STAT 929 Project : Data Analysis of Time Series

Deepak Rishi

August 10, 2016

1 Fit Bit Data

1.1 Task 1
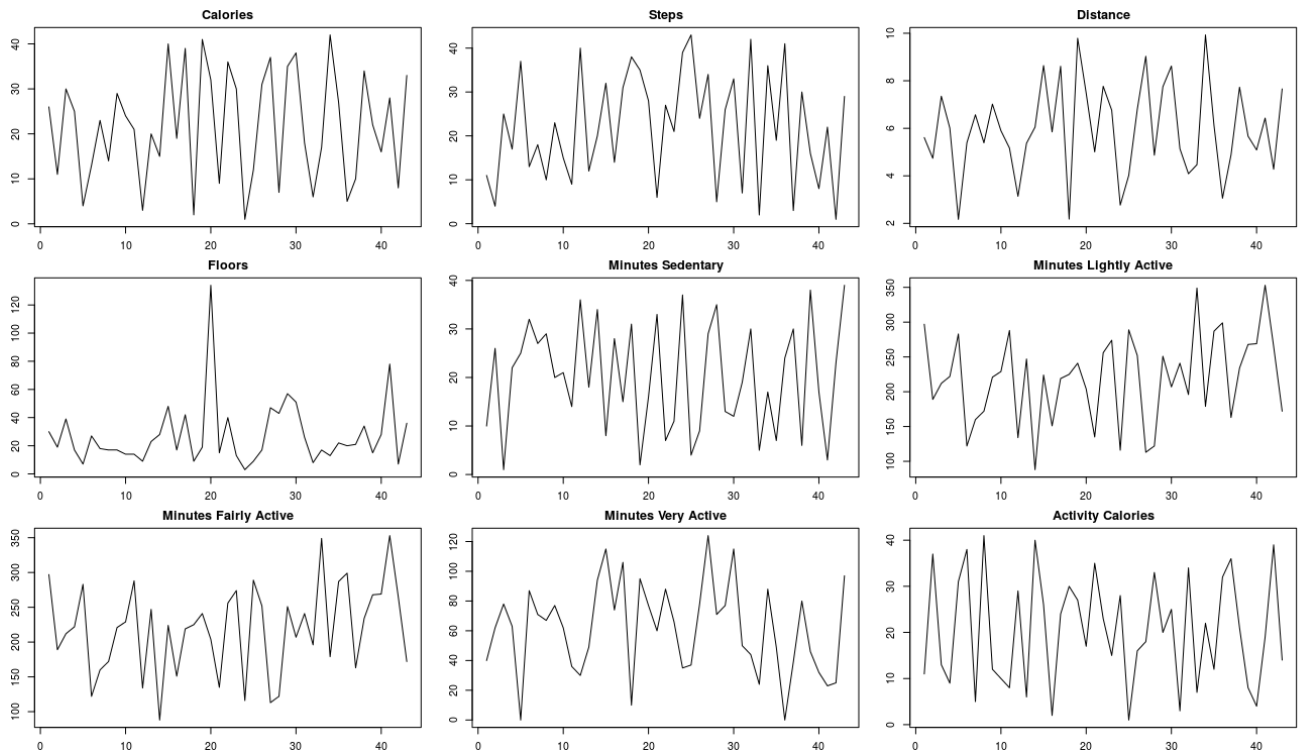
The time series plot of all the attributes is :

Figure 1: Time series plot of all the attributes

Before processing any attribute , the attribute is centered by subtracting the mean.
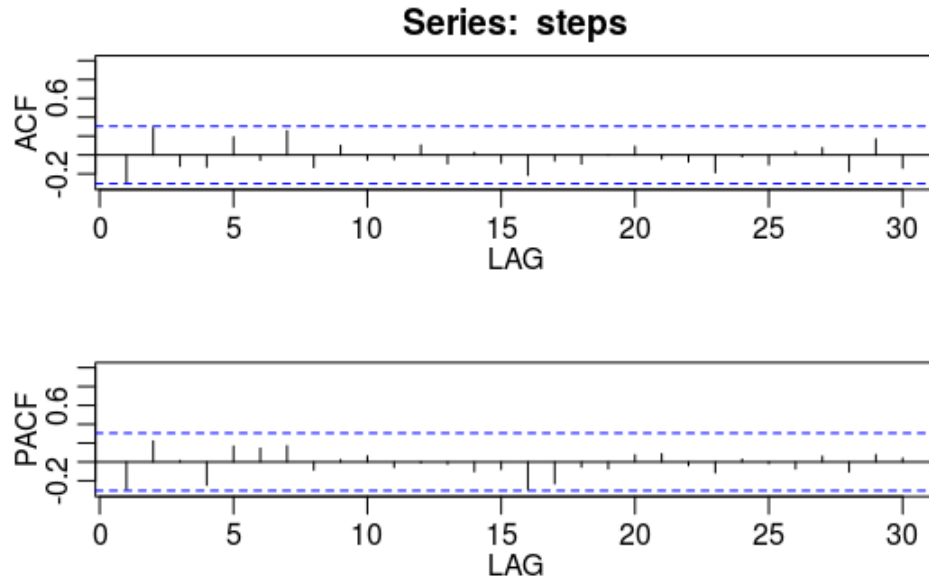
1.1.1 Steps

The ACF and PACF plots for Steps are

Figure 2: ACF and PACF of Steps

From Figure's 1 and 2 it can be seen that the **Steps** is very close to being a white noise series. Although as can be seen in Figure 2 , the ACF is slightly significant at lag 7 and the PACF is significant at lag 1. Hence a **ARMA(1,7)** model is fit to the data.

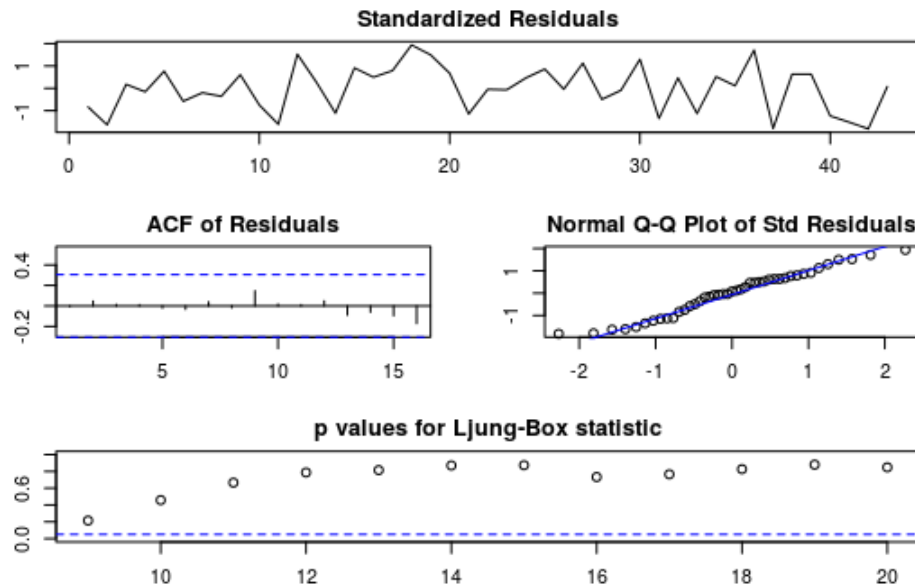The following diagnostic justifies this.

Figure 3: Dignostic Plots for Steps ARMA(1,7)

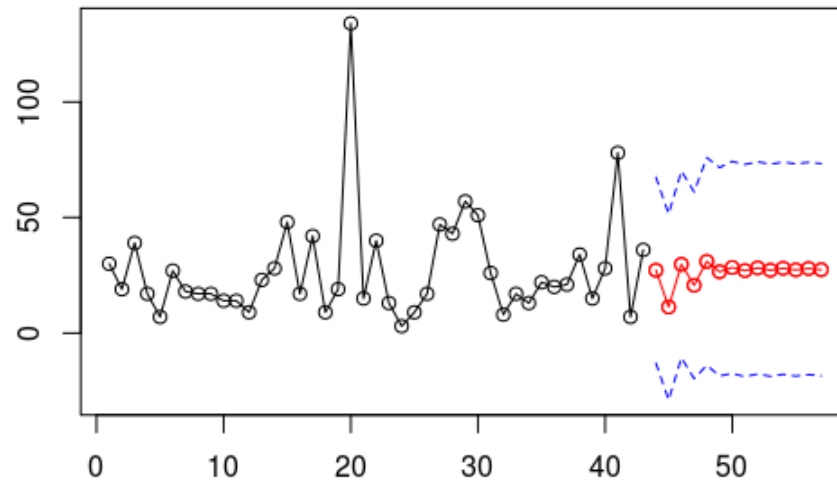The 14 day forecast with 95% prediction bands for the Step Series is :

Figure 4: 14 days forecast for Steps ARMA(1,7) with 95% prediction bands

1.1.2 Floors

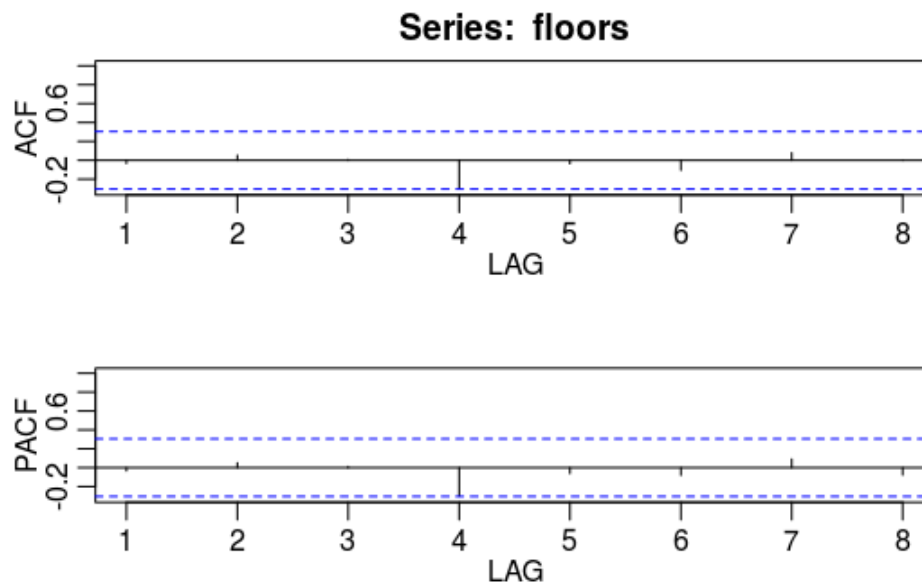The ACF and the PACF of the **Floors** attribute is :

Figure 5: ACF & PACF for Floors

3

As can be seen from Figure 5 we have slightly significant lags at 4 for both ACF and PACF. However, just fitting an **MA(4)** model gave good results in diagnostic plots and also gave the best AICc. The following diagnostic plot shows this.
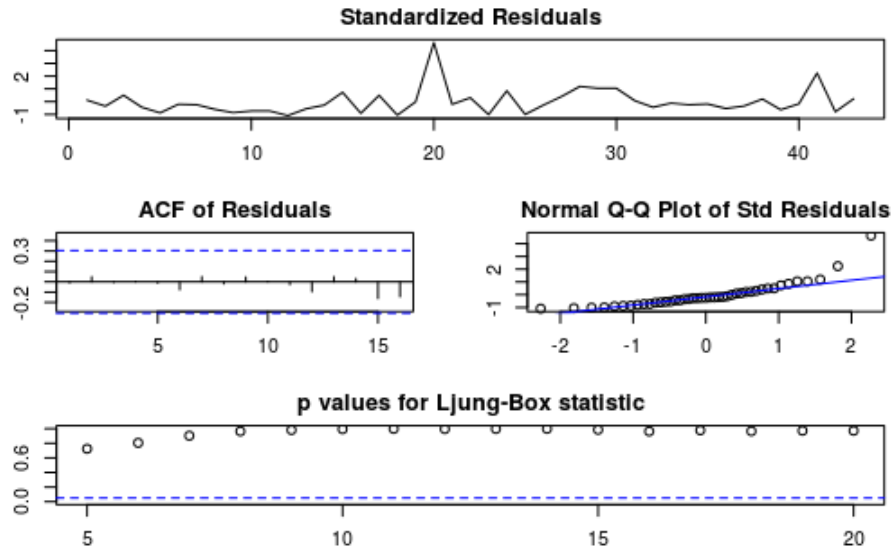

Figure 6: Diagnostic plot for Floors

The 14 day forecast with 95% prediction bands for the Floor Series is :
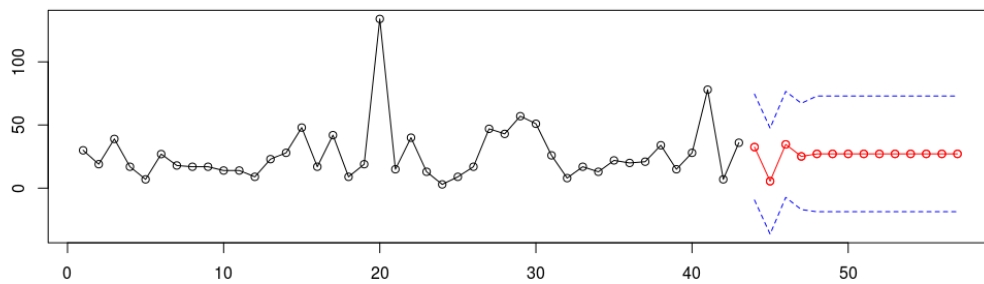

Figure 7: Diagnostic plot for Floors

1.1.3 Activity Calories

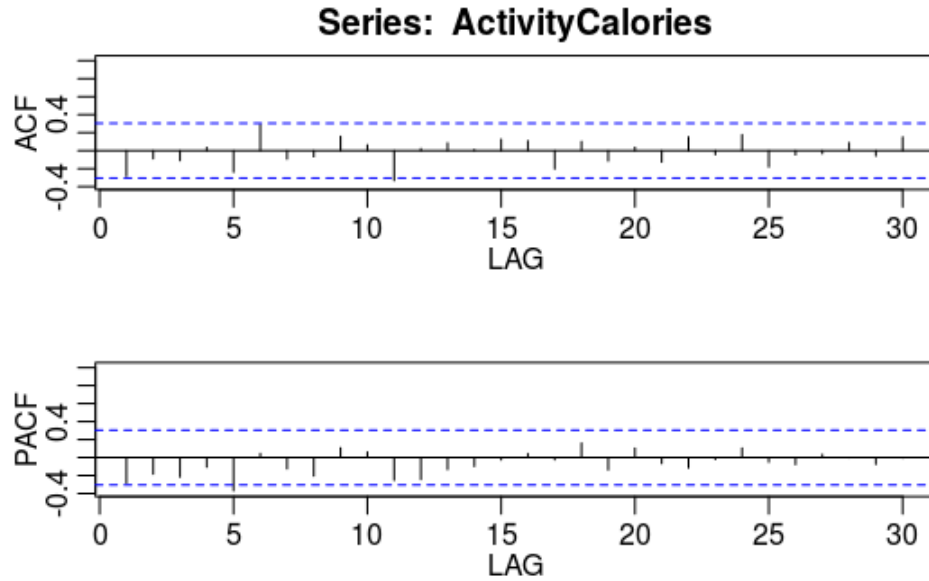The ACF and PACF plots for Activity Calories are :

Figure 8: ACF & PACF for Activity Calories

Differencing the series introduced false autocoreations.

As can be seen in Figure 8 the PACF cuts off at lag 5 and the ACF cyles down to 0. This suggests a **ARMA(5,0)** model. The diagnostic plot shows this.

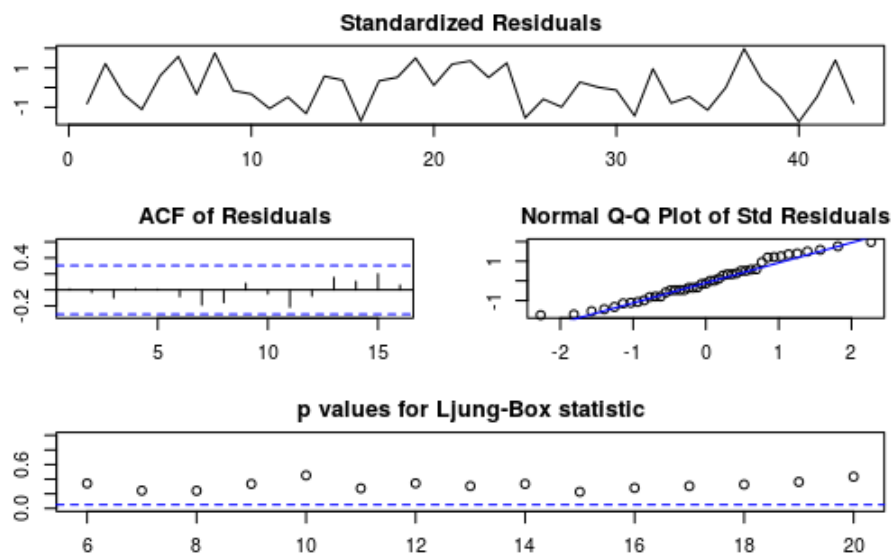The ACF and PACF plots for Activity Calories are :

Figure 9: Diagnostic plot for Activity Calories

The 14 day forecast with 95% prediction bands for the Activity Calories Series is :
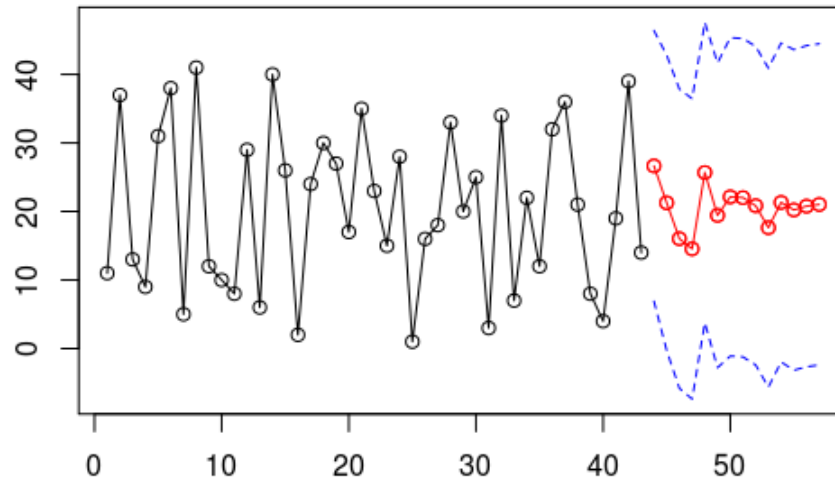
The ACF and PACF plots for Activity Calories are :

Figure 10: Forcast for Activity Calories

1.2 Task 2

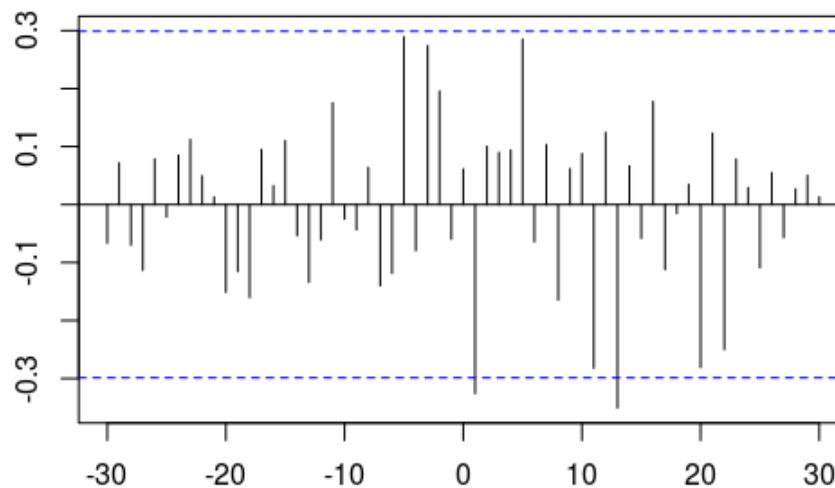The CCF plot between steps and floors is :

Figure 11: CCF between Steps & Floors

Most of the CCF values are insignificant, except for at lags 1 and 11.

The CCF plot between steps and floors and activity calories is :
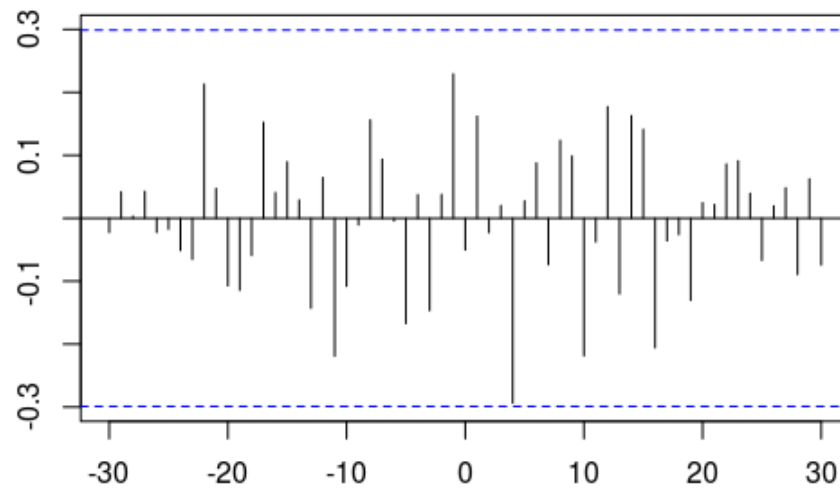

Figure 12: CCF between Floors & Activity Calories

All of the CCF values are insignificant.

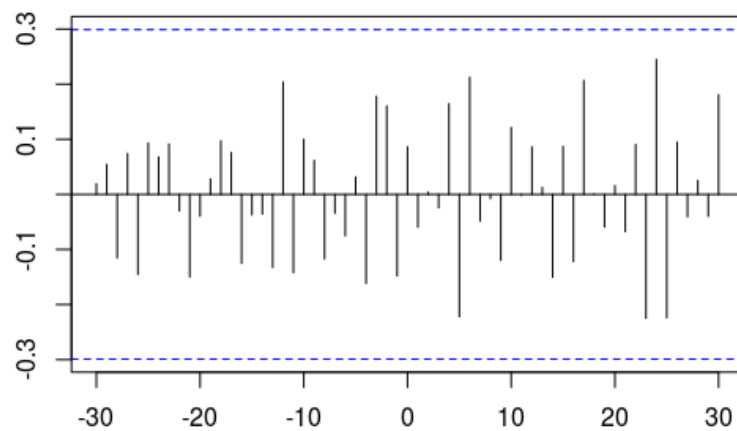The CCF plot between steps and floors and activity calories is :

Figure 13: CCF between Steps & Activity Calories

All of the CCF values are insignificant.

2 Lumber Data and Southern Oscillation Index

2.1 Task 1 Lumber Data

The lumber data was centered before processing it.

(a) Lumber Data

(b) ACF/PACF of original data

(c) ACF/PACF of once differenced data

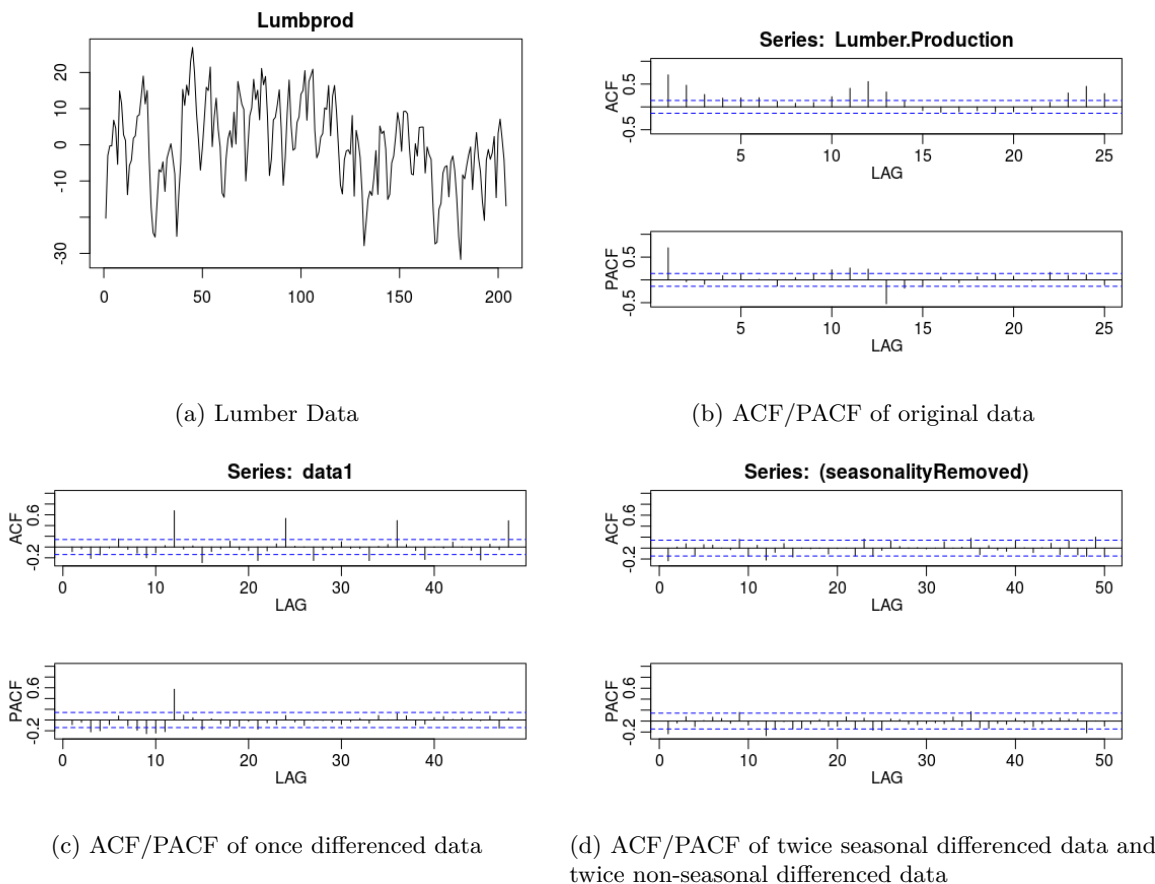(d) ACF/PACF of twice seasonal differenced data and twice non-seasonal differenced data

Figure 14: Lumber Data can various ACF/PACF plots

The following steps are followed to fit a **SARIMA (1,2,1)X(0,2,1)** model to the data.

- As per Figure 14d at seasonal lags the ACF cuts off at lag 1 and PACF tails off. This suggests a SMA of order 1.

- At the non-seasonal lags, it can be seen in Figure 14d that the ACF cuts off at 1 and so does PACF. At this point 3 models were tried.

    MA(1) model, which gave an AICc of 4.781826.

    AR(1) model, which gave an AICc of 5.221908.

    ARMA(1,1) model, which gave an AICc of 4.735571.

Hence, an **ARMA(1,1)** model was used to fit the non-seasonal part. The final model is thus given by **SARIMA (1,2,1)X(0,2,1)**. The following diagnostic plot shows the results of fitting the SARIMA model.
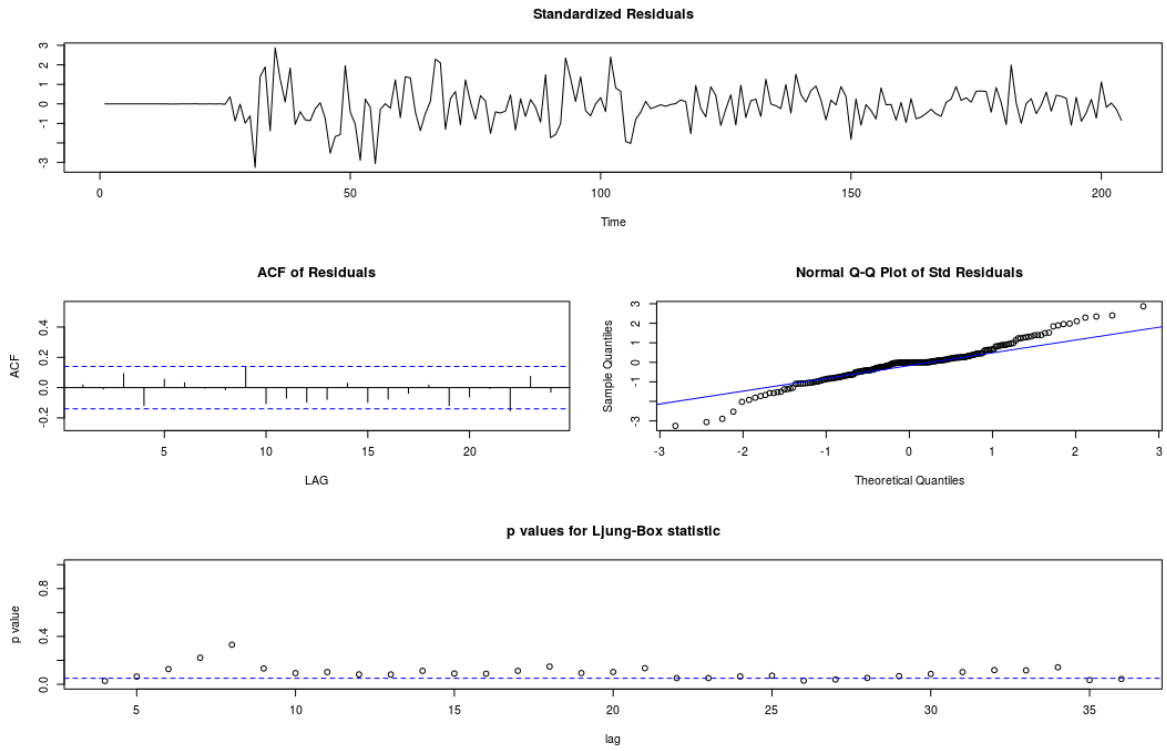
8

Figure 15: **SARIMA (1,2,1)X(0,2,1)** diagnostic plot

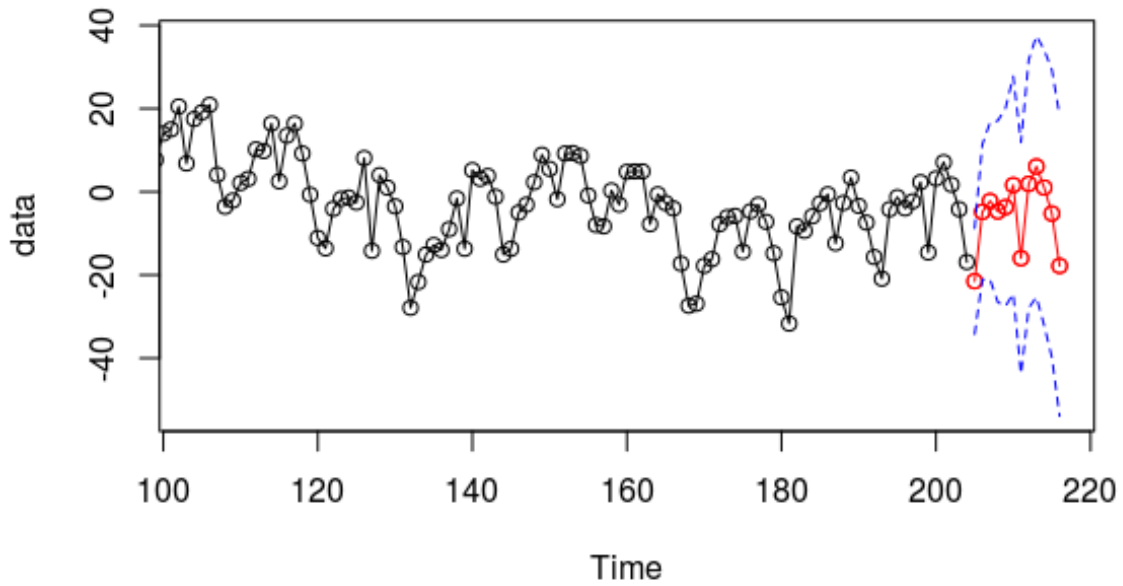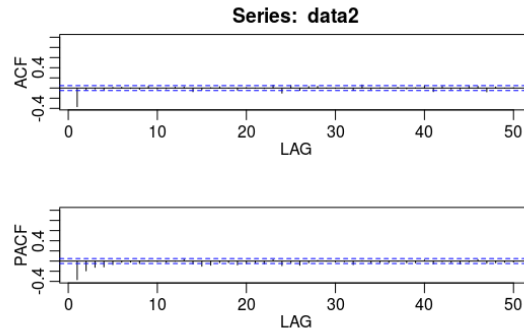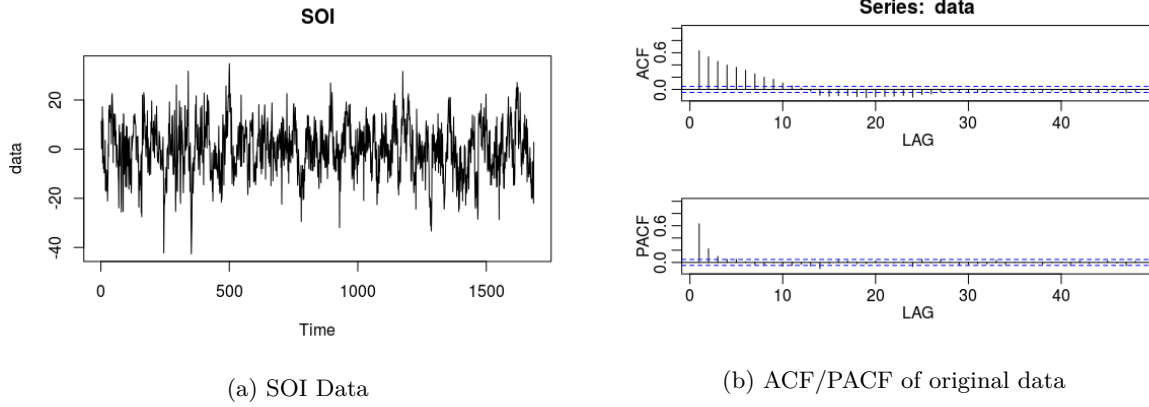The 12 month forecast with 95% prediction bands for the Lumber data is :

Figure 16: **SARIMA (1,2,1)X(0,2,1)** forcast for Lumber data

2.2 Task 1 Southern Oscillation Index data

The soi data was centered before processing it.

(a) SOI Data

(b) ACF/PACF of original data

(c) ACF/PACF of once differenced data

Figure 17: SOI Data can various ACF/PACF plots

Hence, an **ARMA(1,1)** model was used to fit the model. The following diagnostic plot shows the results of fitting the ARIMA model.
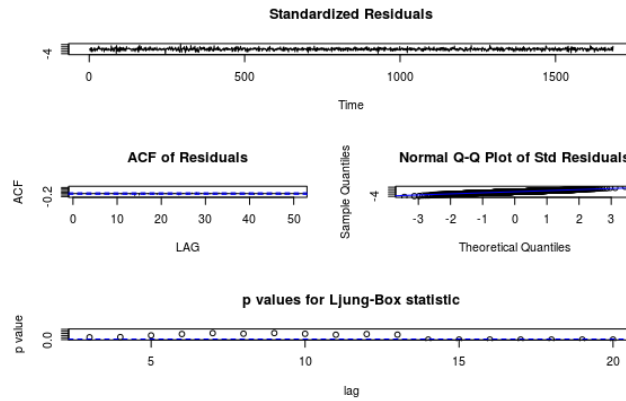

Figure 18: **ARIMA (1,1,1)** diagnostic plot

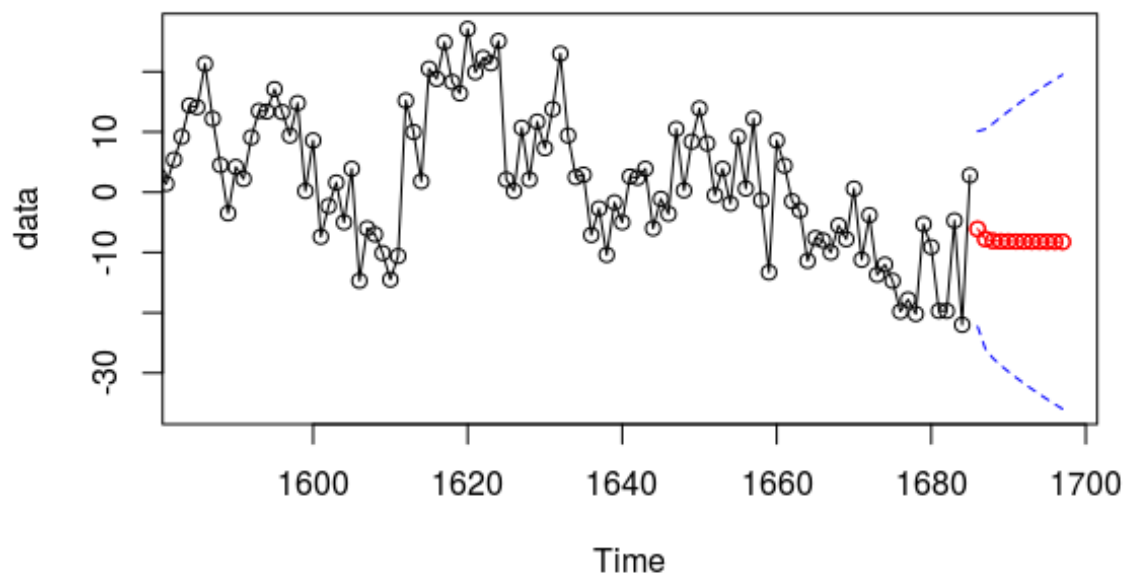The 12 month forecast with 95% prediction bands for the SOI data is :

Figure 19: **ARIMA (1,1,1)** forcast for soi data
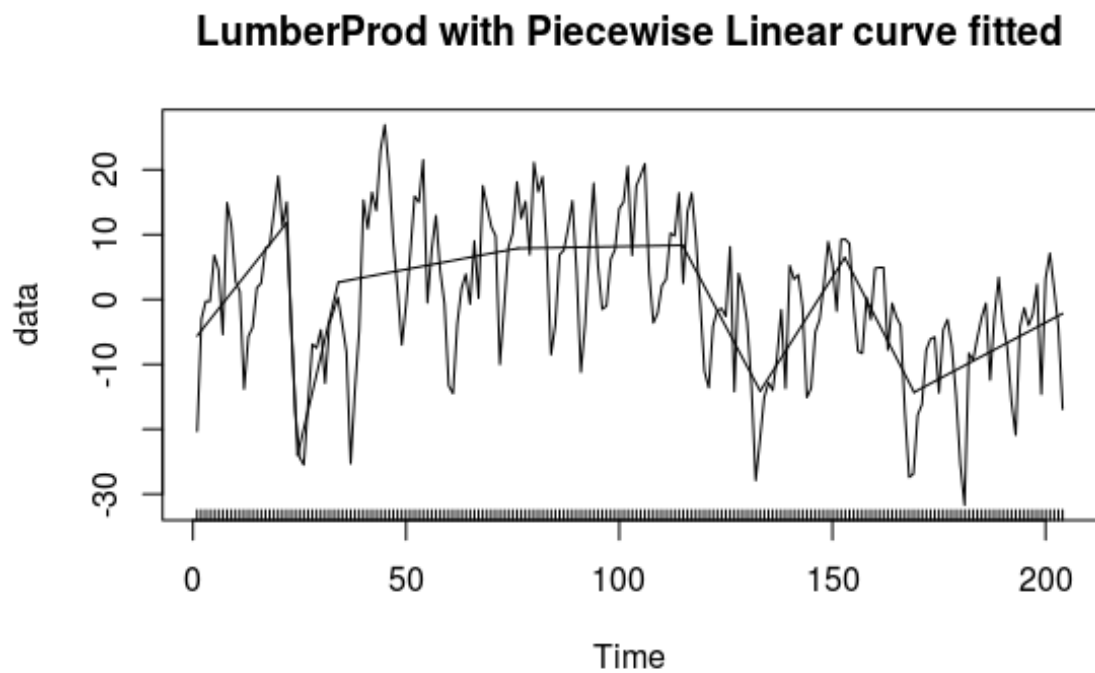
2.2.1 Task 2 Lumberprod Periodogram
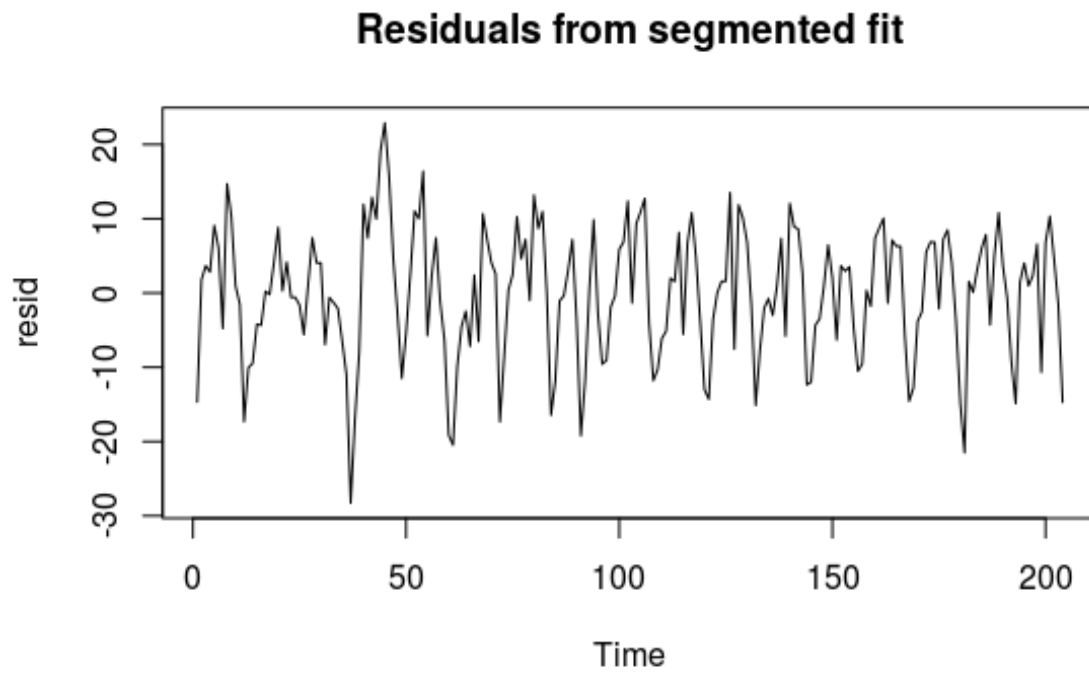

Figure 20: Segmented fit for Lumberprod data

Residuals from segmented fit

Figure 21: Residuals from segmented fit

Series: resid
Smoothed Periodogram
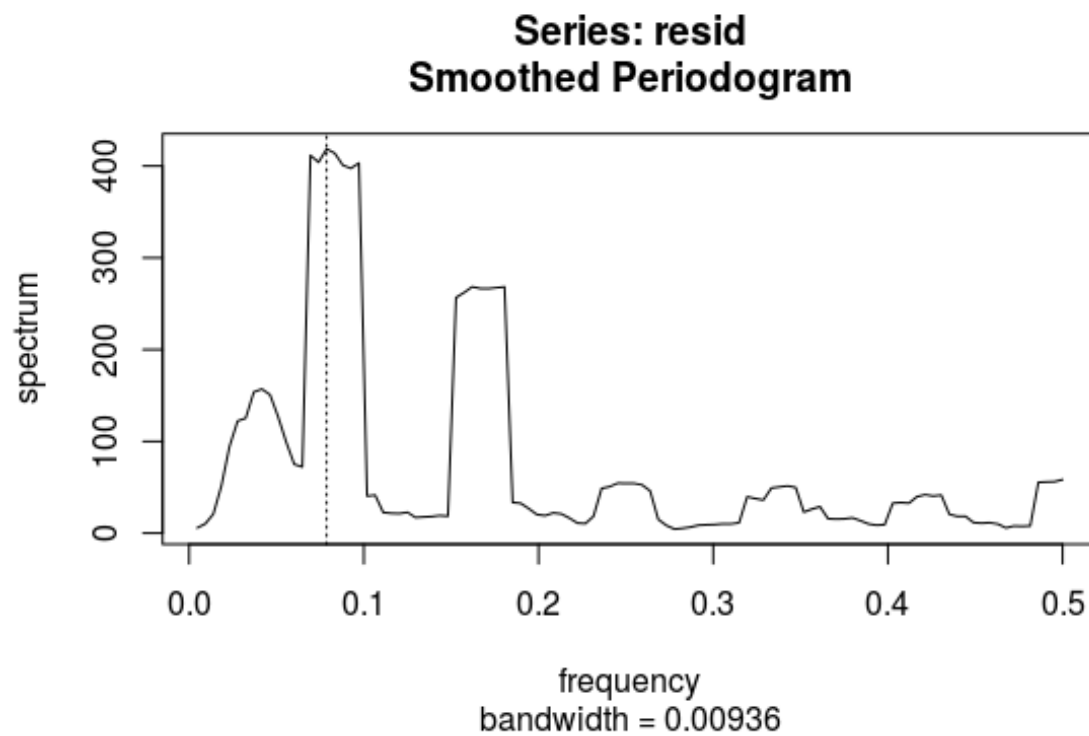

frequency
bandwidth = 0.00936

Figure 22: Residual smoothed periodogram with daniell kernel of order 3

- The 95 % confidence interval is $[221.0659, 1076.022]$ for a maximum spectrum of 418.7224 at a frequency of 0.078703704 . It is an significant interval.
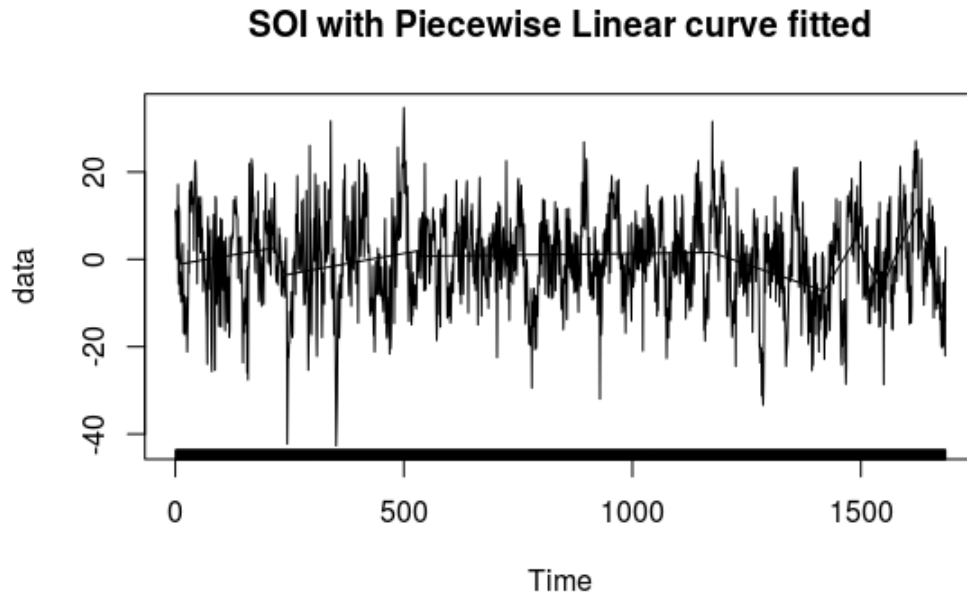
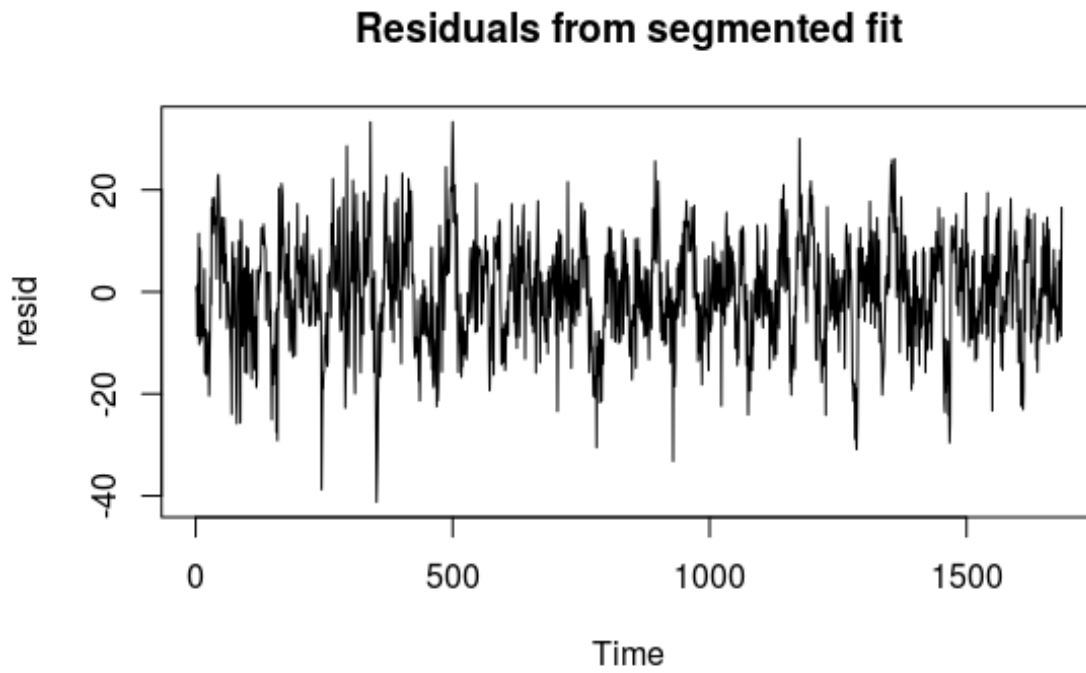### 2.2.2 Task 2 SOI

Figure 23: Segmented fit for SOI data

Figure 24: Residuals from segmented fit
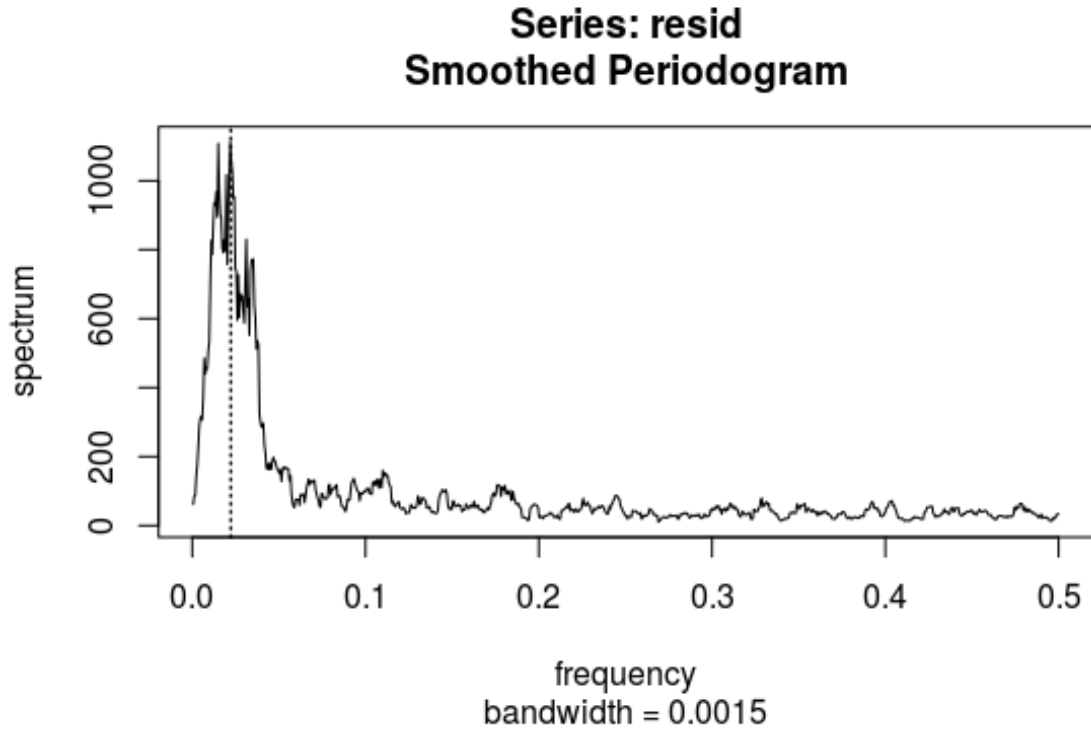
Series: resid
Smoothed Periodogram

Figure 25: Residual smoothed periodogram with daniell kernel of order 4

- The 95 % confidence interval is $[631.4339, 2462.478]$ for a maximum spectrum of 1112.683 at a frequency of 0.0219907407 times the period (1/12). It is an significant interval.

2.2.3   Task 3 Coherence Inference between the series

The following figure shows the lumberprod periodogram at the frequency at which SOI spectrum has the maximum value.
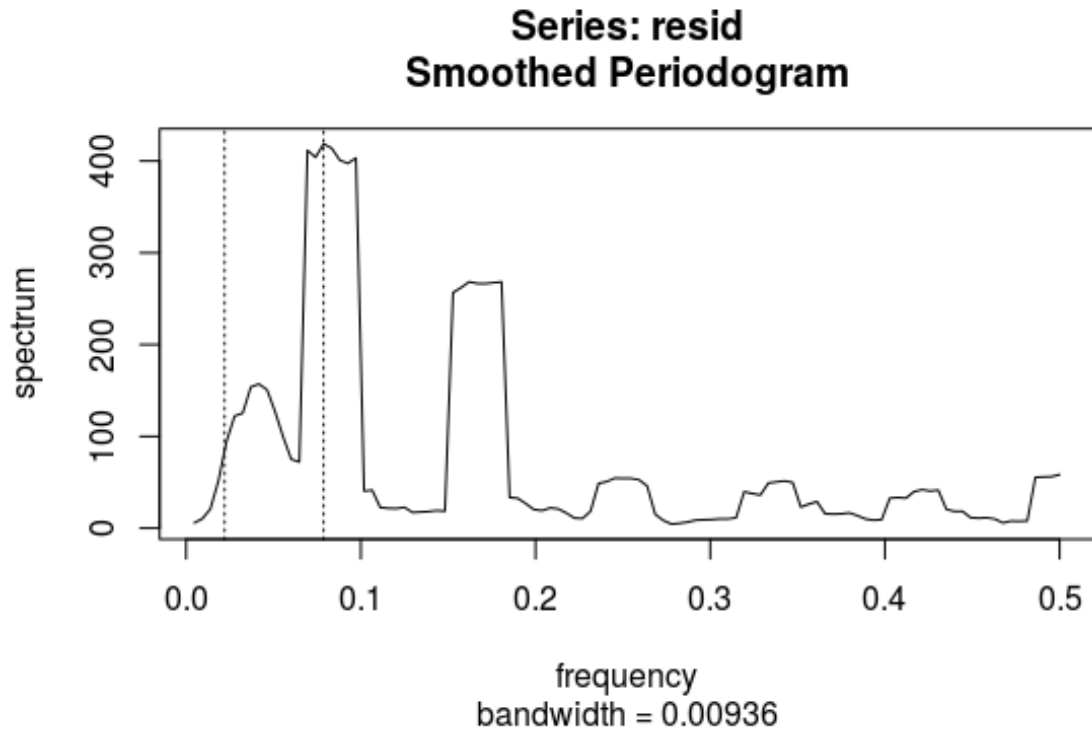
Figure 26: Lumberprod smoothed periodogram at the frequency at which SOI spectrum has the maximum value.

- The lumberprod periodgram has a small peak at at the frequency at which SOI spectrum has the maximum value. This shows a possible correlation between the SOI Lumberprod data.

- The 95 % confidence interval for the SOI freqency spectrum value of 84.00045 (obtained via linear interpolation) is $[44.34832, 215.8622]$. It is an significant interval.