Stats 326 - Assignment 5

Hasnain Cheena

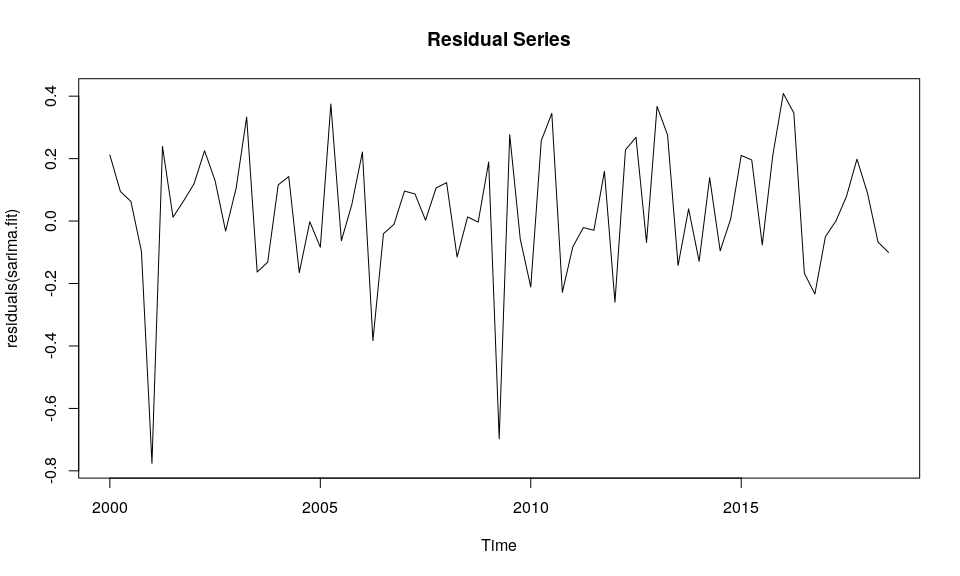
27/05/2020

## Question 1

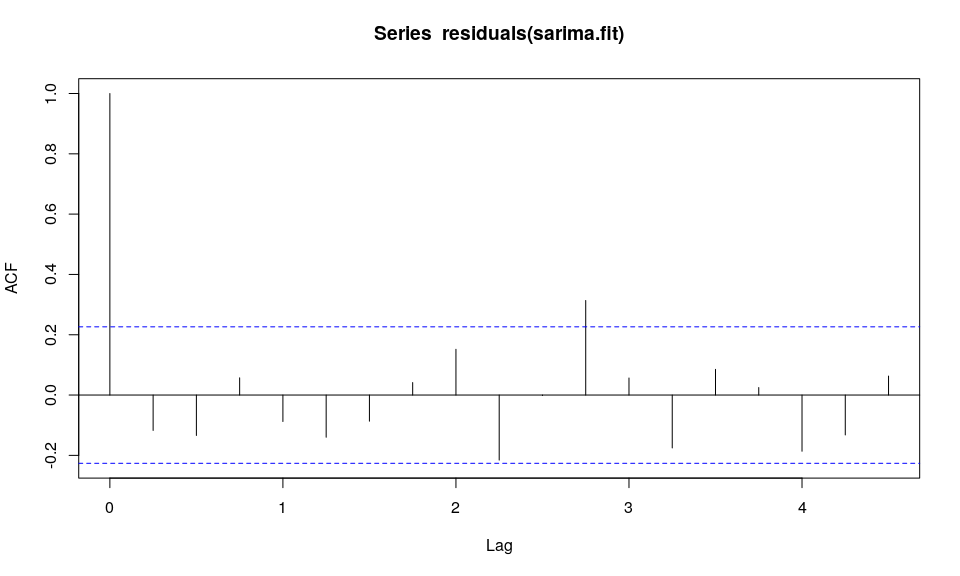
#best fitting SARIMA  
sarima.fit = arima(red.CO2.ts,order=c(0,1,1),  
seasonal=list(order=c(0,1,1),period=4))  
sarima.fit

##   
## Call:  
## arima(x = red.CO2.ts, order = c(0, 1, 1), seasonal = list(order = c(0, 1, 1),   
## period = 4))  
##   
## Coefficients:  
## ma1 sma1  
## 0.5764 -0.8992  
## s.e. 0.1129 0.1326  
##   
## sigma^2 estimated as 0.03919: log likelihood = 10.48, aic = -14.96

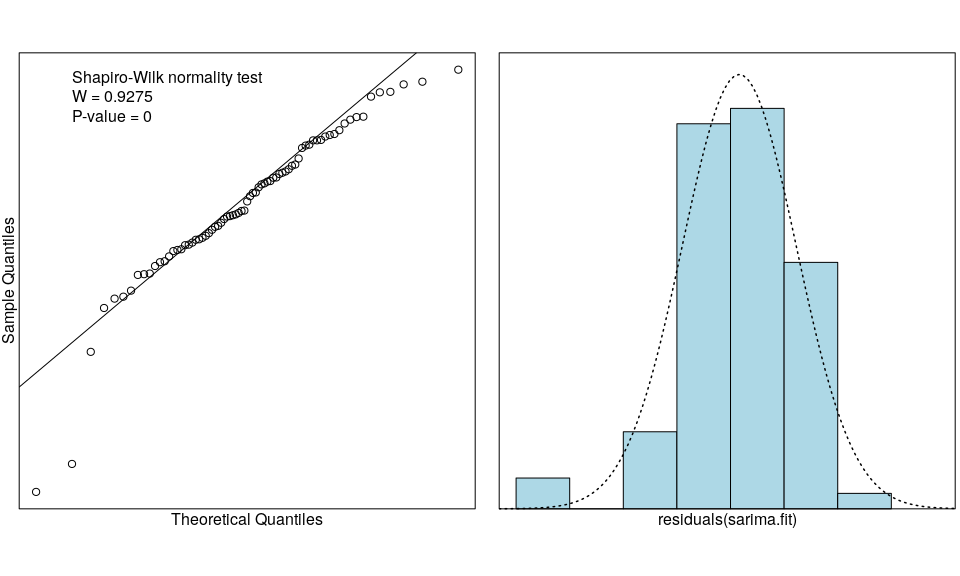
plot.ts(residuals(sarima.fit),main="Residual Series")



acf(residuals(sarima.fit))



normcheck(residuals(sarima.fit), shapiro.wilk = T)

 The Residual Series show reasonably random scatter about 0, although there are two large negative residuals (2001 Quarter 1 and 2009 Quarter 2). The autocorrelation function plot of the Residual Series shows a significant lag at lag 11. This is an unsual lag to be significant in quarterly data. Furthermore, the residuals seems reasonably normally distributed. Therefore, all model assumptions are satifised. Normality is okay with CLT though?

#predictions  
sarima.pred = predict(sarima.fit,n.ahead=4)  
sarima.pred

## $pred  
## Qtr1 Qtr2 Qtr3 Qtr4  
## 2018 405.8136  
## 2019 405.7907 406.3947 407.6694   
##   
## $se  
## Qtr1 Qtr2 Qtr3 Qtr4  
## 2018 0.1983784  
## 2019 0.3702966 0.4845881 0.5766555

#actual  
pred.CO2.ts

## Qtr1 Qtr2 Qtr3 Qtr4  
## 2018 405.83  
## 2019 405.73 406.71 408.25

#RMSEP  
RMSEP.sarima = sqrt(1/4\*sum((pred.CO2.ts-sarima.pred$pred)^2))  
RMSEP.sarima

## [1] 0.3318586

The model had an RMSEP of 0.33 ppm. The best predicting model from previous assginments was the seasonal-trend-lowess seasonally adjusted model as it had the lowest RMSEP (0.2 ppm). Therefore, the SARIMA model is not better predicting the the STL model has SARIMA has a higher RMSEP.

## Question 2

#best fitting SARIMA  
sarima.fit.full = arima(full.CO2.ts,order=c(0,1,1),  
seasonal=list(order=c(0,1,1),period=4))  
sarima.fit.full

##   
## Call:  
## arima(x = full.CO2.ts, order = c(0, 1, 1), seasonal = list(order = c(0, 1, 1),   
## period = 4))  
##   
## Coefficients:  
## ma1 sma1  
## 0.5614 -0.8524  
## s.e. 0.1150 0.1008  
##   
## sigma^2 estimated as 0.04027: log likelihood = 10.98, aic = -15.97

Model in backshift notation:

#### Predictions

Prediction 2019 Q4:

#2019 Q4  
pred.2019Q4 = 408.25 + 405.83 - 405.56 + (0.5614 \* 0.031927800) - (0.8524 \* 0.009630730) - (0.4785 \* -0.101088224)

Prediction 2020 Q1:

#2020Q1  
pred.2020Q1 = pred.2019Q4 + 405.73 - 405.83 - (0.8524 \* -0.106133913) - (0.4785 \* 0.009630730)  
pred.2020Q1

## [1] 408.5639

Prediction 2020 Q2:

#2020Q3  
pred.2020Q2 = pred.2020Q1 + 406.71 - 405.73 - (0.8524 \* 0.402908803) - (0.4785 \* -0.106133913)

Prediction 2020 Q3:

#2020Q4  
pred.2020Q3 = pred.2020Q2 + 408.25 - 406.71 - (0.8524 \* 0.031927800) - (0.4785 \* 0.402908803)

#results  
results.df = data.frame(Time=c("2019.4", "2020.1", "2020.2", "2020.3"),  
 Predictions=c(round(pred.2019Q4,2),round(pred.2020Q1,2),round(pred.2020Q2,2),round(pred.2020Q3,2)))  
   
results.df

## Time Predictions  
## 1 2019.4 408.58  
## 2 2020.1 408.56  
## 3 2020.2 409.25  
## 4 2020.3 410.57

## Question 3 - Executive Summary

The task was to predict the atmospheric concentration of carbon dixiode at Cape Grim, Tasmania, Australia (in parts per million) between 2019 Quarter 4 and 2020 Quarter 3.

We need to be wary of our predictions and their reliability as we have a time series with only 79 observations. However the model that is used is a good model therefore the predictions should be reasonably reliable.

Several different models were built using observations between 2000 Quarter 1 and 2018 Quarter 3 and used to predict 2018 Quarter 4 to 2019 Quarter 3. Each model’s predictions were then compared to the actual values to find the model that produced the most accurate predictions. The best predicting model found from this method was then re-run on all the avaliable data and predictions for 2019 Quarter 4 to 2020 Quarter 3 were produced as per the task.

We predict the carbon dixiode concentration in the atmosphere above Cape Grim in Tasmania, Australia will be: 2019 Quarter 4: 408.60 ppm 2020 Quarter 1: 408.61 ppm 2020 Quarter 2: 409.34 ppm 2020 Quarter 3: 410.76 ppm