Team Number:	apmcm220****
Problem Chosen:	C

Summary

Global warming affects people of all countries. Especially since the Industrial Revolution, the greenhouse effect has become more and more serious. How to predict the future global temperature and analyze the factors affecting the change of global temperature have become an important topic.

Before the formal modeling, we need to preprocess the data, delete the data before the industrial revolution, further supplement and improve the data, and solve the global monthly mean temperature and annual mean temperature at the observation point.

For the first part, it is first demonstrated that the date is not a temperature abrupt point by using the **M-K test** for March 2022, **and that the increase of global temperature in March 2022 did not resulted in a larger increase than observed over any previous 10-year period.** The ARIMA model based on the global annual mean temperature and the LSTM model based on the monthly mean data of the northern and southern hemispheres have been established to describe and forecast the global temperature level. The global mean temperature predicted by the **two models in 2050 and 2100 is greater than 20°C.** Finally, by calculating the **RMSE value** of the two models, it is found that the RMSE of the LSTM neural network model is **0.054**, while that of the ARIMA model is **0.151**, indicating that **the LSTM neural network model is more accurate.**

For the second part, the normal distribution J-B test is conducted on global temperature, time and location respectively, and the null hypothesis is rejected. Spearman's correlation coefficient is used, and **it is found that global temperature has strong positive correlation with time and location.** Through the analysis of natural disasters, forest fires will affect the rise of global temperature, volcanic eruptions will produce a large amount of volcanic ash will lead to local temperature reduction, which will lead to global temperature increase in the long run, and the COVID-19 pandemic will affect the global temperature reduction. Through grey correlation analysis, it is found that **the main factor affecting global temperature change is the concentration of CO_2.**

In the third part, our team put forward our findings in the results and some suggestions on how to deal with global warming in the future.

Keywords: M-K test ARIMA Model LSTM neural network Grey correlation analysis

Contents

I. Introduction

1.1 Background

I believe you have personally felt the unusual temperature this summer, river flow, land drought, abnormal temperature, frequent high temperature warning in some areas, and even two months without rain phenomenon. In addition, due to the continuous high temperature influence, the power generation of hydropower stations has been seriously unable to maintain normal electricity, and large-scale power outages occur from time to time. Will such "extreme" weather continue to occur in the future ...?

CO_2 — the main contributor to the greenhouse effect, was around 280 parts per million before the Industrial Revolution, but with the use of coal, a CO_2 producing energy source, global warming has had to be taken seriously, with the largest 10-year average increase to date recorded in March 2002.

Therefore, how to analyze the current global temperature change and forecast the future global temperature level is the primary problem we should solve, so as to better cope with the future climate warming.

Figure 1 Lake Mead

Figure 2 Poyang Lake

1.2 Our work

Through the data given and collected by our team, we can build a model of temperature prediction and factors affecting temperature change, so as to analyze and prevent the greenhouse effect phenomenon. Our main work is as follows:

1. To analyze whether the rise in global temperatures in March 2022 resulted in a larger increase than was observed in any previous decade.

2. Develop two or more models to describe past and future global temperature levels.

3. See if each model agrees with the prediction that the global average temperature will reach $20°C$ by 2050 or 2100, and compare which model is the most accurate.

4. Establish a mathematical model to analyze the relationship among global temperature, time and location, and collect relevant data to analyze whether the factors of natural disasters have an impact on global temperature, analyze the main causes of global temperature change and some measures to restrain or slow down global warming.

5. Prepare an article explaining the team's findings and recommendations for the future.

II. Problem analysis

2.1 Data pre-processing

Due to the large amount of data given by the topic, it is necessary to analyze the data first. In this case, we delete the data before the Industrial Revolution because the data before the Industrial Revolution has little impact. Considering the large seasonal difference between the north and the south, the data were divided into the northern and southern hemispheres according to latitude and longitude, and then the monthly average temperature and the total annual average temperature were obtained respectively.

2.2 The first part of the analysis

- The M-K test was conducted for the northern and southern hemispheres using the monthly mean data, and the intersection of UF statistics and UB statistics was used to determine whether March 2022 was the abrupt temperature change point.
- For observation, the given data were data series. First, the ARIMA model was established by means of annual average temperature, and then the residual test was conducted on its white noise. Finally, the model was used to predict the data of the following 15 years. Considering the seasonal differences between the northern and southern hemispheres, the LSTM neural network model was used to train the monthly mean temperature of the northern and southern hemispheres respectively. Finally, the model was used to predict the data for the next 15 years to observe the trend.

- The model in the above question is used to predict the data in 2050 and 2100 respectively, and the prediction of whether the global mean temperature at the observation point in 2050 or 2100 will reach 20.00℃ is observed. Then calculate the RMSE value of the two models to determine which one is more accurate.

2.3 The second part of the analysis

- The relationship diagram between global temperature and time was drawn, and then the normal distribution J-B test was conducted, and the result was observed to reject the null hypothesis. Then the Spearman correlation coefficient was used to analyze the relationship between global temperature and time, and the relationship between global temperature and location respectively.
- By observing global temperature changes before and after volcanic eruptions, forest fires, and COVID-19, we can look at the impact of these natural disasters on temperature.
- Grey correlation analysis was used to determine the main causes of global temperature change, and recommended measures to curb or slow global warming.

III. Assumption of model

- Assume that the data given is true and valid and will not cause unnecessary errors
- Assume that only the effect of CO_2 on global temperature is considered in greenhouse gases
- It is assumed that per capita CO_2 concentration is representative of global warming

IV. Symbol Description

Symbol	Description	Unit
T_{gmt}	The global mean temperature	$^\circ C$
$sgn()$	Sign function of M-K test	
$Y(x_0(k), x_i(k))$	grey relational coefficient	

V. Temperature prediction for ARIMA and LSTM

5.1 Data preprocessing

5.1.1 *Data missing value analysis*

Will give the data using the python pandas libraries were analyzed, and found the AverageTemperature and AverageTemperatureUncertainty missing 11002 data, data can be filled in via the Weather Service website or by means of averages using data from the same day in different years.

Table 1 Number of missing data

AverageTemperature	AverageTemperatureUncertainty
11002	11002

The data given in the title is from 1743 to 2013. Considering that after the Industrial Revolution, with the frequent use of coal and oil, CO_2— the main greenhouse gas, was released in large quantities, during the industrial Revolution and its subsequent period, the global temperature rose rapidly and the rate of climate warming accelerated, especially after the second Industrial Revolution. The greenhouse effect of the Earth is becoming more pronounced, and before the Industrial Revolution, there was little difference in the amount of CO_2 gas, so the data before 1900 can be deleted.

5.1.2 *Data increase and processing*

By visualizing different latitude and longitude coordinates of the data set, it can be observed that the data are distributed on all continents, and the observation points are on land, where the temperature measured is higher than the global average temperature. Our team selected some land observation points and ocean observation points from the National Bureau of Statistics, which increased the reliability of the average temperature. Since there is a large difference between the northern and southern hemispheres in the same month of the year, monthly and annual mean temperatures are calculated for the northern and southern hemispheres respectively.

Figure 3 Visualization of different urban locations

5.2 Temperature change point test based on Mann-Kendall method

Mann-Kendall is a nonparametric statistical method. The variables of this method can not obey the normal distribution, the sample size of time series X is n. Where for all $i, j \leq n$, defining statistics φ:

$$\varphi = \sum_{i=1}^{n-1} \sum_{j=1+1}^{n} sgn(X_j - X_i) \tag{1}$$

Where X_i and X_j are observations corresponding to the i-th and j-th time series, and $i<j$, sgn() are symbolic functions , $Var(\varphi)$ is the variance of φ:

$$sgn(X_j - X_i) = \begin{cases} 1 & X_j - X_i>0 \\ 0 & X_j - X_i = 0 \\ -1 & X_j - X_i<0 \end{cases} \tag{2}$$

$$Var(\varphi) = \frac{n(n-1)(2n+5)}{18} \tag{3}$$

The standardized statistic Z is:

$$Z = \begin{cases} \frac{S-1}{\sqrt{Var(\varphi)}} & \text{S} >0 \\ 0 & \text{S} = 0 \\ \frac{S+1}{\sqrt{Var(\varphi)}} & \text{S} <0 \end{cases} \tag{4}$$

We divided the data given in the question into two categories: the southern hemisphere and the northern hemisphere. M-K test was carried out on the two types of data respectively. The graphs obtained through Matlab calculation are shown as follows:

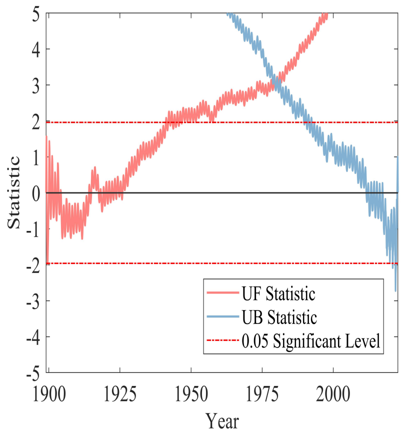


Figure 4 M-K test in the Northern Hemisphere

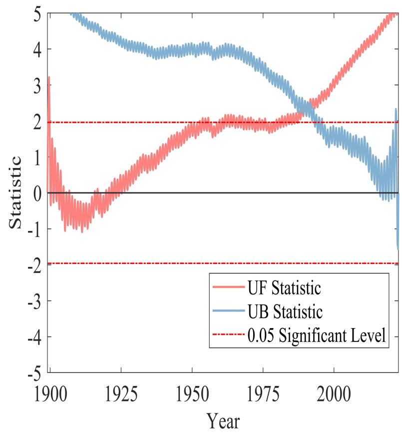


Figure 5 M-K test in the Southern Hemisphere

The coordinates of M-K intersection are in January 1980 in the Northern Hemisphere and the corresponding statistic is 3.112. In the Southern Hemisphere, it is March 1991, corresponding to a statistic of 2.31. The dates of both points are not attached to March 2022, so we can reject the null hypothesis that there is no significant upward or downward trend in the time series data, and that the increase of global temperature in March 2022 did not resulted in a larger increase than observed over any previous 10-year period.

Table 2　The intersection coordinates of the M-K test

	Northern Hemisphere	Southern Hemisphere
Date	In January 1980	In March 1991
Statistic	3.112	2.31

5.3 Projections of global temperature levels

5.3.1 *Global temperature level projections based on ARIMA*

ARIMA(Auto Regressive Integrated Moving Average) model[1] is a very effective model for predicting time series. ARIMA model is particularly suitable for fitting nonstationary data. Since the annual mean temperature data of observation points from 1900 to 2021 are selected, ARIMA can better model establishment and temperature prediction for the data, and in general, it can better describe the description and prediction of global temperature level. Where (p, d, q) is the non-seasonal part, and its form is as follows:

$$(1 - \sum_{i=1}^{p} \alpha_i L^i)(1 - L)^d y_t = \alpha_0 + (1 + \sum_{i=1}^{q} \beta_i L^i)\varepsilon_t \tag{5}$$

The global mean temperature is the mean of measured temperatures in the northern and southern hemispheres, A_north is the mean temperature in the Northern hemisphere, A_south is the mean temperature in the Southern hemisphere, and AT_{gmt} is the global mean temperature.

$$T_{gmt} = T_{north} + T_{sourth} \tag{6}$$

The expert modeler of SPSS was used to predict the annual mean temperature of the observation points in ARIMA time series. The outliers in the series were replaced by the mean value of the sequence. Through analysis, it can be seen that the ARIMA(0,1,2) model can better analyze and predict the time series. It can be seen that the autocorrelation coefficient and partial autocorrelation coefficient of all lag orders are not significantly different from 0, and the P-value obtained by Q test of the residual is 0.746, that is, we cannot reject the null hypothesis and consider the residual as white noise. Therefore, the ARIMA(0,1,2) model can well identify the temperature data, and the predicted temperature data is reliable.

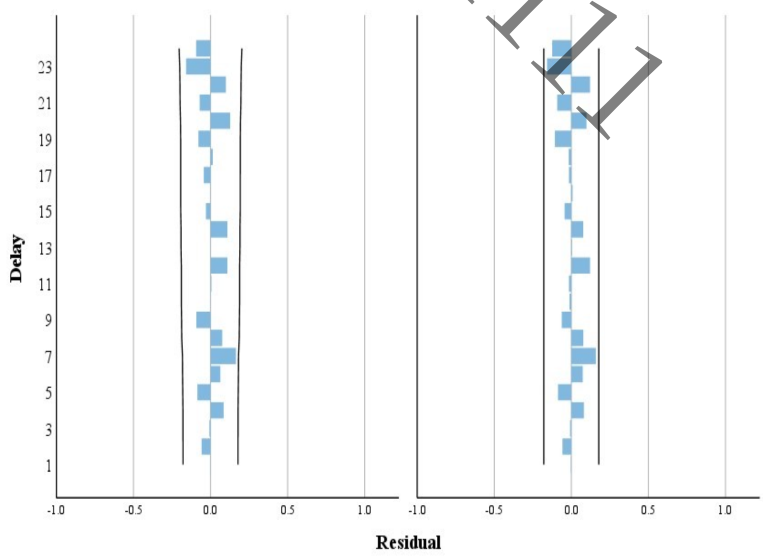


Figure 6 ACF(left)–PACF(right)

The ARIME model predicts the global mean temperature for the next 15 years after 2021 and plots it as follows:

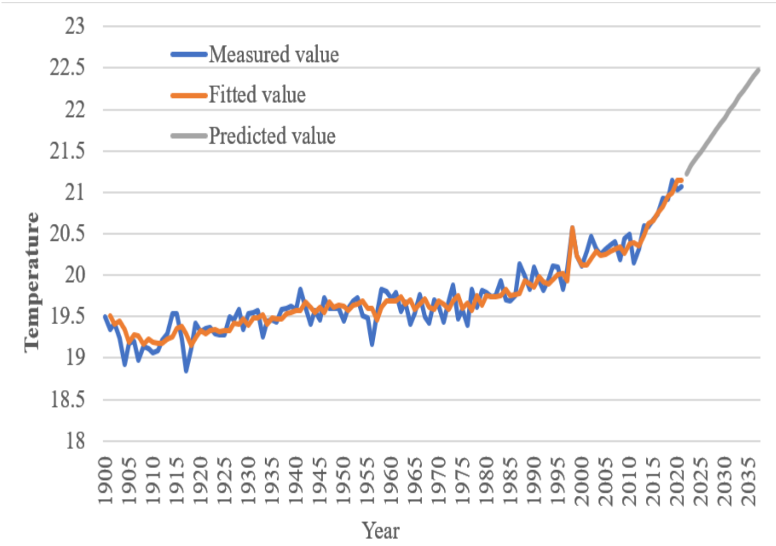


Figure 7 Time series diagram and temperature prediction

5.3.2 *Global temperature level projections based on LSTM*

LSTM(Long short-term memory) is a recursive neural network[2] that can learn long-term dependence and is suitable for extracting time series features from time series. It has great applicability in time series prediction and strong model generalization[3] ability. The following figure is the implementation schematic diagram of the method:

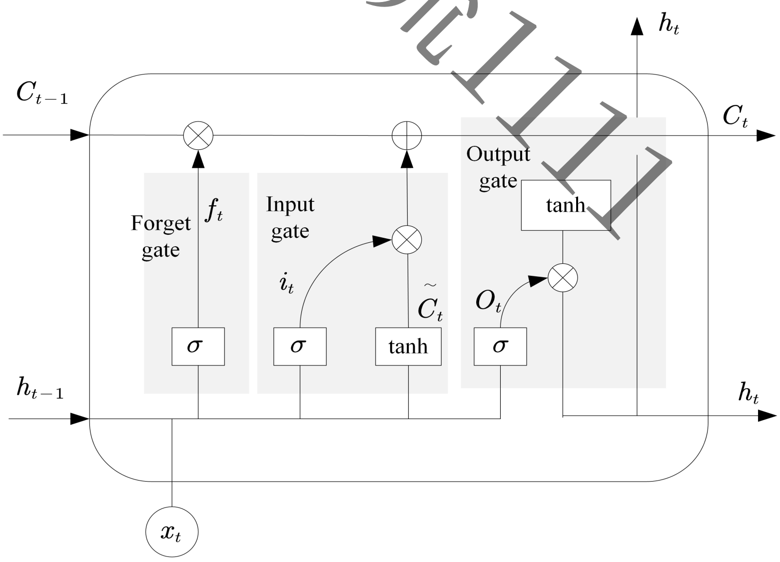


Figure 8 Cell structure of LSTM

LSTM was used to train the monthly mean temperature data of the northern and southern hemispheres from January 1900 to October 2022, and then the trained model was used to predict the 15 years after 2022. The predicted general graph is as follows:

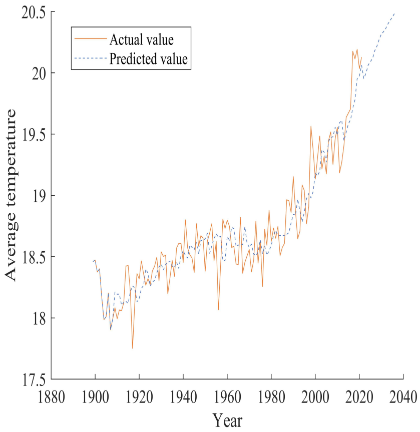


Figure 9 The Northern Hemisphere

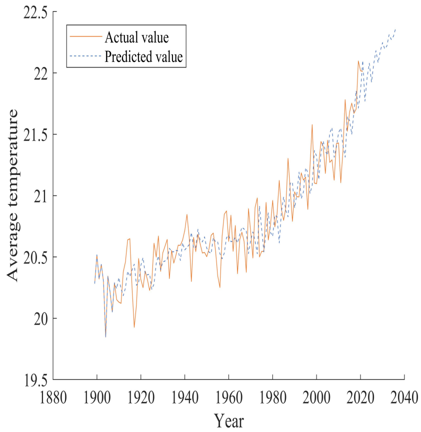


Figure 10 The Southern Hemisphere

5.4 Projections of global mean temperatures in 2050 or 2100

The data of 2050 and 2100 are substituted into the established ARIMA model and LSTM model for prediction, and the results are shown in the following table:

Table 3 Global temperature projections for 2050 and 2100

Year	ARIMA	LSTM	
		Northern Hemisphere	Southern Hemisphere
2050	25.37°C	20.72°C	22.47°C
2100	26.72°C	20.84°C	22.51°C

After screening the data on the topic, it is found that the average temperature of the global observation point has been more than 20°C as early as 1987. It can be seen from the above table that the global average temperature of the two models we established in 2050 and 2100 both exceeded 20°C. According to the predicted data, it can be observed that the general trend of global temperature will be higher and higher. The results predicted by the LSTM model are all smaller than those predicted by the ARIMA model. For the results predicted by the LSTM model, the average temperature of the northern Hemisphere is lower than that of the southern Hemisphere. Both of our models predict global mean temperatures exceeding 20°C in 2050 and 2100.

5.5 Comparison of models

For the comparison between the established ARIMA model and LSTM model, RMSE(Root Mean Square Error) was used for discrimination. The dimension of the evaluation index was the same as the original data. The real and predicted values of each model could be substituted into the formula to calculate the results.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{7}$$

RMSE was performed on the real and predicted values of the two models from 1900 to 2021, and the results of the two models were calculated as follows:

Table 4 RMSE values for ARIMA and LSTM

ARIMA	LSTM
0.151	0.054

By comparing the RMSE values of the two models, we can see that $RMSE_{LSTM}$ $<RMSE_{ARIMA}$, so we believe that the LSTM model is more accurate and the predicted results are more convincing.

VI. Analysis of influencing factors of Spearman and Grey correlation analysis

6.1 Correlation analysis of Spearman coefficient

According to the observation points given in the topic and the data found by oneself, the change of global mean temperature with each year can be made by combining the data predicted by oneself in the 15 years from 2022 to 2037. The figure is shown as follows:

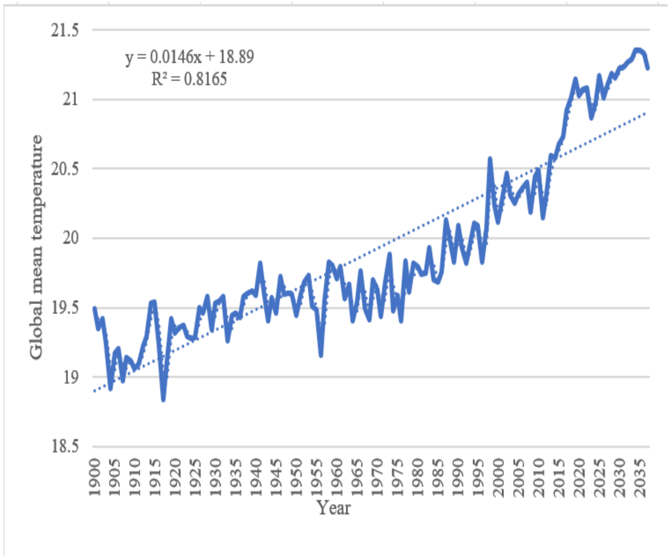


Figure 11 Global mean temperature in relation to time

Through the image, it can be observed that there is a relatively obvious linear relationship between temperature and year. The R-squared result obtained by fitting this line segment is 0.8165, which can be considered to have a strong linear relationship. The annual average temperature sample number is 138 > 30, belonging to a large sample, so we conducted the normal distribution JB test on it and constructed the JB statistic:

$$JB = \frac{n}{6}[S^2 + \frac{(K+3)^2}{4}] \tag{8}$$

JB test was conducted by Matlab's jbtest function, and the p value was 0.0042<0.01, that is, we rejected the null hypothesis at the 99% confidence level, that is, global mean temperature data did not follow the normal distribution, and Spearman correlation coefficient was used in this case, the time and temperature are converted into the form of grades for dimensionless processing, where n is the number of samples and a_i is the grade difference between temperature and time in the same sample.

$$r_s = 1 - \frac{6\sum_{i=1}^{n} a_i^2}{n(n^2-1)} \tag{9}$$

Spss was used to calculate the Spearman correlation coefficient, and the Spearman correlation value was 0.897, $p < 0.001$, that is, the null hypothesis was rejected at the 99% confidence interval. Combined with the graph analysis, it can be considered that the global mean temperature and time have a strong positive correlation.

As for the correlation between global temperature and geographical location, it can be known from consulting the data that temperature has nothing to do with longitude, but mainly has to do with latitude. Therefore, in this question, according to the latitude given by the topic, each city is divided into tropical, South temperate and North temperate zones. Since the data in the cold zone is relatively extreme, the data in the cold zone is not considered in this question. Three different positions were analyzed respectively to obtain the temperature line diagram:

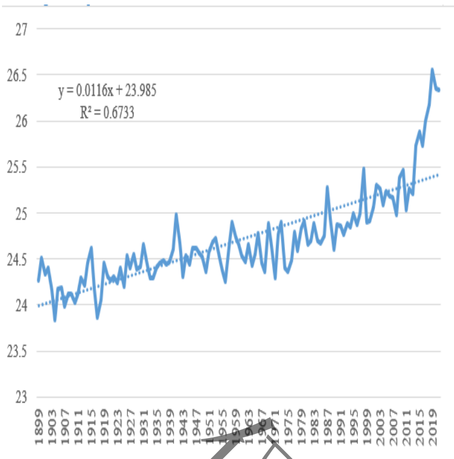


Figure 12 Tropic

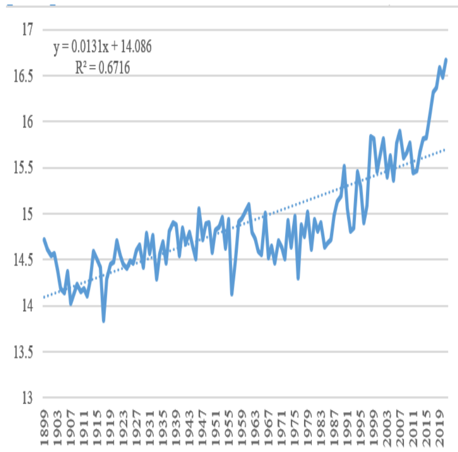


Figure 13 North temperate zone

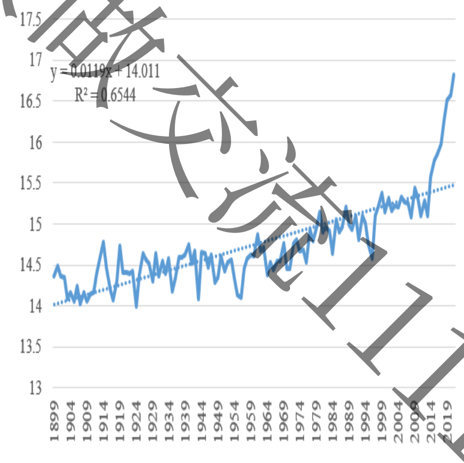


Figure 14 South temperate zone

As can be seen from the figure above, it is not obvious that there is a certain linear relationship for the three different positions. Spearman correlation coefficient is used to explain it. The results are as follows:

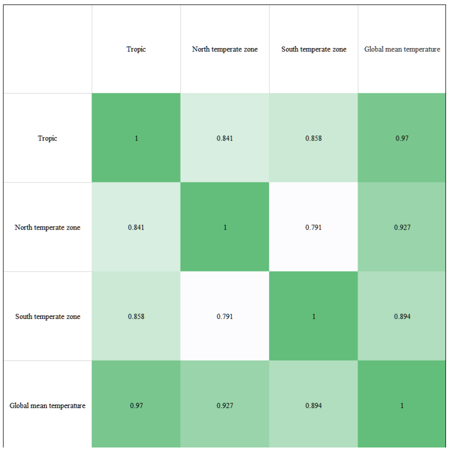


Figure 15 The result of Spearman correlation coefficient

The comparative analysis of the results shows that there is a strong correlation between global temperature and geographical location under the 99% confidence interval, and there is a strong positive correlation between global temperature and geographical location. There is also a strong positive correlation between global temperature and time.

6.2 Analysis of factors affecting temperature change

This paper aims to study the impact of volcanic eruptions, forest fires and COVID-19 on global temperature. By studying the impact of natural disasters on temperature in the most affected countries, this paper analyzes the temperature changes in the past ten years.

6.2.1 *Based on the effect of forest fires on temperature*

This paper analyzes the temperature change by studying the Australian fires (2019.9-2020.2). The data of September, 10,11,12, and 1,2 in recent ten years are averaged. This paper first obtains the average temperature of the month without the fire, and then obtains the data of the month with the fire, as shown in the following table:

Table 5 Whether to consider the average monthly temperature of the fire

	Failure to consider fire	Consider the fire
September	17.80	18.03
October	20.50	20.88
November	23.20	23.49
December	24.32	24.85
January	25.98	26.25
February	25.52	25.76

As can be seen from the above table, the impact of forest fires on the average temperature without forest fires is 0.32°C, so the regional temperature will rise by 0.32°C degrees after forest fires occur. Therefore, it can be considered that forest fires have a significant impact on global temperature.

6.2.2 *Based on the effects of volcanic eruptions on temperature*

This paper analyzes the temperature change by studying the eruption of Krakatoa Volcano in Indonesia. Since the eruption will produce a large amount of volcanic ash, the atmosphere will be damaged and a large amount of sunlight will be reflected, resulting in a small amount of sunlight received by the ground, which will lead to a decrease in temperature. By referring to the data, it can be known that Krakatoa volcano had a major volcanic eruption in 1883. The outbreak caused global temperatures to drop an average of 0.6°C over five years. Based on the data given in the title, the data of Indonesia was analyzed, and the average temperature of two Indonesian cities from 1874-1888 was selected as the average temperature of Indonesia as the basis for analysis. The average temperature is shown in the figure below:

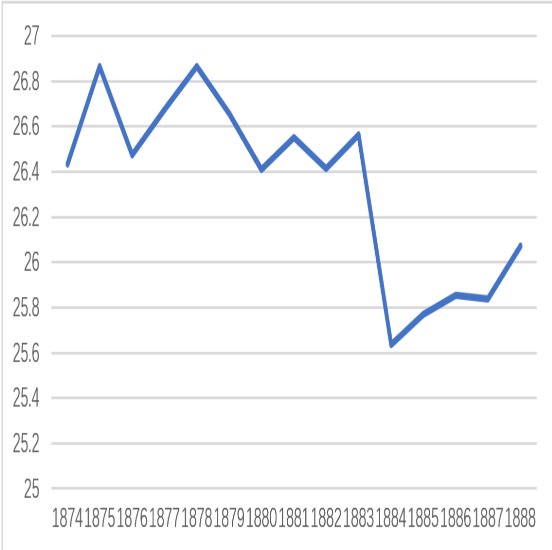


Figure 16 Mean temperature in Indonesia from 1874 to 1888

By comparing and analyzing the average temperature between 1874-1883 and 1874-1888, the average temperature between 1874-1883 and 1874-1888 was 26.59°C and 26.33°C respectively. After the eruption of the volcano, the average temperature in Indonesia dropped by 0.26°C. Therefore, it can be considered that the volcanic eruption has a significant impact on global temperature. The reason for the temperature drop may be a large amount of volcanic ash, which affects the sunlight, so the local temperature will be lowered, but for the world, it will produce a large amount of CO_2.

6.2.3 *Based on the impact of COVID-19 on temperature*

In this paper, temperature changes are analyzed by studying the impacts in China and the United States (December 2019.12 to present). Since COVID-19 started in December 2019, data from 2020 and 2021 are used as observation points.

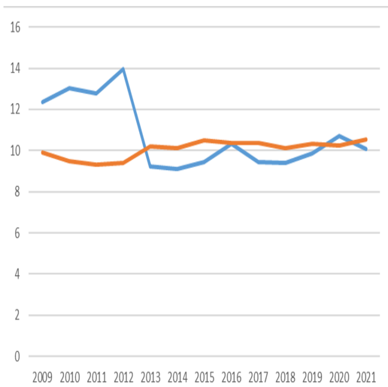


Figure 17 The average temperature of the last ten years in China and the United States

It can be analyzed from the figure that under the influence of COVID-19, major countries, including China and the United States, have shown a downward trend in temperature, which is mainly caused by the impact of COVID-19, which leads to people's home isolation, fewer cars on the highway and less exhaust emissions from factories.

6.3 Grey correlation analysis of temperature change factors

Starting from carbon dioxide concentration, population and forest area, the main causes of global warming are analyzed. This paper conducts a model through grey correlation analysis. In the collected data, global temperature is taken as the parent sequence, carbon dioxide concentration, population and global forest area as the sub-sequence, and the relationship between global temperature and the three sub-sequences is discussed respectively[4]. Solve the difference of the parent sequence corresponding to each sample of the subsequence and find the maximum and minimum difference:

$$z = \min_i \min_k |x_0(j) - x_i(j)| \tag{10}$$

$$m = \max_i \max_k |x_0(j) - x_i(j)| \tag{11}$$

To solve the grey correlation coefficient, where $\rho = 0.5$:

$$Y(x_0(k), x_i(k)) = \frac{z + \rho m}{|x_0(k) - x_i(k) + \rho m|} \tag{12}$$

Compared with the final value of the grey correlation degree of the three sub-sequences, the greater the grey correlation degree, the higher the correlation degree with the parent sequence. matlab was used to analyze and process the data, and the following results were obtained by means of data normalization:

Table 6 The results of grey correlation analysis

	Grey relational degree	Ranking
CO_2 concentrations	0.813	1
Global forest area	0.805	2
Population	0.537	3

It can be seen from the table that the grey correlation degree between carbon dioxide concentration and global temperature is the highest, which is 0.813. Therefore, this paper believes that the main reason affecting the change of global temperature is carbon dioxide concentration. The higher the concentration of carbon dioxide, the greater the change in global temperatures.

6.4 Our Measures

Our team proposed the following two measures to mitigate global warming:

1. Cut emissions directly. Strengthen international cooperation among countries and call for countries to sign a contract to limit their carbon emissions in order to jointly tackle global warming. We can also call on governments of all countries to issue relevant policies, strengthen the publicity of energy conservation and emission reduction, raise the awareness of the whole people to save energy and emission reduction, start from themselves, adopt a green lifestyle, and reduce heat emissions. In addition, existing industrial enterprises can be reformed to fully mobilize their enthusiasm for energy conservation and emission reduction. Replace direct emission equipment with zero-carbon or low-carbon equipment, reduce greenhouse gas emissions in the production process, and avoid excessive burning of mineral energy to help the national economy transition to low-carbon.

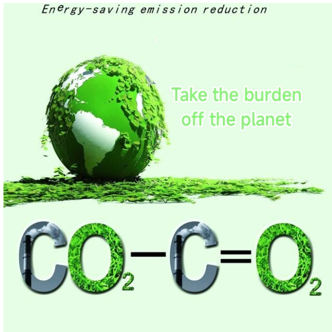
2. Indirect emission reduction. We will vigorously carry out afforestation efforts to increase revenue and reduce expenditure. Through afforestation to improve vegetation coverage to achieve the purpose of expanding carbon storage. We all know that global warming is caused by excessive emissions of carbon dioxide and other greenhouse gases. Through the photosynthesis of vegetation, it can absorb a lot of carbon dioxide, reduce the content of carbon dioxide in the atmosphere, and effectively mitigate global warming. We will accelerate the development of clean and renewable energy, improve the energy mix and develop alternative energy sources. We will set more active targets for new energy development, promote onshore power generation and photovoltaic power generation, develop hydropower in light of local conditions, and vigorously increase the capacity to absorb new energy. We will adhere to and improve the dual control of energy consumption, accelerate the application of low-carbon technologies, and improve energy efficiency.

VII. A letter to the APMCM Organizing Committee

Deer APMCM Organizing Committee,

Our team presents some of our findings on global warming and our views on mitigating the greenhouse effect in the future.

By analyzing the data given by the topic and the data we found ourselves, our team found that global warming is an irresistible fact. The rise of global temperature is closely related to time and location. According to our research, the phenomenon of global warming has become a matter of great concern to all mankind in the future. The greenhouse gases led by CO_2 are the main culprit of the greenhouse effect. The analysis and prediction of global CO_2 can be regarded as a good way to restrain or slow down the greenhouse effect.



For the future, Control the emission of waste gas and waste heat. The greenhouse gases that cause climate warming are released when humans burn fossil fuels. Curbing climate warming requires reducing emissions of these gases. For example, by improving production equipment, reducing industrial use of $CFCs$ and fossil fuels and increasing the frequency of green travel. Second, vigorously develop and utilize new energy sources. With the rapid development of science and technology, we should not just focus on the use of coal and oil. We should take a long-term view and increase the development of pollution-free energy. Such as the wide distribution of solar energy, clean and sanitary wind energy, Marine energy and so on. Finally, protect forests to strengthen greening. The Nemesis of the "greenhouse effect" should be forest vegetation, because forest vegetation can absorb large amounts of carbon dioxide through photosynthesis. To stop deforestation and increase forest coverage is a strategic measure to control the harm of "greenhouse effect". Afforestation is a good way to mitigate climate change. Many developed countries have taken indirect measures to reduce emissions from forests.

I believe that the greenhouse effect can be better solved in the future, and the global temperature will not change too much. After all, the earth is the common home of mankind, and everyone in the "global village" needs to jointly take care of it.

Yours Sincerely,

apmcm2203963

VIII. References

[1] Yao Zhaoming, PENG Shanghai. Based on neural network time series analysis of the temperature field prediction [J]. Journal of henan institute of urban construction, 2022, 31 (3) : 39-44. DOI: 10.14140 / j.carol carroll nki HNCJXB. 2022.03.007.

[2] Chen Bin, Liu Yue, Yin Kailang, Fang Xun. Joint prediction of runway temperature Data and Mechanism based on LSTM in Snow and ice weather [J/OL]. Journal of Beijing University of Aeronautics and Astronautics :1-16[2022-11-28. DOI:10.13700/j.bh.1001-5965.2022.0579. page number, year of publication.

[3] Zhang Weidong. Prediction Method of Axle box Temperature of Subway Vehicles based on LSTM-BPNN-SVR [J]. Computer and Digital Engineering, 2002,50(09):2079-2082+2108.

[4] Chen Yun, Wang Shimei. Study on the Correlation between Higher Vocational Education and Rural Revitalization Development in Fujian Province – Based on the grey correlation analysis of combination empowerment [J]. Tonghua normal university journal, the lancet 2022 (11):132-137.The DOI: 10.13877 / j.carol carroll nki cn22-1284.2022.11.023.

IX. Appendix

Listing 1: Latitude and longitude visualization

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from matplotlib import font_manager
import folium
from tqdm import tqdm

data = pd.read_csv('2022_APMCM_C_Data.csv', encoding='gbk')
data.isnull().sum()
#Combined longitude and latitude
for row in tqdm(range(data.shape[0])):
    tem = []
    if data.loc[row, 'Latitude'][-1] == 'N':
        tem.append(eval(data.loc[row, 'Latitude'][:-1]))
    else:
        tem.append(-eval(data.loc[row, 'Latitude'][:-1]))

    if data.loc[row, 'Longitude'][-1] == 'E':
        tem.append(eval(data.loc[row, 'Longitude'][:-1]))
    else:
        tem.append(-eval(data.loc[row, 'Longitude'][:-1]))
    # print(tem)
    data.loc[row, 'posistion'] = str(tem)
data.to_excel('clean.xlsx', index=None)
posistion_dict = {k: v for k, v in data.groupby('posistion')}
[eval(i) for i in posistion_dict.keys()]
map = folium.Map(location=[31, 120.5], zoom_start=8,min_zoom=2,
    tiles='Stamen Watercolor')
for i in tqdm([eval(i) for i in posistion_dict.keys()]):
    folium.Circle(radius=50, location=i,color='blue').add_to(map)
map.save('map.html')
```

Listing 2: M-K Test

```
%% M-K test
A=xlsread("Mean monthly temperature.csv");
FontSize = 16;
%define picture color
color=[[254 129 125]/255;[130 176 210]/255];
x = zeros(size(A,1),1);
for i = 1:size(A,1)
   x(i) = datenum(A(i,1),A(i,2),1);
end
titleName = {'Northern hemisphere mean temperature','Southern
   hemisphere mean temperature'};
for j=1:2
   y=A(:,2 + j);
   time_series=y;
   n=size(A,1);
   UF=zeros(size(time_series));
   E = n*(n-1)/4;
   Var = n*(n-1)*(2*n+5)/72;
   r1 = zeros(1,n);
   for i= 1:n
      r1(i) = sum(time_series(i)>time_series(1:i));
   end
   s = zeros(size(time_series));
   for k = 2:n
      s(k) = sum(r1(1:k));
      E = k*(k-1)/4;
      Var = k*(k-1)*(2*k+5)/72;
      UF(k) = (s(k)-E)/sqrt(Var);
   end
   time_series2 = zeros(1,n);
   for i=1:n
      time_series2(i)=time_series(n-i+1);
   end
   UB = zeros(size(time_series2));
   r2 = zeros(1,n);
   for i= 1:n
```

```
        r2(i) = sum(time_series2(i)>time_series2(1:i));
    end
    s2 = zeros(size(time_series2));
    for k = 2:n
        s2(k) = sum(r2(1:k));

        E = k*(k-1)/4;
        Var = k*(k-1)*(2*k+5)/72;
        UB(k) = -(s2(k)-E)/sqrt(Var);
    end

   %plotting
    ub_1 = zeros(1,n);
    for i=1:n
        ub_1(i)=UB(n-i+1);
    end
    figure(j)
    plot(x,UF,'color',color(1,:),'linewidth',1.5);
    hold on
    plot(x,ub_1,'color',color(2,:),'linewidth',1.5);
    plot(x,1.96*ones(n,1),'-.r','linewidth',1);
    plot(x,0*ones(n,1),'-','color',[0.2,0.2,0.2],'linewidth',1);
    plot(x,-1.96*ones(n,1),'-.r','linewidth',1);
    datetick('x','yyyy')
    axis([min(x),max(x),-5,5]);
    legend('UF Statistic','UB Statistic','0.05 Significant Level');
    set(gca,'Fontsize',12)
    set(gca,'ytick',-5:1:5)
    set(gca,'FontName','Times New Roman','FontSize',fontSize)
    xlabel('Year','FontSize',fontSize);
    ylabel('Statistic','Fontsize',fontSize);
end
```

Listing 3: JB test was used to determine the normal distribution

```
[h,p] = jbtest(Test(:,1),0.05)
n_c = size(Test,2)
H = zeros(1,6);
```

```
P = zeros(1,6);
for i = 1:n_c
[h,p] = jbtest(Test(:,i),0.05);
H(i)=h;
P(i)=p;
end
disp(H)
disp(P)
```