Time Series Analysis – Final Project

Forecast of unemployment rate in U.S.

Group Details:

Name	Student Id
Vishal Patnaik Damodarapatruni	S3811521
Sri Venkata Manideepu Maddipati	S3820822
Vijay Chalamala	S3817359

Table of Contents

Introduction

The aim of this analysis is to predict the unemployment rate in United States of America based on the data released by U.S Department of Labor, which is also known as the Bureau of Labor Statistics (BLS). The Labor Bureau has been collecting the unemployment rate data since 1948 to present, to analyze the effect of employment due to external factors like recessions. This rate represents the percentage of the labor force who are unemployed and this rate is also defined as the U-3 measure of labor under utilization.

A detailed analysis is done using different time series techniques to identify the best model which predicts the future unemployment rate in U.S.

Scope

In this project, we are going to analyse different trends in the data to identity the suitable model. Also, the orders (p, d, q) if required are identified using the model specification tools. Then, we integrate the GARCH model with the best ordered (p, d, q) model to finalize the best model which fits the data. With this resultant model we are going to predict the forecast of unemployment rate in U.S.

Data

The data used here to predict the future unemployment rate is recorded by U.S Bureau of Labor Statistics, this time series data is released every year by the Labor Bureau since 1948 by conducting Current Population Survey (Household Survey) in United States of America. This data has two variables namely Data and Rate (measured in percentage).

Source: Unemployment Rate. (2021). Retrieved 13 June 2021, from
https://fred.stlouisfed.org/series/UNRATE

Plotting the time series data for better understanding.
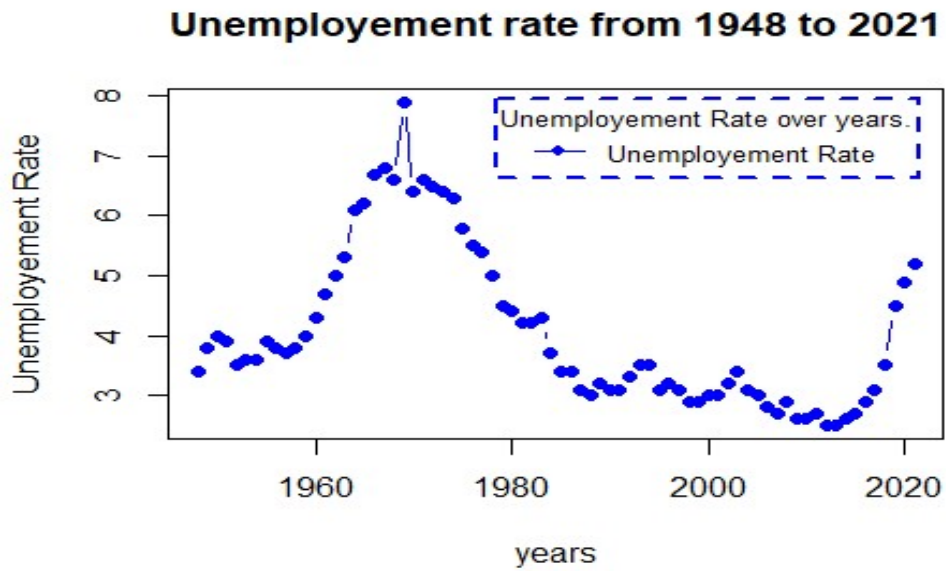
Unemployement rate from 1948 to 2021

Fig 1: Unemployment rate from 1960 to 2021 - Time series plot.

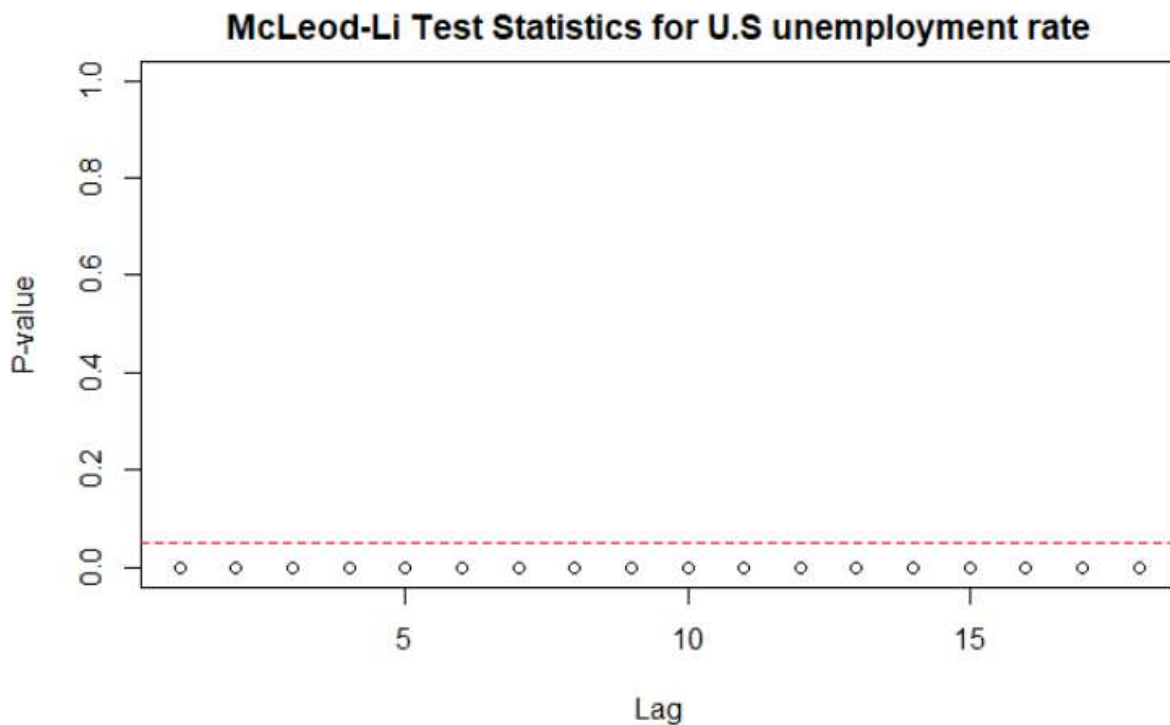## McLeod-Li Test Statistics for U.S unemployment rate

Fig 2: McLeod-Li Test Statistics for U.S unemployment rate.

McLeod-Li test is significant at 5% level of significance for all lags suggesting volatility clustering.

Descriptive analysis

1. From fig1, we can observe an upward trend in the plot until 1970 and from then we can see downward trend. By this we can say that there is no trend in the plot.
2. This series shows Autoregressive and moving average behaviour.
3. From the plot, we can conclude that there is no seasonality in the series.
4. From the time series plot and McLeod-Li test, we can see change in variance.

Model Building

Model Building Strategy:

1. Model Specification
2. Model Fitting
3. Model Diagnostics.

Model Specification:

Since the series has autoregressive and moving average behaviour, AR/MA model fits the series. Also, we choose AR/MA + GARCH model to fit the series as the series has a change in variance. Let us check whether the data is stationary or not because to fit the AR/MA model, the data should be in stationary.

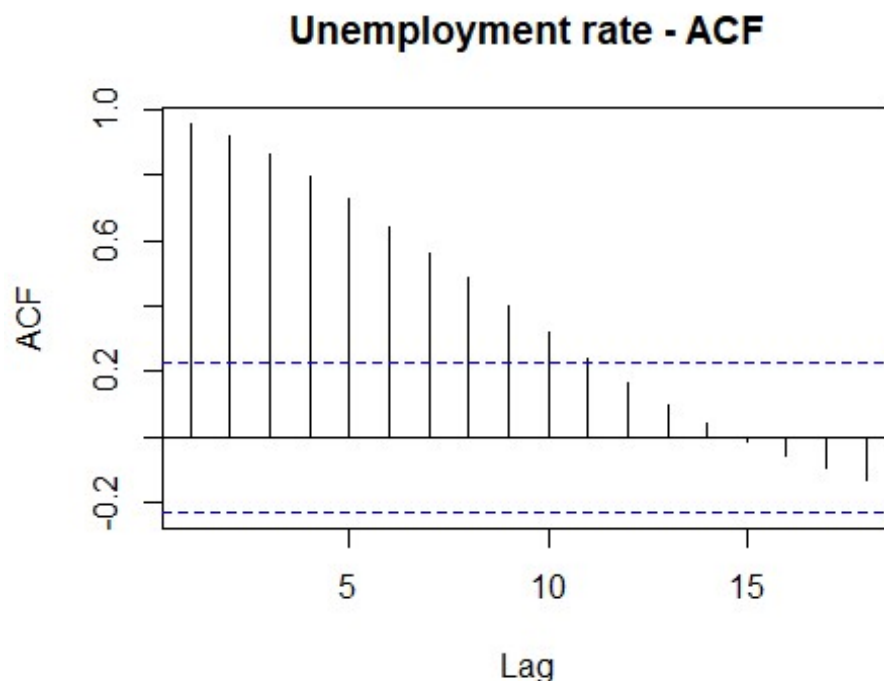**Testing for stationary**

Let us analyse ACF and PACF plots.

Fig 3: ACF plot for U.S unemployment rate.
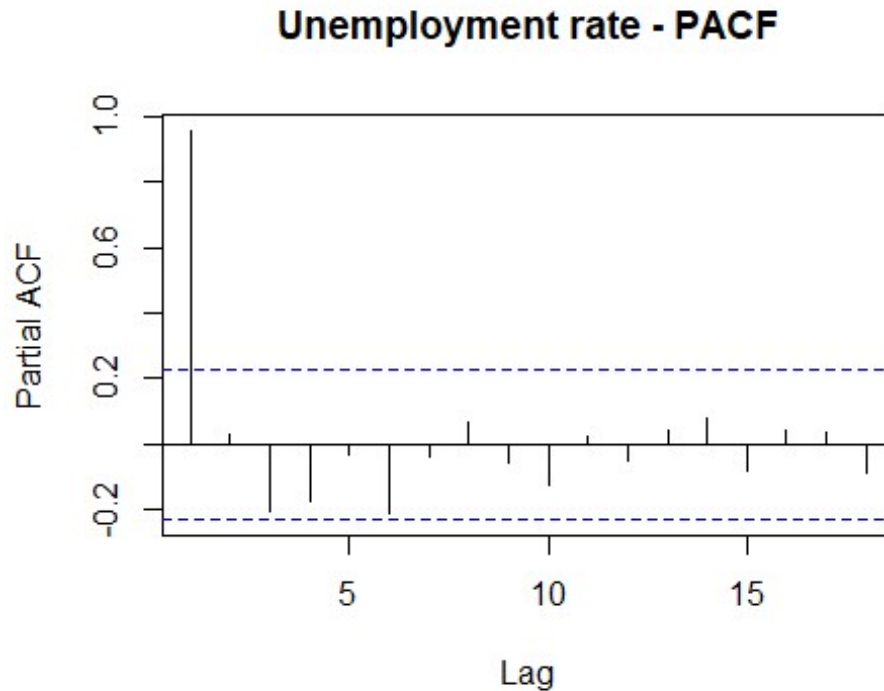
Unemployment rate - PACF

Fig 4: PACF plot for U.S unemployment rate.

In both ACF and PACF, we can observe high peak at the first lag and significant lags are in decomposition pattern in ACF plot.

Running Augmented Dickey-Fuller tests on series data.

```
##
## Title:
##  Augmented Dickey-Fuller Test
##
## Test Results:
##    PARAMETER:
##      Lag Order: 1
##    STATISTIC:
##      Dickey-Fuller: 0.1193
##    P VALUE:
##      0.65
##
## Description:
##  Sun Jun 13 20:11:46 2021 by user: HP
```

Hypotheses:

H0: The data is not stationary.

HA: The data is stationary.

Interpretations:

p - value - 0.65 > 0.05

Since, p - value is greater than 0.05, the test is not statistically significant. Thereby, we fail to reject Null hypothesis i.e., The data is not stationary.

Let's apply transformation on the data to check whether the change in variance is decreasing or not.
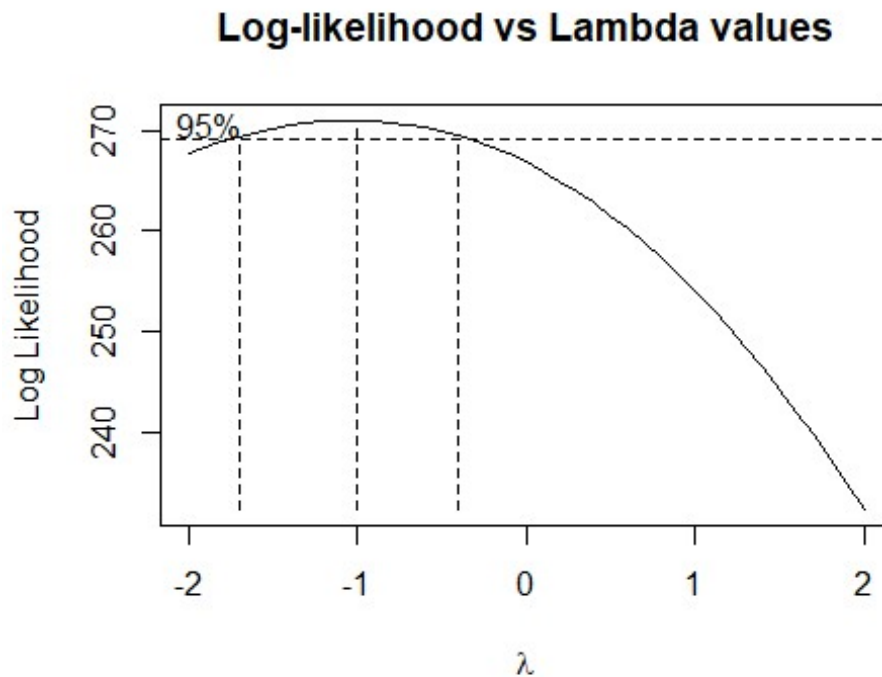
Applying Box_Cox transformation,

Fig 5: Log-likelihood vs Lambda values

Change in variance is decreased. Now, let us calculate the best lambda.

Running shapiro test to check the normality,

```
##
##  Shapiro-Wilk normality test
```

```
##
## data:  UNRATE_TS_BC
## W = 0.96005, p-value = 0.01964
```

p-value = 0.02 < 0.05

Since, p - value is less than 0.05, the test is statistically significant at 95% confidence intervals. Thereby, Null hypothesis can be rejected i.e., data is normalized.

Running ADF test,

```
##
## Title:
##  Augmented Dickey-Fuller Test
##
## Test Results:
##   PARAMETER:
##     Lag Order: 1
##   STATISTIC:
##     Dickey-Fuller: 0.1874
##   P VALUE:
##     0.6716
##
## Description:
##  Sun Jun 13 20:11:47 2021 by user: HP
```

Hypotheses:

  H0: The data is not stationary.

  HA: The data is stationary.

Interpretations:

p - value - 0.67 > 0.05

Since, p - value is greater than 0.05, the test is not statistically significant. Thereby, we fail to reject Null hypothesis i.e., The data is not stationary.

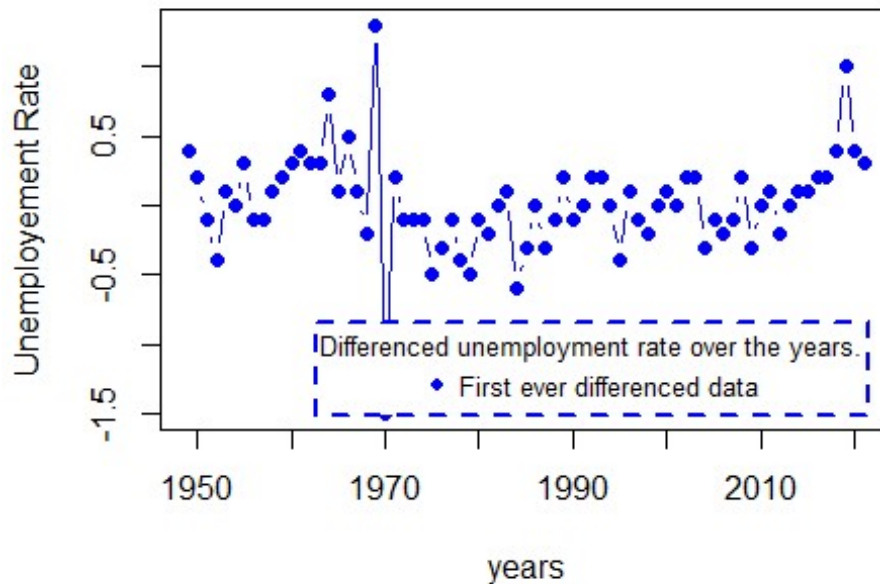Since the data is not stationary, we are differencing the data to order 1 i.e., d=1.

Fig 6: Differenced series of unemployment rate.

Applying ADF test,

```
## Warning in adfTest(UNRATE_diff): p-value smaller than printed p-value

##
## Title:
##  Augmented Dickey-Fuller Test
##
## Test Results:
##   PARAMETER:
##     Lag Order: 1
##   STATISTIC:
##     Dickey-Fuller: -4.3833
##   P VALUE:
##     0.01
##
## Description:
##  Sun Jun 13 20:11:47 2021 by user: HP
```

Hypotheses:

H0: The data is not stationary.

HA: The data is stationary.

Interpretations:

p - value ~ 0.01 < 0.5

P - value rounded to 0.01 as it is very small (exponential value.)

Since, p - value is less than 0.05, the test is statistically significant at 95% confidence intervals. Thereby, Null hypothesis can be rejected i.e., data is stationary.
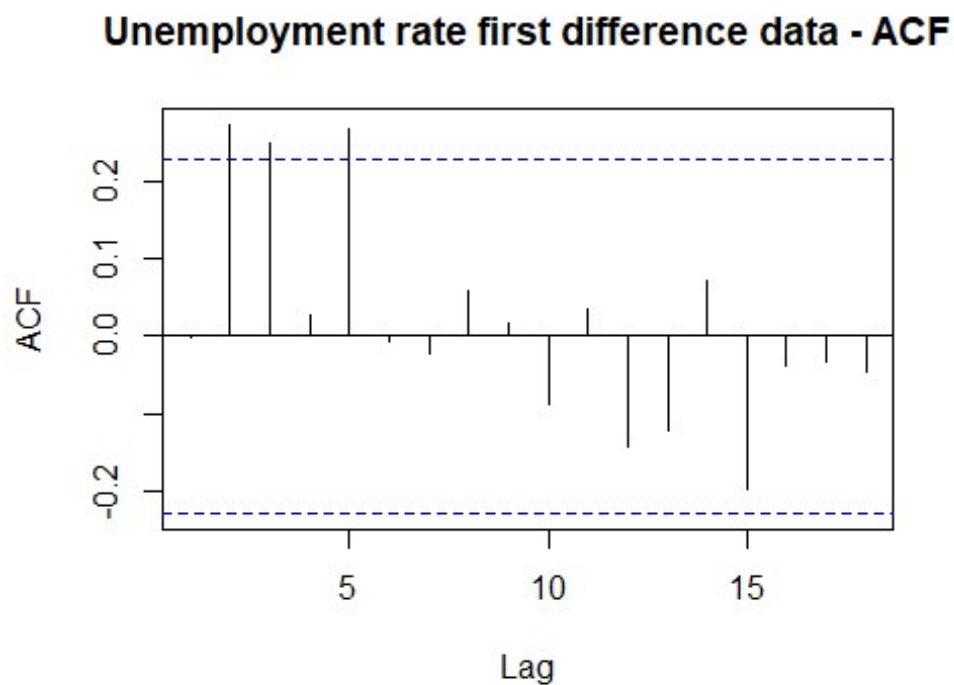
Order Specification

Using ACF and PACF

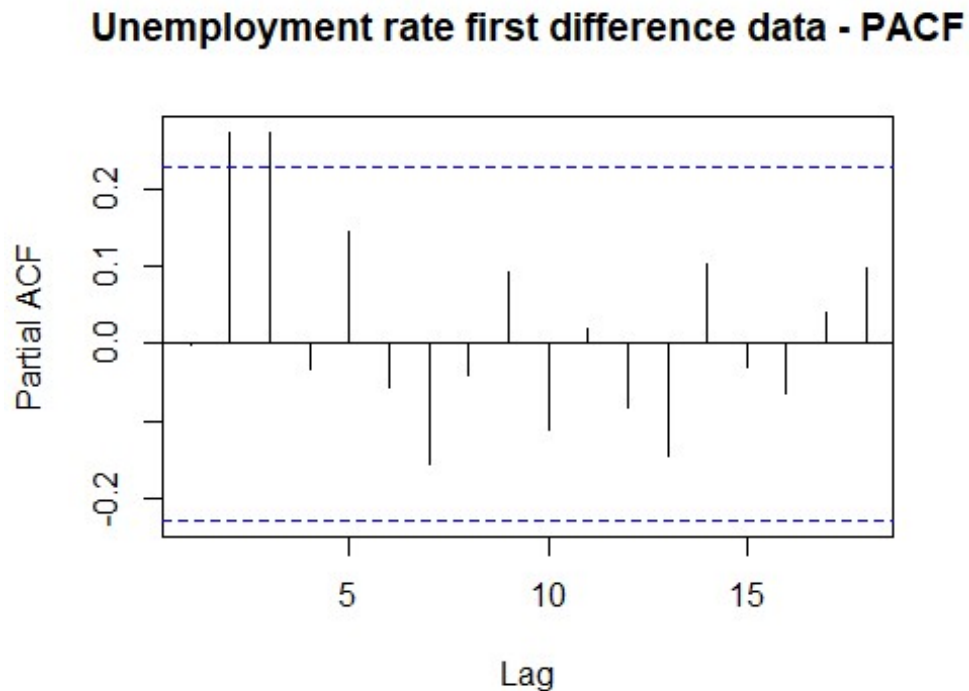Fig 7: Unemployment rate first difference data - ACF

Fig 8: Unemployment rate first difference data - PACF

The possible orders from ACF and PACF are, (0,1), (0,2), (0,3), (1, 0), (1,1), (1,2), (1,3), (2, 0), (2,1), (2,2), (2,3)

Using EACF

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

Possible orders from EACF, (3,2), (3,3), (3,4).

Using BIC

```
## Warning in leaps.setup(x, y, wt = wt, nbest = nbest, nvmax = nvmax, force.
in =
## force.in, : 5 linear dependencies found

## Reordering variables and trying again:
```

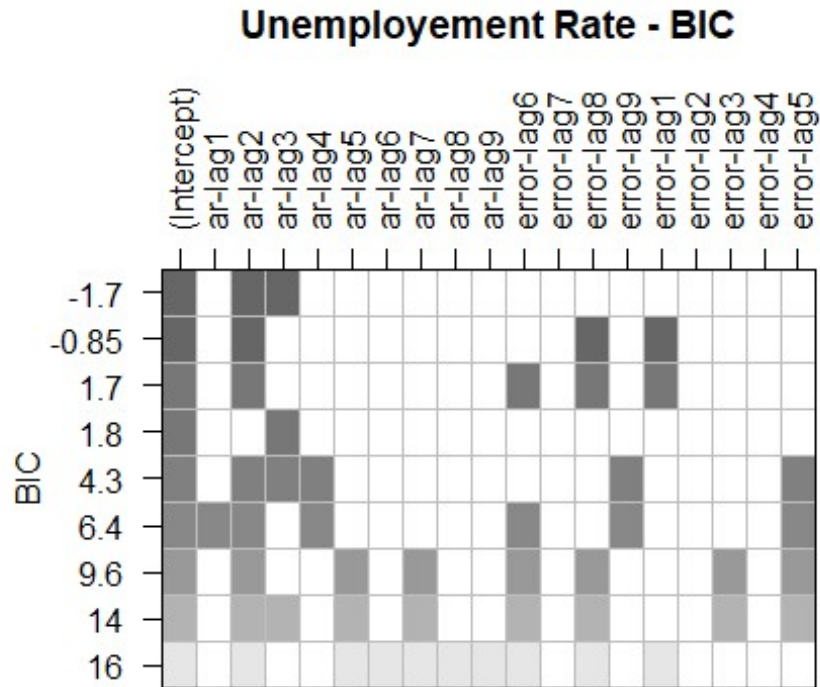## Unemployement Rate - BIC

Fig 9: Unemployment rate first difference data - BIC

Possible orders from BIC, (3, 0), (2, 1).

Overall possible ARMA orders are,

(0,1), (0,2), (0,3), (1, 0), (1,1), (1,2), (1,3), (2, 0), (2,1), (2,2), (2,3), (3,2), (3,3), (3,4) and (3, 0).

Model Fitting and Diagnostics

ARIMA(0,0,1)

```
##
## z test of coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## ma1        -0.0012071  0.0942890 -0.0128   0.9898
## intercept   0.0246484  0.0422593  0.5833   0.5597
```
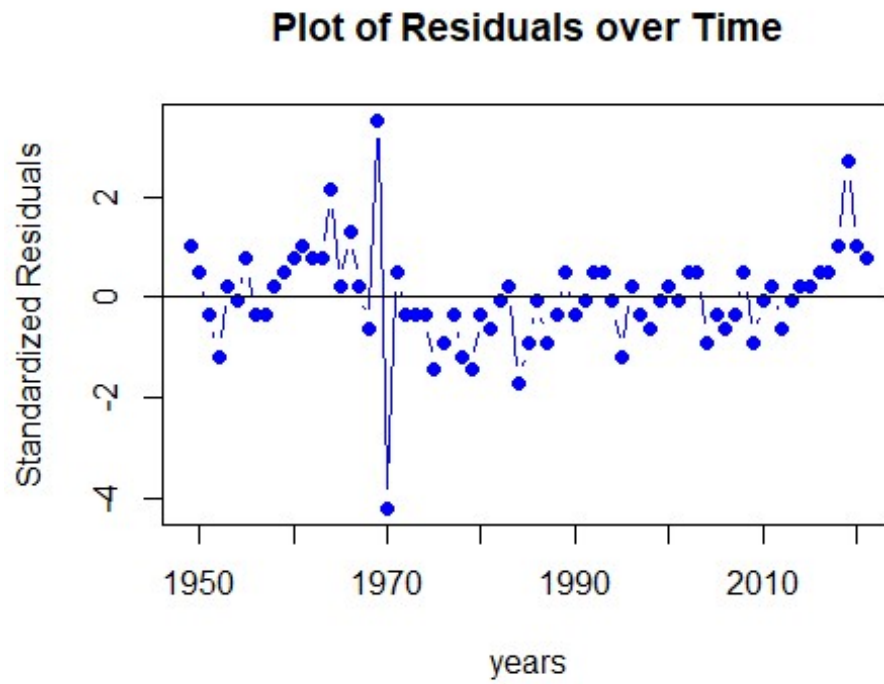
Plot of Residuals over Time

Fig 10: Residual's plot over time for ARIMA(0,0,1)
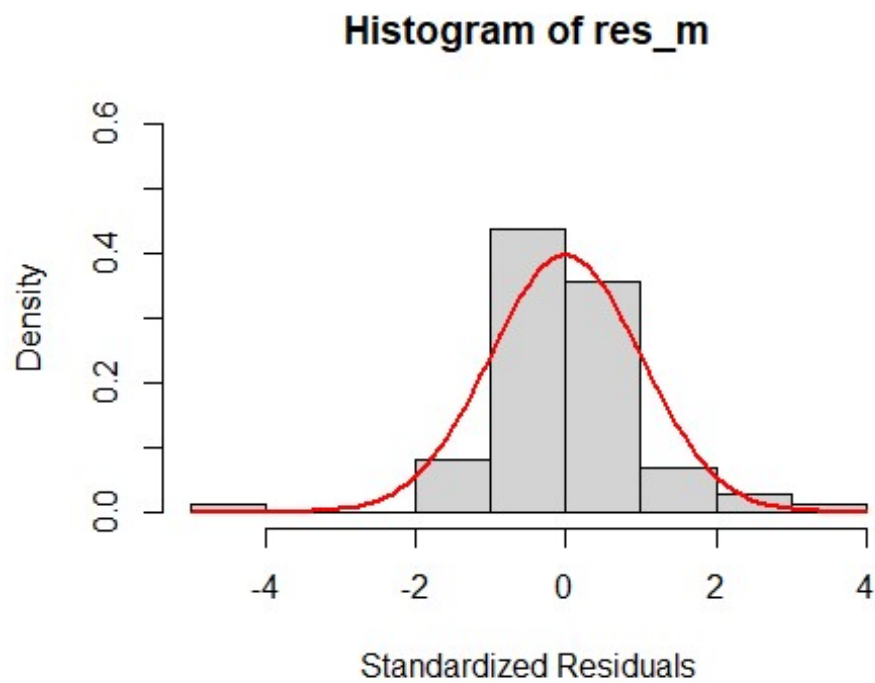
Histogram of res_m

Fig 11: Histogram of ARIMA(0,0,1)
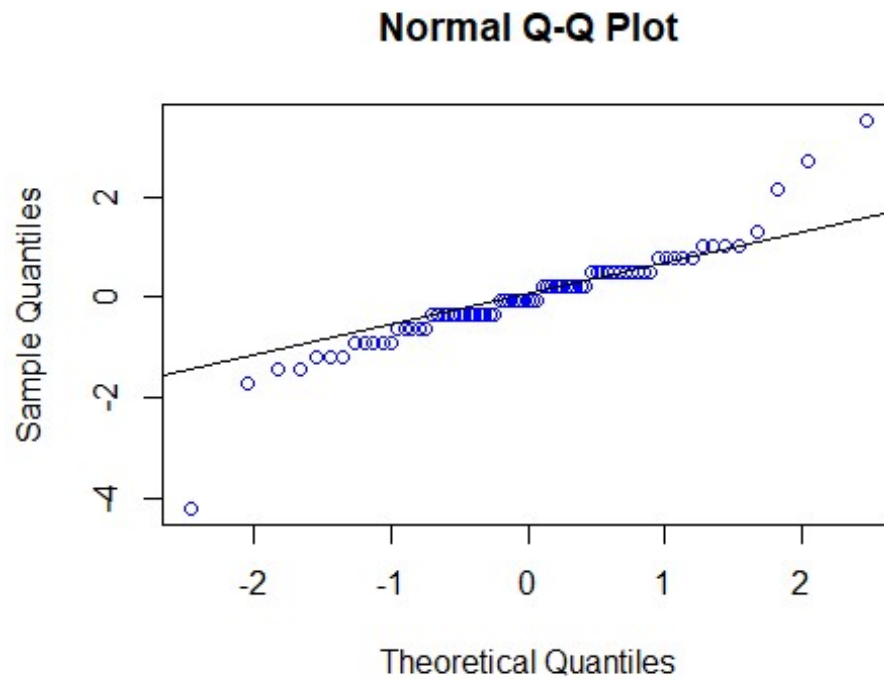
Normal Q-Q Plot

Fig 12: Normal Q-Q plot for ARIMA(0,0,1)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.90137, p-value = 3.091e-05

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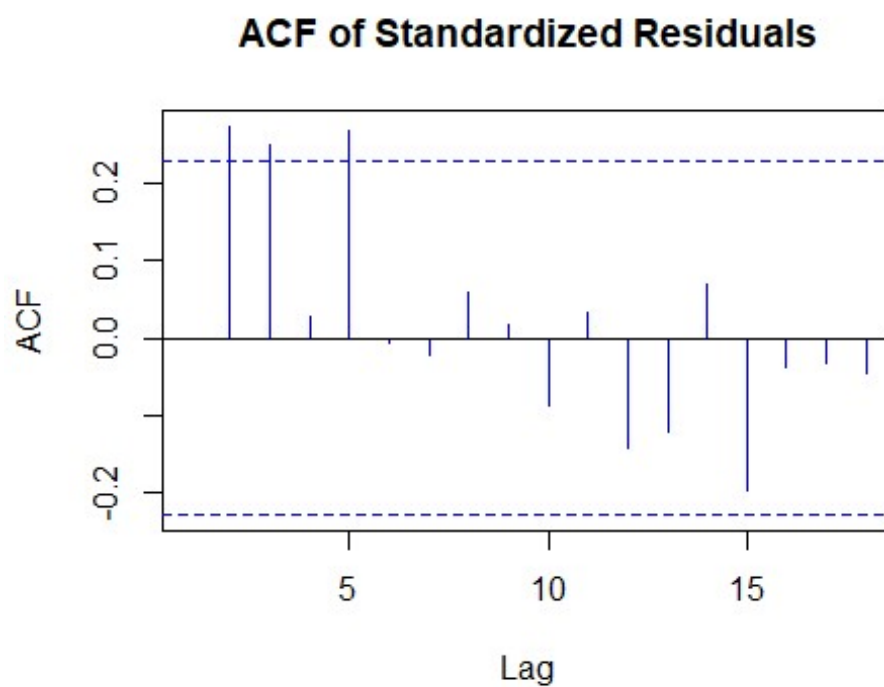
ACF of Standardized Residuals

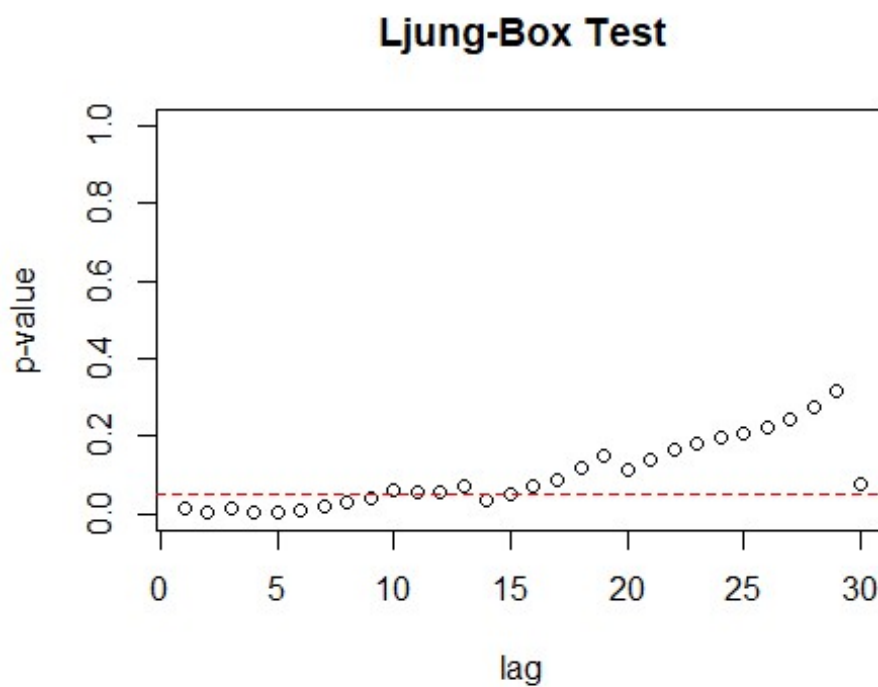Fig 13: Autocorrelation plot for ARIMA(0,0,1)

Ljung-Box Test

Fig 14: Ljung-Box test for ARIMA(0,0,1)

Residual Analysis for ARIMA(0,0,1):

1. p-value is greater than 0.05 for ma1.
2. In Residual plot, the line is passing through 0.
3. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
4. There are significant lags in Autocorrelation plot.
5. From Ljung box test, we can observe few significant lags.

Therefore, it is not a good model for the data.

ARIMA (0,0,2)

```
##
## z test of coefficients:
##
##            Estimate Std. Error z value Pr(>|z|)
## ma1       -0.079466   0.119146 -0.6670 0.504795
## ma2        0.284368   0.101772  2.7942 0.005203 **
## intercept  0.028440   0.048514  0.5862 0.557716
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
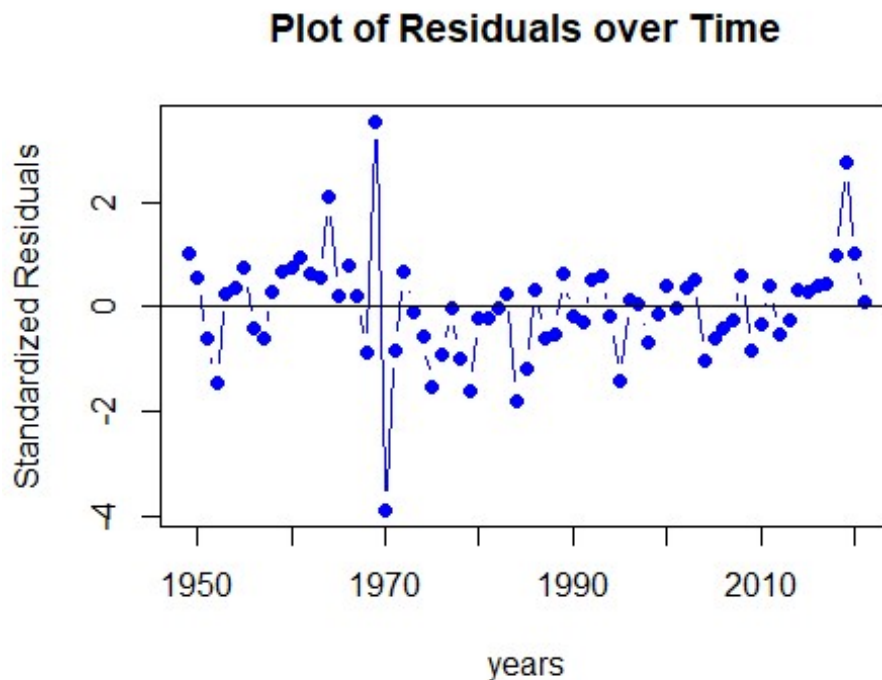

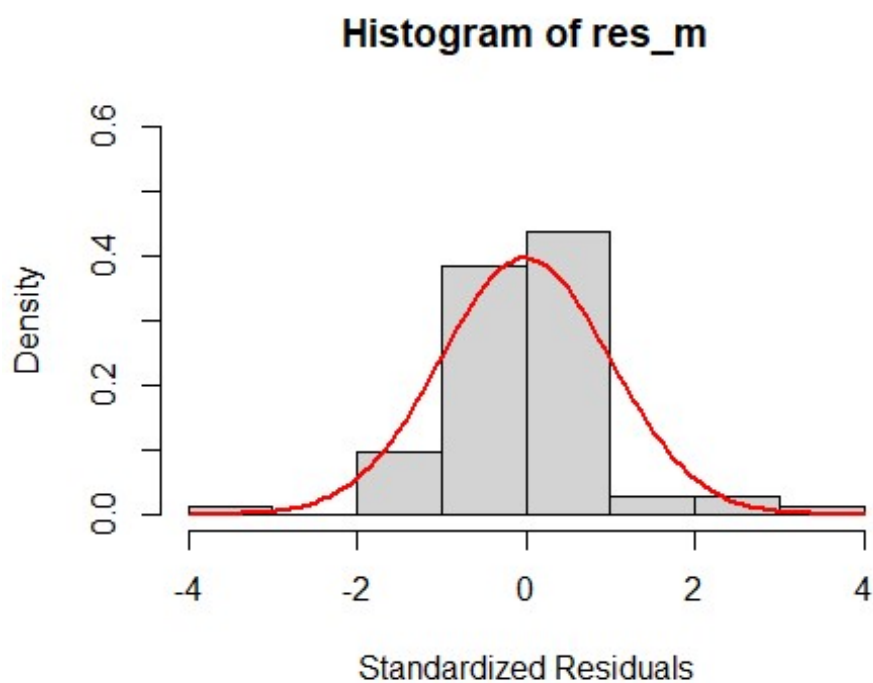Fig 15: Residual's plot over time for ARIMA (0,0,2)

Fig 16: Histogram of ARIMA (0,0,2)
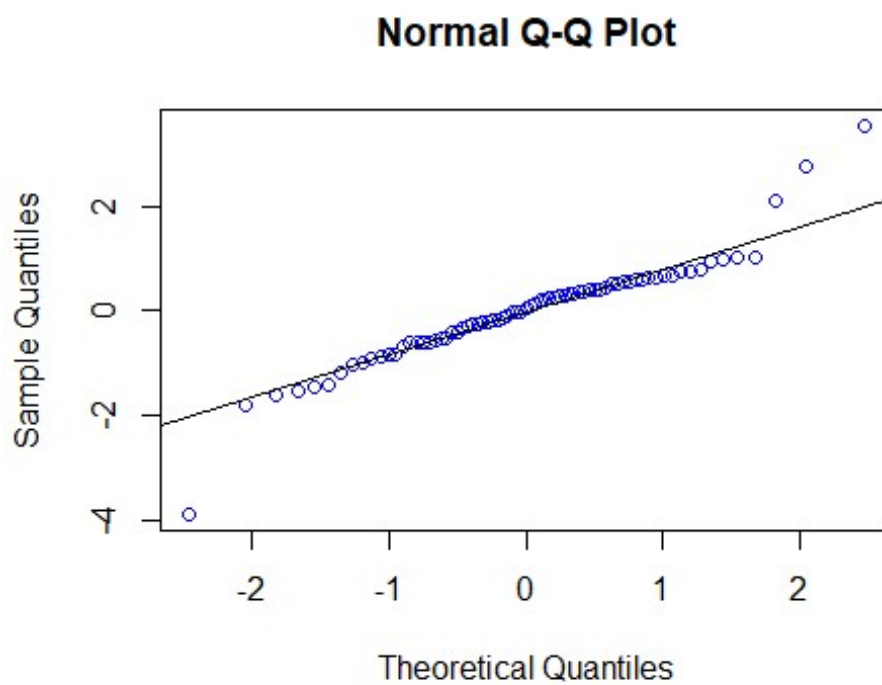

Fig 17: Normal Q-Q plot for ARIMA (0,0,2)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.91889, p-value = 0.0001722

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

ACF of Standardized Residuals

Fig 18: Autocorrelation plot for ARIMA (0,0,2)

Ljung-Box Test

Fig 19: Ljung-Box test for ARIMA (0,0,2)

Residual Analysis for ARIMA(0,0,2):

1. p-value is less than 0.05 for ma2. Thereby, it is significant.
2. In Residual plot, the line is passing through 0.
3. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
4. There is only one significant lag in Autocorrelation plot.
5. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it can be considered as a good model for the data.

ARIMA(0,0,3)

```
##
## z test of coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## ma1         -0.135365   0.131755 -1.0274  0.30423
## ma2          0.158017   0.108373  1.4581  0.14482
## ma3          0.341004   0.170296  2.0024  0.04524 *
## intercept    0.035071   0.053313  0.6578  0.51065
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
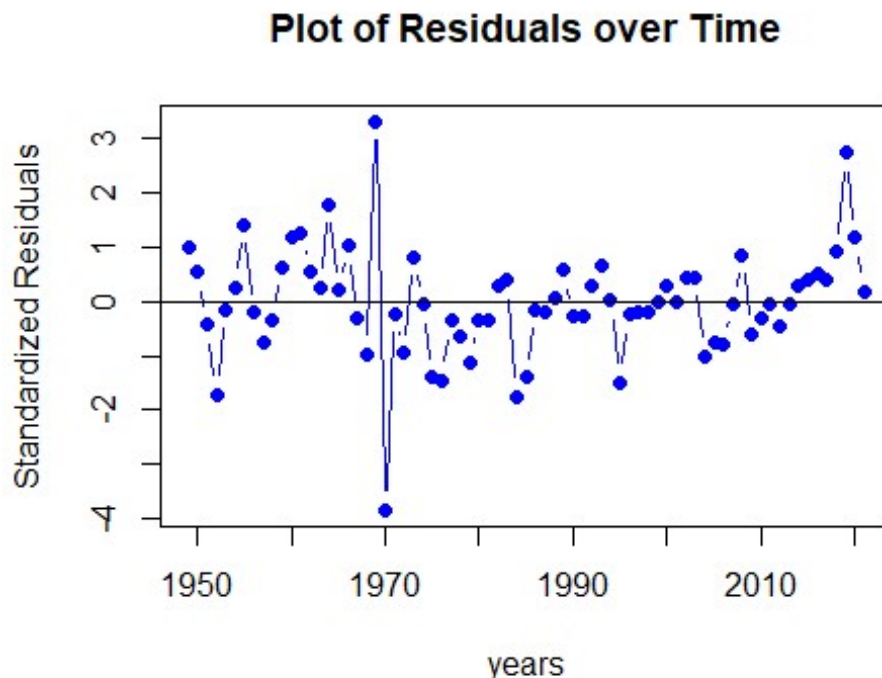

Fig 20: Residual's plot over time for ARIMA(0,0,3)
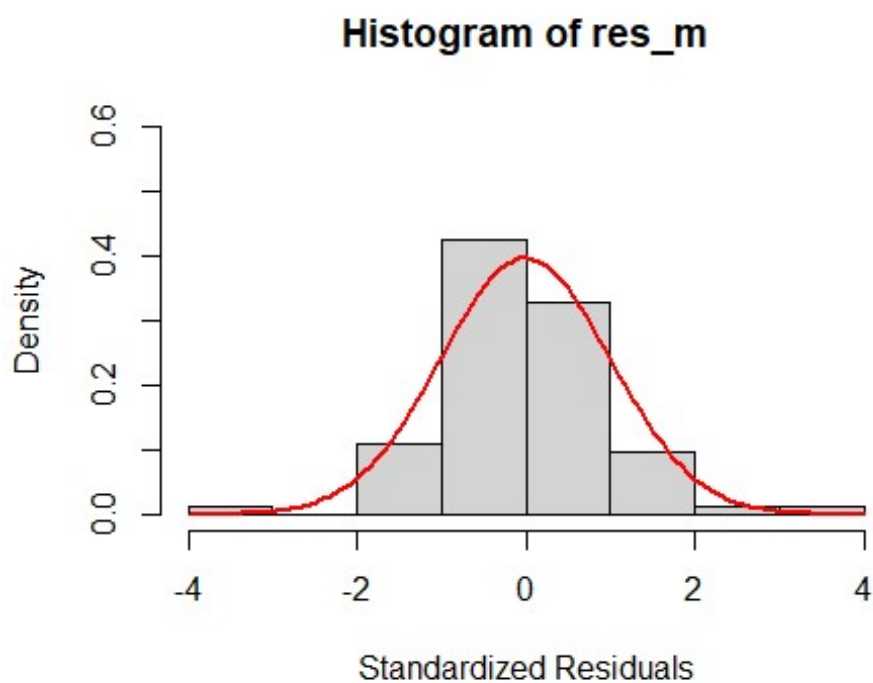
Histogram of res_m

Fig 21: Histogram of ARIMA(0,0,3)
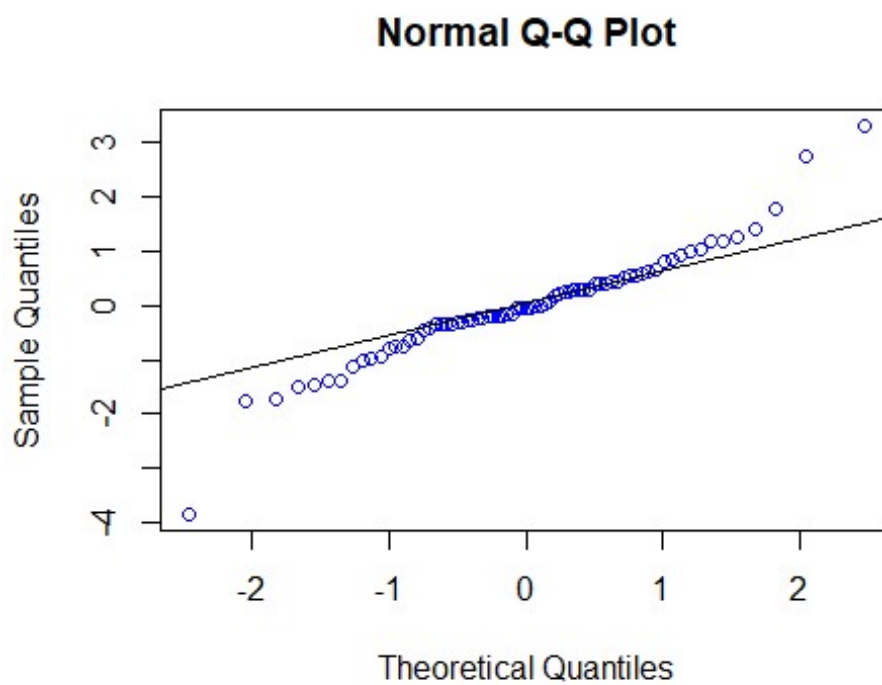
Normal Q-Q Plot

Fig 22: Normal Q-Q plot for ARIMA(0,0,3)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.93929, p-value = 0.001601

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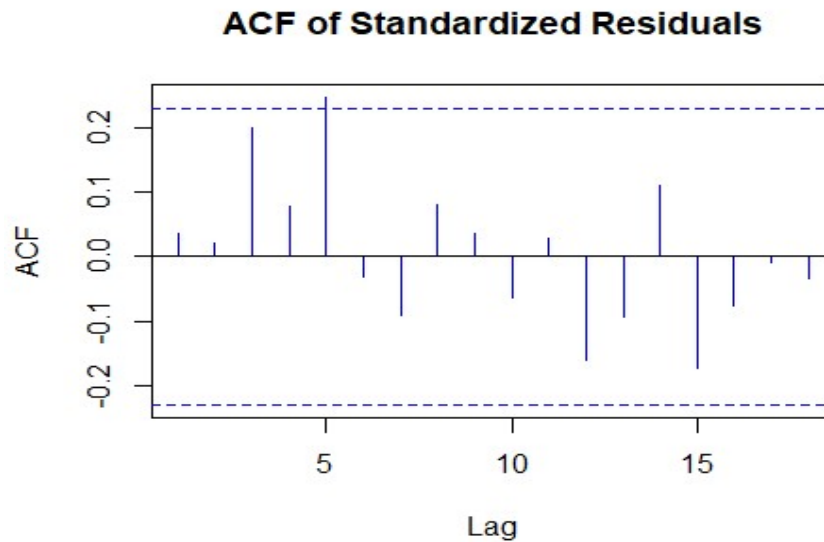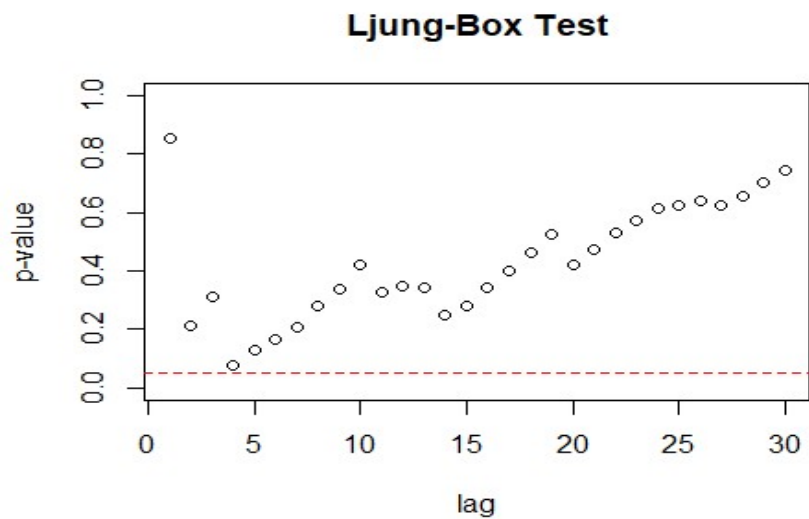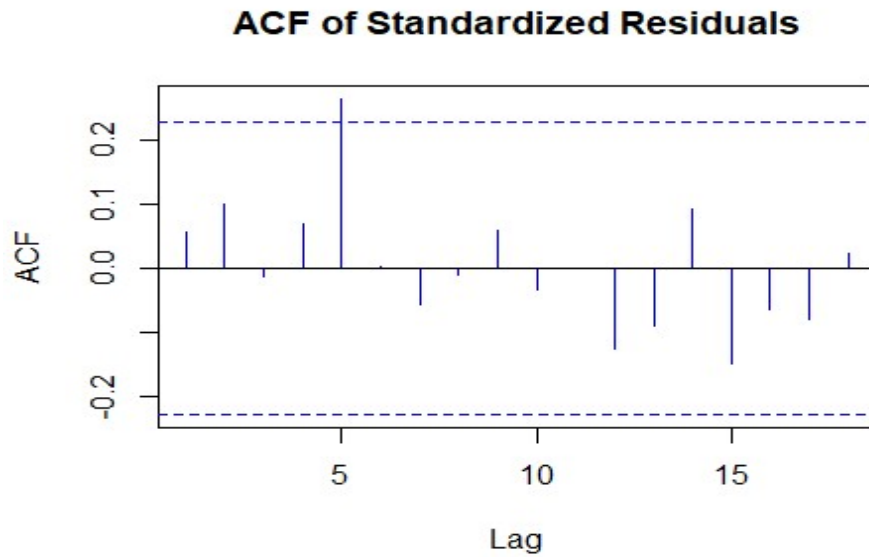

Fig 23: Autocorrelation plot for ARIMA(0,0,3)
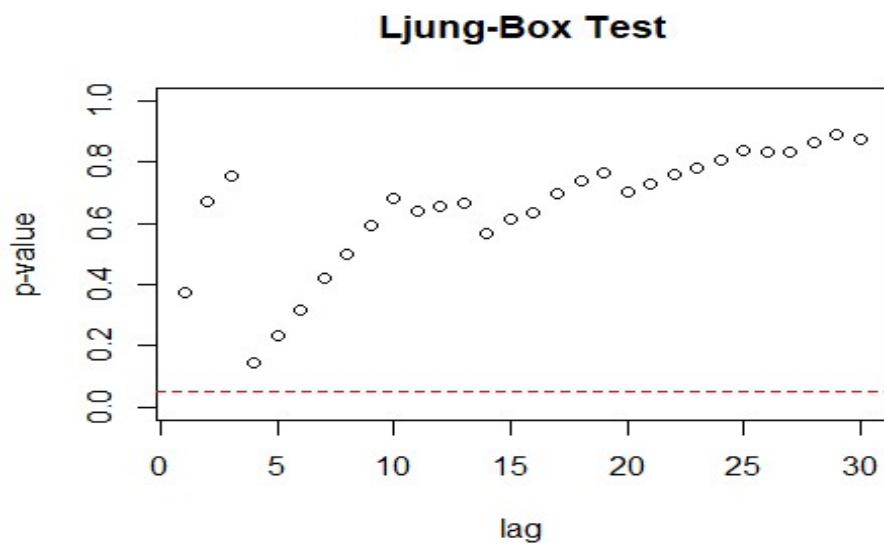
Fig 24: Ljung-Box test for ARIMA(0,0,3)

Residual Analysis for ARIMA(0,0,3):

1. p-value is less than 0.05 for ma3. Thereby, it is significant.
2. In Residual plot, the line is passing through 0.
3. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
4. There is only one significant lag in Autocorrelation plot.
5. Ljung box test shows good results as all the dots are above the dotted line.

Overall, it can be considered as a good model but comparatively, ARIMA(0,0,2) performed better as the coefficients of ARIMA(0,0,3) are less significant than the coefficients of ARIMA(0,0,2).

ARIMA (1,0,0)

```
## 
## z test of coefficients:
## 
##             Estimate Std. Error z value Pr(>|z|)
## ar1        -0.001878   0.117609 -0.0160   0.9873
## intercept   0.024634   0.042236  0.5833   0.5597
```


Fig 25: Residual's plot over time for ARIMA(1,0,0)
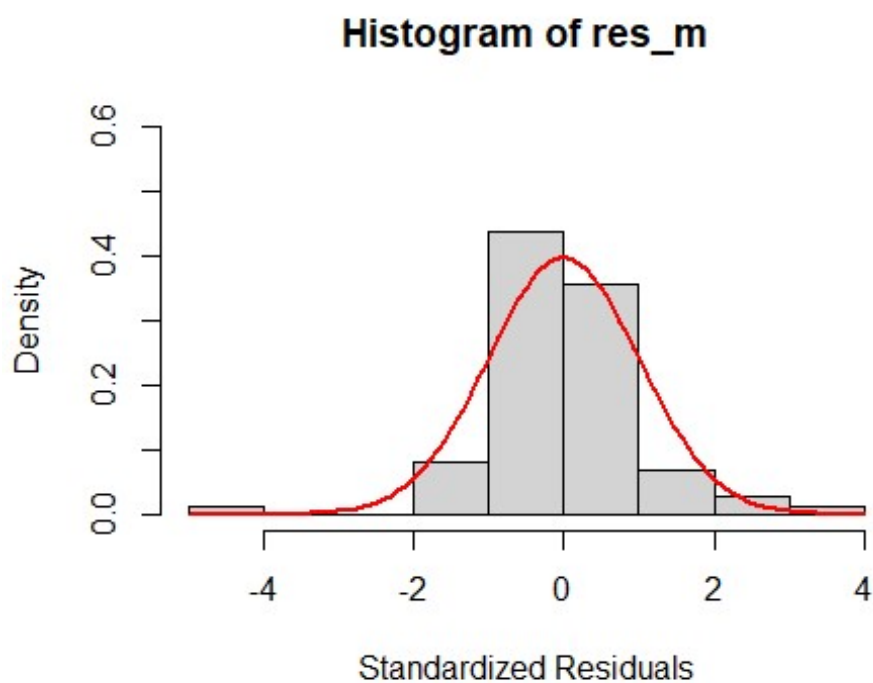
Histogram of res_m

Fig 26: Histogram of ARIMA(1,0,0)
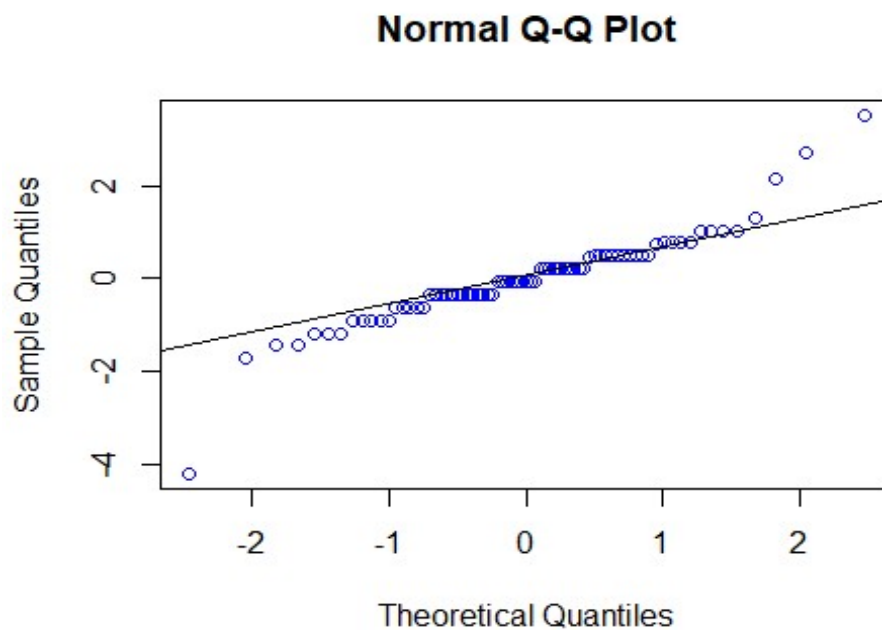
Normal Q-Q Plot

Fig 27: Normal Q-Q plot for ARIMA(1,0,0)

```
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.90161, p-value = 3.162e-05

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

ACF of Standardized Residuals

Fig 28: Autocorrelation plot for ARIMA(1,0,0)

Ljung-Box Test

Fig 29: Ljung-Box test for ARIMA(1,0,0)

Residual Analysis for ARIMA(1,0,0):

1. p-value is greater than 0.05 for ma1.
2. In Residual plot, the line is passing through 0.
3. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
4. There are significant lags in Autocorrelation plot.
5. From Ljung box test, we can observe few significant lags.

Therefore, it is not a good model for the data.

ARIMA (1,0,1)

```
##
## z test of coefficients:
##
##             Estimate Std. Error z value Pr(>|z|)
## ar1       -0.436057   0.502016 -0.8686   0.3851
## ma1        0.382523   0.504920  0.7576   0.4487
## intercept  0.024364   0.040670  0.5991   0.5491
```
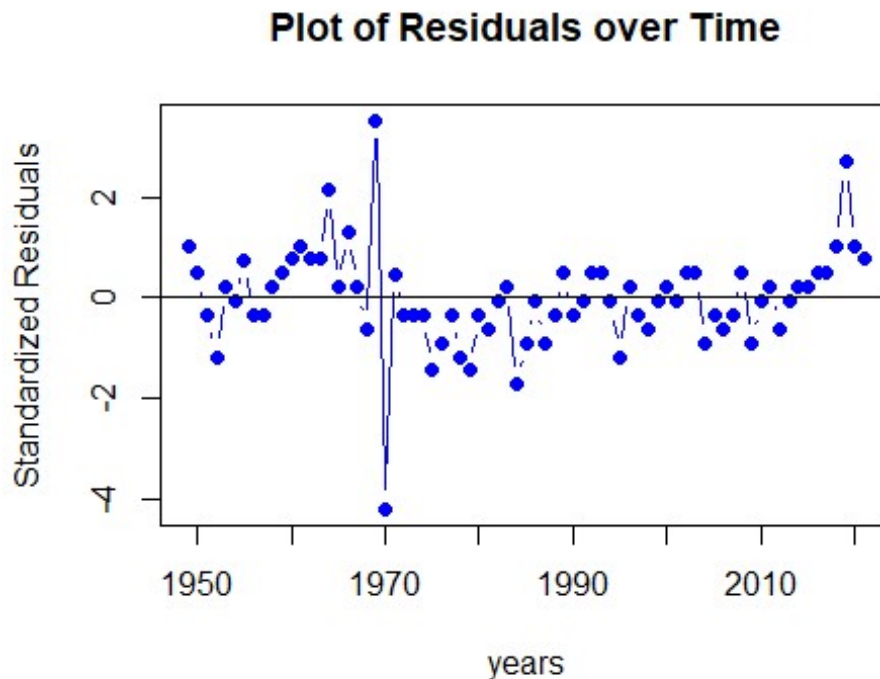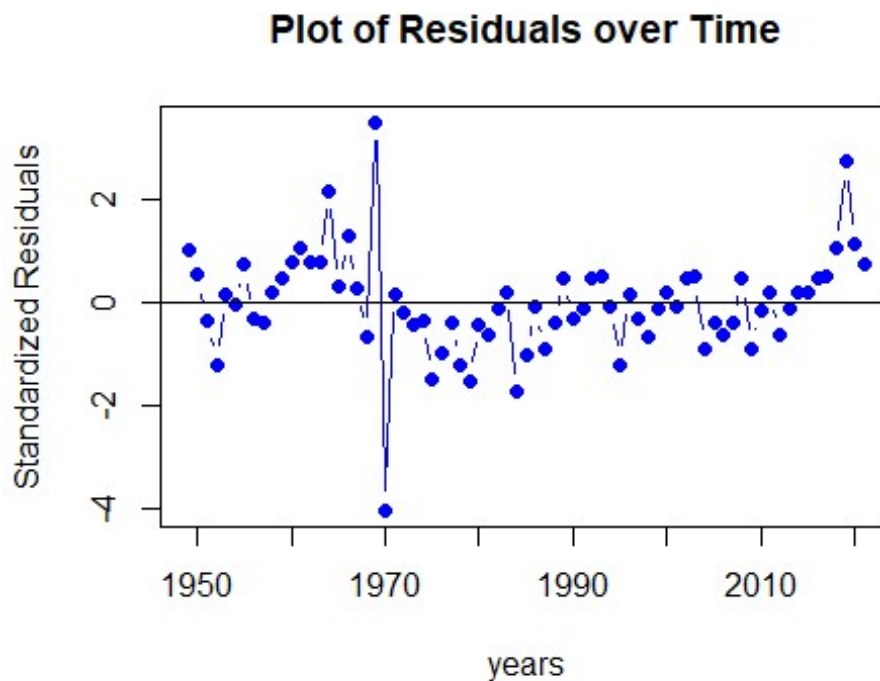

Fig 30: Residual's plot over time for ARIMA(1,0,1)

Fig 31: Histogram of ARIMA(1,0,1)

Fig 32: Normal Q-Q plot for ARIMA(1,0,1)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.91706, p-value = 0.0001428

## Warning in (ra^2)/(n - (1:lag.max)): longer object length is not a multipl
e of
## shorter object length
```
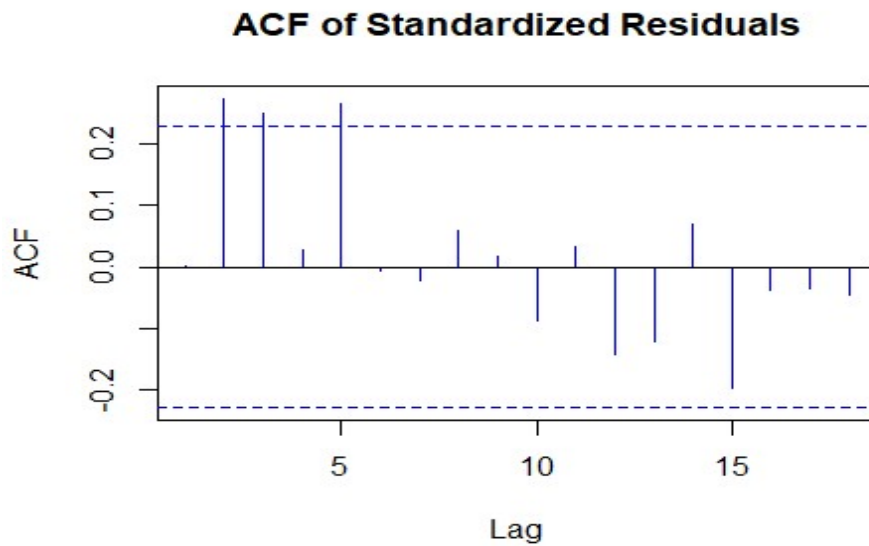
ACF of Standardized Residuals
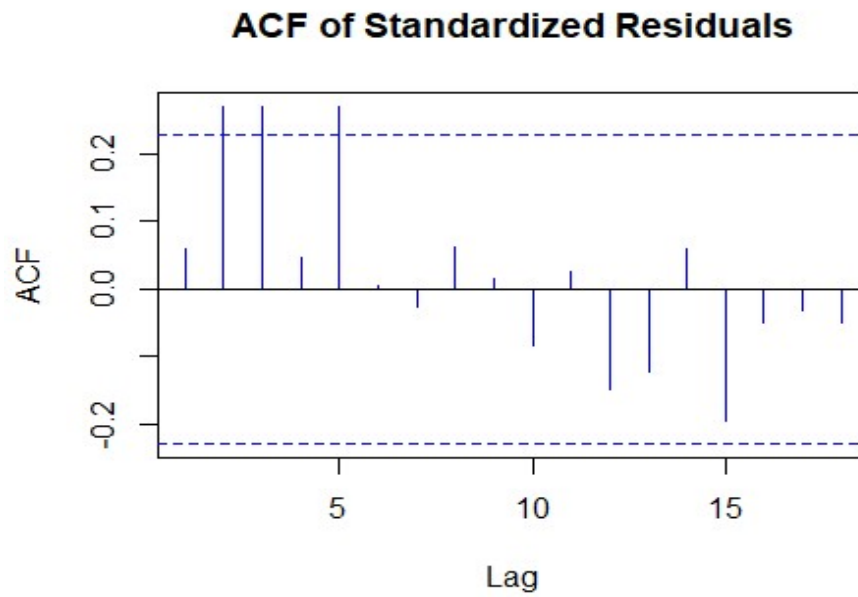
Fig 33: Autocorrelation plot for ARIMA(1,0,1)
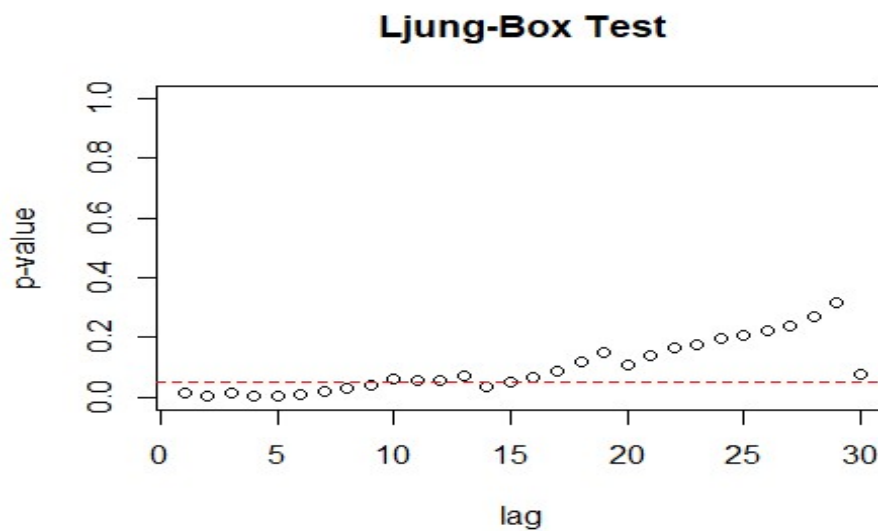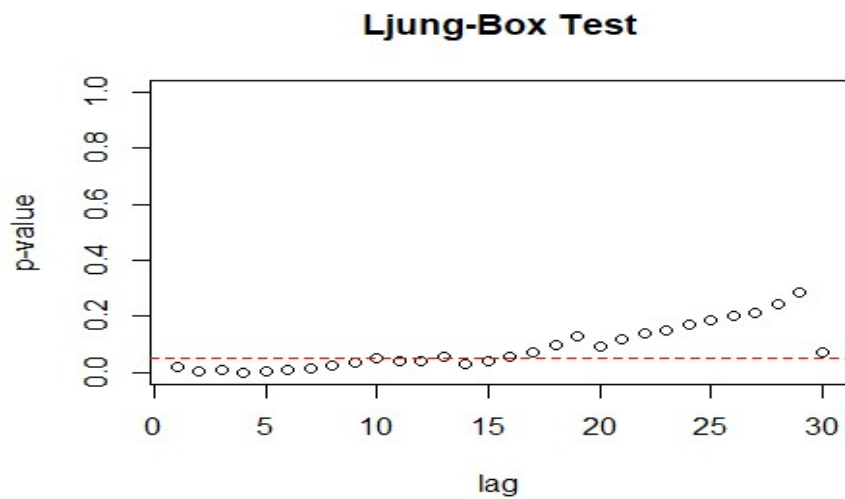
Ljung-Box Test

Fig 34: Ljung-Box test for ARIMA(1,0,1)

Residual Analysis for ARIMA(1,0,1):

1. p-value is greater than 0.05 for ar1 and ma1.
2. In Residual plot, the line is passing through 0.
3. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
4. There are significant lags in Autocorrelation plot.
5. From Ljung box test, we can observe few significant lags.

Therefore, it is not a good model for the data.

ARIMA (1,0,2)

```
##
## z test of coefficients:
##
##            Estimate Std. Error z value  Pr(>|z|)
## ar1        0.798396   0.129194  6.1798 6.418e-10 ***
## ma1       -0.946113   0.152809 -6.1915 5.960e-10 ***
## ma2        0.422548   0.127995  3.3013 0.0009624 ***
## intercept  0.056250   0.088843  0.6331 0.5266427
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Plot of Residuals over Time
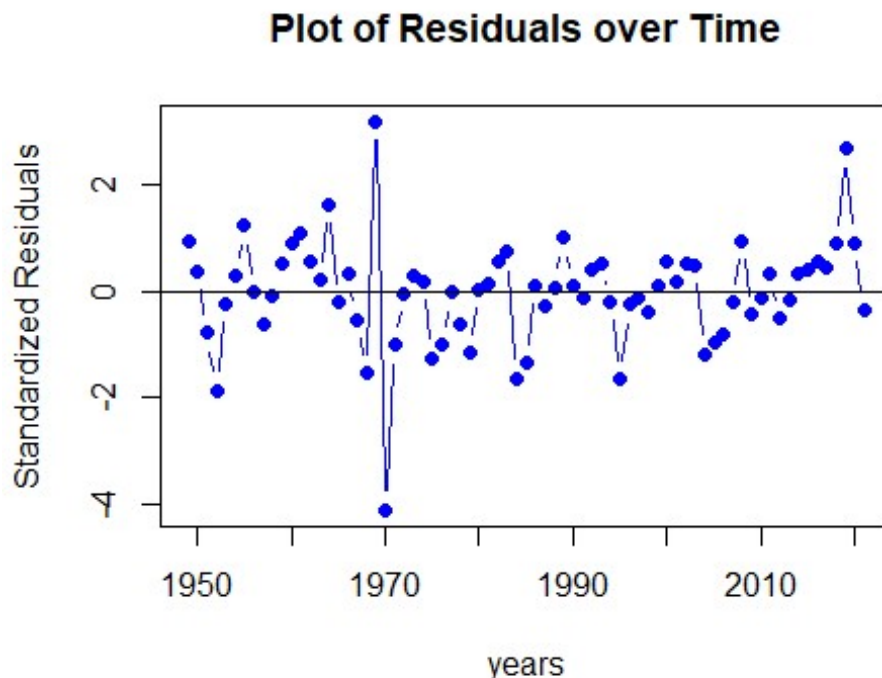
Fig 35: Residual's plot over time for ARIMA(1,0,2)

27

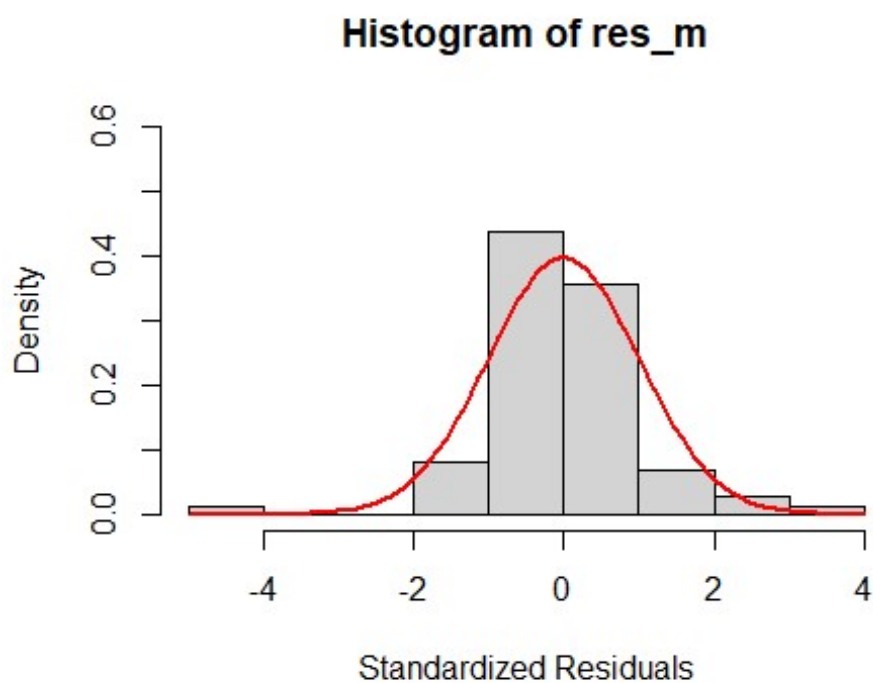## Histogram of res_m

Fig 36: Histogram of ARIMA(1,0,2)
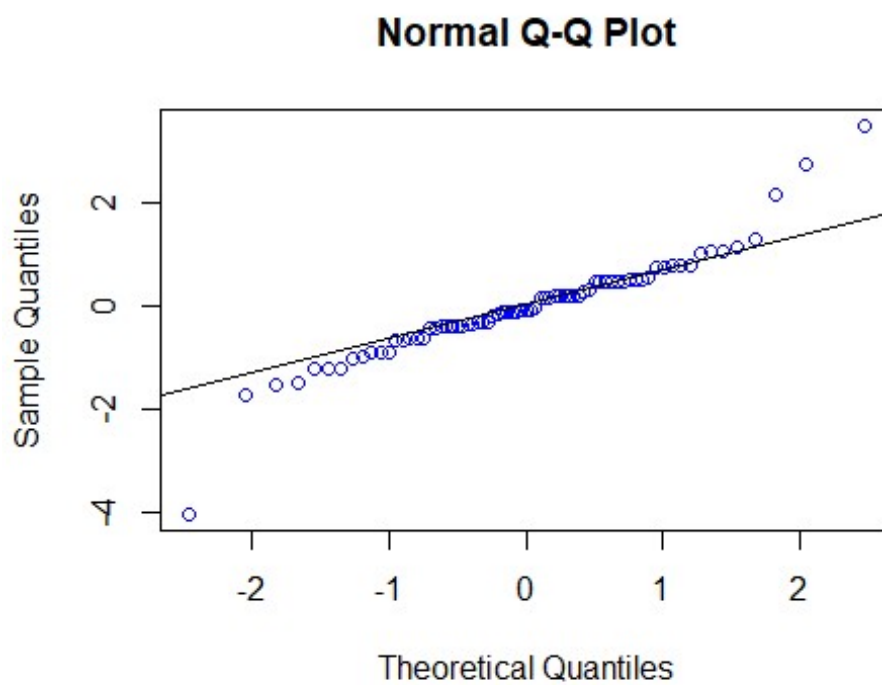
Normal Q-Q Plot

Fig 37: Normal Q-Q plot for ARIMA(1,0,2)

```
## 
##  Shapiro-Wilk normality test
## 
## data:  res_m
## W = 0.92758, p-value = 0.0004307

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

ACF of Standardized Residuals
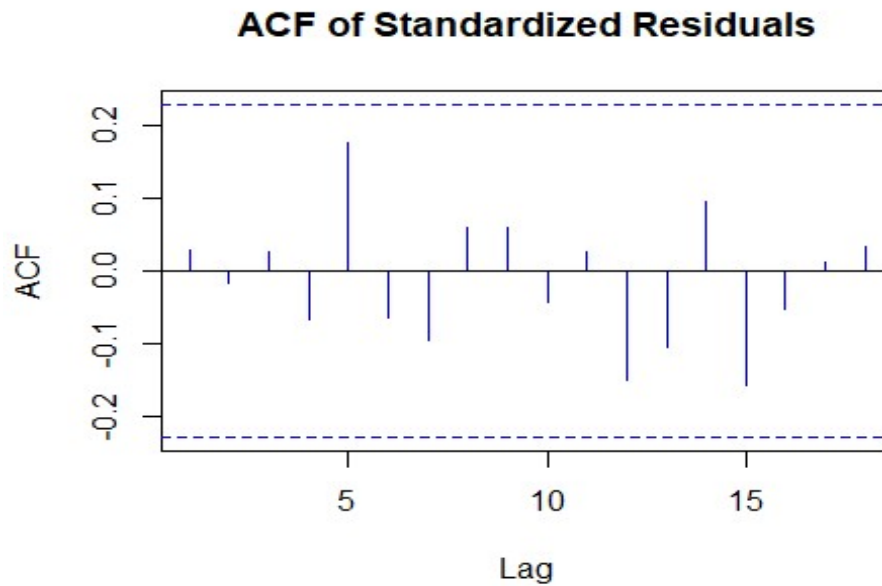
Fig 38: Autocorrelation plot for ARIMA(1,0,2)
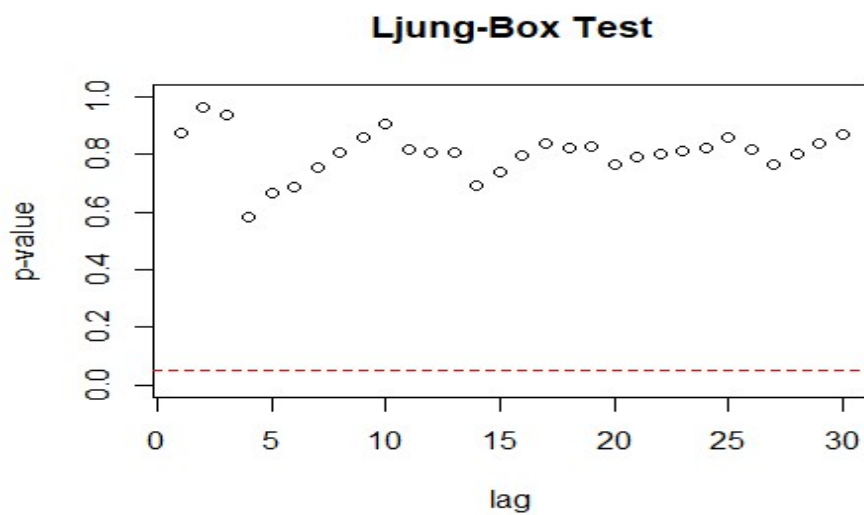
Ljung-Box Test

Fig 39: Ljung-Box test for ARIMA(1,0,2)

Residual Analysis for ARIMA(1,0,2):

1. p-value is less than 0.05 for ar1, ma1, and ma2. Hence, significant.
2. In Residual plot, the line is passing through 0.
3. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
4. There are no significant lags in Autocorrelation plot.
5. Ljung box test shows good results as all the dots are above the dotted line.

Overall, it can be considered as a best model.

ARIMA (1,0,3)

```
##
## z test of coefficients:
##
##            Estimate Std. Error z value Pr(>|z|)
## ar1        0.716644   0.273053  2.6246 0.008676 **
## ma1       -0.846430   0.293113 -2.8877 0.003880 **
## ma2        0.360540   0.169874  2.1224 0.033804 *
## ma3        0.096113   0.211756  0.4539 0.649912
## intercept  0.050341   0.081955  0.6143 0.539046
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
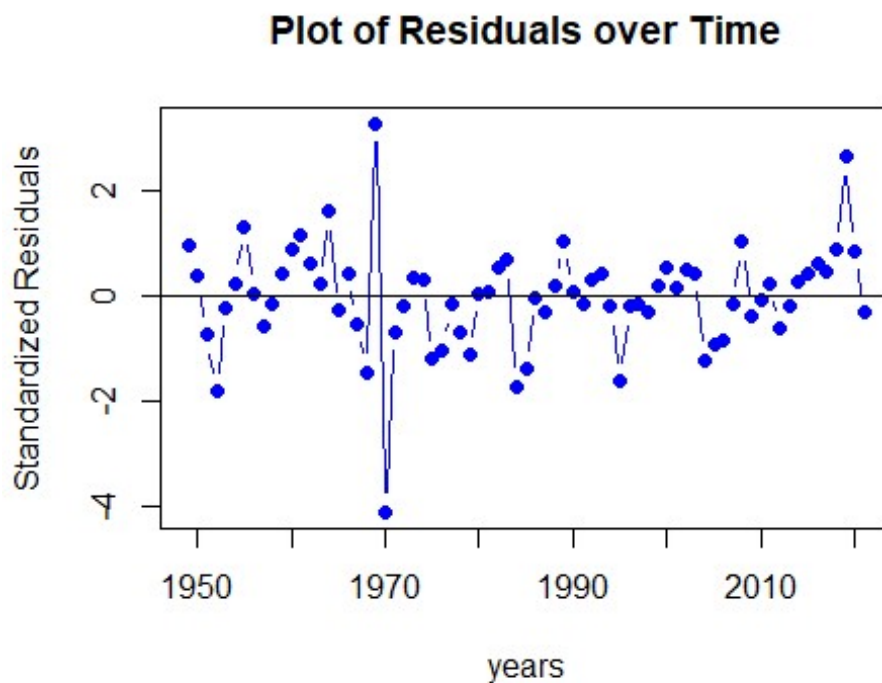

Fig 40: Residual's plot over time for ARIMA(1,0,3)

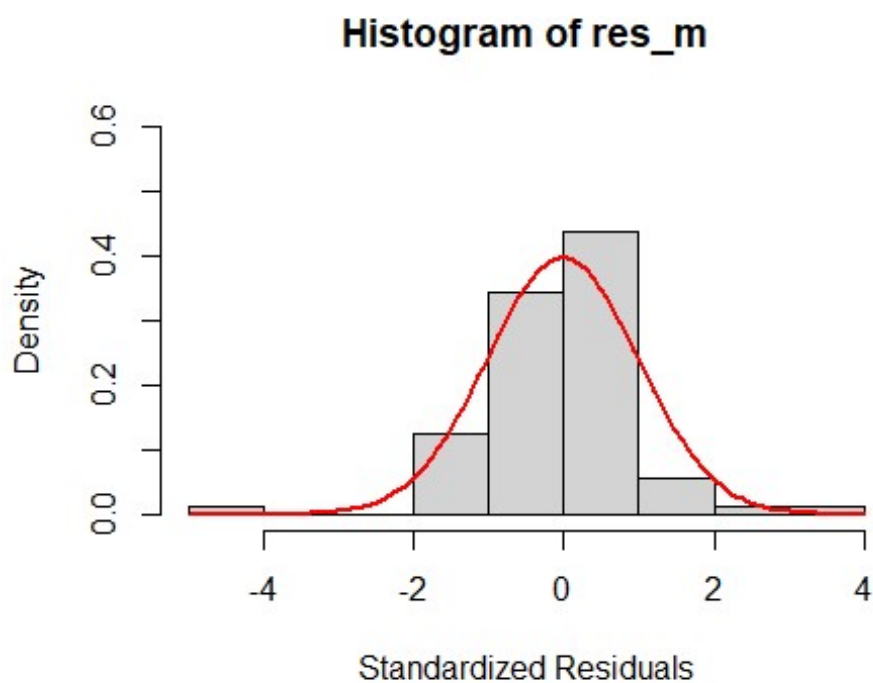## Histogram of res_m

Fig 41: Histogram of ARIMA(1,0,3)

Normal Q-Q Plot
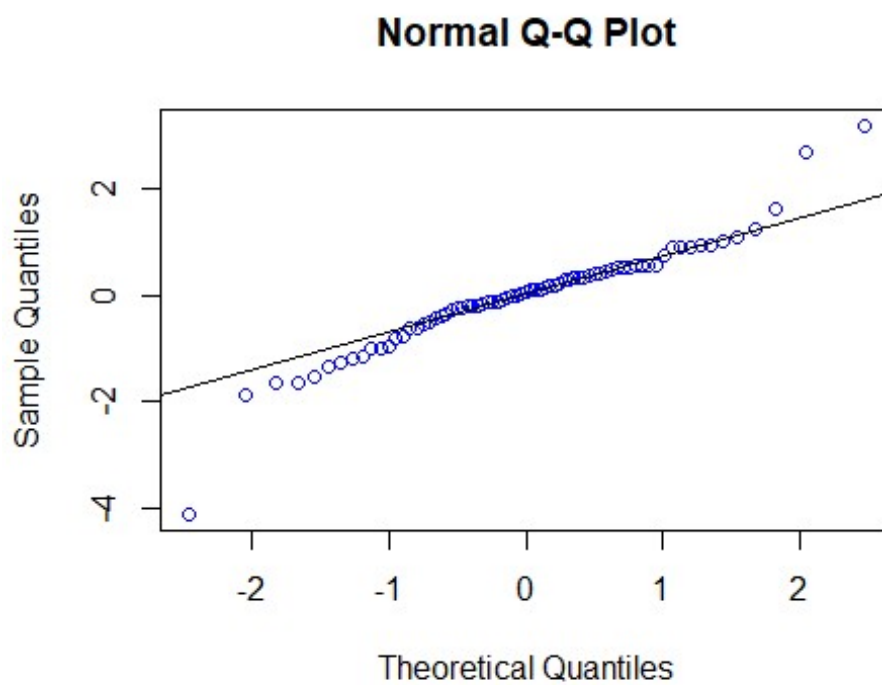

Fig 42: Normal Q-Q plot for ARIMA(1,0,3)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.92607, p-value = 0.0003661

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

ACF of Standardized Residuals

Fig 43: Autocorrelation plot for ARIMA(1,0,3)

Ljung-Box Test

Fig 44: Ljung-Box test for ARIMA(1,0,3)

Residual Analysis for ARIMA(1,0,3):

1. p-value is less than 0.05 for ar1, ma1 and ma2. Hence, significant.
2. For ma3 p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it can be considered as a good model but comparatively, ARIMA(1,0,2) performed better as the coefficients of ARIMA(1,0,3) are less significant than the coefficients of ARIMA(1,0,2).

ARIMA (2,0,0)

```
## 
## z test of coefficients:
## 
##             Estimate Std. Error z value Pr(>|z|)
## ar1        0.0039028  0.1123138  0.0347  0.97228
## ar2        0.2787470  0.1129483  2.4679  0.01359 *
## intercept 0.0309668  0.0561633  0.5514  0.58138
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
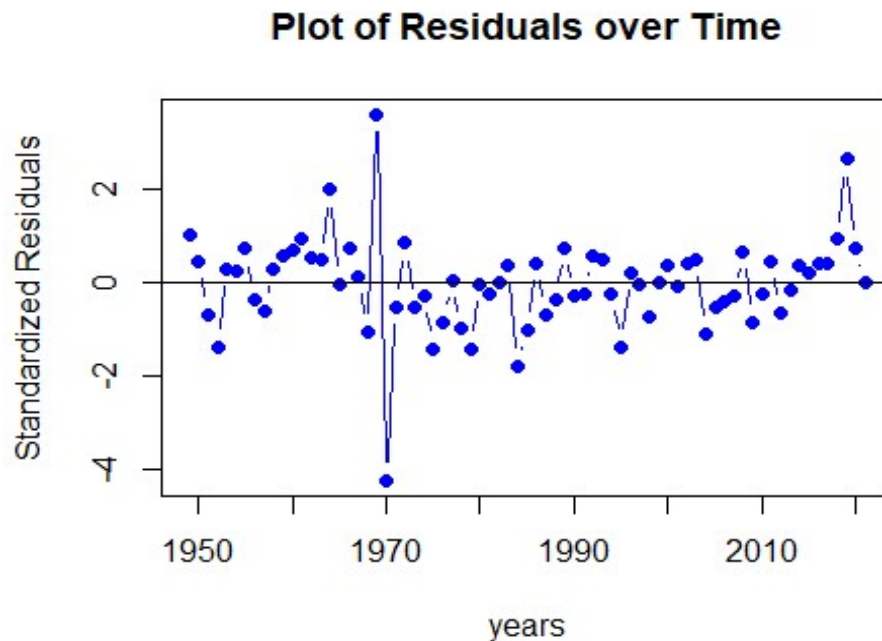

Fig 45: Residual's plot over time for ARIMA(2,0,0)

Histogram of res_m

Fig 46: Histogram of ARIMA(2,0,0)

Normal Q-Q Plot

Fig 47: Normal Q-Q plot for ARIMA(2,0,0)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.89873, p-value = 2.419e-05

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

ACF of Standardized Residuals

Fig 48: Autocorrelation plot for ARIMA(2,0,0)

Ljung-Box Test

Fig 49: Ljung-Box test for ARIMA(2,0,0)

35

Residual Analysis for ARIMA(2,0,0):

1. p-value is greater than 0.05 for ar1 and less than 0.05 for ar2.
2. In Residual plot, the line is passing through 0.
3. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
4. There is one significant lag in Autocorrelation plot.
5. From Ljung box test, we can observe one dot touching the dotted line.

Therefore, it is not a good model for the data.

ARIMA (2,0,1)

```
##
## z test of coefficients:
##
##             Estimate Std. Error z value  Pr(>|z|)
## ar1         0.474855   0.172715  2.7494 0.0059711 **
## ar2         0.333597   0.116953  2.8524 0.0043391 **
## ma1        -0.546236   0.157253 -3.4736 0.0005135 ***
## intercept  0.055383   0.090437  0.6124 0.5402775
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
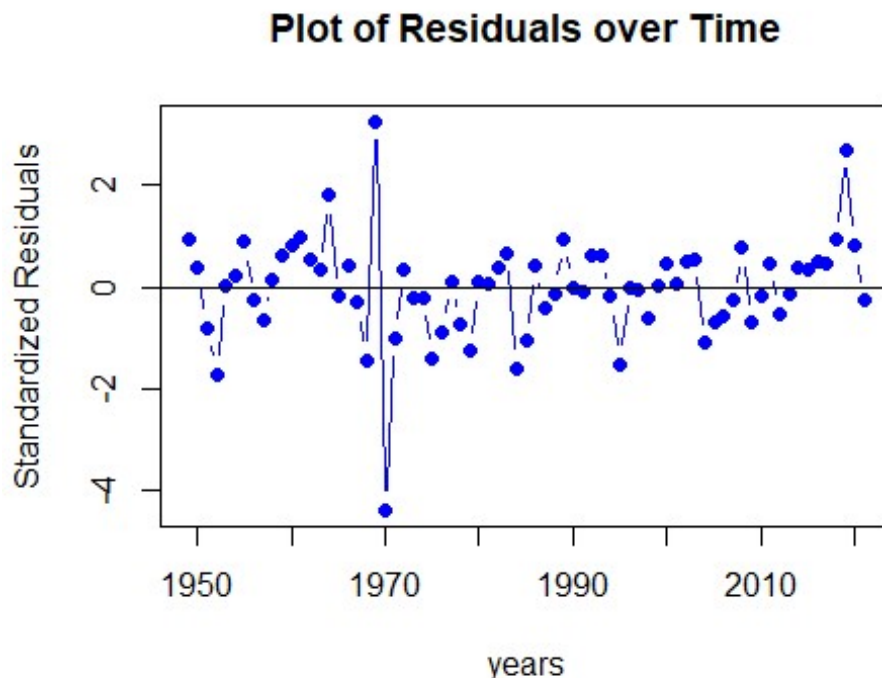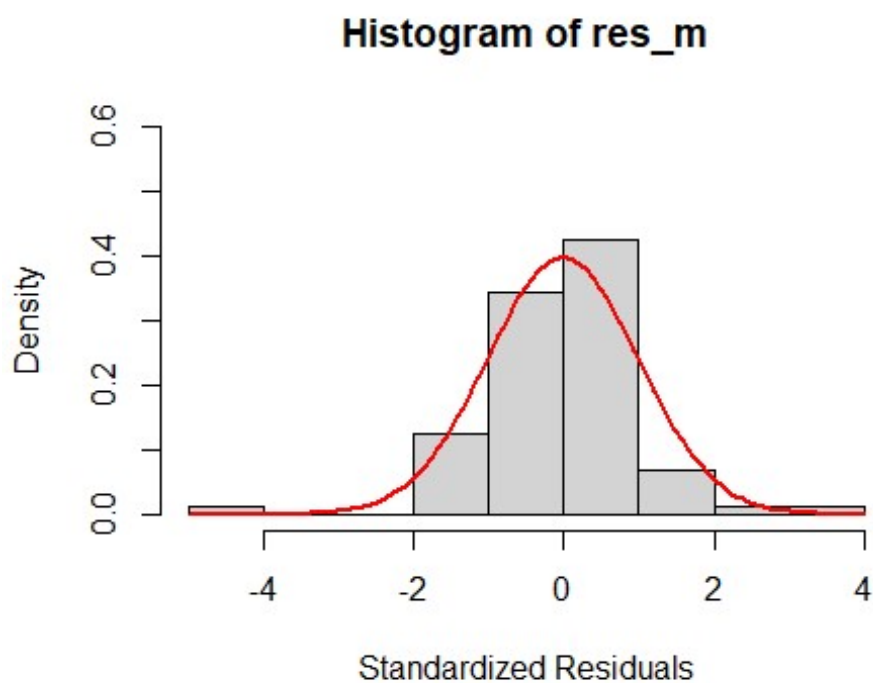

Fig 50: Residual's plot over time for ARIMA(2,0,1)
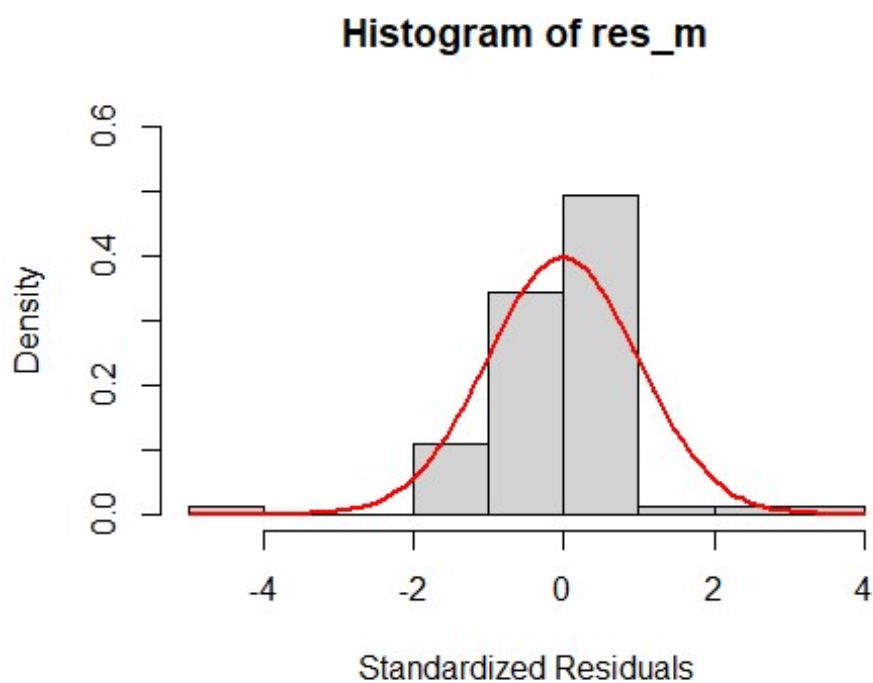
Histogram of res_m

Fig 51: Histogram of ARIMA(2,0,1)
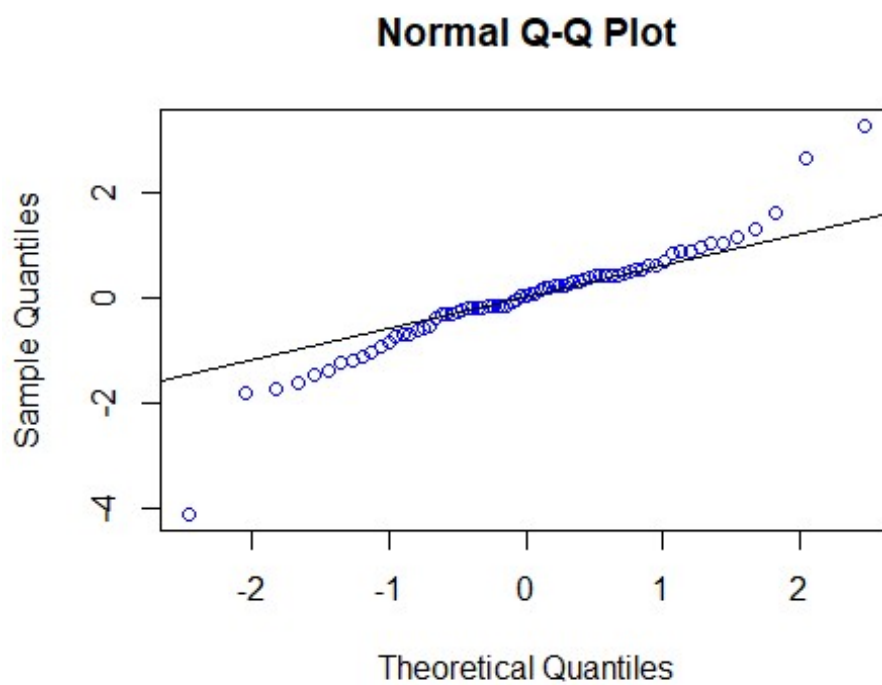
Normal Q-Q Plot

Fig 52: Normal Q-Q plot for ARIMA(2,0,1)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.90391, p-value = 3.926e-05

## Warning in (ra^2)/(n - (1:lag.max)): longer object length is not a multipl
e of
## shorter object length
```
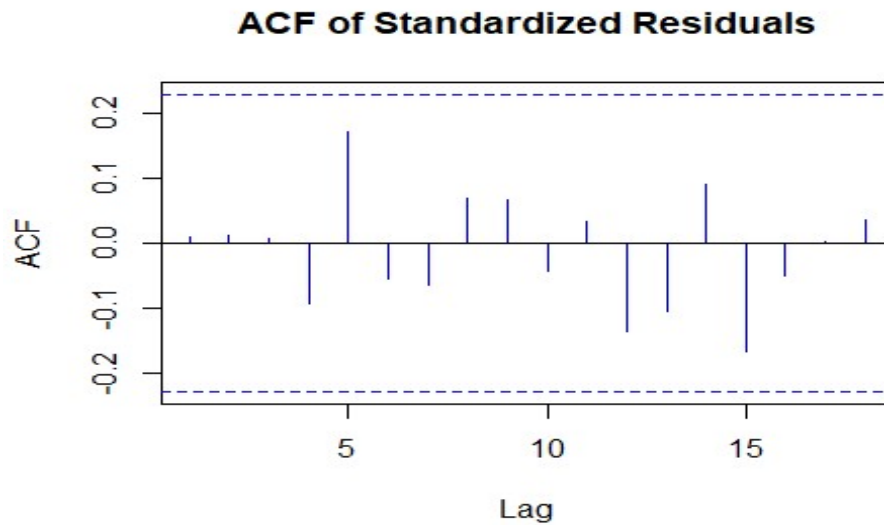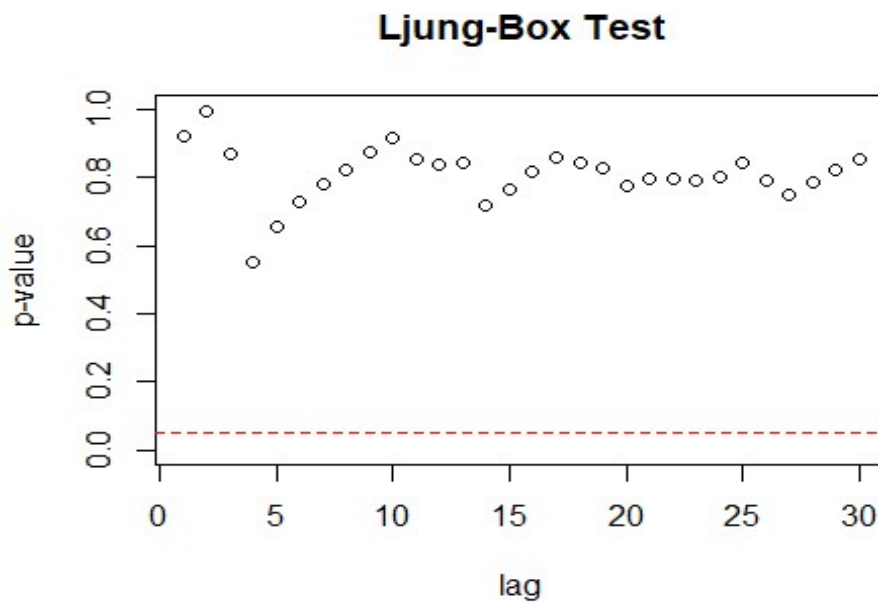
Fig 53: Autocorrelation plot for ARIMA(2,0,1)

Fig 54: Ljung-Box test for ARIMA(2,0,1)

Residual Analysis for ARIMA(2,0,1):

1. p-value is less than 0.05 for ar1, ar2 and ma1. Hence, significant.
2. In Residual plot, the line is passing through 0.
3. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
4. There are no significant lags in Autocorrelation plot.
5. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it can be considered as a good model but comparatively, ARIMA(1,0,2) performed better.

ARIMA (2,0,2)

```
##
## z test of coefficients:
##
##             Estimate Std. Error z value  Pr(>|z|)
## ar1         0.911148   0.241978  3.7654 0.0001663 ***
## ar2        -0.143022   0.285848 -0.5003 0.6168336
## ma1        -1.042590   0.207317 -5.0290 4.932e-07 ***
## ma2         0.540727   0.270317  2.0003 0.0454627 *
## intercept   0.050194   0.082589  0.6078 0.5433476
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Plot of Residuals over Time

Fig 55: Residual's plot over time for ARIMA(2,0,2)
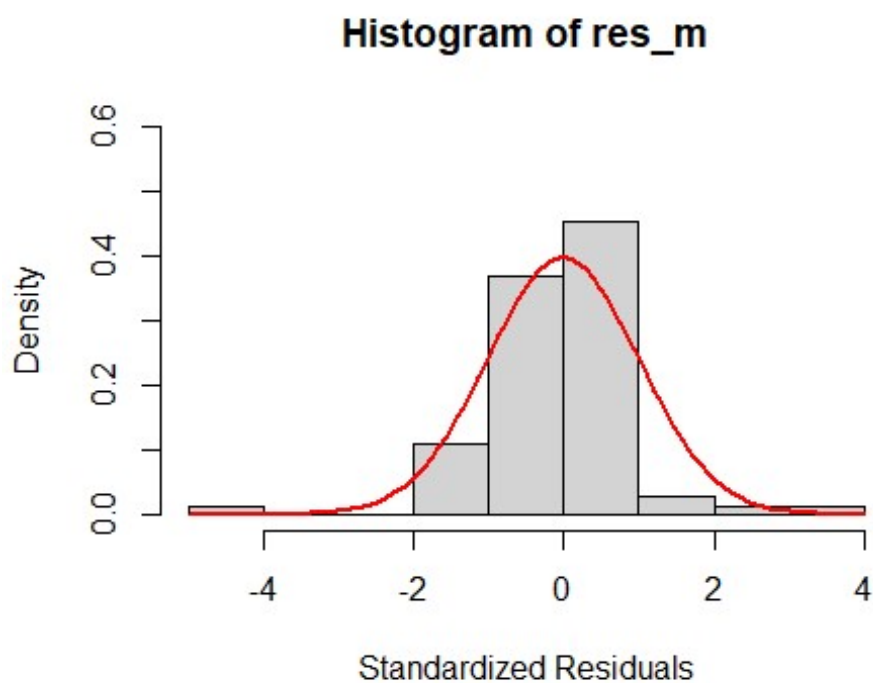
Histogram of res_m

Fig 56: Histogram of ARIMA(2,0,2)
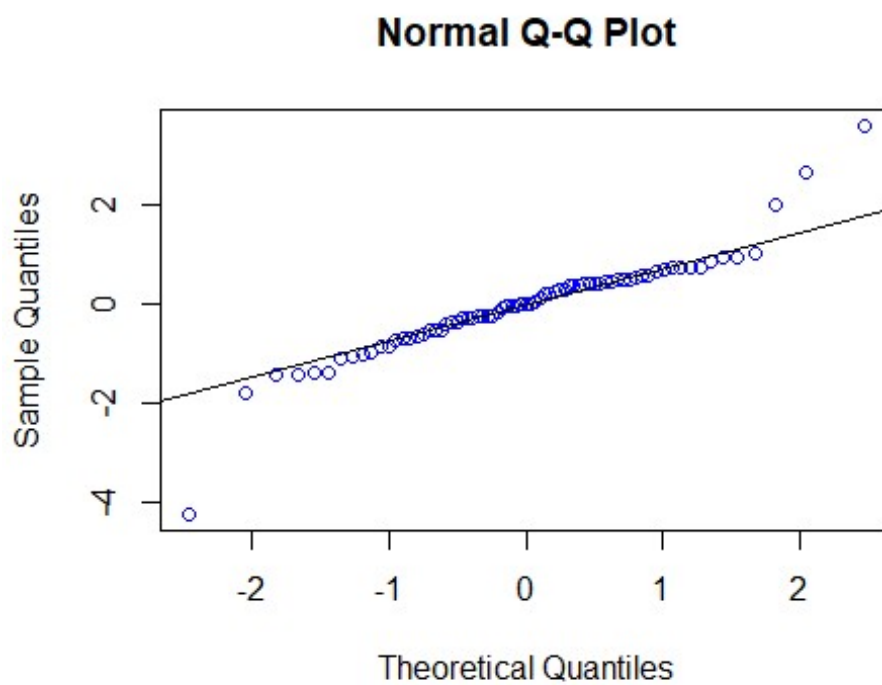
Normal Q-Q Plot

Fig 57: Normal Q-Q plot for ARIMA(2,0,2)

```
##
##   Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.92627, p-value = 0.0003739

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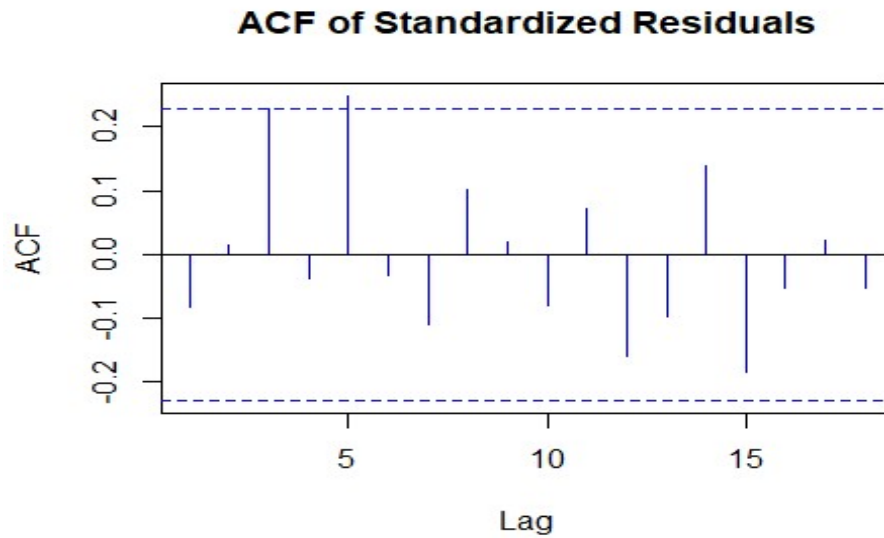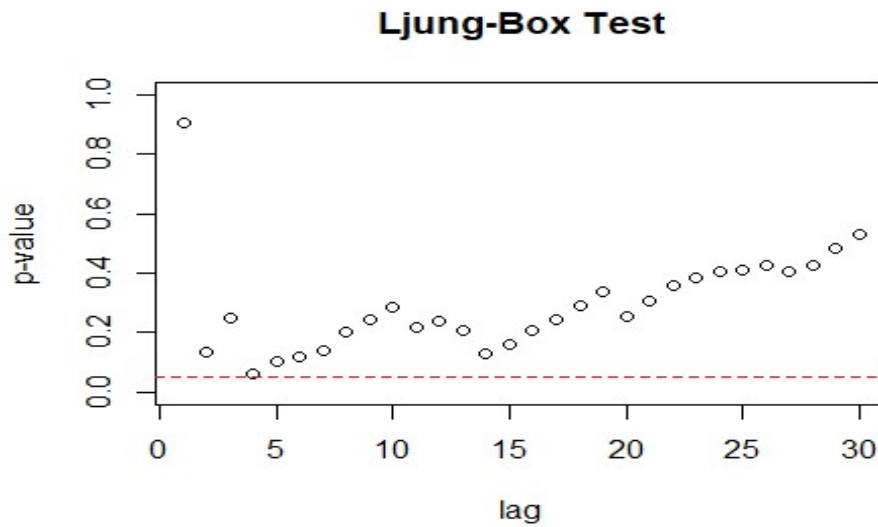

Fig 58: Autocorrelation plot for ARIMA(2,0,2)

Fig 59: Ljung-Box test for ARIMA(2,0,2)

Residual Analysis for ARIMA(2,0,2):

1. p-value is less than 0.05 for ar1, ma1 and ma2. Hence, significant.
2. For ar2 p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it can be considered as a good model but comparatively, ARIMA(1,0,2) performed better.

ARIMA (2,0,3)

```
##
## z test of coefficients:
##
##            Estimate Std. Error z value Pr(>|z|)
## ar1        0.035923   0.253459  0.1417 0.887293
## ar2        0.552169   0.248142  2.2252 0.026067 *
## ma1       -0.148764   0.249274 -0.5968 0.550647
## ma2       -0.318535   0.252178 -1.2631 0.206540
## ma3        0.405351   0.129386  3.1329 0.001731 **
## intercept  0.054010   0.085088  0.6348 0.525587
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Fig 60: Residual's plot over time for ARIMA(2,0,3)
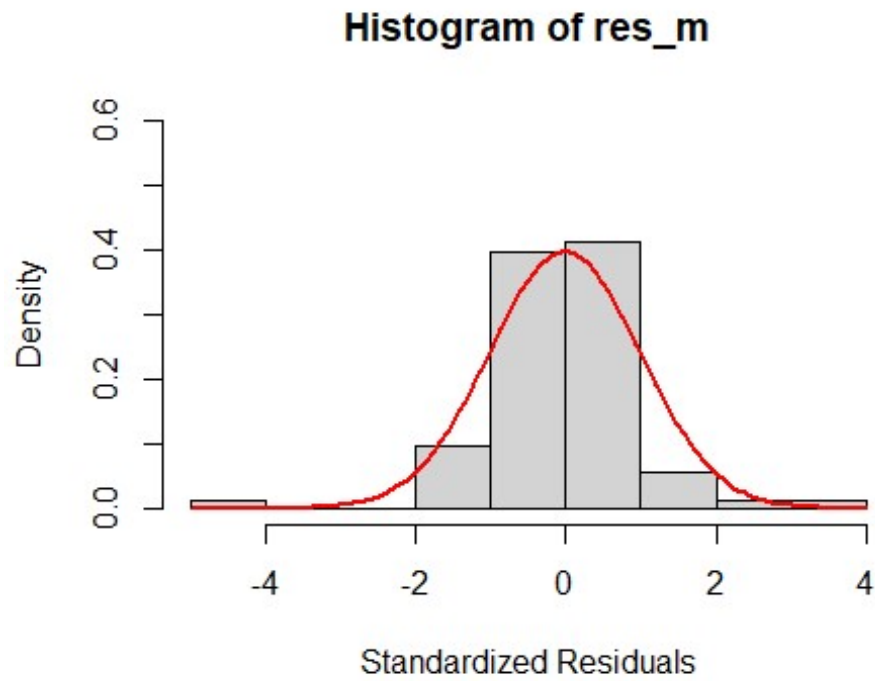
Histogram of res_m

Fig 61: Histogram of ARIMA(2,0,3)
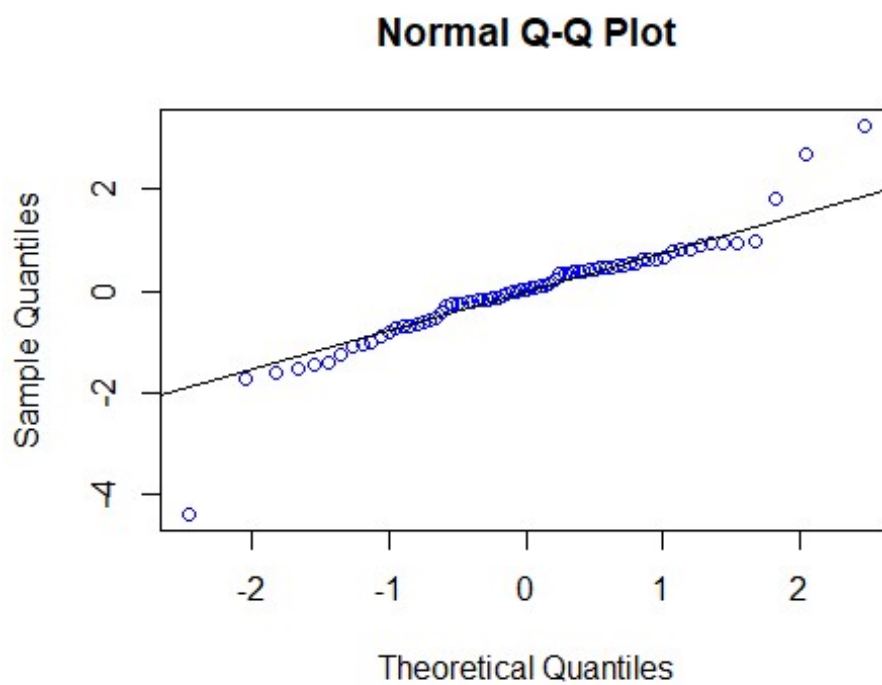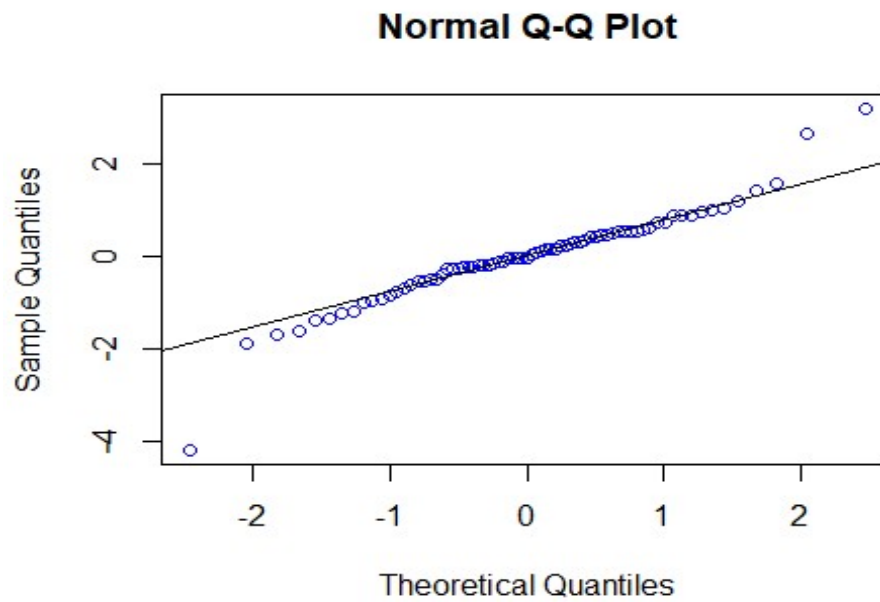
Normal Q-Q Plot

Fig 62: Normal Q-Q plot for ARIMA(2,0,3)

```
##
##   Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.92773, p-value = 0.0004378

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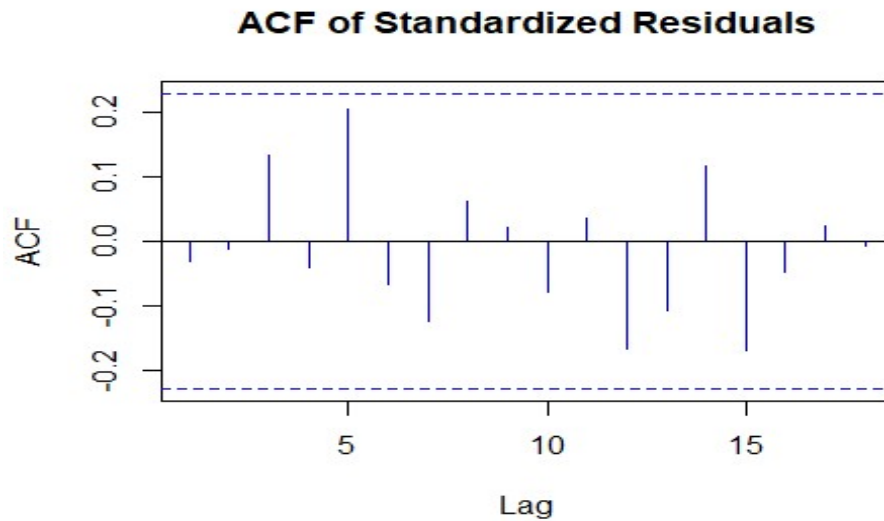
ACF of Standardized Residuals

Fig 63: Autocorrelation plot for ARIMA(2,0,3)
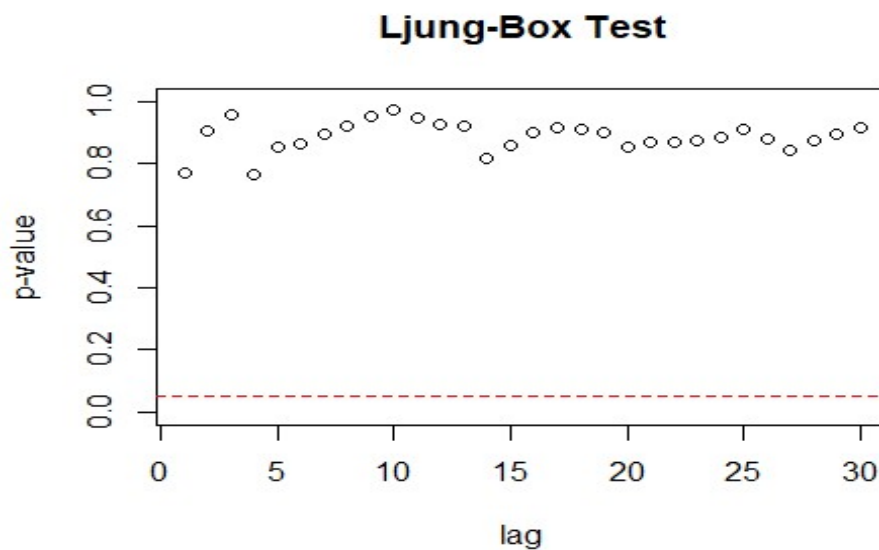
Ljung-Box Test

Fig 64: Ljung-Box test for ARIMA(2,0,3)

Residual Analysis for ARIMA(2,0,3):

1. p-value is less than 0.05 for ar2, and ma3. Hence, significant.
2. For ar1, ma1, and ma2, p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it cannot be considered as a good model.

ARIMA (3,0,0)

```
##
## z test of coefficients:
##
##             Estimate Std. Error z value Pr(>|z|)
## ar1        -0.081572   0.110613 -0.7375 0.460847
## ar2         0.295658   0.106759  2.7694 0.005616 **
## ar3         0.323403   0.116612  2.7733 0.005549 **
## intercept   0.050956   0.081397  0.6260 0.531305
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Fig 65: Residual's plot over time for ARIMA(3,0,0)
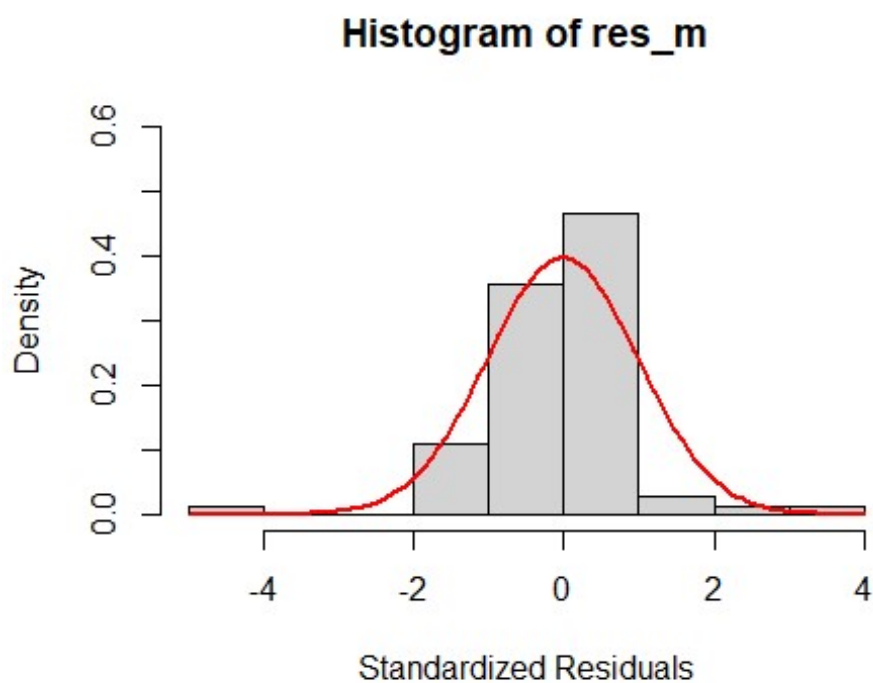
Fig 66: Histogram of ARIMA(3,0,0)
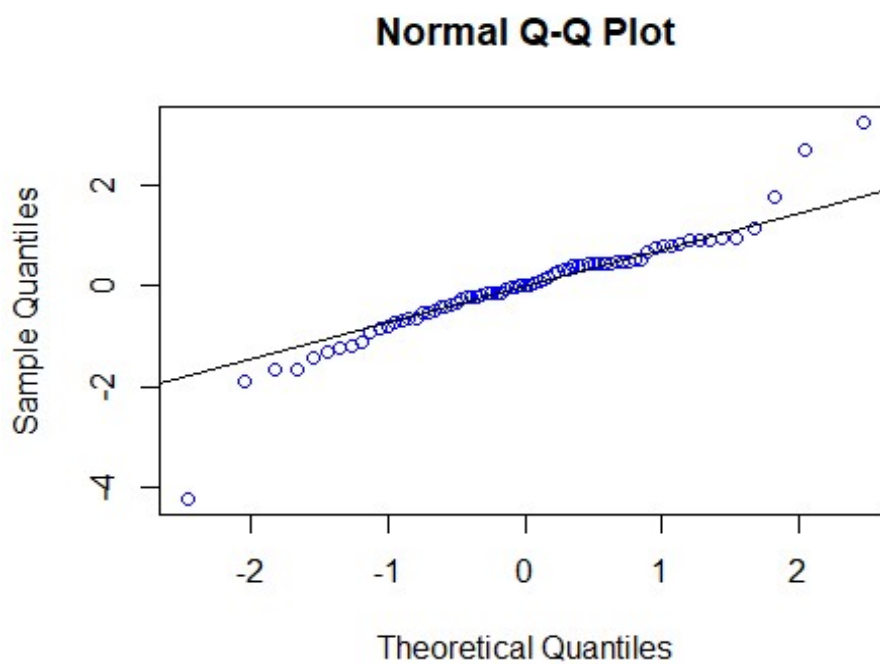
Fig 67: Normal Q-Q plot for ARIMA(3,0,0)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.91737, p-value = 0.0001474

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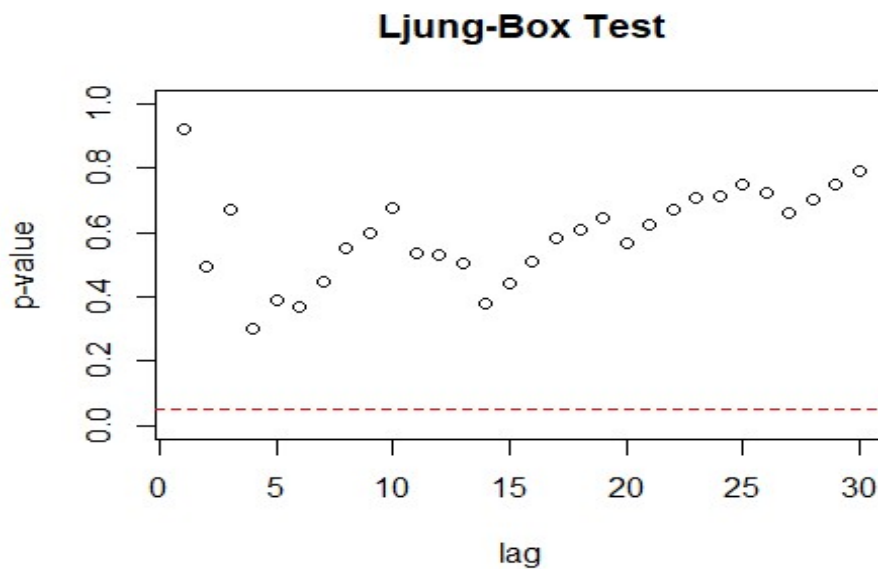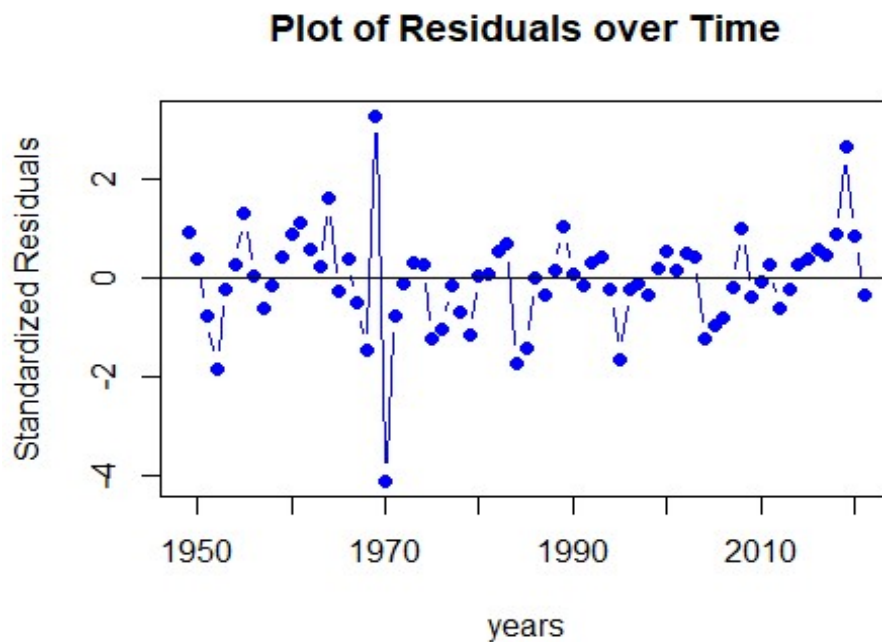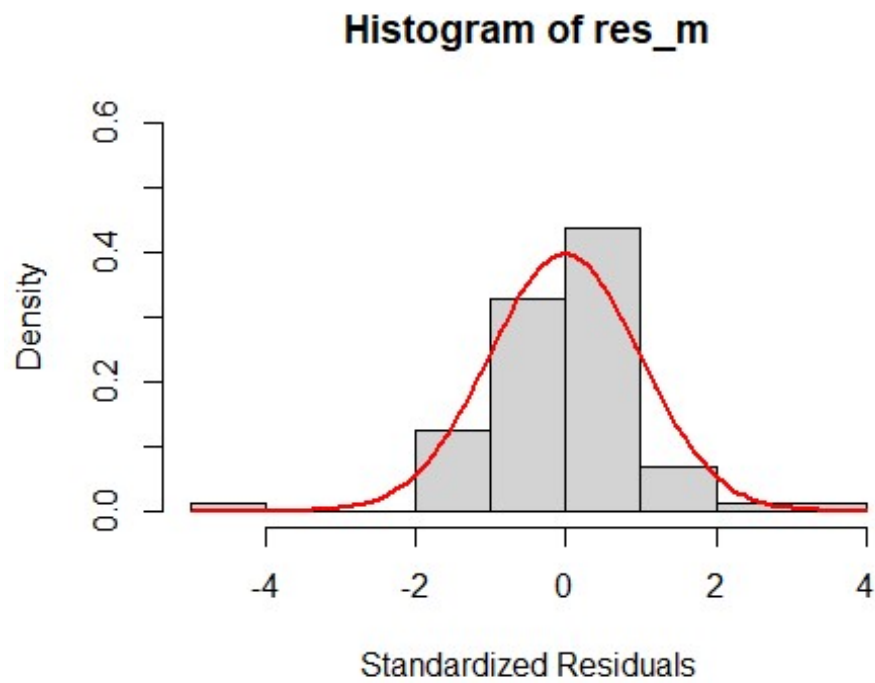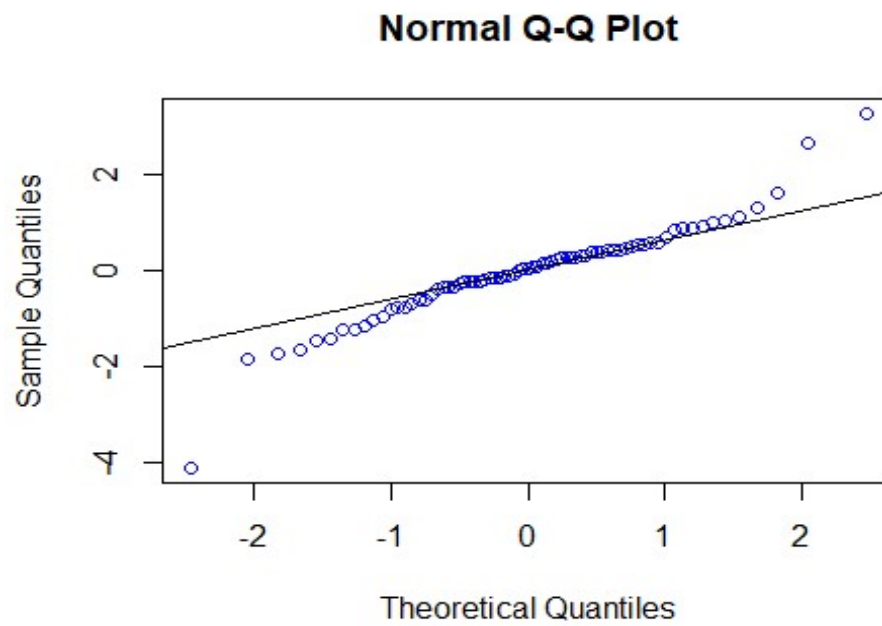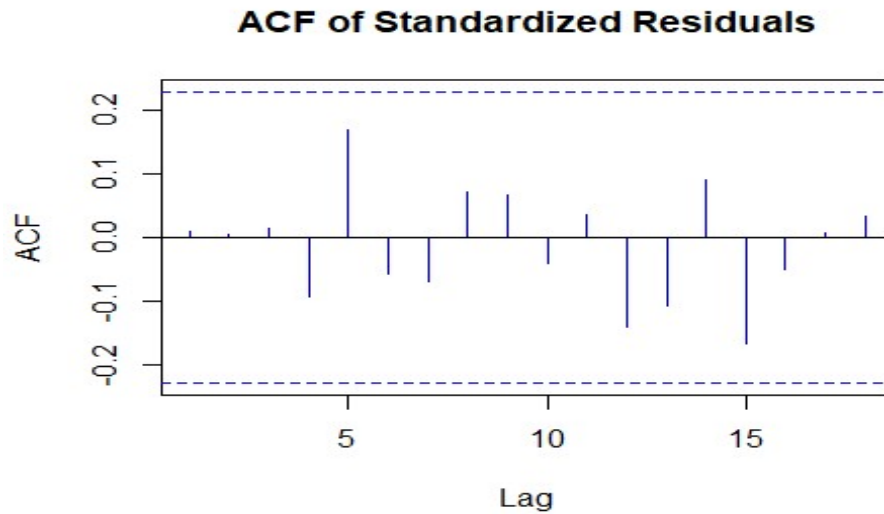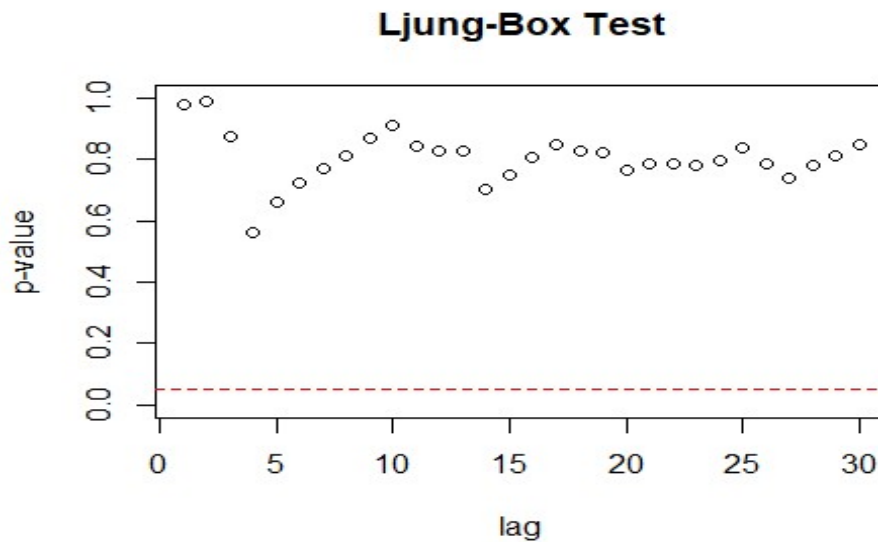

Fig 68: Autocorrelation plot for ARIMA(3,0,0)
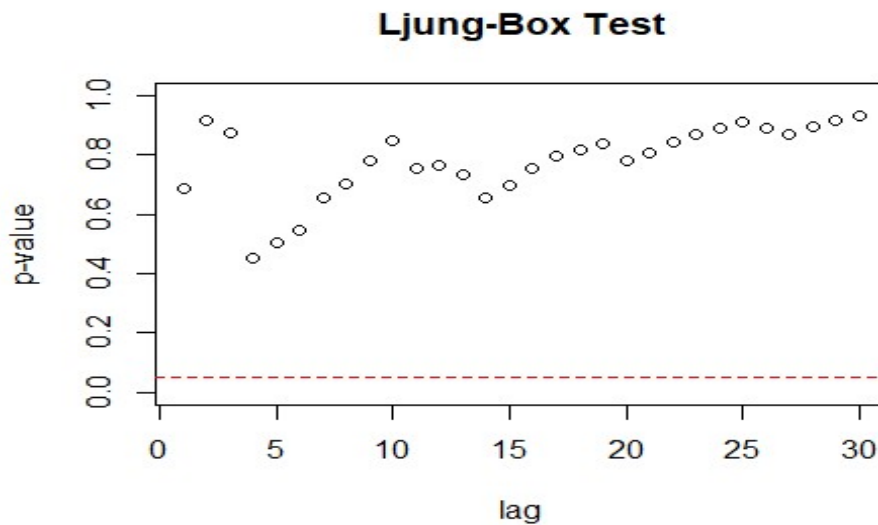
Fig 69: Ljung-Box test for ARIMA(3,0,0)

Residual Analysis for ARIMA(3,0,0):

1. p-value is less than 0.05 for ar2, and ar3. Hence, significant.
2. For ar1, p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it cannot be considered as a good model.

ARIMA (3,0,2)

```
##
## z test of coefficients:
##
##            Estimate Std. Error z value  Pr(>|z|)
## ar1       -0.323554   0.280430 -1.1538 0.2485909
## ar2        0.528993   0.216385  2.4447 0.0144976 *
## ar3        0.396153   0.117745  3.3645 0.0007668 ***
## ma1        0.250286   0.301592  0.8299 0.4066050
## ma2       -0.282254   0.238282 -1.1845 0.2362011
## intercept  0.057545   0.090998  0.6324 0.5271440
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
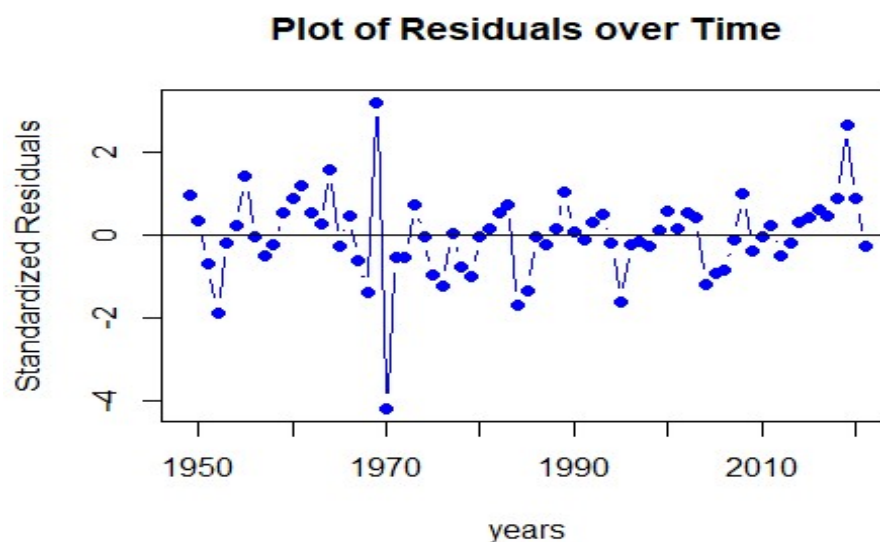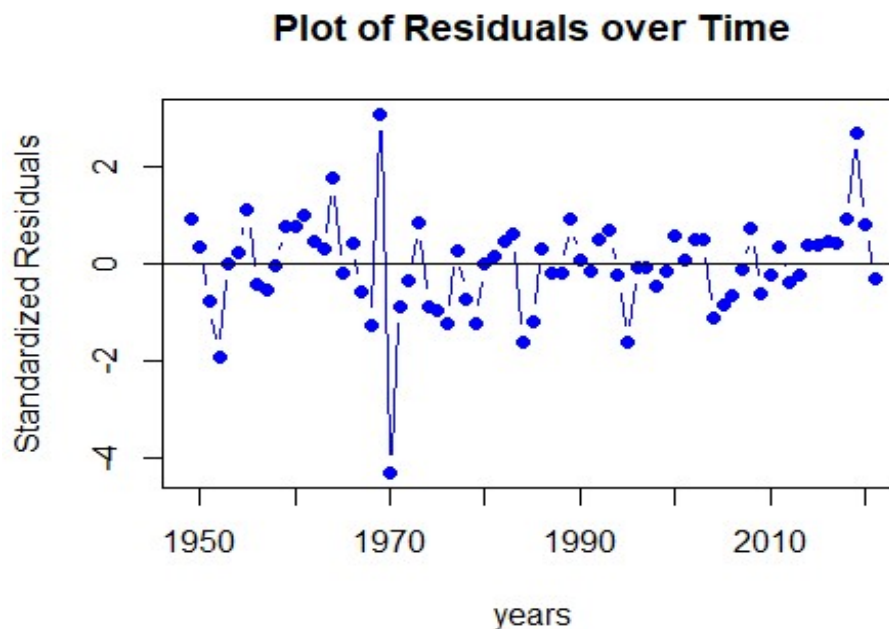

Fig 70: Residual's plot over time for ARIMA(3,0,2)
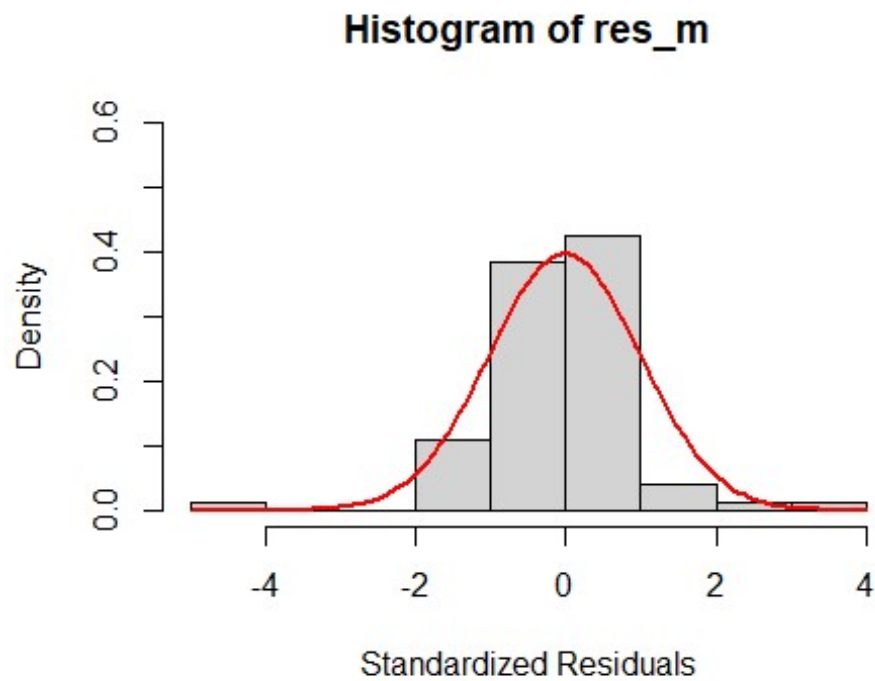
Histogram of res_m

Fig 71: Histogram of ARIMA(3,0,2)
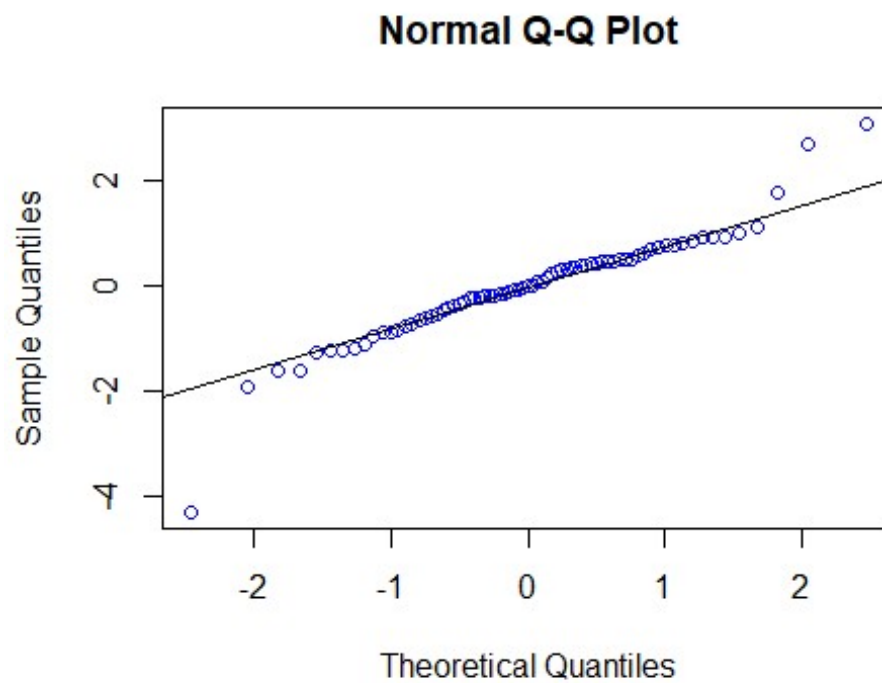
Normal Q-Q Plot

Fig 72: Normal Q-Q plot for ARIMA(3,0,2)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.92004, p-value = 0.0001939

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

ACF of Standardized Residuals

Fig 73: Autocorrelation plot for ARIMA(3,0,2)
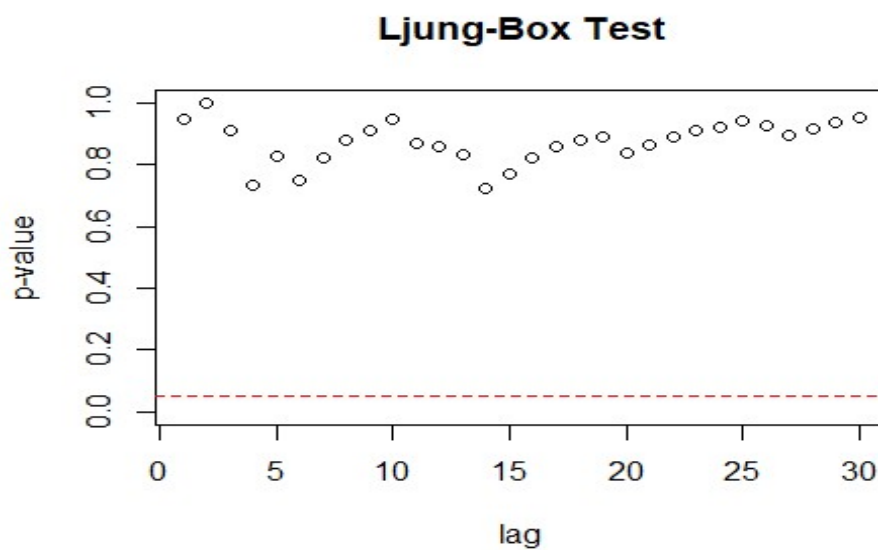
Ljung-Box Test

Fig 74: Ljung-Box test for ARIMA(3,0,2)

Residual Analysis for ARIMA(3,0,2):

1. p-value is less than 0.05 for ar2, and ar3. Hence, significant.
2. For ar1, ma1 and ma2, p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it cannot be considered as a good model.

ARIMA (3,0,3)

```
##
## z test of coefficients:
##
##            Estimate Std. Error z value Pr(>|z|)
## ar1        0.022084   0.328317  0.0673  0.94637
## ar2        0.553295   0.243475  2.2725  0.02306 *
## ar3        0.021414   0.315931  0.0678  0.94596
## ma1       -0.136343   0.316287 -0.4311  0.66641
## ma2       -0.319450   0.247150 -1.2925  0.19617
## ma3        0.387554   0.294861  1.3144  0.18872
## intercept  0.054969   0.087382  0.6291  0.52931
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
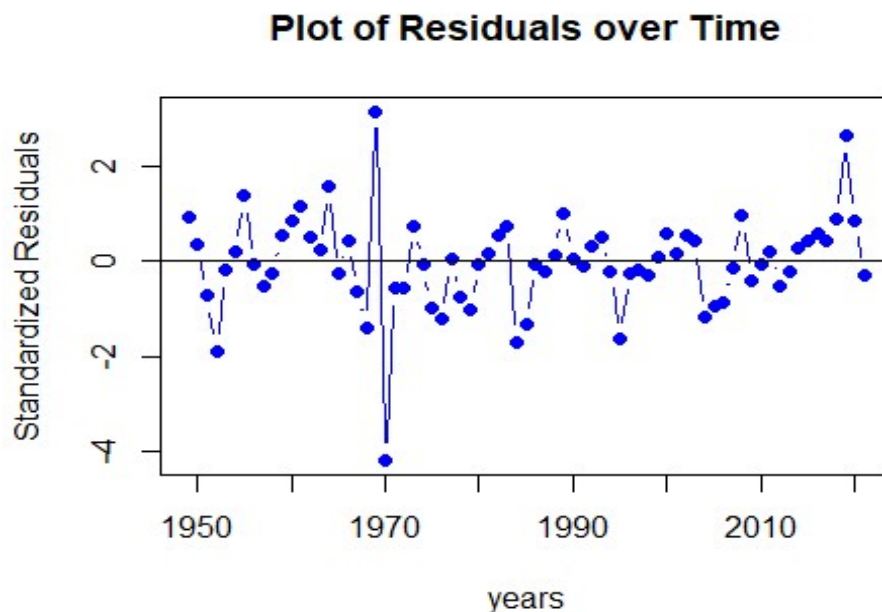

Fig 75: Residual's plot over time for ARIMA(3,0,3)
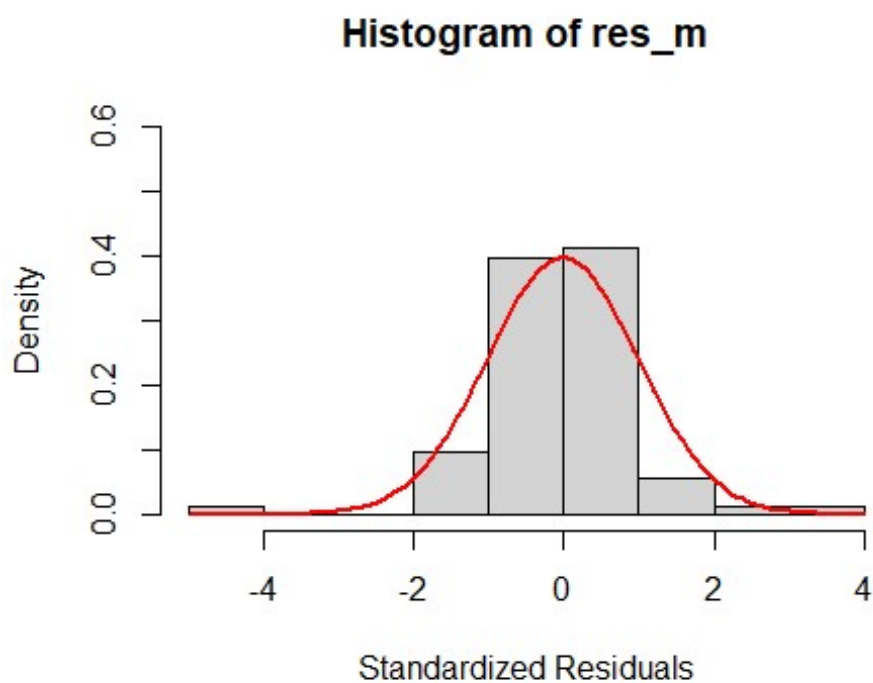
Histogram of res_m

Fig 76: Histogram of ARIMA(3,0,3)
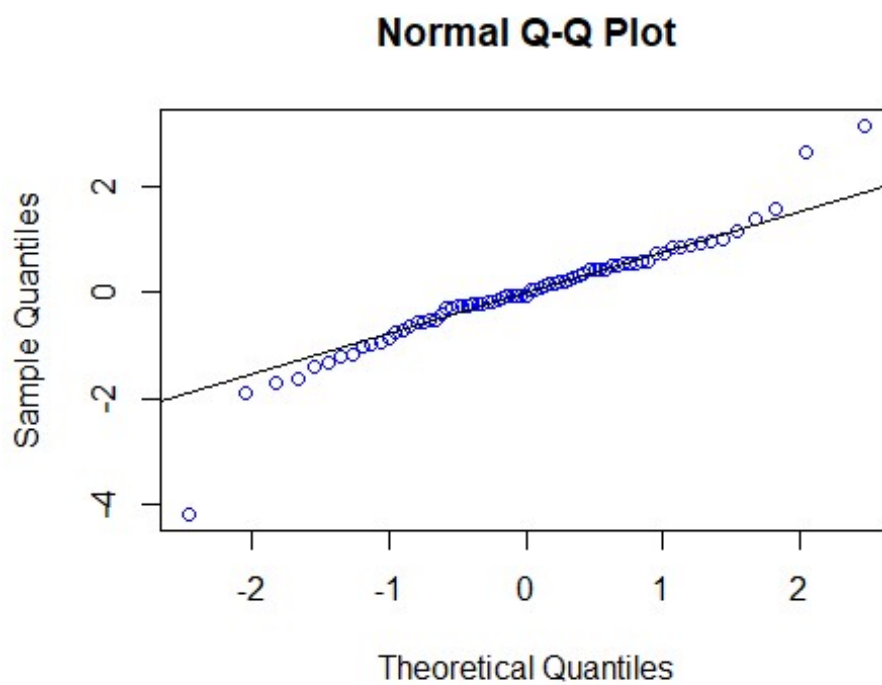
Normal Q-Q Plot

Fig 77: Normal Q-Q plot for ARIMA(3,0,3)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.92803, p-value = 0.0004523

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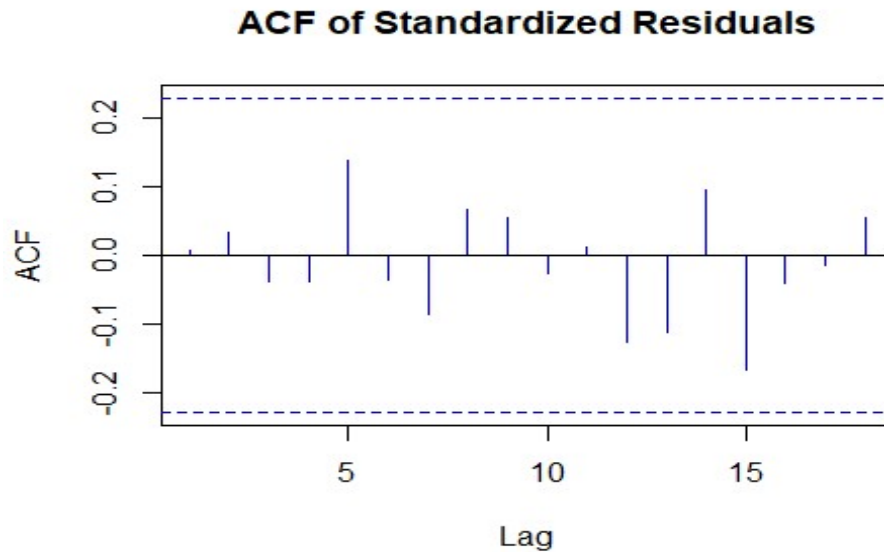
ACF of Standardized Residuals

Fig 78: Autocorrelation plot for ARIMA(3,0,3)
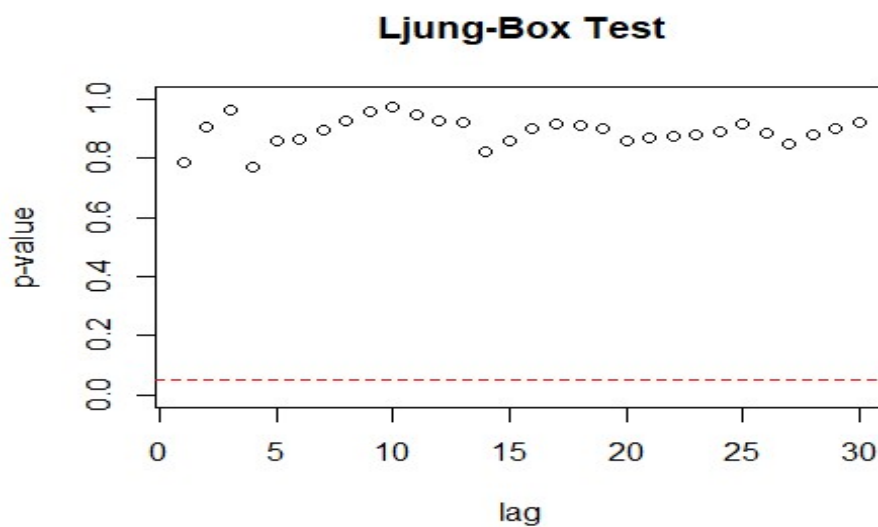
Ljung-Box Test

Fig 79: Ljung-Box test for ARIMA(3,0,3)

53

Residual Analysis for ARIMA(3,0,3):

1. p-value is less than 0.05 for ar2. Hence, significant.
2. For ar1, ar3, ma1, ma2 and ma3, p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it cannot be considered as a good model.

ARIMA (3,0,4)

```
## 
## z test of coefficients:
## 
##             Estimate Std. Error z value  Pr(>|z|)
## ar1        -1.038563   0.203496 -5.1036 3.332e-07 ***
## ar2         0.540636   0.250023  2.1623 0.0305915 *
## ar3         0.696890   0.182612  3.8162 0.0001355 ***
## ma1         0.943105   0.224047  4.2094 2.561e-05 ***
## ma2        -0.418491   0.269339 -1.5538 0.1202383
## ma3        -0.052883   0.261685 -0.2021 0.8398476
## ma4         0.427136   0.144079  2.9646 0.0030309 **
## intercept   0.055116   0.087359  0.6309 0.5280958
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

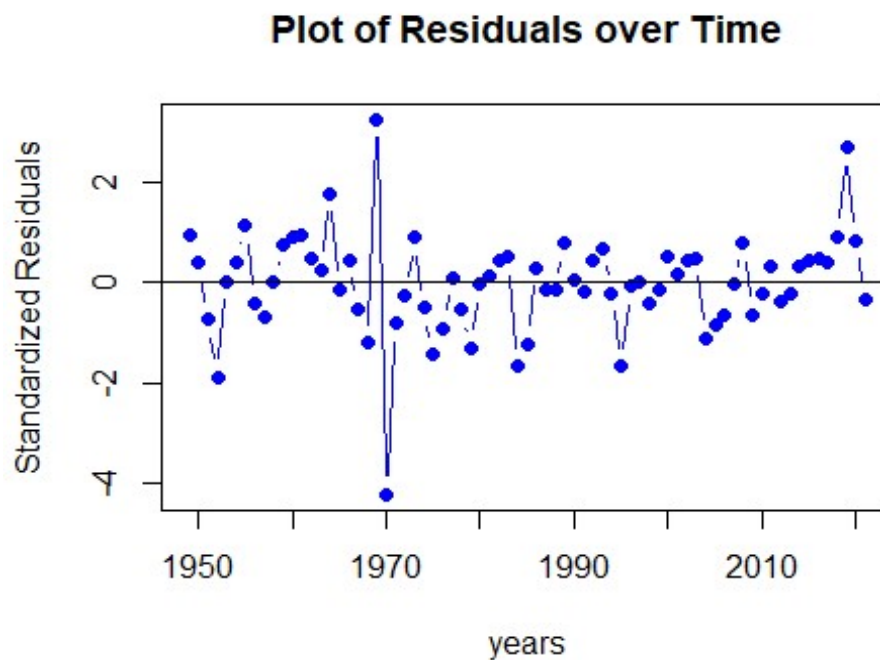## Plot of Residuals over Time
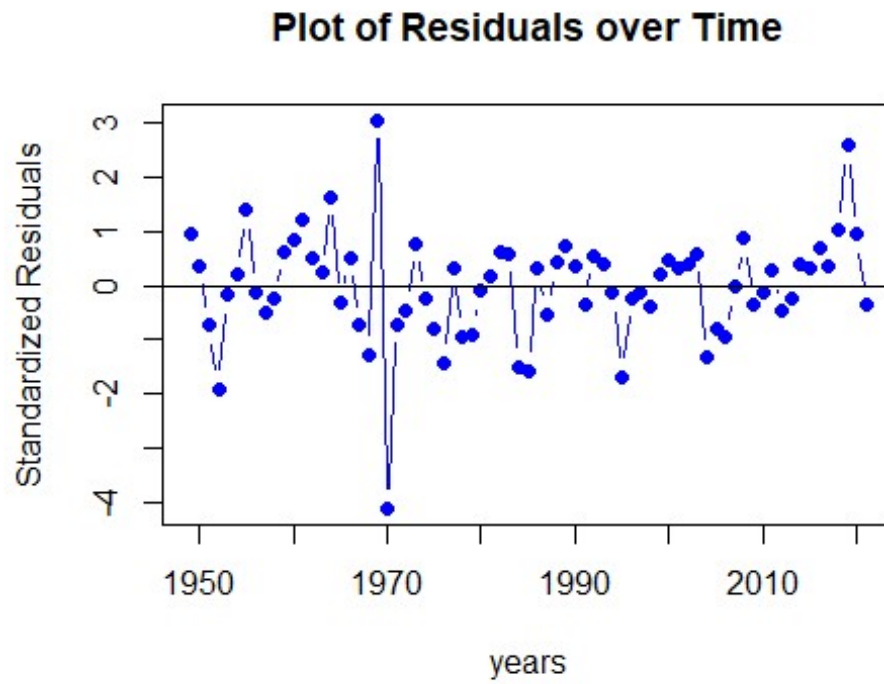
Fig 80: Residual's plot over time for ARIMA(3,0,4)
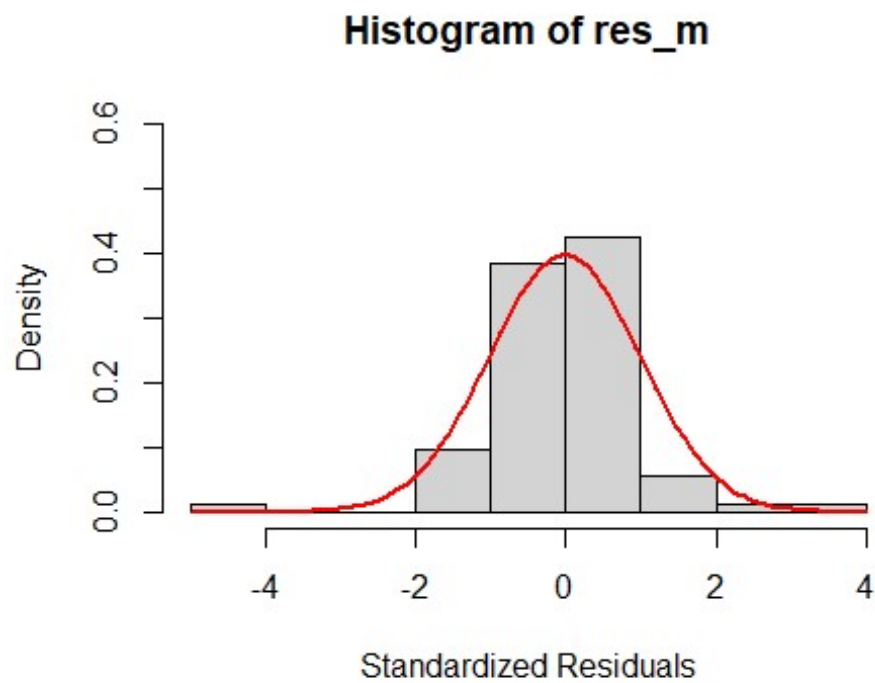
Histogram of res_m

Fig 81: Histogram of ARIMA(3,0,4)
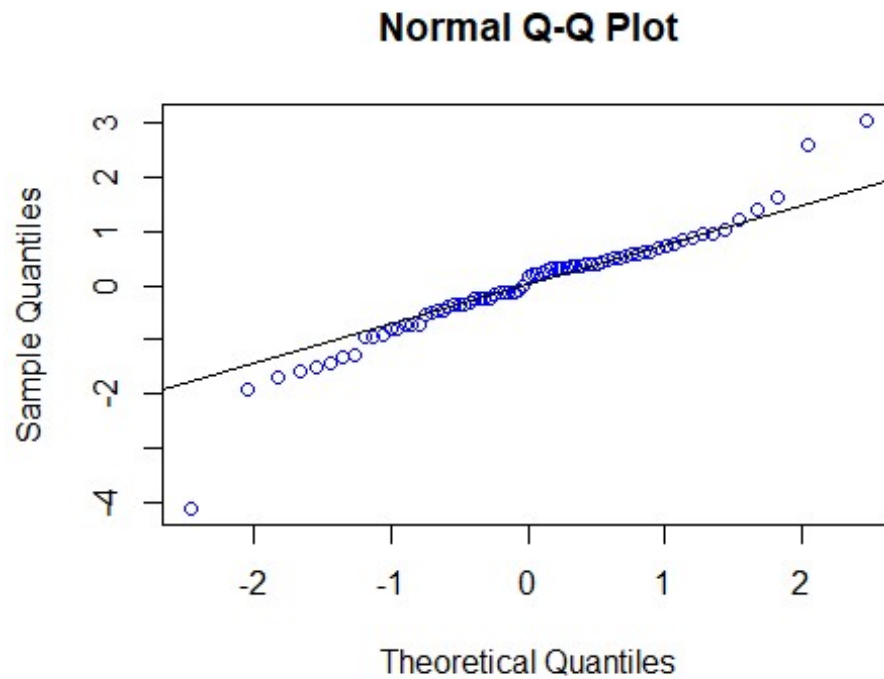
Normal Q-Q Plot

Fig 82: Normal Q-Q plot for ARIMA(3,0,4)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.93499, p-value = 0.0009773

## Warning in (ra^2)/(n - (1:lag.max)): longer object length is not a multipl
e of
## shorter object length
```
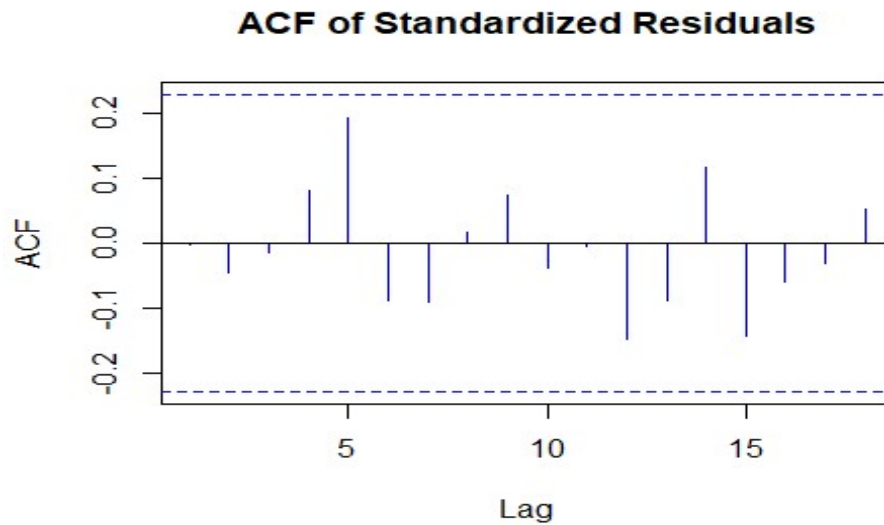
ACF of Standardized Residuals
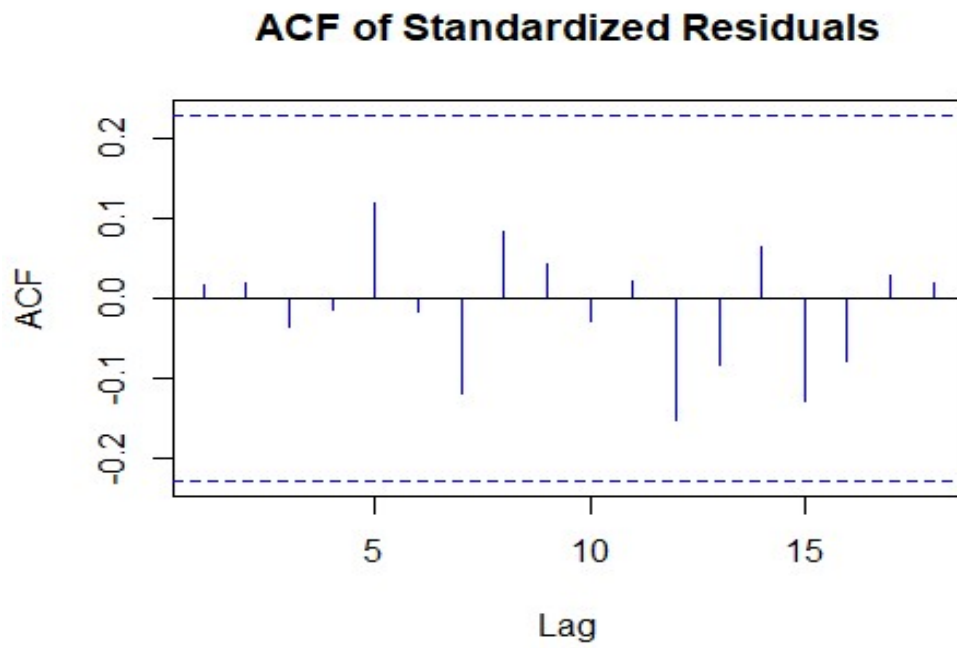
Fig 83: Autocorrelation plot for ARIMA(3,0,4)
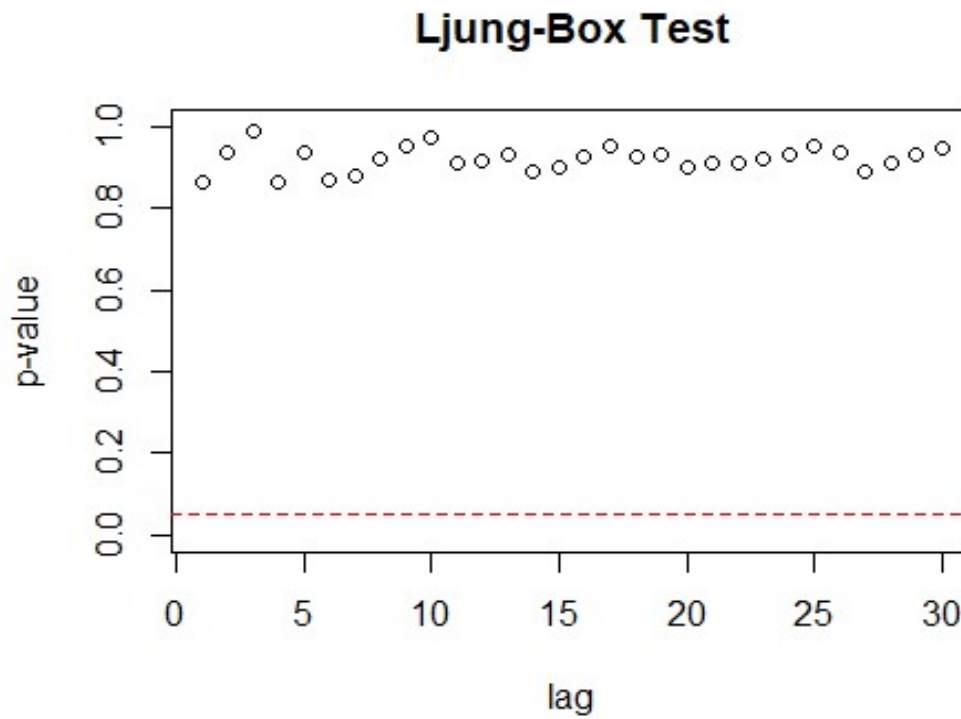
Ljung-Box Test

Fig 84: Ljung-Box test for ARIMA(3,0,4)

Residual Analysis for ARIMA(3,0,4):

1. p-value is less than 0.05 for ar1, ar2, ar3, ma1 and ma4. Hence, significant.
2. For ma2 and ma3, p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it cannot be considered as a good model.

```
##              df      AIC
## model_12_ml  5 55.15990
## model_30_ml  5 55.55828
## model_13_ml  6 56.87477
## model_22_ml  6 56.89851
## model_21_ml  5 57.75292
## model_23_ml  7 57.96646
## model_32_ml  7 58.51401
## model_03_ml  5 58.99181
## model_33_ml  8 59.96165
## model_02_ml  4 60.29462
## model_34_ml  9 60.55548
## model_20_ml  4 60.76434
## model_10_ml  3 64.57871
## model_01_ml  3 64.57880
## model_11_ml  4 66.32107
```

Overall, ARIMA(1,0,2) is the best model.

Now over-fitting AR and MA,

Over-fitting models: AR - ARMA(2, 0, 2) MA - ARMA(1, 0, 3)

Over-fitting model - ARIMA (2,0,2)

```
##
## z test of coefficients:
##
##            Estimate Std. Error z value  Pr(>|z|)
## ar1        0.911148   0.241978  3.7654 0.0001663 ***
## ar2       -0.143022   0.285848 -0.5003 0.6168336
## ma1       -1.042590   0.207317 -5.0290 4.932e-07 ***
## ma2        0.540727   0.270317  2.0003 0.0454627 *
## intercept  0.050194   0.082589  0.6078 0.5433476
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Plot of Residuals over Time

Fig 55: Residual's plot over time for ARIMA(2,0,2)

Histogram of res_m

Fig 56: Histogram of ARIMA(2,0,2)

Normal Q-Q Plot

Fig 57: Normal Q-Q plot for ARIMA(2,0,2)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.92627, p-value = 0.0003739

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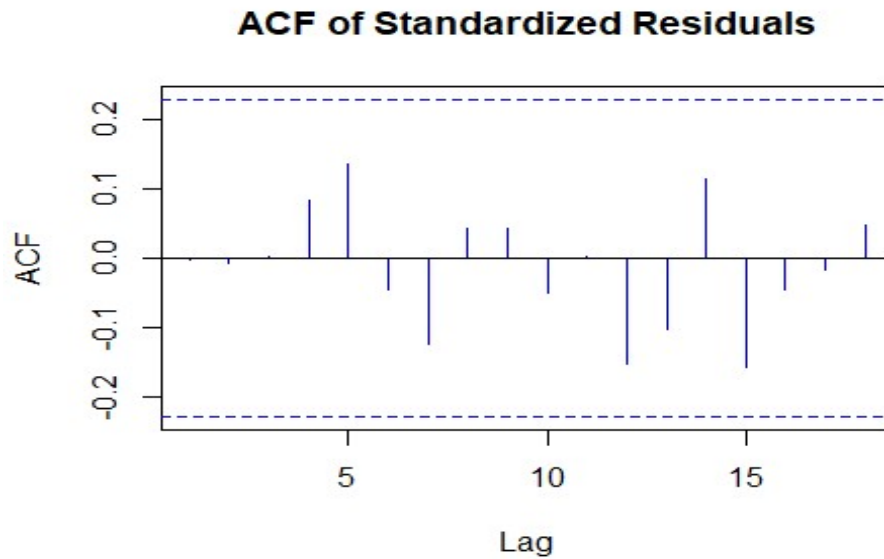
ACF of Standardized Residuals

Fig 58: Autocorrelation plot for ARIMA(2,0,2)

Ljung-Box Test

Fig 59: Ljung-Box test for ARIMA(2,0,2)

Residual Analysis for ARIMA(2,0,2):

1. p-value is less than 0.05 for ar1, ma1 and ma2. Hence, significant.
2. For ar2 p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it can be considered as a good model but comparatively, ARIMA(1,0,2) performed better.

Over-fitting - ARIMA (1,0,3)

```
##
## z test of coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## ar1         0.716644   0.273053  2.6246 0.008676 **
## ma1        -0.846430   0.293113 -2.8877 0.003880 **
## ma2         0.360540   0.169874  2.1224 0.033804 *
## ma3         0.096113   0.211756  0.4539 0.649912
## intercept   0.050341   0.081955  0.6143 0.539046
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
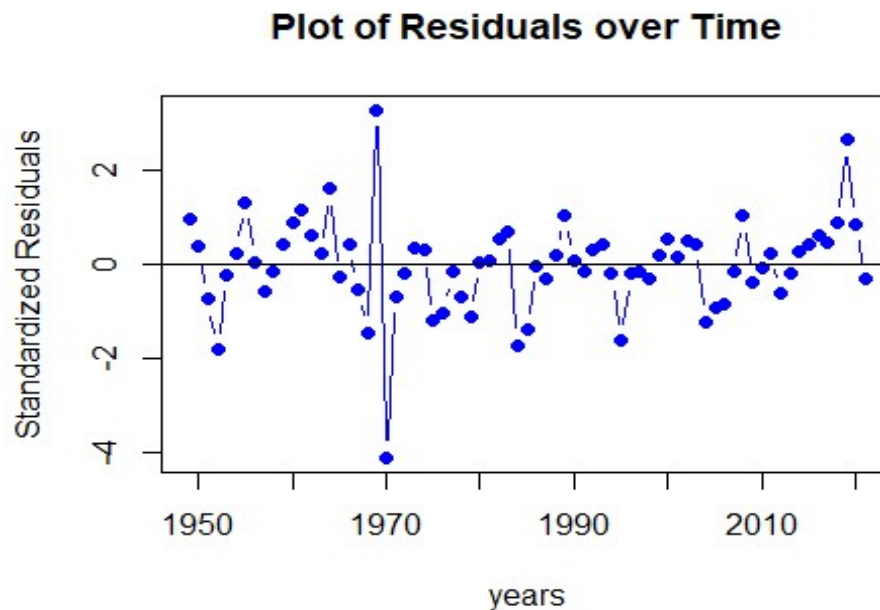

Fig 85: Residual's plot over time for ARIMA(1,0,3)

Histogram of res_m

Fig 86: Histogram of ARIMA(1,0,3)

Normal Q-Q Plot

Fig 87: Normal Q-Q plot for ARIMA(1,0,3)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.92607, p-value = 0.0003661

## Warning in (ra^2)/(n - (1:lag.max)): longer object length is not a multipl
e of
## shorter object length
```
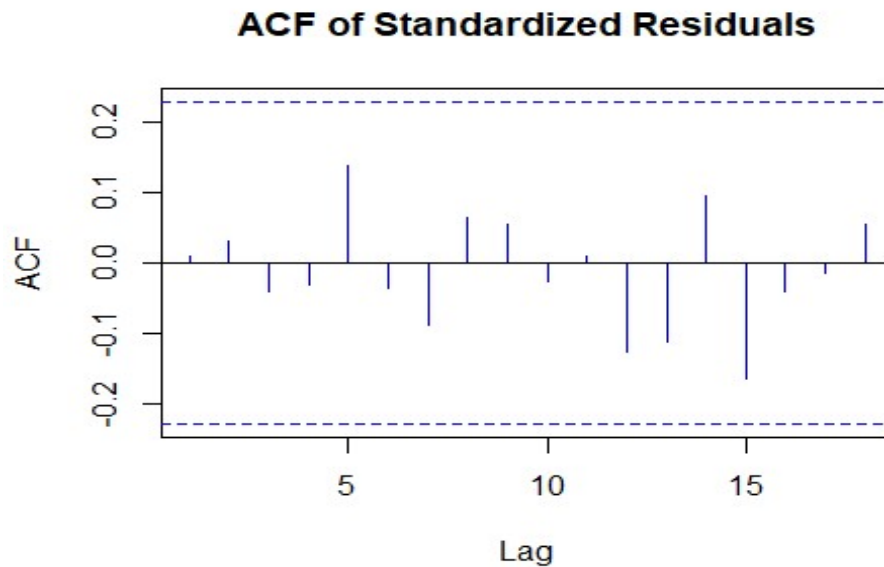
ACF of Standardized Residuals
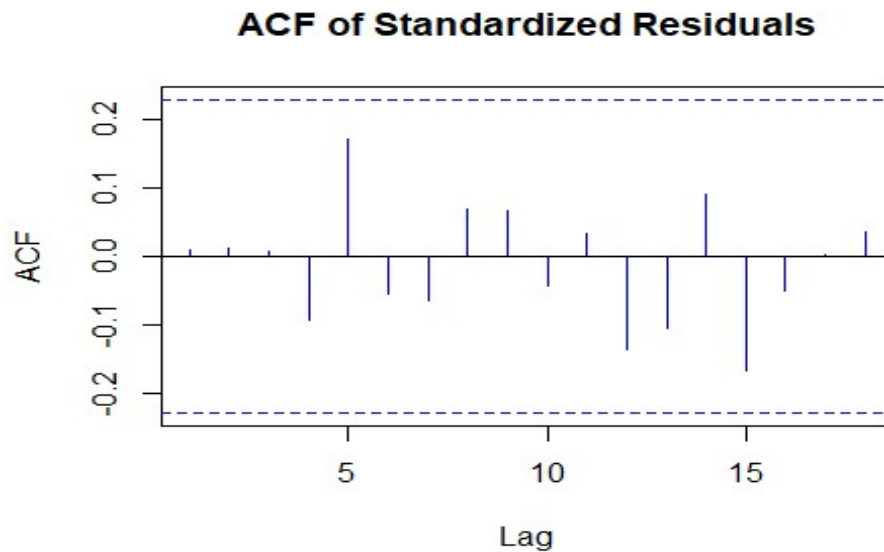
Fig 88: Autocorrelation plot for ARIMA(1,0,3)

Ljung-Box Test

Fig 89: Ljung-Box test for ARIMA(1,0,3)

Residual Analysis for ARIMA(1,0,3):

1. p-value is less than 0.05 for ar1, ma1 and ma2. Hence, significant.
2. For ma3 p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see some normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it can be considered as a good model but comparatively, ARIMA(1,0,2) performed better as the coefficients of ARIMA(1,0,3) are less significant than the coefficients of ARIMA(1,0,2).

ARIMA (1, 0, 2) is best model compared to over-fitting models. Thereby, we are good to use (1, 0, 2) order in ARIMA model.

Absolute residuals to find the order for GARCH model,
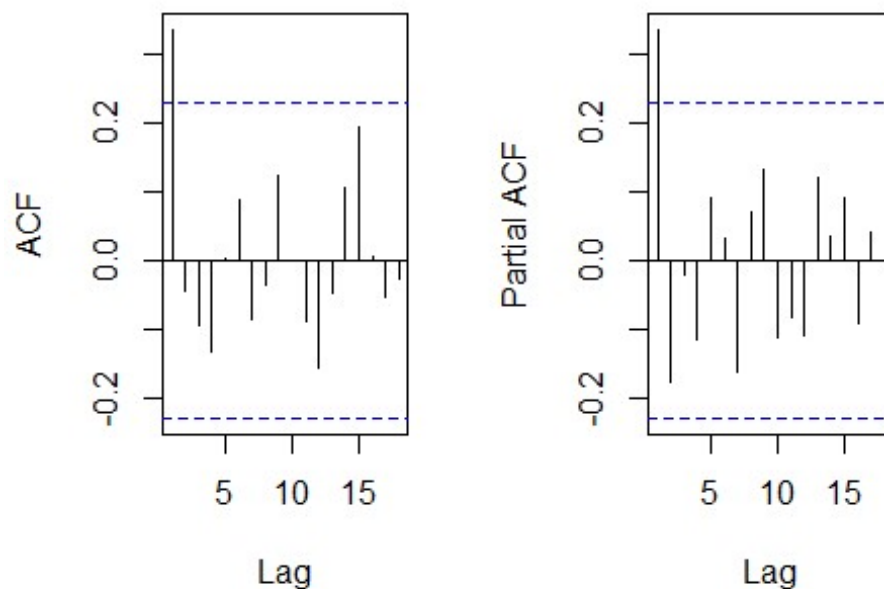
Fig 90: ACF and PCF plots for absolute residuals

Possible orders from ACF and PACF, (0, 1), (1, 0), and (1, 1)

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

Possible orders from EACF are (0, 1), (0, 2), (1, 1), and (1, 2).

```
## Warning in leaps.setup(x, y, wt = wt, nbest = nbest, nvmax = nvmax, force.
in =
## force.in, : 8 linear dependencies found

## Reordering variables and trying again:
```
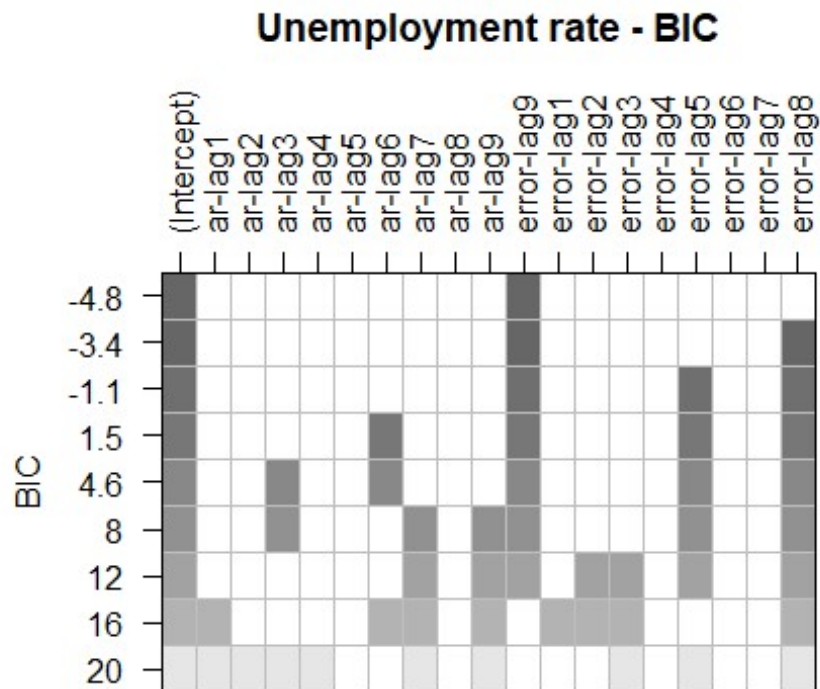

Fig 91: BIC table for absolute residuals

Possible orders from BIC are (0, 9) and (0, 8).

Overall possible ARMA orders are (0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (0, 9) and (0, 8).

GARCH orders for above orders:

max(p,q) = 0 and q = 1 ==> max(p,1) = 0 and q = 1 ==> No models can be identified.
max(p,q) = 0 and q = 2 ==> max(p,2) = 0 and q = 2 ==> No models can be identified.
max(p,q) = 1 and q = 0 ==> max(p,0) = 1 and q = 0 ==> p can only be 1.
max(p,q) = 1 and q = 1 ==> max(p,1) = 1 and q = 1 ==> p can only be 1.

max(p,q) = 1 and q = 2 ==> max(p,2) = 1 and q = 2 ==> No models can be identified.
max(p,q) = 0 and q = 9 ==> max(p,9) = 0 and q = 9 ==> No models can be identified.
max(p,q) = 0 and q = 8 ==> max(p,8) = 0 and q = 8 ==> No models can be identified.

GARCH models - GARCH(1, 0), GARCH(1, 1).

Squared residuals to find the order for GARCH model,
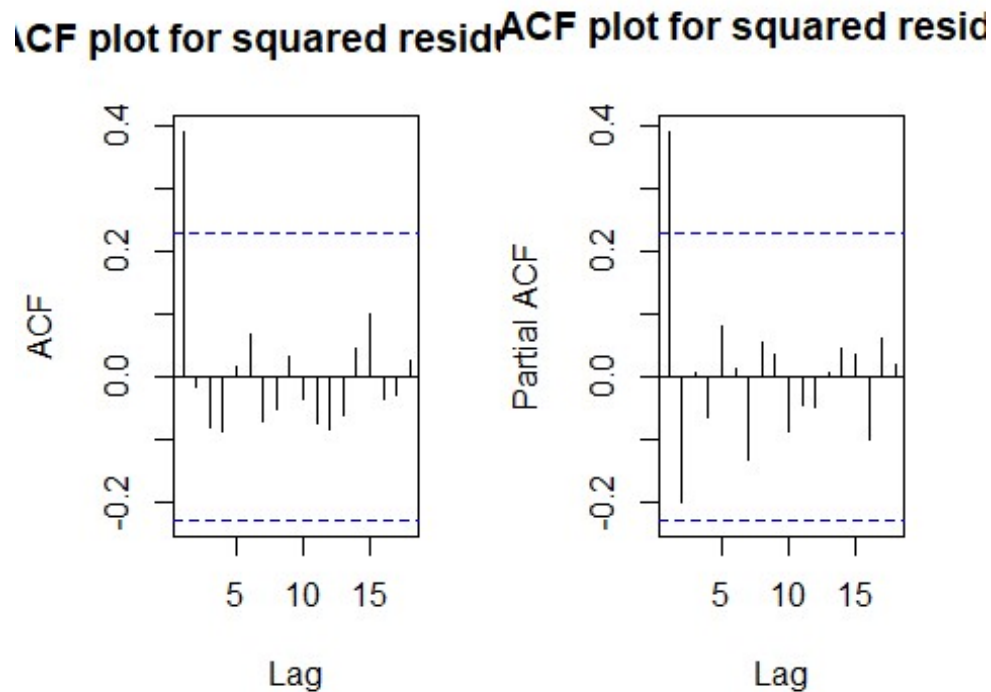
Fig 92: ACF and PACF plots for squared residuals

Possible orders from ACF and PACF, (0, 1), (1, 0), and (1, 1)

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

Possible orders from EACF are (0, 1), (0, 2), (1, 1), and (1, 2).

```
## Warning in leaps.setup(x, y, wt = wt, nbest = nbest, nvmax = nvmax, force.
in =
## force.in, : 7 linear dependencies found

## Reordering variables and trying again:
```
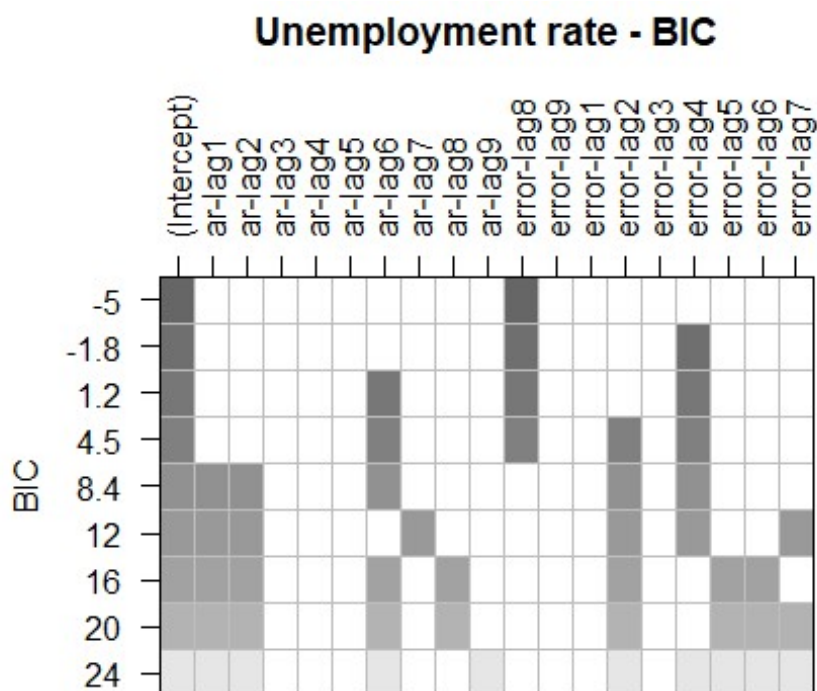

Fig 93: BIC table for squared residuals

Possible orders from BIC are (0, 8), (0, 4) and (6, 4).

Overall possible ARMA orders are (0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (0, 8), (0, 4) and (6, 4).

GARCH orders for above orders:

max(p,q) = 0 and q = 1 ==> max(p,1) = 0 and q = 1 ==> No models can be identified.
max(p,q) = 0 and q = 2 ==> max(p,2) = 0 and q = 2 ==> No models can be identified.
max(p,q) = 1 and q = 0 ==> max(p,0) = 1 and q = 0 ==> p can only be 1.
max(p,q) = 1 and q = 1 ==> max(p,1) = 1 and q = 1 ==> p can only be 1.
max(p,q) = 1 and q = 2 ==> max(p,2) = 1 and q = 2 ==> No models can be identified.
max(p,q) = 0 and q = 8 ==> max(p,8) = 0 and q = 8 ==> No models can be identified.
max(p,q) = 0 and q = 4 ==> max(p,4) = 0 and q = 4 ==> No models can be identified.
max(p,q) = 6 and q = 4 ==> max(p,4) = 6 and q = 4 ==> p can only be 6.

GARCH models - GARCH(1, 0), GARCH(1, 1) and GARCH(6, 4).

Oveall GARCH models are, GARCH(1, 0), GARCH(1, 1) and GARCH(6, 4).

Combining GARCH model with ARMA models we get ARMA + GARCH models, i.e.,
{ARMA(1,2) + GARCH(1,0), ARMA(1,2) + GARCH(1,1), ARMA(1,2) + GARCH(6,4)}

ARMA(1,2) + GARCH(1,0):

```
## Warning in .sgarchfit(spec = spec, data = data, out.sample = out.sample, :
## ugarchfit-->waring: using less than 100 data
##   points for estimation
##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : sGARCH(0,1)
## Mean Model   : ARFIMA(1,0,2)
## Distribution : norm
##
## Optimal Parameters
## -----------------------------------
##         Estimate  Std. Error    t value Pr(>|t|)
## ar1     0.82084    0.131580    6.238327 0.000000
## ma1    -0.93022    0.155780   -5.971397 0.000000
## ma2     0.39653    0.123439    3.212397 0.001316
## omega   0.00000    0.000358    0.000003 0.999998
## beta1   0.99594    0.004370  227.883527 0.000000
##
## Robust Standard Errors:
##         Estimate  Std. Error    t value Pr(>|t|)
## ar1     0.82084    0.167855    4.890153 0.000001
## ma1    -0.93022    0.177523   -5.240009 0.000000
## ma2     0.39653    0.254318    1.559204 0.118948
## omega   0.00000    0.000394    0.000003 0.999998
## beta1   0.99594    0.007069  140.881139 0.000000
##
## LogLikelihood : -22.3546
##
## Information Criteria
## -----------------------------------
##
## Akaike        0.74944
## Bayes         0.90632
## Shibata       0.74084
## Hannan-Quinn 0.81196
##
## Weighted Ljung-Box Test on Standardized Residuals
## -----------------------------------
##                              statistic p-value
```

```
## Lag[1]                          0.008877  0.9249
## Lag[2*(p+q)+(p+q)-1][8]    1.936403  1.0000
## Lag[4*(p+q)+(p+q)-1][14]   3.850541  0.9803
## d.o.f=3
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ------------------------------------
##                           statistic  p-value
## Lag[1]                        10.21 0.001399
## Lag[2*(p+q)+(p+q)-1][2]      10.24 0.001649
## Lag[4*(p+q)+(p+q)-1][5]      10.75 0.005888
## d.o.f=1
##
## Weighted ARCH LM Tests
## ------------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[2]   0.06761 0.500 2.000  0.7948
## ARCH Lag[4]   0.63239 1.397 1.611  0.8251
## ARCH Lag[6]   0.81633 2.222 1.500  0.9277
##
## Nyblom stability test
## ------------------------------------
## Joint Statistic:  0.9053
## Individual Statistics:
## ar1   0.2321
## ma1   0.4364
## ma2   0.1105
## omega 0.1233
## beta1 0.1304
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:        1.28 1.47 1.88
## Individual Statistic:   0.35 0.47 0.75
##
## Sign Bias Test
## ------------------------------------
##                    t-value       prob sig
## Sign Bias           1.6071 1.127e-01
## Negative Sign Bias  0.7077 4.815e-01
## Positive Sign Bias  4.8561 7.346e-06 ***
## Joint Effect       24.3336 2.128e-05 ***
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## ------------------------------------
##   group statistic p-value(g-1)
## 1    20     17.14       0.5806
## 2    30     29.33       0.4480
## 3    40     32.75       0.7493
```

```
## 4      50        34.53          0.9415
##
##
## Elapsed time : 0.109571
```

Plot of Residuals over Time
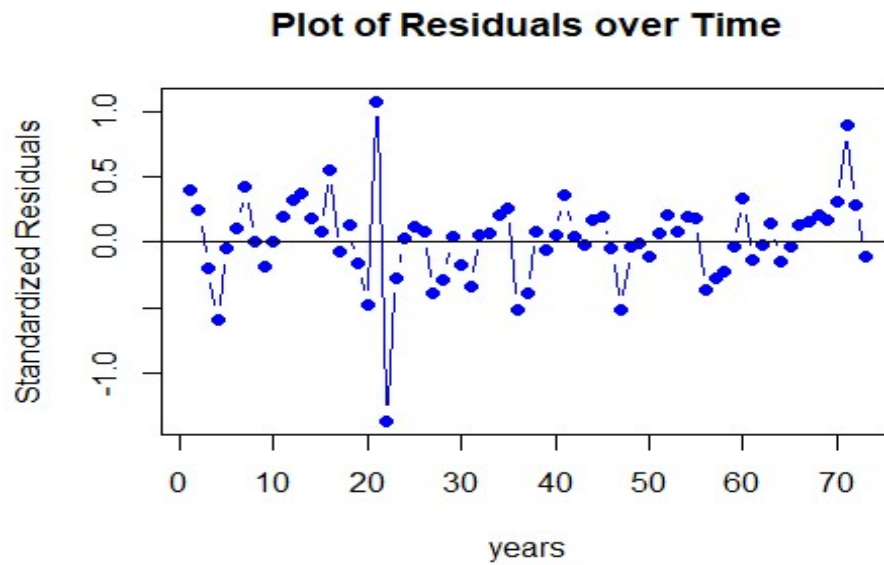
Fig 94: Residual's plot over time for ARMA(1,2) + GARCH(1,0)
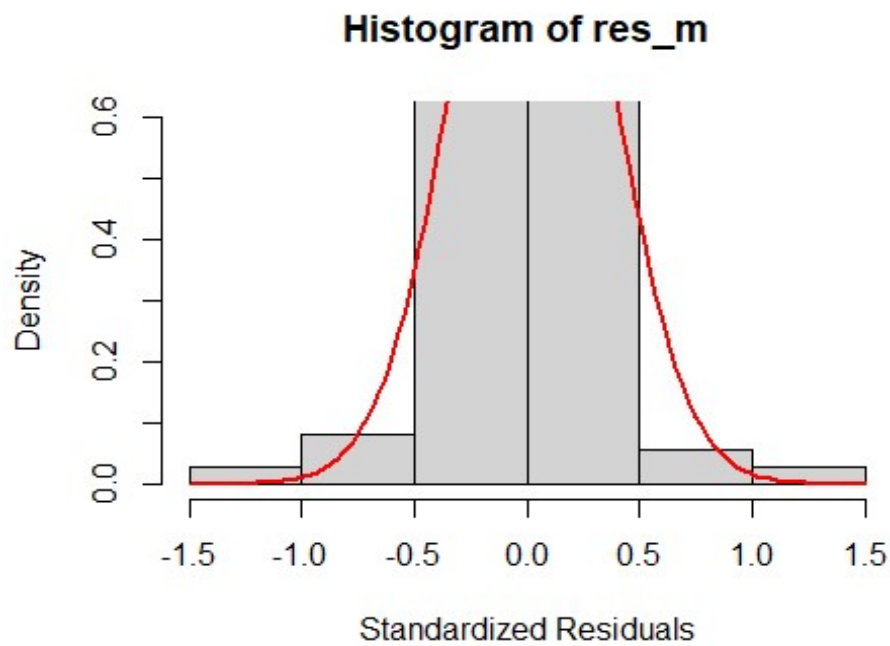
Histogram of res_m

Fig 95: Histogram of ARMA(1,2) + GARCH(1,0)

Normal Q-Q Plot

Fig 96: Normal Q-Q plot for ARMA(1,2) + GARCH(1,0)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.92213, p-value = 0.000241

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

ACF of Standardized Residuals

Fig 97: Autocorrelation plot for ARMA(1,2) + GARCH(1,0)

Ljung-Box Test

Fig 98: Ljung-Box test for ARMA(1,2) + GARCH(1,0)

Residual Analysis for ARMA(1,2) + GARCH(1,0):

1. p-value is less than 0.05 for ar1, ma1, ma2, and beta1. Hence, significant.
2. In Residual plot, the line is passing through 0.
3. From normal distribution curve, qq plot and Shapiro Walk test, we can see normality in the data.
4. There are no significant lags in Autocorrelation plot.
5. Ljung box test shows good results as all the dots are above the dotted line.

Overall, it can be considered as a best model.

ARMA(1,2) + GARCH(1,1):

```
## Warning in .sgarchfit(spec = spec, data = data, out.sample = out.sample, :
## ugarchfit-->waring: using less than 100 data
##  points for estimation

##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : sGARCH(1,1)
## Mean Model   : ARFIMA(1,0,2)
## Distribution : norm
##
## Optimal Parameters
## ------------------------------------
##         Estimate  Std. Error    t value Pr(>|t|)
## ar1      0.82084    0.132559   6.192249 0.000000
## ma1     -0.93022    0.157979  -5.888266 0.000000
## ma2      0.39653    0.131092   3.024848 0.002488
## omega    0.00000    0.000395   0.000017 0.999986
## alpha1   0.00000    0.009565   0.000000 1.000000
## beta1    0.99594    0.008947 111.314582 0.000000
##
## Robust Standard Errors:
##         Estimate  Std. Error    t value Pr(>|t|)
## ar1      0.82084    0.183733   4.467567 0.000008
## ma1     -0.93022    0.203951  -4.561001 0.000005
## ma2      0.39653    0.306328   1.294476 0.195501
## omega    0.00000    0.000310   0.000022 0.999983
## alpha1   0.00000    0.012121   0.000000 1.000000
## beta1    0.99594    0.006366 156.439550 0.000000
##
## LogLikelihood : -22.3546
##
## Information Criteria
```

```
## -------------------------------------
## 
## Akaike        0.77684
## Bayes         0.96510
## Shibata       0.76465
## Hannan-Quinn  0.85186
## 
## Weighted Ljung-Box Test on Standardized Residuals
## -------------------------------------
##                              statistic p-value
## Lag[1]                         0.008877  0.9249
## Lag[2*(p+q)+(p+q)-1][8]        1.936403  1.0000
## Lag[4*(p+q)+(p+q)-1][14]       3.850541  0.9803
## d.o.f=3
## H0 : No serial correlation
## 
## Weighted Ljung-Box Test on Standardized Squared Residuals
## -------------------------------------
##                              statistic  p-value
## Lag[1]                           10.21 0.001399
## Lag[2*(p+q)+(p+q)-1][5]          10.75 0.005888
## Lag[4*(p+q)+(p+q)-1][9]          11.31 0.025944
## d.o.f=2
## 
## Weighted ARCH LM Tests
## -------------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[3]    0.4224 0.500 2.000  0.5157
## ARCH Lag[5]    0.8240 1.440 1.667  0.7858
## ARCH Lag[7]    1.1479 2.315 1.543  0.8884
## 
## Nyblom stability test
## -------------------------------------
## Joint Statistic:  3.6216
## Individual Statistics:
## ar1    0.2321
## ma1    0.4364
## ma2    0.1105
## omega  0.1233
## alpha1 0.1514
## beta1  0.1304
## 
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:        1.49 1.68 2.12
## Individual Statistic:   0.35 0.47 0.75
## 
## Sign Bias Test
## -------------------------------------
##                    t-value      prob sig
## Sign Bias           1.6071 1.127e-01
```

```
## Negative Sign Bias  0.7077 4.815e-01
## Positive Sign Bias  4.8561 7.346e-06 ***
## Joint Effect       24.3336 2.128e-05 ***
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## ------------------------------------
##   group statistic  p-value(g-1)
## 1    20    17.14        0.5806
## 2    30    29.33        0.4480
## 3    40    32.75        0.7493
## 4    50    34.53        0.9415
##
##
## Elapsed time : 0.07097483
```

Plot of Residuals over Time
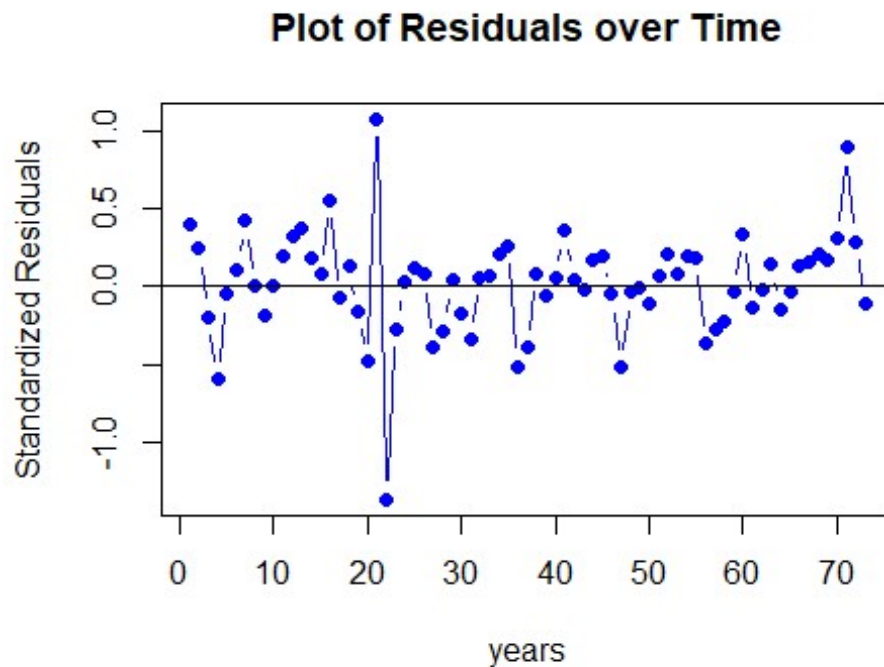
Fig 99: Residual's plot over time for ARMA(1,2) + GARCH(1,1)
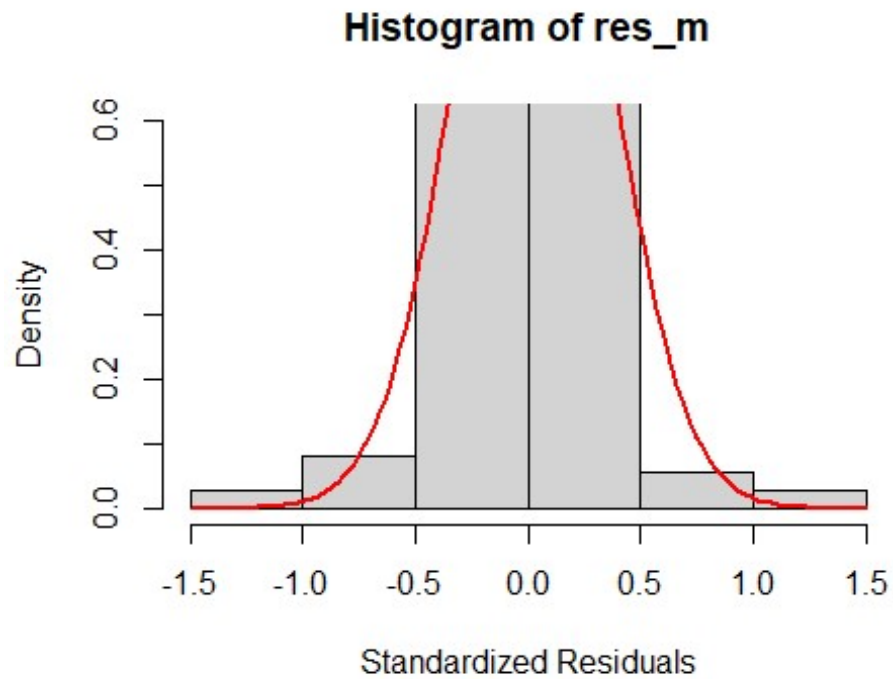
Histogram of res_m

Fig 100: Histogram of ARMA(1,2) + GARCH(1,1)

Normal Q-Q Plot
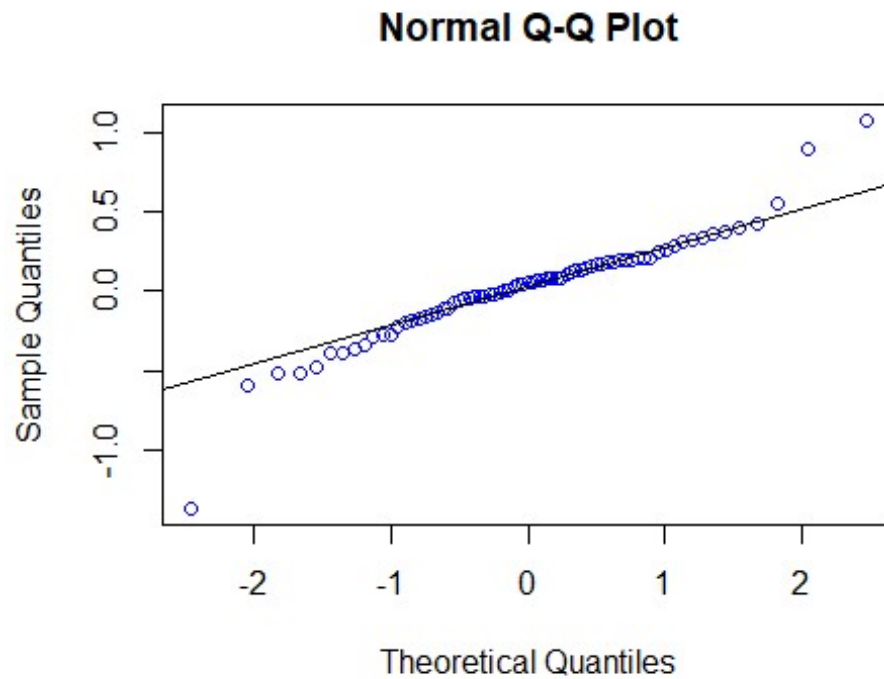

Fig 101: Normal Q-Q plot for ARMA(1,2) + GARCH(1,1)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.92213, p-value = 0.000241

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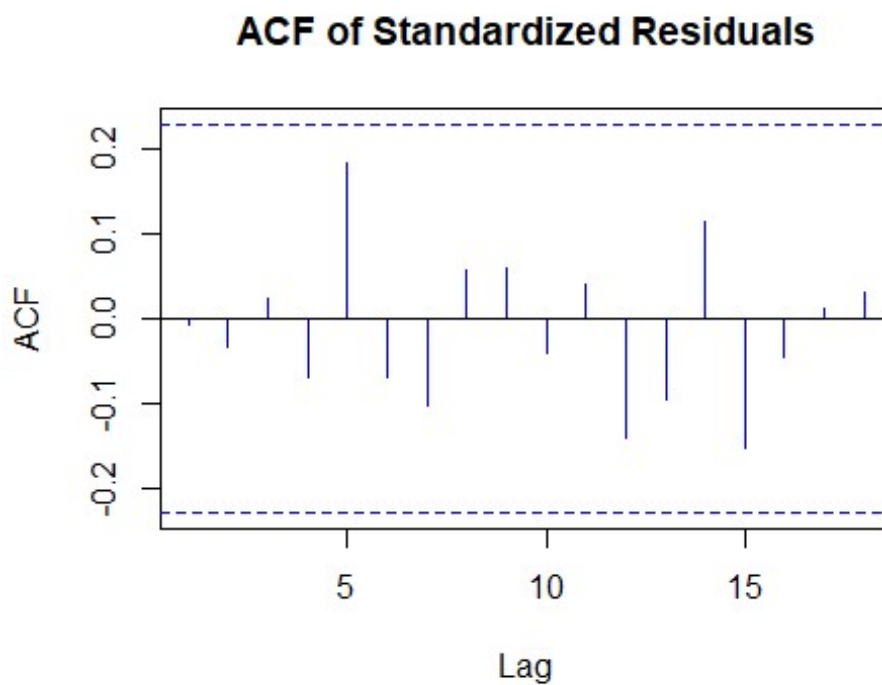

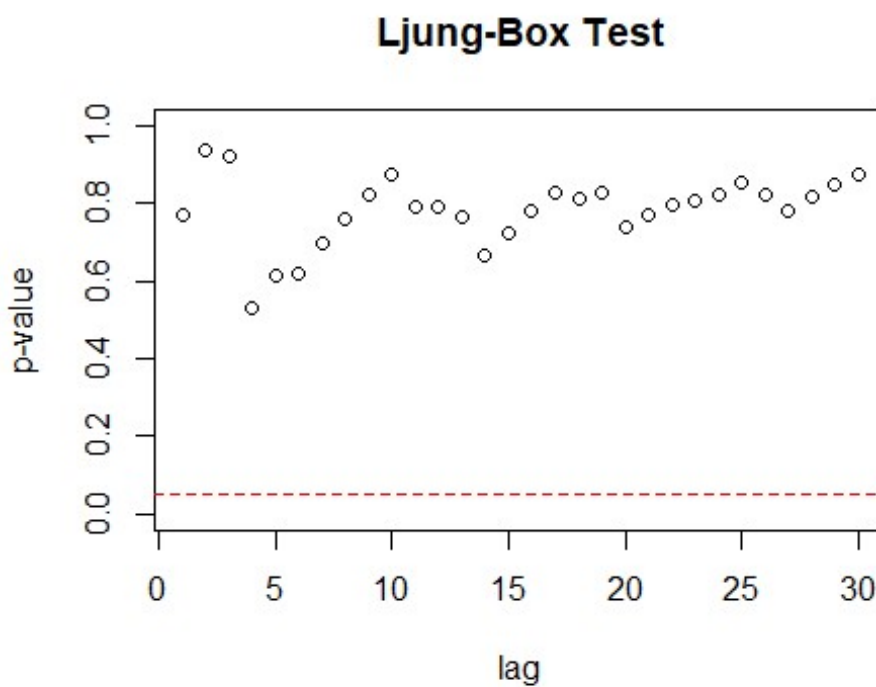Fig 102: Autocorrelation plot for ARMA(1,2) + GARCH(1,1)
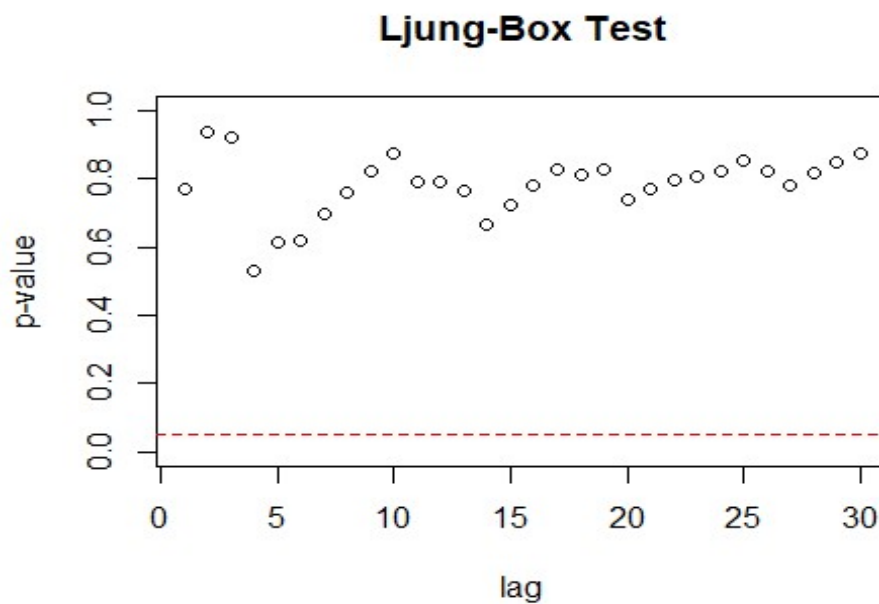
Fig 103: Ljung-Box test for ARMA(1,2) + GARCH(1,1)

Residual Analysis for ARMA(1,2) + GARCH(1,1):

1. p-value is less than 0.05 for ar1, ma1, ma2 and beta1. Hence, significant.
2. For alpha1, p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it is not considered as a good model since alpha1 is not significant.

ARMA(1,2) + GARCH(6,4):

```
## Warning in .sgarchfit(spec = spec, data = data, out.sample = out.sample, :
## ugarchfit-->waring: using less than 100 data
##  points for estimation

##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : sGARCH(4,6)
## Mean Model   : ARFIMA(1,0,2)
## Distribution : norm
##
## Optimal Parameters
## ------------------------------------
##         Estimate  Std. Error    t value Pr(>|t|)
## ar1     0.933069    0.068765  13.568977 0.000000
## ma1    -0.817909    0.107240  -7.626921 0.000000
## ma2     0.149800    0.102695   1.458692 0.144650
## omega   0.039661    0.010318   3.843870 0.000121
## alpha1  0.678364    0.001311 517.255633 0.000000
## alpha2  0.000000    0.144540   0.000000 1.000000
## alpha3  0.000000    0.137700   0.000000 1.000000
## alpha4  0.000000    0.177355   0.000000 1.000000
## beta1   0.000000    0.219305   0.000000 1.000000
## beta2   0.000000    0.195995   0.000000 1.000000
## beta3   0.000000    0.267271   0.000000 1.000000
## beta4   0.036393    0.047856   0.760465 0.446976
## beta5   0.000000    0.072914   0.000002 0.999998
## beta6   0.000000    0.049406   0.000001 1.000000
##
## Robust Standard Errors:
##         Estimate  Std. Error    t value Pr(>|t|)
## ar1     0.933069    0.046509  20.062237 0.000000
```

```
## ma1     -0.817909     0.091187 -8.969610 0.000000
## ma2      0.149800     0.082955  1.805806 0.070949
## omega    0.039661     0.012412  3.195293 0.001397
## alpha1   0.678364     0.023591 28.754763 0.000000
## alpha2   0.000000     0.119428  0.000000 1.000000
## alpha3   0.000000     0.087163  0.000000 1.000000
## alpha4   0.000000     0.152673  0.000000 1.000000
## beta1    0.000000     0.175900  0.000000 1.000000
## beta2    0.000000     0.125331  0.000000 1.000000
## beta3    0.000000     0.235662  0.000000 1.000000
## beta4    0.036393     0.025278  1.439680 0.149958
## beta5    0.000000     0.049348  0.000003 0.999997
## beta6    0.000000     0.037837  0.000001 0.999999
##
## LogLikelihood : -10.51211
##
## Information Criteria
## ---------------------------------
##
## Akaike        0.67156
## Bayes         1.11083
## Shibata       0.61266
## Hannan-Quinn 0.84662
##
## Weighted Ljung-Box Test on Standardized Residuals
## ---------------------------------
##                             statistic p-value
## Lag[1]                        0.002159  0.9629
## Lag[2*(p+q)+(p+q)-1][8]       3.404995  0.9720
## Lag[4*(p+q)+(p+q)-1][14]      7.003266  0.5596
## d.o.f=3
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ---------------------------------
##                             statistic p-value
## Lag[1]                         0.04532  0.8314
## Lag[2*(p+q)+(p+q)-1][29]       9.58378  0.9002
## Lag[4*(p+q)+(p+q)-1][49]      16.10311  0.9541
## d.o.f=10
##
## Weighted ARCH LM Tests
## ---------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[11]   0.08271 0.500 2.000  0.7737
## ARCH Lag[13]   0.92992 1.493 1.855  0.7985
## ARCH Lag[15]   1.66676 2.471 1.754  0.8577
##
## Nyblom stability test
## ---------------------------------
```

```
## Joint Statistic:  7.9713
## Individual Statistics:
## ar1    0.20140
## ma1    0.07810
## ma2    0.24034
## omega  0.08484
## alpha1 0.10355
## alpha2 0.87760
## alpha3 0.85646
## alpha4 0.92669
## beta1  0.69567
## beta2  0.84351
## beta3  0.70532
## beta4  0.05524
## beta5  0.11318
## beta6  0.70596
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:          3.08 3.34 3.9
## Individual Statistic:     0.35 0.47 0.75
##
## Sign Bias Test
## -----------------------------------
##                    t-value    prob sig
## Sign Bias          0.1315 0.8958
## Negative Sign Bias  0.3873 0.6997
## Positive Sign Bias  0.5094 0.6121
## Joint Effect       0.5648 0.9045
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## -----------------------------------
##   group statistic p-value(g-1)
## 1    20    13.85       0.7924
## 2    30    21.11       0.8551
## 3    40    41.52       0.3614
## 4    50    45.49       0.6161
##
##
## Elapsed time : 0.237272
```

Plot of Residuals over Time
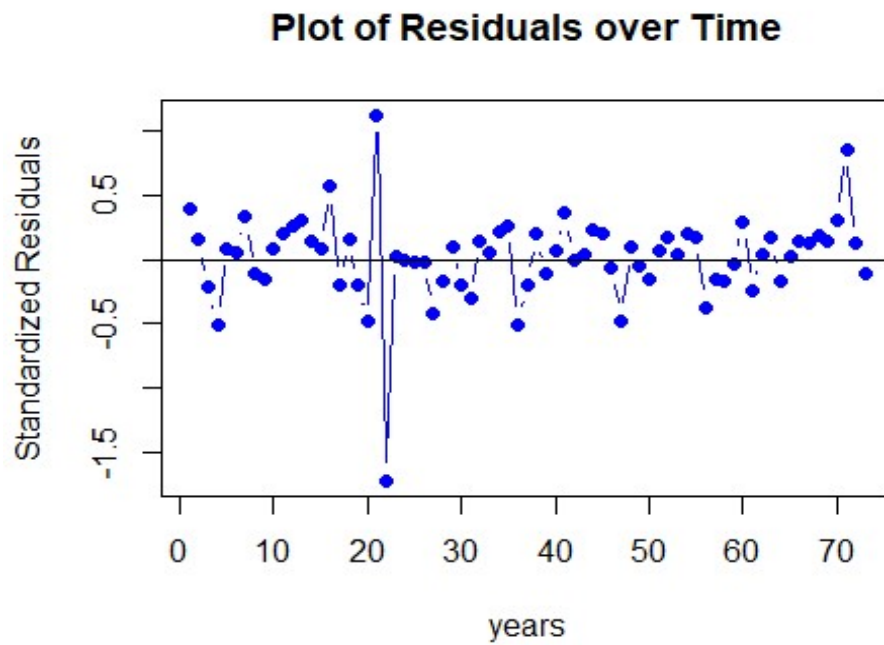
Fig 104: Residual's plot over time for ARMA(1,2) + GARCH(6,4)
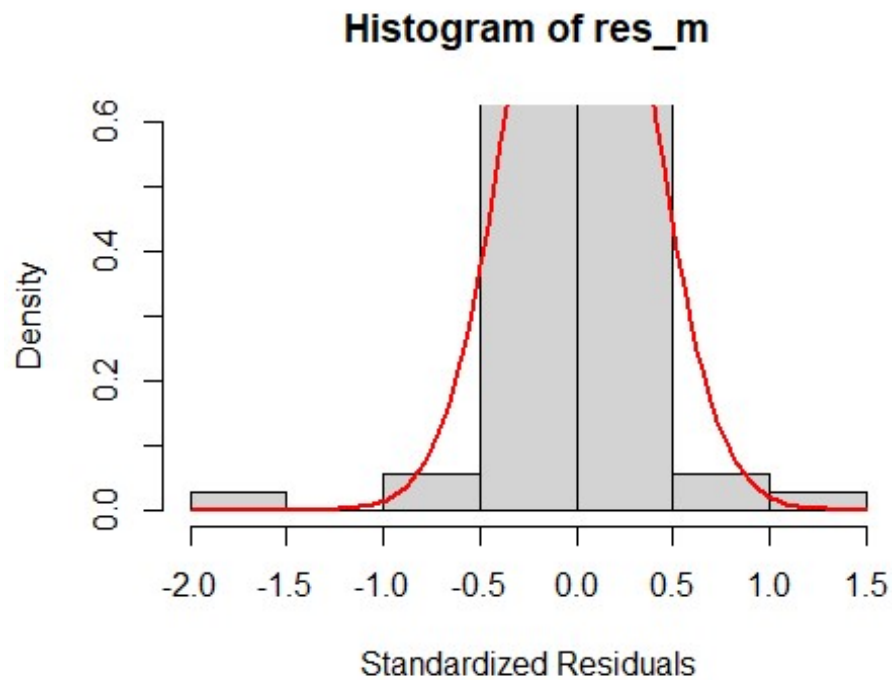
Histogram of res_m

Fig 105: Histogram of ARMA(1,2) + GARCH(6,4)
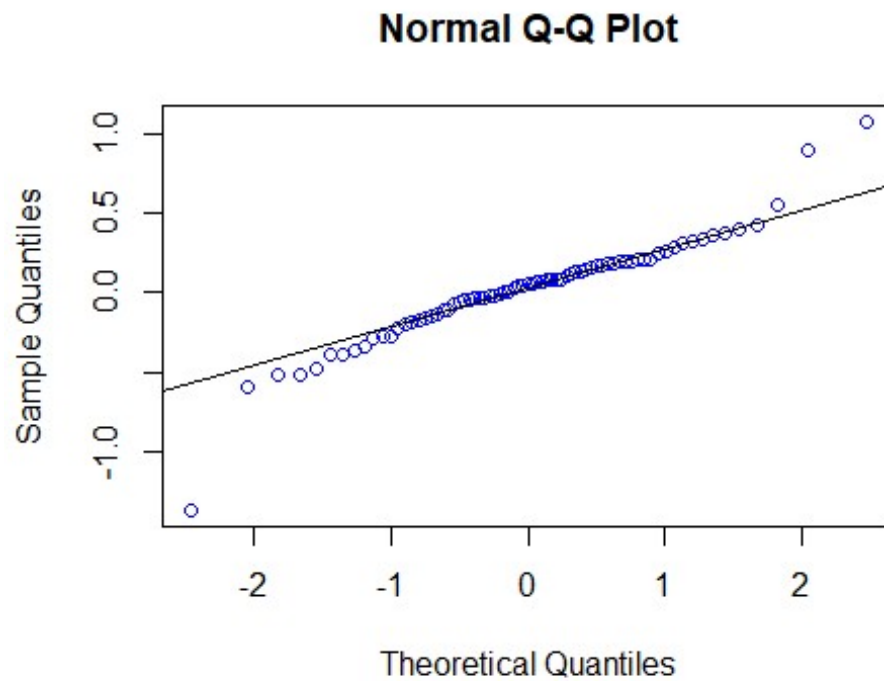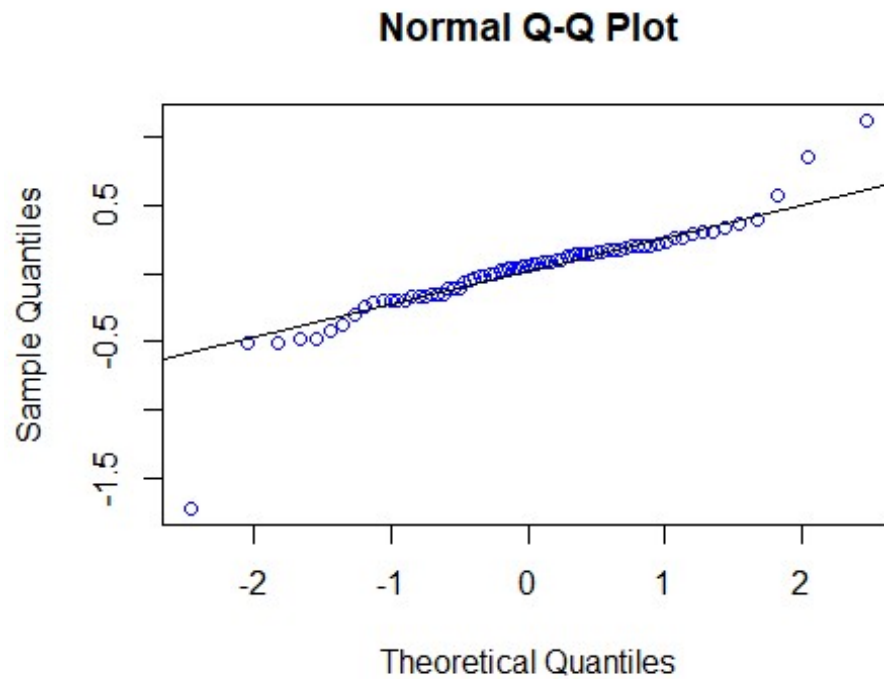
Normal Q-Q Plot

Fig 106: Normal Q-Q plot for ARMA(1,2) + GARCH(6,4)

```
## 
##  Shapiro-Wilk normality test
## 
## data:  res_m
## W = 0.84873, p-value = 3.946e-07

## Warning in (ra^2)/(n - (1:lag.max)): longer object length is not a multipl
e of
## shorter object length
```
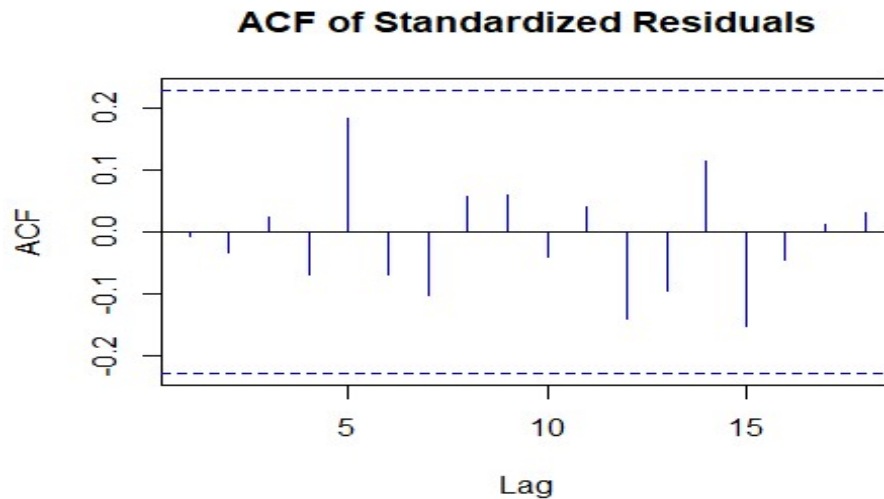
ACF of Standardized Residuals
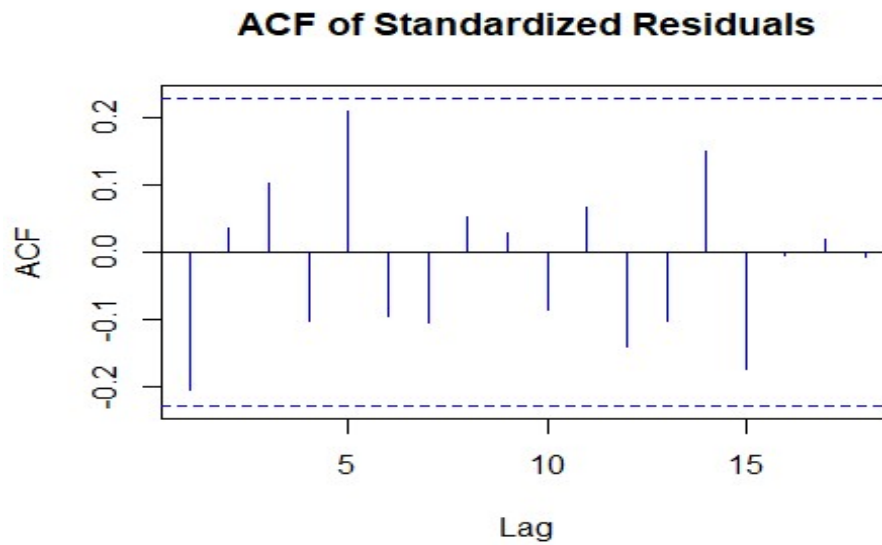
Fig 107: Autocorrelation plot for ARMA(1,2) + GARCH(6,4)
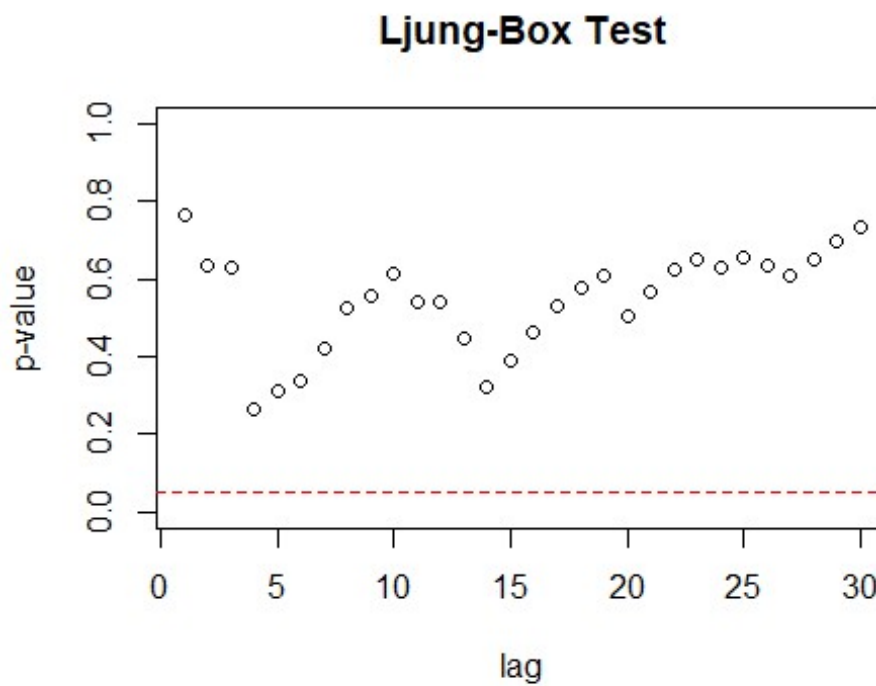
Ljung-Box Test

Fig 108: Ljung-Box test for ARMA(1,2) + GARCH(6,4)

Residual Analysis for ARMA(1,2) + GARCH(6,4):

1. p-value is less than 0.05 for ar1, ma1, and alpha1. Hence, significant.

2. For ma2, alpha2, alpha3, alpha4, beta1, beta2, beta3, beta4, beta5, and beta6, p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it is not considered as a good model since there are lot of insignificant values.

```
##          model       AIC       BIC
## 3 model_12_64 0.6715646 1.1108308
## 1 model_12_10 0.7494412 0.9063220
## 2 model_12_11 0.7768385 0.9650954

##          model       AIC       BIC
## 1 model_12_10 0.7494412 0.9063220
## 2 model_12_11 0.7768385 0.9650954
## 3 model_12_64 0.6715646 1.1108308
```

Overall, ARMA(1,2) + GARCH(1,0) model is best compared to others.

Now we are doing Over-fitting on the best model ARMA(1,2) + GARCH(1,0).

Over-fitting models are,

{ARMA(2, 2) + GARCH(1, 0), ARMA(1, 3) + GARCH(1, 0), ARMA(1, 2) + GARCH(2, 0), ARMA(1, 2) + GARCH(1, 1)}

ARMA(2, 2) + GARCH(1, 0):

```
## Warning in .sgarchfit(spec = spec, data = data, out.sample = out.sample, :
## ugarchfit-->waring: using less than 100 data
##  points for estimation

##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : sGARCH(0,1)
## Mean Model   : ARFIMA(2,0,2)
## Distribution : norm
##
## Optimal Parameters
## -----------------------------------
##        Estimate  Std. Error   t value Pr(>|t|)
## ar1     0.94283    0.297081   3.173649 0.001505
## ar2    -0.13705    0.302879  -0.452485 0.650919
```

```
## ma1    -1.01561    0.253721   -4.002853 0.000063
## ma2     0.48730    0.251722    1.935856 0.052885
## omega   0.00000    0.000329    0.000002 0.999999
## beta1   0.99564    0.004362  228.226931 0.000000
##
## Robust Standard Errors:
##          Estimate  Std. Error      t value Pr(>|t|)
## ar1      0.94283    0.394698    2.388736 0.016906
## ar2     -0.13705    0.266172   -0.514886 0.606633
## ma1     -1.01561    0.165077   -6.152324 0.000000
## ma2      0.48730    0.228218    2.135234 0.032742
## omega    0.00000    0.000222    0.000002 0.999998
## beta1    0.99564    0.007291  136.564587 0.000000
##
## LogLikelihood : -22.36174
##
## Information Criteria
## ---------------------------------
##
## Akaike        0.77703
## Bayes         0.96529
## Shibata       0.76484
## Hannan-Quinn 0.85206
##
## Weighted Ljung-Box Test on Standardized Residuals
## ---------------------------------
##                              statistic p-value
## Lag[1]                         0.03938  0.8427
## Lag[2*(p+q)+(p+q)-1][11]       2.80742  1.0000
## Lag[4*(p+q)+(p+q)-1][19]       5.96041  0.9723
## d.o.f=4
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ---------------------------------
##                              statistic  p-value
## Lag[1]                           9.728 0.001814
## Lag[2*(p+q)+(p+q)-1][2]          9.807 0.002150
## Lag[4*(p+q)+(p+q)-1][5]         10.295 0.007733
## d.o.f=1
##
## Weighted ARCH LM Tests
## ---------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[2]    0.1487 0.500 2.000  0.6998
## ARCH Lag[4]    0.6577 1.397 1.611  0.8169
## ARCH Lag[6]    0.8224 2.222 1.500  0.9267
##
## Nyblom stability test
## ---------------------------------
```

```
## Joint Statistic:  1.3369
## Individual Statistics:
## ar1   0.2364
## ar2   0.1031
## ma1   0.4392
## ma2   0.1075
## omega 0.1243
## beta1 0.1328
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:          1.49 1.68 2.12
## Individual Statistic:     0.35 0.47 0.75
##
## Sign Bias Test
## -----------------------------------
##                     t-value      prob sig
## Sign Bias             1.656 1.022e-01
## Negative Sign Bias    0.525 6.013e-01
## Positive Sign Bias    4.981 4.595e-06 ***
## Joint Effect         25.258 1.364e-05 ***
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## -----------------------------------
##   group statistic p-value(g-1)
## 1    20    17.68       0.5436
## 2    30    26.86       0.5791
## 3    40    32.75       0.7493
## 4    50    46.86       0.5602
##
##
## Elapsed time : 0.1221571
```

Plot of Residuals over Time
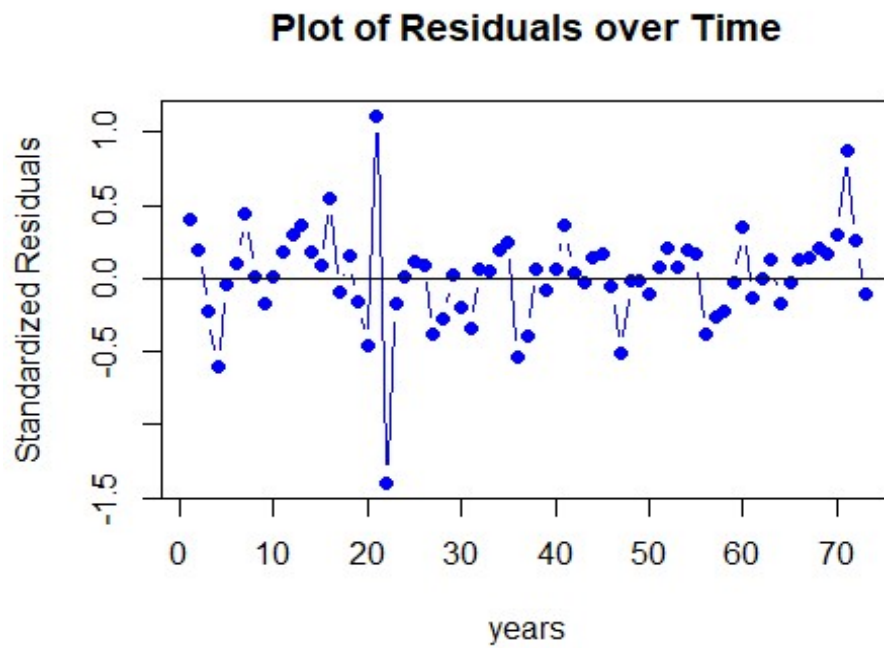
Fig 109: Residual's plot over time for ARMA(2, 2) + GARCH(1, 0)
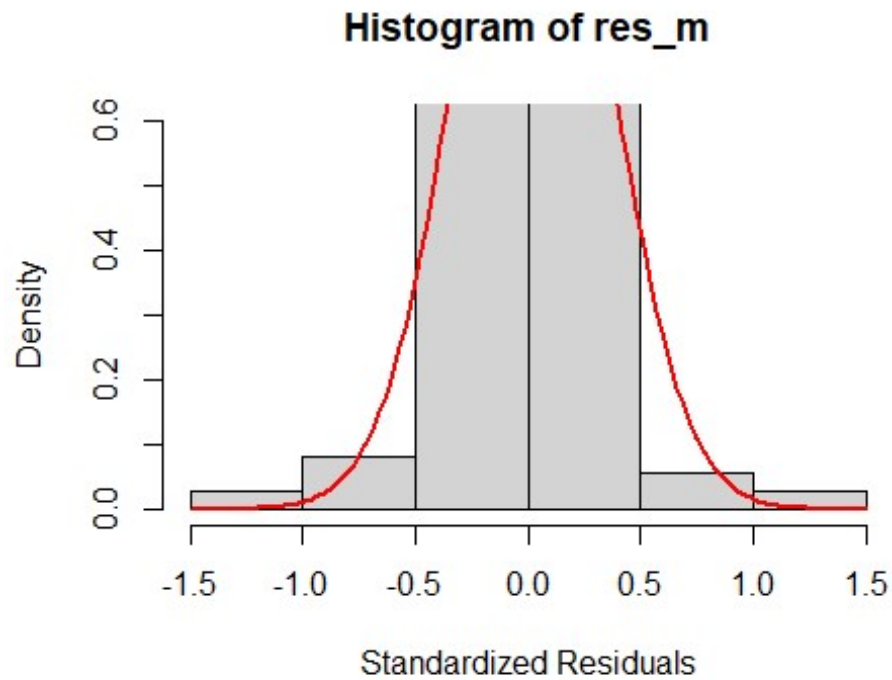
Histogram of res_m

Fig 110: Histogram of ARMA(2, 2) + GARCH(1, 0)
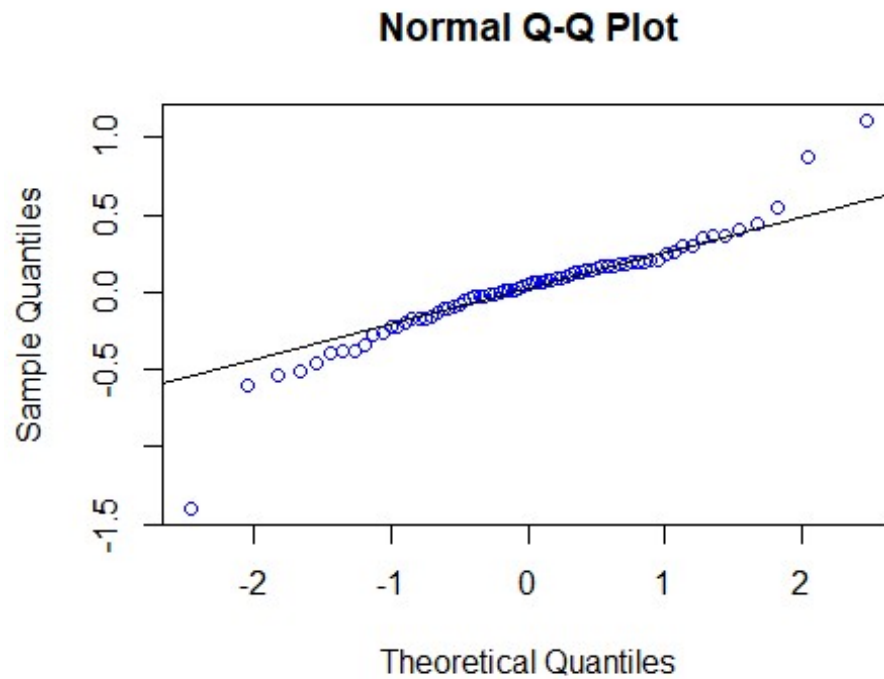
Normal Q-Q Plot

Fig 111: Normal Q-Q plot for ARMA(2, 2) + GARCH(1, 0)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.91365, p-value = 0.0001013

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

ACF of Standardized Residuals
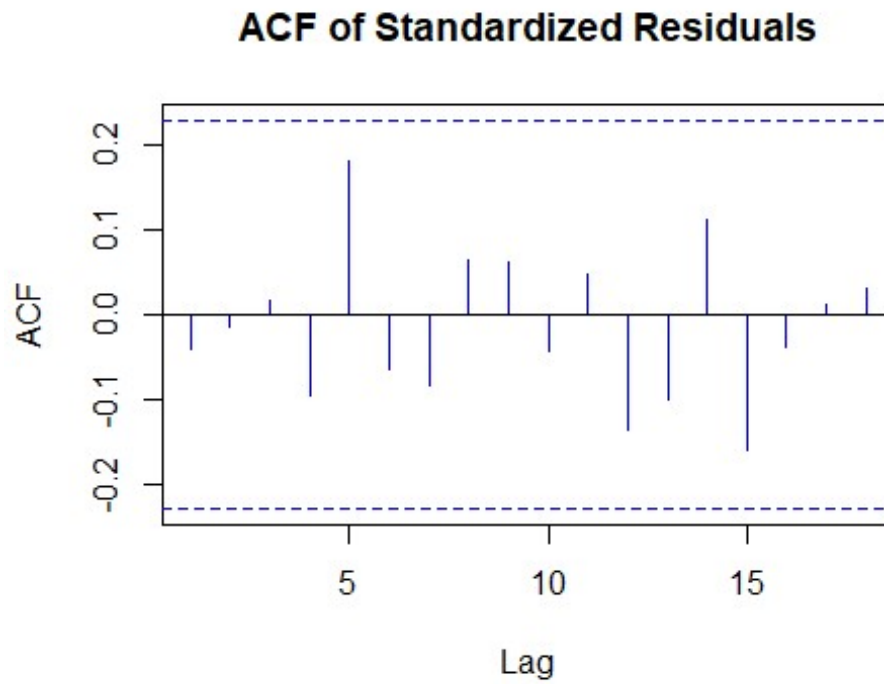
Fig 112: Autocorrelation plot for ARMA(2, 2) + GARCH(1, 0)
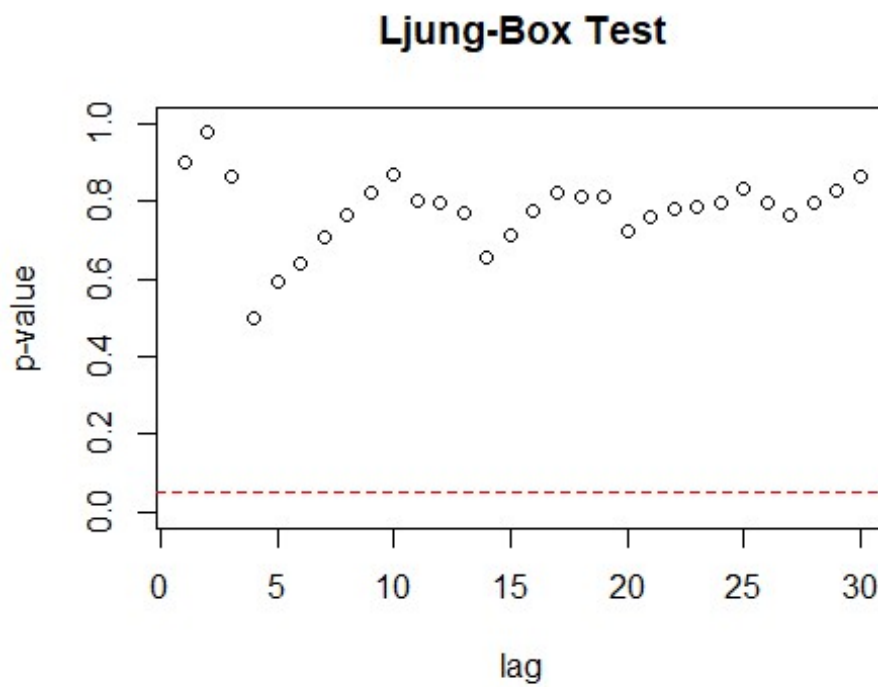
Ljung-Box Test

Fig 113: Ljung-Box test for ARMA(2, 2) + GARCH(1, 0)

Residual Analysis for ARMA(2,2) + GARCH(1,0):

1. p-value is less than 0.05 for ar1, ma1, and beta1. Hence, significant.
2. For ar2 and ma2, p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it is not considered as a good model.

ARMA(1, 3) + GARCH(1, 0):

```
## Warning in .sgarchfit(spec = spec, data = data, out.sample = out.sample, :
## ugarchfit-->waring: using less than 100 data
##  points for estimation

##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : sGARCH(0,1)
## Mean Model   : ARFIMA(1,0,3)
## Distribution : norm
##
## Optimal Parameters
## ------------------------------------
##         Estimate  Std. Error    t value Pr(>|t|)
## ar1     0.761049    0.238477   3.191295 0.001416
## ma1    -0.850391    0.270238  -3.146824 0.001651
## ma2     0.341165    0.164435   2.074774 0.038007
## ma3     0.079606    0.183995   0.432653 0.665267
## omega   0.000000    0.000351   0.000001 1.000000
## beta1   0.995597    0.004521 220.203122 0.000000
##
## Robust Standard Errors:
##         Estimate  Std. Error    t value Pr(>|t|)
## ar1     0.761049    0.278493   2.732742 0.006281
## ma1    -0.850391    0.305443  -2.784122 0.005367
## ma2     0.341165    0.272743   1.250864 0.210984
## ma3     0.079606    0.157622   0.505044 0.613528
## omega   0.000000    0.000174   0.000001 0.999999
## beta1   0.995597    0.007949 125.240678 0.000000
##
## LogLikelihood : -22.21719
##
```

```
## Information Criteria
## ------------------------------------
##
## Akaike       0.77307
## Bayes        0.96133
## Shibata      0.76088
## Hannan-Quinn 0.84810
##
## Weighted Ljung-Box Test on Standardized Residuals
## ------------------------------------
##                             statistic p-value
## Lag[1]                       0.004487  0.9466
## Lag[2*(p+q)+(p+q)-1][11]     2.647551  1.0000
## Lag[4*(p+q)+(p+q)-1][19]     5.691649  0.9812
## d.o.f=4
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ------------------------------------
##                             statistic  p-value
## Lag[1]                          9.728 0.001815
## Lag[2*(p+q)+(p+q)-1][2]         9.810 0.002146
## Lag[4*(p+q)+(p+q)-1][5]        10.297 0.007723
## d.o.f=1
##
## Weighted ARCH LM Tests
## ------------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[2]    0.1543 0.500 2.000  0.6944
## ARCH Lag[4]    0.6615 1.397 1.611  0.8157
## ARCH Lag[6]    0.8248 2.222 1.500  0.9263
##
## Nyblom stability test
## ------------------------------------
## Joint Statistic:  1.056
## Individual Statistics:
## ar1   0.2708
## ma1   0.4897
## ma2   0.1141
## ma3   0.1073
## omega 0.1194
## beta1 0.1271
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:          1.49 1.68 2.12
## Individual Statistic:     0.35 0.47 0.75
##
## Sign Bias Test
## ------------------------------------
##                     t-value      prob sig
```

```
## Sign Bias             1.6733 9.886e-02    *
## Negative Sign Bias  0.5953 5.536e-01
## Positive Sign Bias  4.8814 6.682e-06 ***
## Joint Effect        24.5262 1.940e-05 ***
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## -----------------------------------
##   group statistic p-value(g-1)
## 1    20     17.14        0.5806
## 2    30     21.11        0.8551
## 3    40     36.04        0.6056
## 4    50     44.12        0.6708
##
##
## Elapsed time : 0.199332
```
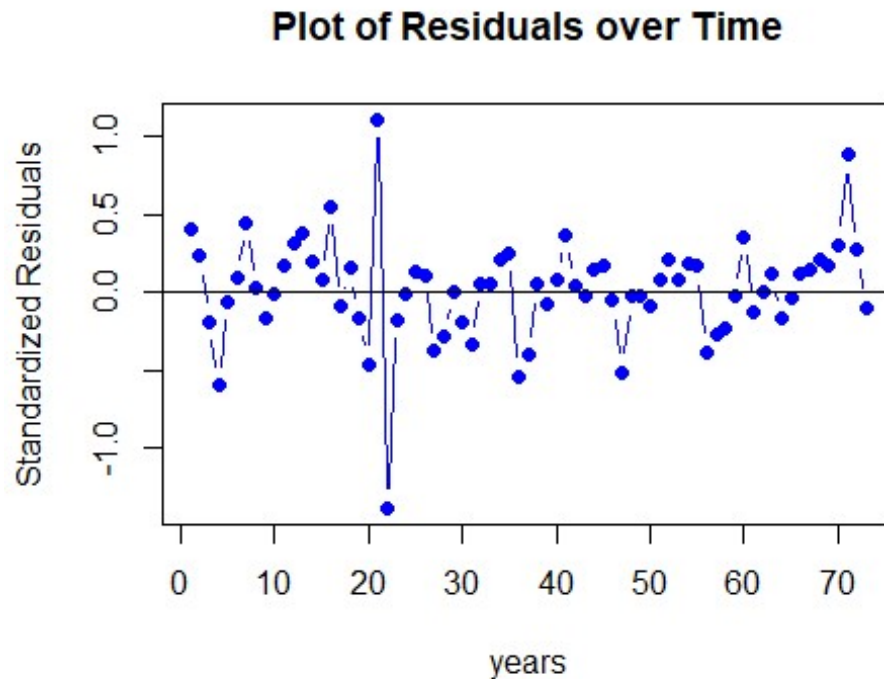

Fig 114: Residual's plot over time for ARMA(1, 3) + GARCH(1, 0)
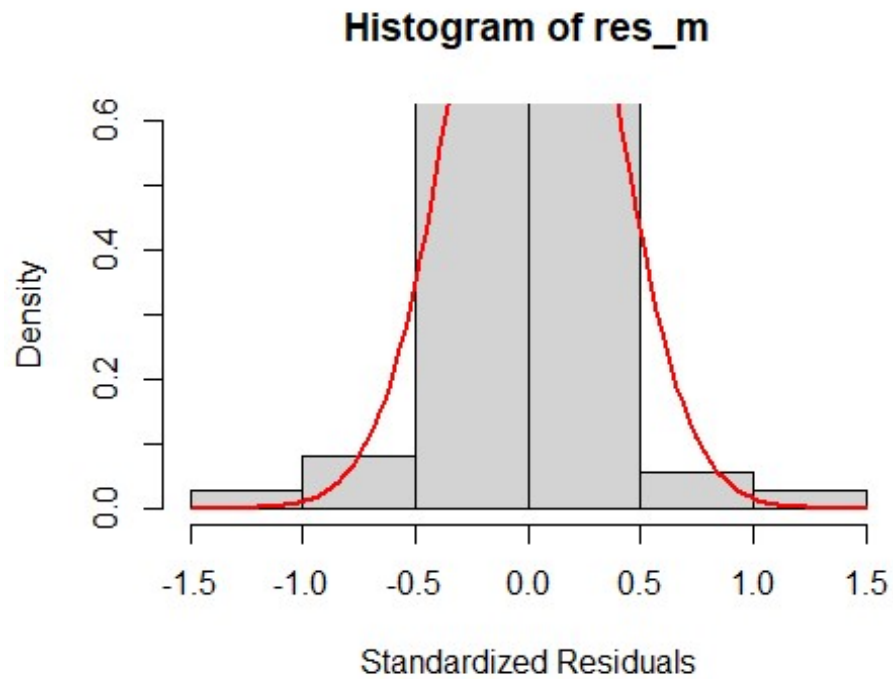
Histogram of res_m

Fig 115: Histogram of ARMA(1, 3) + GARCH(1, 0)
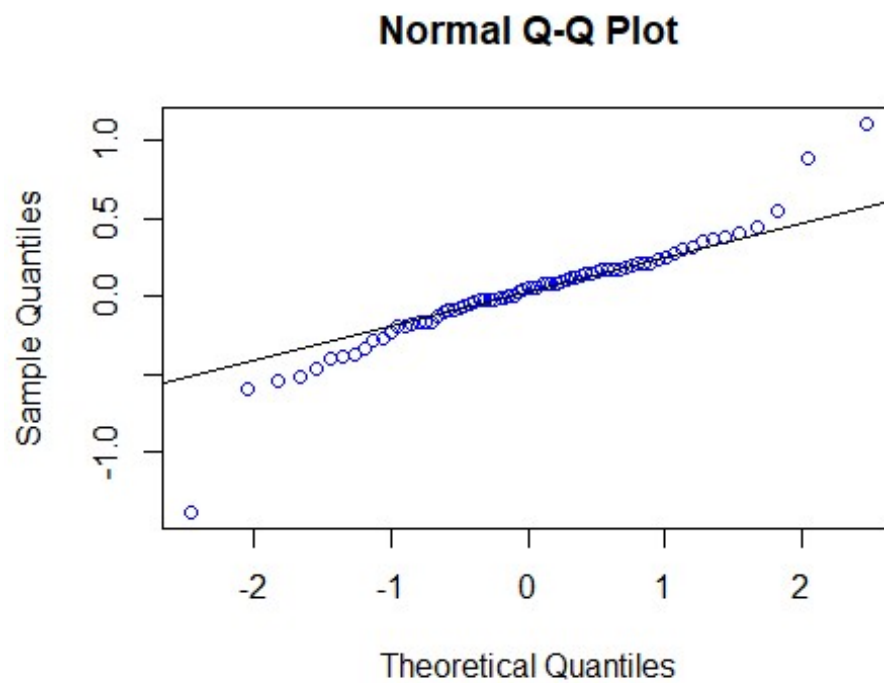
Normal Q-Q Plot

Fig 116: Normal Q-Q plot for ARMA(1, 3) + GARCH(1, 0)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.91881, p-value = 0.0001708

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

ACF of Standardized Residuals
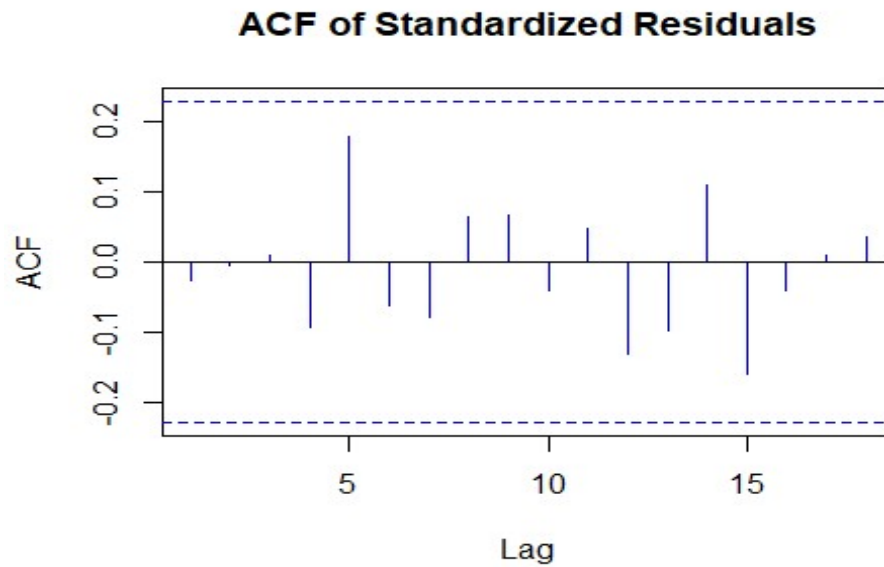
Fig 117: Autocorrelation plot for ARMA(1, 3) + GARCH(1, 0)
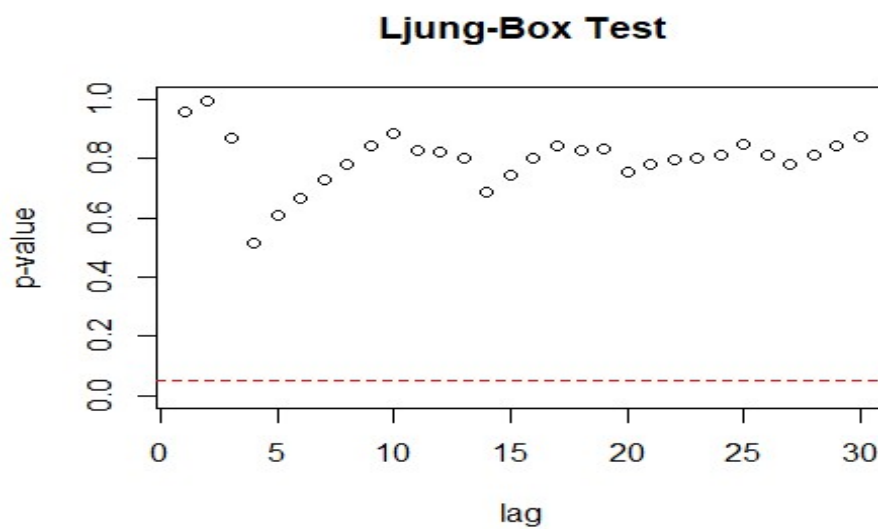
Ljung-Box Test

Fig 118: Ljung-Box test for ARMA(1, 3) + GARCH(1, 0)

95

Residual Analysis for ARMA(1, 3) + GARCH(1, 0):

1. p-value is less than 0.05 for ar1, ma1, ma2 and beta1. Hence, significant.
2. For ma3, p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it is not considered as a good model.

ARMA(1, 2) + GARCH(2, 0):

```
## Warning in .sgarchfit(spec = spec, data = data, out.sample = out.sample, :
## ugarchfit-->waring: using less than 100 data
##  points for estimation

##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : sGARCH(0,2)
## Mean Model   : ARFIMA(1,0,2)
## Distribution : norm
##
## Optimal Parameters
## ------------------------------------
##         Estimate  Std. Error  t value Pr(>|t|)
## ar1     0.820838    0.131565  6.23905 0.000000
## ma1    -0.930222    0.155666 -5.97575 0.000000
## ma2     0.396534    0.123416  3.21297 0.001314
## omega   0.000000    0.000234  0.00000 1.000000
## beta1   0.995901    0.017291 57.59595 0.000000
## beta2   0.000043    0.018326  0.00237 0.998109
##
## Robust Standard Errors:
##         Estimate  Std. Error    t value Pr(>|t|)
## ar1     0.820838    0.167413   4.903086 0.000001
## ma1    -0.930222    0.174017  -5.345586 0.000000
## ma2     0.396534    0.252653   1.569481 0.116536
## omega   0.000000    0.000032   0.000000 1.000000
## beta1   0.995901    0.001205 826.262057 0.000000
## beta2   0.000043    0.006971   0.006231 0.995028
##
## LogLikelihood : -22.3546
##
```

```
## Information Criteria
## ------------------------------------
##
## Akaike        0.77684
## Bayes         0.96510
## Shibata       0.76465
## Hannan-Quinn 0.85186
##
## Weighted Ljung-Box Test on Standardized Residuals
## ------------------------------------
##                              statistic p-value
## Lag[1]                        0.008877  0.9249
## Lag[2*(p+q)+(p+q)-1][8]       1.936403  1.0000
## Lag[4*(p+q)+(p+q)-1][14]      3.850542  0.9803
## d.o.f=3
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ------------------------------------
##                              statistic  p-value
## Lag[1]                           10.21 0.001399
## Lag[2*(p+q)+(p+q)-1][5]          10.75 0.005888
## Lag[4*(p+q)+(p+q)-1][9]          11.31 0.025944
## d.o.f=2
##
## Weighted ARCH LM Tests
## ------------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[3]    0.4224 0.500 2.000  0.5157
## ARCH Lag[5]    0.8240 1.440 1.667  0.7858
## ARCH Lag[7]    1.1479 2.315 1.543  0.8884
##
## Nyblom stability test
## ------------------------------------
## Joint Statistic:  1.3364
## Individual Statistics:
## ar1   0.2321
## ma1   0.4364
## ma2   0.1105
## omega 0.1233
## beta1 0.1304
## beta2 0.1304
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:         1.49 1.68 2.12
## Individual Statistic:    0.35 0.47 0.75
##
## Sign Bias Test
## ------------------------------------
##                       t-value      prob sig
```

```
## Sign Bias           1.6071 1.127e-01
## Negative Sign Bias  0.7077 4.815e-01
## Positive Sign Bias  4.8561 7.346e-06 ***
## Joint Effect        24.3336 2.128e-05 ***
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## ----------------------------------
##   group statistic p-value(g-1)
## 1    20     17.14       0.5806
## 2    30     29.33       0.4480
## 3    40     32.75       0.7493
## 4    50     34.53       0.9415
##
##
## Elapsed time : 0.09893608
```


Plot of Residuals over Time
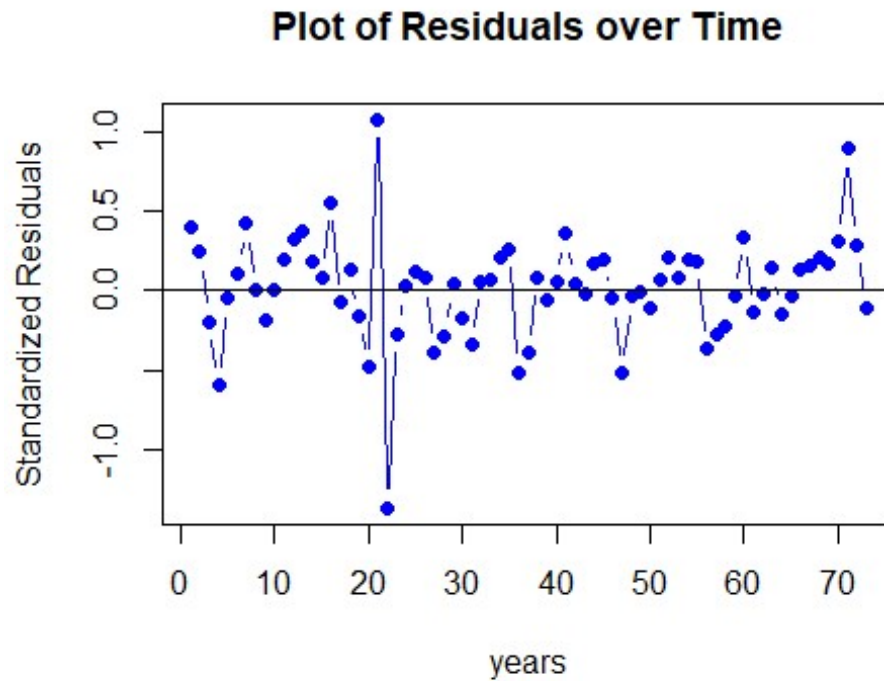
Fig 119: Residual's plot over time for ARMA(1, 2) + GARCH(2, 0)

Fig 120: Histogram of ARMA(1, 2) + GARCH(2, 0)

Fig 121: Normal Q-Q plot for ARMA(1, 2) + GARCH(2, 0)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.92213, p-value = 0.000241

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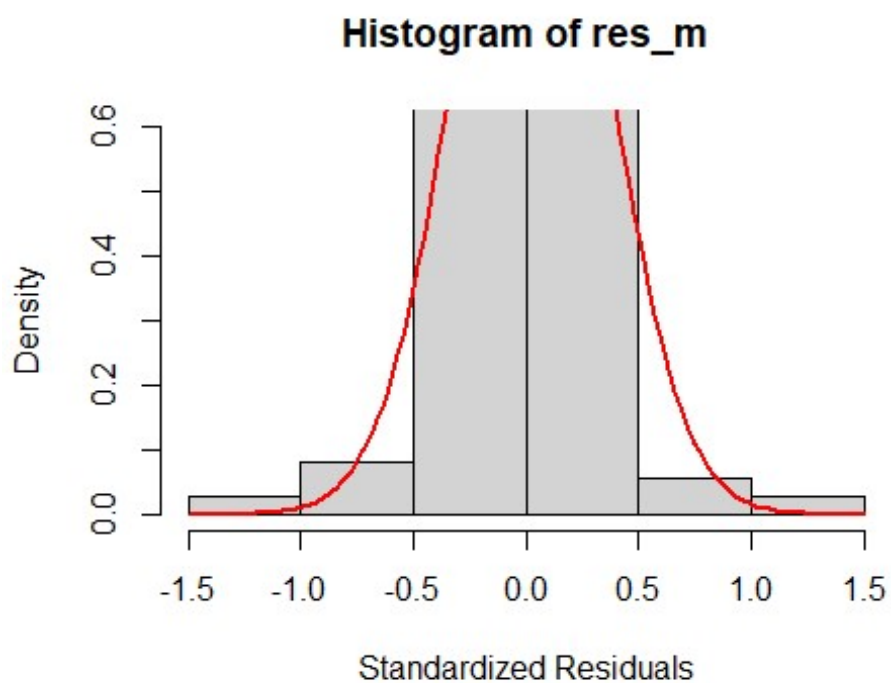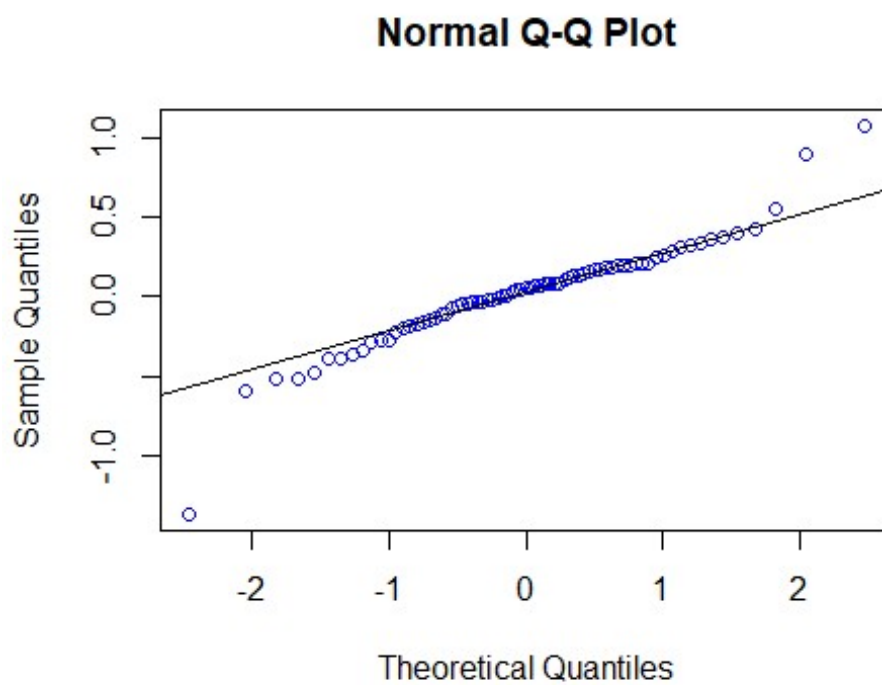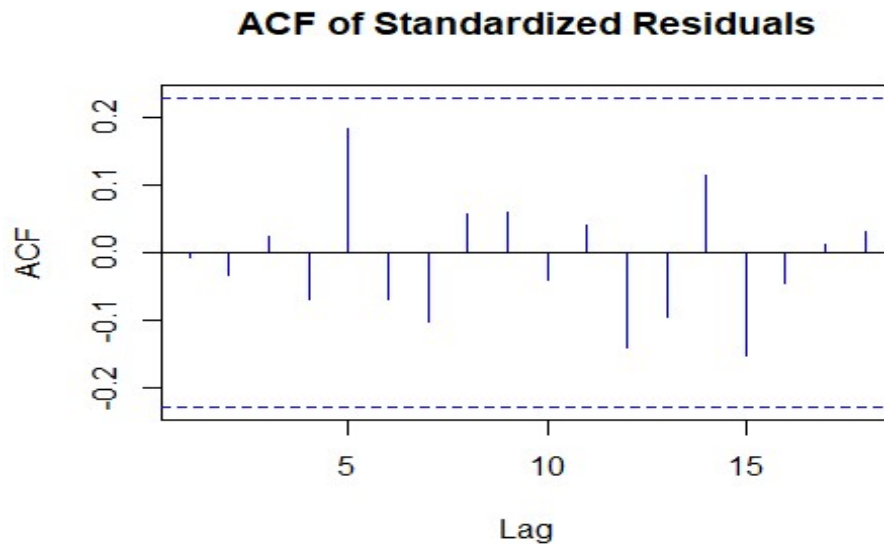

Fig 122: Autocorrelation plot for ARMA(1, 2) + GARCH(2, 0)
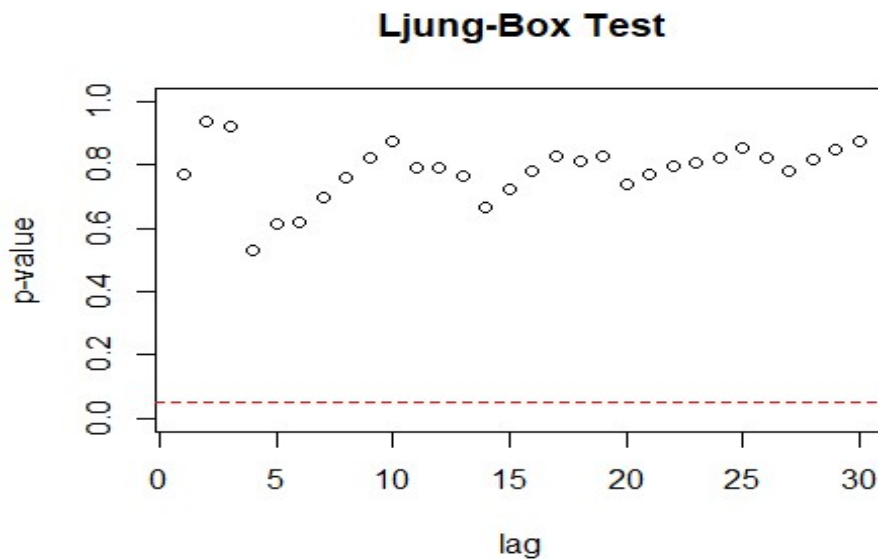
Fig 123: Ljung-Box test for ARMA(1, 2) + GARCH(2, 0)

Residual Analysis for ARMA(1, 2) + GARCH(2, 0):

1. p-value is less than 0.05 for ar1, ma1, ma2 and beta1. Hence, significant.
2. For beta2, p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it can be considered as a good model.

ARMA(1,2) + GARCH(1,1):

```
## Warning in .sgarchfit(spec = spec, data = data, out.sample = out.sample, :
## ugarchfit-->waring: using less than 100 data
##  points for estimation

##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : sGARCH(1,1)
## Mean Model   : ARFIMA(1,0,2)
## Distribution : norm
##
## Optimal Parameters
## ------------------------------------
##         Estimate  Std. Error    t value Pr(>|t|)
## ar1      0.82084    0.132559   6.192249 0.000000
## ma1     -0.93022    0.157979  -5.888266 0.000000
## ma2      0.39653    0.131092   3.024848 0.002488
## omega    0.00000    0.000395   0.000017 0.999986
## alpha1   0.00000    0.009565   0.000000 1.000000
## beta1    0.99594    0.008947 111.314582 0.000000
##
## Robust Standard Errors:
##         Estimate  Std. Error    t value Pr(>|t|)
## ar1      0.82084    0.183733   4.467567 0.000008
## ma1     -0.93022    0.203951  -4.561001 0.000005
## ma2      0.39653    0.306328   1.294476 0.195501
## omega    0.00000    0.000310   0.000022 0.999983
## alpha1   0.00000    0.012121   0.000000 1.000000
## beta1    0.99594    0.006366 156.439550 0.000000
##
## LogLikelihood : -22.3546
##
```

```
## Information Criteria
## ------------------------------------
##
## Akaike        0.77684
## Bayes         0.96510
## Shibata       0.76465
## Hannan-Quinn  0.85186
##
## Weighted Ljung-Box Test on Standardized Residuals
## ------------------------------------
##                               statistic p-value
## Lag[1]                         0.008877  0.9249
## Lag[2*(p+q)+(p+q)-1][8]        1.936403  1.0000
## Lag[4*(p+q)+(p+q)-1][14]       3.850541  0.9803
## d.o.f=3
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ------------------------------------
##                               statistic  p-value
## Lag[1]                            10.21 0.001399
## Lag[2*(p+q)+(p+q)-1][5]           10.75 0.005888
## Lag[4*(p+q)+(p+q)-1][9]           11.31 0.025944
## d.o.f=2
##
## Weighted ARCH LM Tests
## ------------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[3]    0.4224 0.500 2.000  0.5157
## ARCH Lag[5]    0.8240 1.440 1.667  0.7858
## ARCH Lag[7]    1.1479 2.315 1.543  0.8884
##
## Nyblom stability test
## ------------------------------------
## Joint Statistic:  3.6216
## Individual Statistics:
## ar1    0.2321
## ma1    0.4364
## ma2    0.1105
## omega  0.1233
## alpha1 0.1514
## beta1  0.1304
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:        1.49 1.68 2.12
## Individual Statistic:   0.35 0.47 0.75
##
## Sign Bias Test
## ------------------------------------
##                       t-value      prob sig
```

```
## Sign Bias             1.6071 1.127e-01
## Negative Sign Bias  0.7077 4.815e-01
## Positive Sign Bias  4.8561 7.346e-06 ***
## Joint Effect        24.3336 2.128e-05 ***
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## ------------------------------------
##    group statistic p-value(g-1)
## 1     20     17.14        0.5806
## 2     30     29.33        0.4480
## 3     40     32.75        0.7493
## 4     50     34.53        0.9415
##
##
## Elapsed time : 0.07993007
```
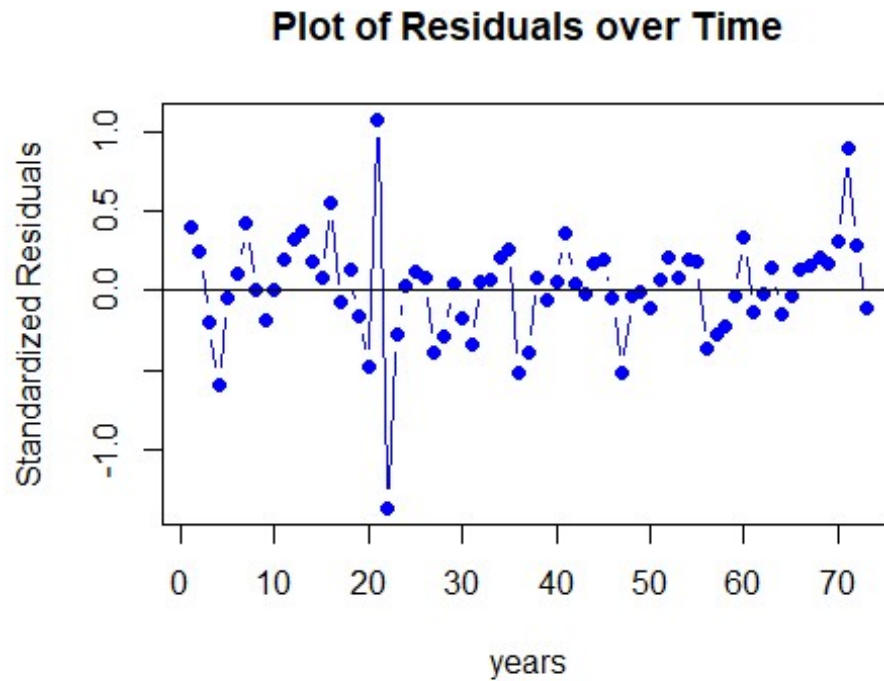

Fig 124: Residual's plot over time for ARMA(1,2) + GARCH(1,1)
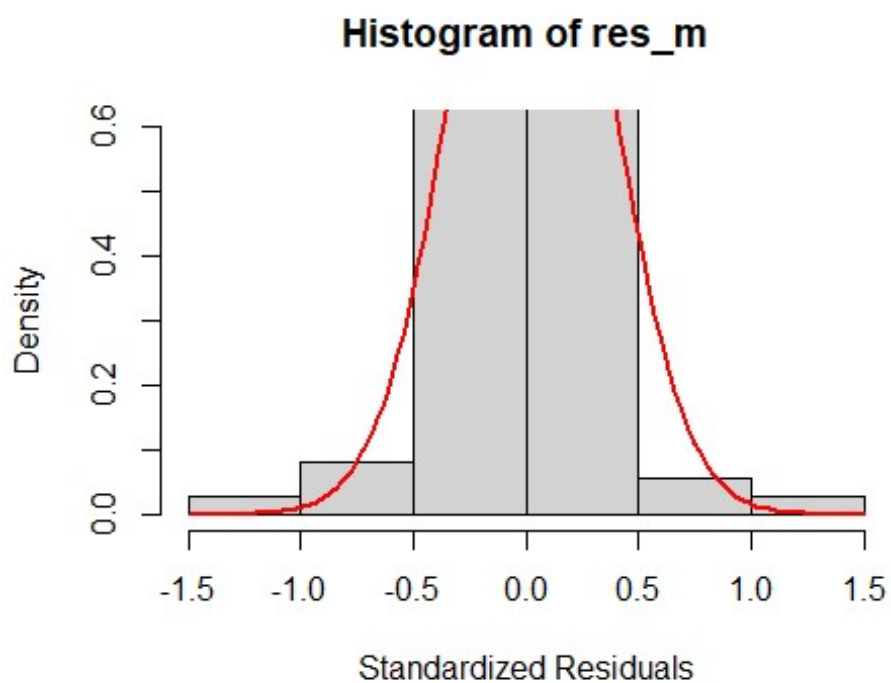
Fig 125: Histogram of ARMA(1,2) + GARCH(1,1)
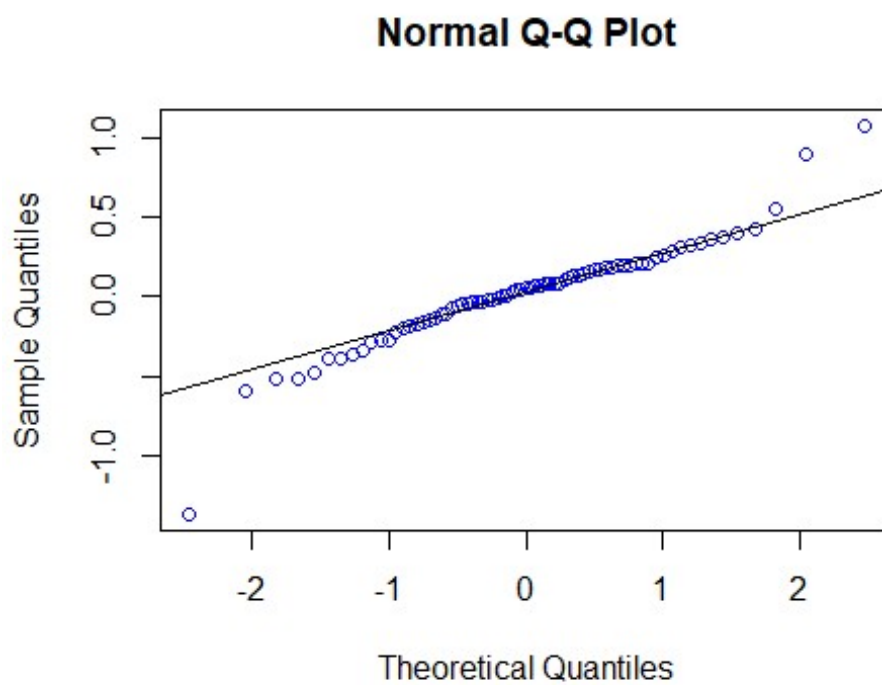
Fig 126: Normal Q-Q plot for ARMA(1,2) + GARCH(1,1)

```
##
##  Shapiro-Wilk normality test
##
## data:  res_m
## W = 0.92213, p-value = 0.000241

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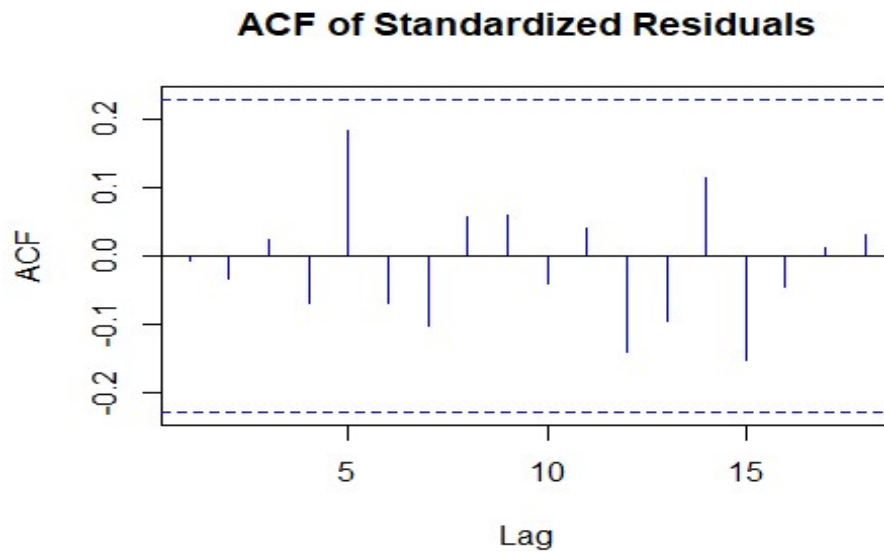

Fig 127: Autocorrelation plot for ARMA(1,2) + GARCH(1,1)
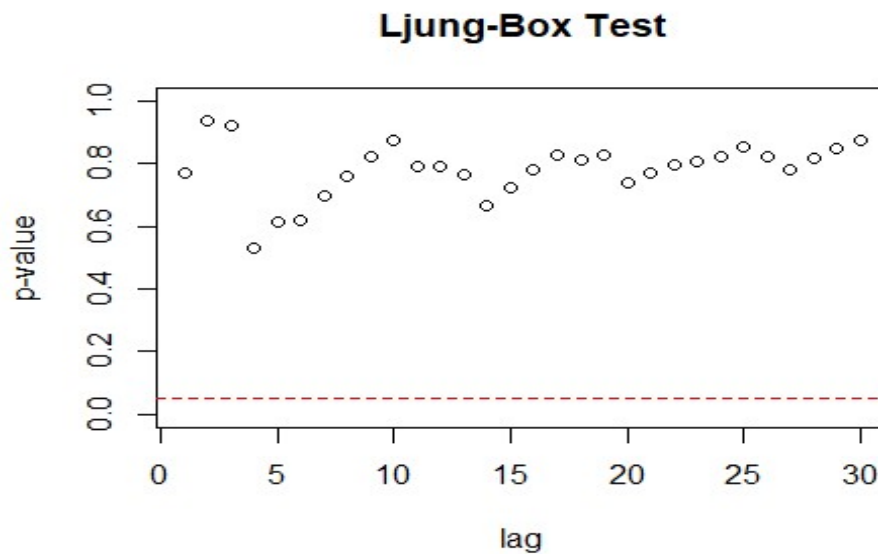
Fig 128: Ljung-Box test for ARMA(1,2) + GARCH(1,1)

Residual Analysis for ARMA(1,2) + GARCH(1,1):

1. p-value is less than 0.05 for ar1, ma1, ma2 and beta1. Hence, significant.
2. For alpha1, p-value is greater than 0.05. Hence, not significant.
3. In Residual plot, the line is passing through 0.
4. From normal distribution curve, qq plot and Shapiro Walk test, we can see normality in the data.
5. There are no significant lags in Autocorrelation plot.
6. Ljung box test shows good results as all the dots are above the dotted line.

Therefore, it is not considered as a good model since alpha1 is not significant.

Overall, ARMA(1,2) + GARCH(1,0) is best model. With this we are going to predict the future unemployment rate of U.S for the next 6 years i.e., till 2027.

Forecasting
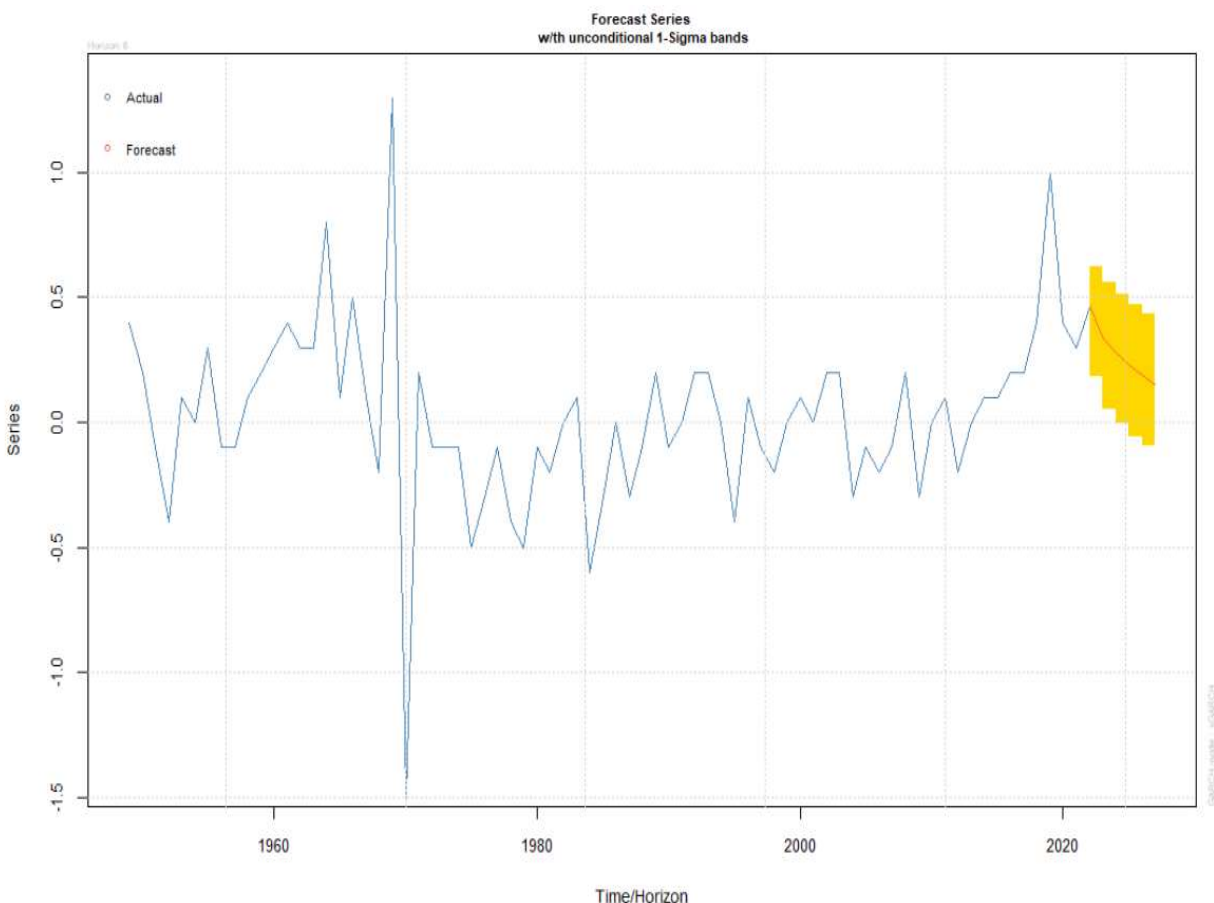
Fig 129: Forecast on unemployment rate for next 6 years

From the forecast graph, we can observe a decrease in the unemployment rate over the next 6 years.

Conclusion

1. The U.S. unemployment rate data set is loaded and converted as Time Series data.

2. Then respective analysis is done, and suitable model is justified.

3. The series is then checked for stationary and required conversions are done.

4. The corresponding ARMA orders and GARCH orders are obtained from the model specification tools.

5. Using these orders, the models are fitted on the series data.

6. Each model's residuals are then diagnosed, and best model is identified.

7. Future predictions are made on the series data with the best ARMA + GARCH model.

8. We observed a decline in the series for the next 6 years of unemployment rate.

9. This decrease in the unemployment rate indicates that there will be a more chance of employment in US in future.

References

1. Unemployment Rate. (2021). Retrieved 13 June 2021, from https://fred.stlouisfed.org/series/UNRATE

2. Time Series Lectures and Tutorials.

Appendix

library(TSA)

library(fUnitRoots)

library(forecast)

library(CombMSC)

library(lmtest)

library(fGarch)

library(rugarch)

library(tseries)

library(FitAR)

```
# Importing the data from CSV file to dataframe.

UNRATE <- read.csv("UNRAT.csv", header = TRUE)


# DATE variable is changed to DATE format and converted as rownames.

UNRATE$DATE <- as.Date(UNRATE$DATE)


rownames(UNRATE) <- UNRATE$DATE

UNRATE <- within(UNRATE, rm(DATE))


# Here, the dataframe is converted to Time Series(ts) object.

UNRATE_TS <- ts(as.vector(as.matrix(t(UNRATE))), start = 1948, end = 2021)


# Function to plot a single data.

v_Plot <- function(v, m){

  plot(v, type = "b", pch = 19, col = "blue", xlab = "years", ylab = "Unemployement Rate",
main = m)

}


# Function to form a legend corresponding to a single data plot.

v_leg1 <- function(t, l, c, p){

  legend("topright", inset = .03, title = t, legend = l, col = c, horiz = TRUE, cex = 0.8, lty = 1,
box.lty = 2, box.lwd = 2, box.col = "blue", pch = p)

}


v_Plot(UNRATE_TS, "Unemployement rate from 1948 to 2021")

v_leg1("Unemployement Rate over years.", c("Unemployement Rate"), c("blue"), c(19))


# Function to plot using two different data (Scatter Plot).
```

```
v_Plot1 <- function(v, v1, x, y, m){

  plot(x = v, y = v1, pch = 19, col = "blue", xlab = x, ylab = y, main = m)

}
```

```
# Function to form a legend corresponding to a two data plot.
```

```
v_leg <- function(t, l){

  legend("bottomright", inset = .03, title = t, legend = c(l), col = c("blue"), horiz = TRUE, cex =
0.8, box.lty = 2, box.lwd = 2, box.col = "blue", pch = c(19))

}
```

```
McLeod.Li.test(y = UNRATE_TS, main = "McLeod-Li Test Statistics for U.S unemployment
rate")
```

```
acf(UNRATE_TS, main = "Unemployment rate - ACF")
```

```
pacf(UNRATE_TS, main = "Unemployment rate - PACF")
```

```
adfTest(UNRATE_TS)
```

```
UNRATE_TS_1 <-  BoxCox.ar(UNRATE_TS + 1)
title(main = "Log-likelihood vs Lambda values")
```

```
# Checking best lambda
```

```
lambda = UNRATE_TS_1$lambda[which(max(UNRATE_TS_1$loglike) ==
UNRATE_TS_1$loglike)]

UNRATE_TS_BC = ((UNRATE_TS^lambda) - 1) / lambda


shapiro.test(UNRATE_TS_BC)


adfTest(UNRATE_TS_BC)


UNRATE_diff = diff(UNRATE_TS)


# Plot - First difference.

v_Plot(UNRATE_diff, "Differenced data of unemployment rate in U.S.")

v_leg("Differenced unemployment rate over the years.", c("First ever differenced data"))


adfTest(UNRATE_diff)


acf(UNRATE_diff, main = "Unemployment rate first difference data - ACF")


pacf(UNRATE_diff, main = "Unemployment rate first difference data - PACF")


#calculating the eacf

eacf(UNRATE_diff, ar.max = 6, ma.max = 9)


#creating the BIC plot

res = armasubsets(UNRATE_diff, nar = 9, nma = 9, y.name = 'ar', ar.method = 'ols')

plot(res)

title(main = "Unemployement Rate - BIC", line = 6)
```

```r
# Function for residual analysis.

res_analysis <- function(res_m) {

  # Scatter plot for model residuals
  plot(res_m, type = "b", pch = 19, col = "blue", xlab = "years", ylab = "Standardized Residuals", main = "Plot of Residuals over Time")

  abline(h = 0)

  # Standard distribution
  hist(res_m, xlab = 'Standardized Residuals', freq = FALSE, ylim = c(0, 0.6))
  curve(dnorm(x, mean = mean(res_m), sd = sd(res_m)), col = "red", lwd = 2, add = TRUE, yaxt = "n")

  # QQplot for model residuals
  qqnorm(res_m, col = c("blue"))
  qqline(res_m)

  # Auto-Correlation Plot
  acf(res_m, main = "ACF of Standardized Residuals",col=c("blue"))

  # Shapiro wilk test
  print(shapiro.test(res_m))

  # Ljung box test
  k=0
  LBQPlot(res_m, lag.max = 30, StartLag = k + 1, k = 0, SquaredQ = FALSE)
}
```

```
model_10_ml = arima(UNRATE_diff,order=c(1,0,0),method='ML')
coeftest(model_10_ml)
res_analysis(rstandard(model_10_ml))




model_01_ml = arima(UNRATE_diff,order=c(0,0,1),method='ML')
coeftest(model_01_ml)
res_analysis(rstandard(model_01_ml))




model_02_ml = arima(UNRATE_diff,order=c(0,0,2),method='ML')
coeftest(model_02_ml)
res_analysis(rstandard(model_02_ml))




model_03_ml = arima(UNRATE_diff,order=c(0,0,3),method='ML')
coeftest(model_03_ml)
res_analysis(rstandard(model_03_ml))

model_11_ml = arima(UNRATE_diff,order=c(1,0,1),method='ML')
coeftest(model_11_ml)
res_analysis(rstandard(model_11_ml))




model_12_ml = arima(UNRATE_diff,order=c(1,0,2),method='ML')
coeftest(model_12_ml)
```

```
res_analysis(rstandard(model_12_ml))


model_13_ml = arima(UNRATE_diff,order=c(1,0,3),method='ML')

coeftest(model_13_ml)

res_analysis(rstandard(model_13_ml))


model_20_ml = arima(UNRATE_diff,order=c(2,0,0),method='ML')

coeftest(model_20_ml)

res_analysis(rstandard(model_20_ml))


model_21_ml = arima(UNRATE_diff,order=c(2,0,1),method='ML')

coeftest(model_21_ml)

res_analysis(rstandard(model_21_ml))


model_22_ml = arima(UNRATE_diff,order=c(2,0,2),method='ML')

coeftest(model_22_ml)

res_analysis(rstandard(model_22_ml))


model_23_ml = arima(UNRATE_diff,order=c(2,0,3),method='ML')

coeftest(model_23_ml)

res_analysis(rstandard(model_23_ml))
```

```
model_30_ml = arima(UNRATE_diff,order=c(3,0,0),method='ML')

coeftest(model_30_ml)

res_analysis(rstandard(model_30_ml))
```

```
model_32_ml = arima(UNRATE_diff,order=c(3,0,2),method='ML')

coeftest(model_32_ml)

res_analysis(rstandard(model_32_ml))
```

```
model_33_ml = arima(UNRATE_diff,order=c(3,0,3),method='ML')

coeftest(model_33_ml)

res_analysis(rstandard(model_33_ml))
```

```
model_34_ml = arima(UNRATE_diff,order=c(3,0,4),method='ML')

coeftest(model_34_ml)

res_analysis(rstandard(model_34_ml))
```

```
sc.AIC=AIC(model_01_ml, model_02_ml, model_03_ml, model_10_ml, model_11_ml,
model_12_ml, model_13_ml, model_20_ml, model_21_ml, model_22_ml, model_23_ml,
model_30_ml, model_32_ml, model_33_ml, model_34_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(sc.AIC, score = "aic")
```

```
model_22_ml = arima(UNRATE_diff,order=c(2,0,2),method='ML')
coeftest(model_22_ml)
res_analysis(rstandard(model_22_ml))
```

```
model_13_ml = arima(UNRATE_diff,order=c(1,0,3),method='ML')
coeftest(model_13_ml)
res_analysis(rstandard(model_13_ml))
```

```
model_12_res = model_12_ml$residuals
model_12_abs = abs(model_12_res)
```

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

```
acf(model_12_abs, main = "The ACF plot for absolute residual series")
pacf(model_12_abs, main = "The PACF plot for absolute residual series")
```

```
eacf(model_12_abs, ar.max = 6, ma.max = 9)
```

```
res = armasubsets(model_12_abs, nar = 9, nma = 9, y.name = 'ar', ar.method = 'ols')

plot(res)

title(main = "Unemployment rate - BIC", line = 6)



model_12_sq = model_12_res ^ 2



par(mfrow=c(1,2))



acf(model_12_sq, main = "The ACF plot for squared residual series")

pacf(model_12_sq, main = "The PACF plot for squared residual series")



eacf(model_12_sq, ar.max = 6, ma.max = 9)



res = armasubsets(model_12_sq, nar = 9, nma = 9, y.name = 'ar', ar.method = 'ols')

plot(res)

title(main = "Unemployment rate - BIC", line = 6)



model_1210 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(0, 1)),

            mean.model = list(armaOrder = c(1, 2), include.mean = FALSE),

            distribution.model = "norm")



model_12_10 <- ugarchfit(spec = model_1210, data = UNRATE_diff)
```

```
model_12_10
```

```
res_analysis(model_12_10@fit$residuals)
```

```
model_1211 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1, 1)),
            mean.model = list(armaOrder = c(1, 2), include.mean = FALSE),
            distribution.model = "norm")
```

```
model_12_11 <- ugarchfit(spec = model_1211, data = UNRATE_diff)
model_12_11
```

```
res_analysis(model_12_11@fit$residuals)
```

```
model_1264 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(4, 6)),
            mean.model = list(armaOrder = c(1, 2), include.mean = FALSE),
            distribution.model = "norm")
```

```
model_12_64 <- ugarchfit(spec = model_1264, data = UNRATE_diff)
model_12_64
```

```
res_analysis(model_12_64@fit$residuals)
```

```
AICs = c(infocriteria(model_12_10)[1], infocriteria(model_12_11)[1],
infocriteria(model_12_64)[1])
```

```
BICs = c(infocriteria(model_12_10)[2], infocriteria(model_12_11)[2],
infocriteria(model_12_64)[2])


UN = data.frame(model = c("model_12_10", "model_12_11", "model_12_64"))

UN$AIC = AICs

UN$BIC = BICs


UN[order(UN$AIC), ]

UN[order(UN$BIC), ]



model_2210 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(0, 1)),
            mean.model = list(armaOrder = c(2, 2), include.mean = FALSE),
            distribution.model = "norm")


model_22_10 <- ugarchfit(spec = model_2210, data = UNRATE_diff)

model_22_10


res_analysis(model_22_10@fit$residuals)



model_1310 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(0, 1)),
            mean.model = list(armaOrder = c(1, 3), include.mean = FALSE),
            distribution.model = "norm")


model_13_10 <- ugarchfit(spec = model_1310, data = UNRATE_diff)

model_13_10
```

```
res_analysis(model_13_10@fit$residuals)


model_1220 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(0, 2)),
            mean.model = list(armaOrder = c(1, 2), include.mean = FALSE),
            distribution.model = "norm")


model_12_20 <- ugarchfit(spec = model_1220, data = UNRATE_diff)
model_12_20


res_analysis(model_12_20@fit$residuals)


model_1211 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1, 1)),
            mean.model = list(armaOrder = c(1, 2), include.mean = FALSE),
            distribution.model = "norm")


model_12_11 <- ugarchfit(spec = model_1211, data = UNRATE_diff)
model_12_11


res_analysis(model_12_11@fit$residuals)


UN_FUTURE = ugarchforecast(model_12_10, data = UNRATE_diff, n.ahead = 6)


v_Plot(UN_FUTURE, "Forecast for next 6 years")
```