Final project

Seongu Lee

3/3/2022

Abstract

This project is about "How much will be our bills for next 1 year" I used data set about average bill. I used boxcox transformation and differencing to make data stationary. I used acf, pacf analysis and auto.arima to determine best sarima model. I used diagnostic tests to verify if the model is good to be used for forecast. Finally, I used prediction function to predict the future bills. Based on the model selecting process, I indentified two models SARIMA $(2, 0, 1)$ x $(1, 1, 0)_{12}$ and SARIMA $(1, 0, 1)$ x $(2, 1, 1)_{12}$. I examined plots of residuals, diagnostic tests,Aic and causality and invertibility. From those process, I identified model 1 was good to use for forecast. Finally, I got best model which was first one and I predicted the next 12 values(12 months)

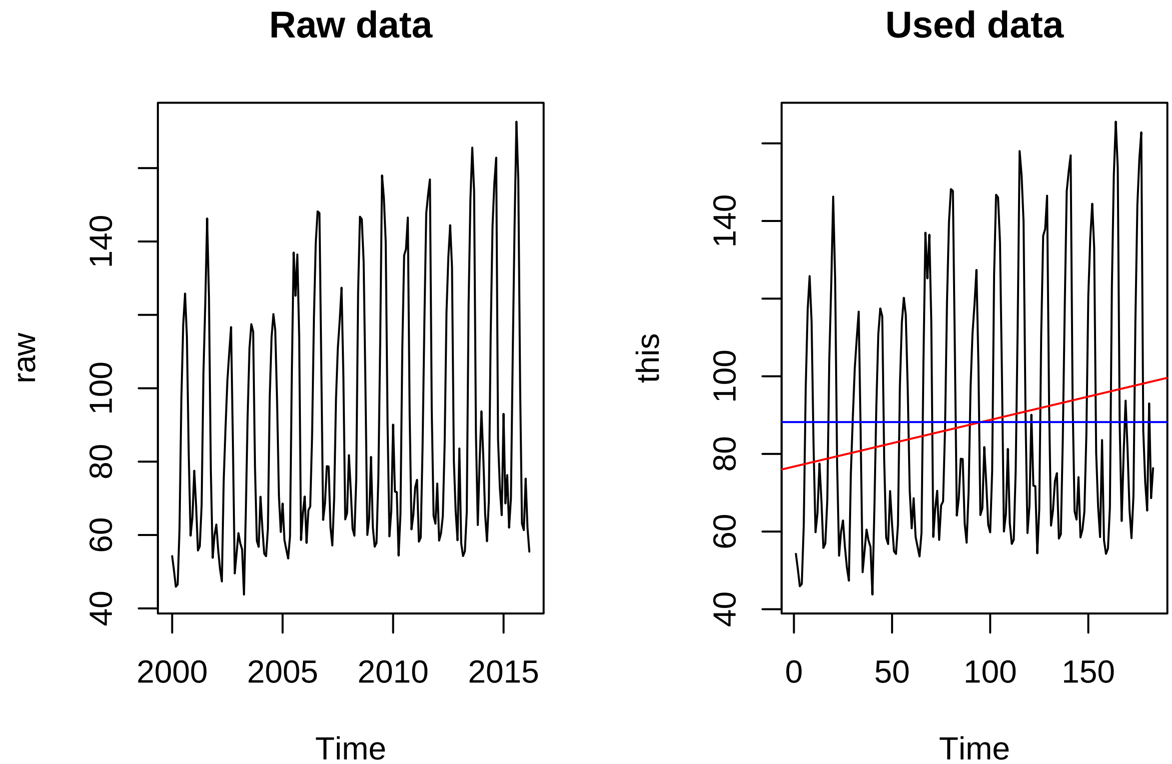
Introduction

I was curious about electric bills by time. If the weather is hot or cold, I need to use more electricity and It will cost me more. This dataset is about the electric bills and usage of electricity in Kwh from 2000 to 2015 in Austin Tx. Because people will use more electricity in summer or winter, I will expect that it will be high bills in summer and winter. Or, it will be different. My goal is to forecast the next one year electric bills using the previous data to prepare to pay the next bills.

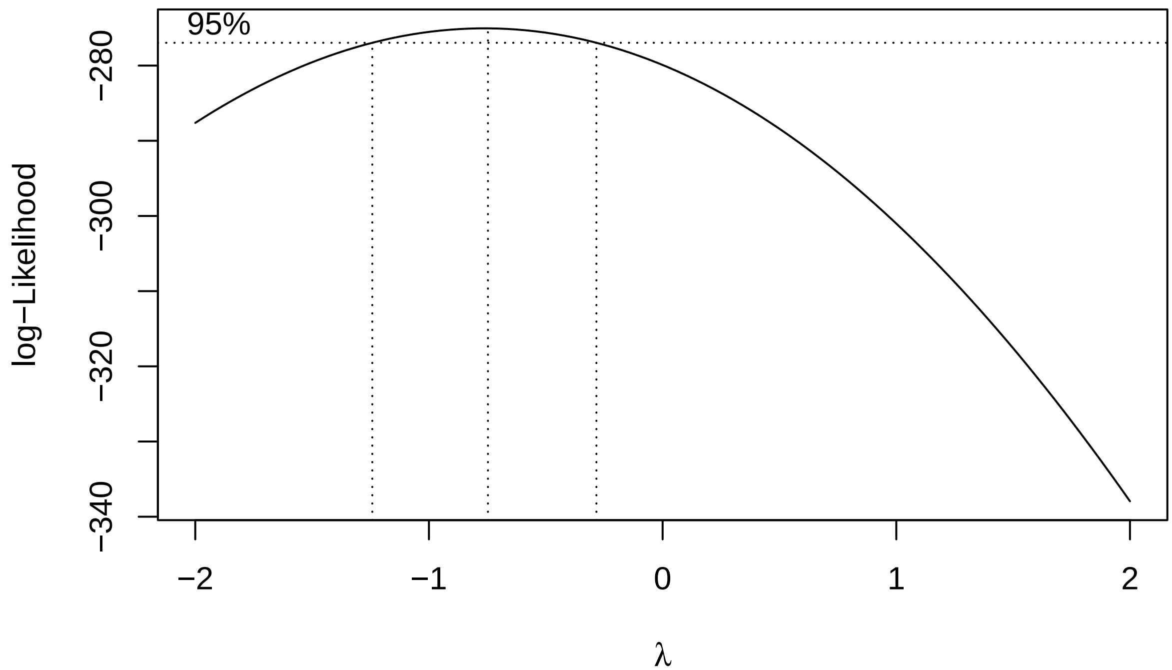
I plotted data without 12 values to check the forecasted values. I used the plot to determine if the dataset has trend, seasonality or variance changing. Next, I used box-cox transform to make data to stabilized the variance and differenced the data to remove the seasonality and trend. Also, I plotted pacf/acf and used auto arima function to determine best two models SARIMA $(2, 0, 1)$ x $(1, 1, 0)_{12}$ and SARIMA $(1, 0, 1)$ x $(2, 1, 1)_{12}$. And I used diagnostic tests("Box-Pierce, Ljung, shapiro), Aic and resdiuals plot to verify the model 1 was more good to use for forecasting. And model 1 was able to be used to forecast next 12 values successfully with confidence intervals. My goal is to predict the next 12 values and it will match to the real values.

The data can be obtained on data.world. And I used rstudio for all processes

Analysis



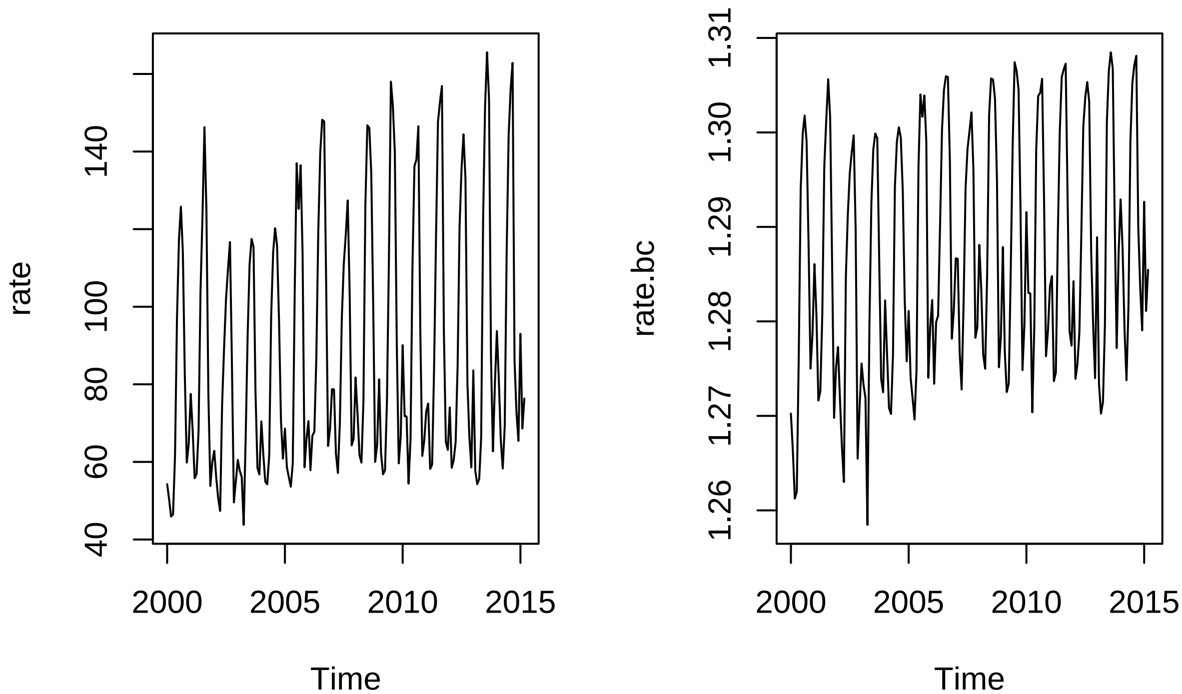
This dataset has increasing, decreasing trend, volatility of the variance and strong seasonality. Since this data set doesn't look like normally distributed, I need use Box-Cox transformation to solve volatility of the variance. I don't see any heteroskedasticity, so I don't need to use any transforms(log, square).




```
## lambda =  -0.7474747
```

I applied boxcos to find lambda. From the lambda plot, the 95% of the confidence interval doesn't include 0, so I can use boxcox transformation to make variance stable lambda doesn't include zero. So, it is good to use boxcox transform. I also used the formula (1/lambda)*(x^lambda-1) for the boxcox transformation

```r
#compare plots
par(mfrow=c(1, 2))
plot(rate)
plot(rate.bc)
```



plots look similar

```
#Compare var
var(rate)
```

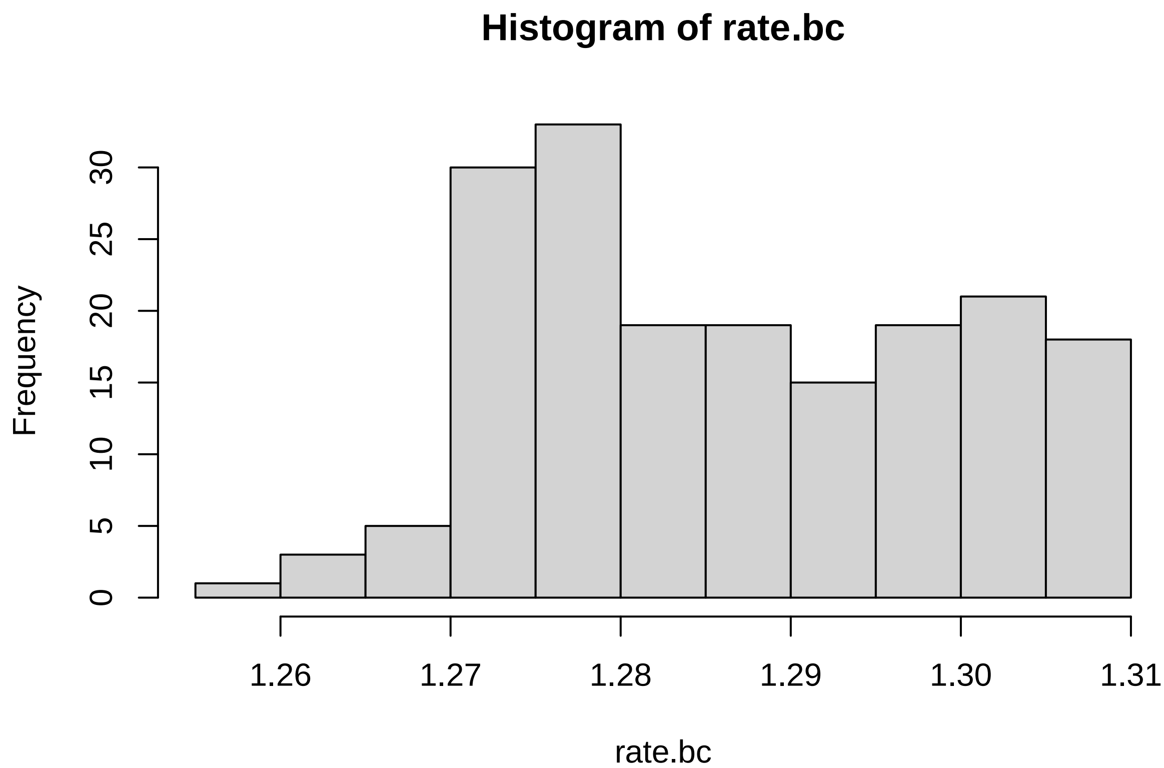
```
## [1] 1043.242
```

```
var(rate.bc)
```

```
## [1] 0.0001587282
```

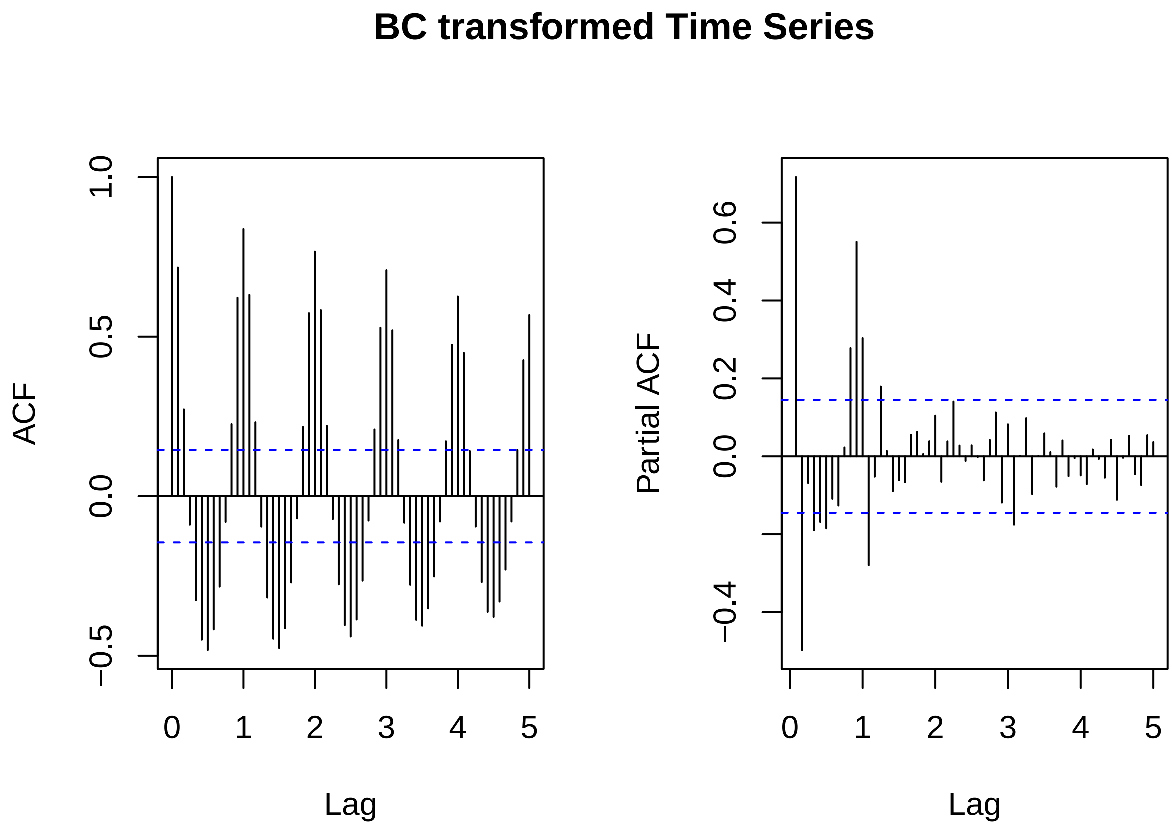
variance was reduced a lot with Box Cos transform so I was good to use the boxcox transform.

```
hist(rate.bc)
```

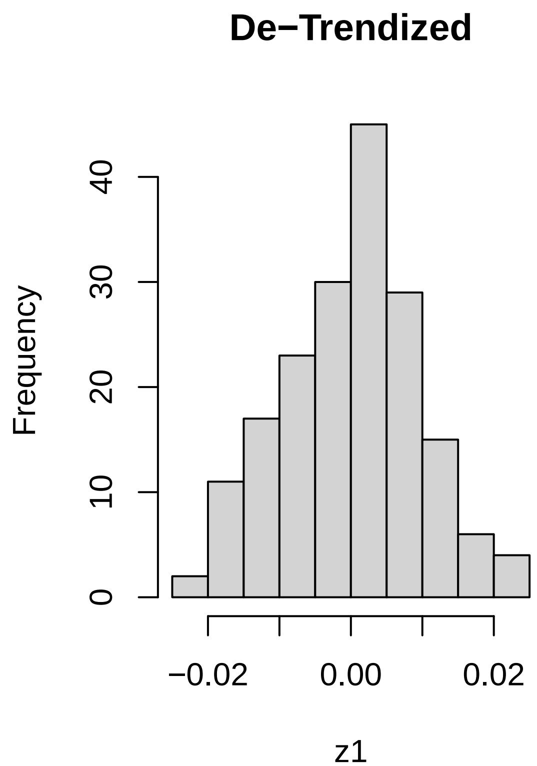


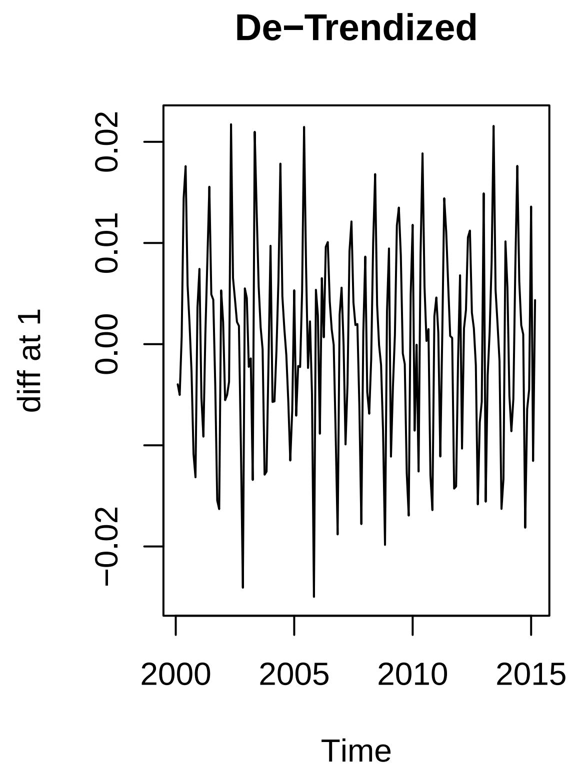
Hist still doesn't look normally distributed. I need to make this more look normally distributed.

```
# acf and pacf for transformed dataset.
par(mfrow=c(1, 2))
acf(rate.bc,lag.max = 60,main = "")
pacf(rate.bc,lag.max = 60,main = "")
title("BC transformed Time Series", line = -1, outer=TRUE) #
```

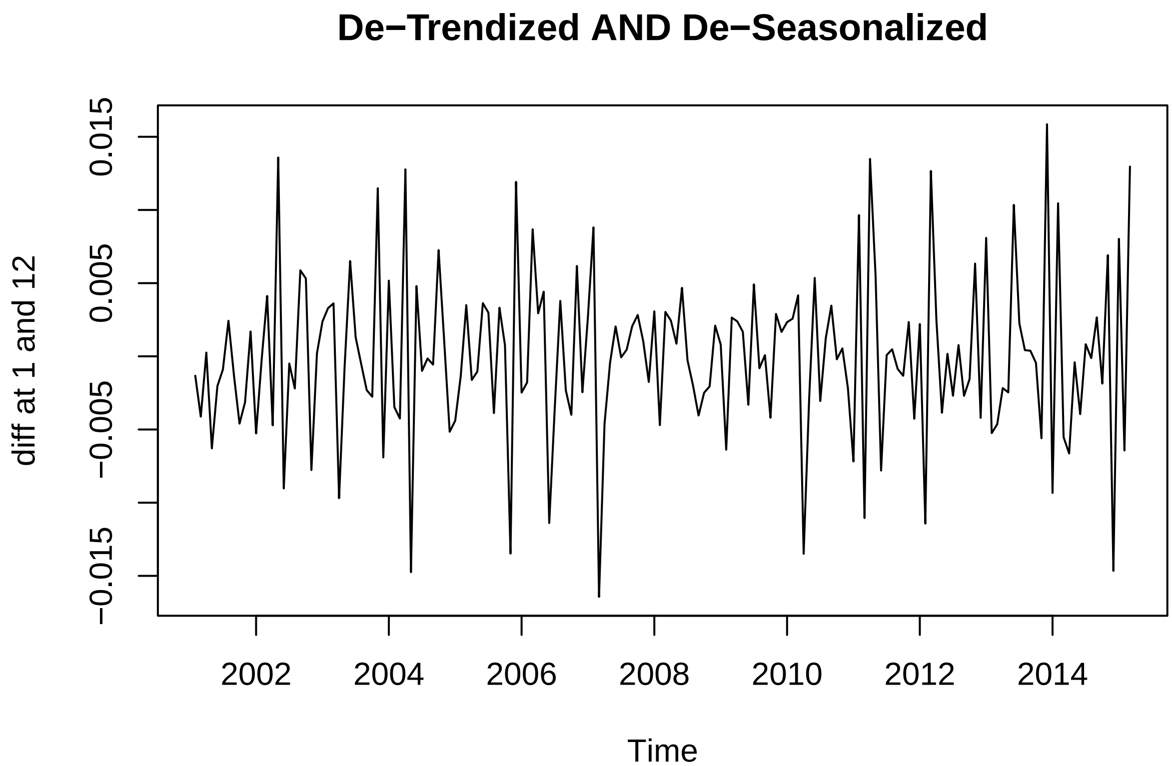


The acf and pacf plots show variance became stable.This shows by years, but my dataset is months data. So, the lag shows different numbers. Also, by plots, there is significant corrections every 12 lags. Also, plots show seasonality.

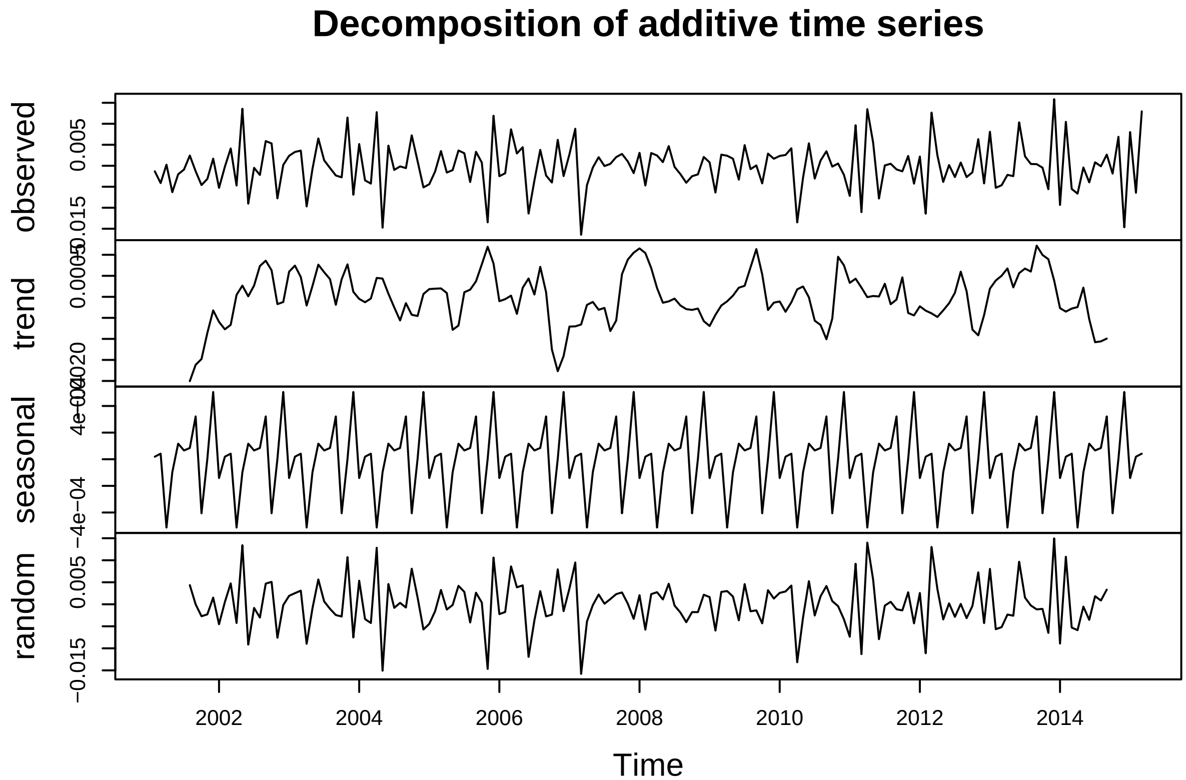




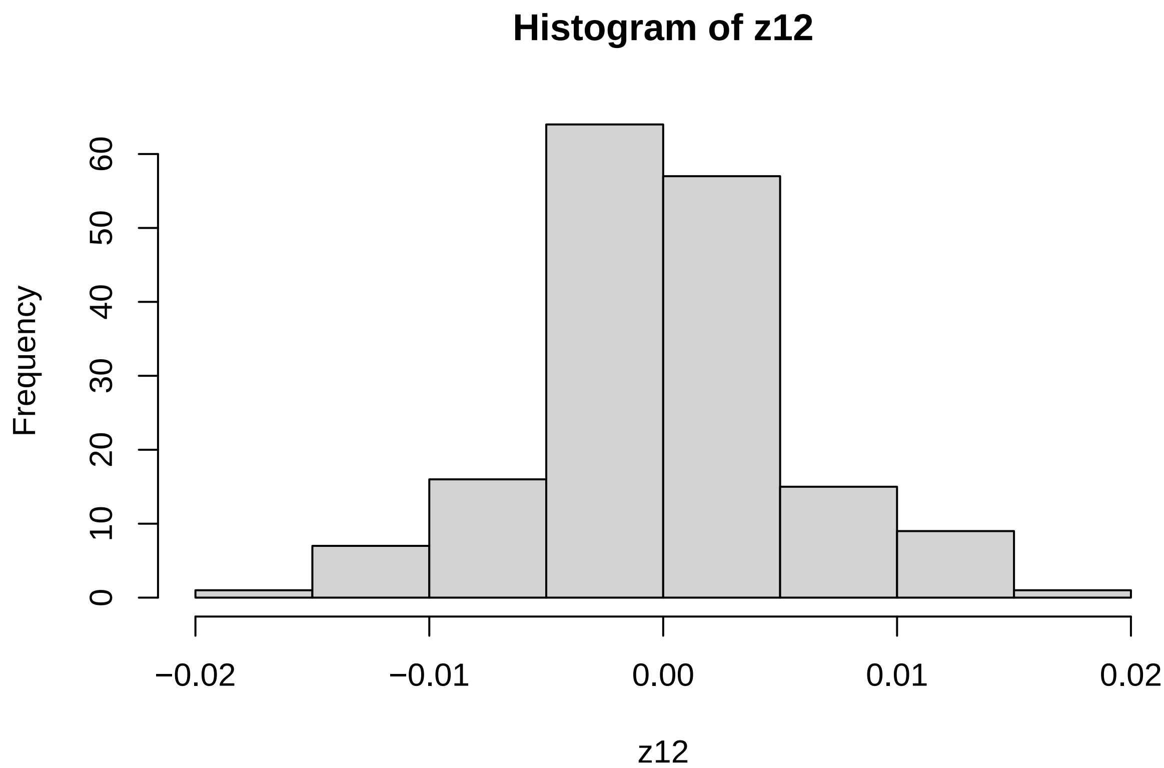
I differenced the box transformed data at lag 1 to remove trend. And trend was removed and looks normally distributed



The plot looks stationary



Decomposition shows the de seasonalized and de trended data in one plot. Trend is gone and seasonal was reduced



No more trend, reduced seasonality and no variance changing and normally distributed. Looks like little bit concentrate on the middle. But it is good to use.

```
## Warning: package 'tseries' was built under R version 4.1.2
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
## Warning in adf.test(z12): p-value smaller than printed p-value
```

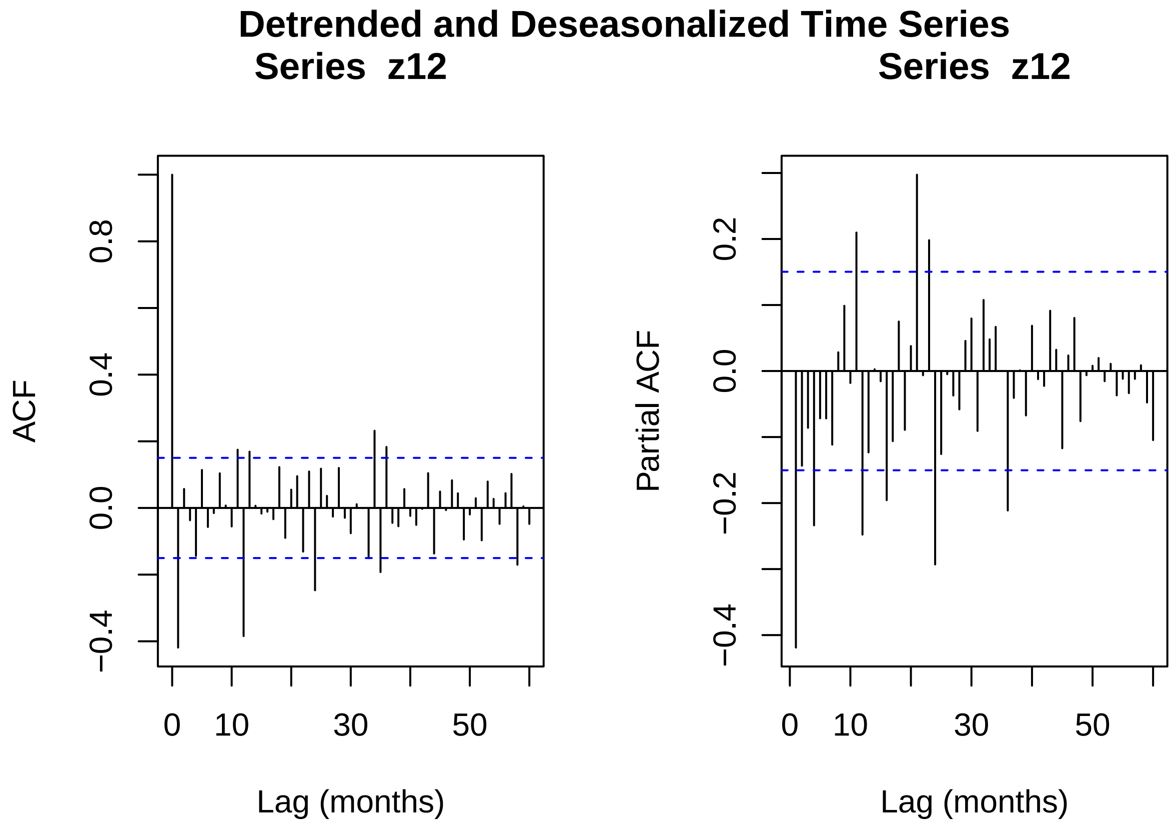
```
##
##  Augmented Dickey-Fuller Test
##
## data:  z12
## Dickey-Fuller = -7.8508, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary
```

```
## Variance:  3.210415e-05
```

```
## mean =  -5.614088e-05
```

I differenced again the box transformed data at lag 12 to remove seasonality Variance decreased and mean is almost zero. pvalue is 0.01 which is less then 0.05 from adf.test. The plots look good. The data is stationary

I will use acf and pacf plots of stationary data to determine the models



Based on the acf and pacf plots, d = 0 and D =1 Also, there is significant peak at lag 12 and cut off after 0 from acf and significant peaks at 12,24,36 from pacf. So, P can be 1,2,3 and Q can be 0,1 There are significant peaks at lag 1,4,11 from acf and significant peaks at lag 2,4 from pacf, so p can be 2,4 and q can be 1,4,11

```
## Warning: package 'forecast' was built under R version 4.1.2
```

```
## Series: rate.bc
## ARIMA(1,0,1)(2,1,1)[12] with drift
##
## Coefficients:
##          ar1      ma1     sar1     sar2     sma1  drift
##       0.7645  -0.3789  -0.142  -0.2916  -0.6594  1e-04
## s.e.  0.1129   0.1565   0.139   0.1201   0.1654  1e-04
##
## sigma^2 = 1.336e-05:  log likelihood = 713.69
## AIC=-1413.38   AICc=-1412.69   BIC=-1391.39
```

Also ARIMA(1,0,1)(2,1,1)[12] can be a candidate based on auto. arima function

Thus I identifed these two models for forecasting and I should decide the best one.

- 1. SARIMA $(4, 0, 1)$ x $(1, 1, 0)_{12}$
    - AIC $= -1363.35$
    - $((1 + 0.5622B\breve{}0.5335B^2 - 0.1655B^3 + 0.085B^4)(1 - B^{12})Y_t = (1 + 0.9642B)Z_t)$
- 2. SARIMA $(1, 0, 1)$ x $(2, 1, 1)_{12}$
    - AIC $= -1409.07$
    - $((1 - 0.8693B)(1 + 0.2144B^{12} + 0.3446B^{24})(1 - B^{12})Y_t = (1 - 0.4665B)(1 - 0.5469B^{12})Z_t)$

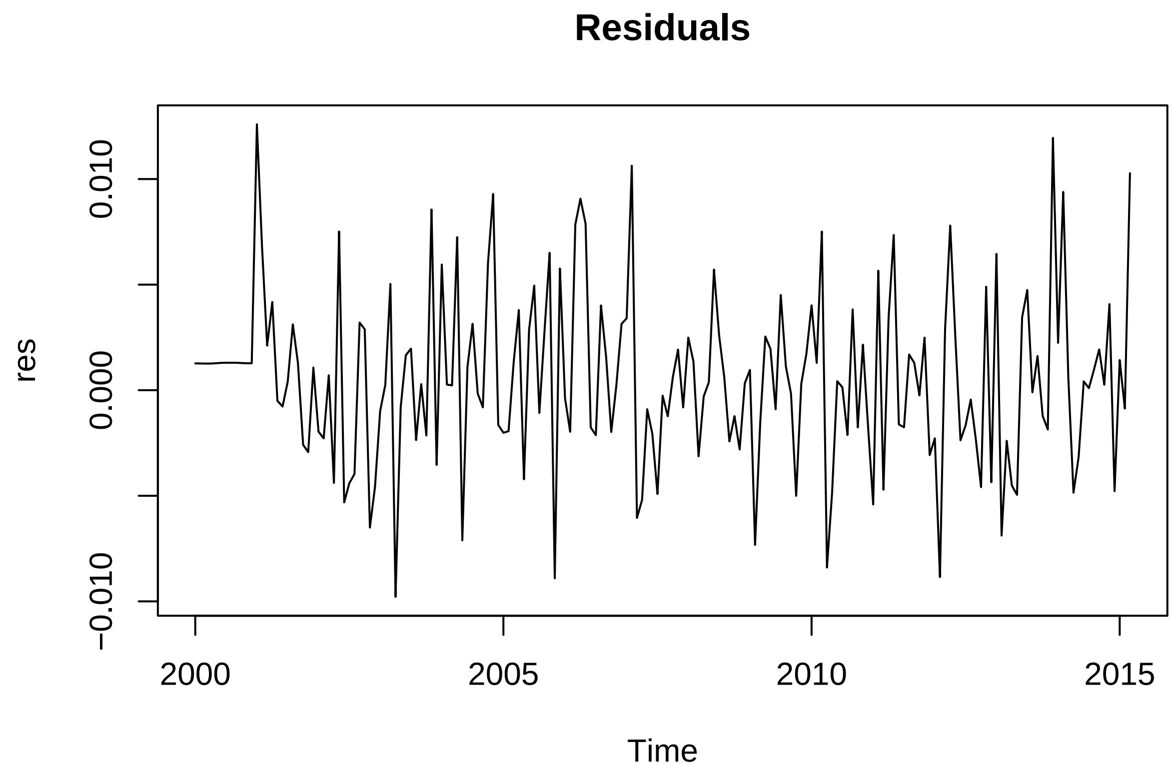
```
## AR4 roots  =  -1.154151-0i 1.550605+2.790879i 1.550605-2.790879i
```

```
## MA1 roots =  -1.037129+0i
```

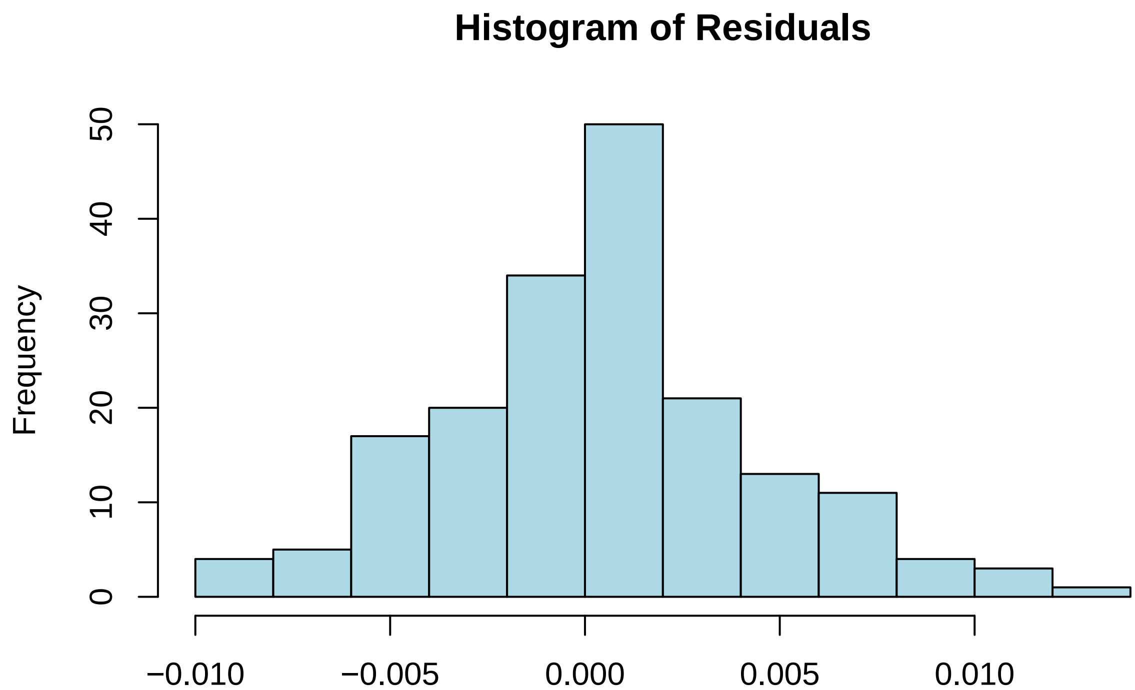
```
## SAR1 roots =  2.557545+0i
```

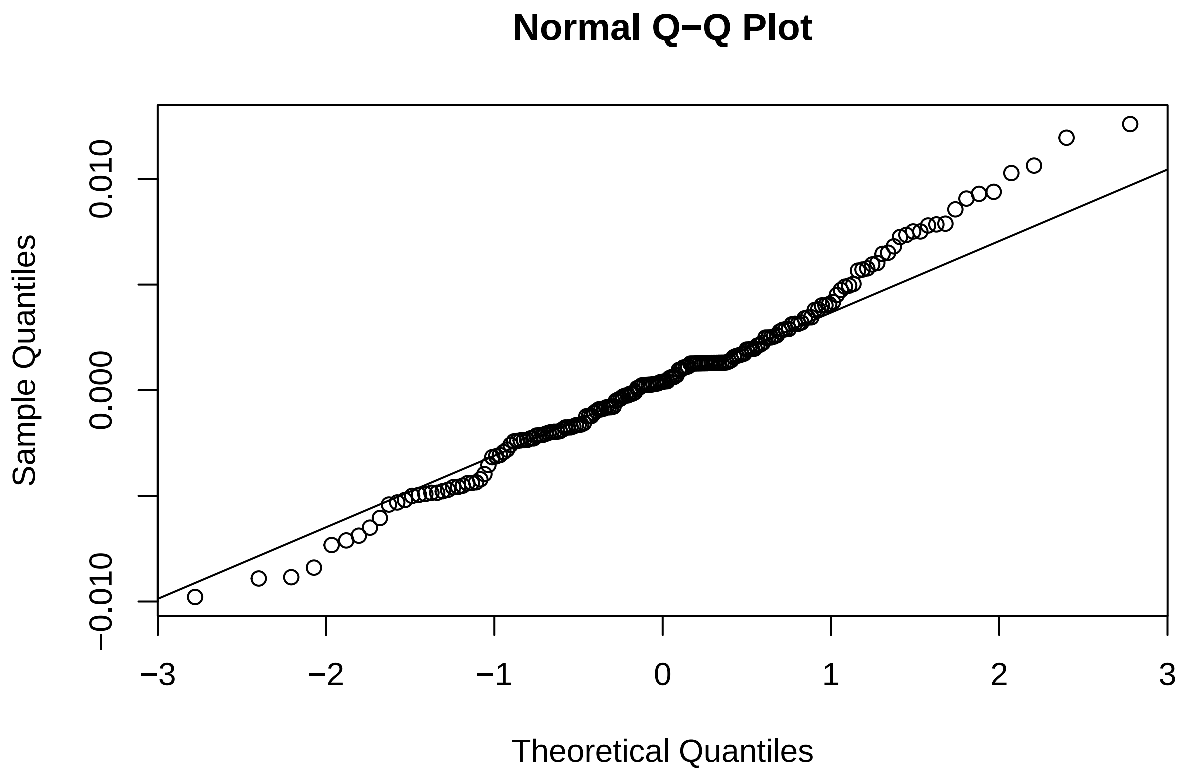
Roots are outside of unit circle Therefore Model 1 is causal and invertible.

```
##
## Call:
## arima(x = rate.bc, order = c(4, 0, 1), seasonal = list(order = c(1, 1, 0), period = 12),
##     method = "ML")
##
## Coefficients:
##           ar1     ar2     ar3      ar4     ma1     sar1
##       -0.5622  0.5335  0.1655  -0.0805  0.9642  -0.3910
## s.e.   0.0873  0.0902  0.0915   0.0814  0.0455   0.0734
##
## sigma^2 estimated as 1.816e-05:  log likelihood = 689.11,  aic = -1364.22
```



The plot of residuals look stationary


Histogram of Residuals
Frequency
0
10
20
30
40
50
−0.010
−0.005
0.000
0.005
0.010



The plot of the histogram looks symmetric and the QQ plot seems to be normally distributed.

```
## Mean  0.000525096
```

```
## var  1.689618e-05
```

```
## lag  13.96424
```

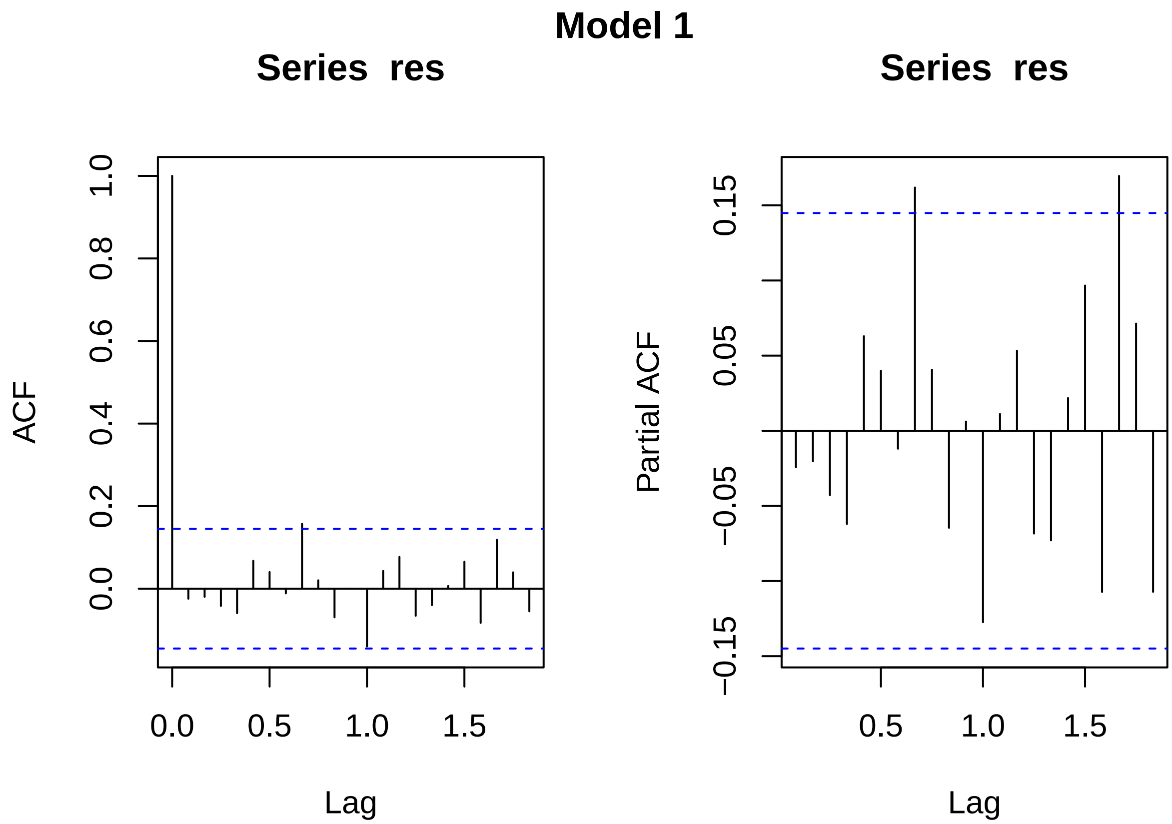
lag should be square root of length of data Mean is almost zero and variance is small

```
##
##  Box-Pierce test
##
## data:  res
## X-squared = 12.779, df = 9, p-value = 0.1729
```

```
##
##  Box-Ljung test
##
## data:  res
## X-squared = 13.61, df = 9, p-value = 0.1369
```

```
##
##  Box-Ljung test
##
## data:  res^2
## X-squared = 16.769, df = 14, p-value = 0.2687
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res
## W = 0.98572, p-value = 0.05995
```



Residuals for model 1 shows no trend, variance changing, no seasonal component. Histrogram and QQ plot shows that this is normally distributed. By the disagnostic tests, this model passed all test which shows greater than 0.05. ACF and Pacf show that there is a point beyond the interval. But except this every points are inside the interval so I can say this is white noise. Also, roots are outside unit circle. Therefore Model 1 is causal and invertible.Mean and var are also good looking from data

```
## [1] "Analysis for model 2"
```

```
## AR1 roots  =  1.150351+0i
```

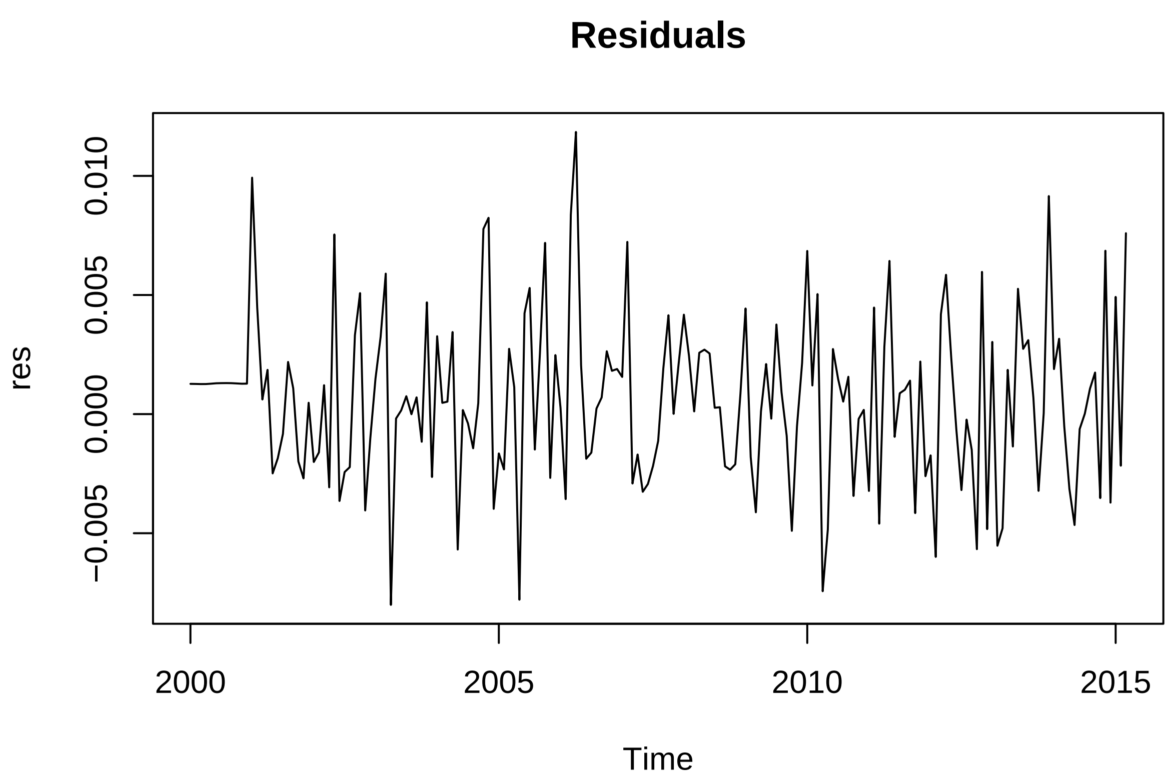
```
## MA1 roots =  -0.311085+1.674856i -0.311085-1.674856i
```

```
## SAR1 roots =  2.143623+0i
```

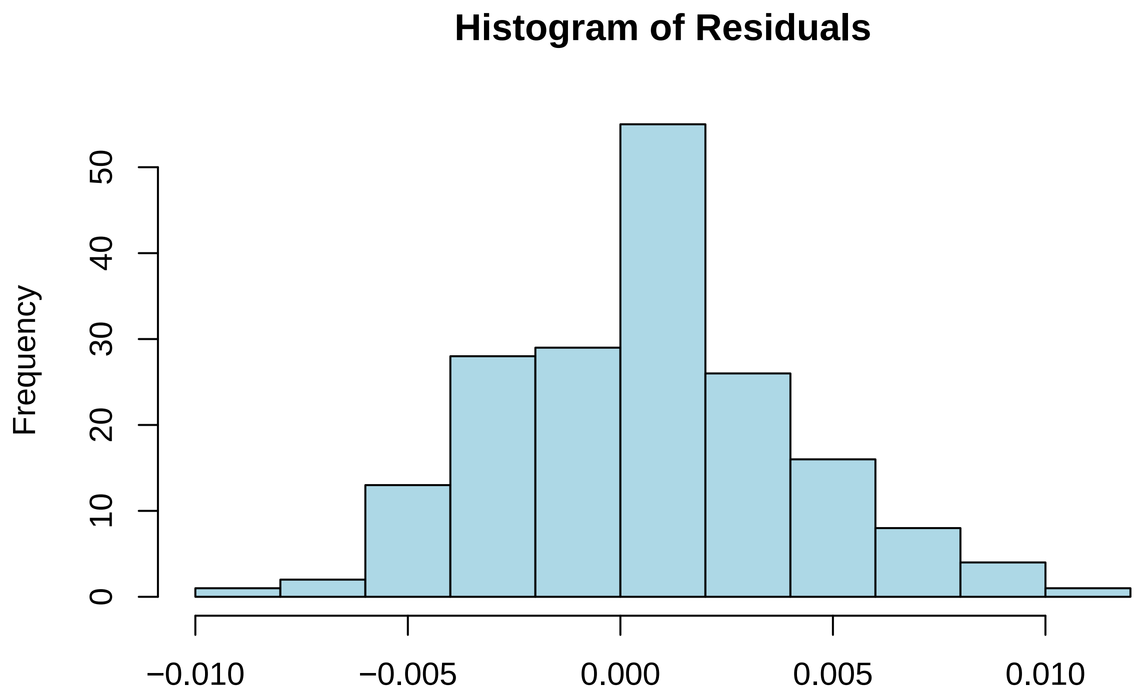
```
## SAR1 roots =  1.828488+0i
```

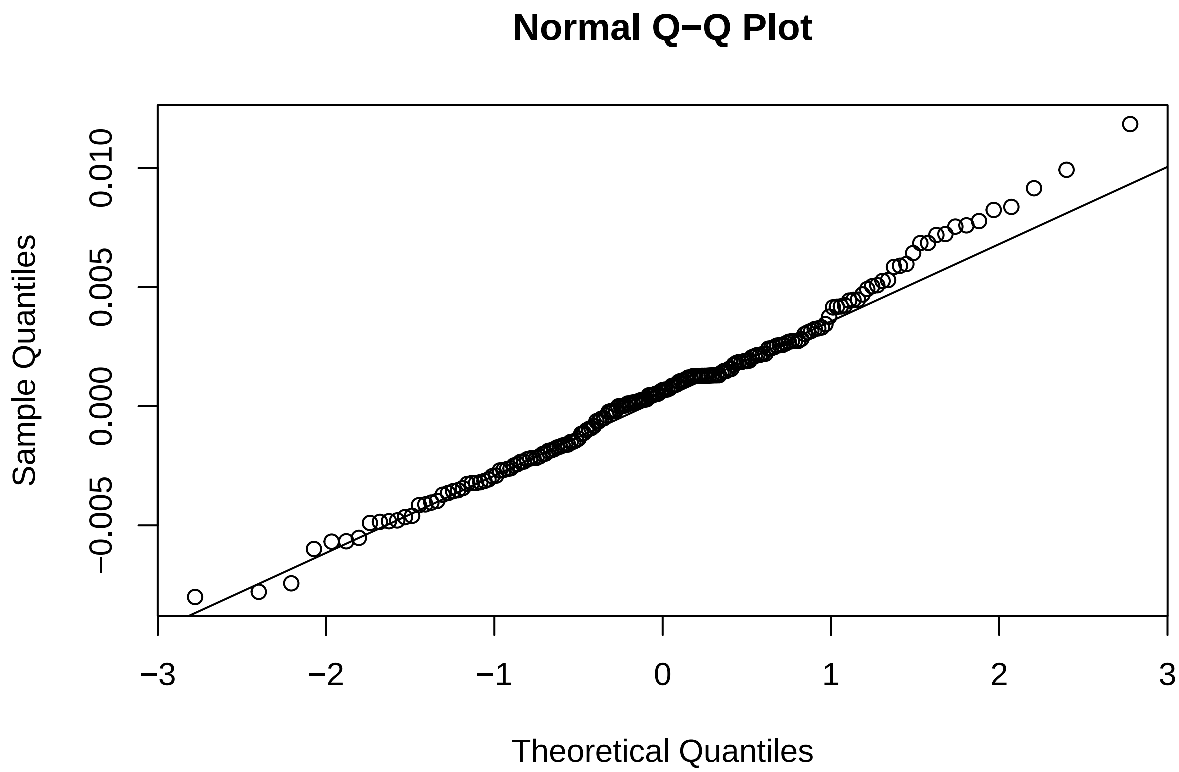
model 2 is also causal and invertible.The roots are outside of unit circle

```
##
## Call:
## arima(x = rate.bc, order = c(1, 0, 1), seasonal = list(order = c(2, 1, 1), period = 12),
##     method = "ML")
##
## Coefficients:
##          ar1      ma1     sar1     sar2     sma1
##       0.8693  -0.4665  -0.2144  -0.3446  -0.5469
## s.e.  0.0867   0.1446   0.1315   0.1061   0.1561
##
## sigma^2 estimated as 1.337e-05:  log likelihood = 710.53,  aic = -1409.07
```



The plot of residuals look stationary


Histogram of Residuals
Frequency
0
10
20
30
40
50
−0.010
−0.005
0.000
0.005
0.010



The plot of the histogram looks symmetric and the QQ plot seems to be normally distributed.

```
## Mean  0.0005937182
```

```
## var  1.231702e-05
```

```
## lag  13.96424
```

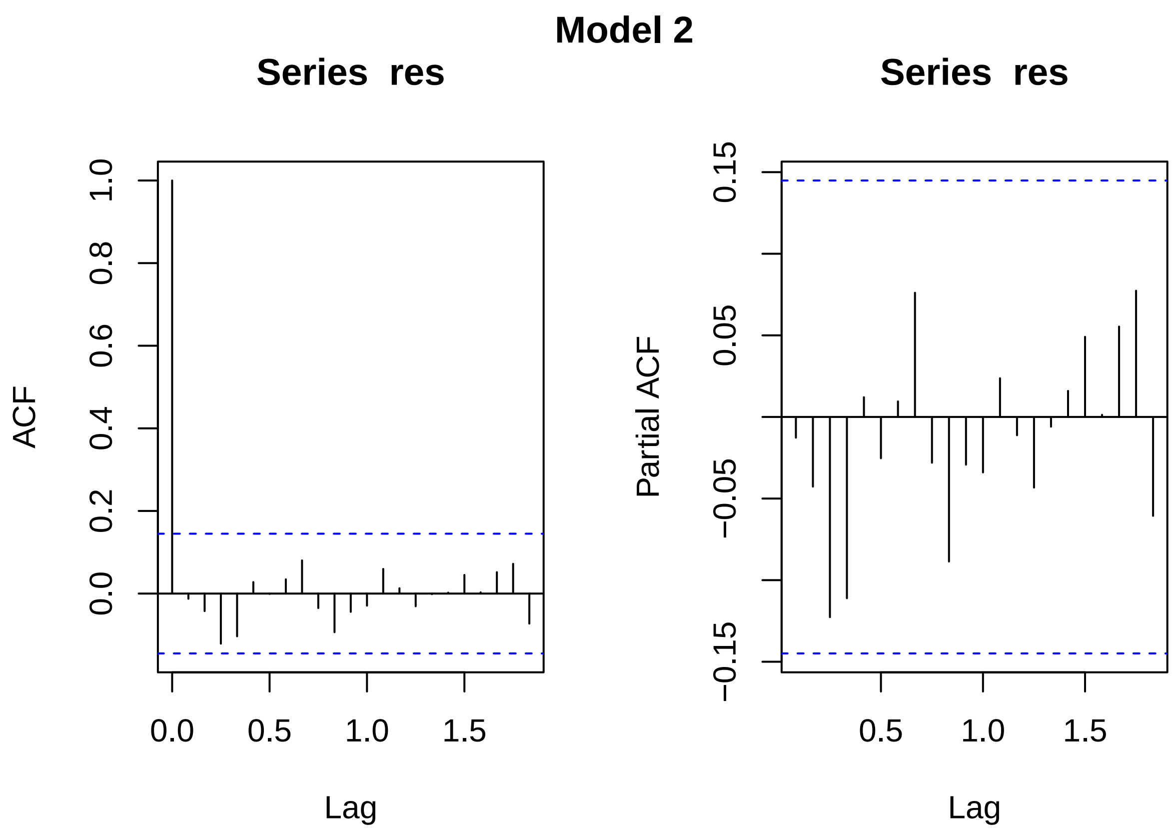
lag should be square root of length of data Mean is almost zero and variance is small

```
##
##  Box-Pierce test
##
## data:  res
## X-squared = 9.5799, df = 9, p-value = 0.3856
```

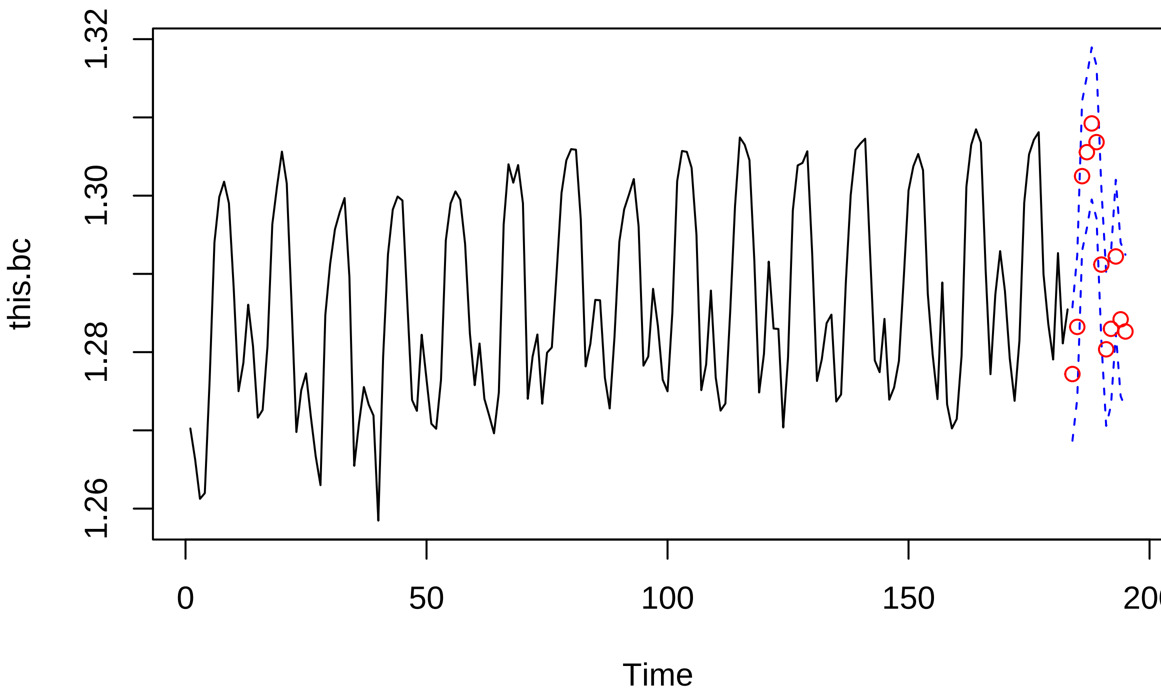
```
##
##  Box-Ljung test
##
## data:  res
## X-squared = 10.038, df = 9, p-value = 0.3474
```

```
##
##  Box-Ljung test
##
## data:  res^2
## X-squared = 24.517, df = 14, p-value = 0.03964
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res
## W = 0.98912, p-value = 0.1756
```



Residuals for model 2 shows no trend, variance changing, no seasonal component. Histrogram and QQ plot shows that this is normally distributed. However, by the disagnostic tests, this model didnt pass Box-Ljung test which shows less than 0.05. Every points are inside the interval so I can say this is white noise. Also, roots are outside unit circle. Therefore Model 1 is causal and invertible. Even though, model 2 has smaller AIC, model 1 passed all test and other factors satisfied to be use for forecast such as more normally distributed. I will use model 1 to forecast the next 12 values.

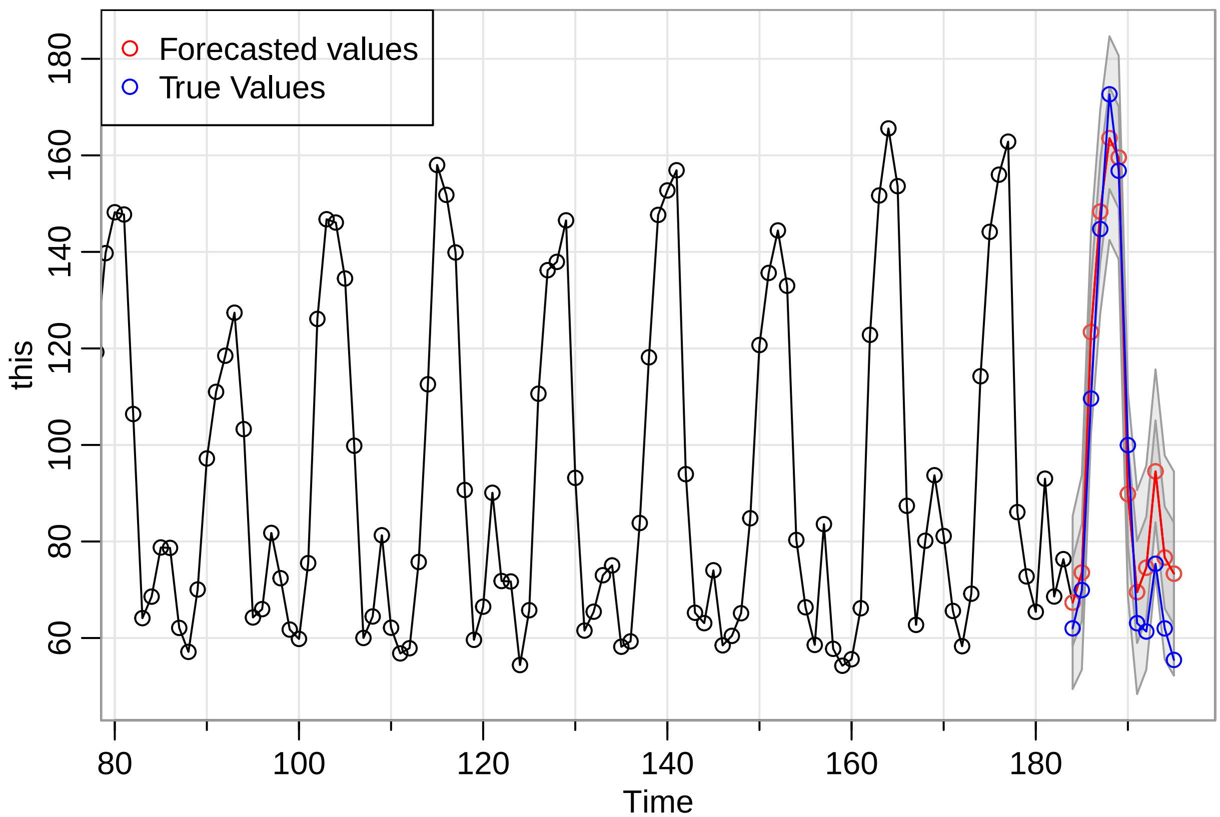


This is the predicted plot of boxcox transformed data. The blue line shows the interval of predicted data values. This shows the predicted values well

```
## Warning: package 'astsa' was built under R version 4.1.1
```

```
##
## Attaching package: 'astsa'
```

```
## The following object is masked from 'package:forecast':
##
##     gas
```



This is the plot of forecasted values and real values. The real values are in the intervals of the forecast. This shows the model 1 forecasting is mostly accurate, since the value is in the intervals. Blue is the rue values and red is forecasted values.

Conclusion

So, my goal was to predict the last 12 values with this process. And the goal was achieved since the prediction was working well. My model to predict the values was SARIMA $(4, 0, 1)$ x $(1, 1, 0)_{12}$

detailed

$((1 + 0.5622B\breve{}0.5335B^2 - 0.1655B^3 + 0.085B^4)(1 - B^{12})Y_t = (1 + 0.9642B)Z_t)$

and this shows the white noise of residuals and passed all of the diagnostic tests.This was the reason I picked model 1 to forecast the data.

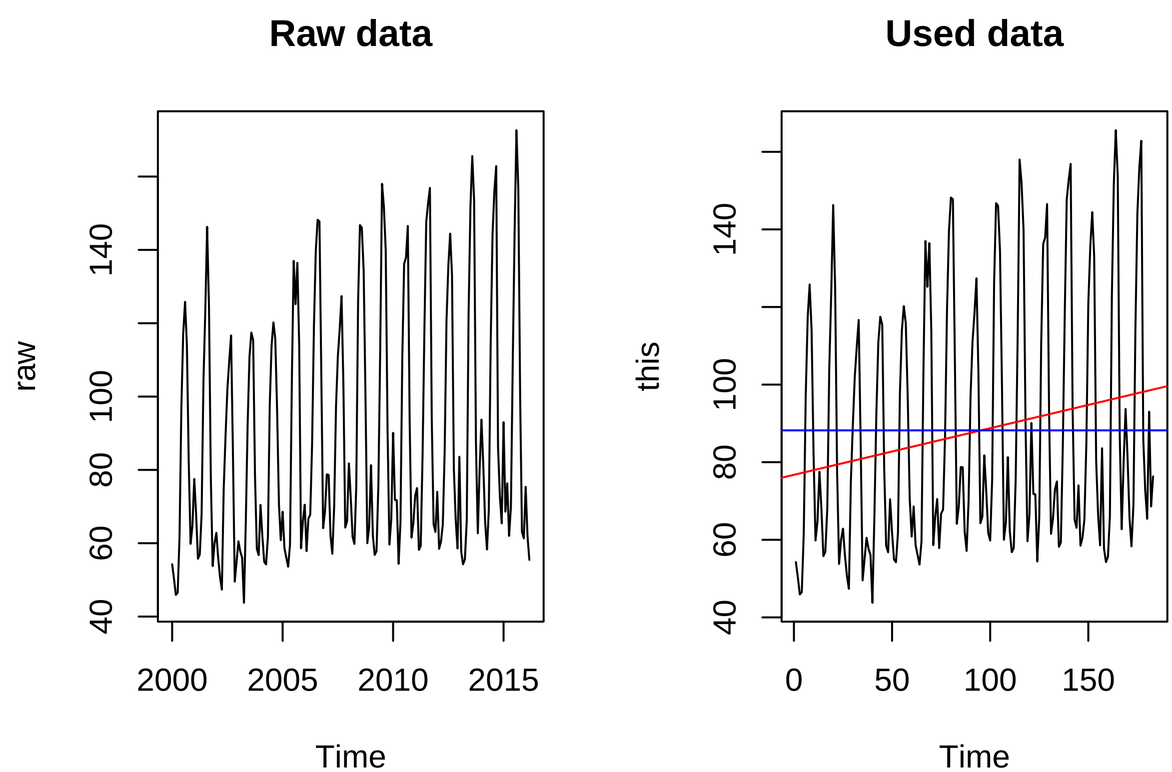
Professor Raya Feldman and TA Youhong provide me information and lectures to finish my project.

Reference

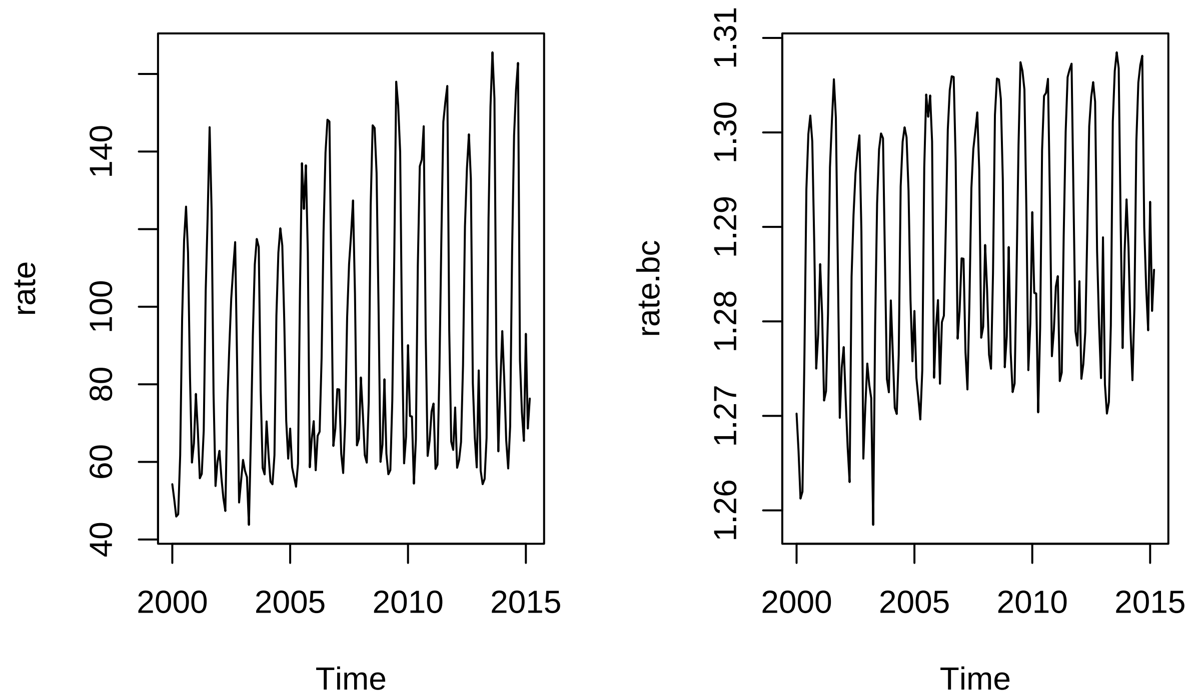
Data from https://data.world/cityofaustin/d9pb-3vh7

Appendix

```
data = read.csv("C:/Users/sungu/Desktop/kwh.csv") # Electricity bills in the Austin
data$Average.Bill = as.numeric(gsub("\\$", "", data$Average.Bill)) # Average bill in Austin Tx
employ= rev(data$Average.Bill) #dataset shows counterorder of time so reversed the order

rate = ts(employ, start = c(2000,1),end = c(2015,3),frequency = 12) # dataset from Jan 2000 to Dec 2014
test.rate = ts(employ, start = c(2015,4), end = c(2016,3), frequency = 12) # test dataset for forecast.
this = employ[c(1: 183)]
this.test = employ[c(184:195)]

par(mfrow=c(1, 2))
raw = ts(employ, start = c(2000,1),frequency = 12) # raw dataset 2000 to 2015 monthly average bill
plot(raw,main = "Raw data")
ts.plot(this,main = "Used data")
fitt <- lm(this ~ as.numeric(1:length(this)))
abline(fitt, col="red")
abline(h=mean(raw), col="blue")
```




```
par(mfrow=c(1, 2))
plot(rate)
plot(rate.bc)
```



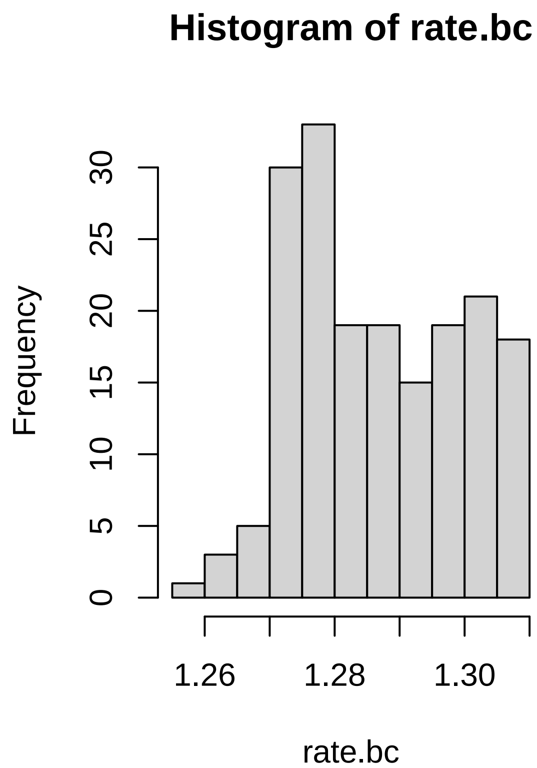

```r
#Compare var
var(rate)
```

```
## [1] 1043.242
```

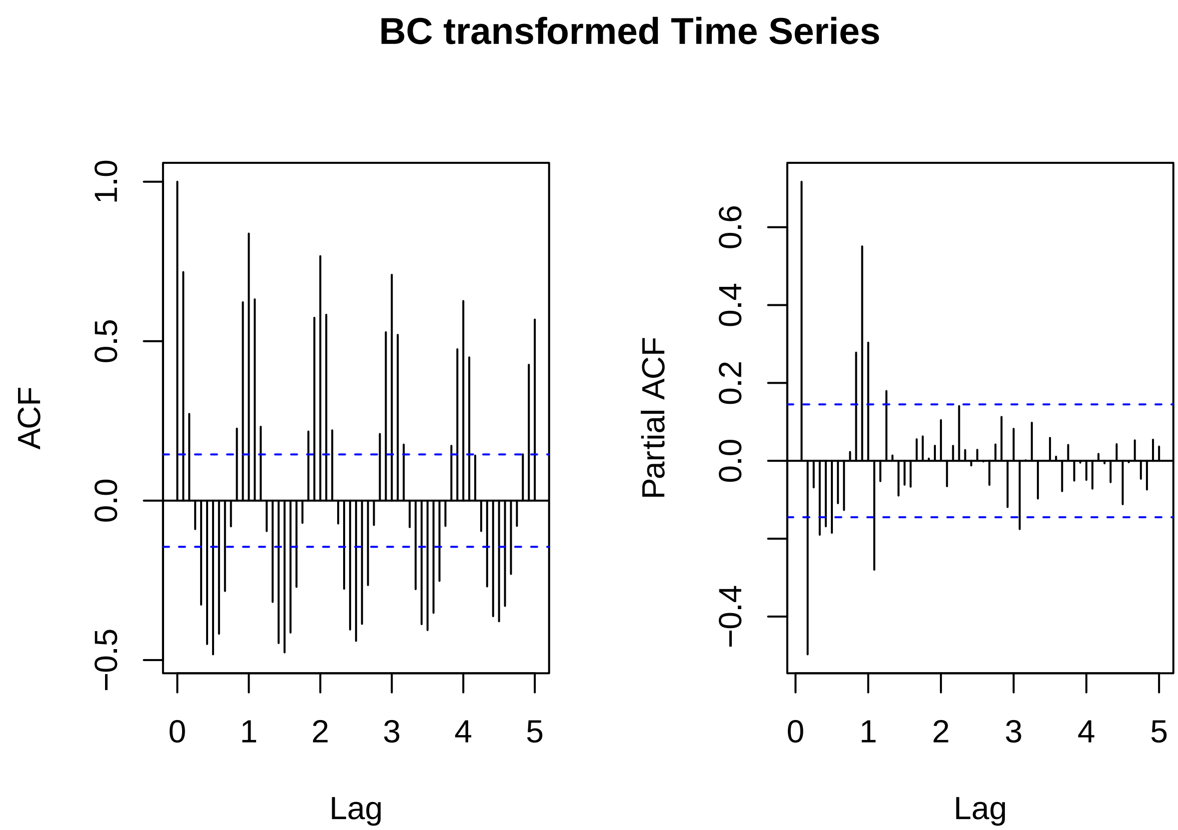
```r
var(rate.bc)
```

```
## [1] 0.0001587282
```

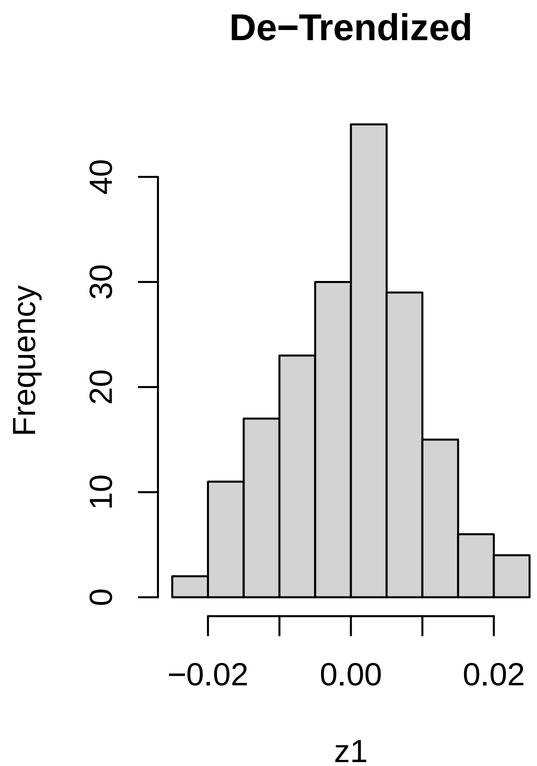
```r
hist(rate.bc)

# acf and pacf for transformed dataset.
par(mfrow=c(1, 2))
```

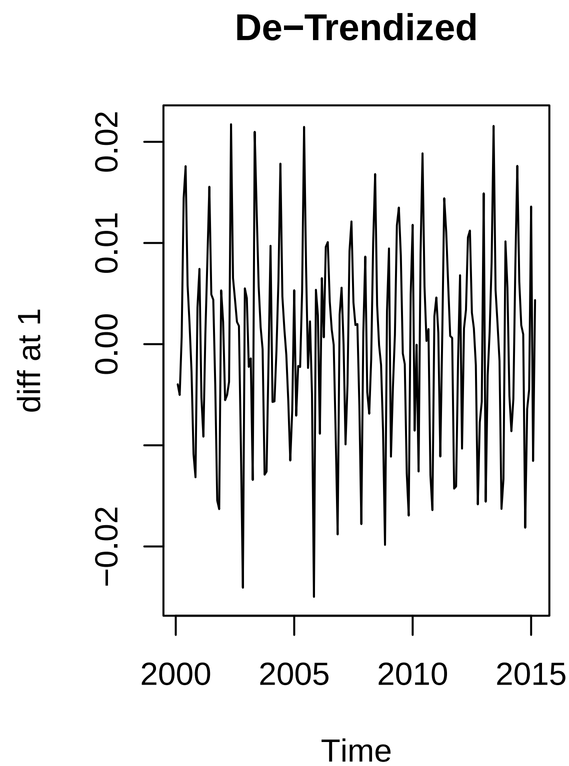



```
acf(rate.bc,lag.max = 60,main = "")
pacf(rate.bc,lag.max = 60,main = "")
title("BC transformed Time Series", line = -1, outer=TRUE) #
```

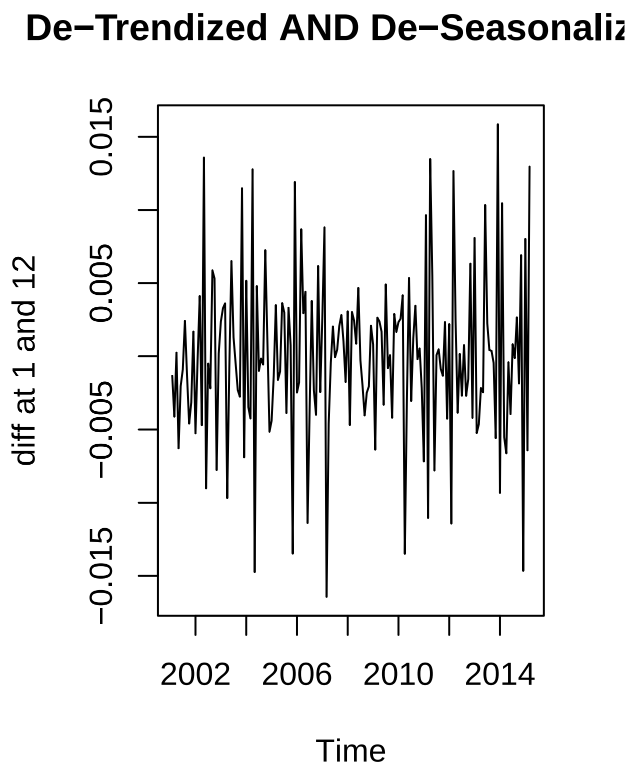



```
z1 = diff(rate.bc, 1)

par(mfrow=c(1, 2))
hist(z1,main = "De-Trendized")
ts.plot(z1,main = "De-Trendized",ylab = expression('diff at 1'))
```

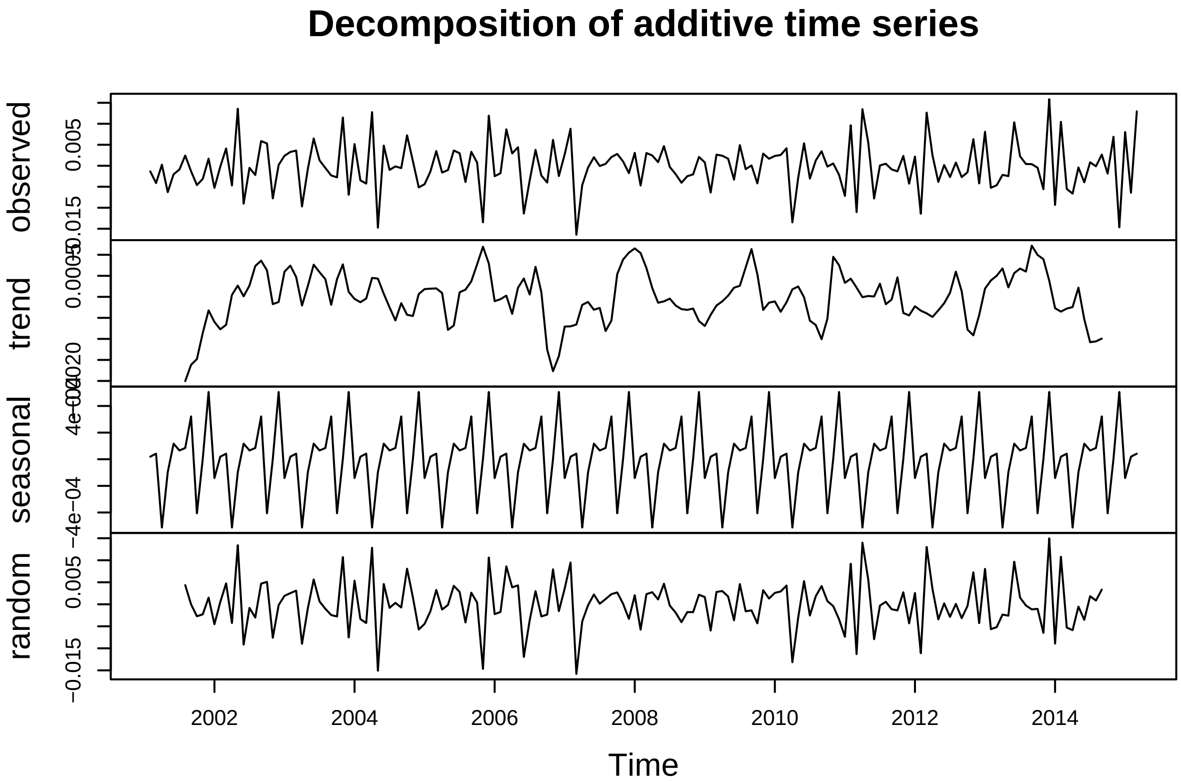





```
z12 = diff(z1,12)
ts.plot(z12,main = "De-Trendized AND De-Seasonalized",ylab = expression("diff at 1 and 12"))

par(mfrow=c(1, 2))
```




```
plot(decompose(z12))
```




```
hist(z12)

library(tseries)
adf.test(z12)
```

```
## Warning in adf.test(z12): p-value smaller than printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  z12
## Dickey-Fuller = -7.8508, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary
```

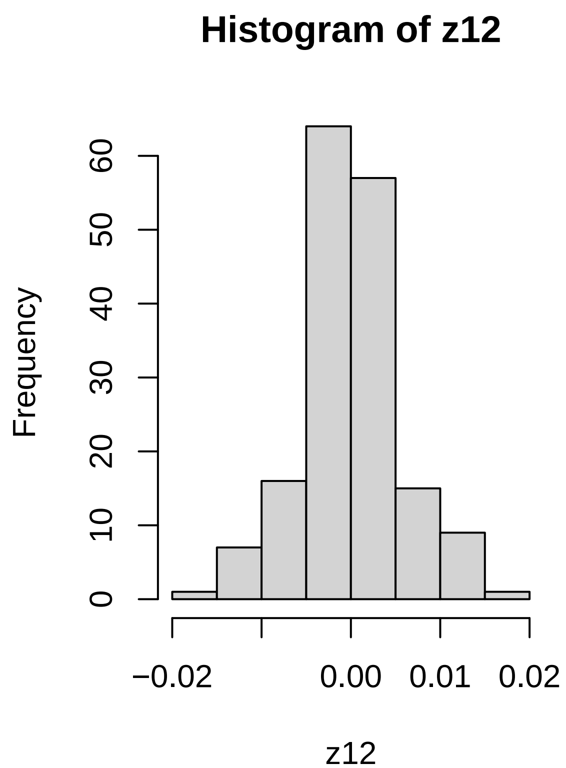
```
cat("Variance: ", var(z12),"\n")#creased
```

```
## Variance:  3.210415e-05
```

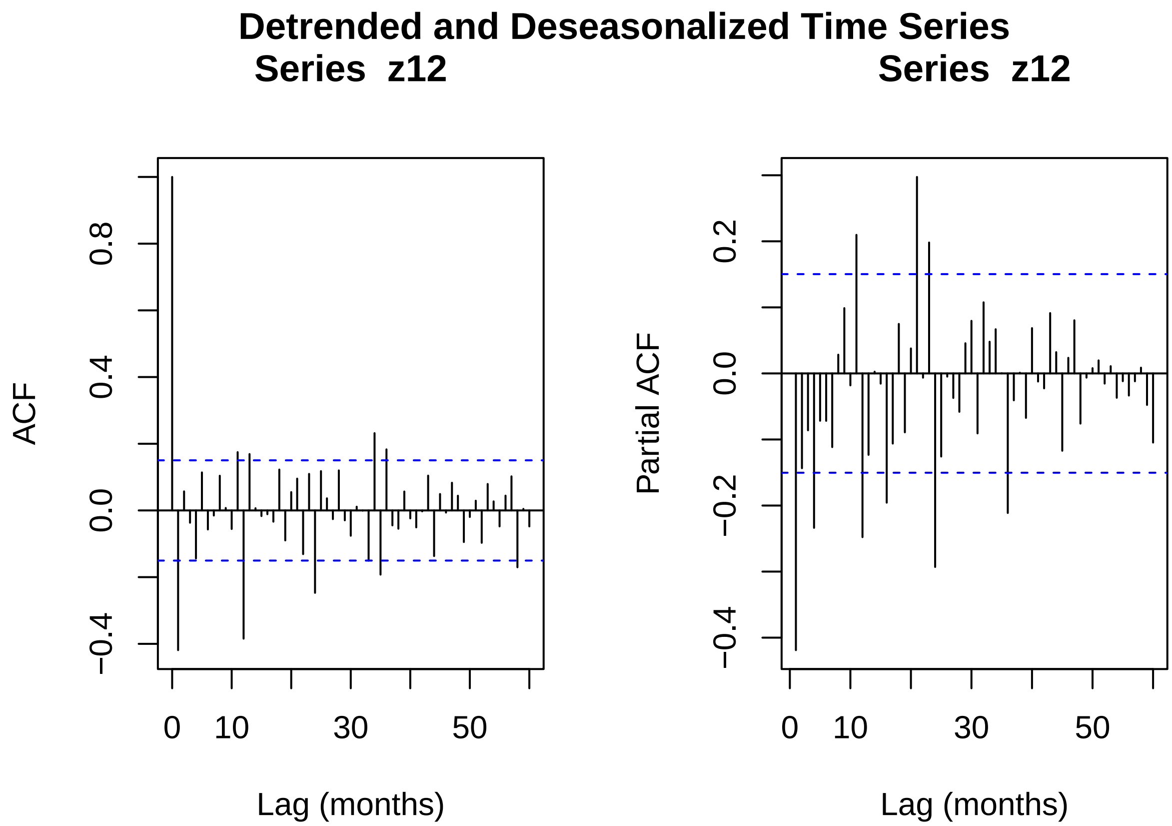
```
cat('mean = ',mean(z12)) # mean almost zero
```

```
## mean =  -5.614088e-05
```

```r
acfplz <- acf(z12, lag.max = 60,plot=FALSE)
pacfplz<- pacf(z12,lag.max = 60,plot=FALSE)

## Transform the lags from years to months
acfplz$lag <- acfplz$lag * 12
pacfplz$lag <- pacfplz$lag * 12

par(mfrow=c(1, 2))
```




```r
plot(acfplz, xlab="Lag (months)")
plot(pacfplz, xlab="Lag (months)")
title("Detrended and Deseasonalized Time Series", line = -1, outer=TRUE)
```




```
library(forecast)
auto.arima(rate.bc)
```

```
## Series: rate.bc
## ARIMA(1,0,1)(2,1,1)[12] with drift
##
## Coefficients:
##          ar1      ma1     sar1     sar2     sma1   drift
##       0.7645  -0.3789  -0.142  -0.2916  -0.6594  1e-04
## s.e.  0.1129   0.1565   0.139   0.1201   0.1654  1e-04
##
## sigma^2 = 1.336e-05:  log likelihood = 713.69
## AIC=-1413.38   AICc=-1412.69   BIC=-1391.39
```

```
fit1<-arima(rate.bc, order=c(4,0,1), seasonal=list(order=c(1,1,0), period=12),method="ML")
fit2<- arima(rate.bc, order=c(1,0,1), seasonal=list(order=c(2,1,1), period=12),method="ML")

cat("AR4 roots  = ", polyroot(c(1, 0.5622,-0.1655,0.085)),"\n")# AR4 model 1
```

```
## AR4 roots  =  -1.154151-0i 1.550605+2.790879i 1.550605-2.790879i
```

```
cat("MA1 roots = ", polyroot(c(1, 0.9642 )),"\n") #MA1 for Model 1
```

```
## MA1 roots =  -1.037129+0i
```

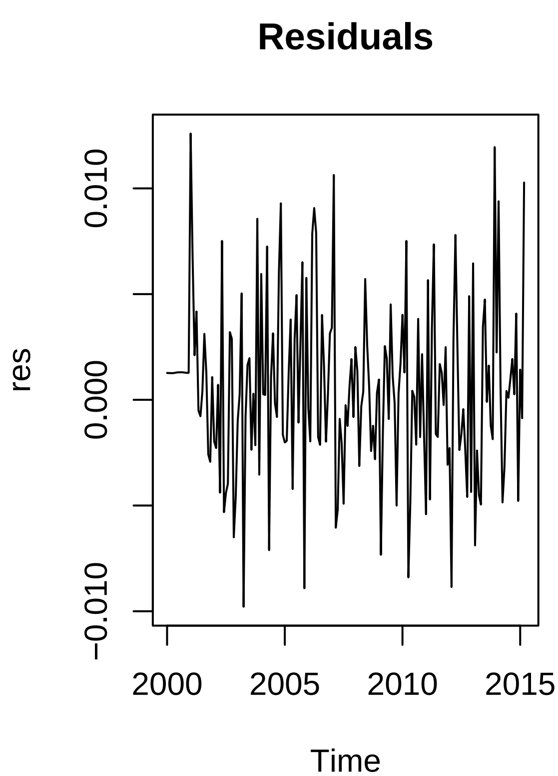
```r
cat("SAR1 roots = ", polyroot(c(1, -0.3910 )),"\n") #Sar model 1
```

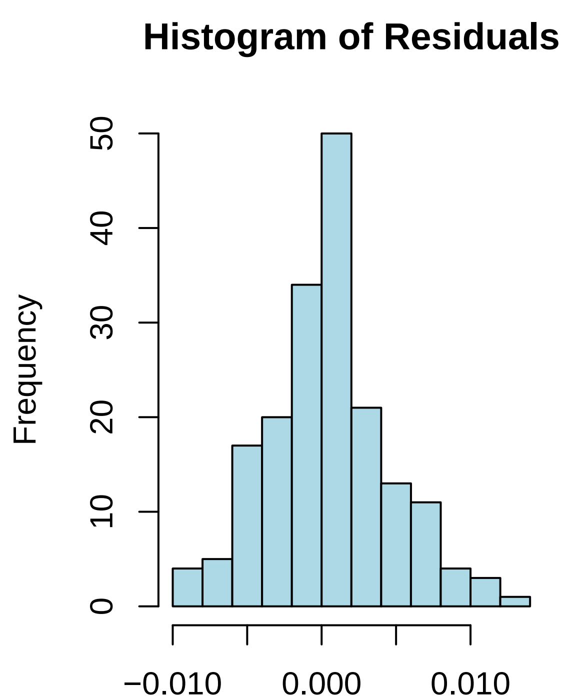
```
## SAR1 roots =  2.557545+0i
```

```r
print(fit1)
```

```
##
## Call:
## arima(x = rate.bc, order = c(4, 0, 1), seasonal = list(order = c(1, 1, 0), period = 12),
##     method = "ML")
##
## Coefficients:
##           ar1     ar2     ar3      ar4     ma1      sar1
##       -0.5622  0.5335  0.1655  -0.0805  0.9642  -0.3910
## s.e.   0.0873  0.0902  0.0915   0.0814  0.0455   0.0734
##
## sigma^2 estimated as 1.816e-05:  log likelihood = 689.11,  aic = -1364.22
```

```r
res<-residuals(fit1)
plot(res, main="Residuals")

hist(res, col="light blue", xlab="", main="Histogram of Residuals")
```





```r
qqnorm(res);qqline(res)

cat("Mean ",mean(res),"\n")
```

```
## Mean  0.000525096
```

```r
cat("var ",var(res),"\n")
```

```
## var  1.689618e-05
```

```r
cat("lag ",sqrt(length(rate)+12),"\n")
```

```
## lag  13.96424
```

```r
#Diagnostic test
Box.test(res,lag=14, type = c("Box-Pierce"), fitdf=5)
```

```
##
##  Box-Pierce test
##
## data:  res
## X-squared = 12.779, df = 9, p-value = 0.1729
```

```r
Box.test(res,lag=14, type = "Ljung", fitdf=5)
```

```
##
##  Box-Ljung test
##
## data:  res
## X-squared = 13.61, df = 9, p-value = 0.1369
```

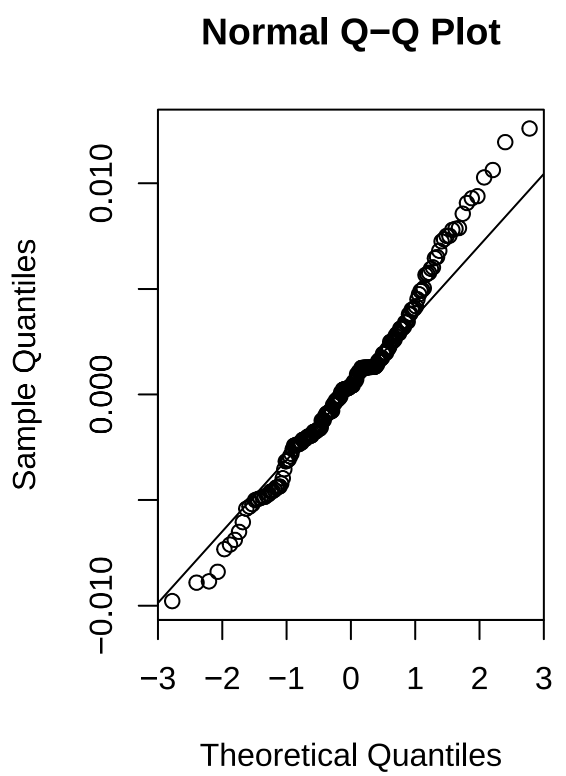
```r
Box.test(res^2, lag=14, type = c("Ljung-Box"), fitdf=0)
```

```
##
##  Box-Ljung test
##
## data:  res^2
## X-squared = 16.769, df = 14, p-value = 0.2687
```

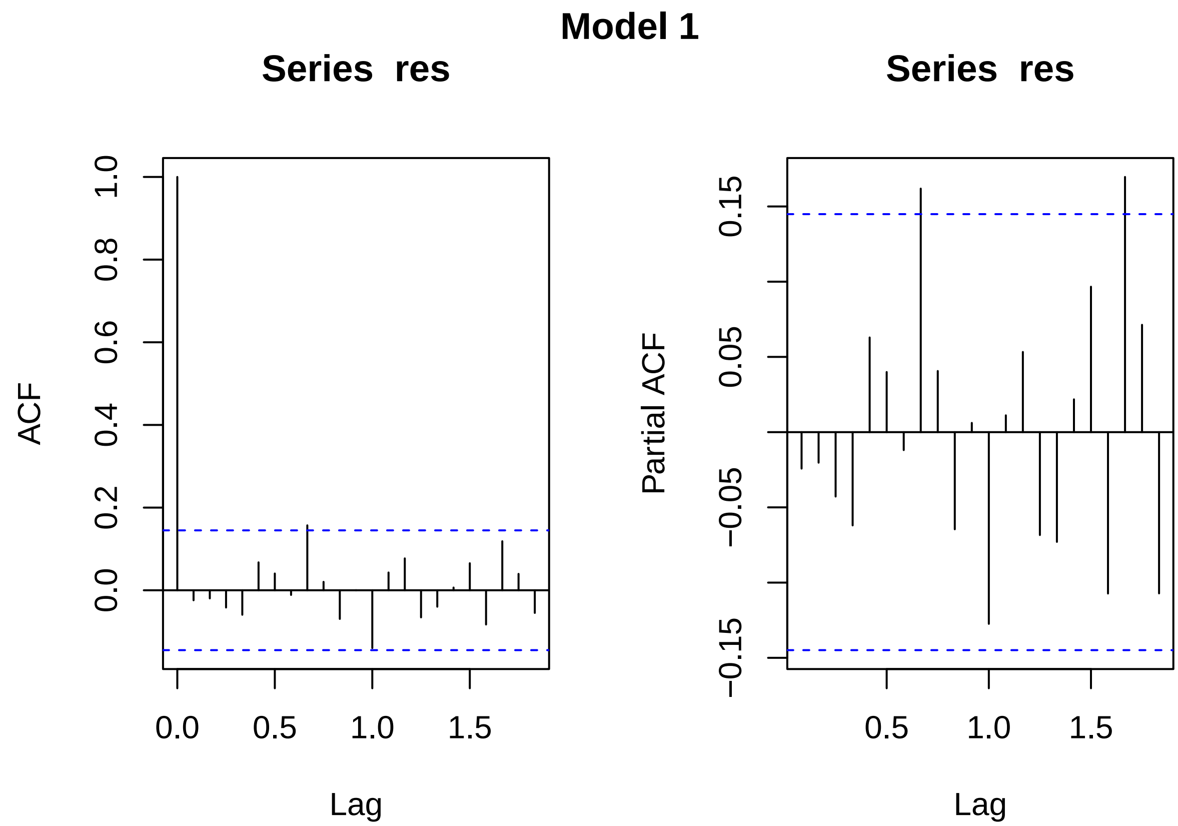
```r
shapiro.test(res)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res
## W = 0.98572, p-value = 0.05995
```

```r
par(mfrow=c(1, 2))
```




```r
acf(res)
pacf(res)
title("Model 1", line = -1, outer=TRUE)
```




```r
cat("AR1 roots  = ", polyroot(c(1,-0.8693)),"\n")#AR1 model 2
```

```
## AR1 roots  =  1.150351+0i
```

```r
cat("MA1 roots = ", polyroot(c(1, 0.2144,0.3446 )),"\n") #SAR3 for Model 2
```

```
## MA1 roots =  -0.311085+1.674856i -0.311085-1.674856i
```

```r
cat("SAR1 roots = ", polyroot(c(1, -0.4665 )),"\n") #Ma1model 2
```

```
## SAR1 roots =  2.143623+0i
```

```r
cat("SAR1 roots = ", polyroot(c(1, -0.5469 )),"\n") #SMA1 model 2
```

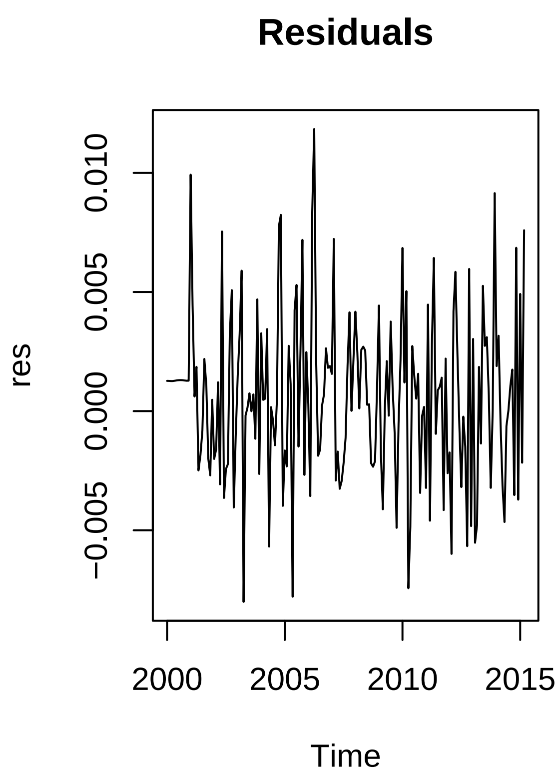
```
## SAR1 roots =  1.828488+0i
```

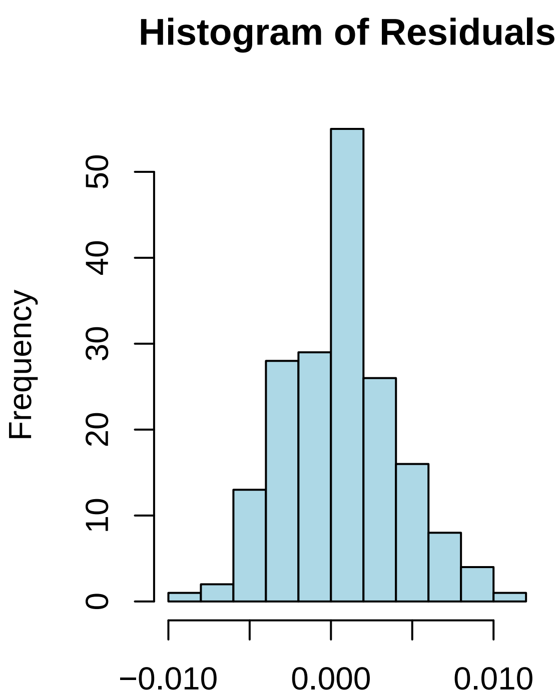
```r
print(fit2)
```

```
##
## Call:
## arima(x = rate.bc, order = c(1, 0, 1), seasonal = list(order = c(2, 1, 1), period = 12),
##     method = "ML")
```

```
##
## Coefficients:
##          ar1      ma1     sar1     sar2     sma1
##       0.8693  -0.4665  -0.2144  -0.3446  -0.5469
## s.e.  0.0867   0.1446   0.1315   0.1061   0.1561
##
## sigma^2 estimated as 1.337e-05:  log likelihood = 710.53,  aic = -1409.07
```

```
res<-residuals(fit2)
plot(res, main="Residuals")

hist(res, col="light blue", xlab="", main="Histogram of Residuals")
```






```
qqnorm(res);qqline(res)

cat("Mean ",mean(res),"\n")
```

```
## Mean  0.0005937182
```

```
cat("var ",var(res),"\n")
```

```
## var  1.231702e-05
```

```r
cat("lag ",sqrt(length(rate)+12),"\n")
```

```
## lag  13.96424
```

```r
#Diagnostic test
Box.test(res,lag=14, type = c("Box-Pierce"), fitdf=5)
```

```
##
##  Box-Pierce test
##
## data:  res
## X-squared = 9.5799, df = 9, p-value = 0.3856
```

```r
Box.test(res,lag=14, type = "Ljung", fitdf=5)
```

```
##
##  Box-Ljung test
##
## data:  res
## X-squared = 10.038, df = 9, p-value = 0.3474
```

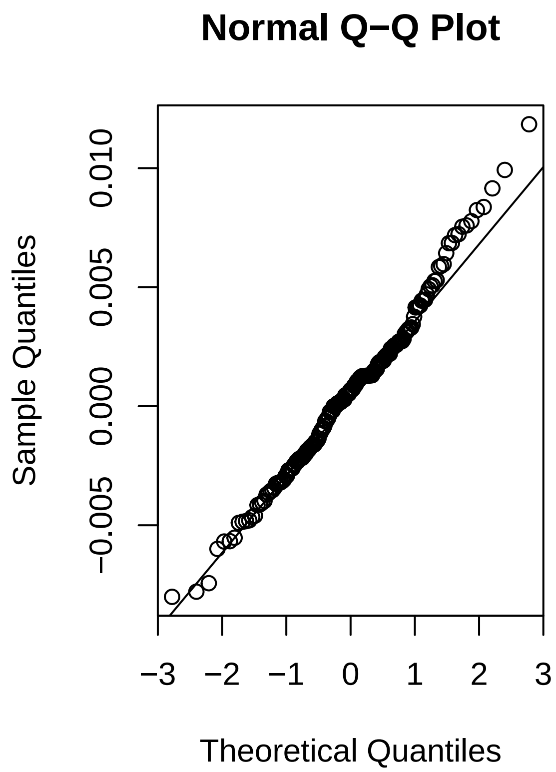
```r
Box.test(res^2, lag=14, type = c("Ljung-Box"), fitdf=0)
```

```
##
##  Box-Ljung test
##
## data:  res^2
## X-squared = 24.517, df = 14, p-value = 0.03964
```

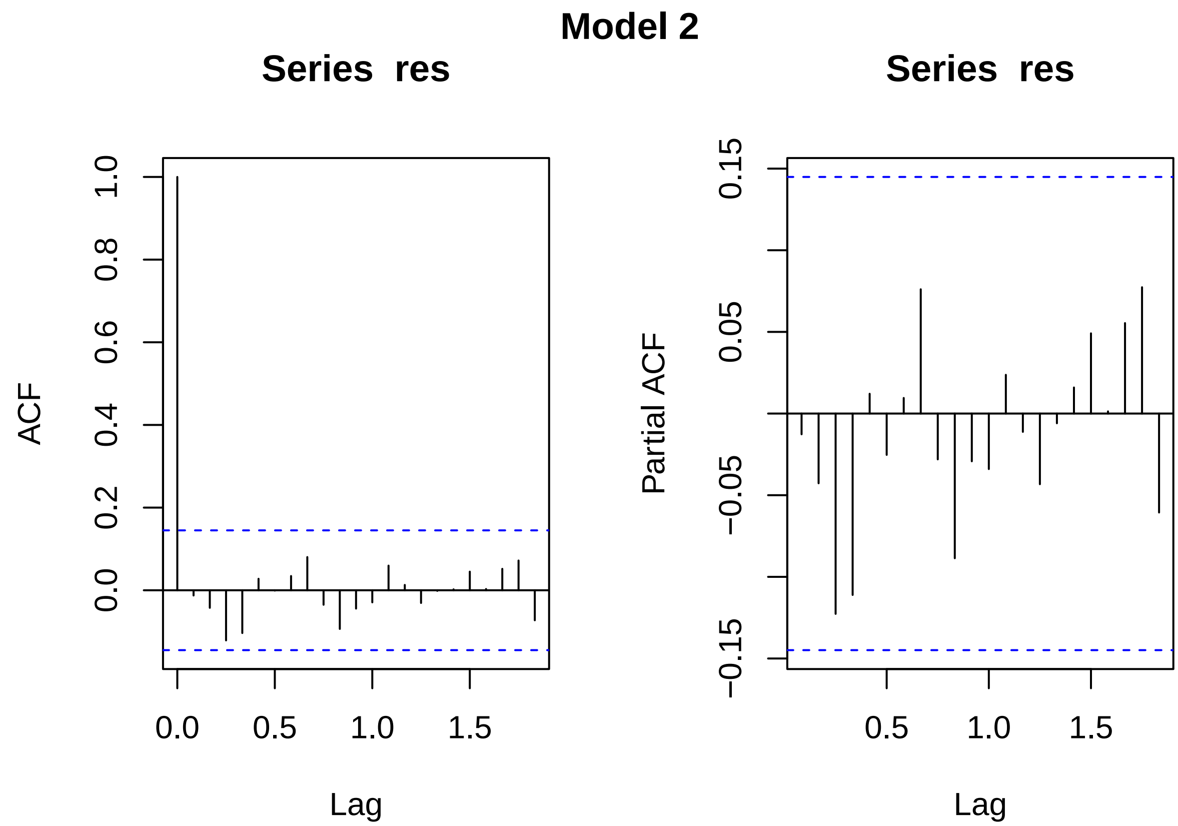
```r
shapiro.test(res)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res
## W = 0.98912, p-value = 0.1756
```

```r
par(mfrow=c(1, 2))
```




```
acf(res)
pacf(res)
title("Model 2", line = -1, outer=TRUE)
```




```r
forecast(fit1)
```

```
##          Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## Apr 2015       1.277199 1.271738 1.282661 1.268846 1.285552
## May 2015       1.283227 1.277340 1.289113 1.274224 1.292229
## Jun 2015       1.302488 1.296367 1.308609 1.293126 1.311850
## Jul 2015       1.305553 1.299328 1.311778 1.296033 1.315074
## Aug 2015       1.309228 1.303000 1.315455 1.299703 1.318752
## Sep 2015       1.306837 1.300580 1.313093 1.297268 1.316405
## Oct 2015       1.291190 1.284930 1.297449 1.281617 1.300763
## Nov 2015       1.280360 1.274090 1.286630 1.270771 1.289950
## Dec 2015       1.282979 1.276705 1.289253 1.273384 1.292574
## Jan 2016       1.292225 1.285946 1.298503 1.282623 1.301827
## Feb 2016       1.284190 1.277909 1.290471 1.274583 1.293796
## Mar 2016       1.282637 1.276354 1.288920 1.273028 1.292246
## Apr 2016       1.276176 1.269126 1.283227 1.265394 1.286959
## May 2016       1.282280 1.275083 1.289477 1.271273 1.293287
## Jun 2016       1.301373 1.294115 1.308631 1.290274 1.312473
## Jul 2016       1.305268 1.297971 1.312566 1.294108 1.316429
## Aug 2016       1.308557 1.301259 1.315854 1.297396 1.319718
## Sep 2016       1.307215 1.299906 1.314525 1.296036 1.318395
## Oct 2016       1.290808 1.283496 1.298119 1.279626 1.301989
## Nov 2016       1.281528 1.274212 1.288844 1.270339 1.292716
## Dec 2016       1.281510 1.274193 1.288828 1.270319 1.292702
## Jan 2017       1.292339 1.285019 1.299658 1.281144 1.303533
## Feb 2017       1.283030 1.275709 1.290351 1.271834 1.294226
```

```
## Mar 2017       1.283707 1.276385 1.291028 1.272510 1.294904
```

```
pred.tr <- predict(fit1, n.ahead = 12)
U.tr= pred.tr$pred + 2*pred.tr$se
L.tr= pred.tr$pred - 2*pred.tr$se
ts.plot(this.bc, xlim=c(1,length(this.bc)+12), ylim = c(min(this.bc),max(U.tr)))
lines((length(this.bc)+1):(length(this.bc)+12),U.tr, col="blue", lty="dashed")
lines((length(this.bc)+1):(length(this.bc)+12),L.tr, col="blue", lty="dashed")
points((length(this.bc)+1):(length(this.bc)+12), pred.tr$pred, col="red")

library(astsa)
pred.tr <- sarima.for(this, n.ahead=12, plot.all=F,
p=4, d=0, q=1, P=1, D=1, Q=0, S=12)
lines(184:195, pred.tr$pred, col="red")
lines(184:195, this.test, col="blue")
points(184:195, this.test, col="blue")
legend("topleft", pch=1, col=c("red", "blue"),
legend=c("Forecasted values", "True Values"))
```