TIME SERIES PROJECT: Electricity Consumption in Guangzhou
Summer 2014
Student Name: RUHAI YUAN
Registration ID: 71786771

Introduction

Electricity is a very important resource, especially in the modern society. It is hard for people to live without electricity. Furthermore, in some big cities in China, the development of electricity plays an important role in the economic development. Guangzhou is one of the biggest cities in China, and analyze the electricity consumption of the citizens in Guangzhou is meaningful.

All the raw data used in this project was found at http://www.gzstats.gov.cn. The dataset can be downloaded from http://data.gzstats.gov.cn/gzStat1/chaxun/ydk.jsp, by select the Electricity Consumption of Citizens in Guangzhou series.

analysis

the following graph(Fig.1) shows the plotting of the original time series.

First by looking at Fig.1, we can easily tell that the value of the data grow with the time, the upward trend alone would lead us to specify a non-stationary model. Moreover, its seasonal characteristic is obvious, it has higher records in summer months.

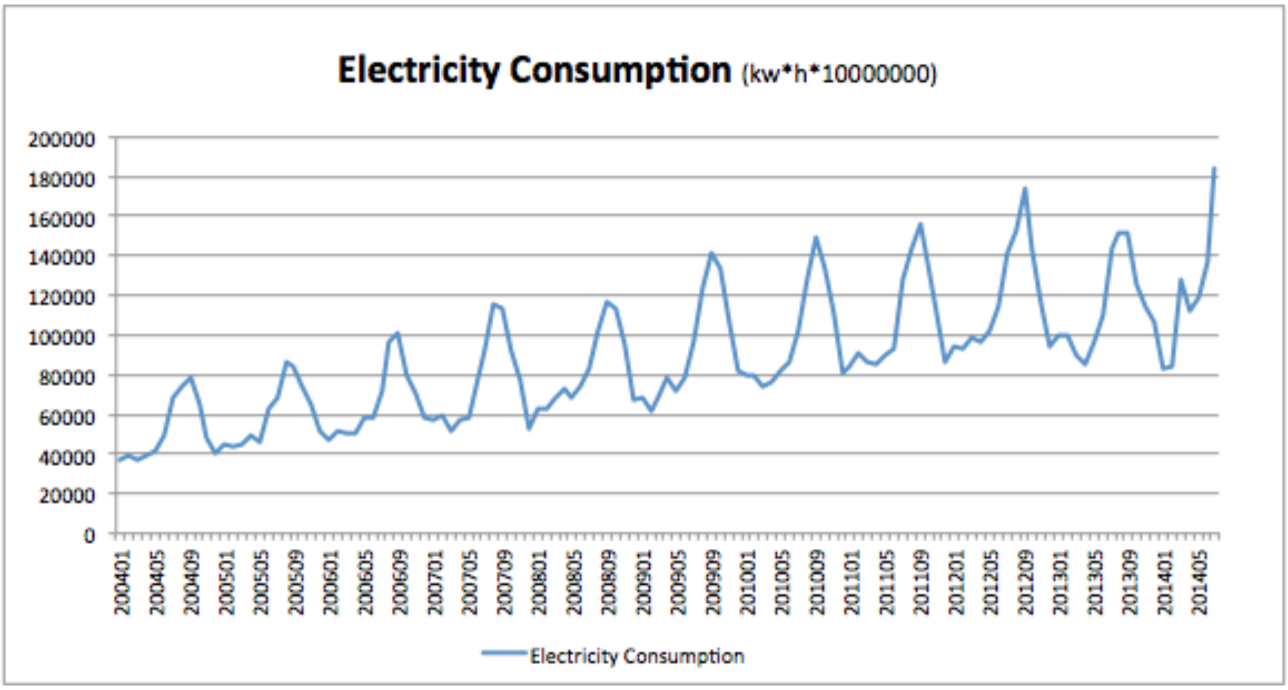

Fig.1

Fig.2 shows the sample autocorrelation function for this series.The seasonal autocorrelation relationships are shown quite prominently in this display. Notice the strong correlation at lags 12, 24, 36, and so on. So, I prefer to chose a seasonal model to fit this series.

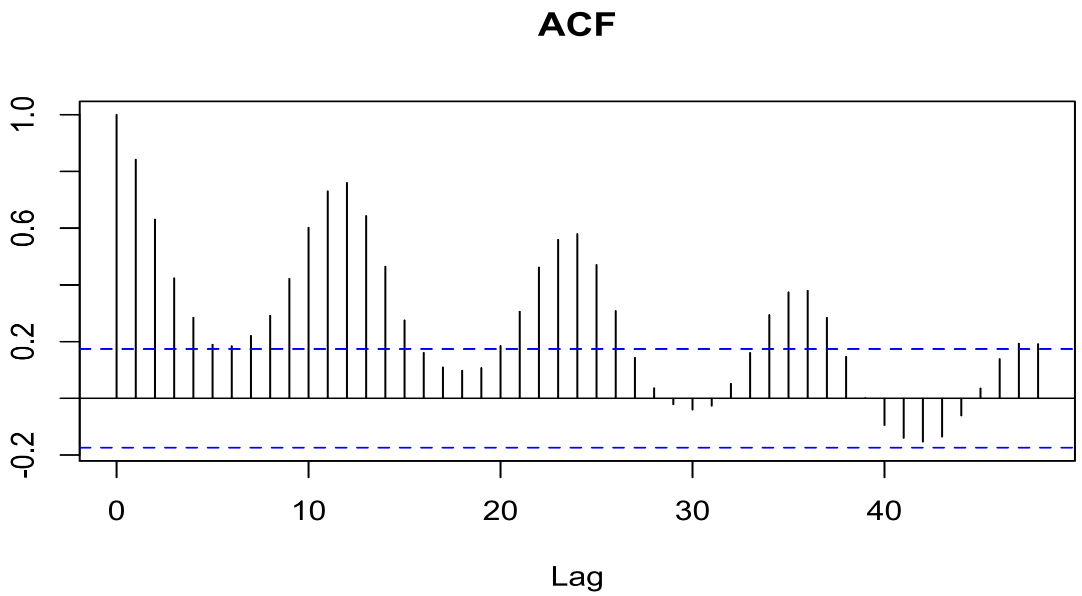

Fig.2

Firstly, we use first difference to eliminate the growing trend. Fig.3 shows the time series plot of the electricity consumption after we take the first difference. The trend almost disappeared.

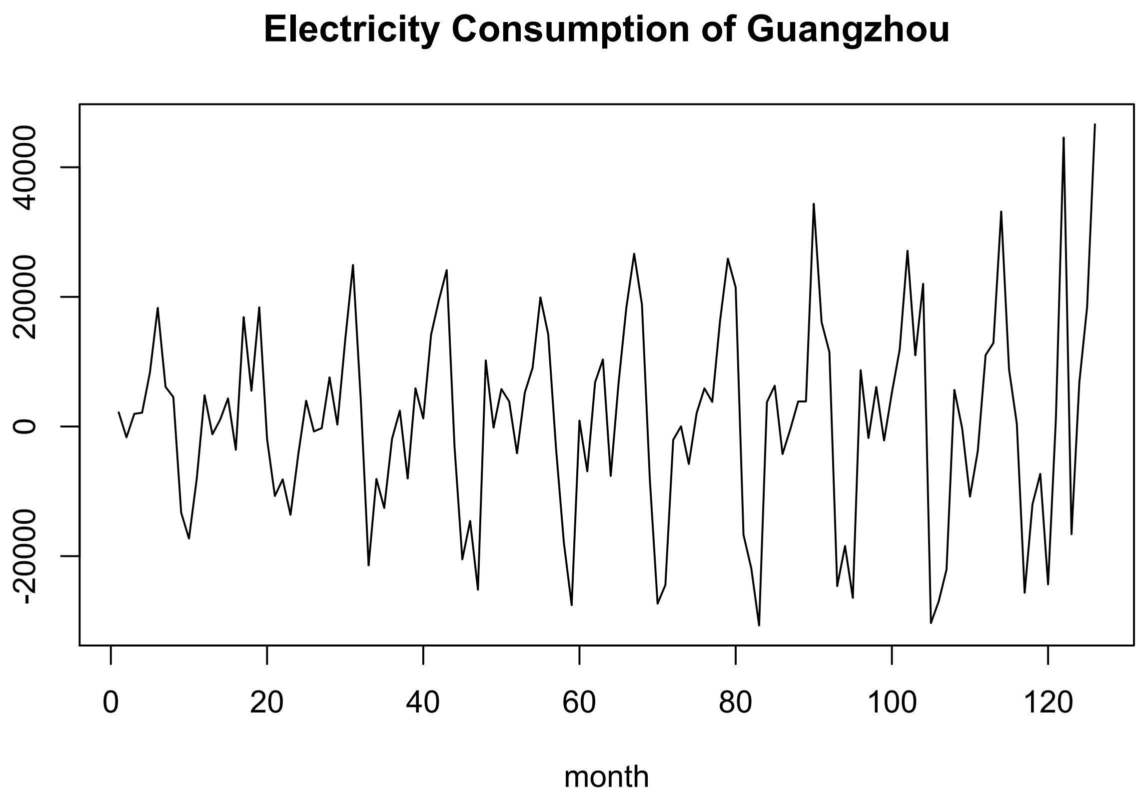

Fig.3

It looks much more stationary than before, but also have strong seasonality. However,we should look at the autocovariance function of the data. Fig.4 shows the ACF of the data after first difference. The positive seasonality is more obvious. There are strong correlation at every 12 lags, and weaker correlation next to the 12, 24, 36 lags. After all, this time series are still non-stationary.

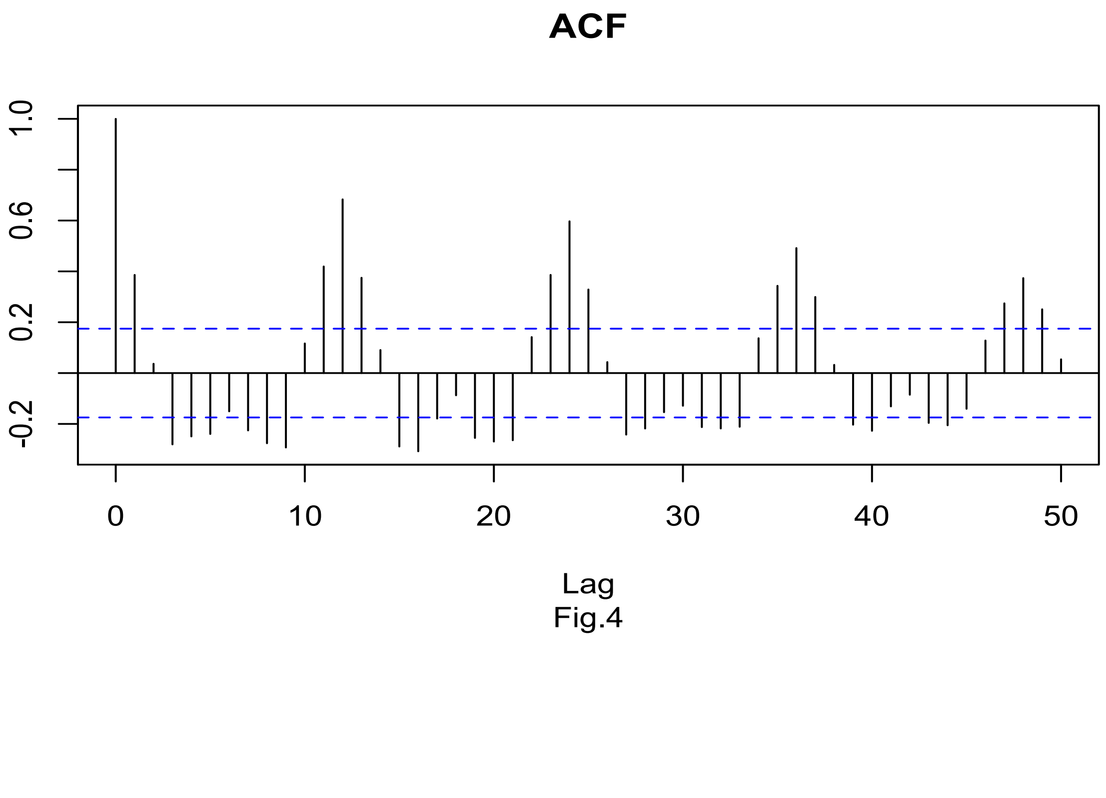

Fig.4

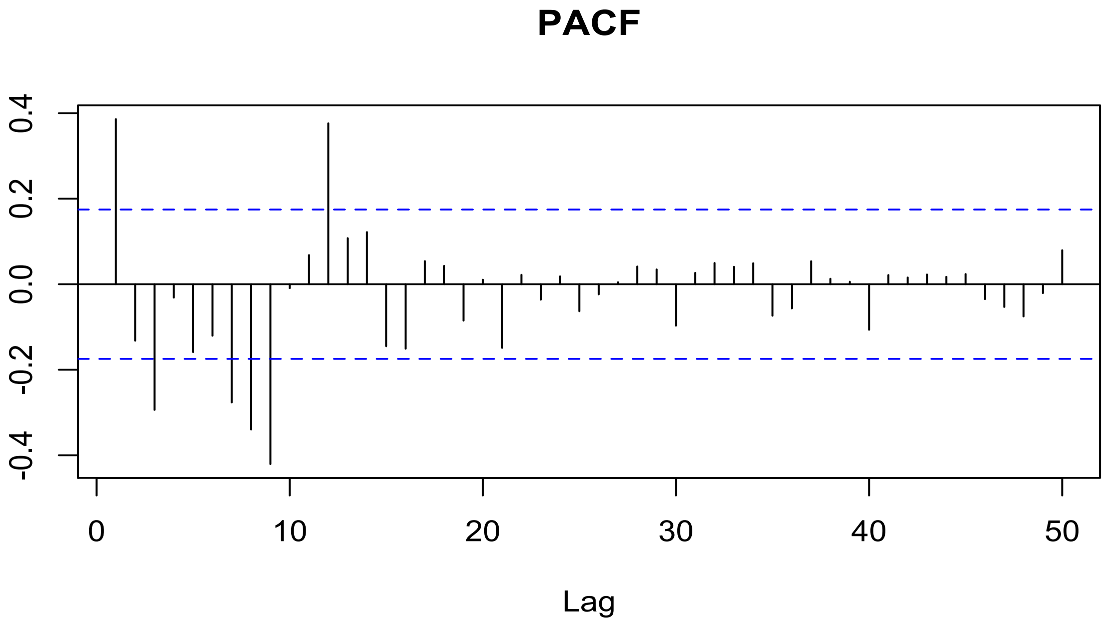

Fig.5

To make a seasonal difference is useful, Fig.6 shows the data after both first and seasonal difference. It appears that most, if not all, of the seasonality and the trend are gone now. Furthermore, Fig.7 the ACF picture shows the data are almost stationary.

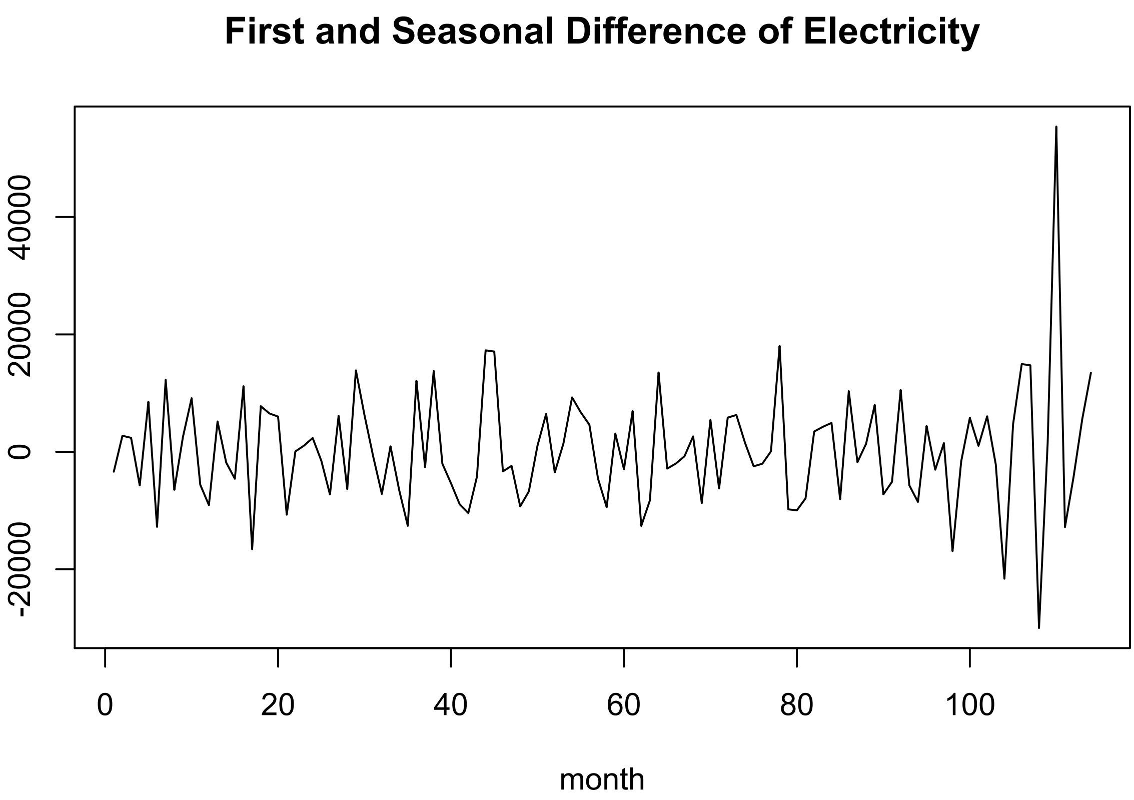

Fig.6

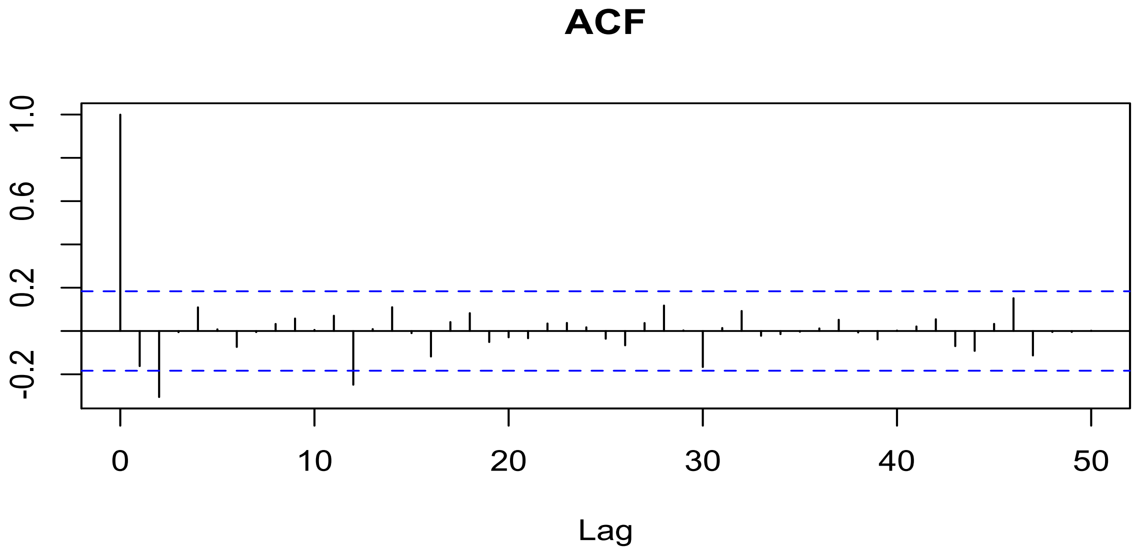

Fig.7

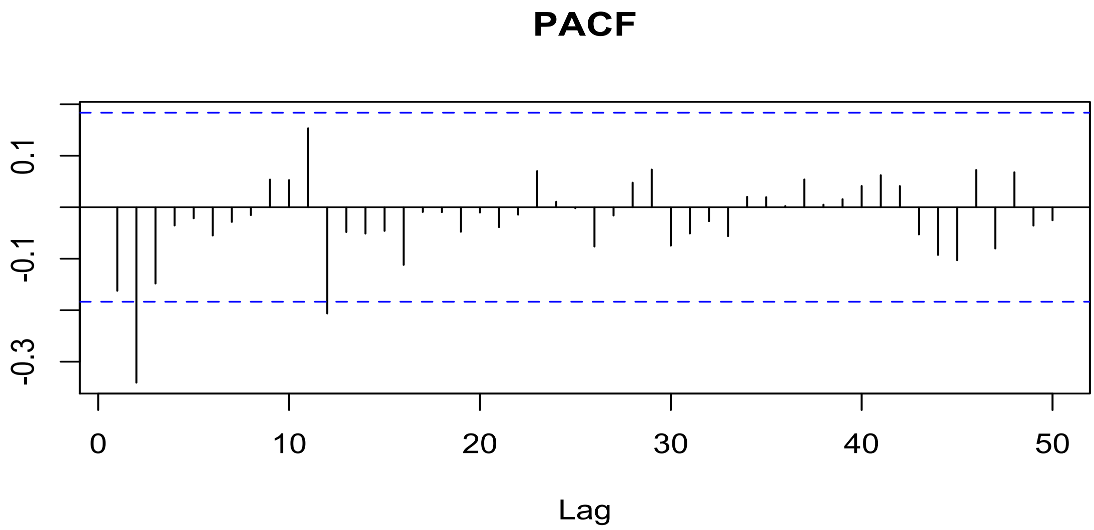

Fig.8

Fig.7 and Fig.8 confirms that very little autocorrelation remains in the series after these two differences have been taken. Also suggests that a model which incorporates the lag 1 and lag 12 autocorrelations might be adequate.

Now, we will consider a seasonal ARIMA model to fit this time series. Firstly, we consider the ACF and PACF at every 12 lag. We can identify a possible case: the ACF cut of at lag 12 and the PACF cut of at lag 12. So we can consider an ARMA$(1,1)_{12}$ model. Secendly, we consider the ACF and PACF at lags 1 to 11. Both the ACF and PACF are decaying quickly, we can consider a possible case: ARMA(2,2).

Hence, the fit model is $\mathrm{ARIMA}(2,1,2)\times(1,1,1)_{12}$. The estimated model for the electricity data series is

$$\left(1 + 0.3377B^{12}\right)\left(1 - 0.1198B - 0.0986B^{2}\right)\Delta_{12}\Delta x_t$$
$$= \left(1 - 0.1453B^{12}\right)\left(1 - 0.5721B - 0.4729B^{2}\right)\varepsilon_t$$

We can see the residuals, the ACF plot of the residual series, the Q-Q plot of the residuals and the LBP tests at various lags with p-values shown in Fig.9 and Fig.10. There are a few large residuals, but the diagnostic plots seem fine, the residuals have no relationship with each other, and it is white noice with no infromations.

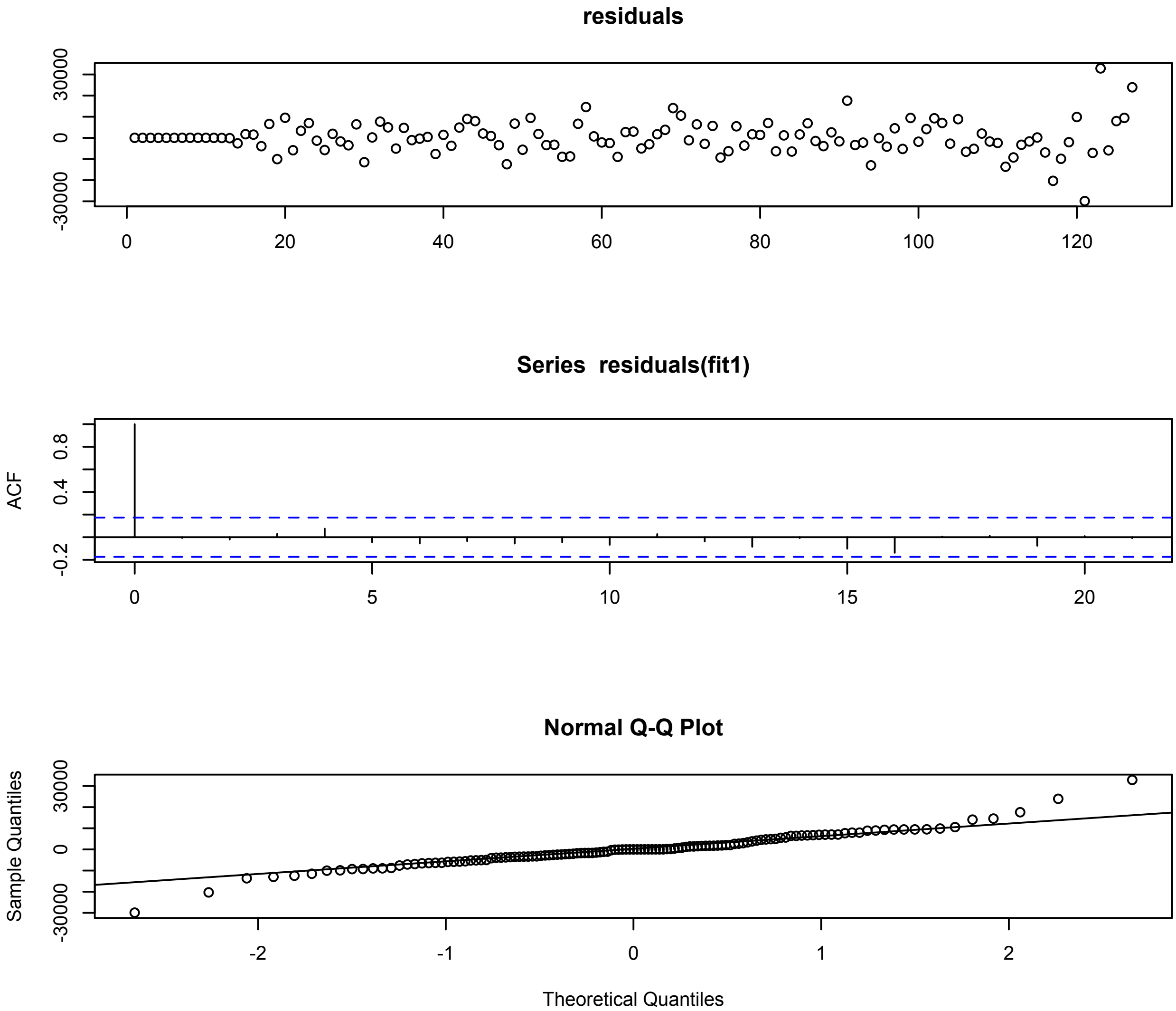

Fig.9

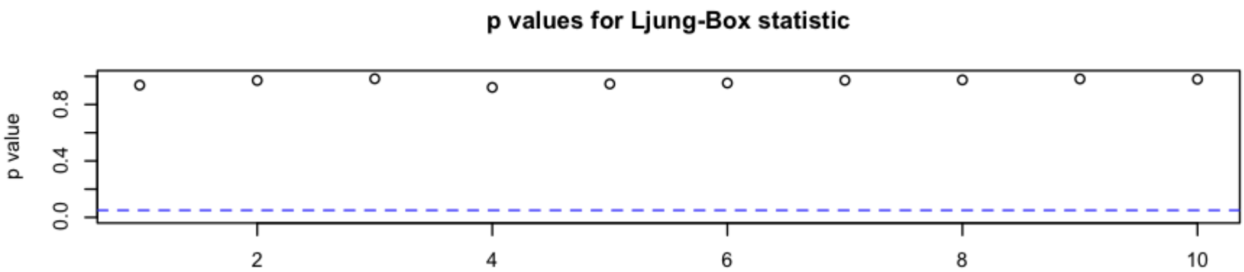

Fig.10

We compare the original data and the model shown in Fig.11, they are almost the same. And we can do further predict using this model, and find the 95% confidence interval of our prediction, it is shown in Fig,12 .

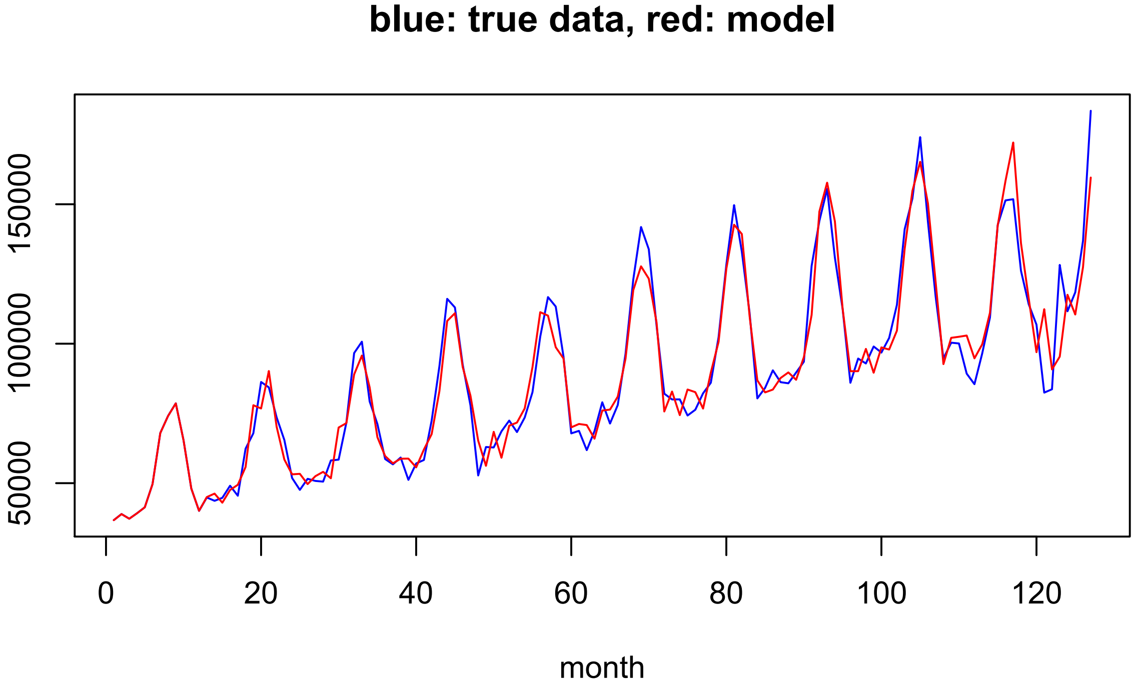

Fig.11

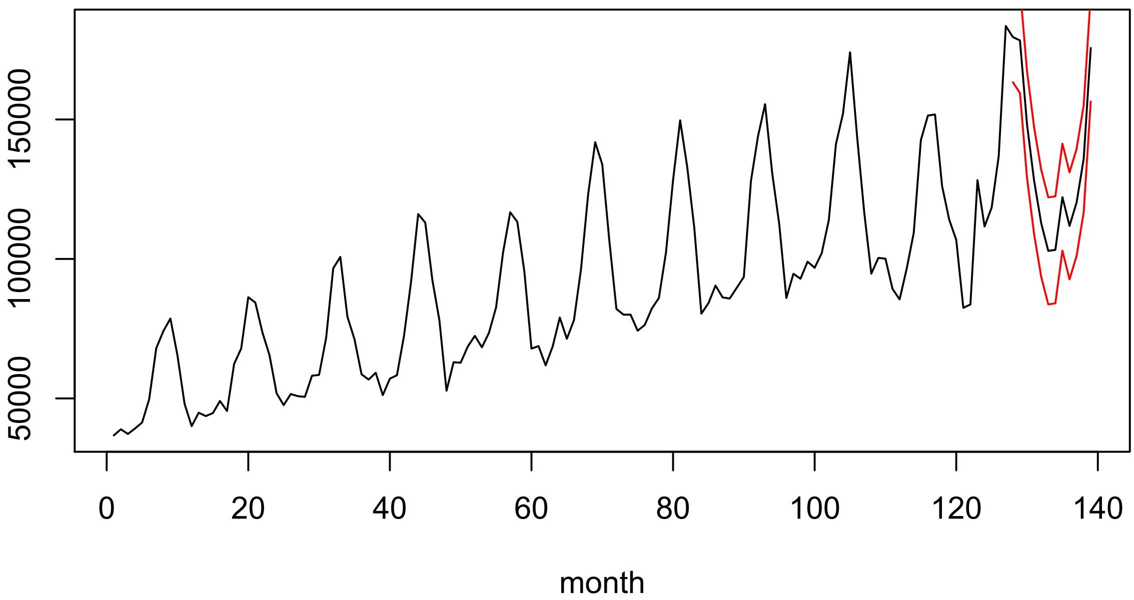

Fig.12