption, world oil consumption, global niobium minerals, world proven natural gas reserves, global crude oil production, and global cadmium production, these indicators will have an impact on the global temperature changes.

First for natural gas, oil, coal resources mining will further strengthen the methane of carbon dioxide greenhouse gas emissions control work, improve the methane of carbon dioxide emissions monitoring and statistical system, strengthen the energy activities and industrial production process of methane of carbon dioxide greenhouse gas emissions control, in the coal resources and industry is more concentrated areas to carry out demonstration projects, efforts to reduce methane emissions;

Second for energy use to further strengthen the energy revolution and low carbon development strategy deployment and practical action, must control and reduce the total amount of consumption of fossil energy, reduce pollutant emissions from the source, and increase the intensity of the energy low carbon transformation, increase the proportion of non-fossil energy, mainly increase wind power, solar power and nuclear power supply.

## 5. Model Evaluation and Further Discussion

### 5.1. Advantages

Based on the basis of data collection, this paper selects 10 indicators related to global temperature and establishes a comprehensive evaluation system using principal component analysis. In this evaluation system, the causes of the global temperature rise are deeply analyzed, and the main causes are combined with the targeted treatment, so as to put forward effective suggestions.

The time series prediction model is comprehensive, starting from the collected start time, conducts the stability test, the prediction is practical, and has a high reference value.

Model 1 When using the correlation analysis, we combine the idea of clustering to consider more comprehensively.

### 5.2. Shortcomings and Improvements

This paper is limited by the amount of data obtained. If more effective data are obtained, the comprehensive evaluation system can be further improved to improve the analysis of the factors affecting the global temperature warming.

This model mainly considers dominant loss and does not explore implicit losses.

### 5.3. Model Promotion

This model can predict the future global temperature to increase the awareness of environmental protection. This model can provide feasible measures to the phenomenon of global temperature warming.

## References

- [1] Ding Song, Li Ruojin, Guo Junha. An entropy-based TOPSIS and optimized grey prediction model for spatiotemporal analysis in strategic emerging industry[J]. *Expert Systems With Applications*, 2023, 213(PC).
- [2] Li Peng, Xu Xiaosu, Yang Shirui, Jiang Xuefeng. Open circuit fault diagnosis strategy of PMSM drive system based on grey prediction theory for industrial robot[J]. *Energy Reports*, 2023, 9(S1).
- [3] Cusworth George, Brice Jeremy, Lorimer Jamie, Garnett Tara. When you wish upon a (GWP) star: Environmental governance and the reflexive performativity of global warming metrics.[J]. *Social studies of science*, 2022.
- [4] Sgubin Giovanni, Swingedouw Didier, Mignot Juliette, Gambetta Gregory Alan, Bois Benjamin, Loukos Harilaos, Noël Thomas, Pieri Philippe, García de Cortázar Aauri Iñaki, Ollat Nathalie, van Leeuwen Cornelis. Non-linear loss of suitable wine regions over Europe in response to increasing global warming.[J]. *Global change biology*, 2022.
- [5] Perazzolli Michele, Vicelli Bianca, Antonielli Livio, Longa Claudia M. O., Bozza Elisa, Bertini Laura, Caruso Carla, Pertot Ilaria. Simulated global warming affects endophytic bacterial and fungal communities of Antarctic pearlwort leaves and some bacterial isolates support plant growth at low temperatures[J]. *Scientific Reports*, 2022, 12(1).
- [6] Jenkins Katie, Kennedy-Asser Alan, Andrews Oliver, Lo Y T Eunice. Updated projections of UK heat-related mortality using policy-relevant global warming levels and socio-economic scenarios[J]. *Environmental Research Letters*, 2022, 17(11).
- [7] Yan An, Yang Yuqi, Jiang Xinyang, Mo Hanxiang, Duan Boya. A Study on Prediction and Scheduler for Foot Traffic of University Canteens Based on Gray Prediction and BP Neural Network[J]. *Academic Journal of Computing & Information Science*, 2022, 5(11).
- [8] Lindsly Stephen, Chen Can, Liu Sijia, Ronquist Scott, Dilworth Samuel, Perlman Michael, Rajapakse Indika. 4DNvestigator: time series genomic data analysis toolbox.[J]. *Nucleus (Austin, Tex.)*, 2021, 12 (1).
- [9] Savadkoobi M., Oladunni T., Thompson L.A.. Deep neural networks for human's fall-risk prediction using force-plate time series signal[J]. *Expert Systems With Applications*, 2021, 182.
- [10] Amigó José M., Dale Roberto, Tempesta Piergiulio. Complexity-based permutation entropies: From deterministic time series to white noise[J]. *Communications in Nonlinear Science and Numerical Simulation*, 2022, 105.