ISSN: 1813-4890

Path analysis and study of global temperature change

Mengting Lu[1, a], Zijian Chen[2], Wenshuo Li[3], Ziheng Liu[4], Huiqin Deng [5]

[1]School of Anhui University of Finance and Economics ,Anhui ,China;

[2]School of Anhui University of Finance and Economics ,Anhui ,China;

[3]School of Anhui University of Chinese Medicine,Anhui ,China.

[4]School of Anhui University of Chinese Medicine,Anhui ,China.

[5]School of Anhui University of Chinese Medicine,Anhui ,China.

[a]3368096562@qq.com

Abstract

In this paper, the path of global temperature change and its influencing factors, Using the grey predictions, Multiple linear regression and the ARIMA-LSTM combined predictive model, A grey prediction global temperature level prediction model based on time series, a combined global temperature level prediction model based on differential integration moving average autoregressive based on time series-long short-time memory neural network (ARIMA-LSTM), a global temperature influence factor analysis model based on multiple regression were constructed, Using MATLAB, Python, STATA and other software programming solution, Conclusion at the global temperature change law, the factors affecting the global temperature change and the most important factors, Finally, the model and the survey data are combined to provide suggestions to suppress global warming. The highlight of this paper is the construction of a combined ARIMA-LSTM model, which combines the advantages of both models to obtain the optimal prediction value.Process the data and use MATLAB software to draw a scatter plot of global temperature versus time to find the maximum of the global temperature increase rate in 2022 and the temperature increase rate in the previous decade, respectively. First, the data are discretized and a gray prediction model is built, and the error test is performed with the posterior difference test; second, a combined ARIMA-LSTM model is built and the error test is performed with MSE, MAPE. Comparing the error magnitudes of the two models according to the error test, predicting the global temperature in 2050 and 2100 and predicting the time when the global average annual temperature reaches 20.00℃, the ARIMA-LSTM is found to be better. A multiple linear regression model is constructed for the three variables of global temperature, time, and region.Collect relevant data to analyze the impact of natural disasters on global temperature. Collect data to refine the multiple linear regression model of global temperature through variable selection and stepwise regression. We get that geographical location is the main cause of global change, and propose targeted measures to curb global warming. The paper also concludes with an objective evaluation of the strengths and weaknesses of the model and an improvement of the model.

Keywords

GM(1,1) model; ARIMA-LSTM model; Multiple linear regression; Global annual mean temperature; Time series.

1. Restatement of the problem

1.1. Background knowledge

CO_2 emissions have been stabilized before the industrial revolution, but in 2004, CO_2 emissions have increased significantly, and this change has caused great concern. Therefore, itis important to study the trend of global climate temperature change and the factors influencing the global climate temperature change.

1.2. Questions to be addressed

Problem 1: Compare the rate of increase in global temperature in March 2022 with the rate of increase observed during any previous 10-year period.

Problem 2: Based on historical data, develop two or more mathematical models to describe the past and predict when the global temperature will reach 20.00°C in the future, and compare the accuracy of the models.

Problem 3: Using the results in Problem 1, the Annex 2022 APMCM C data .csv and data from other datasets collected by your team, build a mathematical model to analyze the relationship between global temperature, time, and location and explain the relationship between them.

Problem 4: Collect data to analyze the effects of natural disasters (e.g., volcanic eruptions, forest fires, and Neocrown pneumonia) on global temperatures.

Problem 5: Analyze the most significant factors affecting global temperature change and also give measures to curb global warming.

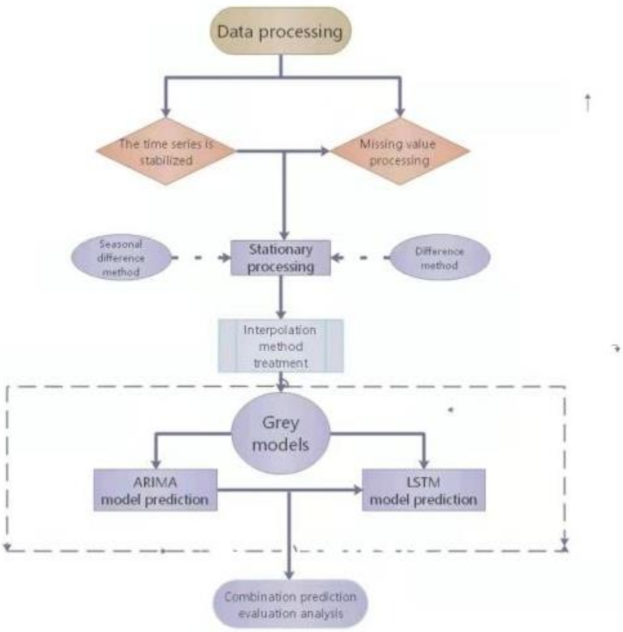

Fig. 1: Flow chart of problem solving a impact factors

2. Model assumptions

① Assume that the missing data in the data pre-processing global temperature change is consistent with the overall trend of data change

② Assume that no other factors are associated with the global temperature time and location relationship when studying them

③ Assume that no large perturbations of global temperature change will be generated in the long-term global future time

④ Assume that the collected global temperature change data are true and reliable

ISSN: 1813-4890

⑤ Assume that the effects of natural disasters on global temperature are negligible except for volcanic eruptions, forest fires, and neocrown pneumoni

3. Symbols

Table 1: Symbols.

Serial number	Symbols	Definition
1	e_t	Prediction error
2	y_t	Temperature change value
3	X_t	Time series
4	ε_t	Residual
5	L	Factor of lag
6	∇^d	Backward difference of order d
7	$\varphi(L)$	Autocorrelation coefficient
8	θ_i	Moving average coefficient

4. Data processing

4.1. Missing value processing

The missing data in this dataset are presented as null values, and instead of removing them in order not to affect the integrity of this time series, we decided to use the mean-fill method for such data. The data for this experiment are regional temperature data with a time series, and it is crucial to carry forward and backward, so it is necessary to take into account the before and after part of the missing values when dealing with this type of data, and to fill the vacant part by averaging them. For this data set, we list the time-series data with the same decision attribute values, and calculate the average of the before and after time-series data to fill the missing values, so that they can be more closely related to the trend of the time series.

4.2. Data classification results

Classification processing of the data

We import the data in Annex 2022 APMCM C data .csv into the Python library, classify the data by year, and average the temperature for all regions for different years, and use this value to replace the global average temperature to make a line graph, as shown in Figure 2.

Through Figure 2, we can see that the temperature from 1743 to 1800 is significantly lower than that from 1800 to 2013.

We found that this is due to the missing data of many cities from 1743 to 1800 and the increase of global temperature after the industrial revolution, so we further processed the data to get the scatter plot of global annual average temperature before and after the industrial revolution, as shown in Figure 2.1

ISSN: 1813-4890

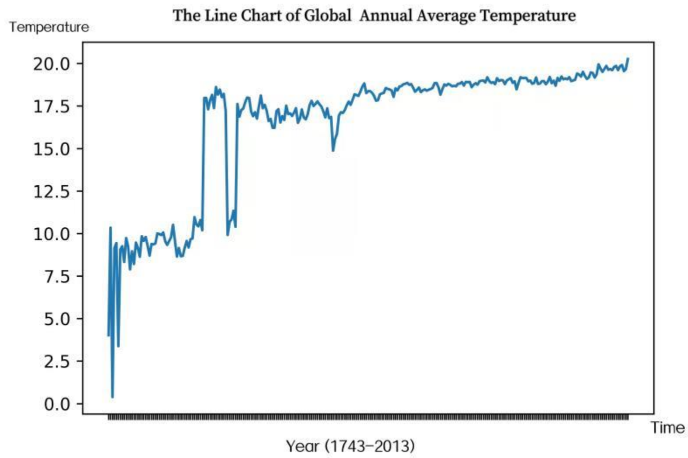

Fig.2: Global average annual temperature line graph

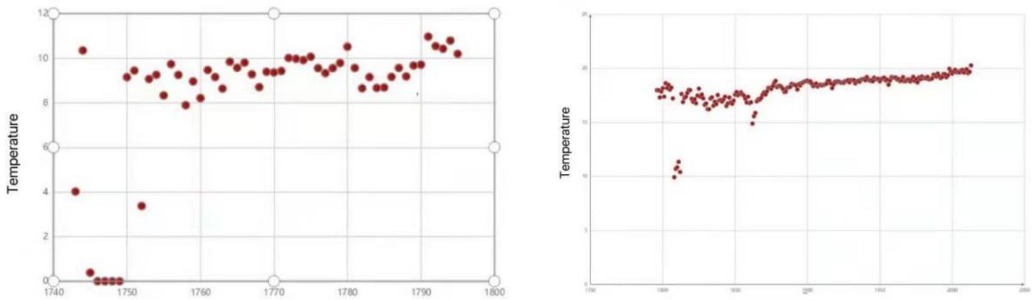

Fig. 3: Scatter plot of global average annual temperature before and after the industrial revolution

5. Model Solution

5.1. Analysis and solution of Problem 1 a)

Using python, the maximum temperature growth rate for the decade before 2013 is solved as shown in Figure 4.

The maximum growth rate is 75%, while the growth rate from 2012 to 2022 is 9.92%. It is clear that the global temperature growth rate in March 2022 is not the maximum and is not greater than that observed at any previous time.

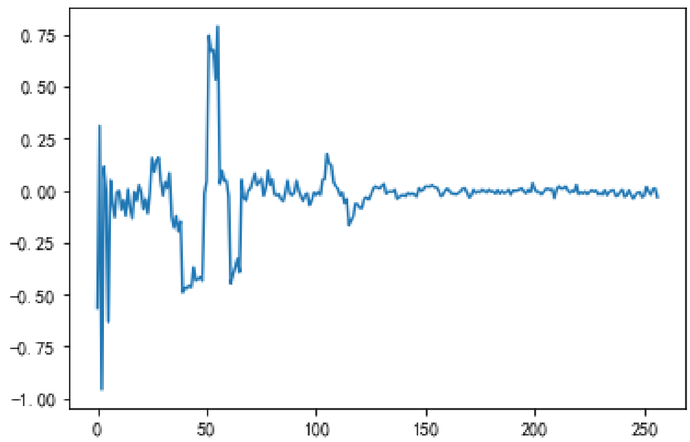

Fig. 4: Decadal growth rate of global temperature interval

5.1.1. Analysis of the problem

We were asked to compare the rate of temperature increase in 2022 with any observed rate of increase in the previous decade, based on historical data, build two or more mathematical models to describe the past and predict future global temperature levels, and use each of the 1(b) to predict global temperatures in 2050 and 2100, and find the time when global temperatures reach 20.00°C based on your models, along with an indication of which model is more accurate in its predictions.

To address this issue, first, we preprocess the data in Appendix 2022 APMCM C data.csv to ensure the accuracy of the global annual average temperature by removing all data for the year in which the data are missing, based on missing data for some years or months in some countries. Secondly, a time series was constructed based on the pre-processed data, and a gray prediction model and an ARIMA-LSTM model were built to predict the time when the global temperature reached 20.00°C. The time when the global temperature reached 20.00°C was then predicted based on the predicted 20.00°C. Then, the global temperature growth rate from 2012 to 2022 is calculated based on the predicted temperature level in 2022, and the maximum value of the temperature growth rate in the decade before 2013 is solved by python, and the magnitudes of the two growth rates are compared. Finally, the two models are tested separately to compare the magnitude of the error and illustrate the accuracy of the model prediction.

5.2. Analysis and Solution of Problem 1 b), c) and d)

5.2.1. Introduction to the gray GM(1,1) prediction model

The GM(1,1) model is a model based on the basic principle of gray prediction, in which the original data series are processed by a series of operations such as accumulation and subtraction, and the generated gray series are used to find the inherent patterns among the data series, so as to predict the future. Its principle and modeling process are as follows:

(1) Selecting the original data series GM(1,1) model has a short term memory in general sense and is suitable for dealing with a small number of data series.

We select some non-negative data as the original data series m(0) , denoted as :

$$m^{(0)} = \{m^{(0)}(1), m^{(0)}(2), \ldots m^{(0)}(i) \ldots m^{(0)}(n)\}$$

(2)Level Ratio Judgment

Define the level ratio of the data series as

$$\sigma(j) = \frac{m^{(0)}(j)}{m^{(0)}(j-1)}$$

The smooth ratio of the sequence is defined as

$$\rho(j) = \frac{m^{(0)}(j)}{\sum_{i=1}^{j-1} m(i)}$$

For a sequence, if its level ratio satisfies.

For any j， 1> $\sigma(j)$ >0， Then the sequence has a negative gray index law.

For any j， s> $\sigma(j)$ >0， and s>1, Then the sequence has a positive gray exponential law.

For any j， s> $\sigma(j)$ >q， znd $s-q=\psi$， Then the sequence is a sink index law with absolute gray level of

If ψ <0.5， Then the sequence is said to have a quasi-exponential law.

For the smooth ratio, a sequence is called quasi-smooth if it satisfies the following three conditions.

① $\frac{\rho(j+1)}{\rho(j)}<1$ $j=2,3,\ldots,n$

② $0\le\rho(j)\le\partial$ $j=2,3,\ldots,n$

③ $\partial<0.5$

Since the variation of the original data series in the system may not have certain regularity, it is necessary to use gray processing to reduce the randomness of the original data series by subjecting the original data series to one accumulation process to obtain.

$$m^{(1)}=\{m^{(1)}(1),m^{(1)}(2),\ldots m^{(1)}(i),m^{(1)}(n)\}$$

Among them,

$$m^{(1)}(k)=\sum_{k=1}^{n}m^{(0)}(k)$$

is called the immediate neighbor generating sequence of

$$z^{(1)}=\{z^{(1)}(1),z^{(1)}(2),\ldots z^{(1)}(i)\ldots z^{(1)}(n)\}$$

where

$$z^{(1)}(k)=0.5(m^{(1)}(k)+m^{(1)}(k))$$

Then the following gray differential equation is referred to as the GM(1,1) model. $m^{(0)}(k)+\alpha z^{(1)}(k)=\mu$

(4) For establishing the first-order whitening differential equation, it is obtained that $\frac{dm^{(1)}}{dt}+\alpha m^{(1)}=\mu$

where α is the development coefficient and u is the amount of gray effect.

(5Based on the least squares principle for solving) α and u

$$\text{let}\ \hat{c}=\begin{pmatrix}\alpha\\ \mu\end{pmatrix}\text{and}\ A=\begin{pmatrix}-Z^{(1)}(2) & 1\\ -Z^{(1)}(3) & 1\\ . & .\\ . & .\\ -Z^{(1)}(i) & 1\\ . & .\\ -Z^{(1)}(n) & 1\end{pmatrix}\quad Y_N=\begin{pmatrix}m^{(0)}(2)\\ m^{(0)}(3)\\ .\\ .\\ m^{(0)}(i)\\ .\\ m^{(0)}(n)\end{pmatrix},$$

we have,

$$\hat{c}=\begin{pmatrix}\alpha\\ \mu\end{pmatrix}=\left[(A^TA)^{-1}A^TY_N\right],$$

and

$$\frac{dm^{(1)}}{dt}+\alpha m^{(1)}=\mu.$$

Therefore, the solution of the first-order whitening equation can be obtained

$$m^{(1)}(k+1)=(m^{(0)}(1)-\frac{\mu}{\alpha})e^{-\alpha k}+\frac{\mu}{\alpha}$$

(6)Data reduction.

Since the prediction mechanism of the gray GM(1,1) model is generated by accumulation, the generated prediction value is a sequential additive quantity, and the inverse generation reduction of the generated results is needed, after which the results can be applied in the decision making, based on the prediction mechanism of the gray GM(1,1) model, the inverse generation accumulation reduction of the generated is needed, and the reduction is.

The prediction generation sequence is.

$$m^{(1)} = \{m^{(1)}(1), m^{(1)}(2), \ldots m^{(1)}(i), \ldots m^{(1)}(n)\}$$

The above sequence is cumulated and subtracted to obtain:

$$m^{(0)} = \{m^{(0)}(1), m^{(0)}(2), \ldots m^{(0)}(i), \ldots m^{(0)}(n)\}$$

then is the predicted value that can be used as a basis for decision making, where:

$$m^{(0)}(k) = m^{(0)}(k) - m^{(0)}(k-1)$$

5.2.2. Solution of the results

We import the global average temperature data from 1840 to 2013 in chronological order as a raw series into Python software and get the following prediction results as in Figure:4. The global average temperature reaches 20.00°C in 2056 as obtained from the image.

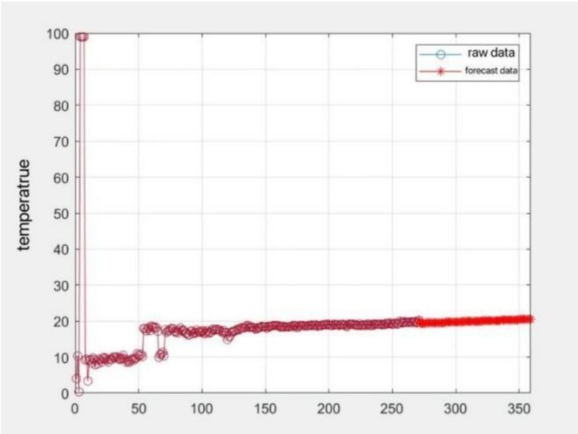

Fig. 5: Fitting prediction curve of global annual average temperature

As shown in Figure 5, the average relative residual is 0.026687, which indicates that the model fits the original data very well; the average stepwise deviation is 0.1395, which indicates that the model fits the original data to the general requirement.

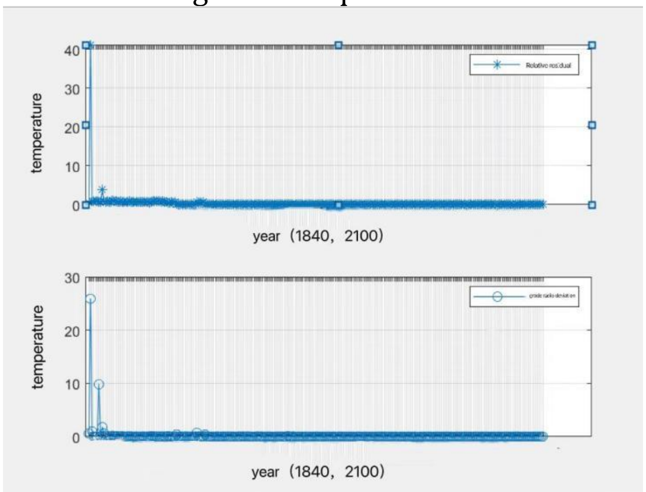

Fig. 6: Grade deviation and relative residuals of model predictions

5.2.3. Construction of ARIMA-LSTM model

ARIMA model, also known as autoregressive integrated sliding average model, is a widely used time series model. When using ARIMA model for forecasting, the original data series needs to be tested for smoothness first, and if the original data series is not smooth, it needs to be transformed into smooth data.

ISSN: 1813-4890

The principle of ARIMA model and the modeling process are as follows.
Use ADF unit root test whether the signal is a smooth series.
If the signal is non-stationary, it needs to be differenced to the d-order.
③ The order q of the AR model and MA model is determined by the truncation or tailedness of the smooth signal using the ACF function and PACF function, as shown in Table II.
④ Determine the arima model parameters p,d,q, and then the prediction function can be used to predict the values of the test data set.
The ARIMA(p,d,q) model expression is:

$$\begin{cases} \varphi(L)\nabla^d X_t = \theta(L)\varepsilon_t \\ \varphi(L) = 1 - \varphi_1 L - \varphi_2 L^2 - \cdots \varphi_P L^P \\ \theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots \theta_P L^P \end{cases}$$

where $\varphi_i(i = 1, 2, \cdots, p)$ and $\theta_i(i = 1, 2, \cdots, p)$ are autocorrelation and moving average coefficients, respectively; $\varphi(L)$ is the autocorrelation coefficient polynomial; $\theta(L)$ is the moving average coefficient polynomial; L is the lag operator; ∇^d refers to the d-order backward differential; t is the period; X_t is the time series; ε_t is the residual term in period t.

Table 2: Order determination method

Model	ACF	PACF
AR(p)	Nature of tailing	Property of truncation
MA(q)	Property of truncation	Nature of tailing
ARMA(p,q)	Nature of tailing	Nature of tailing

The LSTM model is the result of the improvement of the traditional RNN, which has certain advantages over the traditional neural network. the LSTM model is a control gate mechanism with four parts: memory cell, input gate, output gate, and forgetting gate. The principle is shown in the figure

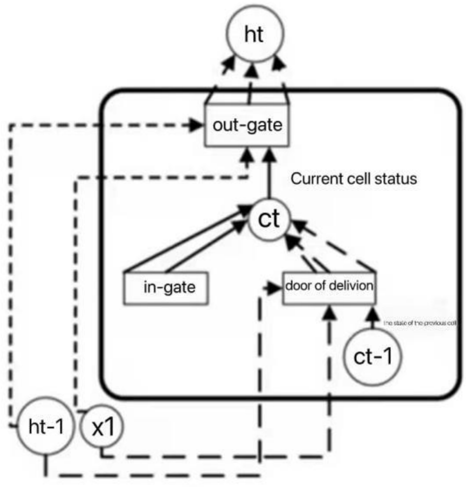

Figure 7: LSTM schematic

Principle explanation: The input gate i_t receives the current input x_t and the final hidden state h_{t-1} as input and is calculated according to the following equation.

$$i_t = \sigma(w_{ix} + w_{ih} h_{t-1} + b_i)$$

After the calculation, when i_t is 0, it means that any information entered will not enter the unit state, when i_t is 1, it means that all information currently entered enters the unit state.

The formula will calculate another value, called the candidate value, for the current unit state, $c_t = \tanh(w_{cx}x_t + w_{ch}h_{t-1} + b_c)$.

The oblivion gate will perform the following operations: an oblivion gate value of 0 means that no information is passed to the computation c_{t-1}, a value of 1 means that all information is propagated to the

$f_t = \sigma(w_{fx}x_t + w_{fh}h_{t-1} + b_f)$

Final state h_t of the LSTM cell

$o_t = \sigma(w_{ox}x_t + w_{oh}h_{t-1} + b_o)$

$h_t = o_t \tanh(c_t)$

（1）～（5） w_{ix}、w_{ih}、w_{cx}、w_{ch}、w_{fx}、w_{fh}、w_{ox}、w_{oh} denotes the weight matrix of each control gate； and b_i、b_c、b_f、b_o Indicates the bias of each control gate； σ、tanh are the activation function sigmoid function and tanh function, respectively.In response to the long-term temperature change time series is disturbed by multiple factors, its non-linearity is high and the accuracy of the prediction results is low, so this paper improves the combination of ARIMA and LSTM models, using series-parallel weighting combination, plus the model calibration in combination method, using fixed weights to weight the combination of two sets of prediction results, the accuracy of the prediction results will be relatively improved.

The optimal combination of forecasts uses the criterion of minimum combined forecast error, and then finds the weighting coefficients of each model, the following is the ARIMA-LSTM model combination modeling process.

① Let there be i single forecasting models, then the combined forecasting model formula is

$$p_t = \sum_{k=1}^{i} \lambda_k x_{kt} \quad (k=1,2,3,...i;\ t=1,2,3,...I)$$

where p_t is the prediction value of the combined prediction model; λ_k is the weighting coefficient of the kth prediction model $\sum^{i} \lambda_k = 1$; x_{kt} is the prediction result of the kth model at moment t.

② The prediction error expression of the combined prediction model is

$$e_t = y_t - p_t = \sum \lambda_k e_{kt}$$

e_t is the prediction error value of the combined prediction model; y_t is the temperature change measurement; and e_{kt} is the prediction error value of the kth model at moment t.

③ The squared value of the prediction error for the combined prediction model

$$e_t^2 = \left(\sum_{k=1}^{i} \lambda_k e_{kt}\right)^2 = \left([\lambda_1, \lambda_2, \cdots, \lambda_n][e_{1t}, e_{2t}, \cdots e_{nt}]^T\right)^2 = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_i \end{bmatrix}^T \begin{bmatrix} e_{1t}^2 & e_{2t}e_{1t} & \cdots & e_{1t}e_{it} \\ e_{2t}e_{1t} & e_{2t}^2 & \cdots & e_{2t}e_{it} \\ \vdots & \vdots & \ddots & \vdots \\ e_{it}e_{1t} & e_{it}e_{2t} & \cdots & e_{it}^2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_i \end{bmatrix}$$

④ Let the prediction error sum of squares of the combined prediction model be.

$$S=\sum_{t=1}^{I} e_t^{\ 2}=\sum_{t=1}^{I}\sum_{j=1}^{i}\left[\lambda_k\lambda_j\left(\sum_{t=1}^{I} e_{kt}e_{jt}\right)\right]=\begin{bmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_i\end{bmatrix}^T\begin{bmatrix} e_{1t}^{\ 2} & e_{2t}e_{1t} & \cdots & e_{1t}e_{it}\\ e_{2t}e_{1t} & e_{2t}^{\ 2} & \cdots & e_{2t}e_{it}\\ \vdots & \vdots & \ddots & \vdots\\ e_{it}e_{1t} & e_{it}e_{2t} & \cdots & e_{it}^{\ 2}\end{bmatrix}\begin{bmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_i\end{bmatrix}$$

⑤Let the weighted coefficient vector of the combined prediction model be: $\vec{\lambda}=[\lambda_1,\lambda_2,\cdots,\lambda_n]^T$, The error information matrix is $E(n)$. We have

$$E(i)=\begin{bmatrix} E_{11} & E_{12} & \cdots & E_{1i}\\ E_{21} & E_{22} & \cdots & E_{2i}\\ \vdots & \vdots & \ddots & \vdots\\ E_{i1} & E_{i2} & \cdots & E_{ii}\end{bmatrix}$$

In it, $E_{kj}=E_{jk}=\sum_{t=1}^{I} e_{kt}e_{jt}$, $E_{ki}=\sum_{t=1}^{I} e_{kt}^{\ 2}$, Then the combined prediction model predicts the sum of squared errors

$$S=\vec{\lambda}E(n)\vec{\lambda}^T$$

If the unit column vector is $K_n=[1,1,\cdots,1]^T$, Then the constraint of the weighting factor $\sum_{k=1}^{i}\lambda_k=1$: It can be written as $I_n^T\lambda=1$.

If the combined prediction model error variance model is required to be minimized, go to the smallest value under the constraint $S=\vec{\lambda}E(n)\vec{\lambda}^T+k(I_n^T\vec{\lambda}-1)$.

The above equation is derived by taking the partial derivative and multiplying both sides of the equation by $I_n^T E^{-1}(n)$, Find the value of the smallest variance where

$$S=\frac{1}{I_n^T E^{-1}(n)I_n}$$

Therefore, the formula for calculating the sum of squares of weighting coefficients and optimal combination of prediction methods can be derived

The weighting coefficients and the sum of squares of the prediction errors are solved by the following formulas

① The formula for solving the weighting coefficients.

$$\lambda_1=\frac{E_{22}-E_{12}}{E_{11}+E_{22}-2E_{12}}=\frac{\sum_{t=1}^{i}e_{2t}^2-\sum_{t=1}^{i}e_{1t}e_{2t}}{\sum_{t=1}^{i}e_{1t}^2+\sum_{t=1}^{i}e_{2t}^2-2\sum_{t=1}^{i}e_{1t}e_{2t}}$$

5.2.4. Solving the model

The annual average temperature from 1840 to 2013 was selected for prediction, with the first 80% of the data as the training set and the last 20% of the data as the test set, and the following are the model prediction results, as shown in Figure 7: ARIMA fit and difference data

$$\lambda_2 = \frac{E_{11} - E_{12}}{E_{11} + E_{22} - 2E_{12}} = \frac{\sum_{t=1}^{i} e_{1t}^2 - \sum_{t=1}^{i} e_{1t}e_{2t}}{\sum_{t=1}^{i} e_{1t}^2 + \sum_{t=1}^{i} e_{2t}^2 - 2\sum_{t=1}^{i} e_{1t}e_{2t}}$$

② Prediction error sum of squares solution

$$S = \frac{E_{11}E_{22} - E_{12}^{\ 2}}{E_{11} + E_{22} - 2E_{12}} = \frac{\sum_{t=1}^{i} e_{1t}^2 \sum_{t=1}^{i} e_{2t}^2 - \left(\sum_{t=1}^{i} e_{1t}e_{2t}\right)^2}{\sum_{t=1}^{i} e_{1t}^2 + \sum_{t=1}^{i} e_{2t}^2 - 2\sum_{t=1}^{i} e_{1t}e_{2t}}$$

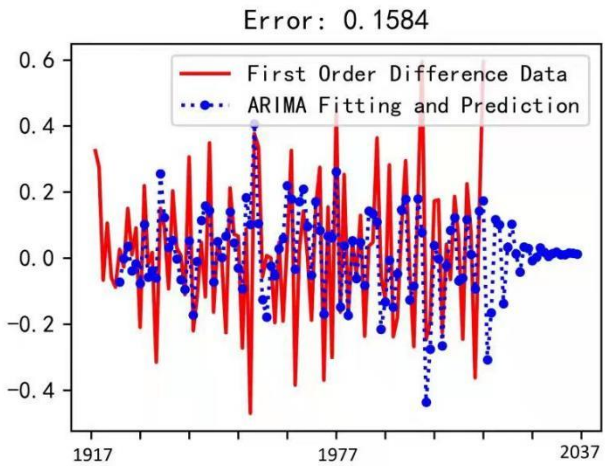

Fig. 8: ARIME fit and differential data

As shown in Figure 8: Trend of autocorrelation and partial autocorrelation

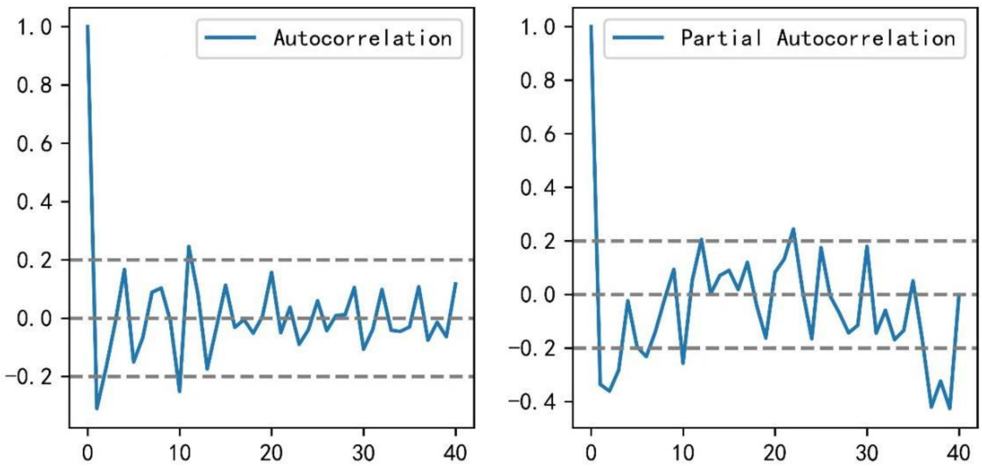

Fig. 9: Autocorrelation and partial autocorrelation

Based on the compiled global average annual temperature, the following ARIMA model predictions are obtained by taking the model into solution

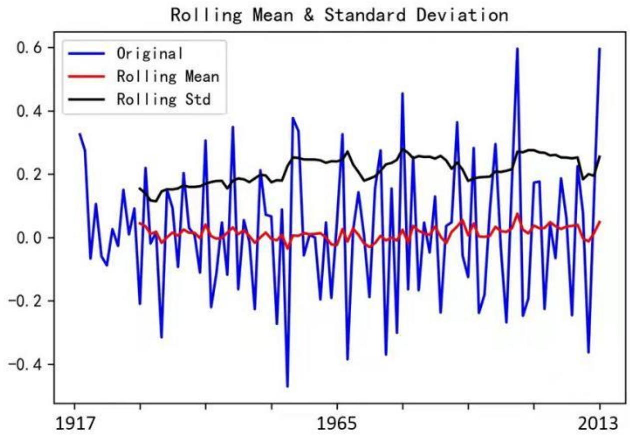

Fig. 10: ARIMA model results prediction

Based on the results of the model, we perform an error test and find MSE=10.12 and MASE=7.24%. This indicates that the combined ARIMA-STLM model predicts better and has less error. The model prediction to reach a global average annual temperature of 20.00°C in 2070 compared with model 1 shows that the ARIMA-LSTM model prediction is more accurate

5.3. Analysis and Solution of Problem II a), b) and c)

5.3.1. Modeling

Develop a mathematical model to analyze the relationship between global temperature, time, and location. Collect relevant data and analyze natural

The impact of natural disasters (e.g., volcanic eruptions, forest fires, and Neoconiosis) on global temperature. Find out the main factors affecting global temperature change

The purpose of this study is to identify the main causes of global temperature change and to propose measures to curb global warming. First, global temperature is treated as the dependent variable, and time and location are treated as independent variables.

We also collect data on volcanic eruptions, forest fires and new crown epidemics, and further refine the global temperature equation through variable selection and stepwise

We also collect data on volcanic eruptions, forest fires and the New Crown epidemic, and further improve the multiple linear regression model of global temperature through variable selection and stepwise regression to obtain their effects on global temperature.

Analysis and Solution of Problem II a), b) and c)

5.3.2. Solving the problem

Model III - Multiple regression-based model of global temperature influencing factors

(1) Preparation of the model

The global temperature is denoted as y, the time is denoted as x1, and the location is denoted as x2. Obviously, x1 and x2 are independent of each other, and then the following

Multiple linear regression model.

$$y = b_0 + b_1 x_1 + b_2 x_2 + \varepsilon$$

(2) Solution of the model

The data of the variables are imported into STATA14.0 software, and the following regression results are obtained in Table 3.

Table 3: Calculation results of the model

Parameter	Estimated value	t	p	F test P	GFT
b_0	71.53	131.2	0	0	0.269
b_1	0.317	2.69	0.007	0	0.269
b_2	-0.3	-1.2	0.132	0	0.269

The estimate of y is obtained as $y = b_0 + b_1 x_1 + b_2 x_2 + \varepsilon$.

This shows that time is positively correlated with global temperature and location is negatively correlated with global temperature, the larger the dimension, the lower the temperature.

Consider the effect of volcanic eruption, forest fire and new crown epidemic on temperature, and note that volcanic eruption is x3, forest fire is x4 and new crown epidemic is x5, and define.

$$x_3 = \begin{cases} 1, \text{exit} \\ 0, \text{ not} \end{cases}, \quad x_4 = \begin{cases} 1, \text{exit} \\ 0, \text{ not} \end{cases}, \quad x_5 = \begin{cases} 1, \text{exit} \\ 0, \text{ not} \end{cases}$$

The variable data were again imported into STATA 14.0 software and the following regression results were obtained in Table 4 below.

Table 4： data

Parameter	Estimated value	Cofidence interval
b_0	110	[101,131]
b_1	42	[32,50]
b_2	67	[51,77]
b_3	21	[7,32]
b_4	14	[7,22]
b_5	11	[1,31]

From this, the estimate of y is obtained as

$$y = 110 + 42x_1 + 67x_2 + 21x_3 + 14x_4 + 11x_5$$

It shows that volcanic eruptions, forest fires and the New Crown epidemic all have effects on global temperature, and volcanic eruptions > forest fires > New Crown epidemic have effects on global mean annual temperature.

6. Model evaluation and improvement

6.1. Advantages of the model.

ARIMA has better linear trend prediction with fast convergence and robustness, and LSTM, as a new prediction method, has the ability to memorize patterns for a long time, and its advanced structure relies on feedback connections, which is capable of long and short term memory of global annual average temperature time series characteristics, ARIAM-LSTM combines the advantages of both and has better stability.

6.2. Disadvantages of the model.

The gray GM (1, 1) forecasting model is suitable for a small amount of forecasting data, while the data volume in this paper is large.

6.3. Model improvements.

Model predictions are not only related to the model itself, but also to the data. There are many other factors that affect global temperature, and more comprehensive data can make global temperature calculations more accurate.

7. Causal analysis and suggestions

7.1. Analysis of the causes of global temperature change

According to the data we have compiled and its prediction results, we can see that the annual average temperature of the world is rising seriously.

The causes are the combined effect of volcanic eruptions, new crown epidemics, and forest fires, and the relationship between temperature change, geographic location, and time change:

A significant positive spatial correlation between changes in atmospheric pollutant concentrations and global temperature

Emissions of atmospheric pollutants lead to an increase in carbon dioxide and other gases. The atmospheric greenhouse effect is the effect of atmospheric substances on the near-Earth atmosphere, i.e., the increase of warming substances such as carbon dioxide in the atmosphere will make it possible to block more long-wave radiation energy expenditures from the near-Earth atmosphere and the ground to the cosmic space, thus causing global warming.

The thermal energy released from energy sources that are not fully utilized contributes to global warming to some extent

Since the heat energy cannot be fully utilized, the heat energy in the atmosphere will become unrecoverable cosmic radiation and electromagnetic waves scattered into space, the heat energy is not transferred to other places, the heat energy is still in the universe as an isolated thermodynamic system, which accelerates the entropy increase of the system, thus increasing the temperature.

Geological activities have a great influence on global temperature

The most obvious impact of geological activities on global temperature is volcanic eruptions, which send large amounts of sulfur dioxide and hydrogen sulfide into the atmosphere. This aerosol creates a "parasol effect" that helps the Earth reflect sunlight away from the surface, cooling it.

Ocean activity has an impact on global temperatures

Tidal changes in the oceans cause cold water to rise, which lowers the temperature of the oceans and further lowers the global temperature. Secondly, because the oceans are a huge heat capacity, the Earth's temperature does not rise too fast. Finally, the oceans provide a huge living place for a large number of organisms, including algae that carry out photosynthesis to convert large amounts of carbon dioxide, the most important place on Earth for producing oxygen.

Temperature changes are also closely related to solar activity

According to current research, solar activity may affect Earth's temperature through a variety of pathways, and there is a strong correlation between solar activity and global temperature, although it may be unrelated or even negative some of the time.

7.2. Recommendations and Measures

The global trend of global warming has become more and more obvious, and the impact on the Earth and on human beings is becoming more and more serious, if we do not take measures now to implement effective and

If we do not take measures now to implement effective and long-term solutions, the life of the earth and the survival of human beings will be greatly threatened, and we will give the following suggestions and measures:

Adjust the energy structure and develop clean energy such as hydropower, photoelectricity, and wind power. Human activities use a lot of fossil energy, which releases a lot of heat during the use of fossil energy and causes the global temperature to rise. The clean energy such as hydropower and photoelectricity is essentially a part of the solar energy on the surface of the earth, which is always in the system of the earth's surface and circulates within the system, transforming from one form to another without increasing the heat load on the surface.

The photosynthesis of green plants is a large-scale heat-absorbing reaction, and the burning of fossil energy is a process of converting organic matter into inorganic matter.

The opposite is true for photosynthesis, which converts inorganic material into organic material, absorbing solar energy and consolidating it, and they have an unparalleled role in curbing global temperature growth.

Promote convection and infrastructural water storage

The temperature of latitudinal humidity is significantly lower than that of deserts at the same latitude because water vapor can carry surface heat to higher altitudes. In recent years, due to increasing urbanization, many surfaces have been hardened extensively and heat is left on the ground. Therefore, collecting and purifying rainwater and other water on the ground is a means of cooling the ground, and it is necessary to form an efficient municipal pipe network. Large-scale water pools or water towers can be built to spray the sun when necessary.

Large-scale industrial thermal conversion Converting surface heat into chemical energy through large-scale industrial models, electrolysis of water. The products of electrolysis of water are oxygen and hydrogen.

The use of light to electrolyze water essentially converts surface heat energy into chemical energy in the form of hydrogen, after which the product hydrogen is released, and because of its extremely low density, it is emitted high into space. This energy is dissipated at the surface, which has a positive effect on curbing global warming.

8. Future development and suggestions

Through our team's analysis of the global temperature dataset and the related content collected, we found that the global temperature is on the rise. This increase in global temperature has the potential to have a significant impact on the future. There are many consequences of global temperature increase, such as the melting of glaciers, which will cause the sea level to rise, thus endangering the safety of the oceans; it will affect the safety of the oceans by causing the acidification of sea water due to the emission of large amounts of carbon dioxide, which will affect marine biodiversity; it will also increase the humidity of the air, which will increase the

ISSN: 1813-4890

amount of rainfall, thus aggravating the haze phenomenon; and it will even lead to climate disruption and frequent extreme weather.

This will lead to climate disruptions and frequent occurrence of extreme weather. In order to reduce the damage caused by global warming, we suggest controlling greenhouse gas emissions, adjusting the energy consumption structure, adopting new or clean energy sources, and planting forests to increase the global forest cover.

Acknowledgements

Natural Science Foundation.

References

[1] Qing Pan.Research on prediction model of industrial pollution gas concentration based on deep learning and time series.Zhejiang Gongshang University,2022.

[2] Yating Gao.Research on Time Series Trend Prediction Method and its Application to Stock Market.Shandong Technology and Business University,2022.

[3] Hanchao Zhu.GOLD PRICE FORECAST BASED ON GREY NEURAL NETWORK MODEL.Shanghai Jiao Tong University,2018.

[4] Zhou Yin-ming.Prediction of PM2.5 concentration based on Grey Theory and support vector machine.ZhejiangA&FUniversity,2018.

[5] Dong Yun Cheng.Research on the combined forecast model of urban water supply based on ARIMA-LSTM.Kunming University of Science and Technology,2021.

[6] Song Han. Research on Prediction Model of Wall Thickness of Natural Gas Storage Pipeline Based on Grey Theory and Neural Network.China University of Petroleum,2020.

[7] Jin Yongchao;Wang Renfang;Zhuang Xiaodie;Wang Kenan;Wang Honglian;Wang Chenxi;Wang Xiyin.Prediction of COVID-19 Data Using an ARIMA-LSTM Hybrid Forecast Model.Mathematics, 2022,10(21):4001-4001.

[8] Rasaki O.Olanrewaju;Mamadou A.Jallow;Sodiq A.Olanrewaju.An Analysis of the Atlantic Ocean Wave Via Random Cosine and Sine Alternate Wavy ARIMA Functions.International Journal of Intelligent Systems and Applications,2022,14(5):22-34.

[9] Tang Xiaoyu;Xu Sijia;Ye Hui.The Way to Invest: Trading Strategies Based on ARIMA and Investor Personality.Symmetry,2022,14(11):2292.

[10] Lu Haoyang;Tong Jing;Tang Yajiao.Analysis of Green Credit and the Ecological Welfare Performance Based on Empirical Models and ARIMA(2,3,2):Taking China as an Example. Sustainability, 2022,14(19),119.