ISSN: 2790-5209 Vol 2, No.4, 2024

Global Warming OR Not?

Haoyang Tu

Millbrook school

15618517686@163.com

Abstract

Recent years have seen rising Earth's temperatures and a shrinking Antarctic land area. Studying factors affecting global warming and finding solutions to slow it is crucial. Using the Pettitt mutation point detection method, we examined March temperatures over the past decade and identified a significant change in March 2015. Predictive models (ARIMA and LSTM) forecast similar future temperatures, with the LSTM model showing slightly higher accuracy. Gray correlation and multiple linear regression analyses identified CO2 as the primary factor influencing global temperatures. Reducing CO2 emissions is essential to curb global warming.

Keywords

Global Warming, LSTM, ARIMA, The Gray Correlation Analysis, The Multiple Linear Regression Model, the Pearson correlation, Pettitt Mutation Detection.

1. Introduction

1.1. Background

Global warming, a long-term rise in the average temperature of the Earth's climate system, has become a critical challenge for our generation. This phenomenon is largely attributed to increased concentrations of greenhouse gases produced by human activities such as deforestation, fossil fuel burning, and industrial processes. In recent years, discussions and studies on global warming have triggered different views and interpretations in the scientific community. As greenhouse gas emissions continue to rise, researchers warn that the effects of global warming may become more pronounced, such as an increase in the number of extreme heat events, which have been reported frequently in recent years, and the possibility of even more dramatic temperature rises predicted for 2022 and beyond. According to the IPCC report and studies by organizations such as Berkeley Earth, global average temperatures are expected to rise by about1.5 to 2 degrees Celsius by 2050, and could still rise by about 3.5 degrees Celsius by 2100.

The IPCC Special Report on Global Warming of 1.5°C emphasizes the critical need for rapid and far-reaching transitions across energy, land, urban, and industrial systems to limit global warming to 1.5°C above pre-industrial levels. Achieving this target requires a substantial reduction in CO2 emissions, approximately 45% from 2010 levels by 2030 and reaching net zero around 2050. Predicting the future state of global warming is not only scientifically challenging but also crucial for policy-making decisions and future planning. The significance

of accurate model predictions can help to incorporate a better understanding of future climate scenarios, policy formulation, and decision making, and raising public awareness.

1.2. Research objective

The research objectives of this paper include the following three points:

(1) To analyze global temperature trends between 2013 and 2022 to determine whether 2022 experienced a more significant temperature increase compared to other years in that time period.

(2) Construct global temperature projection models to explore the levels to which global average temperatures are likely to rise by 2050 and 2100 in the absence of any climate governance measures, and to project when global temperatures will reach 20 degrees Celsius.

(3) To study the various relevant factors affecting global temperature fluctuations and to explore their interrelationships.

1.3. Our work

A graphical summary of the research methodology is shown in Figure 1.

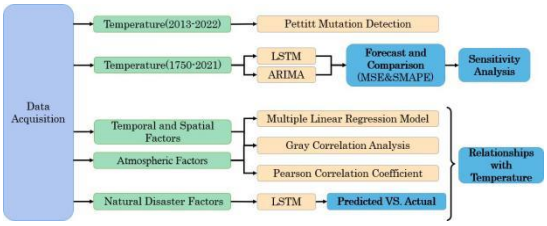


Figure 1: Our work

We first used the Pettitt mutation detection method in order to analyze whether there is a significant temperature mutation in 2022 relative to the past decade. Subsequently, we built an ARIMA model and an LSTM model, respectively, which are designed to make projections of future global land-averaged temperatures. We then used these two models to predict global temperature values for 2050 and 2100. To compare the performance of the two models, we train the models on known data and calculate the MSE and SMAPE of the model predictions for comparison. Additionally, we performed a sensitivity analysis of the models. Finally, we explored the factors affecting global temperature in three dimensions: spatio-temporal factors, atmospheric composition factors and natural disaster factors. In studying the influence of spatial and temporal factors on global temperature, we used multiple linear regression and gray correlation analysis to reveal the intrinsic relationship between longitude, latitude and time series and global temperature. For atmospheric factors, we use gray correlation analysis and Pearson's correlation coefficient to explore the correlation between changes in the concentration of various greenhouse gases and their impact on global temperature fluctuations. For natural disasters, we selected three representative examples of natural disasters and their locations, such as the Australian region suffering from severe forest fires, the Icelandic region experiencing volcanic eruptions, and the city of Wuhan after the outbreak of the COVID-19 epidemic. In order to quantify the specific impacts of natural disasters on local temperatures, we constructed an LSTM model to predict the post-disaster temperature trends using pre-disaster temperature data as inputs, and compared the predicted results with the actual observed temperature data, so as to analyze the potential effects of the natural

disaster events on regional temperatures.

2. Assumptions and Justifications

Considering that practical problems always contain many complex factors, first of all, we need to make reasonable assumptions to simplify the model, and each hypothesis is closely followed by its corresponding explanation:

Assumption1: The data used in the article are true and reliable.

Justification1: The article results are analyzed according to the data searched, so the analysis results are meaningful only when the data is reliable.

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

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

Assumption 3: Our data sources are reliable and accurate.

Justification 3: Since our data comes from international data sites, we assume that the data is reliable. On this basis, objective and accurate results can be obtained by applying the data to the establishment of our model.

Additional assumptions are made to simplify analysis for individual sections. These assumptions will be discussed at the appropriate locations.

3. Notation

Some important mathematical notations used in this paper are listed in Table 1.

Table 1: Notations Used in This Paper

Symbol	Meaning
$U_{t,n}$	Mutation point calculation
$sgn(.)$	Symbolic functions
K_t	The most likely mutation point in the sequence
ε	Random error
α	Fitting coefficient
β	Fitting coefficient
L	Extremely large likelihood function
e	Residual
p	ARIMA model order
q	ARIMA model order
MSE	Mean squared error
$SMAPE$	Symmetric mean absolute percentage error
dw	Calculated value of Durbin-Watson test

4. Data Acquisition

To ensure the authority and consistency of evaluation data, we used data published by international authorities and official statistical agencies. The data source for each metric is displayed in a label.

Table 2: Data Source Table

Database Names	Database Websites Data
NOAA	https://www.esrl.noaa.gov/gmd/ccgg/trends/
NASA GISTEMP	https://data.giss.nasa.gov/gistemp/
HadCRUT	https://www.metoffice.gov.uk/hadobs/hadcrut4/data/current/download.html
Berkeley Earth	http://berkeleyearth.org/data/
Google Scholar	https://scholar.google.com/

5. Analysis of Global Temperature Change

5.1. Pettitt mutation detection

Pettitt mutation detection is a nonparametric test that can test whether there is a mutation point in the time series when the mutation time is unknown, and determine whether the mutation is significant or not. The specific implementation steps of the Pettitt mutation detection are as follows.

(1) Determining the hypothesis test.

The original hypothesis H_0 of Pettitt's mutation detection is: there is no mutation in the time series at time t ; the alternative hypothesis H_1 is: there is a mutation in the time series at time t.

(2) Searching for possible mutation sites

For a time series $X = (x_1, x_2, \ldots, x_n)$, assume that its mutation point is x_t, then the original time series can be divided into two parts: $x_1, x_2, \ldots, x_t$ and $x_{t+1}, x_{t+2}, \ldots, x_n$. The defined statistic $U_{t,n}$ is calculated as follows.

$$U_{t,n} = U_{t-1,n} + \sum_{j=1}^{n} sgn(x_t - x_j) \quad t = 2, \ldots, n \tag{1}$$

Where sgn is the symbolic function, defined as follows.

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

For the time at which the mutation point may occur, define the statistic Kt to find the most likely mutation point.

$$K_t = \max_{1 \le t \le n} |U_{t,n}| \tag{3}$$

(3) Significance test

After finding the optimal possible mutation point according to Formula 3, its significance level Pt is calculated by the following equation.

$$P_t = 2e^{\frac{-6K_t^2}{n^3+n^2}} \quad (4)$$

For a given confidence level , if $Pt > \alpha$, the original hypothesis is accepted, and there is no significant mutation at time t; if $Pt < \alpha$, the original hypothesis is rejected., there is a significant mutation at time t.

We investigated the Northern Hemisphere land mean temperature, Southern Hemisphere land mean temperature, global land mean temperature change and global land plus ocean mean temperature change for the month of March over the last decade. We treat these four time-seriesseparately using the Pettitt mutation point detection to explore whether there is a greater increase in global temperature in March 2022.

5.2. ARIMA

The ARMA model can be derived from the AR model and the MA model. the AR model describes the relationship between current values and history, and can use historical data for forecasting. Its formula is expressed as follows.

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + \mu_t \quad (5)$$

where Xt denotes the value at moment t, αp denotes the coefficient of the pth predictor variable, μt is the random disturbance term, and p denotes the order. If μt is white noise, i.e., μt=εt, then the formula is a pure AR model. It can be seen from the formula that the historical value of the p moments before the t moment is used to predict the current value. And when μt is not white noise, the perturbation is considered as a qth order moving average, i.e., the errors of the first q moment in the autoregressive model are summed up, and the formula is expressed as follows.

$$\mu_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \ldots + \beta_q \varepsilon_{t-q} \quad (6)$$

where ε denotes the white noise series, εt denotes the noise value at moment t, and βq denotes the coefficient of the qth predictor variable. The formula indicates that the historical white noise in the AR model can indirectly affect the current predicted value. When the AR model perturbation term is combined with the MA model, the ARMA model can be obtained, and the formula of ARMA (p,q) is expressed as follows.

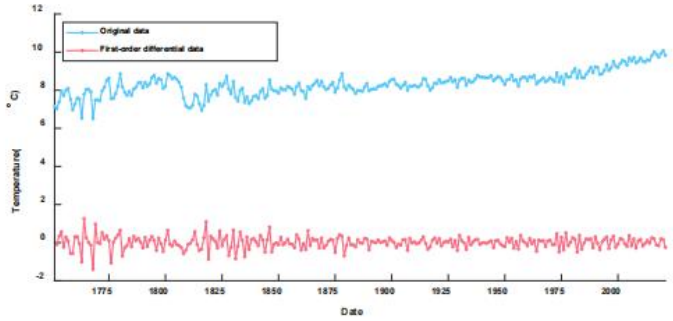
$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \ldots + \beta_q \varepsilon_{t-q} \quad (7)$$

If the ARMA model is applied to a smooth time series transformed into a non-stationary series by a difference operation, the ARIMA model is obtained, which is described as ARIMA (p,d,q), where d denotes the order of the difference transformation.

Taking the global average land temperature changes over the years as an example, the ARIMA model is established as follows:

(1) Determining the differential order

Determine the order of the data for the difference operation by the smoothness test. Determine whether the original data is a smooth time series according to the ADF test, and if not, do the difference operation on the original data until the ADF test is satisfied.



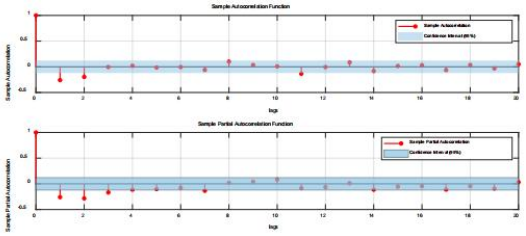


Figure 3: Original data and differential operation results.
Figure 4: Autocorrelation and bi-correlation plots after differential operations

The ADF test is used to determine the difference order of 1; the results of the difference operation are shown in Figure 3. The blue line indicates the original data, and the red line is the result of the First-order difference operation.

(2) Determining the ARIMA model order

The order of the ARIMA model can be selected by manually judging the autocorrelation and partial autocorrelation plots of the smooth time series. As shown in Figure 4, where Lag indicates the order of the delay and the blue line indicates the 95% confidence interval.

The convergence in this question is not typical, and the manual judgment may be erroneous, so the AIC and BIC criteria are chosen to assist in the selection. the formulas of AIC and BIC are as follows.

$$AIC = -2\,ln(L) + 2k \tag{8}$$

$$BIC = -2\,ln(L) + 2k\,ln(n) \tag{9}$$

where L denotes the maximum likelihood function, k is the number of parameters, and n is the series width. the ARIMA model order calculated by the AIC and BIC criteria is ARIMA (1, 1, 1).

(3) Residual Analysis

Assuming that the residuals are et, the correlation equation for each residual is:

$$e_t = \rho e_{t-1} + v_t \tag{10}$$

The original hypothesis ρ and alternative hypothesis ρ are tested, and the test statistics are calculated as follows.

$$dw = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2} \tag{11}$$

The DW test and QQ plot results are shown in Figure 5.

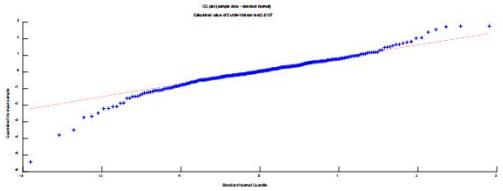


Figure 5: Residual analysis graph

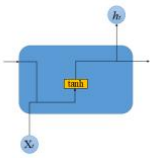


Figure 6: RNN unit.

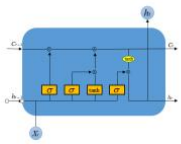


Figure 7: LSTM unit.

The obtained statistic dw is 2.0127, which is between 1 and 3, indicating that the residuals are not autocorrelated. Therefore, the residuals of the proposed model are randomly and normally distributed and not autocorrelated, indicating that the residuals are a segment of white noise signal and the useful signal has been extracted into the ARMR model.

5.3. LSTM

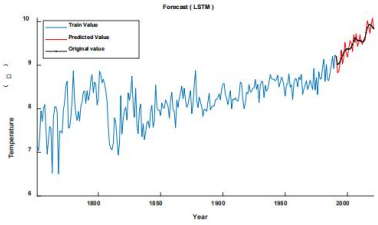
The motivation for the LSTM proposal is to solve the long-term dependency problem we mentioned above. Conventional RNN node outputs are determined only by the weights, bias, and activation functions (Figure 6). The LSTM is composed of a series of LSTM Units , whose chain structure is shown below (Figure 7).

Figure 6: RNN unit. **Figure 7:** LSTM unit.

- Training of LSTM Models

We used the annual global land mean temperature from 1750 to 2021, which was imported into the LSTM algorithm for training. In which, MaxEpochs is set to 630 times, InitialLearnRate is set to 0.9%, GradientThreshold is set to 0.01.

The first 95% of these 271 years, i.e., the data points from 1750 to 2006, were used as the training set for training the LSTM model. The data points from 2007 to 2021 were used as the test set for testing the gap between the predicted data points and the true values. The predicted data are shown in Figure 8.



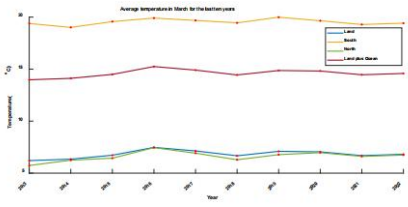


Figure 8: Prediction of the trained LSTM model on the test set.

Figure 9: Average temperature in March for the last ten years.

As can be seen in Figure 8, the data predicted by the model can simulate the distribution of future data points to a certain extent. The effect of data shows that the model predicts better results and can reflect the real trend of data to a certain extent.

5.4. Model Implementation and Results

5.4.1. Evaluation Indicators

To compare the performance of the two models, we used two metrics commonly used in time

series forecasting, MSE and SMAPE.

MSE (mean squared error), is the mean value of the sum of squares of the errors of the observed and predicted values. It is the second-order moment of the error, containing the variance of the estimate and its deviation, and is a measure of the quality of the estimate, whose formula is defined as follows.

$$MSE(y,\hat{y}) = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{12}$$

SMAPE (symmetric mean absolute percentage error), is a measure of the predictive accuracy of a forecasting method. SMAPE is very intuitive in interpreting relative errors, and its formula is defined as follows.

$$SMAPE(y,\hat{y}) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{\frac{|y_i|+|\hat{y}_i|}{2}}\right| \times 100\% \tag{13}$$

5.4.2. Pettitt mutation detection

We investigated the Northern Hemisphere land mean temperature, Southern Hemisphere land mean temperature, global land mean temperature and global land plus ocean mean temperature for the month of March over the last decade. The results are shown in Figure 9.

As can be seen in Figure 9, in March, the Southern Hemisphere land mean temperature is higher than the Northern Hemisphere land mean temperature, which is consistent with our expectation. The global land plus ocean average temperature is also higher than the global land average temperature, which is consistent with the theory of specific heat capacity. After that, we did Pettitt mutation point test for these four data respectively, and the results are shown in Figure 10.

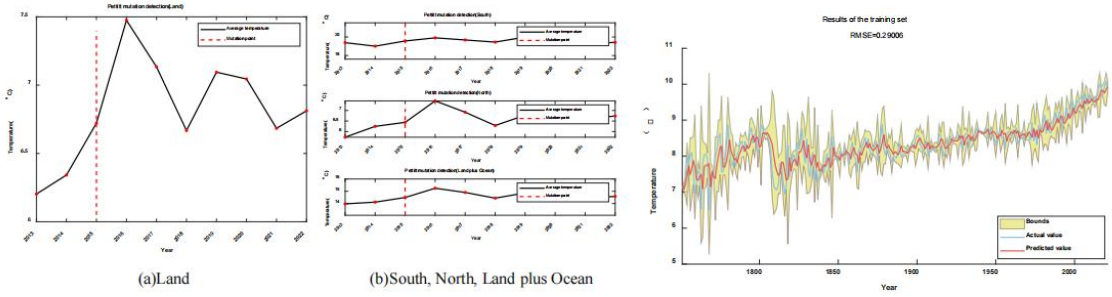


Figure 10: Pettitt mutation detection.

Figure 11: The results of ARIMA model prediction (1800-2021).

As can be seen in Figure 10(a), the global average land temperature in March for the last decade has a sudden increase in 2015. And the global land temperature in March 2022 does not have a higher increase than before. To make our conclusions more reasonable, we further analyse the changes in other average temperatures. As can be seen in Figure10(b), both the Southern Hemisphere land mean temperature, the Northern Hemisphere land mean temperature and the global land plus ocean mean temperature share the same characteristics. They all increase abruptly in 2015, while there is no significant change in March 2022. Therefore, we do not agree that there is a higher increase in global temperature in March

2022 than in the past.

5.4.3. Results of the ARIMA model

We used the global average land temperature from 1800 to 2021 as the independent variable and used the ARIMA model to make predictions, which are shown in Figure 11.

As can be seen in Figure 11, the predictions of the ARIMA model are very close to the original data for known data. the predictions of the ARIMA model have the same trend as the original data. In order to objectively describe the strengths and weaknesses of the prediction results, we calculated the MSE and SMAPE of ARIMA. 0.2901 and 2.6436% are the results of MSE and SMAPE calculations, respectively.

Finally, we use the global average land temperature from 1800 to 2021 as the training set and use the ARIMA model to forecast the future temperature. The coefficients of the ARIMA model were calculated as shown in Table 1.

Table 1: The coefficients of the ARIMA model

Coefficient	α_0	α_1	β_1
Value	0.0063591	0.33684	-0.84058

We used ARIMA model to predict the global temperature in 2050 and 2100 respectively, and the results are shown in Figure 12. From Figure 12, we can see that the global temperature cannot reach 20 degrees in 2050 or in 2100. According to the ARIMA model the global temperature projections for 2050 and 2100 are 10.05 and 10.5 degrees Celsius, respectively.

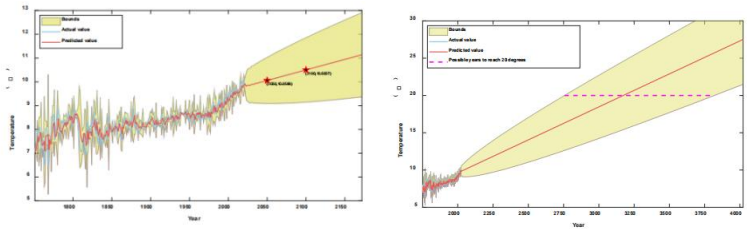


Figure 12: The results of ARIMA model prediction.
Figure 13: The results of ARIMA model prediction (2022-4021).

We then used the LSTM model to predict the global temperature change from 2022 to 4021, and then found the years when the global temperature reached 20 degrees. Since the model predictions have errors, the actual years that reach 20 degrees may vary within a certain range. We used a 95% confidence interval to determine this range. The results are shown in Figure 13.

We predict that global temperatures will reach 20 degrees in 3081. With a 95% confidence interval, the globe will probably reach 20 degrees between [2756, 3819].

5.4.4. Results of the LSTM model

Using the already trained model, the entire data from 1750 to 2021 was used as the test set and the model was allowed to predict the temperature change from 1750 to 2021, and it can be seen in Figure 14 that the model fits well with the real data. The MSE and SMAPE between the real data and the predicted data are also calculated using Equation (12), which yields an MSE of0.2888 and a SMAPE of 2.6302%. It can be seen that the model is well trained and the

future temperature change can be accurately predicted using this model.

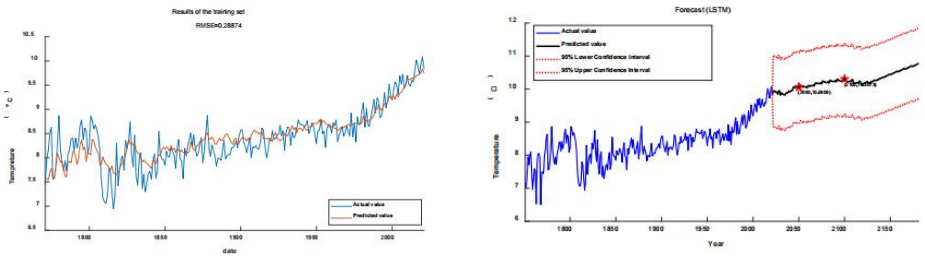


Figure 14: Results of prediction of the full set of data using the trained model.
Figure 15: The results of LSTM model prediction.

We used LSTM model to predict the global temperature in 2050 and 2100 respectively, and the results are shown in Figure 15. From Figure 15, we can see that the global temperature cannot reach 20 degrees in 2050 or in 2100. According to the LSTM model the global temperature projections for 2050 and 2100 are 10.06 and 10.31 degrees Celsius, respectively.

The predicted temperature for the future 2100 years using the model that has been predicted is shown in Figure 16. In the figure, we have used the purple line to mark the range where the global average land temperature reaches 20°C, and because the model's predictions have errors, the years we predict should fall in an interval. We used a 95% confidence interval to determine our interval.

As can be seen from the Figure 16, we predict that the global average land temperature will reach 20°C in 3094. Within the 95% confidence interval, the global average land temperature will reach 20°C between the years [2988, 3199].

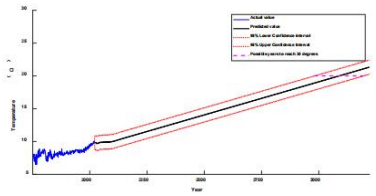


Figure 16: Results of global average land temperature predictions for the future 2100 years.

5.4.5. Comparison of models

The differences between the ARIMA model and the LSTM model are shown in Table 2.

Table 2: Differences between ARIMA and LSTM models.

Models	ARIMA	LSTM
MSE	0.29006	0.28874
SMAPE	2.6436%	2.6302%
Forecast interval	[2756, 3819]	[2988, 3199]

(1) Comparison of MSE and SMAPE between models

A comparison of the above results shows that the MSE and SMAPE predicted by the ARIMA model for known data are 0.2901 and 2.6436%, respectively, while the MSE and SMAPE predicted by the LSTM model for known data are 0.2888 and 2.6302%, respectively, which are smaller than the results of ARIMA, which can indicate that the LSTM model is better than

the ARIMA model.

(2) Comparison between model prediction results

As can be seen in the discussion in Section 5.4.3, the ARIMA model predicts that the global mean land temperature will reach 20°C in 3288. Within the 95% confidence interval, the global land mean temperature will reach 20°C between [2756, 3819] years. The LSTM model predicts that the global land mean temperature will reach 20°C in 3094. At 95% confidence interval, the global land mean temperature will reach 20°C between [2988, 3199] years. The future interval between the two predictions intersects [2988, 3199], so we can assume that the global land mean temperature is more likely to reach 20°C between [2988, 3199].

6. Analysis of Global Temperature Impact Factors

6.1. Impact of Temporal and Spatial Factors

6.1.1. Data description

Since north and south latitudes and east and west longitudes represent different geographical locations, we set the data for west longitude to negative values in order to describe the differences in geographical locations with the data. The increase of global temperature data before 1900 is not very obvious, so we select the data from January 1900 to August 2013 for analysis. The temperatures in each city are shown in Figure 17, and only 15 cities are shown here.

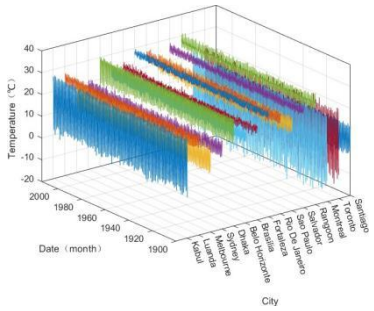


Figure 17: Temperatures by city.

As shown in the Figure 17, the temperature varies greatly from city to city, e.g., the temperature in Kabul varies around 20 degrees Celsius and varies more, while the temperature in Fortaleza varies around 30 degrees Celsius and varies less. For each city, the temperature increases slowly over time. Therefore, we believe that temperature may be correlated with latitude, longitude, and time. We used grey correlation analysis with multiple linear regression to explore the association between temperature and these factors.

6.1.2. Grey correlation analysis

In order to understand the degree of correlation between temperature, time and location, we used gray correlation analysis to determine the gray correlation between time, longitude and latitude and temperature. The specific steps of the gray correlation analysis are as follows.

(1) Data Standardization

Since temperature, time, and location are different orders of magnitude, in order to eliminate the effect of the magnitude, we need to standardize the data and transform them into the

same range.

$$x_{ij} = \frac{x_{ij}^{'}}{mean(x_j^{'})} \tag{14}$$

(2) Calculating the grey correlation

The formula for calculating the gray correlation is as follows.

$$\zeta_i(k) = \frac{min_i min_k |x_0(k) - x_i(k)| + \rho \cdot max_i max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \cdot max_i max_k |x_0(k) - x_i(k)|} \tag{15}$$

where is the discrimination coefficient. It ranges from (0 to 1), and its role is to control the degree of discrimination. The smaller its value, the greater the discrimination, and the larger its value, the smaller the discrimination. It is often taken as 0.5.

6.1.3. Multiple linear regression model

Since gray correlation analysis is only able to determine the degree of correlation between factors, it is not able to determine what kind of relationship the factors have with each other. The specific steps for building a multiple linear regression model are as follows.

(1) Data Standardization

As with gray correlation analysis, we first need to standardize the data.

(2) Determining the multiple linear regression equation

The expression of the multiple linear regression equation is given below.

$$\begin{cases} Y_1 = \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_{p-1} X_{1,p-1} + \varepsilon_1 \\ Y_2 = \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_{p-1} X_{2,p-1} + \varepsilon_2 \\ \vdots \\ Y_3 = \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_{p-1} X_{n,p-1} + \varepsilon_n \end{cases} \tag{16}$$

where ε is the random error and follows a normal distribution, i.e., $\varepsilon \sim N(0, \sigma^2)$. The estimated value of β is obtained by the least squares method and is calculated as follows.

$$\hat{\beta} = (X^T X)^{-1} X^T Y \tag{17}$$

The values of β reflect the relationship of each factor to the temperature.

(3) Significance testing

To verify the validity of the multiple linear regression, we calculated F -values with significance level P to analyze.

6.1.4. Results

We obtained the gray correlation between temperature, time, longitude and latitude using gray correlation analysis as shown in Table 3.

Table 3: Grey correlation analysis.

	Time	Latitude	Longitude
Temperature	0.7919	0.7514	0.6818

The results in Table 3 show that the temperature variation is highly correlated with time and latitude and less correlated with longitude.

After that, we described the relationship between the factors by multiple linear regression equations, and the regression equation coefficients and related statistical indicators are shown in Table 4.

Table 4: Multiple linear regression analysis.

Item	β_0	β_1	β_2	β_3	F	P
Value	1.5345	0.0308	-0.5996	0.0343	26648.5	0

As seen in Table 4, the fitted F-value of the model is 26648.5 and the p-value is approximately 0. This indicates that the model fits well. In order to visualize the fitting effect and the role of different factors on the temperature, we plotted the temperature versus time and the temperature versus latitude and longitude, respectively, as shown in Figure 18, Figure 19.

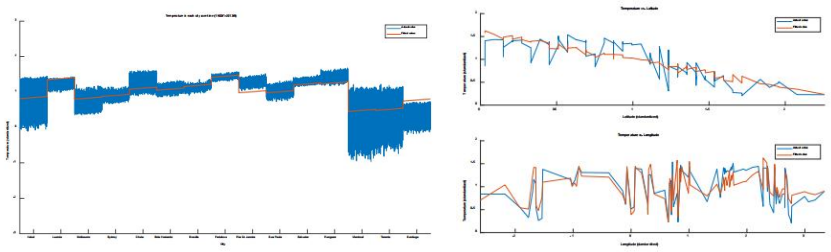


Figure 18: Multiple linear regression analysis (Temperature vs. Time).

Figure 19: Multiple linear regression analysis (Temperature vs. Latitude & Temperature vs. Latitude).

As can be seen in Figure 18, the average temperature in each city has increased over time. The trends of both the actual and the fitted values reflect this feature. As can be seen in Figure 19, the temperature gradually decreases when the latitude goes from low to high.

6.2. Impact of Natural Disaster Factors

Since both MSE and SMAPE metrics of LSTM model are better than ARIMA model, the following problem will be solved by LSTM model.

(1) Impact of forest fires on temperature

By observing the difference between the predicted and actual values, we can determine the impact of forest fires in Australia on the local temperature, and thus on the global temperature impact. The prediction results are shown in Figure 20. The graph shows that the predicted values are slightly higher than the actual values, which means that if there were no forest fires, the future temperatures would be higher than they would have been if the fires had occurred. It shows that forest fires will lower the temperature.

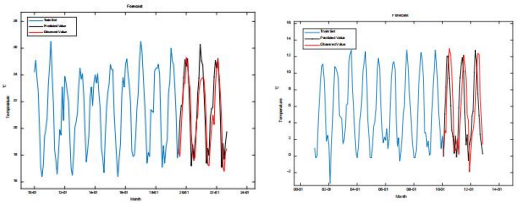


Figure 20: Forecast results for local temperatures in Australia.

Figure 21: Predicted results for temperatures after the Icelandic volcano eruption.

(2) Effect of volcanic eruptions on temperature

In March 2010, the Icelandic volcano erupted. We selected the local monthly average temperature of Iceland from January 2001 to December 2012 and used the data before March 2010 as the training set to train the model. The trained model was used to predict the future temperature and compare it with the actual value, as shown in Figure 21.

As can be seen from the graph, the predicted value is slightly higher than the actual value, which indicates that the local temperature in Iceland has decreased after the volcanic eruption. Therefore, we can know that the volcanic eruption will lower the temperature in the future.

(3) Effect of COVID-19 on temperature

In December 2019, a new type of pneumonia broke out in the world, and the main virus causing this pneumonia is COVID-19. We selected the monthly average temperature of Wuhan from February 2013 to November 2022. The data before December 2019 is used as the training set, and the effect of COVID-19 on the temperature can be known by comparing the predicted data with the actual data, as shown in Figure 22.

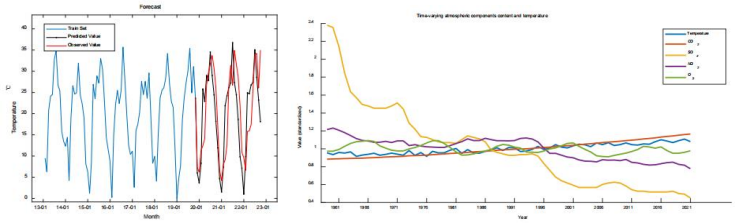


Figure 22: Predicted results for temperature after COVID-19 outbreak.
Figure 23: Time-varying atmospheric components content and temperature

Observing the above graph, we can see that the predicted values are generally higher than the actual values, which indicates that the local temperature in Wuhan dropped after the COVID-19 outbreak. Therefore, we believe that COVID-19 will cause the temperature to drop.

6.3. Impact of Atmospheric Factors

We investigated the atmospheric composition content and temperature from 1959 to 2021, and the results are shown in Figure 23.

It can be seen from Figure 23 that the trends of CO_2, O_3, NO_2 are relatively close to the temperature trends, while the trends of SO_2 differ from the temperature trends. In order to objectively describe the relationship between each factor and temperature, we applied a gray correlation analysis and ranked them from largest to smallest by degree of association. The results are shown in Table 5.

Table 5: Grey correlation analysis（Atmospheric composition and temperature）.

Rank	Component	Gray correlation value
1	CO_2	0.9592
2	O_3	0.9169
3	NO_2	0.8300
4	SO_2	0.6794

From Table 5, it can be seen that CO_2 and O_3 are highly correlated to temperature, NO_2 is

second to temperature, and SO_2 is less correlated to temperature. This is consistent with what we observed from the trend of the graph.

After that, we calculated the Pearson correlation coefficients between each factor to investigate the positive and negative correlations of each atmospheric component on the words of temperature and whether there is a correlation between each atmospheric component. The result is shown in Figure 24.

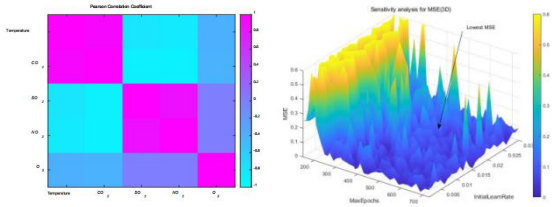


Figure 24: Pearson Correlation Coefficient
Figure 25: Sensitivity analysis for MSE(3D).

As can be seen in Figure 24, temperature has the highest correlation with CO_2, while the other atmospheric components have a lower correlation with temperature.

Combining these two approaches, we believe that the main cause of global temperature change is CO_2.

7. Sensitivity Analysis

By changing the hyperparameters, different models can be trained, and by comparing the MSE and SMAPE metrics of each model, the most suitable hyperparameter setting for predicting that type of data can be derived. As shown in Figure 25, these are the 3D distributions of the MSE metrics derived by varying the MaxEpochs and InitialLearnRate.

From the Figure 25, it can be observed that the lowest MSE metric can be obtained when MaxEpochs is taken as 540 and InitialLearnRate is taken as 0.021. In this paper, the model is trained using this hyperparameter and the model results are robust

8. Conclusion

Our analysis shows that carbon dioxide is one of the factors that have the greatest impact on global warming, so we should first consider how to reduce carbon dioxide emissions. We should develop and apply advanced technologies to reduce carbon emissions, such as large-scale use of renewable energy technologies such as solar energy, wind energy, biomass, and hydropower. We should also develop advanced clean coal technology, fuel cell technology, advanced nuclear power technology, advanced natural gas power generation technology, unconventional energy utilization technology, synthetic fuel utilization technology, decarbonization and carbon sequestration technology, etc.

At the same time, the public will be educated about environmental protection so that every citizen will know how to pay attention to environmental protection in their daily lives and change bad social behavior habits to reduce emissions. It is also important for everyone to

learn about climate change and to consciously pay attention to climate change in industry development, business activities and daily life. For example, when you are ready to upgrade your home appliances, choose energy-efficient appliances; turn off your TV and stereo instead of leaving them on standby; and make full use of solar energy, as solar panels can also generate considerable energy. All of these small aspects of life can make a big contribution to saving energy.

Acknowledgements

Natural Science Foundation.

References

[1] Tedesco Marco,Keenan Jesse M.,Hultquist Carolynne. Measuring, mapping, and anticipating climate gentrification in Florida: Miami and Tampa case studies[J]. Cities,2022,131.

[2] Li Han,Grant Richard J.. Climate gentrification in Miami: A real climate change-minded investment practice?[J]. Cities,2022,131.

[3] Wang Liang. Exploring a knowledge map for urban resilience to climate change[J]. Cities,2022,131.

[4] Reu Junqueira Juliana,Serrao-Neumann Silvia,White Iain. Using green infrastructure as a social equity approach to reduce flood risks and address climate change impacts: A comparison of performance between cities and towns[J]. Cities,2022,131.

[5] Almeida Bernardo,Jacobs Carolien. Land expropriation – The hidden danger of climate change response in Mozambique[J]. Land Use Policy,2022,123.

[6] Athy Ariana E.,Milojev Petar,Gray Nathan Hoturoa,Osborne Danny,Sibley Chris G.,Milfont Taciano L.. Clarifying longitudinal relations between individuals' support for human rights and climate change beliefs[J]. Journal of Environmental Psychology,2022,84.

[7] Alfonso S, Gesto M, Sadoul B. Temperature increase and its effects on fish stress physiology in the context of global warming[J]. Journal of Fish Biology, 2021, 98(6): 1496-1508.

[8] Albouy C, Delattre V, Donati G, et al. Global vulnerability of marine mammals to global warming[J]. Scientific Reports, 2020, 10(1): 548.