Report I of Advance Signal Processing

Name: Liu Zihua

**I. Dataset Selection and Introduction**

图表

中度可信度描述已自动生成 In this report, I selected Hakodate, Otaru, and Osaka tide data in March 2021 from the historical tide data provided by the Japan Meteorological Administration[1] as the research objects for analysis. Among them, Hakodate and Otaru belong to the Hokkaido region, and their data can reflect to a certain extent the tidal activity patterns in the waters around Hokkaido in March. Osaka belongs to the Kansai region and is a certain distance away from Hokkaido. The reason for choosing Osaka is to explore the similarities and differences between Osaka and Hokkaido to extract the tidal motion laws.

The Overall data of the three city can be seen in the following fig.(Fig1)

Fig1. The overall data of Hakotade, Otaru and Osaka

In this report, the following data and signal processing methods are used:

* Autocorrelation analysis
* Partial autocorrelation analysis
* Sliding smoothing
* Smoothness detection
* ARIMA time series model construction and forecast
* Amplitude spectrum calculation
* Power spectrum calculation
* Cross-correlation calculation
* R value calculation

All data processing is performed by Python's third-party libraries: numpy, statsmodel, matlibpyplot for calculation and image drawing. The relevant code has been uploaded to Github and added to this report as an attachment of .zip file.

Github repo url: <https://github.com/Magicboomliu/Japan-Tide-analysis-forcast>

**Question 1**

1. **ACF and PACF**
2. Introduction

图表, 折线图

描述已自动生成Here the data I choose to be processed and analyzed is the tide data of Hatakode( A port city in Hokkaido) in March 2021. Firstly, I preprocessed the data by removing the bias and trend of the data for better statistics, Then I applied moving average( Simple Moving average, Cumulative Moving Average, Exponential Moving Average), as you can see in fig1 to make the data smoothly. And use the simple moving average data(the yellow one) as the data sources.

Fig2. Original data and after moving averaged Data

For the time lag of the data, I choose range(0-30) to discover the laws of the data,

and the same processing method is also applied to the PACF diagram.

图形用户界面

描述已自动生成Next Figs show the auto correlation graph and partial correlation graph:

图表

描述已自动生成图表, 直方图

描述已自动生成

**Analysis:**

Firstly, as we can see in the time interval is in a integer of multiple of **24** hours, the position can reach to a peak, which means to have the highest auto correlation.And we can also see that when the time represented by the x-axe is a multiple of **12**, the autocorrelation is also a pretty high positive value.

The reason why this phenomenon occurs can be inferred like this : It indicates that the tidal fluctuations at a fixed time point every day (24 hours) are highly positively correlated. In addition, the observational data every half day(12 hours) also has a high positive correlation. This can be inferred that the waves of tidal fluctuations may be caused by it may consists of periodic waves with a period of 12 hours and 24 hours.

Second, we can also see in the Autocorrelation graph is that when the time integer of multiple of **6** hours(but not the multiple of 12 hours), and auto-correlation value trends to be a very high negative value. The same phenomenon also occurs at the time interval of every **18** hours.

The reason why this phenomenon occurs can be inferred like this: It indicates that when we observe the data in a time of 1/4 day(6 hours), the tide data changes randomly without have a relationship with the former data. This can be a part of the cycle data.

1. **Amplitude Spectrum and Power Spectrum**

To calculate the amplitude and power spectrum of the tide data of Hakotade, I use the python numpy API to calculate the Fast Fourier Transformation (FFT), and then use the absolute value of the data as the shown amplitude spectrum data. The data sampled is setted into 3600s and set DC component to 0hz.

**图表, 直方图

描述已自动生成**And the result can be seen as follows:

1. **(b) (c)**

Fig.4 (a)Original Hakotade Tide data (b) Amplitude Spectrum (c) The power spectrum of Hakotade.

**Analysis:**

As we can see in a the amplitude spectrum graph, we can see that the components of the tides wave have some main frequencies: they are 2.20e-5Hz, 2.30e-5Hz,1.12e-5Hz,and 1,16e-5Hz, If we change the frequency into the time domine, the corresponding time period is 12.32 Hour, 12.01 Hour, 25.45 Hour and 24.02 Hour. And the strong peak is in 2.30e-5Hz, which is 12.01 Hour in time domine, the actual tidal period is between these two numbers and is closer to 12.01 hours, which also explains why the peaks and troughs gradually shift away from the integer multiple of 6. This result is also consistent with the analysis result of ACF which I analyzed above.

Finally , I also train a ARIMA Model to fit the following laws I mentioned above by doing one order differential to the data. Related codes and model can be see on my github repo: <https://github.com/Magicboomliu/Japan-Tide-analysis-forcast>.

**Question 2**

In order to discover the relationship between different regions in Japan.

I mainly selected 3 regions:

|  |  |  |
| --- | --- | --- |
| **City** | **Region** | **Date** |
| Hakotade | Hokkaido | 2021/04/01(whole March) |
| Otaru | Hokkaido | 2021/04/01(whole March) |
| Osaka | Kansei | 2021/04/01(whole March) |

Table 1. Basic Information

图表, 折线图

描述已自动生成First let us see the cross correlation graph between the Hakotade and Otaru( Both are Cities in Hokkaio). Because the time ranges is too big, I just select the lags to 96, which means (0-96 hours) to see the relation between two data.

1. (b) (c)

Fig 5 : (a)Original data of Hakotade, (b) Original Data of Otaru (c) Cross Correlation of Hakotade and Otaru

**Analysis:**

By looking at the cross correlation map , we can see that when the time interval is multiple at 12 or 24, the correlation value of Hakotade and Otaru is a high positive value. And in the time interval of multiple 6 or 18, Consider that the auto-correlation map of Hakotade is also have high auto correlation value at 12 and 24, and low in 6 or 18, they are likely the same. In can inferred that the Hakotade and Otaru have the similar tide waves laws. It is also reasonable because they are both in the region of Hakkaido.

图表, 折线图

描述已自动生成Second let us see the cross correlation graph between the Hakotade and Osaka(One in Hokkaido, One is the Kansei region). Because the time ranges is too big, I just select the lags to 96, which means (0-96 hours) to see the relation between two data.

1. (b) (c)

Fig 6 : (a)Original data of Hakotade, (b) Original Data of Osaka (c) Cross Correlation of Hakotade and Osaka

**Analysis:**

By looking at the cross correlation map , we can see that when the time interval is multiple at 14 or 28, the correlation value of HAKOTADE and OSAKA have very high negative value. And when the time interval is at the multiple of 8 or 20 hours, the correlation value of HAKOTADE and OSAKA have very high positive value.

This is not same with the auto correlation data of the HAKOTADE itself. This shows that the phases of the tide cycles in the two regions are staggered, and the amount is about half a cycle (that is, about 8 hours). As a result, when the tide in Hakotade reaches the peak, that of Osaka is near the trough, and when the tide in Osaka reaches the peak, that of Hakotade is near the trough. It is very different with the laws of the Hakotade-Otaru. It maybe because that Osaka and Hakotade are in different regions.