

abstract

apmcmthesis \LaTeX template is designed by <https://www.latexstudio.net> for <http://www.apmcm.org>. The template is designed to let everyone focus on the content writing of the paper, without spending too much effort on the customization and adjustment of the format.

Note that users need to have some experience with \LaTeX , at least some of the features of common macro packages, such as references, math formulas, image usage, list environment, etc. Templates have added commonly used macros. Package, no additional user added.

This template is located on <https://github.com/latexstudio/APMCMthesis>. You can update files from the repository.

Keywords: Keywords1 Keywords2 Keywords3

Contents

1. Introduction	1
2. Question restatement	1
3. Terms, definitions and symbols	2
4. Models assumptions	2
5. Solution for question 1	2
5.1 (a)	2
5.1.1 <i>Analysis of problem 1(a)</i>	2
5.1.2 <i>Solution of problem</i>	2
5.2 (b)	4
5.2.1 <i>Analysis of problem 1(b)</i>	4
5.2.2 <i>Model 1: SARIMA</i>	4
5.2.3 <i>Model 2:</i>	7
5.2.4 <i>Description of global temperature in recent 100 years</i>	7
5.3 (c)	7
5.4 (d)	7
6. Solution for question 2	7
6.1 (a)	7
6.1.1 <i>Analysis of problem 2(a)</i>	7
6.1.2 <i>Test of Pearson correlation coefficient</i>	8
6.1.3 <i>Test of Grey relation analysis</i>	8
6.2 (b)	8
6.2.1 <i>Data of natural disasters</i>	8
6.3 (c)	8
6.3.1 <i>Data of more factor</i>	8
6.3.2	8
6.4 (d)	8
7. Solution for question 3	9
8. Future Work	9
8.1 Another model	9
8.1.1 <i>The limitations of queuing theory</i>	9
8.1.2	9
8.1.3	9
8.1.4	9
9. References	9
10. Appendix	10

I. Introduction

In order to indicate the origin of problems, the following background is worth mentioning.

Since this year, we have seen a large number of amazing temperature reports. There is no doubt that the earth is burning. Following the terrible high temperature in these areas from the end of June to the beginning of July, Italy set a European temperature record again, reaching an astonishing 48.8 ° C, and many countries declared a state of emergency.

Global warming is a phenomenon related to nature. It is precisely because of the continuous accumulation of the greenhouse effect that the energy absorbed and emitted by the Earth's atmospheric system is unbalanced. The continuous accumulation of the energy of the Earth's atmospheric system leads to temperature rise and global warming.

Before the industrial revolution, the carbon dioxide (CO₂) in the atmosphere was about 280 parts per million. In March 2004, the atmospheric CO₂ concentration reached 377.7ppm, the largest average increase in 10 years so far. According to the statistics of scientists from the National Oceanic and Atmospheric Administration (NOAA) and the Scripps Institution of Oceanography (SIO), the monthly average concentration of CO₂ reached a peak of 421ppm in May 2022. An OECD report predicted that the CO₂ concentration would reach 685ppm by 2050.

II. Question restatement

1) Do you agree with the global temperature statement? Data given in attachment and other data sets collected by the team analyze changes in global temperature. Are you suggesting that the rise in global temperature in March 2022 has caused more growth than any decade ago. Based on historical data, please develop two or more mathematical models to show the past and predict the future level of global temperature. Use each model at 1(b) to estimate global temperature in 2050 and 2100. Which model of 1(b) do you think is the right one, and why.

2) What are the causes that affect temperature change? Use the results of question 1 and annex Data. Data from other data sets collected by csv and your team, developed a mathematical model to verify the relationship between global temperature, time and location (whatever), and explain the relationship between them or prove that there is no relationship between them. Collect native data and check actions that have an impact on natural disasters (such as volcanic eruptions, forest fires and COVID-19). Does it affect global temperature? What do you think is the main reason that is affecting global temperature change? Do you think there are steps that can stop or slow global warming?

3) Prepare a non-technical article (mostly on 1 page). Write a non-technical article (mostly on 1 page) to the APMCM Organization Committee to explain your team's findings and suggestions for the future. The PDF solution with a total number of not more than 25 pages contains: a full table of one page; and complete solution catalogue; A non-technical article page.

III. Terms, definitions and symbols

IV. Models assumptions

V. Solution for question 1

In whole question 1, we need to analyze global temperature changes using dataset in the attachment and other datasets we collect.

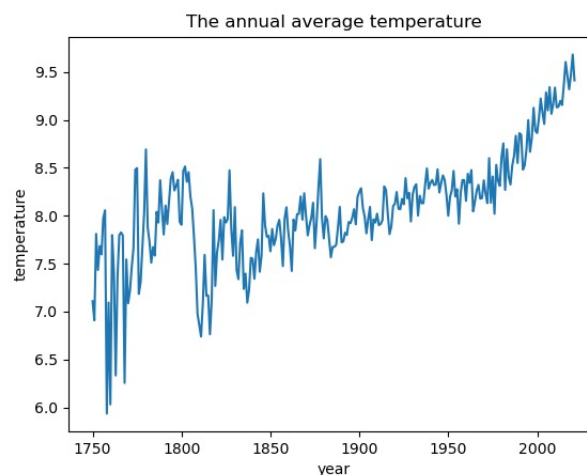
5.1 (a)

5.1.1 *Analysis of problem 1(a)*

In this question, we need to judge if the increase of global temperature in March 2022 resulted in a larger increase than observed over any previous 10-year period. So we can plot the global temperature in 2022. Firstly, consider if global temperature in March 2022 is an abnormal value. Then calculate the average temperature between 2011 to 2021 and its rising rate. We let them be visual and analyze their trend and differences. Finally we conclude a judgement of the statement.

5.1.2 *Solution of problem*

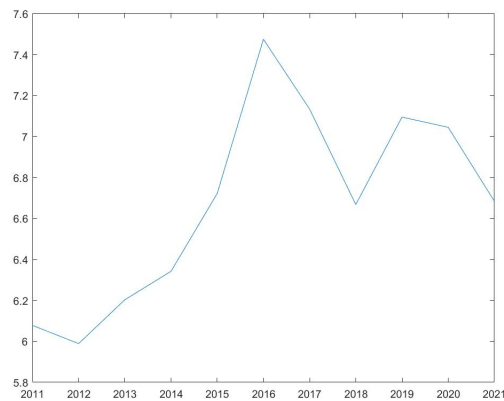
We download global temperature dataset from <https://berkeleyearth.org/data/>. And we choose land only monthly average temperature to plot the figure. The picture about annual average temperature between 1760 to 2022 is the following.



picture 1

Then we get the average temperature in March between 2011 to 2021 is the following.

[6.0780, 5.9890, 6.2030, 6.3420, 6.7220, 7.4750, 7.1340, 6.6680, 7.0950, 7.0450, 6.6840].



picture 2

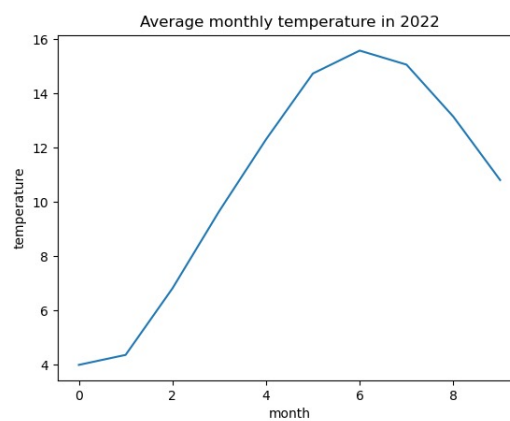
The total increasing rate is 1.100%, and increase 0.606°C in value. Increasing rate respectively is given by

$[-0.0941, -0.0146, 0.0357, 0.0224, 0.0599, 0.1120, -0.0456, -0.0653, 0.0640, -0.0070, -0.0512]$.

Past 10 years average temperature is 6.6759°C .

And the monthly average temperature between January 2022 to October 2022 except March we can calculate is

$[3.989, 4.358, 6.812, 9.665, 12.313, 14.742, 15.588, 15.07, 13.151, 10.811]$.



picture 3

Especially, the increasing rate between 2022 to 2021 is 0.0192. This is much smaller than previous 10 years.

5.2 (b)

5.2.1 Analysis of problem 1(b)

Question (b) require us to build 2 or more mathematical models to describe the past and predict the future global temperature level. We can pick a observation as example to analyze the data annually. That is we build models using yearly average temperature. We decide to use Autoregressive integrated moving average (ARIMA) model and Gray forecast model(GFM) to describe the global temperature.

5.2.2 Model 1: SARIMA

ARIMA model, differential integrated moving average autoregression model, also known as integrated moving average autoregression model, is one of the time series prediction and analysis methods. In ARIMA (p, d, q), AR is 'autoregressive', and p is the number of autoregressive terms; MA is the 'moving average', q is the number of moving average terms, and d is the number of differences (orders) made to make it a stationary sequence. This model is built to describe the relationship between current value and historical value using the historical time data of the variable itself to predict itself. So it fit to the question in 1(b).

Firstly, we need to determine the order p which means the lagged order i.e how many previous variable related to the current value. This leads an autoregressive model(AR):

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \epsilon_t$$

where y_t is current value, μ is constant.

Then we consider more about the error along the time series. We get the moving average model(MA):

$$y_t = \mu + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \epsilon_t$$

This can reduce the random fluctuation when estimating.

Finally, we combine previous 2 models then obtain the model ARMA

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \epsilon_t$$

But this model has some limitations. The results that ARMA obtains might not be stationary, then some test will not be credible. So we can add a difference order d to let them be stationary.

We load the dataset 'land year std data'

Before we obtain the expression of ARIMA, we need to determine the value of p, q and d. In general, we choose low value of d i.e low order for the model. So it is important to find a reasonable value of p, q. We can check the Akaike information criterion(AIC) and Schwarz criterion(SC) given

by respectively:

$$AIC = \ln\left(\frac{SSE}{T}\right) + \frac{2K}{T}, \quad K = p + q + 2.$$

$$SC = \ln\left(\frac{SSE}{T}\right) + \frac{K \ln(T)}{T}.$$

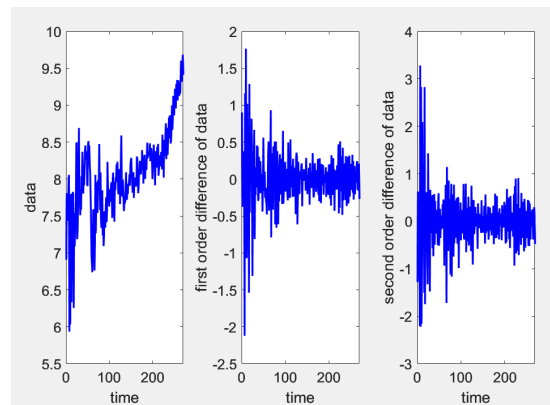
If AIC and SC are small enough, then the choices of p and q are preferred. So we can use the force algorithm to find the best combination of p, q.

Moreover we need to check which model we should use to estimate the data. We can check the autocorrelation function(ACF):

$$ACF(k) = \rho_k = \frac{Cov(y_t, y_{t-k})}{Var(y_t)}.$$

k is the lagged index. We calculate the statistic ACF and plot a scatter so that we can really determine which model is the better one.

After we get the annual average temperature, we first perform a stationarity check on the data. The most intuitive way is to use the data to make a sequence diagram and then observe the trend of the image. At the same time, we calculate the first-order difference and the second-order difference of the data and make a graph to observe the trend.



picture 4

It is not obvious from the image to confirm the stationary relationship of the data, so we use the mathematical test method——ADF test and KPSS test.

In the ADF test, if the return value is 0, the data is not stationary, and if the return value is 1, the data is stationary. Then in KPSS test, if the return value is 0, the data is stationary, and if the return value is 1, the data is not stationary.

From the test result we can know the raw data is not stationary, so we need to calculate the first order difference of data, then we do the same test to the first order difference of data. We can get

Therefore, the data after the first order difference can be used for time series modeling analysis. And then what we need to do is to confirm the parameters p q d. The first step is to examine the autocorrelation and partial autocorrelation plots of a stationary time series.

We can't get the values of p and q from the figure, so we adopt the AIC and BIC criteria to find the most suitable p and q values within a certain range. When we have confirmed the appropriate

```
d_adf =
    logical
    0

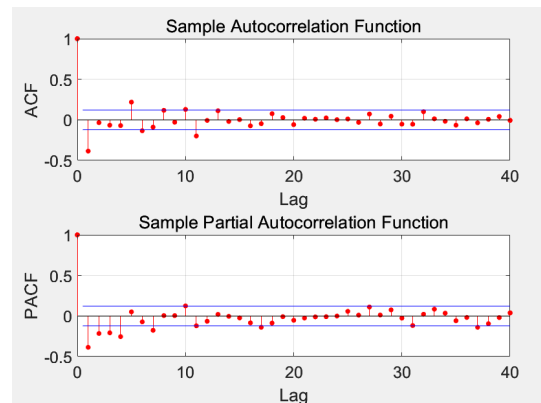
d_kpss =
    logical
    1
```

picture 5

```
d1_adf =
    logical
    1

d1_kpss =
    logical
    0
```

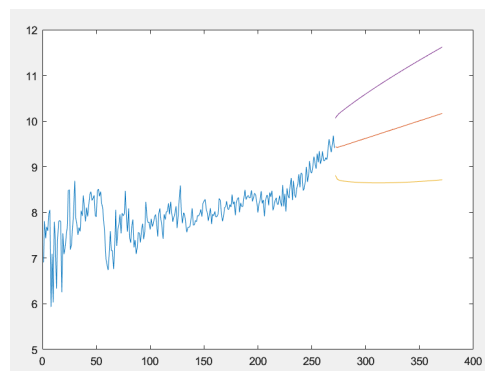
picture 6



picture 7

p q d value, we can build a model to predict the subsequent data. In the process of forecasting, we also set the upper and lower forecast limits under the 95

First, let's predict the annual average temperature for the next 100 years.

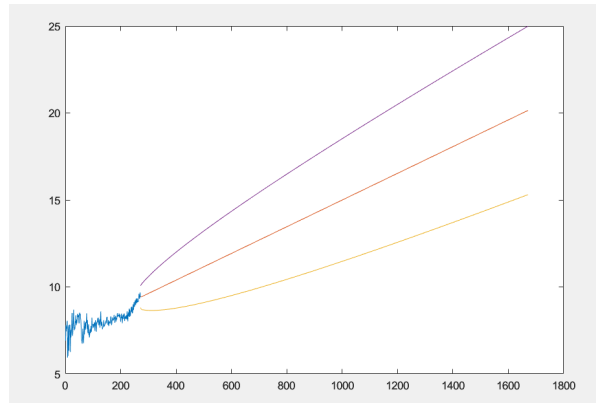


picture 8

Then, we choose seasonal ARIMA to estimate the average global temperature. This is because we obtain only stable prediction by ARIMA, while ARIMA cannot estimate seasonal data well. So we do some extension to ARIMA, add a seasonal factor to ARIMA. That is:

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \sum_{i=1}^p \Gamma_i y_{t-is} + \sum_{i=1}^q \theta_i \epsilon_{t_i} + \sum_{i=1}^q \Theta_i \epsilon_{t-is} + \epsilon_t$$

We can find that it is the multiple of ARIMA and seasonal expression $s(p, q, d)$. This step does not be like obtaining ARIMA from ARMA, which just adds a differential operator d. It multiply



picture 9

seasonal factors to ARIMA to get SARIMA.

5.2.3 Model 2:

5.2.4 Description of global temperature in recent 100 years

Using the models that we build in previous section, we draw the picture of past 50 years and future 50 years global temperature level.

5.3 (c)

According to the model in question 1(b), we can easily predict global temperatures in 2050 and 2100 respectively using different models. Moreover we need to find when the average temperature of observation points in our prediction models will reach 20.00°C .

5.4 (d)

We need to calculate important statistics to determine which model is better. We compare SE, MSE and RMSE and draw a figure as the following.

VI. Solution for question 2

In question 2, we are required to build a global temperature model not only including the value of temperature but also more other reasons such as position, time, volcanic eruptions, forest fires and the COVID-19 etc.

6.1 (a)

6.1.1 Analysis of problem 2(a)

According to the data given in attachment 1, we analyze the correlation between global temperature, time and location. Draw a heat map, and use ridge regression in multiple linear regression to describe the relationship between variables above. We decide to use Pearson correlation coefficient to describe the relationship between temperature, time and location. This is because the values of

them do affect each other, and the order of temperature is an important factor that judge the relationship. Moreover, time temperature and location must satisfy some linear relationship, so we decide to use Pearson correlation coefficient to describe.

And we want to analyze further to prove the correctness of the relationship. So we use Grey relation analysis(GRA) to make sure that they really have some relationship.

6.1.2 Test of Pearson correlation coefficient

6.1.3 Test of Grey relation analysis

6.2 (b)

6.2.1 Data of natural disasters

In this section, we should get data of the natural disasters, we find part of data(such as volcanic eruptions, forest fires) in <https://www.emdat.be/>, and part of data(COVID-19) in <https://covid19.who.int/data>.

6.3 (c)

6.3.1 Data of more factor

We find datasets of world energy, world population and Global forest resource. <https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy/downloads.html> <https://population.un.org/wpp/Download/Standard/MostUsed/> <https://fra-data.fao.org/W0/fra2020/dataDownload/>

6.3.2

6.4 (d)

VII. Solution for question 3

VIII. Future Work

8.1 Another model

8.1.1 *The limitations of queuing theory*

8.1.2

8.1.3

8.1.4

IX. References

- [1] Author, Title, Place of Publication: Press, Year of publication.
- [2] author, paper name, magazine name, volume number: starting and ending page number, year of publication.
- [3] author, resource title, web site, visit time (year, month, day).
- [4] L^AT_EX 资源和技巧学习 <https://www.latexstudio.net>
- [5] L^AT_EX 问题交流网站 <https://wenda.latexstudio.net>
- [6] 模板库维护 <https://github.com/latexstudio/APMCMThesis>

X. Appendix

Listing 1: The matlab Source code of Algorithm

```

kk=2; [mdd, ndd]=size(dd);
while ~isempty(V)
    [tmpd, j]=min(W(i, V)); tmpj=V(j);
    for k=2:ndd
        [tmp1, jj]=min(dd(1, k)+W(dd(2, k), V));
        tmp2=V(jj); tt(k-1, :)= [tmp1, tmp2, jj];
    end
    tmp=[tmpd, tmpj, j; tt]; [tmp3, tmp4]=min(tmp(:, 1));
    if tmp3==tmpd, ss(1:2, kk)=[i; tmp(tmp4, 2)];
    else, tmp5=find(ss(:, tmp4)~=0); tmp6=length(tmp5);
    if dd(2, tmp4)==ss(tmp6, tmp4)
        ss(1:tmp6+1, kk)=[ss(tmp5, tmp4); tmp(tmp4, 2)];
    else, ss(1:3, kk)=[i; dd(2, tmp4); tmp(tmp4, 2)];
    end; end
    dd=[dd, [tmp3; tmp(tmp4, 2)]]; V(tmp(tmp4, 3))=[];
    [mdd, ndd]=size(dd); kk=kk+1;
end; S=ss; D=dd(1, :);

```

Listing 2: The lingo source code

```

kk=2;
[mdd, ndd]=size(dd);
while ~isempty(V)
    [tmpd, j]=min(W(i, V)); tmpj=V(j);
    for k=2:ndd
        [tmp1, jj]=min(dd(1, k)+W(dd(2, k), V));
        tmp2=V(jj); tt(k-1, :)= [tmp1, tmp2, jj];
    end
    tmp=[tmpd, tmpj, j; tt]; [tmp3, tmp4]=min(tmp(:, 1));
    if tmp3==tmpd, ss(1:2, kk)=[i; tmp(tmp4, 2)];
    else, tmp5=find(ss(:, tmp4)~=0); tmp6=length(tmp5);
    if dd(2, tmp4)==ss(tmp6, tmp4)
        ss(1:tmp6+1, kk)=[ss(tmp5, tmp4); tmp(tmp4, 2)];
    else, ss(1:3, kk)=[i; dd(2, tmp4); tmp(tmp4, 2)];
    end;
end
dd=[dd, [tmp3; tmp(tmp4, 2)]]; V(tmp(tmp4, 3))=[];
[mdd, ndd]=size(dd);
kk=kk+1;
end;
S=ss;
D=dd(1, :);

```