

MATH 1318- Time Series Analysis Analysis and Forecast of Ozone Layer Thickness 2017- 2021

By Avinash Matani

S3692165

Table of Contents

List of Figures	3
List of Exhibits	3
Introduction –	4
Methods –	4
Data Interpretation –	4
Time Series of the dataset –	4
Section 1.....	6
Characteristics and Model Selection.....	6
Linear Regression Model.....	6
Quadratic Regression Model -	7
Residual Analysis.....	9
Time Series plot of Standardized residuals.....	9
Histogram of Standardized Residuals	10
Q-Q Plot to check the normality of Standardized residuals.....	10
Normality Test – Shapiro Wilk Test.....	11
ACF plot of Standardized residuals	11
Forecast Analysis.....	12
Section 2.....	12
ACF and PACF Plot of the dataset –	13
Stationary Test of the dataset –.....	13
Q-Q Plot to check the normality of series.....	14
Lag 1 Differenced series	14
ADF Unit Root Test.....	15
Normality Test of Differenced Series	15
ACF and PACF plot of Differenced dataset	16
EACF Plot of the differenced dataset.....	16
BIC Table	17
Conclusion.....	17
References	18

List of Figures

Figure 1 Time Series Plot of Ozone Layer Thickness in Dobson units from 1927 - 2016	4
Figure 2: Scatter Plot of Current Year Ozone Layer Thickness v/s Previous Year Thickness	5
Figure 3 ACF plot of time based Ozone Layer Thickness dataset	5
Figure 4: Time series superimposed with least square regression trend line	7
Figure 5: Quadratic Regression Model fitted with quadratic trend line	8
Figure 6: Standard Residual Plot of Quadratic Regression Model	9
Figure 7: Histogram plot of Standard Residual Model	10
Figure 8: QQ Plot of Standardised Residual of Quadratic Trend Model	10
Figure 9: ACF plot for standardized residual of Quadratic Trend Model	11
Figure 10: Prediction plot of Ozone Layer Thickness from year 2017 to 2021	12
Figure 11: ACF and PACF Plot of Ozone Layer Thickness	13
Figure 12: QQ Plot of Ozone Layer Thickness Dataset	14
Figure 13: Time series plot of first order differenced series	14
Figure 14: ACF and PACF plot of differenced series	16

List of Exhibits

Exhibit 1: Linear Regression statistical output	6
Exhibit 2: Quadratic Regression Model Output	8
Exhibit 3: Normality Test of Standard Residual of Quadratic Trend Model	11
Exhibit 4: Ozone Layer Thickness Forecast of from year 2017 to 2021	12
Exhibit 5: DK Fuller Test for normality test of Ozone Layer Thickness Dataset	13
Exhibit 6: ADF test of Differenced Series	15
Exhibit 7: Shapiro- Wilk Normality test of differenced series	15
Exhibit 8: EACF Plot of differenced series	16
Exhibit 9: BIC plot differenced series	17

Introduction –

The report is mainly divided into two sections. The first section aims to find the best fitted deterministic trend model on the dataset consisting of yearly changes in the thickness of the Ozone layer from the year 1927 to 2016 in Dobson units. Further, we are aimed at predicting the yearly changes in the thickness of the ozone layer for 5 years i.e., from 2017 to 2021 based on the best-fitted model. The second section is aimed at proposing a set of possible ARIMA (p, d, q) models using suitable specification tools to select the appropriate values of p, d, and q. This section is primarily concerned with the identification and selection of the stochastic models.

Methods –

We will be using R software to identify and select the appropriate models. The first section will mainly employ the use of residual analysis, statistical significance test of coefficients, and deterministic trend analysis. The second section will use specification tools including ACF, PACF, EACF, BIC table, differencing and transformation techniques to select appropriate models and their orders.

Data Interpretation –

The dataset represents the change in the thickness of the Ozone layer from the year 1927 to 2016 in terms of Dobson units. An increase and decrease in the thickness level of the ozone layer are represented by the positive and negative sign of thickness level respectively. Further, the plot shows autoregressive and moving average behaviour with some seasonality.

Time Series of the dataset –

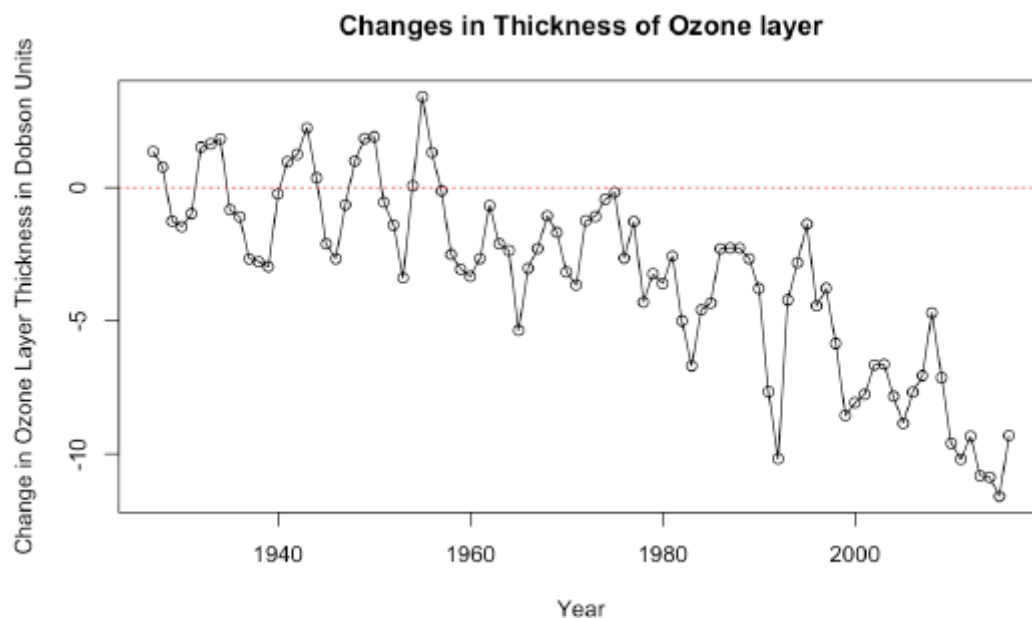


Figure 1 Time Series Plot of Ozone Layer Thickness in Dobson units from 1927 - 2016

Figure 01: shows the time series plot of the ozone layer thickness. Here the values that are neighbours in time, tend to relate with one another. This can be better represented by constructing the scatter plot of values at consecutive time points.

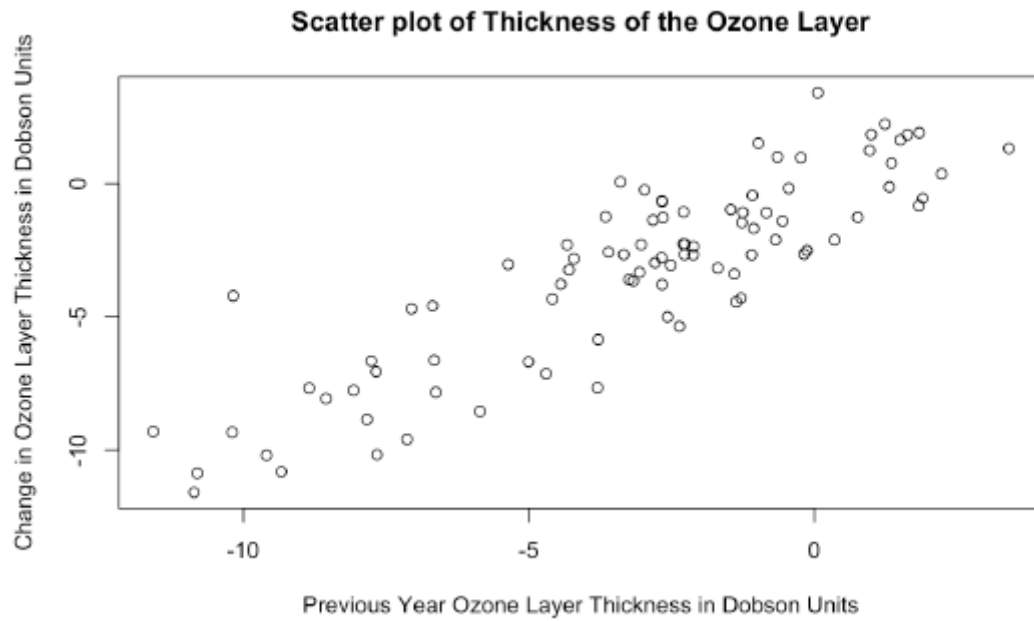


Figure 2: Scatter Plot of Current Year Ozone Layer Thickness v/s Previous Year Thickness

Figure 02 displays the scatter plot of the neighbouring pair of ozone layer thickness values. Its shows a slight downward trend – higher values of thickness tend to be followed in the next batch by lower values. The downward trend is apparent and the scatter plot represents the strong correlation between the previous year and current year values.

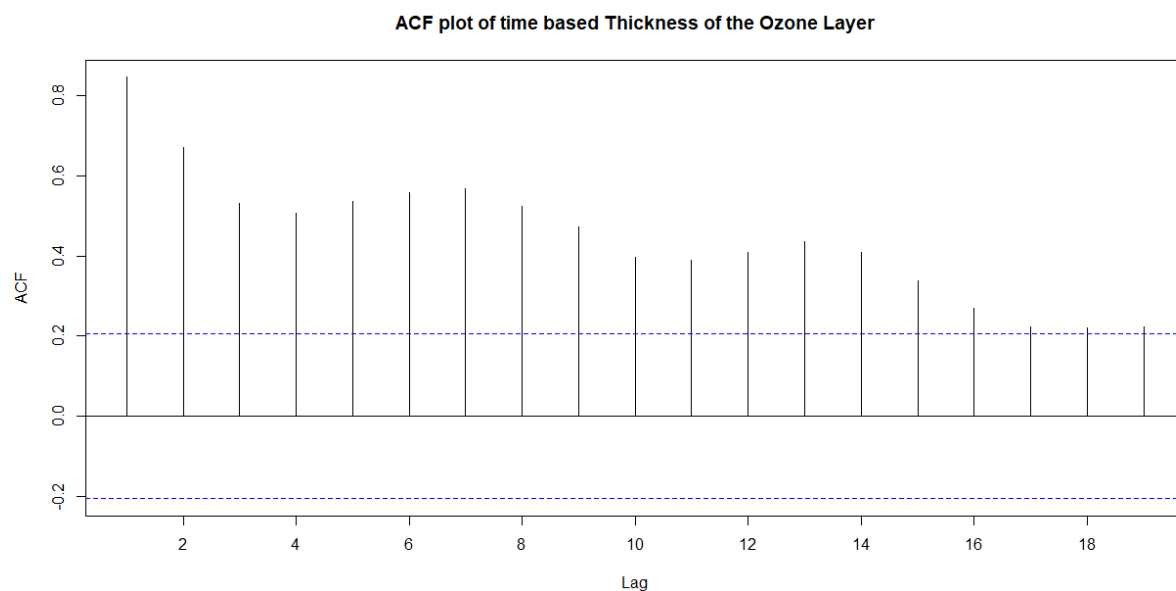


Figure 3 ACF plot of time based Ozone Layer Thickness dataset

Section 1

Characteristics and Model Selection

Since the dataset represents the ozone layer thickness from the year 1927 to 2016, it is apparent that the nature of the data is non-seasonal. Further, the time series plot in figure 01 represents the non-cyclic nature of the graph. This factor rules out the use of cosine, cyclic or seasonal trend model of deterministic nature.

We observe that in figure 01, there is a negative trend from the above graph. It also says that it has a negative correlation in the linear trend. As p-value is less than 5% the model is significant hence we fail to reject the null hypothesis.

Also, the ACF plot in figure 03 represents strong autocorrelation values and slow decaying ACF indicates the presence of a trend. In such circumstance, regression methods that are linear or quadratic in time can be considered as competing methods to estimate the parameters of non-constant mean trend model (D.Cryer & Chan, 2008).

Linear Regression Model

Exhibit 1 shows the output of linear time trend model from the statistical software and Figure 03 represents time series superimposed with least square regression trend line –

Call:

```
lm(formula = Ozone_thickness.ts ~ time(Ozone_thickness.ts))
```

Residuals:

Min	1Q	Median	3Q	Max
-4.7165	-1.6687	0.0275	1.4726	4.7940

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	213.720155	16.257158	13.15	<2e-16 ***
time(Ozone_thickness.ts)	-0.110029	0.008245	-13.34	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.032 on 88 degrees of freedom

Multiple R-squared: 0.6693, Adjusted R-squared: 0.6655

F-statistic: 178.1 on 1 and 88 DF, p-value: < 2.2e-16

Exhibit 1: Linear Regression statistical output

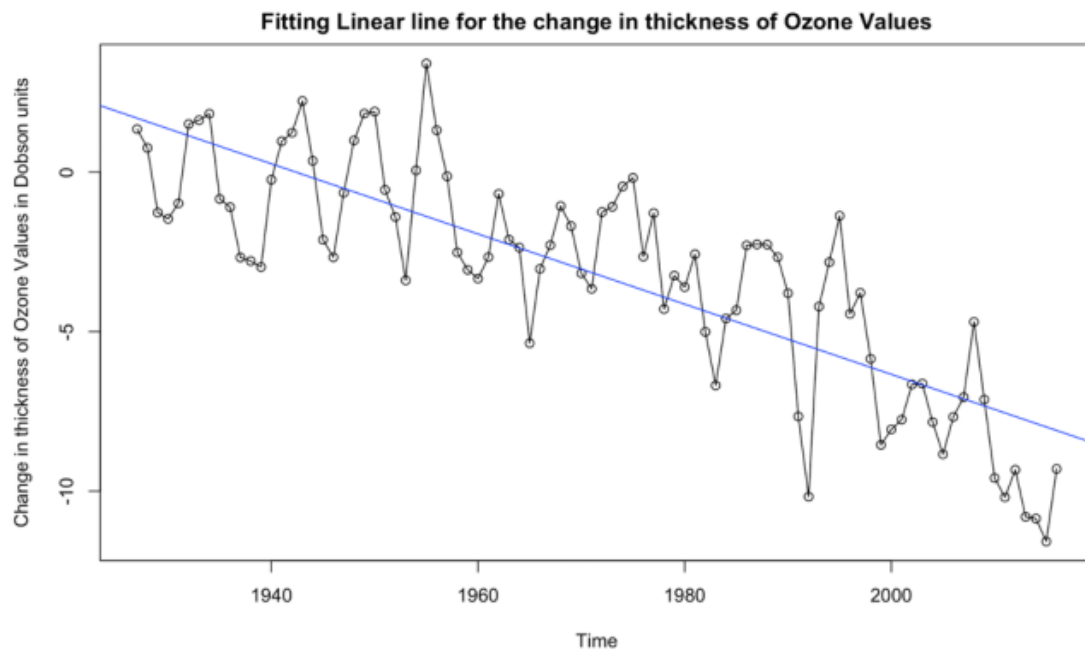


Figure 4: Time series superimposed with least square regression trend line

Based on Exhibit 1, the intercept and slope are estimated as $\beta_1 = 213.72$ and $\beta_0 = -0.110$ respectively. The slope and intercept are statistically significant with 0.05 as the significance level. Further, a glimpse of trend line fitted on the series in figure 03 together with an adjusted R^2 value of 0.665 from exhibit 1 shows that 66.5% of the variation in the series is explained by the linear time trend fitted above. This is a decent level of accuracy, however, as seen in figure 03 the linear trend line is not able to fully capture the mean level in the series with large residual value as some points.

Fitting a quadratic time trend might improve the value of the R^2 for better predictions. However, it should be noted that a quadratic trend adds multi-collinearity of the time variable and may lead to overfitting issues which may lead to misleading R^2 values, regression coefficients and p-values (D.Cryer & Chan, 2008).

Quadratic Regression Model -

Exhibit 2 shows the output of quadratic time trend model from the statistical software and Figure 05 represents time series superimposed with quadratic time trend model –

```

Call:
lm(formula = Ozone_thickness.ts ~ t + t2)

Residuals:
    Min       1Q   Median       3Q      Max
-5.1062 -1.2846 -0.0055  1.3379  4.2325

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -5.733e+03  1.232e+03  -4.654 1.16e-05 ***
t             5.924e+00  1.250e+00   4.739 8.30e-06 ***
t2          -1.530e-03  3.170e-04  -4.827 5.87e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.815 on 87 degrees of freedom
Multiple R-squared:  0.7391,    Adjusted R-squared:  0.7331
F-statistic: 123.3 on 2 and 87 DF,  p-value: < 2.2e-16

```

Exhibit 2: Quadratic Regression Model Output

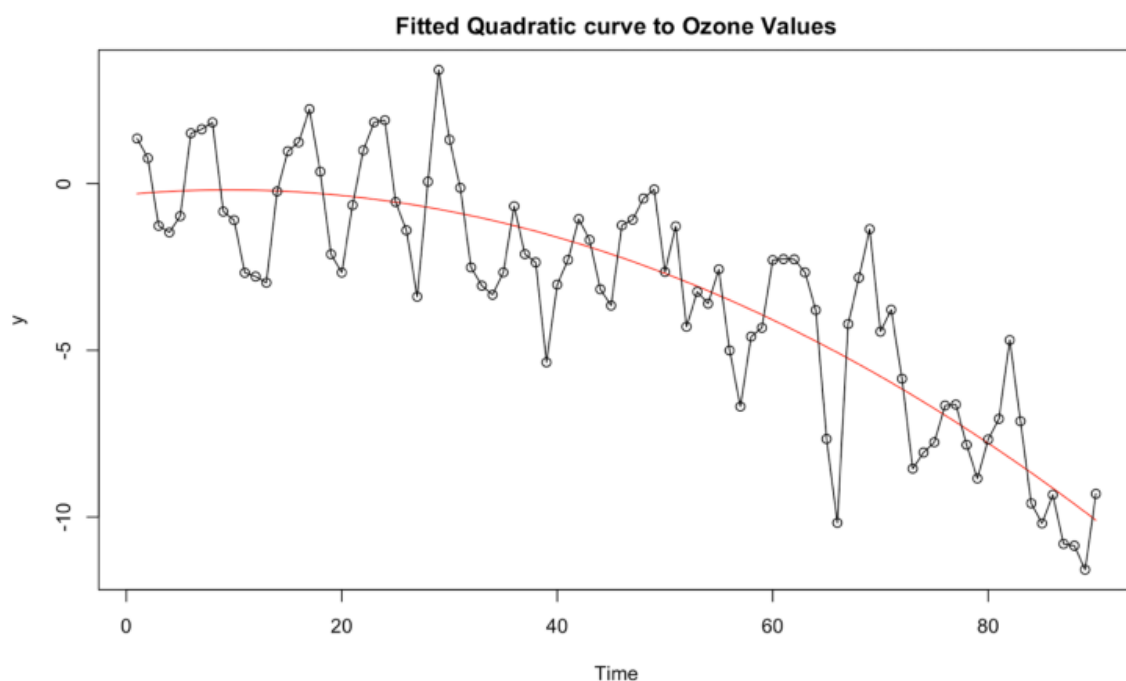


Figure 5: Quadratic Regression Model fitted with quadratic trend line

Based on Exhibit 02, the slope and intercept are estimated as $\beta_1 = -5.733$ and $\beta_0 = 5.924$, $\beta_2 = -1.530$ respectively. The slope and intercept i.e. the coefficient of t and t^2 are statistically significant with 0.05 as the significance level. Further, a glimpse of quadratic trend line fitted on the series in figure 05 together with an adjusted R^2 value of 0.733 from exhibit 2 shows that 73.3% of the variation in the series is explained by the quadratic time trend fitted above. This is a slight improvement in the level of accuracy over the linear regression model. However, as seen in figure 05 there still exists large residual values at some points which indicates that the quadratic trend line is

also not able to fully capture the mean level in the series. With accuracy slightly improved, it is necessary to conduct an appropriate residual to warranty if the trend model is reasonably correct. Kindly note that this improved residual value comes with a cost of multi-collinearity due to the presence of t^2 component in the model (D.Cryer & Chan, 2008).

Residual Analysis

If the quadratic trend model is reasonable, the residual analysis should roughly behave like a true stochastic component. Following the nature of white noise series, the residual must roughly behave like an independent random variable with zero means. This requires a close examination of the standardized residual plots.

Time Series plot of Standardized residuals

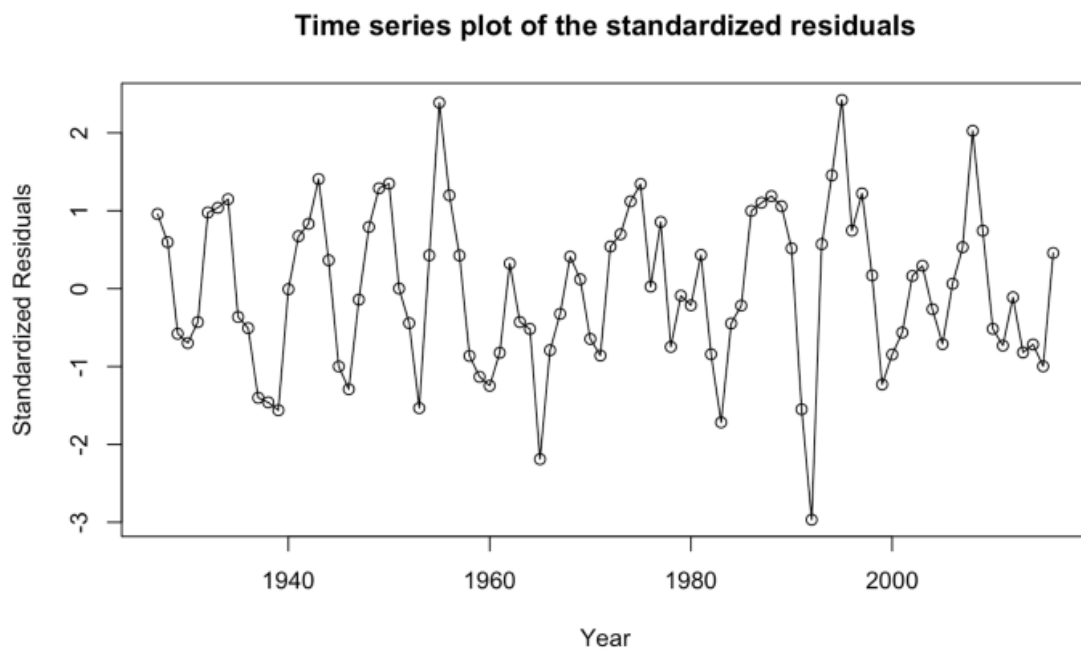


Figure 6: Standard Residual Plot of Quadratic Regression Model

As per figure 06, the standard residuals of the quadratic trend model hang together too much for the white noise – the plot seems smooth over the first half of the residual series. Further, more variation is observed in the last third of the series than first two-third of the series. This can be considered as the limitation of the model with standard residuals not exactly following the stochastic trend. We expect the ACF plot to show some significant lags.

Histogram of Standardized Residuals

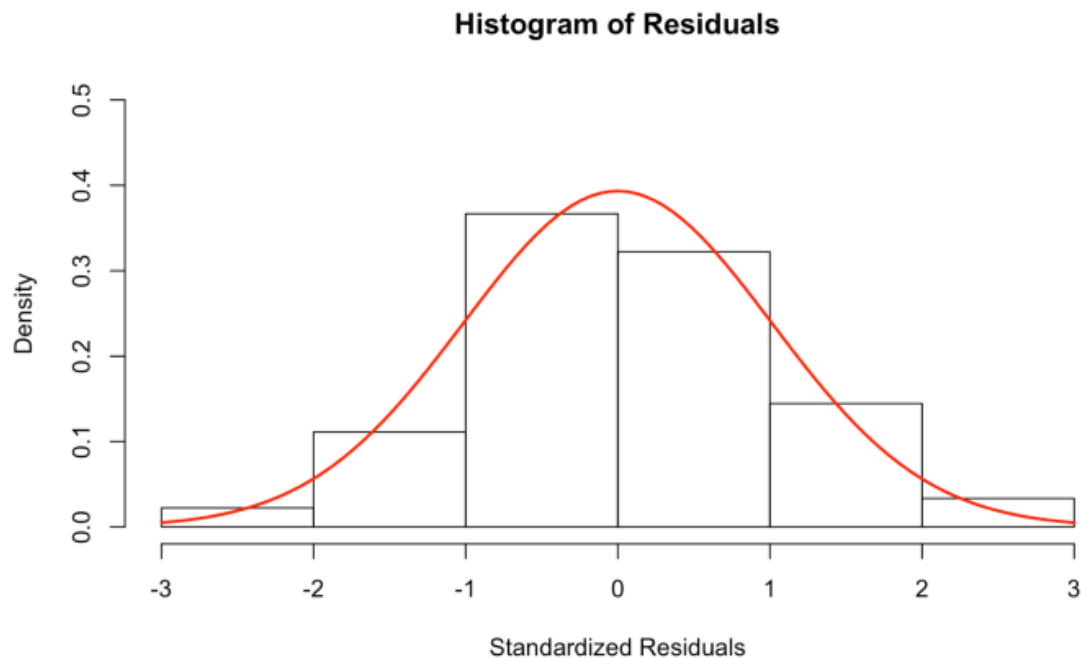


Figure 7: Histogram plot of Standard Residual Model

As in Figure 07, the histogram plot of standard residuals seems somewhat symmetric and tails off at both the high and low ends as observed in the normal distribution. Further, a closer assessment of normality can be conducted by Q-Q plot in figure 8 below.

Q-Q Plot to check the normality of Standardized residuals

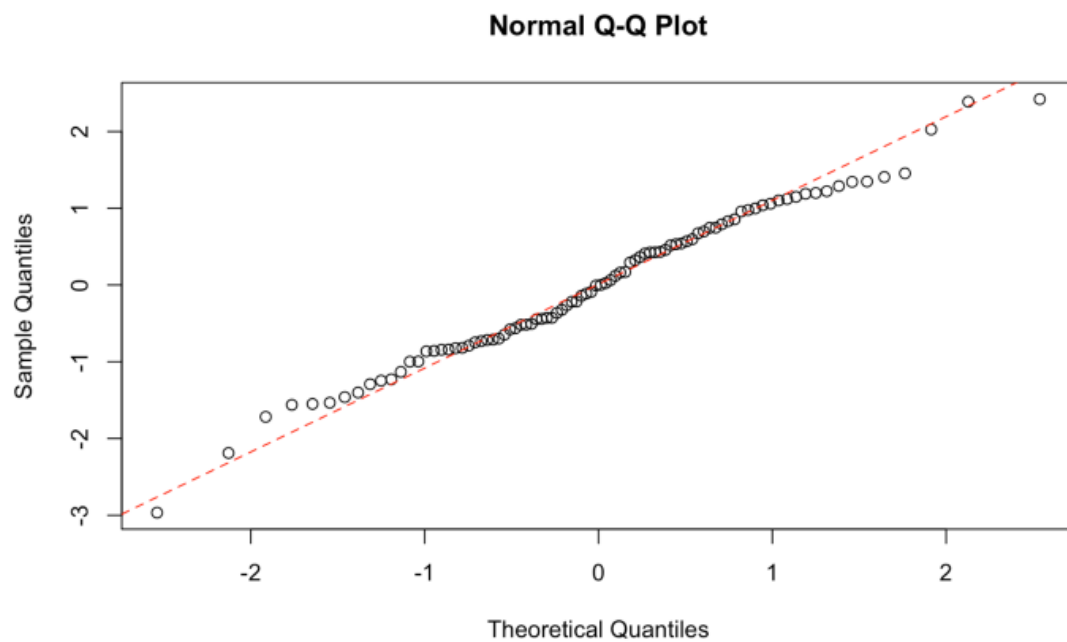


Figure 8: QQ Plot of Standardised Residual of Quadratic Trend Model

The Q-Q plot of standard residuals of quadratic trend model in figure 8 follows approximately a straight line as observed in the normally distributed data. This further support the normal distribution assumption of the stochastic component in the model. Finally, to avoid graphical interpretation error and more strong evidence of normality distribution, the Shapiro Wilk Test below provides an excellent test of normality.

Normality Test – Shapiro Wilk Test

Shapiro-wilk normality test

```
data: res.model.Ozone_thickness.qa  
W = 0.98889, p-value = 0.6493
```

Exhibit 3: Normality Test of Standard Residual of Quadratic Trend Model

The p-value of the Shapiro Wilk normality test in exhibit 3 is 0.64 at 5% significance level. This indicates that we fail to reject the null hypothesis that the stochastic component of this model is normality distributed.

ACF plot of Standardized residuals

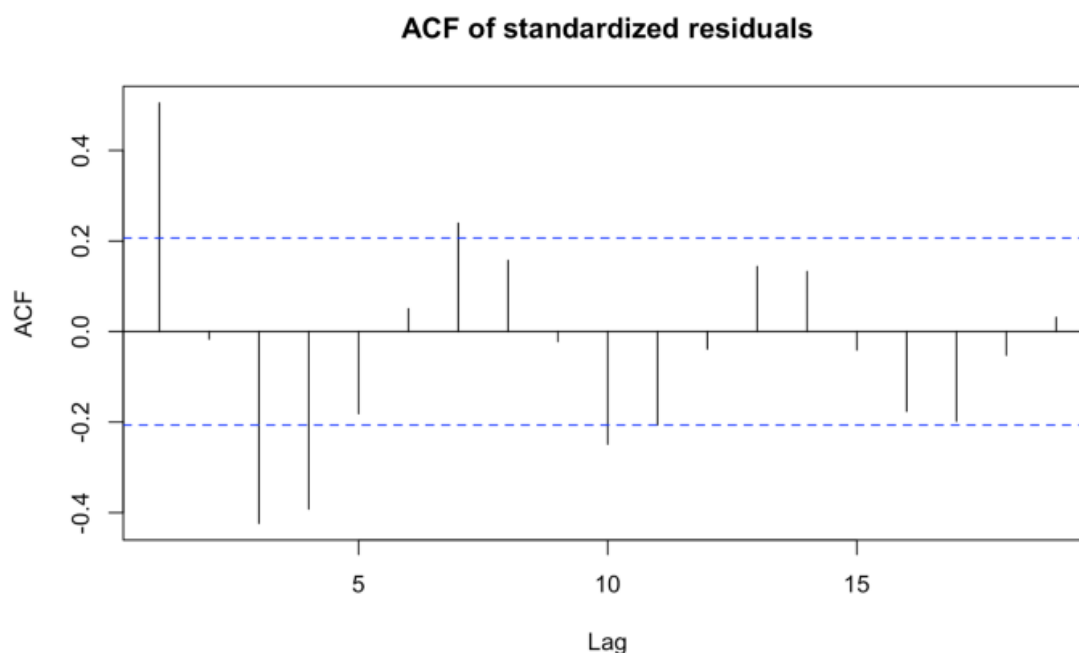


Figure 9: ACF plot for standardized residual of Quadratic Trend Model

From figure 09, as expected the ACF plot shows significant correlation at lag1, lag 3 and lag 4. This is not what we expect from the white noise series. However, considering close to normal distribution from Q-Q plot, histogram plot and normality test of standard residual from the above analysis and almost similar residual nature of linear trend model with improved R2 value we tend to move forward with quadratic trend model for forecast analysis.

Forecast Analysis

	fit	lwr	upr
1	-10.34387	-14.13556	-6.552180
2	-10.59469	-14.40282	-6.786548
3	-10.84856	-14.67434	-7.022786
4	-11.10550	-14.95015	-7.260851
5	-11.36550	-15.23030	-7.500701

Exhibit 4: Ozone Layer Thickness Forecast of from year 2017 to 2021

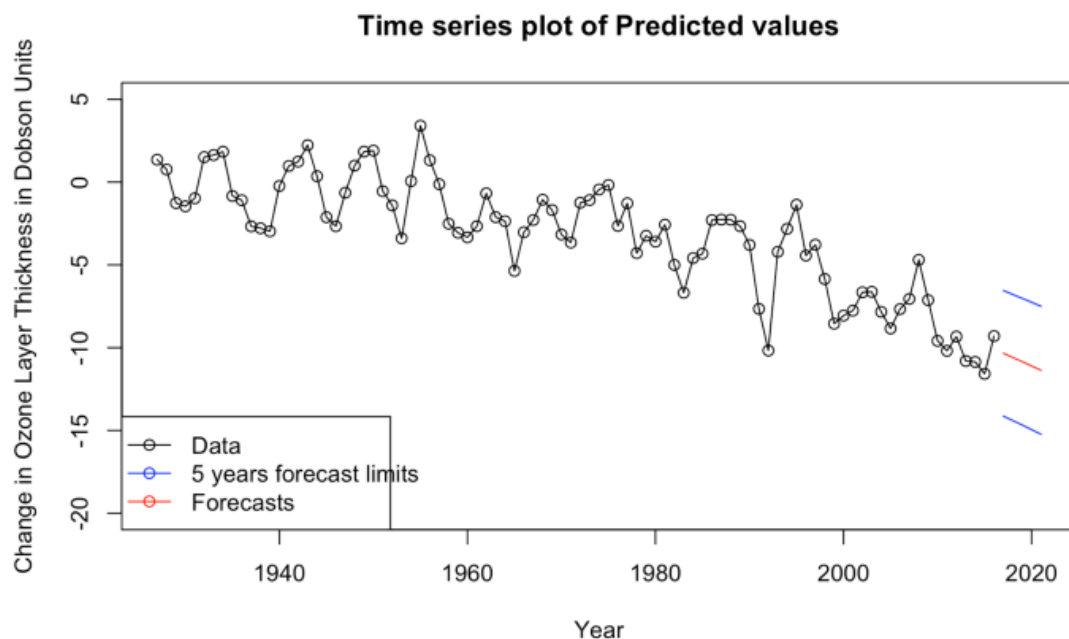


Figure 10: Prediction plot of Ozone Layer Thickness from year 2017 to 2021

Exhibit 4 and figure 10 shows the forecast and prediction plot of the ozone layer thickness respectively from year 2017-2021. Fit indicated the predicted value with lwr and upr indicating the 95% confidence interval of the predicted values.

Section 2

This section is based on identification, selection and interpretation of stochastic trend models. The section mainly demonstrates the use of specification tools such as ACF, PACF, EACF, and BIC table to select the order of ARIMA (p, d, q) models.

Referring to the plot of the Ozone layer thickness dataset in Figure 01 and 02, we observe changing variance around the 0 point in the series, a non-stationary behaviour with a declining trend. Further, the series plot shows a little cyclic nature. It also lacks the presence of significant intervention point. Trend, non-stationary behaviour and seasonality is more apparent from the ACF plot in Figure 11.

ACF and PACF Plot of the dataset –

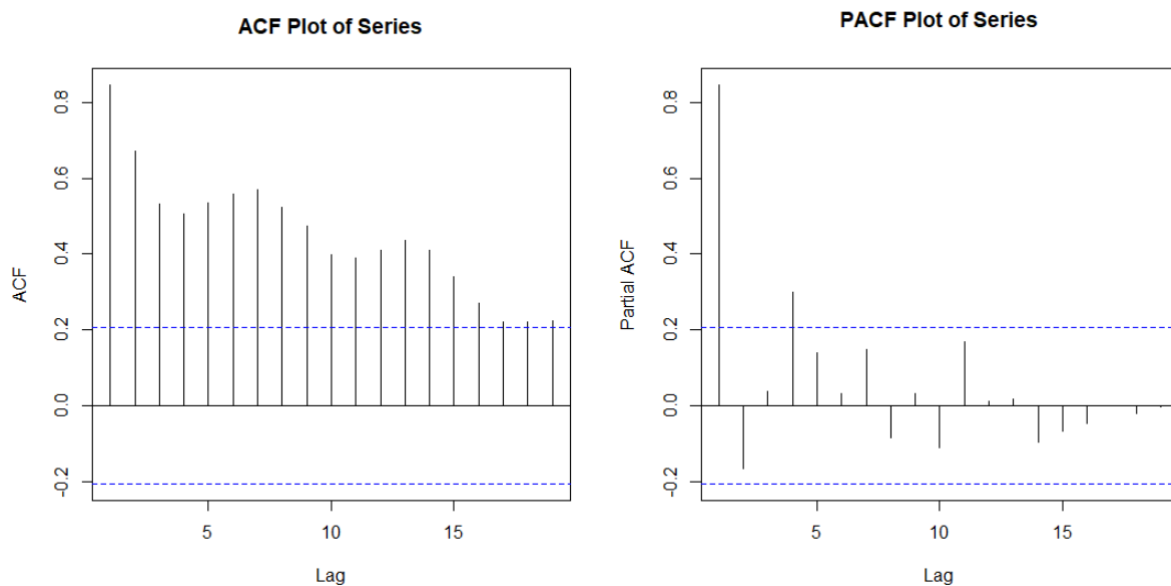


Figure 11: ACF and PACF Plot of Ozone Layer Thickness

Slop decay in ACF plot in figure 11 represents non-stationarity which shows declining correlations. The non-stationarity is confirmed by the augmented DK Fuller test below in exhibit 5.

Stationary Test of the dataset –

Augmented Dickey-Fuller Test

```
data: Ozone_thickness.ts2
Dickey-Fuller = -3.2376, Lag
order = 4, p-value = 0.0867
alternative hypothesis: stationary
```

Exhibit 5: DK Fuller Test for normality test of Ozone Layer Thickness Dataset

The DK- Fuller test in exhibit 5 shows p-value = 0.0867 at 5% significance level. This shows that we can we cannot reject the null hypothesis stating that the series is non-stationary.

Q-Q Plot to check the normality of series

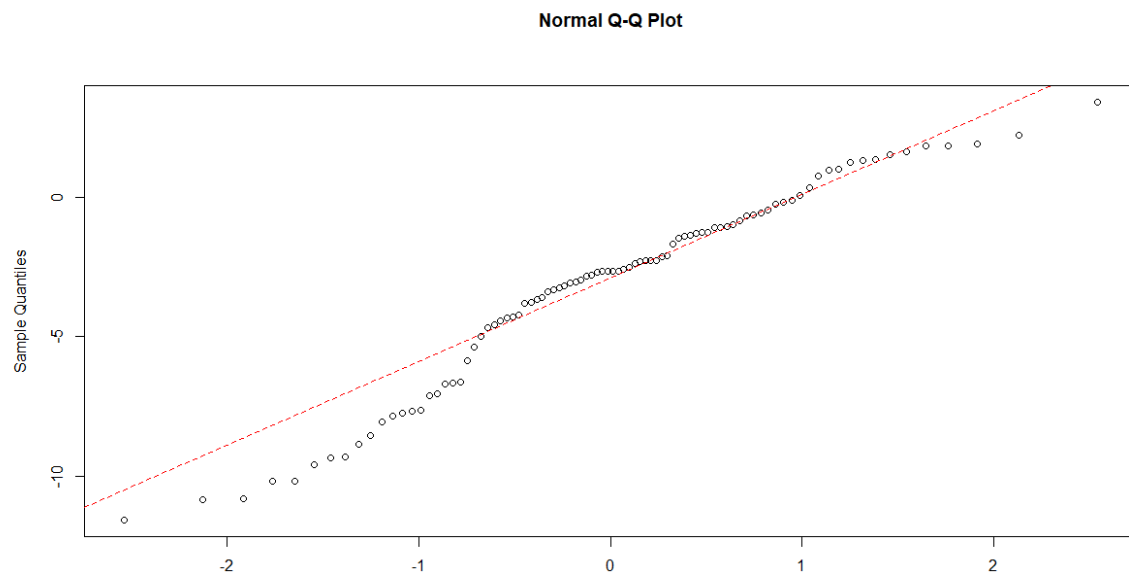


Figure 12: QQ Plot of Ozone Layer Thickness Dataset

Apart from non-stationary behaviour in the ACF plot in Figure 11, the QQ plot of the series in figure 12 also represents a slight deviation of data from normality distribution. Thus requires us to conduct the power transformation namely Box-Cox transformation to handled skewed data. However, the lambda value of 1.14 obtained from Box-Cox transformation disregards the need for power transformation. We are now required to difference the series to handle non-stationarity and find the appropriate order to the model.

Lag 1 Differenced series

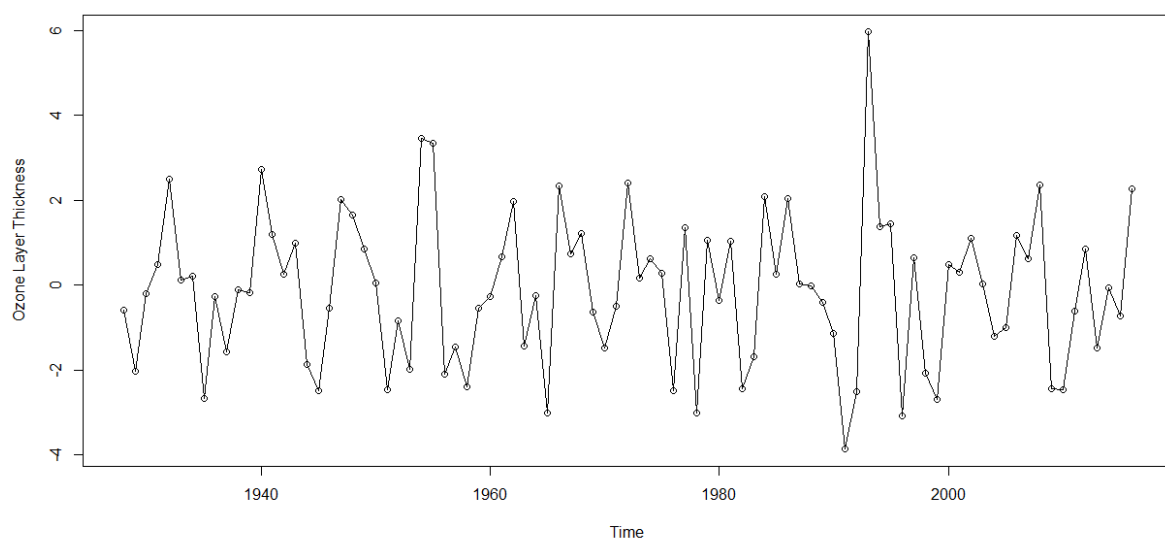


Figure 13: Time series plot of first order differenced series

Plot of the first order differenced series in figure 13 indicates that the non-stationary in the series is removed which can be further tested by ADF unit root test in exhibit 6.

ADF Unit Root Test

Title:
Augmented Dickey-Fuller Test

Test Results:
PARAMETER:
Lag Order: 6
STATISTIC:
Dickey-Fuller: -5.0757
P VALUE:
0.01

Description:
Tue Apr 14 16:30:45 2020 by user: Avinash Matani

Exhibit 6: ADF test of Differenced Series

To consider AR for the once differenced series, we used lag order from ar() function and applied to augmented DK fuller test as shown in exhibit 6. The p-value of 0.01 at 5% significance level indicates that we have rejected the null hypotheses which are series non-stationary. This confirms that the differenced series is now stationary and can be used for estimating the order to models.

Normality Test of Differenced Series

Shapiro-wilk normality test

data: diff.Ozone_thickness.ts2
W = 0.97907, p-value = 0.1606

Exhibit 7: Shapiro- Wilk Normality test of differenced series

The p-value of the normality test of differenced series in exhibit 7 is 0.16 at 5% significance. This indicates that we cannot reject null hypothesis which is series is normal. This indicates that first order differencing was able to make the distribution normal.

ACF and PACF plot of Differenced dataset

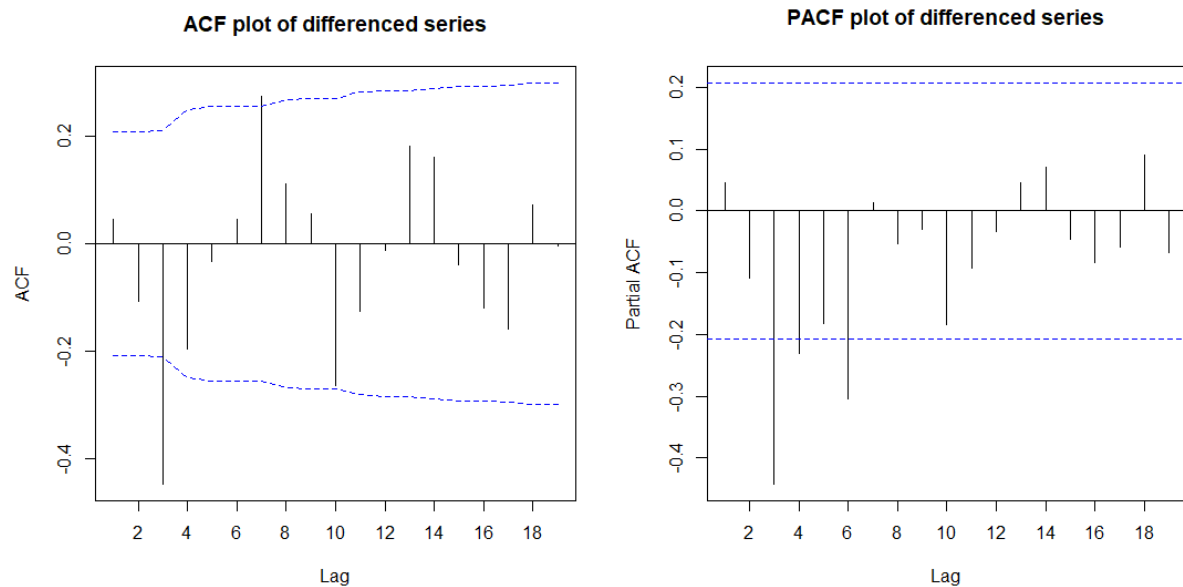


Figure 14: ACF and PACF plot of differenced series

The ACF plot in Figure 14 is constructed using the reference critical bounds taking into account that uncertainty increases with higher lags. The ACF plot indicates the presence of significant ACF at lag 3 and the rest of the correlation are insignificant which represents cut-off, thus we consider IMA (1,3) as a preferred model.

Further, the PACF represents 2 significant lags at lag 3 and lag 4. Thus, we chose ARI (3,1) and ARI (4,1) and ignoring all significant lags after this due to presence of insignificant lag at lag 5.

EACF Plot of the differenced dataset

AR/MA		0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	o	o	x	o	o	o	x	o	o	x	o	o	o	o	o
1	x	o	x	o	o	o	o	o	o	x	o	o	o	o	o
2	x	o	x	o	o	o	x	o	o	x	o	o	o	o	o
3	x	o	x	o	o	x	o	o	o	o	o	o	o	o	o
4	x	o	o	x	o	x	o	o	o	o	o	o	o	o	o
5	x	x	x	x	o	x	o	o	o	o	o	o	o	o	o
6	o	o	o	x	x	o	o	o	o	o	o	o	o	o	o
7	o	o	o	x	o	o	o	o	o	o	o	o	o	o	o

Exhibit 8: EACF Plot of differenced series

Although there is no clear vertex in the EACF plot in exhibit 8, we will take the row corresponding to $p = 4$ as the vertex and include ARIMA(4,1,1), ARIMA(3,1,1), ARIMA(4,1,2), and ARIMA(3,1,3) as well into the set of possible models.

BIC Table

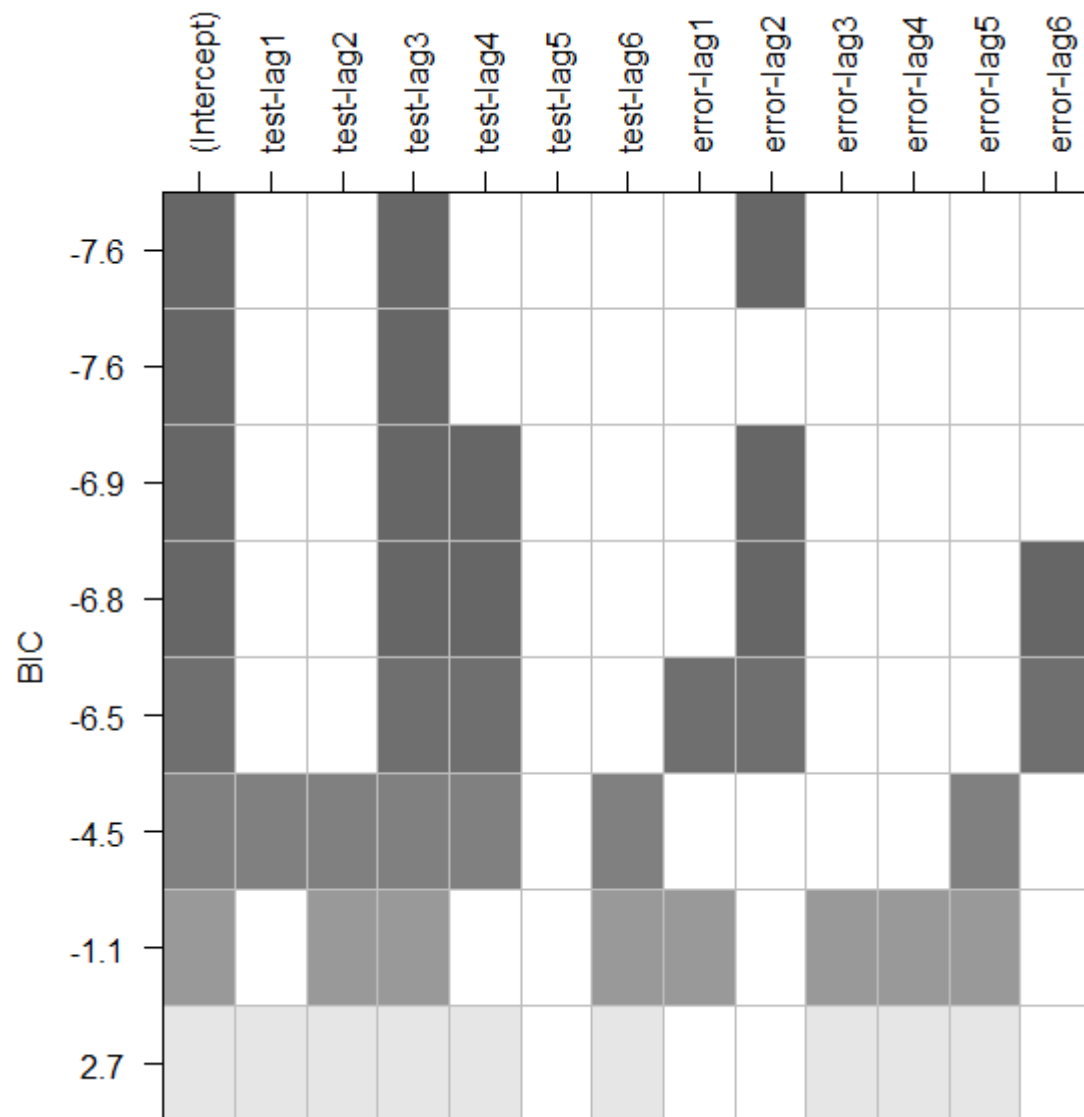


Exhibit 9: BIC plot differenced series

Conclusion

We successfully described, modelled and estimated the deterministic and stochastic trend models for time series dataset which represent ozone layer thickness from the year 1927 to 2016. For the deterministic trend model analysis, we finalized the quadratic term based regression model among other deterministic trend models and predicted the ozone layer thickness values from the year 2017 – 2021. Further, as a part stochastic trend model analysis, we successfully demonstrated the use of identification tools including ACF, PACF, EACF and BIC plots and differencing and series transformation concepts to suggest ARIMA (4,1,1), ARIMA (3,1,1), ARIMA (4,1,2), and ARIMA (3,1,3) as a set of possible models for forecasting applications.

References

D.Cryer, J., & Chan, K.-S. (2008). Time Series Analysis with applications in R. Springer.

Dr. Kalaylioglu. Module 1 - Basic Plots, Examples, and Fundamental Concepts, RMIT University, School of Science

Dr. Kalaylioglu. Module 2 - Analysis of Trends, RMIT University, School of Science

Dr. Kalaylioglu. Module 3 - Models for Stationary Time series, RMIT University, School of Science

Dr. Kalaylioglu. Module 4 - Models for Nonstationary Time Series, RMIT University, School of Science

Dr. Kalaylioglu. Module 5 - Model Specification, RMIT University, School of Science

Assignment1- Time Series

s3692165-Avinash Matani

14 April 2020

Importing Dataset and converting into time series class

```
library(TSA)
library(tseries)
library(fUnitRoots)
Ozone_thickness<-read.csv("data1.csv", header = FALSE)
class(Ozone_thickness)
```

```
## [1] "data.frame"
```

```
Ozone_thickness.ts<- ts(Ozone_thickness, start = 1927)
class(Ozone_thickness.ts)
```

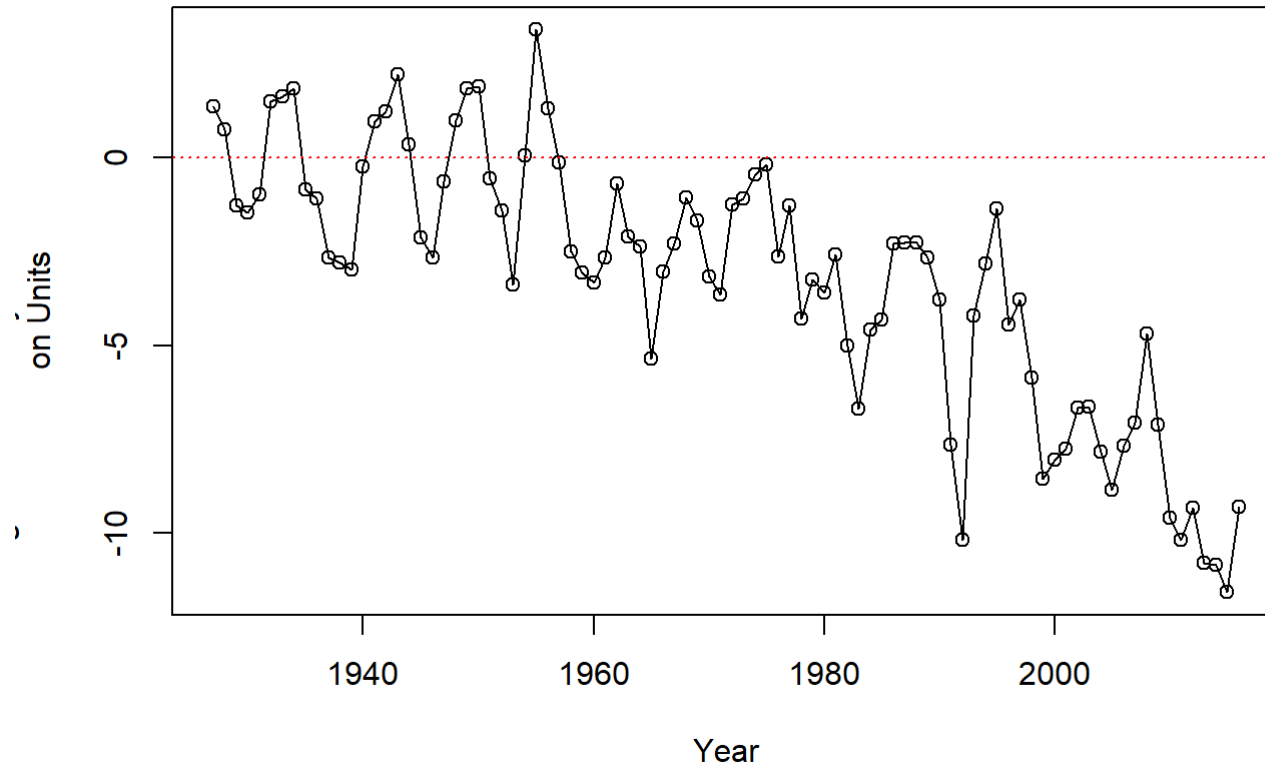
```
## [1] "ts"
```

```
View(Ozone_thickness.ts)
```

Time series plot

```
plot(Ozone_thickness.ts,type="o",xlab="Year", ylab="Change in Ozone Layer Thickness in Dobson Units",main="Changes in Thickness of Ozone layer")
abline(h=0, col="red",lty=3)
```

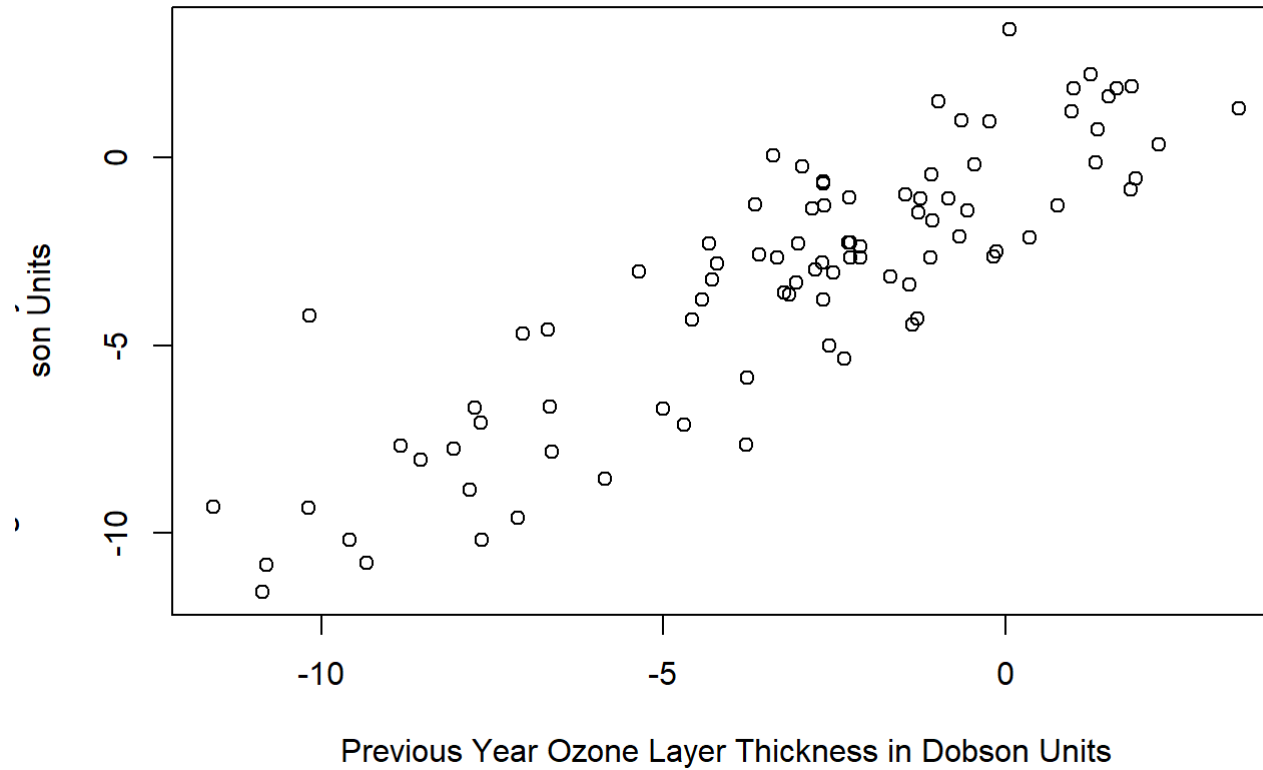
Changes in Thickness of Ozone layer



Scatter Plot

```
plot(y=Ozone_thickness.ts,x=zlag(Ozone_thickness.ts),ylab="Change in Ozone Layer Thickness in Dobson Units",  
     xlab="Previous Year Ozone Layer Thickness in Dobson Units",  
     main = "Scatter plot of Thickness of the Ozone Layer")
```

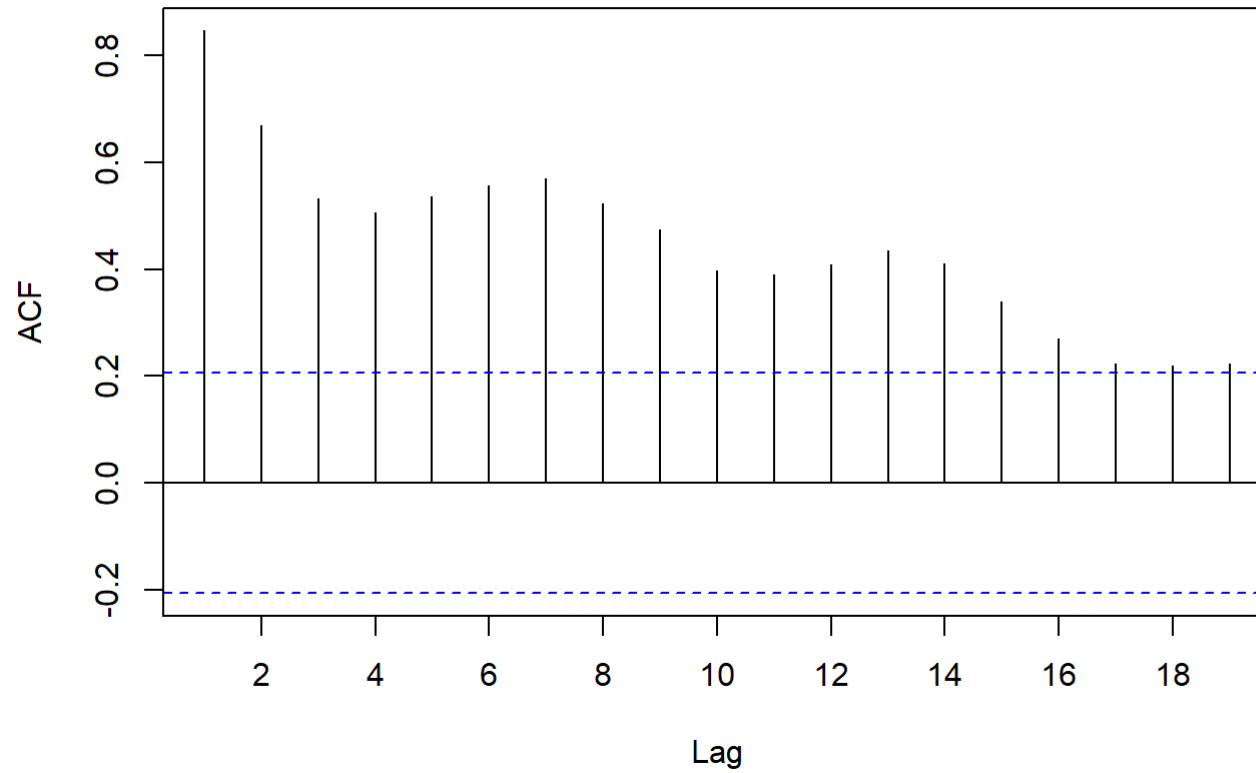
Scatter plot of Thickness of the Ozone Layer



ACF Plot of Ozone Layer Thickness Dataset

```
acf(Ozone_thickness.ts, xaxp=c(0,20,10))
```

Series Ozone_thickness.ts



```
model_ln = lm(Ozone_thickness.ts~time(Ozone_thickness.ts)) # label the linear trend model as model1
summary(model_ln)
```

```
##
## Call:
## lm(formula = Ozone_thickness.ts ~ time(Ozone_thickness.ts))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.7165 -1.6687  0.0275  1.4726  4.7940
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    213.720155   16.257158    13.15  <2e-16 ***
## time(Ozone_thickness.ts) -0.110029    0.008245   -13.34  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.032 on 88 degrees of freedom
## Multiple R-squared:  0.6693, Adjusted R-squared:  0.6655
## F-statistic: 178.1 on 1 and 88 DF, p-value: < 2.2e-16
```

Fitting linear trend model

```
model_ln = lm(Ozone_thickness.ts~time(Ozone_thickness.ts)) # label the linear trend model as model1
summary(model_ln)
```

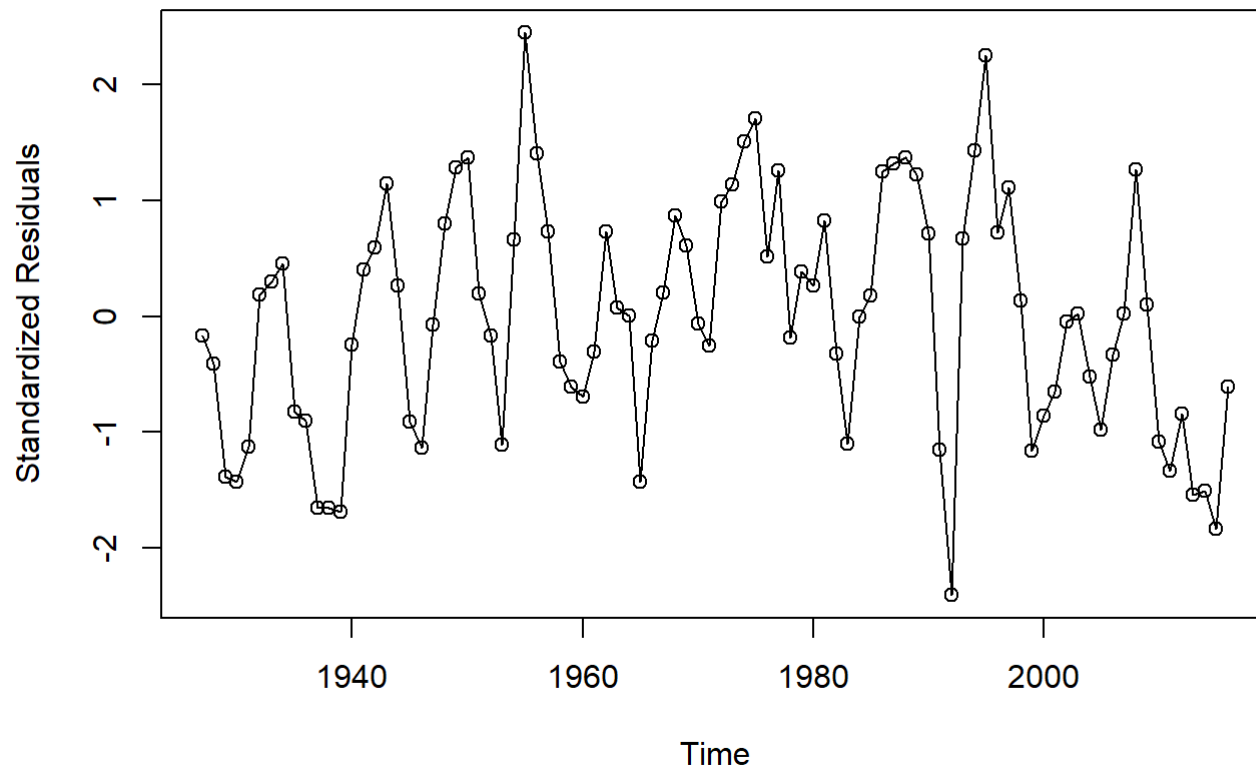
```
##
## Call:
## lm(formula = Ozone_thickness.ts ~ time(Ozone_thickness.ts))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.7165 -1.6687  0.0275  1.4726  4.7940
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    213.720155   16.257158    13.15  <2e-16 ***
## time(Ozone_thickness.ts) -0.110029    0.008245   -13.34  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.032 on 88 degrees of freedom
## Multiple R-squared:  0.6693, Adjusted R-squared:  0.6655
## F-statistic: 178.1 on 1 and 88 DF, p-value: < 2.2e-16
```

Checking standardized residuals of linear trend model

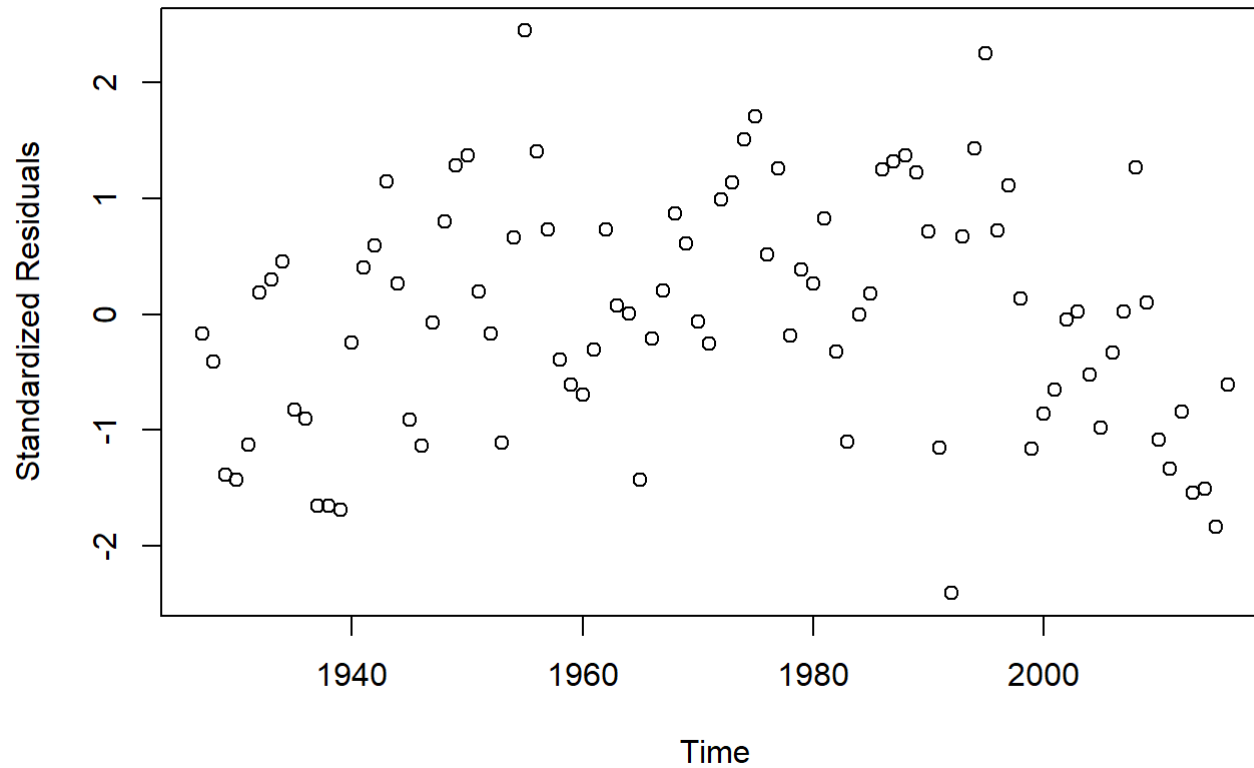
```
res.model_ln = rstudent(model_ln)
```

Time series plot of residual

```
plot(y = res.model_ln, x = as.vector(time(Ozone_thickness.ts)),
     xlab = 'Time', ylab='Standardized Residuals',type='o')
```

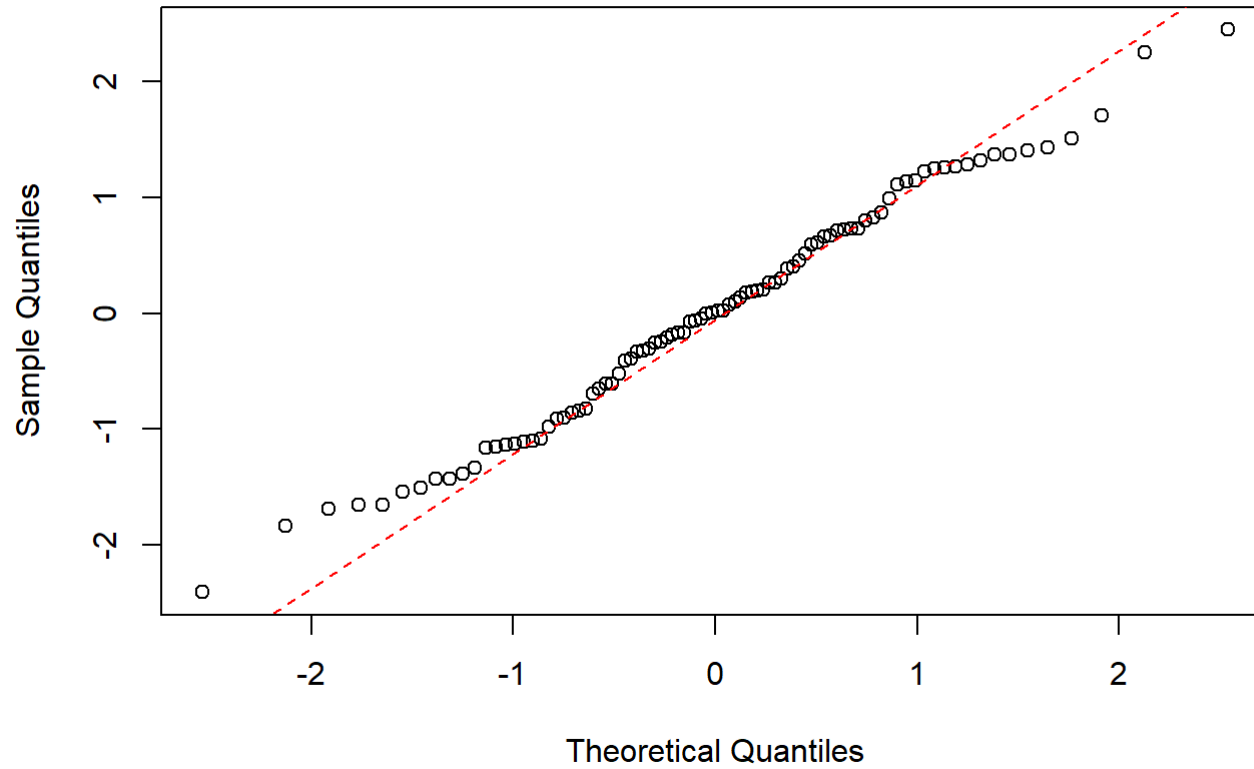
```
plot(y = res.model_ln, x = as.vector(time(Ozone_thickness.ts)),xlab = 'Time', ylab='Standardized Residuals',type='p')
```



Checking normality of standard residuals using qqplot

```
qqnorm(res.model_ln)  
qqline(res.model_ln, col = 2, lwd = 1, lty = 2)
```

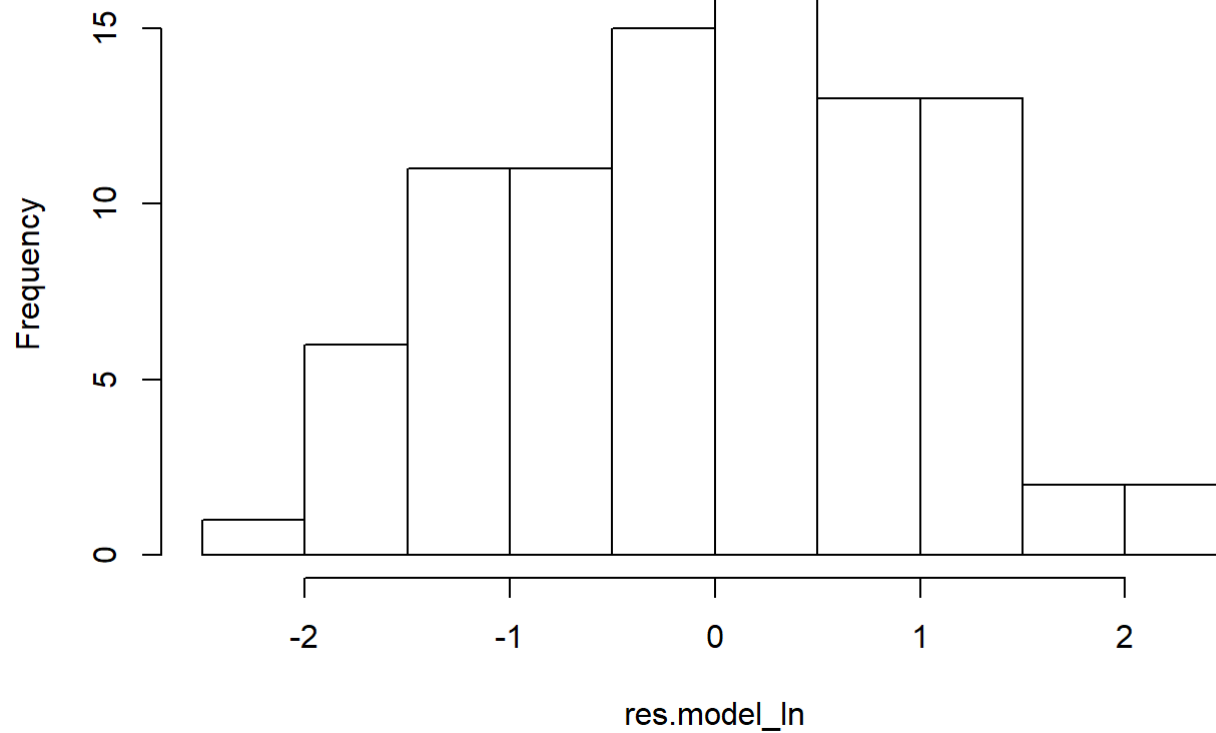
Normal Q-Q Plot



Histogram plot of linear trend model residual to check normality

```
hist(res.model_ln)
```

Histogram of res.model_ln



Normality check test of residual plot

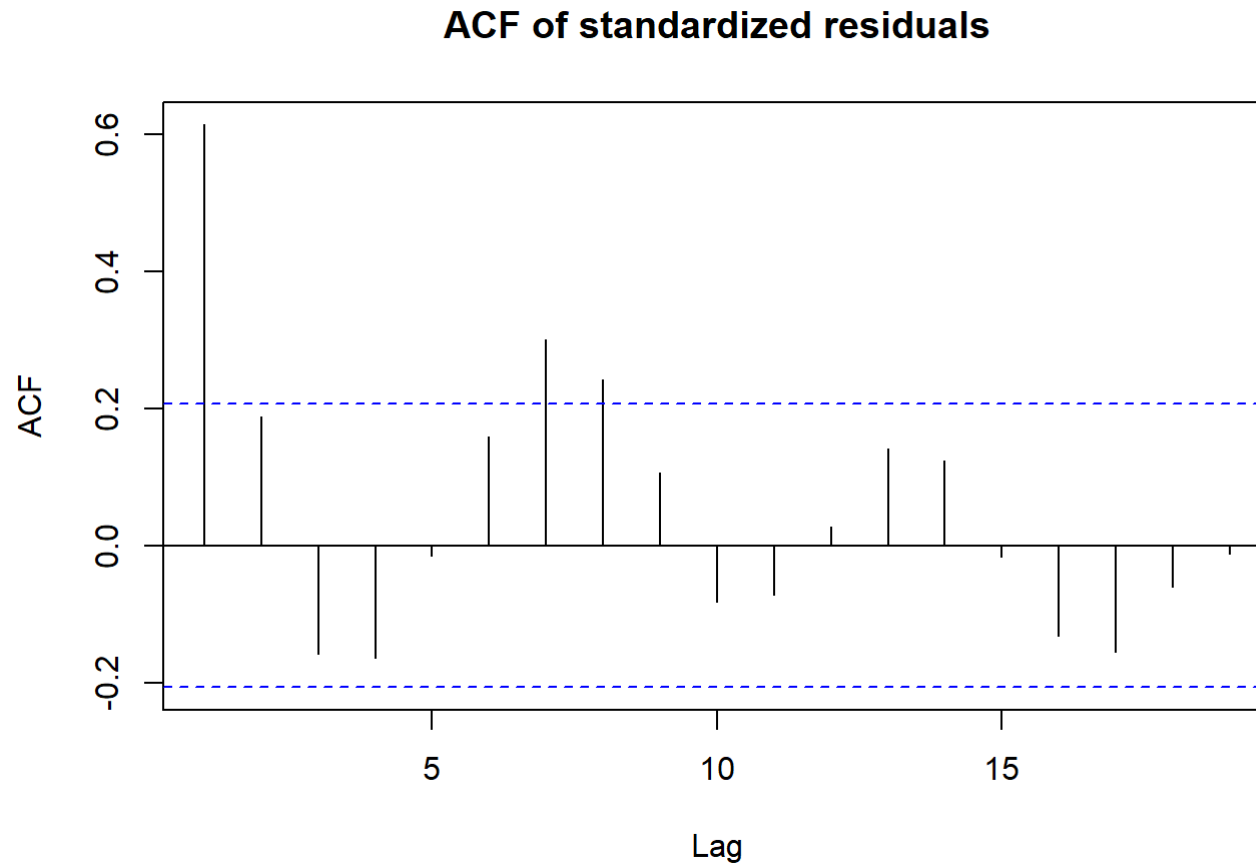
```
shapiro.test(res.model_ln)
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  res.model_ln  
## W = 0.98733, p-value = 0.5372
```

```
# The test also reject the normality assumption
```

ACF plot of linear trend model residual

```
acf(res.model_ln, main = 'ACF of standardized residuals')
```



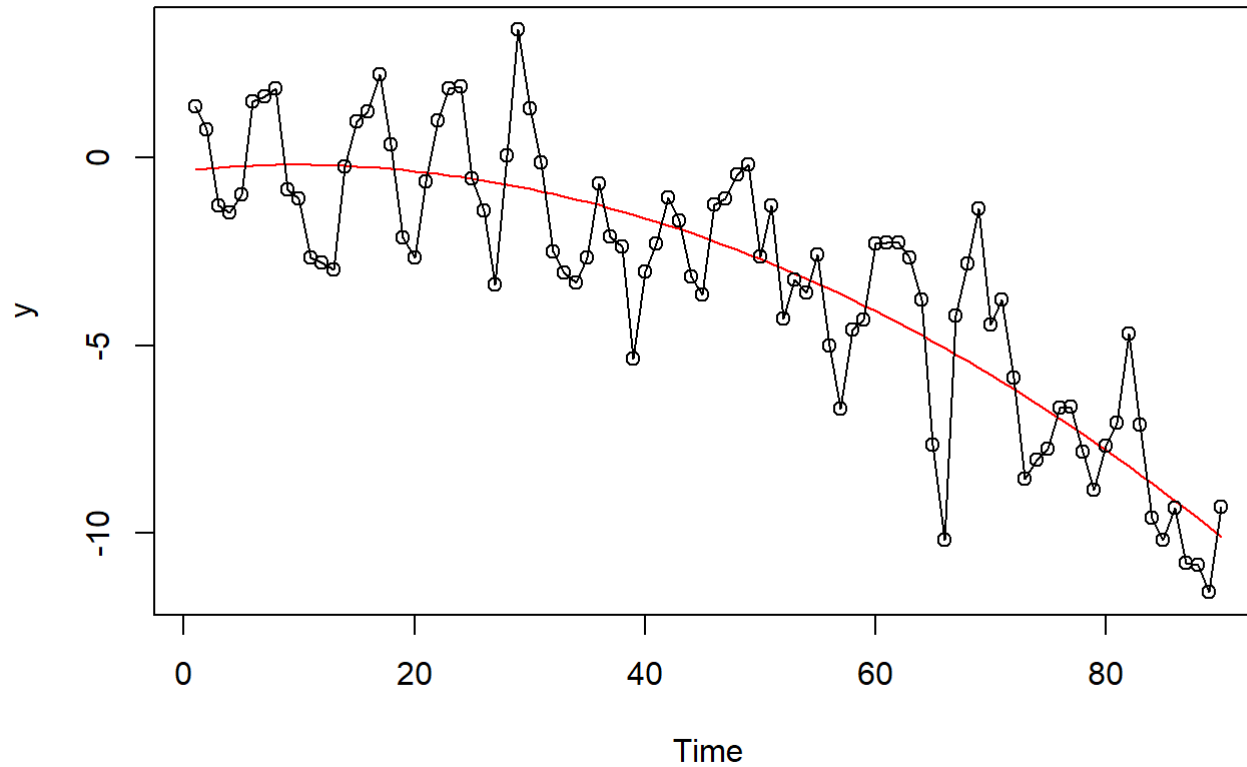
Quadratic Trend Model

```
t = time(Ozone_thickness.ts)
t2 = t^2
model.Ozone_thickness.qa = lm(Ozone_thickness.ts~ t + t2) # labelling the quadratic trend model
summary(model.Ozone_thickness.qa)
```

```
##
## Call:
## lm(formula = Ozone_thickness.ts ~ t + t2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.1062 -1.2846 -0.0055  1.3379  4.2325
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.733e+03  1.232e+03  -4.654 1.16e-05 ***
## t              5.924e+00  1.250e+00   4.739 8.30e-06 ***
## t2            -1.530e-03  3.170e-04  -4.827 5.87e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.815 on 87 degrees of freedom
## Multiple R-squared:  0.7391, Adjusted R-squared:  0.7331
## F-statistic: 123.3 on 2 and 87 DF,  p-value: < 2.2e-16
```

```
# We got an improved R2 value with coefficient of t and t2 being statistically significant
plot(ts(fitted(model.Ozone_thickness.qa)), ylim = c(min(c(fitted(model.Ozone_thickness.qa),as.vector(Ozone_thickness.ts))),
, max(c(fitted(model.Ozone_thickness.qa),as.vector(Ozone_thickness.ts)))),ylab='y',main = "Fitted Quadrati
c curve to Ozone Values",col="red")
lines(as.vector(Ozone_thickness.ts),type="o")
```

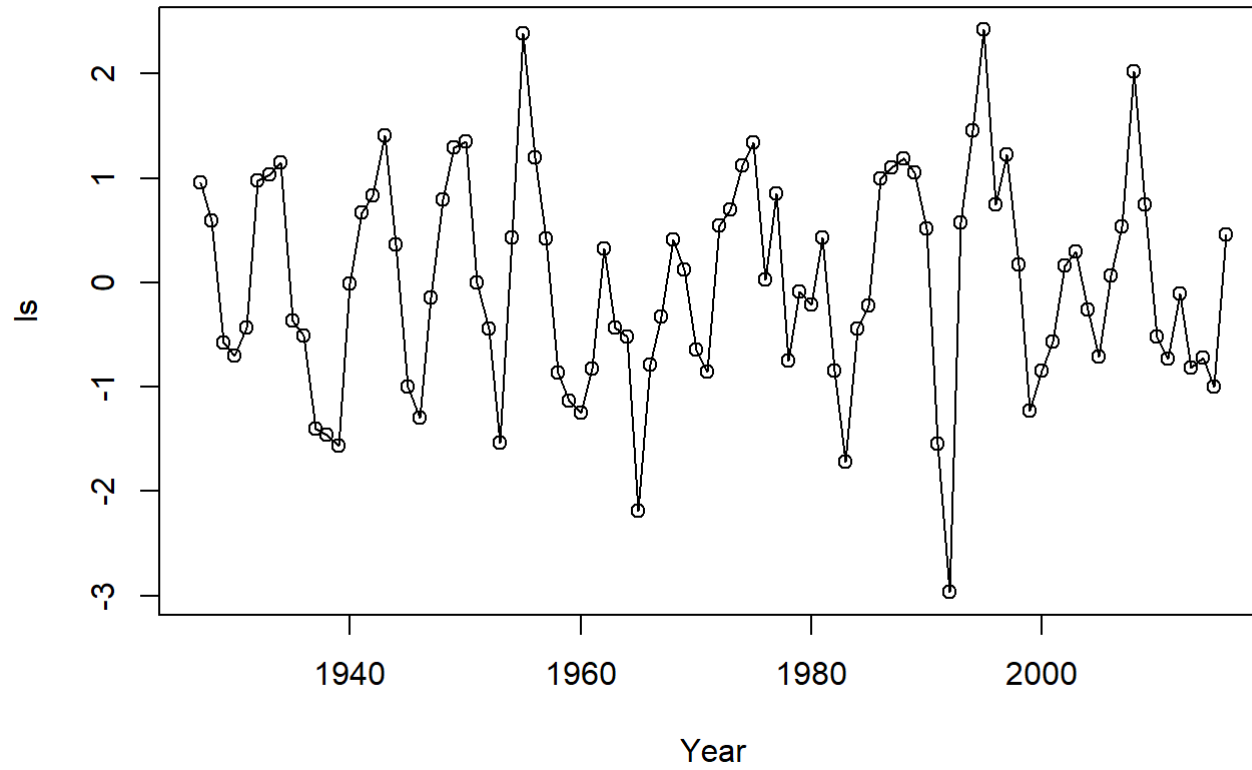
Fitted Quadratic c curve to Ozone Values



Checking standardized residuals of Quadratic Trend Model

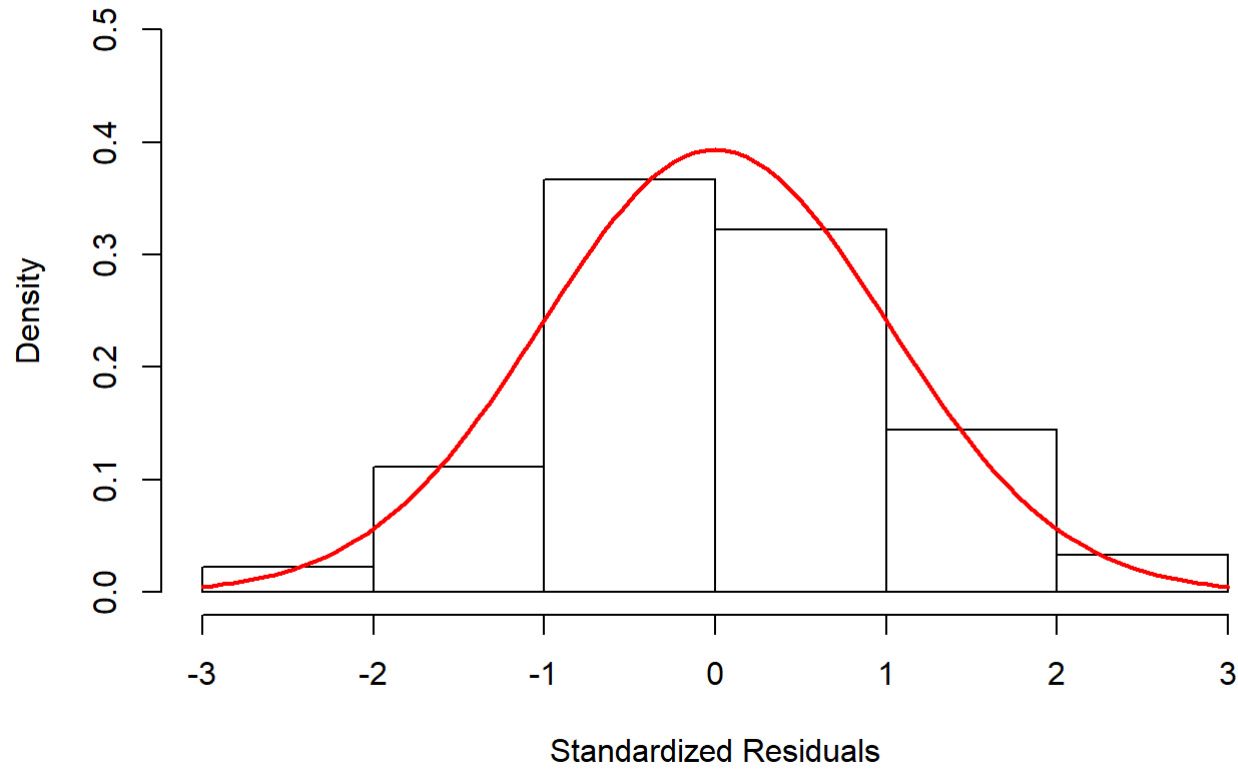
```
res.model.Ozone_thickness.qa = rstudent(model.Ozone_thickness.qa)
plot(y=res.model.Ozone_thickness.qa ,x=as.vector(time(Ozone_thickness.ts)),ylab='Standardized Residuals',xlab='Year',type='o', main = "Time series plot of the standardized residuals")
```

Time series plot of the standardized residuals



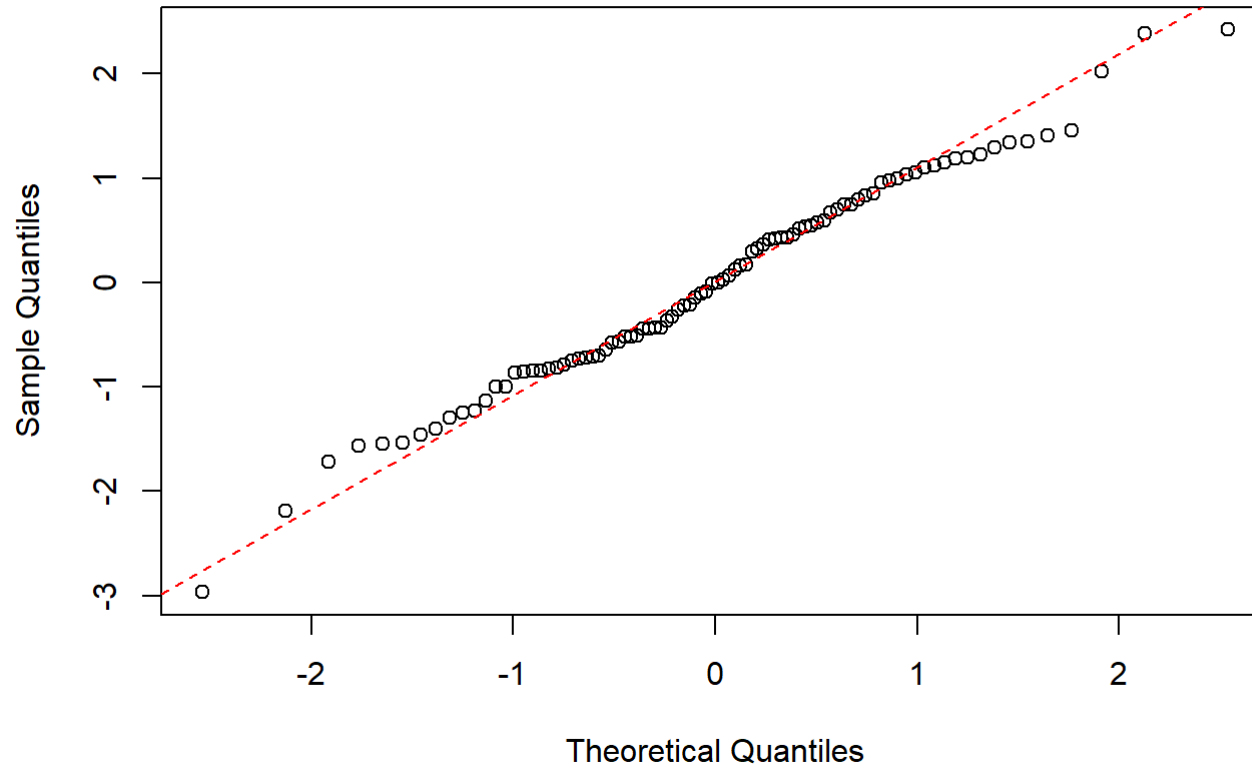
Histogram plot of residual to check normality

Hi
stogram of Residuals



QQplot of residuals of Quadratic Trend model

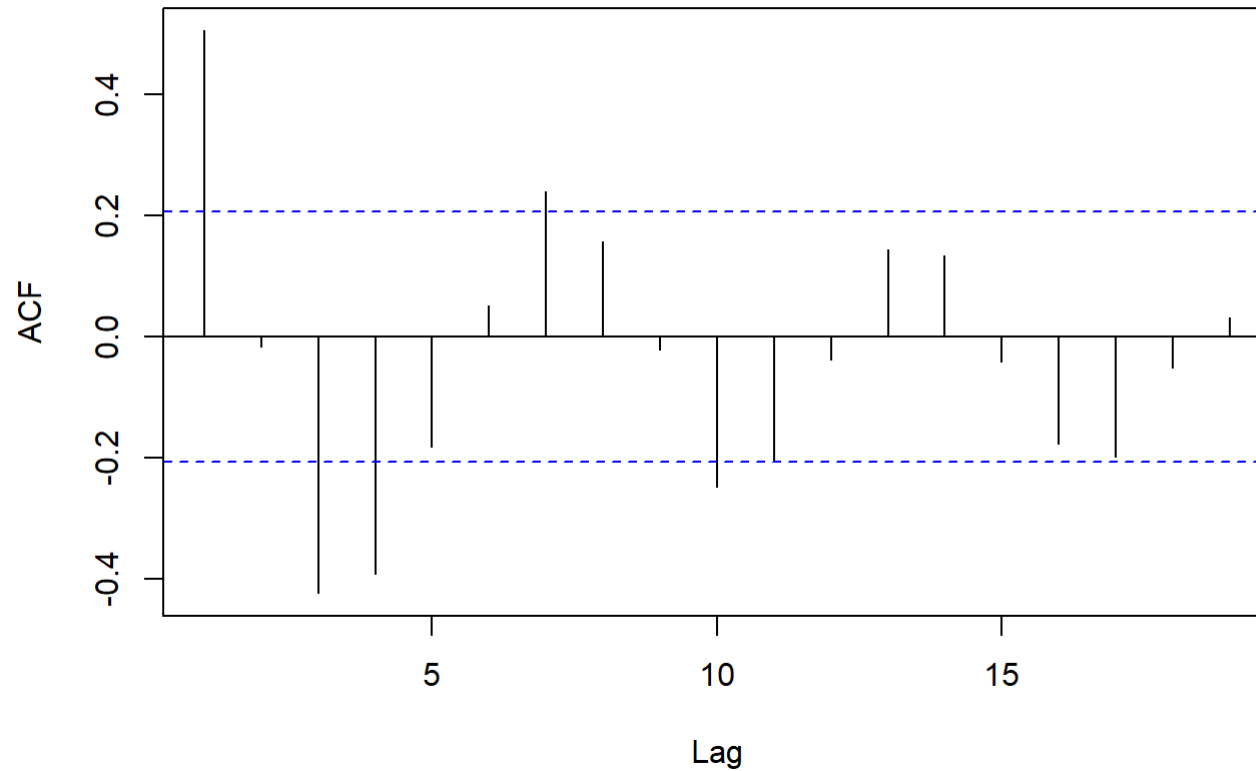
Normal Q-Q Plot



Normality test of Residual of Quadratic Trend Model

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  res.model.Ozone_thickness.qa  
## W = 0.98889, p-value = 0.6493
```

Series res.model.Ozone_thickness.qa

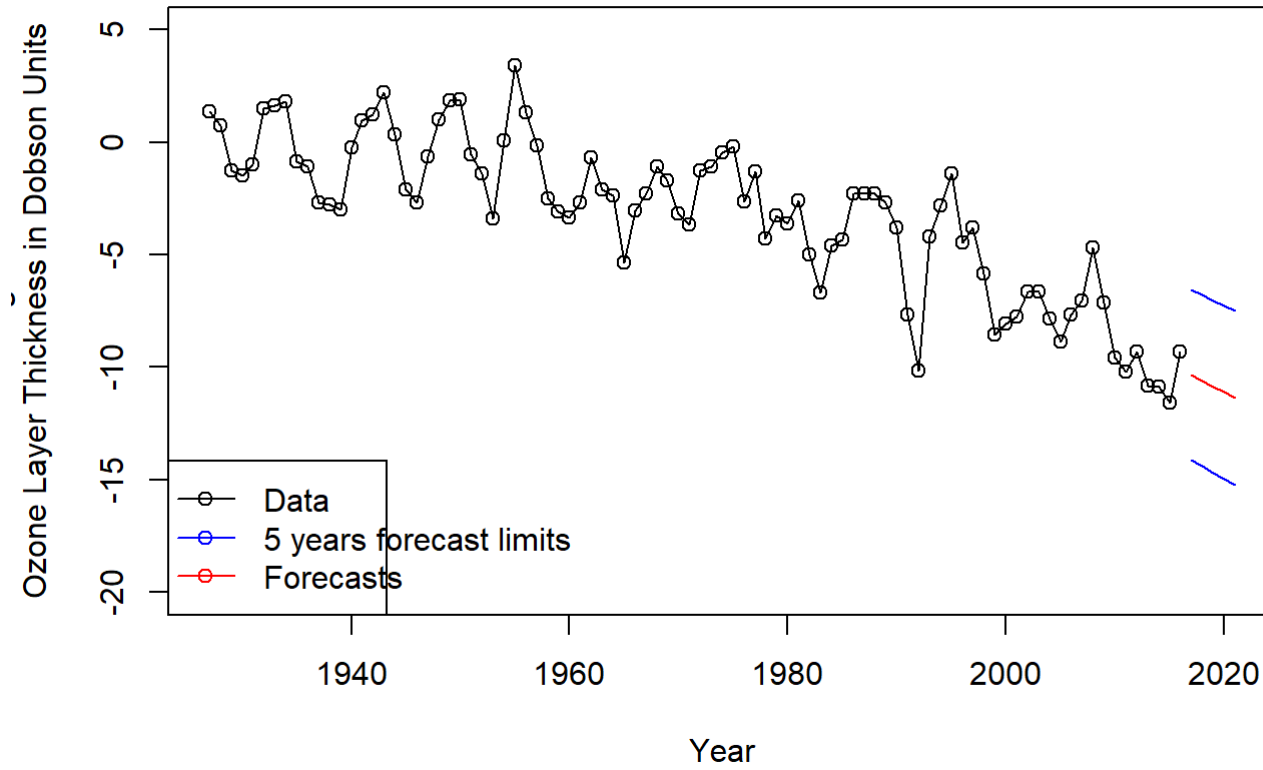


Prediction based on chosen Quadratic Trend Model

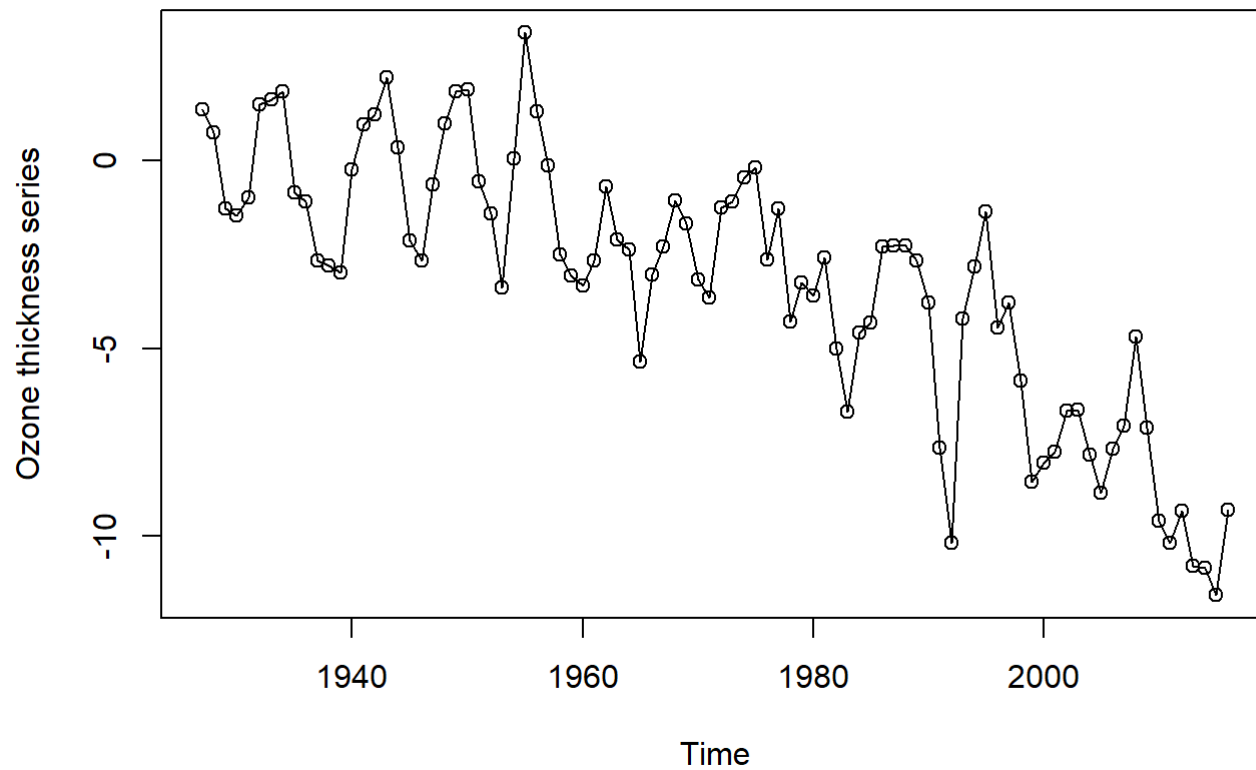
```
##      fit      lwr      upr
## 1 -10.34387 -14.13556 -6.552180
## 2 -10.59469 -14.40282 -6.786548
## 3 -10.84856 -14.67434 -7.022786
## 4 -11.10550 -14.95015 -7.260851
## 5 -11.36550 -15.23030 -7.500701
```

Prediction plot of Quadratic Trend Model

Time series plot of Predicted values

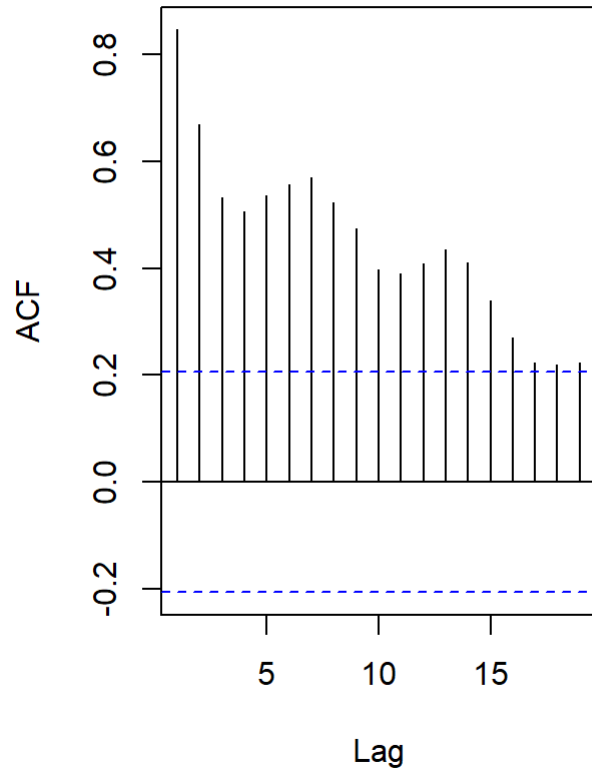


Creating copy of time series for task 2

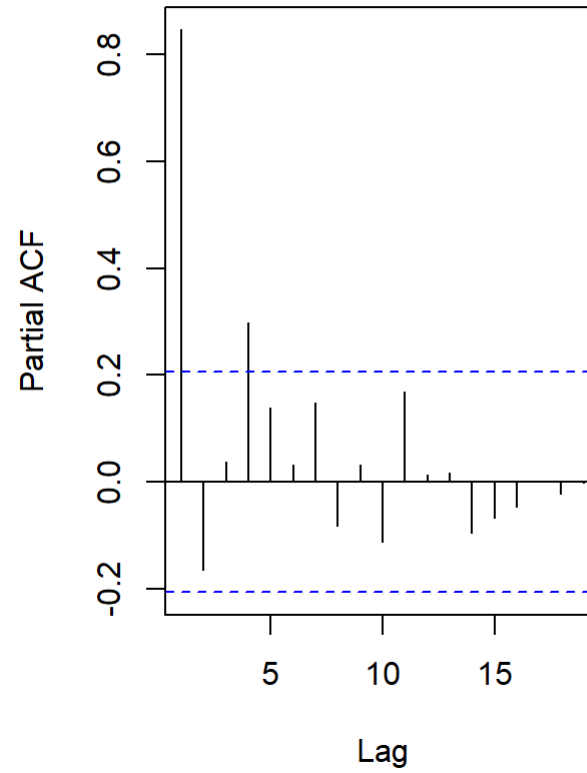


ACF and PACF plot of Ozone Layer Thickness Dataset

ACF Plot of Series

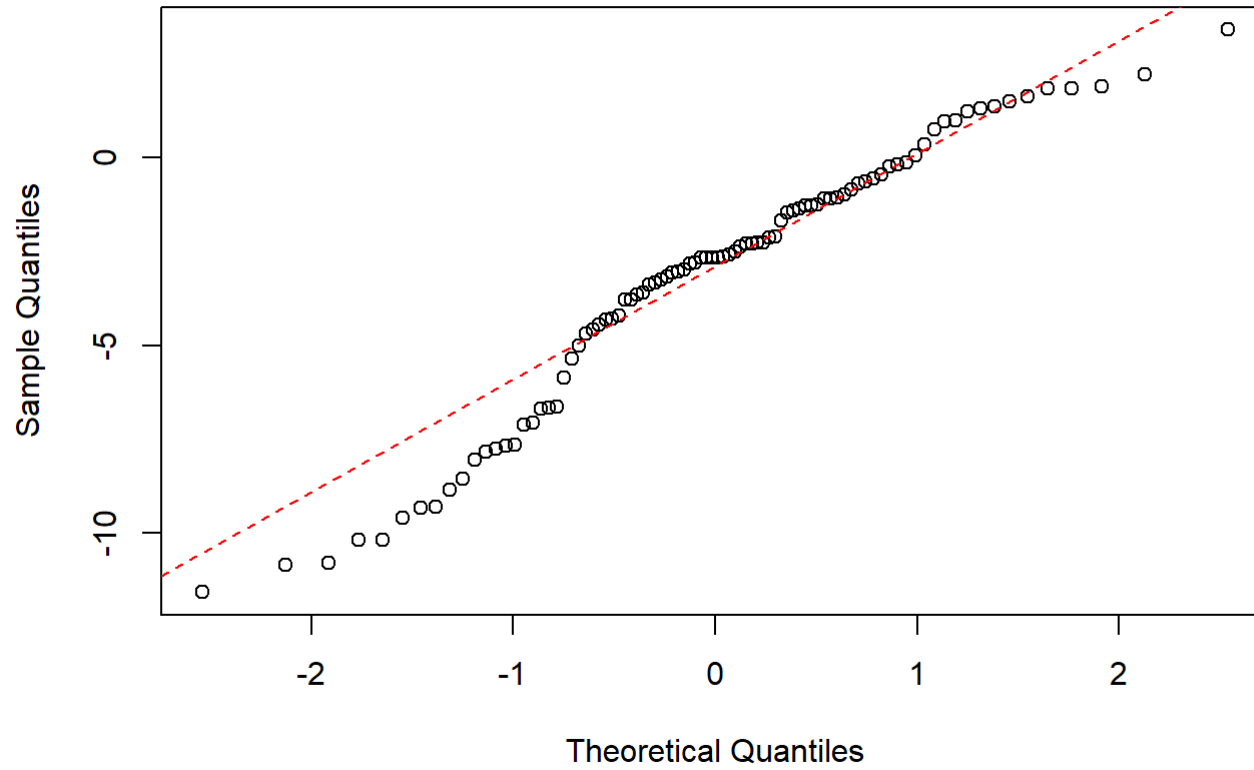


PACF Plot of Series

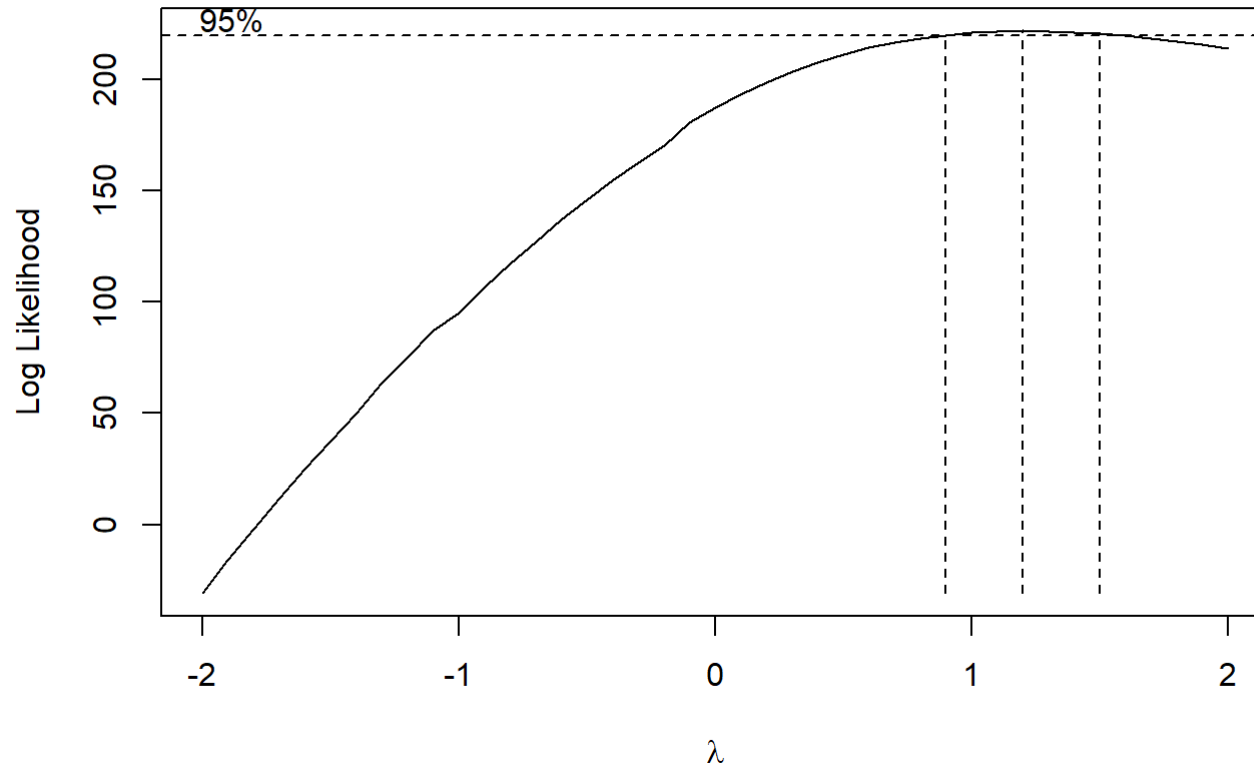


QQ plot of time series dataset

Normal Q-Q Plot



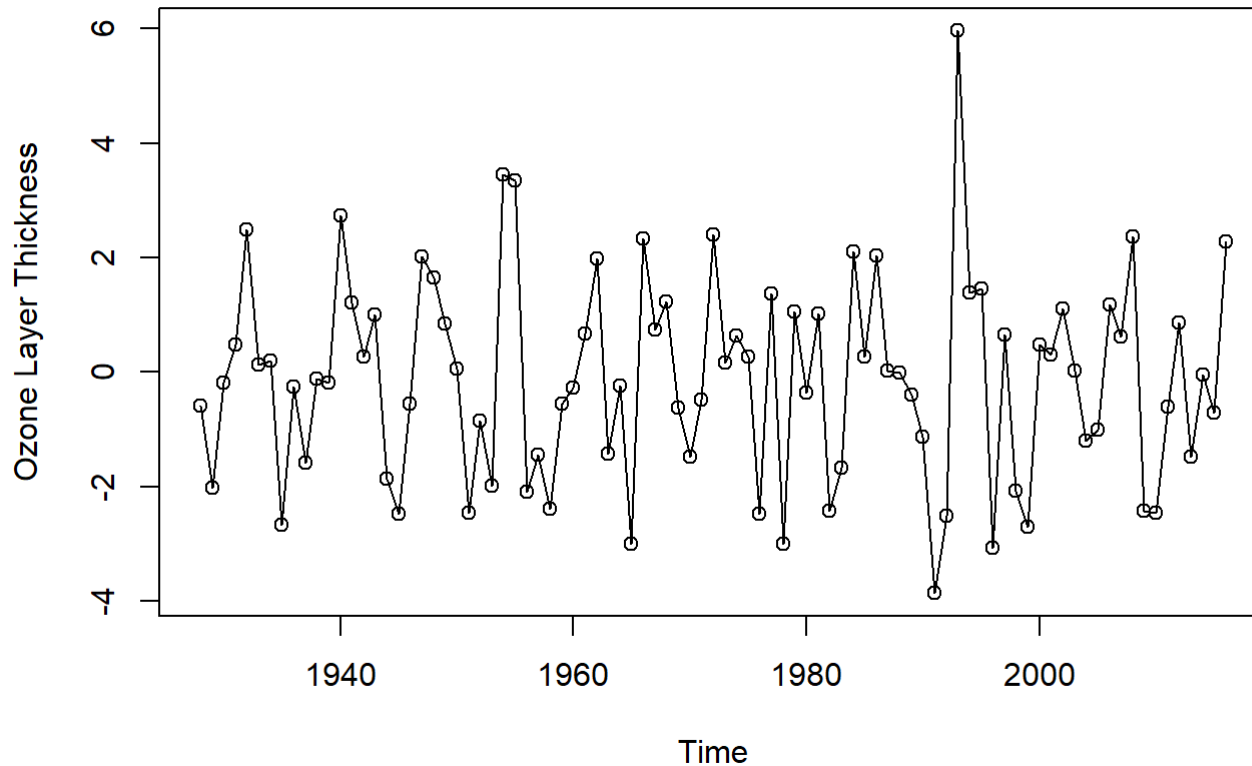
Box Transformation to capture lamba value



```
## [1] 0.9 1.5
```

Differencing without Box Cox Transformation

Times series plot of differenced series



Using ar() to capture lag order values for DK Fuller normality test

```
##  
## Call:  
## ar(x = diff.Ozone_thickness.ts2)  
##  
## Coefficients:  
##      1      2      3      4      5      6  
## -0.1976 -0.2628 -0.6019 -0.3064 -0.2253 -0.3045  
##  
## Order selected 6  sigma^2 estimated as  2.232
```

Augmented DK Fuller test to test stationarity of non transformed differenced series

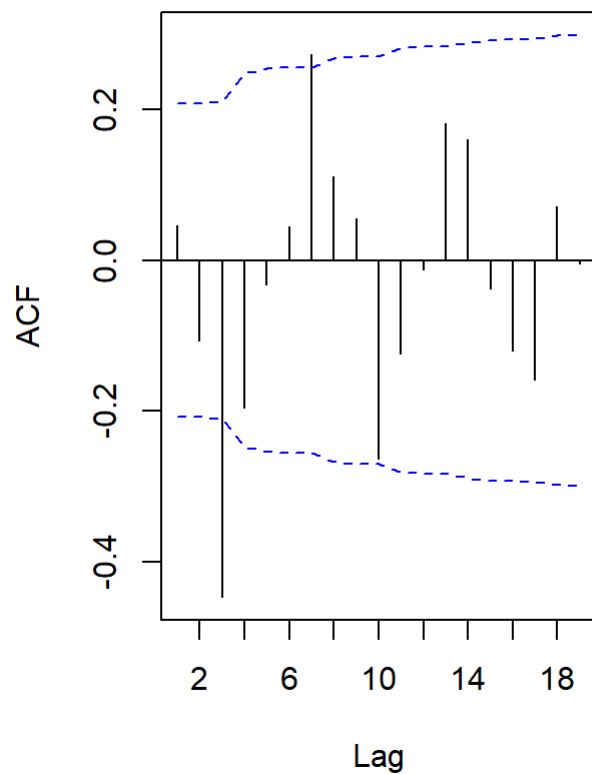
```
##
## Title:
## Augmented Dickey-Fuller Test
##
## Test Results:
## PARAMETER:
## Lag Order: 6
## STATISTIC:
## Dickey-Fuller: -5.0757
## P VALUE:
## 0.01
##
## Description:
## Tue Apr 14 23:07:40 2020 by user: Avinash Matani
```

Normality test of differenced series

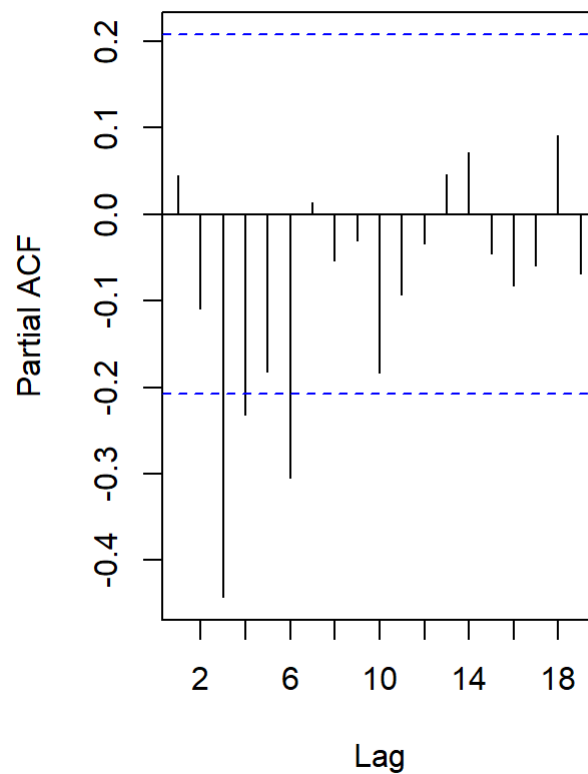
```
##
## Shapiro-Wilk normality test
##
## data: diff.Ozone_thickness.ts2
## W = 0.97907, p-value = 0.1606
```

Specify the orders using ACf, PACF, EACF and BIC over the differenced series

ACF plot of differenced series



PACF plot of differenced series



```
## AR/MA
##   0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 o o x o o o x o o x o o o o
## 1 x o x o o o o o x o o o o
## 2 x o x o o o x o o x o o o
## 3 x o x o o x o o o o o o o
## 4 x o o x o x o o o o o o o
## 5 x x x x o x o o o o o o o
## 6 o o o x x o o o o o o o o
## 7 o o o x o o o o o o o o o
```

