Time Series Analysis- Assignment 2

Tarana Tabassum, s3802509

09/05/2020

EGG DEPOSITIONS OF AGE 3 LAKE HURON BLOATERS, 1981-1996

Abstract

The following study is an analysis of the egg depositions (in millions) of Lake Huron Bloaters (Coregonus hoyi) between years 1981 and 1996 by using various analysis methods and forecasting the depositions for the next 5 years. We initiate the task by understanding the data and convert it to a time series. Later we transform it and difference it. Certain models were identified through EACF, AIC and BIC functions; then tested through the z coefficient test and the residuals were analysed. We finally chose ARIMA(0,1,1) and demonstrate forecast for the next 5 years for the Egg Depositions.

Table of contents

1 Introduction

Coregnous Hoyi is a species of fish which is also knows as “Bloater” and generally found in freshwater. In this report, we study the Egg depositions of age-3 Lake Huron Bloaters has changed from 1981 to 1996. We also aim to predict the changes in Egg depositions in the next 5 year using relevant model fitting and functions on R studio.

1.1 Objective

The objective of this analysis is to analyze the egg depositions of Lake Huron Bloasters by using various time series analysis methods and finally to choose the best model among a set of possible models for this dataset and give forecasts of egg depositions for the next 5 years.

1.2 About the Dataset

Egg depositions (in millions) of age-3 Lake Huron Bloaters (Coregonus hoyi) between years 1981 and 1996 are compiled in our dataset named eggs.

1.3 Loading the Dataset

We read the data into R first and save it as “data”.

```
## Parsed with column specification:
## cols(
##   year = col_double(),
##   eggs = col_double()
## )
```

We observe that it consists of two variables – year and egg Egg Depositions (in millions) over the years.

```
## # A tibble: 16 x 2
##     year   eggs
##    <dbl>  <dbl>
##  1  1981 0.0402
##  2  1982 0.0602
##  3  1983 0.120
##  4  1984 0.181
##  5  1985 0.723
##  6  1986 0.532
##  7  1987 0.432
##  8  1988 0.482
##  9  1989 1.15
## 10  1990 2.10
## 11  1991 1.57
## 12  1992 1.56
## 13  1993 0.763
## 14  1994 0.843
## 15  1995 1.01
## 16  1996 1.02
```

Table 1 : Variables of Data - Egg depositions in millions of age-3 Lake Huron Bloaters between 1981 and 1996

Table 1 here shows the two variable of the dataset that we have on hand.

1.4 Converting the Dataset

Next we convert it to a time series object.

```
## [1] "spec_tbl_df" "tbl_df"      "tbl"         "data.frame"
```

Checking the class, we see that the conversion was successful.

2. Methodology

We will use R studio to analyse this dataset and a set of possible models will be identified after analyzing and transforming of this time series data. By fitting appropriate model to the time series data and examining residuals, a set of eligible models will be used for further model selection using AIC and BIC scored. Finally the best model will be used to predict the changes for the next 5 years.

3. Results and Inferences

3.1 Analysing and transforming the data

3.1.1 Analysing

We plot the data to identify the characteristics of the data.

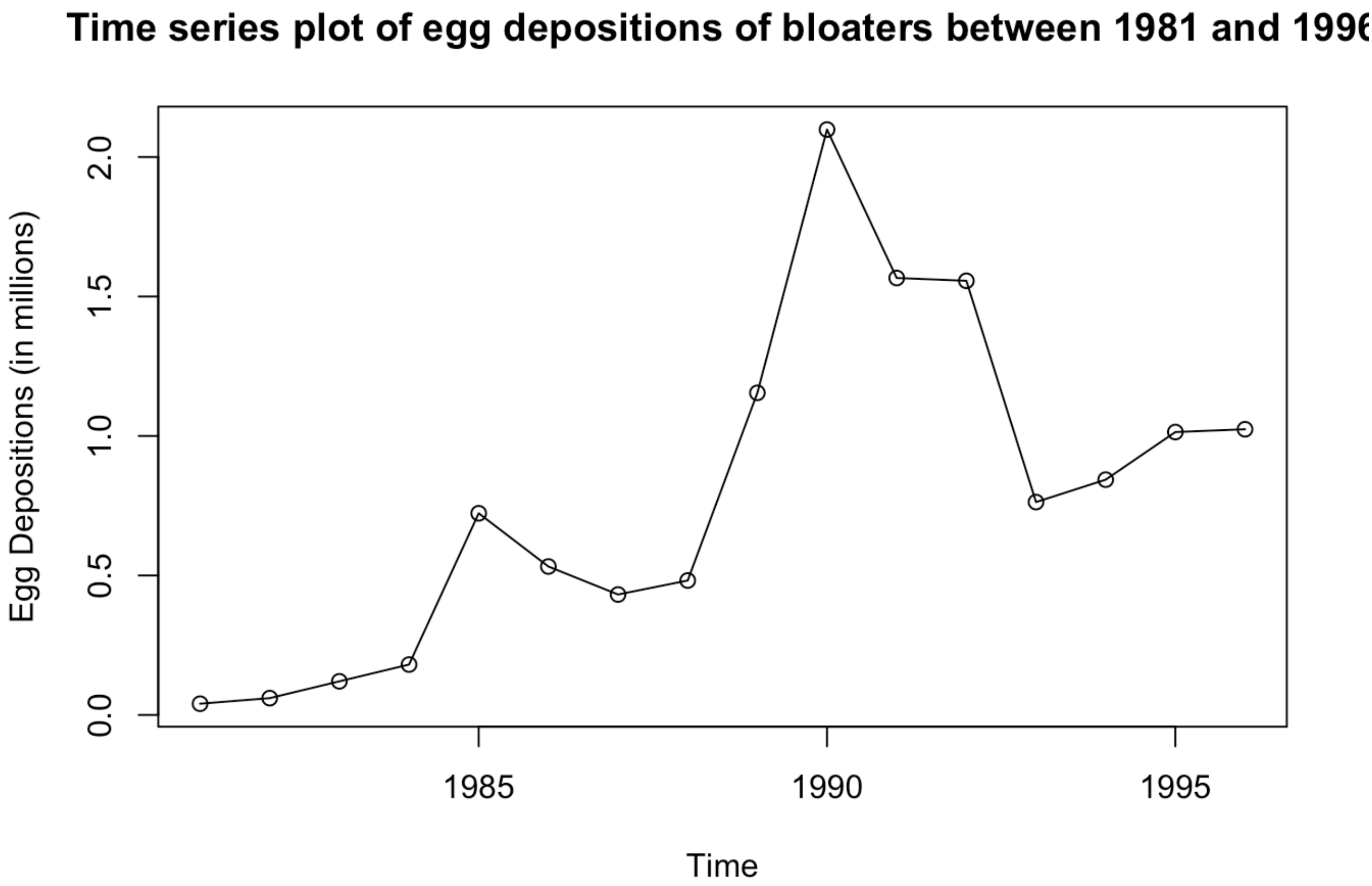


Figure 1 : Time series plot of egg depositions of bloaters between 1981 and 1996

From Figure 1, we find the 5 main characteristics of the data:

1. Trends: Allthough there are some small changing local trends, the overall series seems to have an upward trend.
2. Seasonality: The data does not look like to have a constant seasonality.
3. Intervention/Changing point:The data has upward and downward movements throughout the series. Like in the year 1984 we can observe an upward movement followed by a steep downward movement in the year 1985. Again in the year 1988, we can see an upward movement followed by another downward movement in 1990. But the data does not look to have any changing point as there is no drastic shift in the flow which is followed by constant data.
4. Changing Variance: The data seems to have changing variance.
5. Behaviour: The series looks like non-stationary and has a non constant mean which changes with respect to time. The series has Auto Regressive as well as Moving Average characteristics. So we think an ARIMA model will be a good fit for the series.

Since this is a stochastic trend, we are not going to do any trend analysis for deterministic trend, we'll check straight for ARIMA model.

First we have plotted the ACF and the PACF and then we have done the ADF test on the original to test whether trend is present in the data or not.

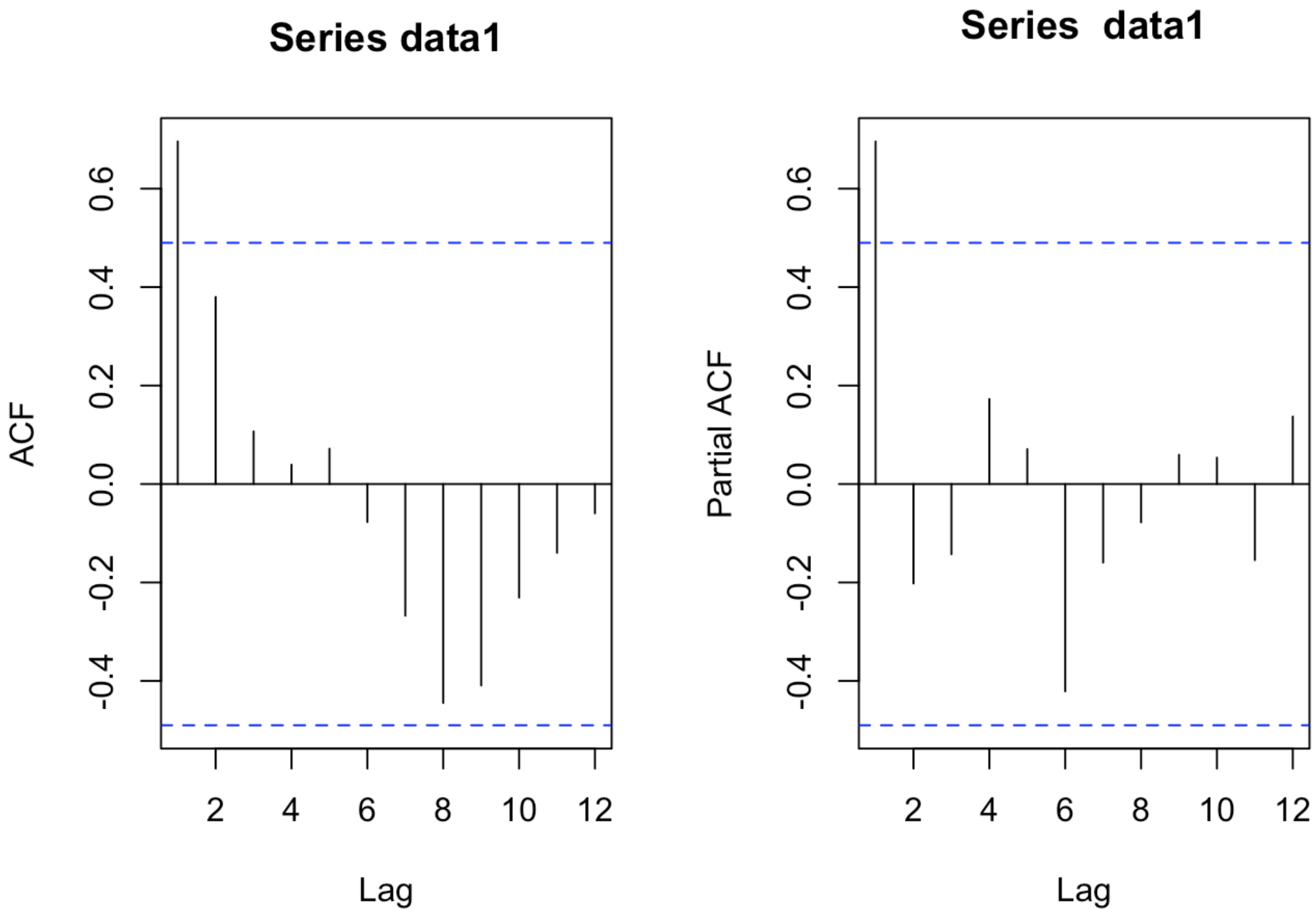


Figure 2 : ACF & Partial ACF

```
##
##  Augmented Dickey-Fuller Test
##
## data:  data1
## Dickey-Fuller = -2.0669, Lag order = 2, p-value = 0.5469
## alternative hypothesis: stationary
```

Table 2: Augmented Dickey-Fuller Test

Figure 2 shows the slowly decaying pattern of the ACF and a very high first correlation in the PACF indicates that there is trend and non stationarity present in the series. Furthermore the p-value generated from the ADF test confirms the presence of the non stationarity in the series. If it had been less than 0.05, we could say that there is no non stationarity present in the series.

3.1.2 Box-Cox Transformation

Now, to further proceed with the ARIMA model, we need to get rid of the changing variance first. So to do so, we have applied the Box-Cox transformation.

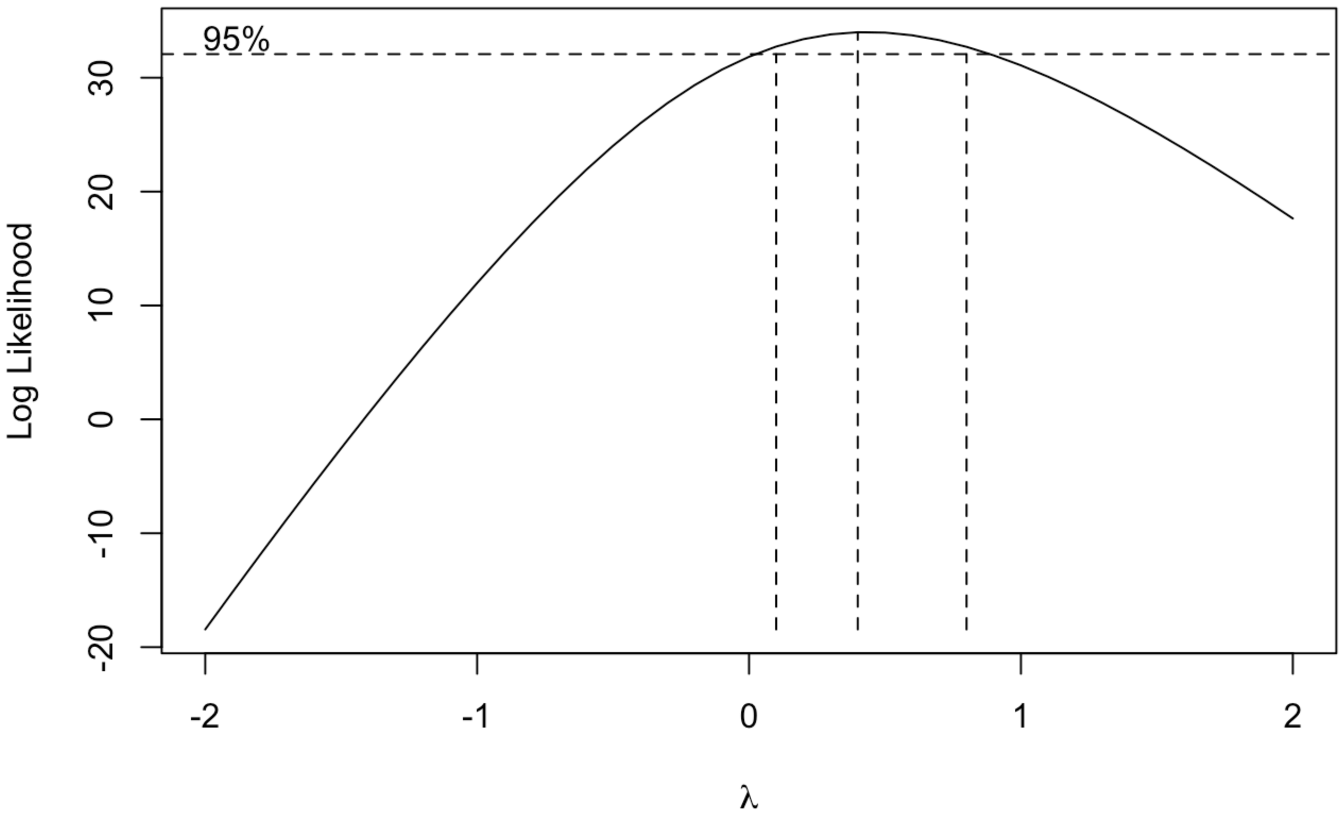


Figure 3 : Box-Cox Transformation

```
## [1] 0.1 0.8
```

The confidence interval comes out to be (01, 0.8). so we choose lambda value as 0.45 and perform the analysis.

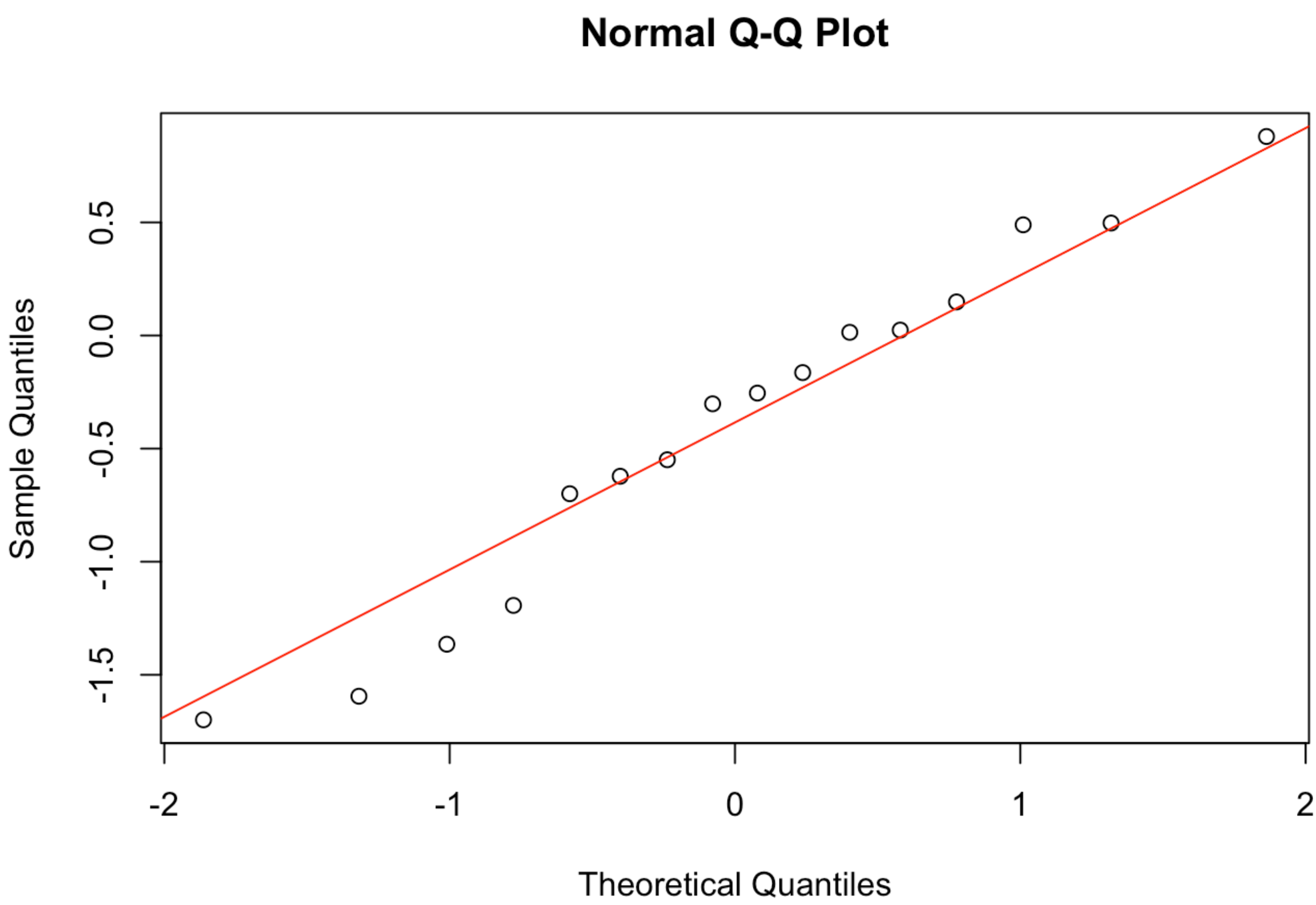


Figure 4: Q-Q Plot

```
##
##  Shapiro-Wilk normality test
##
## data:  data1.bc
## W = 0.96269, p-value = 0.7107
```

Table 3: Shapiro Wilk Test

From the Shapiro Wilk test shown in Table 3, we have failed to reject the null hypothesis which is the series is normal. But we can assume normality based on qq-plot.

3.1.3 Differencing

Running an ADF test to check presence of trend -

```
##
##  Augmented Dickey-Fuller Test
##
## data:  data1.bc
## Dickey-Fuller = -1.6769, Lag order = 2, p-value = 0.6955
## alternative hypothesis: stationary
```

Table 4: Augmented Dickey-Fuller Test

From Table 4 we see that the p-value is greater then 0.5. As a result we cannot reject the null hypothesis that says there is a trend within the series.

Now in order to remove the trend we have done differencing on the data and then have done an ADF test on it again to check whether it has been detrended or not.

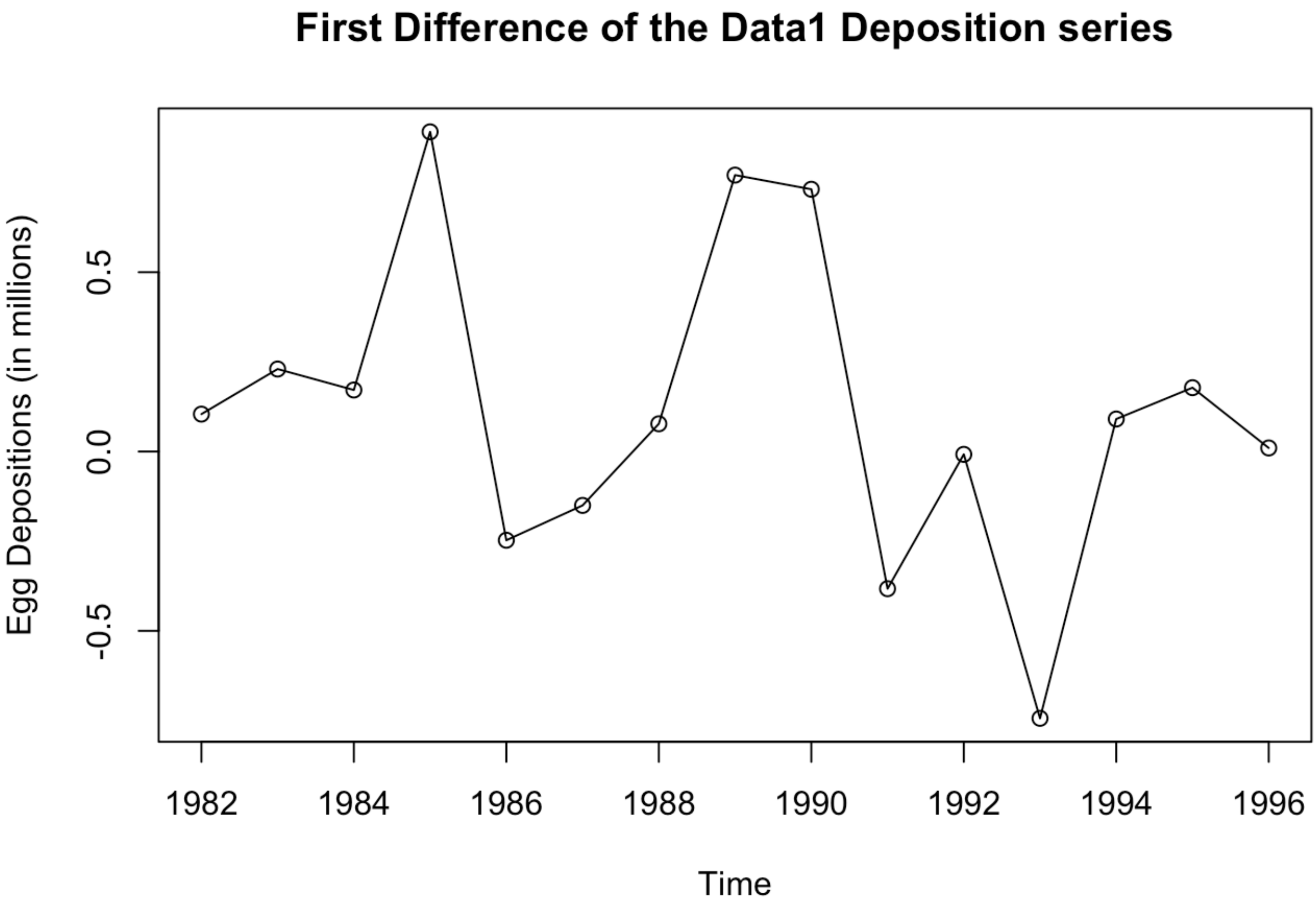


Figure 5: First Difference of the Data1 Deposition series

```
##
##  Augmented Dickey-Fuller Test
##
## data:  data1.bc.diff
## Dickey-Fuller = -3.6798, Lag order = 2, p-value = 0.0443
## alternative hypothesis: stationary
```

Table 5: Augmented Dickey-Fuller Test

As the p-value of the ADF test is below 0.05, we can safely say that there is no trend present in the series anymore.Thus we have been able to successfully reject the null hypothesis which says that the series is in non-stationarity.Now we can say our series is stationary.

3.2 Model fitting

3.2.1 ACF,PACF & EACF Table

To determine our probable ARIMA model, we will now plot the the ACF and PACF.

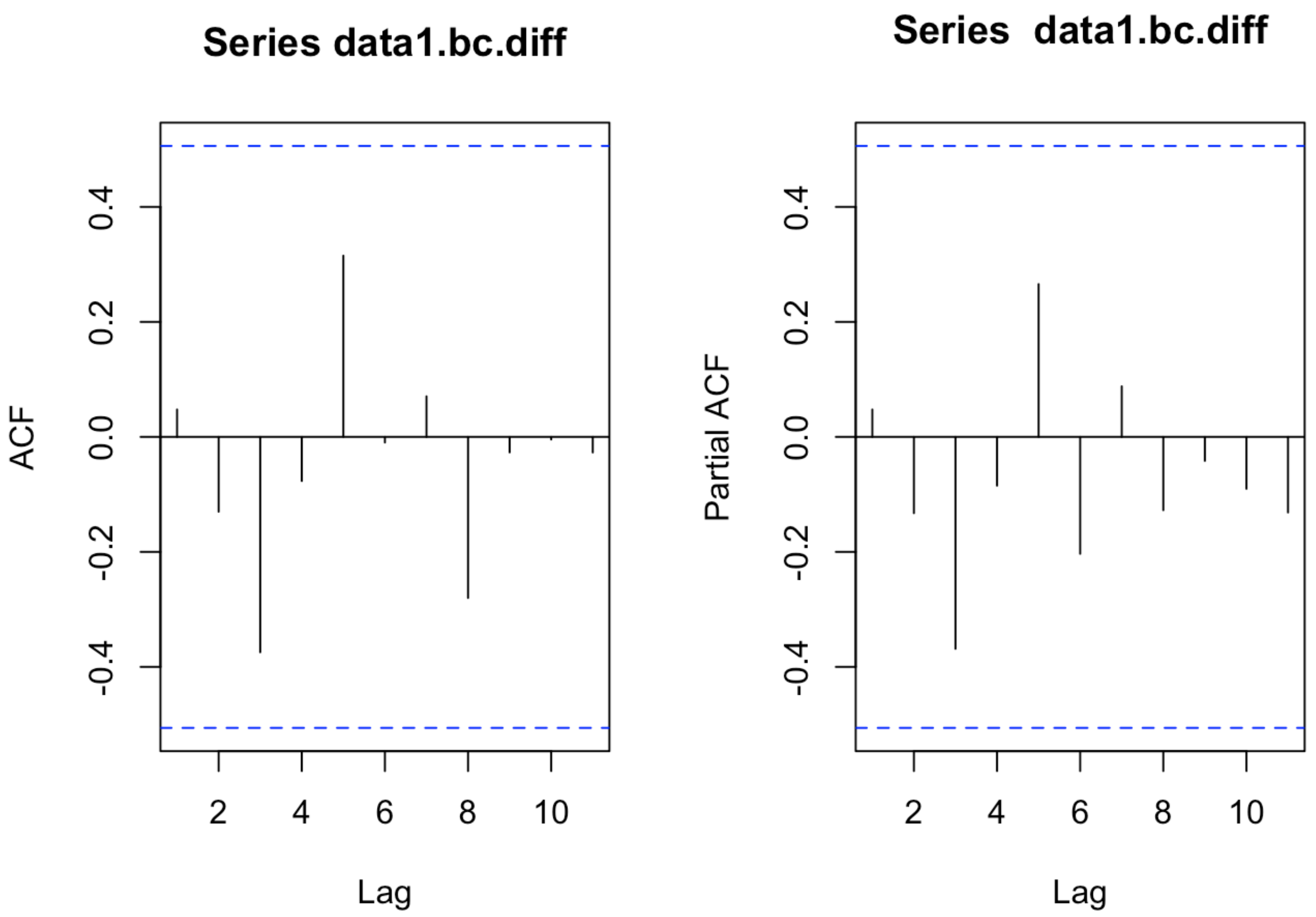


Figure 6: ACF & Partial ACF after differencing

There is no significant lag in both the ACF and the PACF, so we could not find any value for p and q. Also the series looks like a white noise. So now we will plot EACF table, to get our probable models.

```
## AR/MA
##   0 1 2 3
## 0 o o o o
## 1 o o o o
## 2 o o o o
## 3 o o o o
```

Table 6: EACF Table

From the EACF table we have got three probable models which are- ARIMA(0,1,1),ARIMA(1,1,0),ARIMA(1,1,1).The ‘1’ in the middle is the value of ‘d’, as we have done only the first order differencing.

3.2.2 BIC Table

Now we will chart the BIC table to get some more probable models.

```
## Reordering variables and trying again:
```

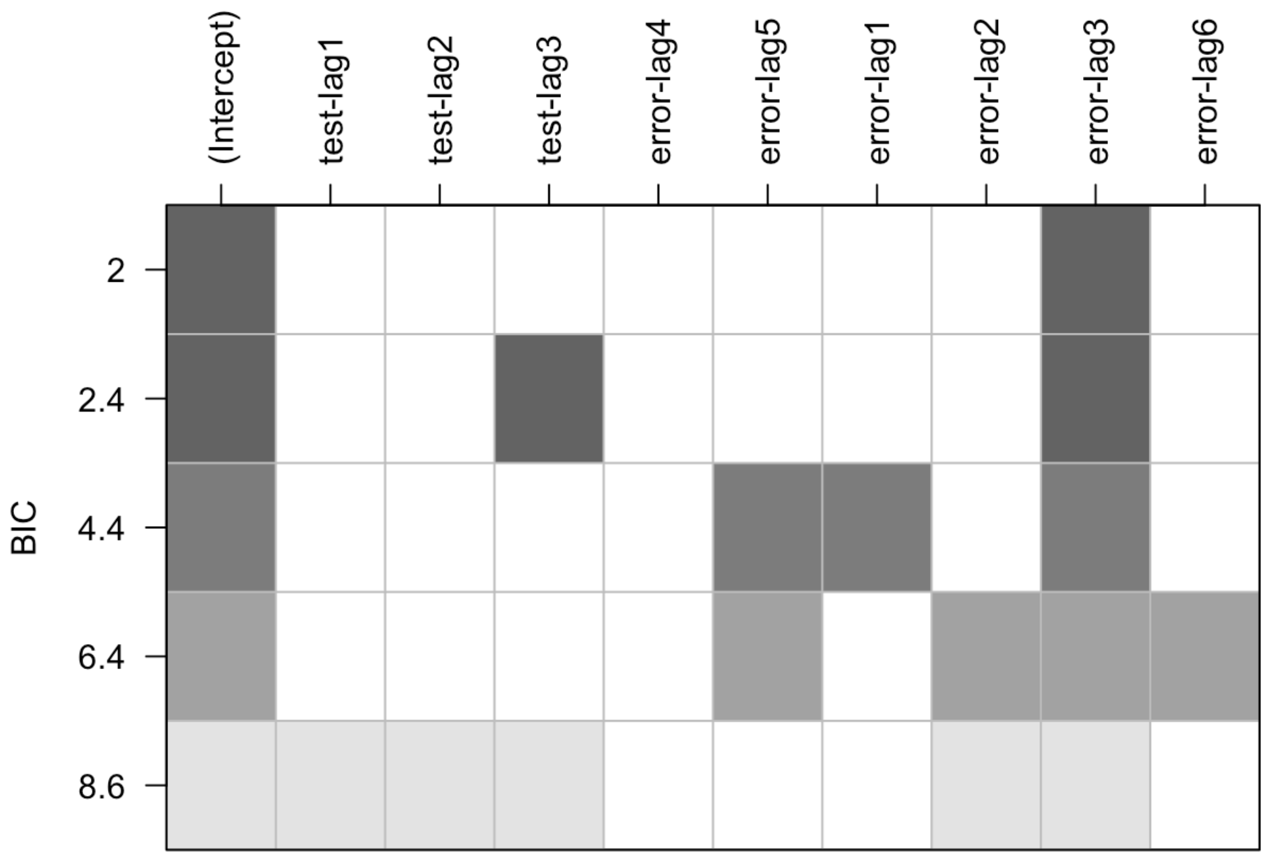


Table 7: BIC Table

From the BIC table we have taken probable models which are- ARIMA(0,1,3), ARIMA(3,1,3), ARMIA(0,1,5)

3.2.3 Parameter Estimation of the Probable Models

So our final set of probable models are ARIMA(0,1,1),ARIMA(1,1,0),ARIMA(1,1,1),ARIMA(0,1,3), ARIMA(3,1,3) and ARIMA(0,1,5). Now we will do the parameter estimation of each model. We have checked the conditional sum of squares (CSS) and the maximum likelihood estimation (ML) of each model here and recognizing whether they are significant or not.If the value of Pr(>|z|) comes under .05, we can say that model is significant and can proceed for the final modelling.

ARIMA(0,1,1)

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value  Pr(>|z|)
## ma1 -1.093338   0.056532  -19.34 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 8: ARIMA(0,1,1) - CSS : z test of coefficients

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ma1 -0.99994    0.66766 -1.4977   0.1342
```

Table 9: ARIMA(0,1,1) - MLE : z test of coefficients

According to table 8 & 9, MA(1) is not significant.

ARIMA(1,1,0)

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.40609    0.24460 -1.6602  0.09687 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 10: ARIMA(1,1,0) - CSS : z test of coefficients

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.38056    0.23493 -1.6199   0.1053
```

Table 11: ARIMA(1,1,0) - MLE : z test of coefficients

According to table 10 & 11, AR(1) is not significant.

ARIMA(0,1,3)

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value  Pr(>|z|)
## ma1 -1.10117    0.29725 -3.7045 0.0002118 ***
## ma2 -0.11909    0.33351 -0.3571 0.7210194
## ma3  0.14673    0.34631  0.4237 0.6717891
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 12: ARIMA(0,1,3) - CSS : z test of coefficients

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ma1 -0.852252   0.536615 -1.5882   0.1122
## ma2 -0.102618   0.286176 -0.3586   0.7199
## ma3 -0.045118   0.512224 -0.0881   0.9298
```

Table 13: ARIMA(0,1,3) - MLE : z test of coefficients

From tables 12 & 13 we can not conclude whether ARIMA(0,1,3) is significant or not since CSS is telling us that MA(1) is significant but MLE is showing us it is not significant.

ARIMA(3,1,3)

```
##
## z test of coefficients:
##
##       Estimate Std. Error   z value Pr(>|z|)
## ar1 -1.2016973  0.0067545 -177.9105  < 2e-16 ***
## ar2 -0.8077002  0.0231155  -34.9419  < 2e-16 ***
## ar3 -0.1124583  0.0998924   -1.1258  0.26025
## ma1  1.9195706         NA        NA       NA
## ma2  1.9992451         NA        NA       NA
## ma3 -1.3935177  0.6294154   -2.2140  0.02683 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 14: ARIMA(3,1,3) - CSS : z test of coefficients

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1  0.35329    0.41854  0.8441 0.398620
## ar2 -0.68002    0.26353 -2.5804 0.009868 **
## ar3 -0.41454    0.32147 -1.2895 0.197226
## ma1 -1.31528    0.56912 -2.3111 0.020828 *
## ma2  1.42450    0.89398  1.5934 0.111063
## ma3 -0.59677    0.51624 -1.1560 0.247685
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 15: ARIMA(3,1,3) - MLE : z test of coefficients

Again, from table 14 & 15, we can not conclude whether ARIMA(3,1,3) is significant or not since CSS is showing some other values and MLE is showing some other.

ARIMA(0,1,5)

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ma1 -1.016594   0.375033 -2.7107 0.006715 **
## ma2  0.013365   0.364525  0.0367 0.970752
## ma3 -0.753976   0.239229 -3.1517 0.001623 **
## ma4  1.192805   0.411640  2.8977 0.003759 **
## ma5 -0.538537   0.506024 -1.0643 0.287215
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 16: ARIMA(0,1,5) - CSS : z test of coefficients

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ma1 -0.824906   0.739713 -1.1152   0.2648
## ma2  0.038341   0.314834  0.1218   0.9031
## ma3 -0.732093   0.840191 -0.8713   0.3836
## ma4  1.065250   0.965774  1.1030   0.2700
## ma5 -0.089888   0.354698 -0.2534   0.7999
```

Table 17: ARIMA(0,1,5) - MLE : z test of coefficients

Again,from table 16 & 17, we can not conclude whether ARIMA(0,1,5) is significant or not since CSS is showing some other values and MLE is showing some other.

ARIMA(1,1,1)

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value  Pr(>|z|)
## ar1  0.030947   0.290460  0.1065    0.9152
## ma1 -1.021947   0.201926 -5.0610 4.171e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 18: ARIMA(1,1,1) - CSS : z test of coefficients

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1  0.11474    0.26992  0.4251 0.670764
## ma1 -1.00000    0.33205 -3.0116 0.002599 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 19: ARIMA(1,1,1) - MLE : z test of coefficients

Again,according to table 18 & 19, we can not conclude whether ARIMA(1,1,1) is significant or not since CSS and MLE are showing different results.

3.2.4 Sorting AIC and BIC values

To conclude on our best ARIMA model, we will now sort the AIC and BIC value of all the probable models. The model which will have the least value among all the models will be our best fit.

```
##              df      AIC
## model_011_ml  2 23.12446
## model_013_ml  4 26.92313
## model_110_ml  2 27.05798
## model_015_ml  6 27.39690
## model_313_ml  7 29.34748
```

Table 20: AIC Table

```
##              df      BIC
## model_011_ml  2 24.40258
## model_110_ml  2 28.33609
## model_013_ml  4 29.47936
## model_015_ml  6 31.23124
## model_313_ml  7 33.82088
```

Table 21: BIC Table

According to table 20 & 21, we can conclude that ARIMA(0,1,1) is the best fit to predict on this particular data.

3.2.5 Model Overfitting

Now we will test the overfitting of our best model which is ARIMA(0,1,1). To do so we will check whether ARIMA(1,1,1) and ARIMA(0,1,2) is significant or not. We have already checked ARIMA(1,1,1) is insignificant. So now we will only test whether ARIMA(0,1,2) is significant or not. If ARIMA(0,1,2) comes insignificant, we can proceed with ARIMA(0,1,1) model to do the forecast.

```
##
## z test of coefficients:
##
##       Estimate Std. Error z value  Pr(>|z|)
## ma1 -1.017026   0.274754 -3.7016 0.0002143 ***
## ma2 -0.087178   0.308221 -0.2828 0.7772974
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 22: ARIMA(0,1,2) - CSS : z test of coefficients

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ma1 -0.88465    0.42637 -2.0748   0.0380 *
## ma2 -0.11535    0.25484 -0.4527   0.6508
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 23: ARIMA(0,1,2) - MLE : z test of coefficients

Both tables 22 & 23 are indicating that ARIMA(0,1,2) is insignificant rather overfitted. Hence we can say that only ARIMA(0,1,1) is our best fit to do the prediction.

3.3 Residual Analysis

We do the residual analysis for the shortlisted model (0,1,1) that we haave found out to be normal and can be a best fit to the data. We do this after having a look at the p-value in shapiro wilk test and those with greater then 0.5 can be considered as normal.

```
##
##  Shapiro-Wilk normality test
##
## data:  res.model
## W = 0.95773, p-value = 0.653
```

```
## Warning in (ra^2)/(n - (1:lag.max)): longer object length is not a multiple
## of shorter object length
```

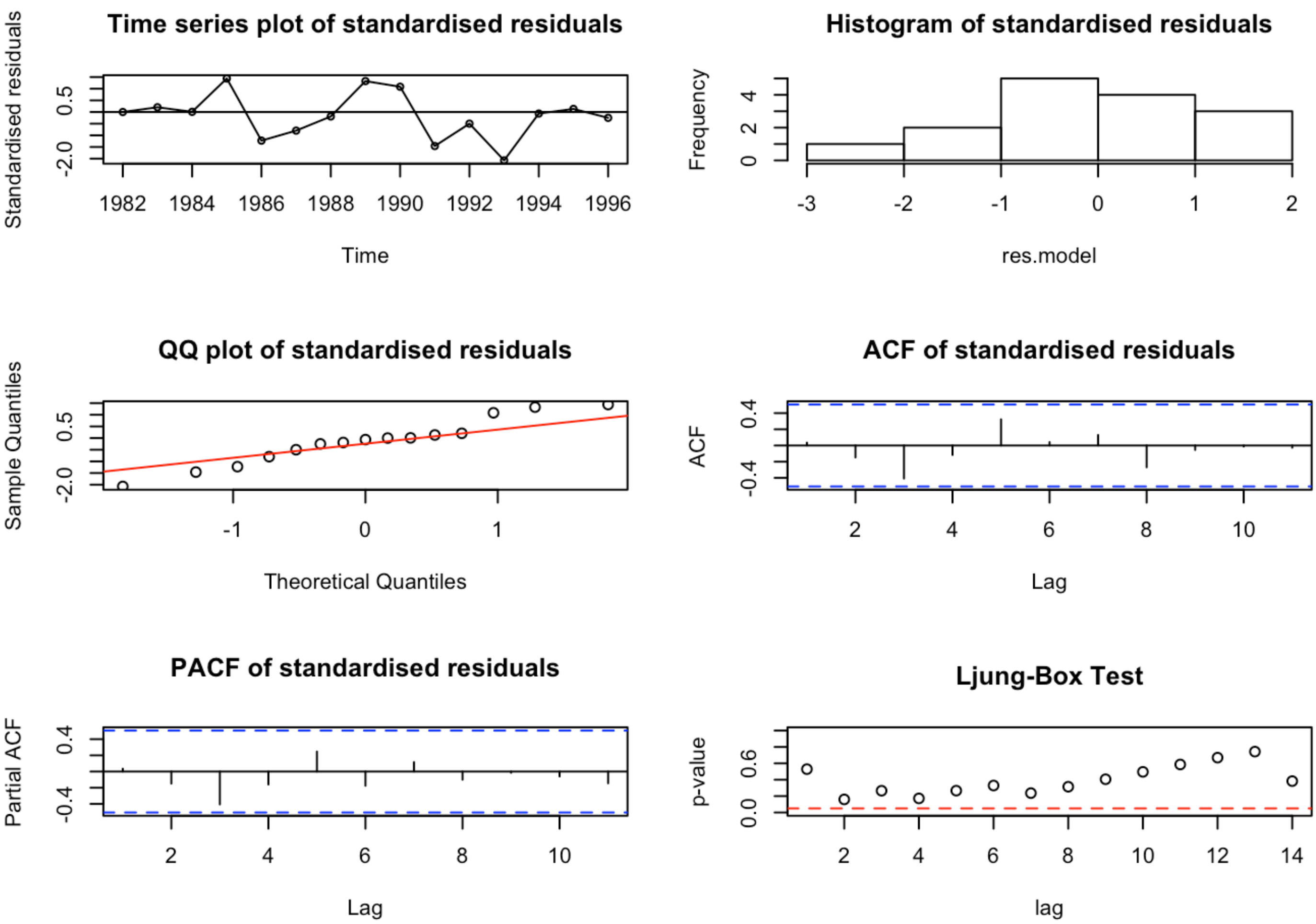


Figure 7: Residual Analysis

The residual analysis indicates

-standardised normals are randomly distributed over time. -histogram looks normally distributed and form a bell shaped curve. -QQ plot looks not very alligned. -ACF plot does not have any significant lags -Shapiro-Wilk normality test indicates that residuals are random with the p-value of 0.3994 -All p-values in Ljung-Box test plot are above the alpha value.

Although QQplot does not look very alligned and the p-value of the Shapiro-Wilk test is not very high, but the other analysises are satisfying enough to go ahead with the model ARIMA(0,1,1) for the final forecasting.

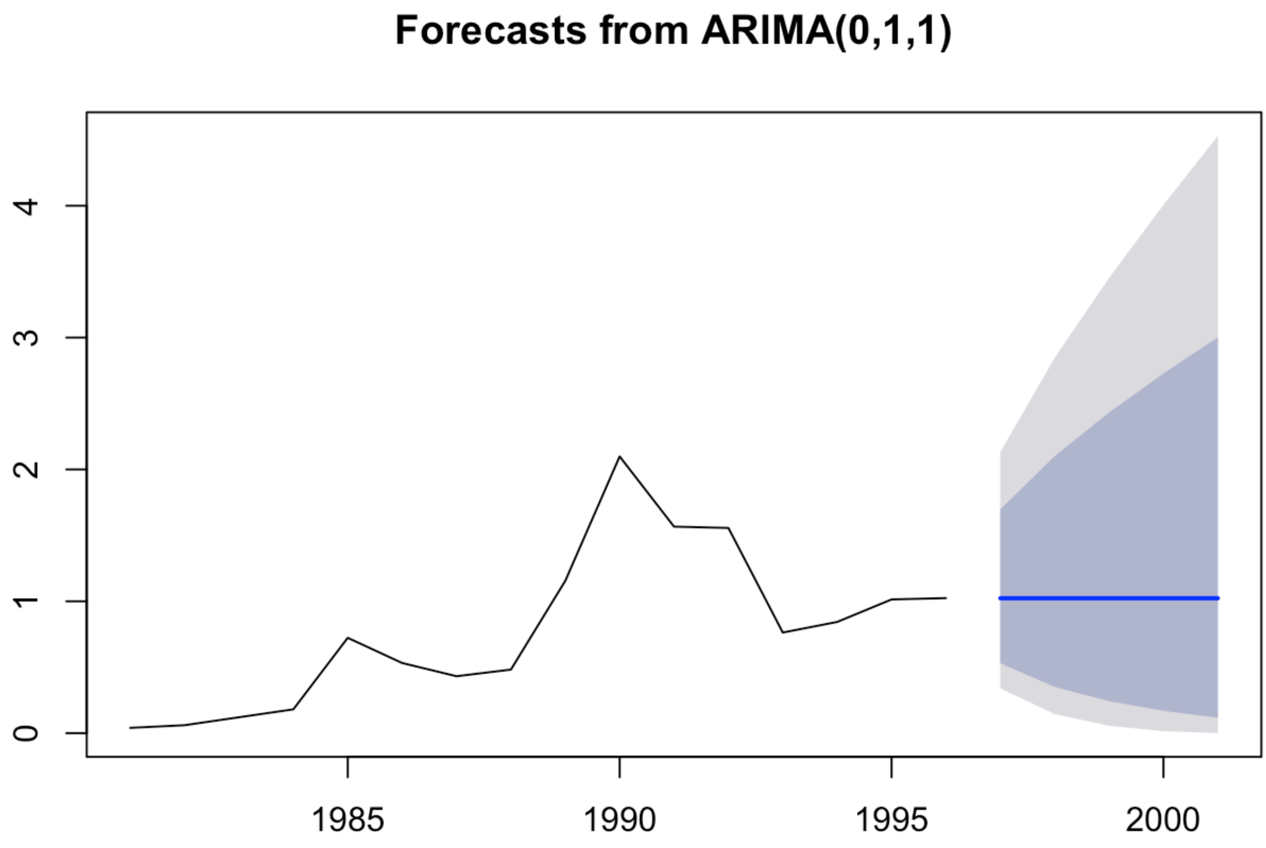
3.4 Predicting and forecasting values

We will now do the final forecastong based on the ARIMA(0,1,1) model for the upcoming 5 years.

```
## $pred
## Time Series:
## Start = 1997
## End = 2001
## Frequency = 1
## [1] 0.1148626 0.1148626 0.1148626 0.1148626 0.1148626
##
## $se
## Time Series:
## Start = 1997
## End = 2001
## Frequency = 1
## [1] 0.4491633 0.4491633 0.4491633 0.4491633 0.4491633
```

Table 24: Prediction of Egg deposition for next 5 years

3.5 Forecasting plot



```
## Series: data1
## ARIMA(0,1,1)
## Box Cox transformation: lambda= 0.45
##
## Coefficients:
##          ma1
##       0.1117
## s.e.  0.2419
##
## sigma^2 estimated as 0.2004:  log likelihood=-8.72
## AIC=21.44   AICc=22.44   BIC=22.85
```

Figure 8: Forecast of Egg Depositions of Coregonus Hoyi for next 5 years

4. Conclusion:

Looking at the forecasts it seems that the value of egg depositions is going to follow a decreasing trend in coming 5 years. To summarize,in this report we used a data representing egg depositions (in millions) of age-3 Lake Huron Bloaters (Coregonus hoyi) between years 1981 and 1996. The goal was to find the best fitting model to the dataset and give predictions of yearly changes for the next 5 years. Linear, quadratic and ARIMA models were applied and their statistical outputs and the plots were obtained. Based on them the suitable model was picked and taken forward for diagnostic testing. As seen, trend models were voted out in the initial phase only because they failed the diagnose testing. So we further based our analysis on the ARIMA models. After using various model specification methods and applying various diagnostics tests we chose ARIMA(0,1,1) as the most suitable model. Our diagnostic phase for chosen ARIMA model involved following tests- - Residual analysis - Histogram of residuals - ACF and PACF plots - Shapiro Wilk test - Ljung-Box test ARIMA(0,1,1) successfuly clears all the diagnostic tests (Figure 7) and comes up as the best fit model for the given dataset of egg deposition for age 3 Lake Huron Bloaters. Finally the forecast is shown for the series for next 5 years (Figure 8).

5. Recommendations

As pollution has played a major role in the depletion and extinction of thousands of flora and fauna, steps must be taken in order to balance the ecosystem for a long-term safe space for every species on the planet.

6. Appendices

```r
library(readr)
library(TSA)
library(forecast)
library(FitAR)
library(lmtest)
library(tseries)
library(fUnitRoots)
data <- read_csv("eggs.csv")
```

```
data
data1 <- ts(as.vector(data$eggs), start=1981, end=1996, frequency = 1)
class(data)
plot(data1, ylab = "Egg Depositions (in millions)", main = "Time series plot of egg
depositions of bloaters between 1981 and 1996", type = "o",)

par(mfrow = c(1,2))
acf(data1)
pacf(data1)
adf.test(data1)
data1.tr = BoxCox.ar(data1, method = "yule-walker")
data1.tr$ci
lambda = 0.45 # The mid point of the interval
data1.bc = ((data1^lambda)-1)/lambda
par(mfrow=c(1,1))
qqnorm(data1.bc)
qqline(data1.bc, col = 2)
shapiro.test(data1.bc)

adf.test(data1.bc)
data1.bc.diff = diff(data1.bc, differences = 1)
plot(data1.bc.diff, type = "o", ylab = "Egg Depositions (in millions)", main = "Fir
st Difference of the Data1 Deposition series")
adf.test(data1.bc.diff)
par(mfrow = c(1,2))
acf(data1.bc.diff)
pacf(data1.bc.diff)
eacf(data1.bc.diff, ar.max = 3, ma.max = 3)

res = armasubsets(y = data1.bc.diff, nar = 3, nma = 6, y.name = 'test', ar.method =
'yw')
plot(res)
# ARIMA(0,1,1) - CSS
model_011_css = arima(data1.bc.diff, order = c(0,1,1), method = 'CSS')
coeftest(model_011_css)
# ARIMA(0,1,1) - ML
model_011_ml = arima(data1.bc.diff, order = c(0,1,1), method = 'ML')
coeftest(model_011_ml)
# ARIMA(1,1,0) - CSS
model_110_css = arima(data1.bc.diff, order = c(1,1,0), method = 'CSS')
coeftest(model_110_css)
# ARIMA(1,1,0) - ML
model_110_ml = arima(data1.bc.diff, order = c(1,1,0), method = 'ML')
coeftest(model_110_ml)
# ARIMA(0,1,3) - CSS
model_013_css = arima(data1.bc.diff, order = c(0,1,3), method = 'CSS')
coeftest(model_013_css)
# ARIMA(0,1,3) - ML
model_013_ml = arima(data1.bc.diff, order = c(0,1,3), method = 'ML')
coeftest(model_013_ml)
```

```
# ARIMA(3,1,3) - CSS
model_313_css = arima(data1.bc.diff, order = c(3,1,3), method = 'CSS')
coeftest(model_313_css)
# ARIMA(3,1,3) - ML
model_313_ml = arima(data1.bc.diff, order = c(3,1,3), method = 'ML')
coeftest(model_313_ml)
# ARIMA(0,1,5) - CSS
model_015_css = arima(data1.bc.diff, order = c(0,1,5), method = 'CSS')
coeftest(model_015_css)
# ARIMA(0,1,5) - ML
model_015_ml = arima(data1.bc.diff, order = c(0,1,5), method = 'ML')
coeftest(model_015_ml)
# ARIMA (1,1,1) - CSS
model_111_css = arima(data1.bc.diff, order = c(1,1,1), method = 'CSS')
coeftest(model_111_css)

# ARIMA (1,1,1) - ML
model_111_ml = arima(data1.bc.diff, order = c(1,1,1), method = 'ML')
coeftest(model_111_ml)
sort.score <- function(x, score = c("bic", "aic")){
  if (score == "aic"){
    x[with(x, order(AIC)),]
  } else if (score == "bic") {
    x[with(x, order(BIC)),]
  } else {
    warning('score = "x" only accepts valid arguments ("aic","bic")')
  }
}


sort.score(AIC(model_110_ml, model_011_ml, model_013_ml, model_313_ml, model_015_ml
), score = "aic")
sort.score(BIC(model_110_ml, model_011_ml, model_013_ml, model_313_ml, model_015_ml
), score = "bic")
model_012_css = arima(data1.bc.diff,order=c(0,1,2),method='CSS')
coeftest(model_012_css)
model_012_ml = arima(data1.bc.diff,order=c(0,1,2),method='ML')
coeftest(model_012_ml)
residual.analysis <- function(model, std = TRUE){
  library(TSA)
  library(FitAR)
  if (std == TRUE){
    res.model = rstandard(model)
  }else{
    res.model = residuals(model)
  }
  par(mfrow=c(3,2))
  plot(res.model,type='o',ylab='Standardised residuals', main="Time series plot of
standardised residuals")
  abline(h=0)
  hist(res.model,main="Histogram of standardised residuals")
```

```
  qqnorm(res.model,main="QQ plot of standardised residuals")
  qqline(res.model, col = 2)
  acf(res.model,main="ACF of standardised residuals")
  pacf(res.model,main="PACF of standardised residuals")
  print(shapiro.test(res.model))
  k=0
  LBQPlot(res.model, lag.max = length(model$residuals)-1 , StartLag = k + 1, k = 0,
SquaredQ = FALSE)
  par(mfrow=c(1,1))
}
residual.analysis(model = model_011_ml)

predict(model_011_ml,n.ahead = 5,newxreg = NULL,se.fit=TRUE)
fit=Arima(data1,c(0,1,1),lambda = 0.45)
plot(forecast(fit,h=5))
fit
```

7. References

- Demirhan H. (2020). Math 1318 Time Series Analysis. [Notes] RMIT University.