Student: HILL JIGISHKUMAR MODI (S3827516)

I certify that this is all my own original work. If I took any parts from elsewhere, then they were non-essential parts of the assignment, and they are clearly attributed in my submission. I will show I am agree to this honour code by typing "Yes": *Yes*.

PROJECT – TIME SERIES ANALYSIS

Time Series Analysis to predict 10 month Waste Collection (in Tonnes) of Melbourne

Hill Jigishkumar Modi s3827516@student.rmit.edu.au

RMIT University, City Campus

Melbourne, Australia

Date: 12/06/2021

CONTENTS

FIGURES

INTRODUCTION

Data is taken from Assignment Section of Time Series Analysis.

Here, we have yearly dataset of Changes in Ozone Layer Thickness in Dobson Units from year 1927 to 2016. As dataset is about change, value will be respective to the previous year. If value is positive, then it represents increase in the Ozone Layer Thickness, if negative, it represents decrease in the Ozone Layer Thickness.

AIM

The Main objective/aim of report is to determine which Model is best suitable for the data, if we observe deterministic trend. For, stochastic trend, our main goal will be to identify set of Possible ARIMA(p,d,q) Models for the annual time series data of Ozone Layer Thickness.

IMPORTANT TERMINOLOGY

MA

MA is moving average. If time-series is fluctuates between time, then we can say, behaviour of time-series is MA.

AR

AR is Auto-Regressive. If time-series has successive time-points, then it's behaviour is AR.

ACF

ACF is Auto-Correlation Function. We can determine q(order of MA) at the lag from the ACF plot. But, if there is pattern in ACF plot, then we can say q(order of MA) = 0. If there is wave like pattern, we can say there is seasonality in the data, if slowly decaying pattern, there is trend.

PACF

PACF is Partial Auto-Correlation Function. We can determine p(order of AR) at the lag from PACF plot.

Both, ACF and PACF tells us autocorrelation at lag.

DETERMINISTIC TREND

Deterministic Trend is the trend in which we can obtain trend by straight-forwardly looking at the equation. It only depends on the time.

$y_t = \beta_0 + \beta_1 t$

Where y_t = current time-point value

β_0 = intercept,

β_1 = slope

t = time point

Deterministic trend (DT) : yt = βt + t Stochastic trend (ST) : yt = β + yt−1 + t ,

STOCHASTIC TREND

Stochastic Trend is the trend in which current time-point value is also dependent on the previous time-point value.

$y_t = \beta_0 + \beta_1 t + y_{t-1}$

Where y_t = current time-point value

y_{t-1} = previous time-point value

β_0 = intercept,

β_1 = slope

t = time point

STATIONARITY / NON- STATIONARITY

If mean of Time-series is same through the time points, then it is stationary.

If mean of time-series plot is changes through the time points, then series is non-stationary.

SHAPIRO-WILK TEST

Shapiro-Wilk test will check the normality.

Hypothesis

Null Hypothesis – data is normally distributed.

Alternate Hypothesis – data is not normally distributed.

Conclusion

if p-value is greater than 0.05, then we fail to reject null hypothesis.

Data is normally distributed.

AIC

AIC is Akaike's Information Criterion. AIC calculates the relative quality of each model and give as output as model with lowest AIC value among them.

BIC

BIC is Bayesian Information Criterion. It gives us model with lowest BIC value.

SARIMA(p,d,q)X(P,D,Q)

ARIMA is AutoRegressive Integrated Moving Average model.

Where p = order of AR (AutoRegressive)

d = number of difference

q = order of MA (Moving Average)

SARIMA(p,d,q)x(P,D,Q)

ADF TEST

ADF Test is Augmented Dicky-Fuller Test. It is used to check the stationarity of time-series.

Hypothesis

Null Hypothesis – Series is Non-stationary

Alternate Hypothesis – Series is stationary.

Conclusion

if p-value is greater than 0.05, then we fail to reject null hypothesis.

Series is non-stationary.

EACF

EACF is Extended AutoCorrelation Function.

We will look for top-most 0s with vertex. Then we will include that order of p and q to ARIMA(p,d,q) and we will also include its neighbour orders.

BIC TABLE

BIC Table is Bayesian Information Criterion Table.

It compares the BIC value with significant possible ARIMA order. If respective box is darker, then it is significant.

Residuals

Ljung-Box Testing

METHODOLOGY

I will check if time-series data has changing variance and stationarity. If it is there, then I'll deal with it with help of log or Box-Cox transformation and with taking the differences till I found stationary time-series.

Then, I'll find set of possible ARIMA(p,d,q) model with help of ACF, PACF, EACF and BIC table.

DESCRIPTIVE STATISTICS

	month	year	residential
1	04	2009	1304.31
2	05	2009	1249.74
3	06	2009	1333.32
4	07	2009	1405.98
5	08	2009	1364.62
6	09	2009	1505.93

Figure 1 – Waste Collection of First 5 months

	month	year	residential
1	10	2019	2742.06
2	11	2019	2662.94
3	12	2019	2738.82
4	01	2020	2608.78
5	02	2020	2382.83
6	03	2020	2627.32

Figure 2 - Waste Collection of Last 5 months

```
> wasteTS = ts(waste$residential, start = c(2009,4), end = c(2020,3), frequency = 12)
> wasteTS
         Jan     Feb     Mar     Apr     May     Jun     Jul     Aug     Sep     Oct     Nov     Dec
2009                         1304.31 1249.74 1333.32 1405.98 1364.62 1505.93 1449.20 1578.91 1637.58
2010 1385.06 1371.76 1753.64 1488.46 1552.58 1538.28 1245.02 1591.70 1587.01 1630.89 1811.23 2005.04
2011 1660.16 1644.12 1704.54 1557.64 1567.16 1541.58 1516.18 1669.48 1577.94 1601.54 1716.64 1650.24
2012 1618.54 1605.88 1678.14 1580.10 1645.68 1523.50 1669.32 1645.42 1473.22 1732.58 1760.86 1842.44
2013 1735.06 2173.34 1669.94 1752.37 1737.86 1434.32 1813.02 1757.26 2026.20 1790.98 1730.50 1876.50
2014 1867.96 1785.60 1837.66 1880.06 1912.48 2093.58 1955.66 1836.81 1989.18 1955.32 1877.48 2162.64
2015 1958.38 1867.96 2020.24 1940.78 1822.60 1896.00 1927.88 1952.66 2028.04 1940.72 2139.99 2206.70
2016 1987.13 2202.28 2450.44 2051.10 2418.20 2090.13 1993.94 2244.62 2071.20 2117.99 2163.72 2195.00
2017 2196.99 2067.04 2333.42 2037.14 2329.93 2172.61 2208.52 2353.64 2553.79 2507.50 2620.04 2472.72
2018 2558.41 2375.20 2491.24 2379.96 2505.84 2289.50 2404.80 2441.82 2239.68 2697.46 2546.07 2542.94
2019 2549.50 2730.18 2752.70 2536.18 2479.20 2262.16 2794.24 3298.58 2932.60 2742.06 2662.94 2738.82
2020 2608.78 2382.83 2627.32
```

Figure 3 - Waste collection as Time-Series Object

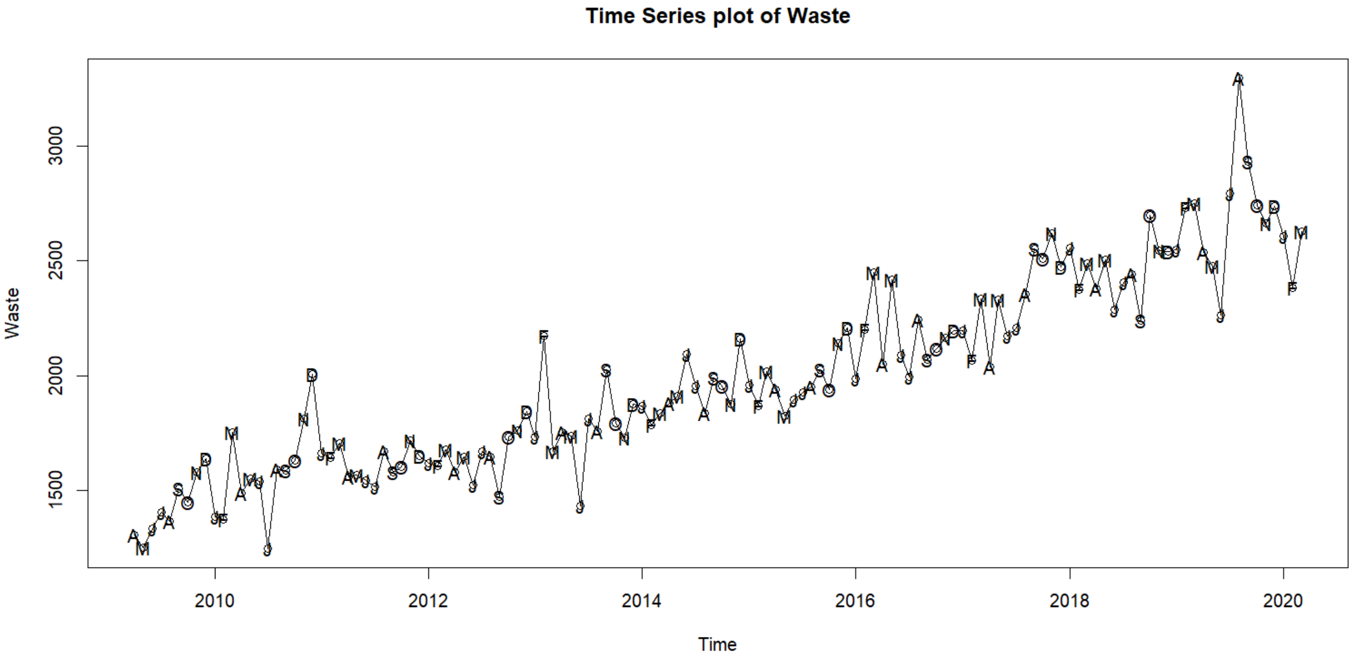


Figure 4 - Time Series Plot

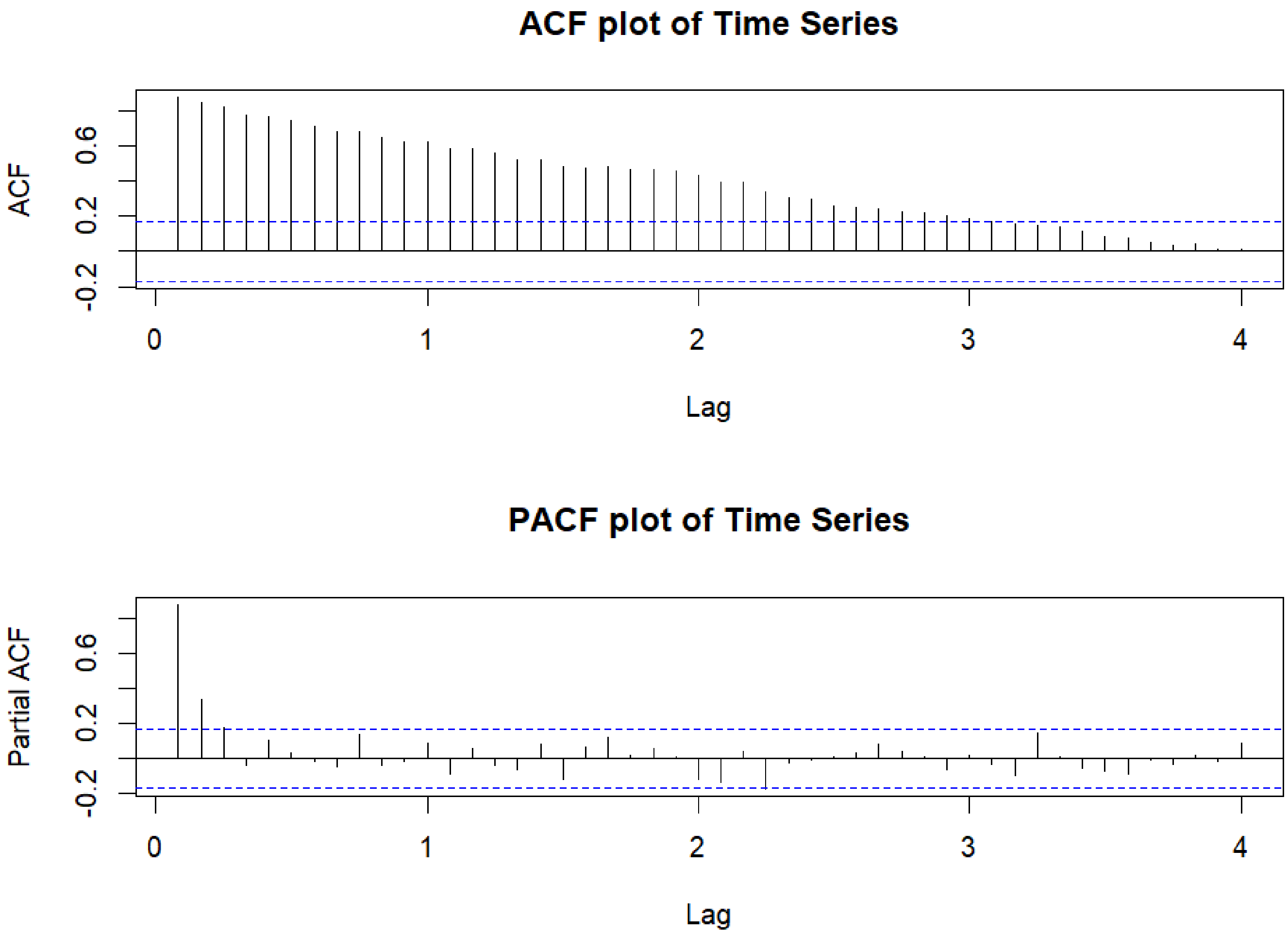


Figure 5 - Auto-correlation of Time Series

TREND

By observing the above graph it is detected that there is an upwards trend with a changing mean level. Hence, the series could be considered a non-stationary series and slowly decaying pattern from ACF plot also confirms our assumption of trend

CHANGING VARIANCE

SEASONALITY

Seasonality refers to repeating patterns in time intervals. In this series it is not possible to identify a clear seasonality. Further analysis needs to be conducted to identify any seasonality if it exists.

BEHAVIOUR

INTERVENTION POINT

RELATION BETWEEN CONSECUTIVE YEARS

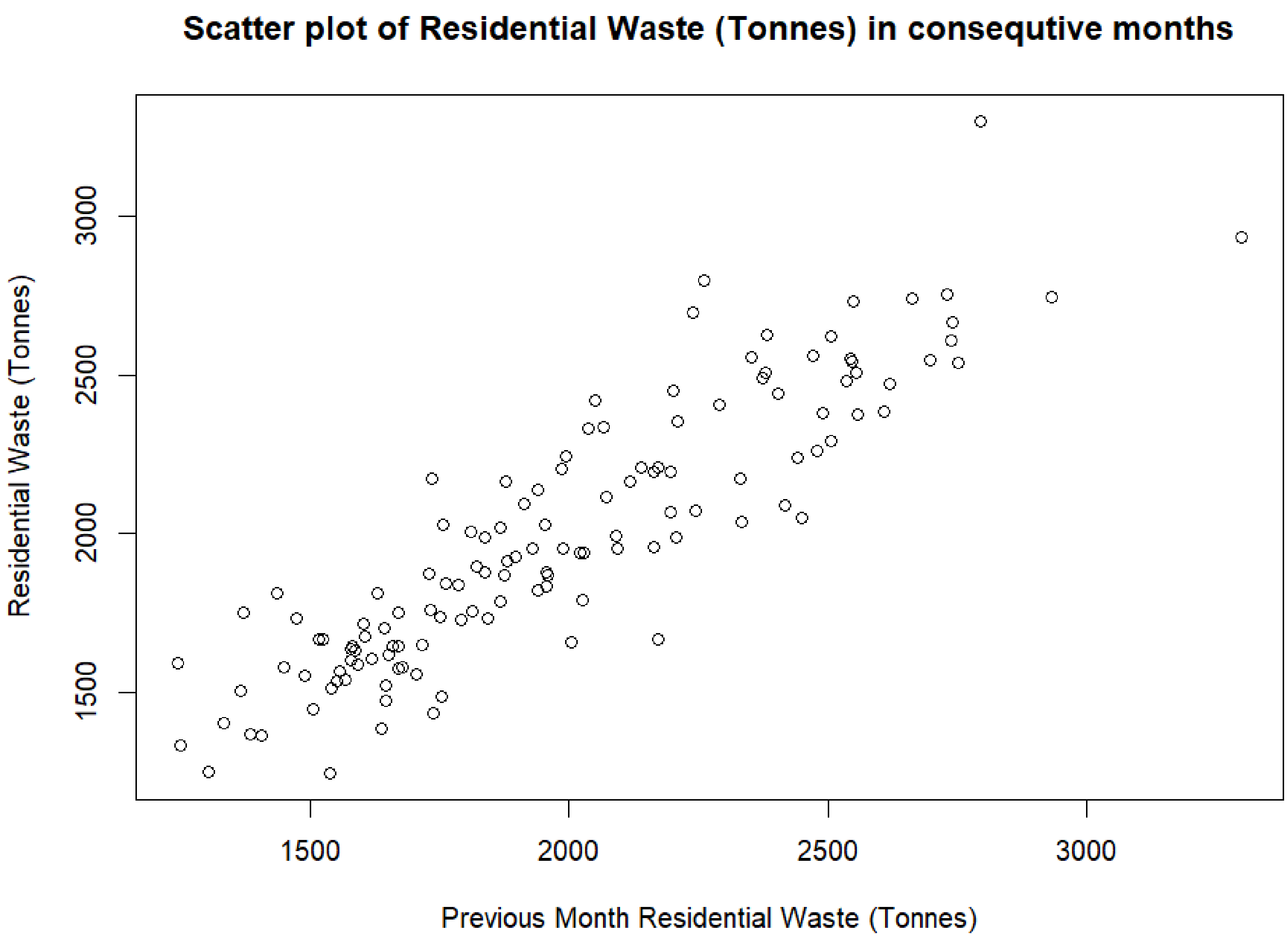


Figure 6 - Scatterplot of Residential waste in consecutive months

```
> cor(y[index],x[index]) %>% round(3)
[1] 0.897
```

Figure 7- Correlation value of Residential waste in consecutive months

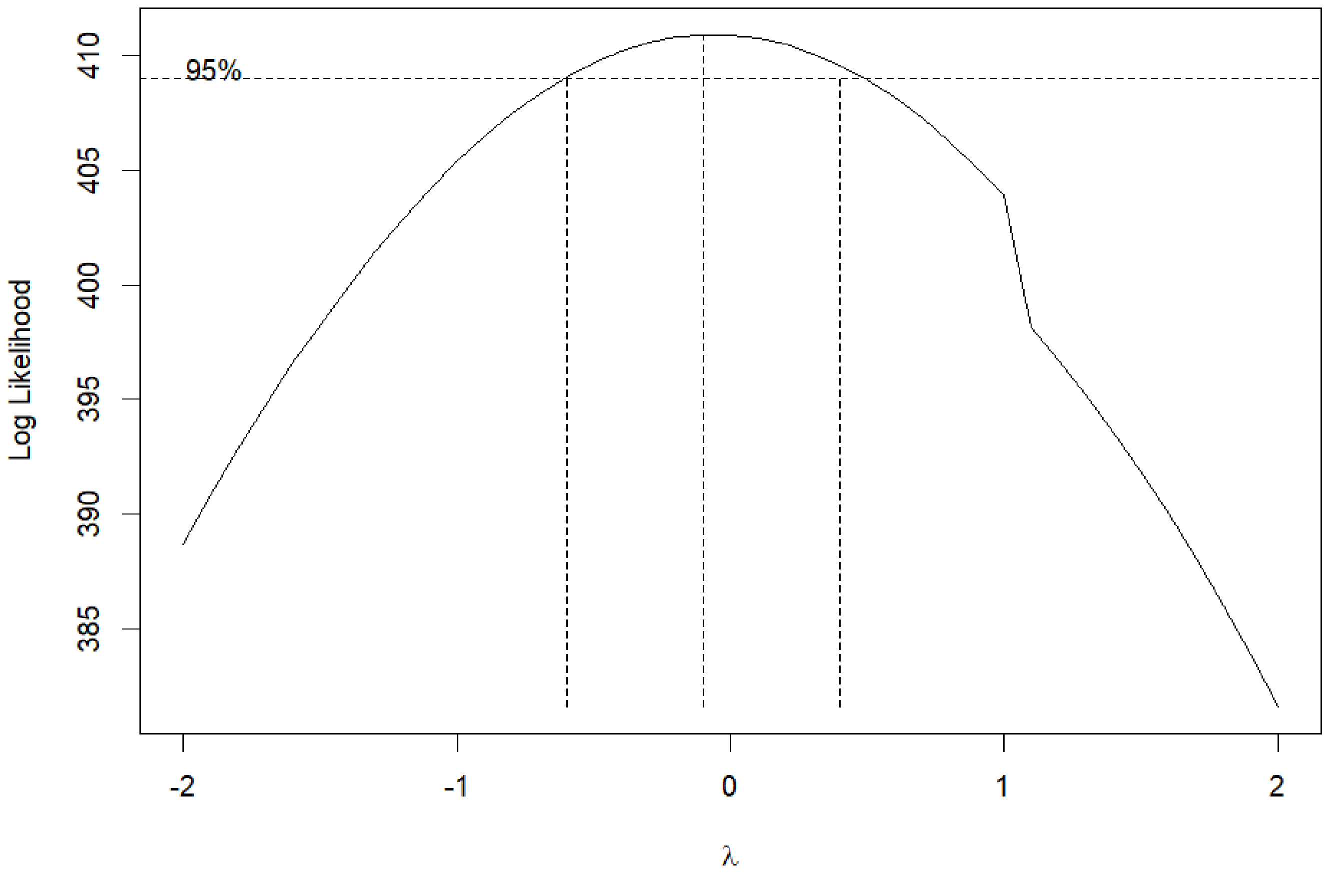


Figure 8- Box-Cox Curve of maximum likelihood

```
> BC$ci
[1] -0.6  0.4
```

Figure 9 - Confidence Interval of lambda

```
> lambda
[1] -0.1
```

Figure 10- lambda value

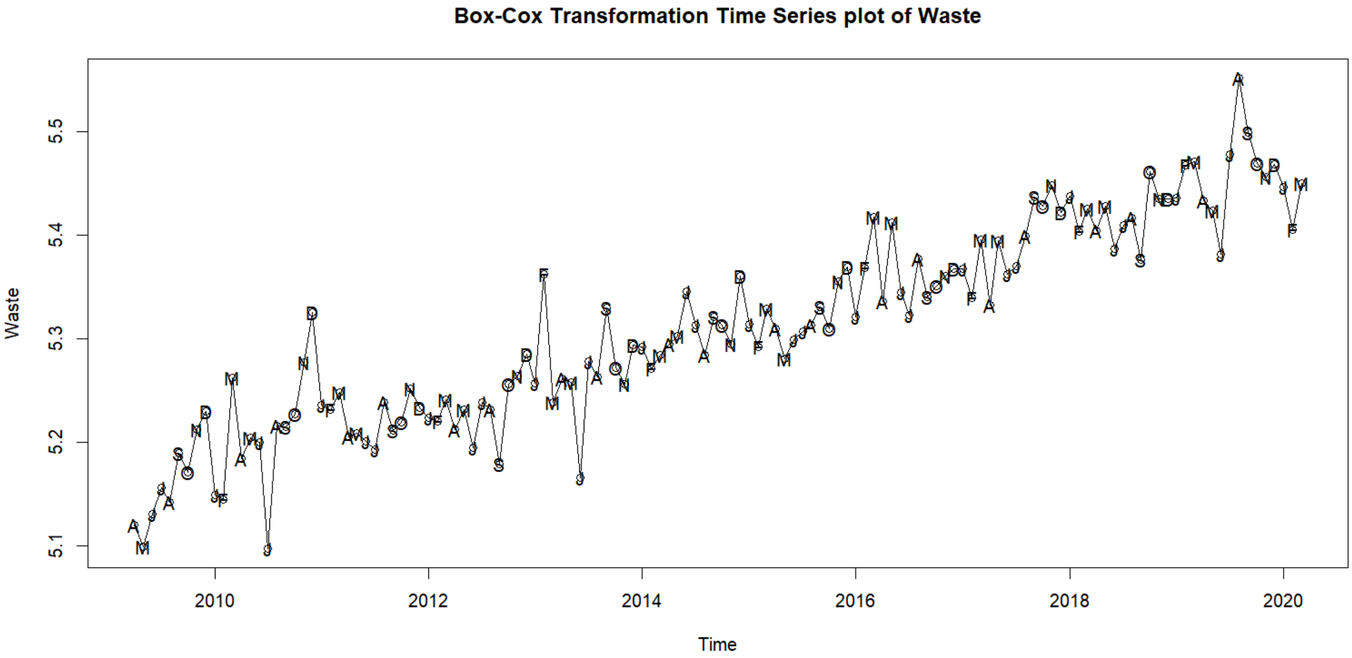


Figure 11 – Box-Cox Transformed Time Series

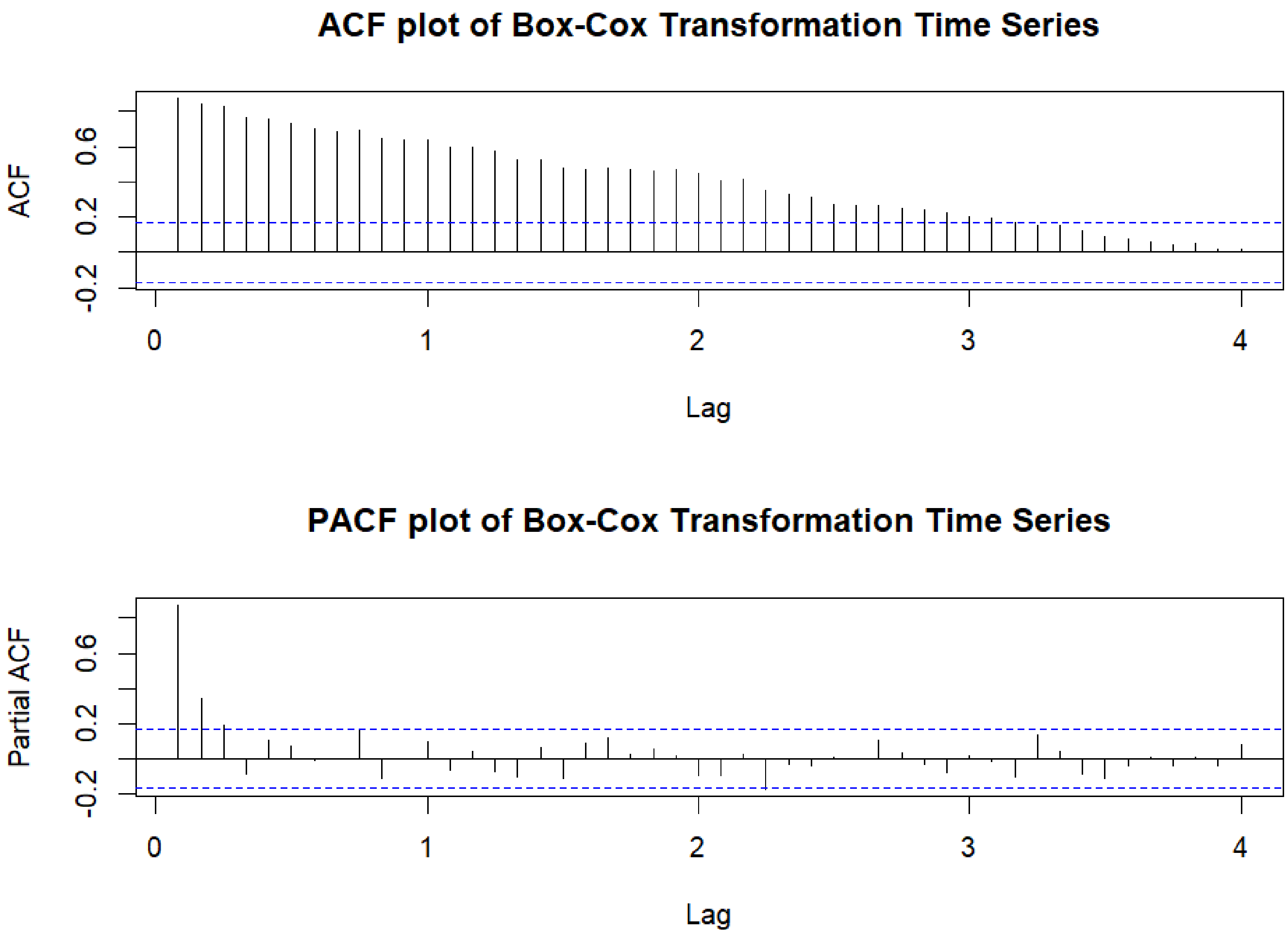


Figure 12 - Auto-correlation of Box-Cox transformed time Series

DETERMINING SET OF POSSIBLE MODELS

SARIMA(0,0,0)X(0,1,0)

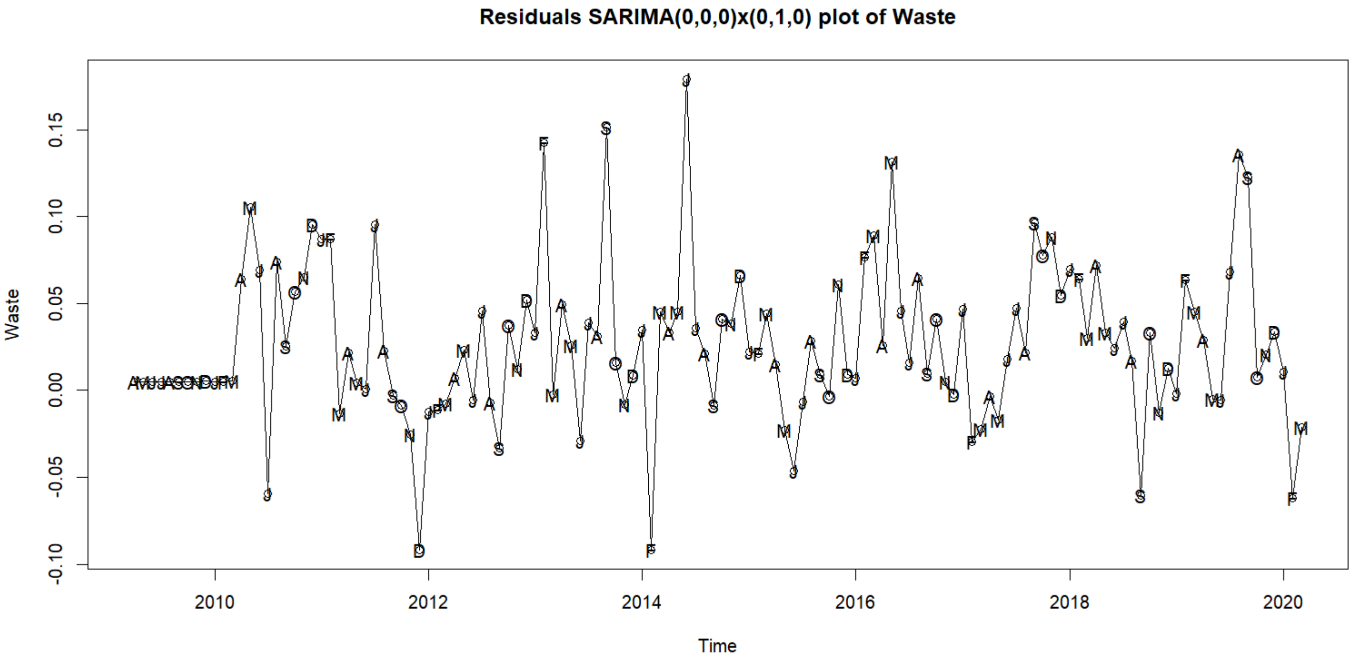


Figure 13 - Residual time-series plot of SARIMA(0,0,0) x (0,1,0)

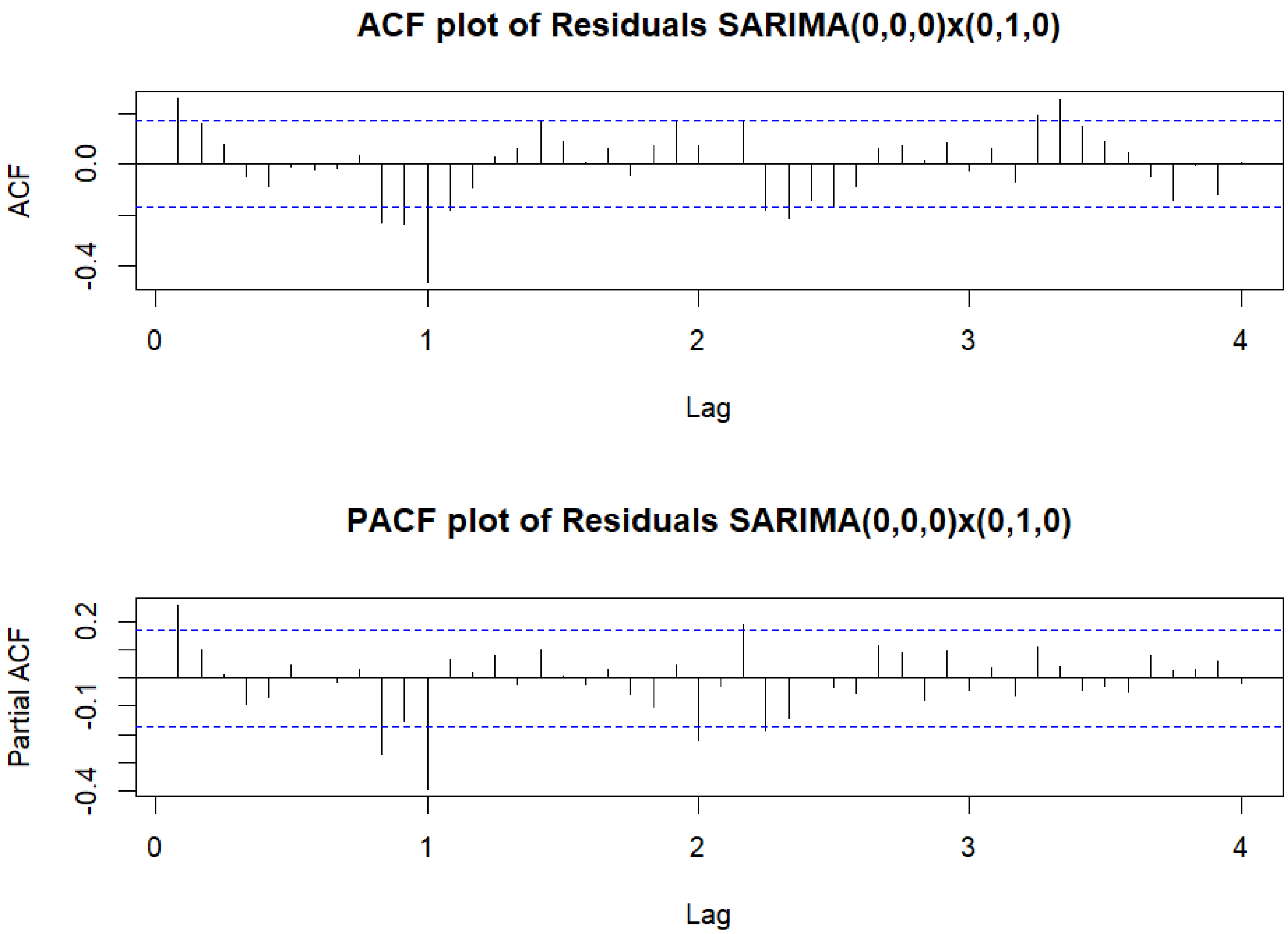


Figure 14- Residual auto-correlation plot of SARIMA(0,0,0) x (0,1,0)

ACF

PACF

SARIMA(0,0,0)X(2,1,1)

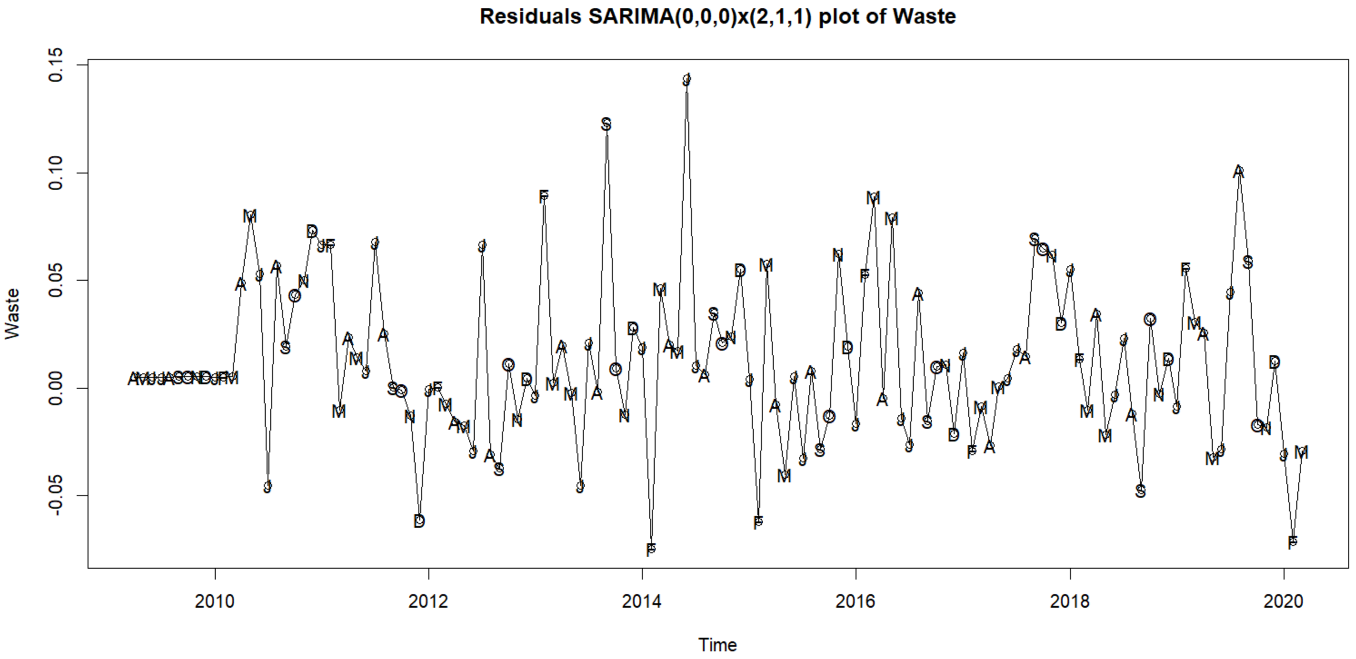


Figure 15- Residual time-series plot of SARIMA(0,0,0) x (2,1,1)

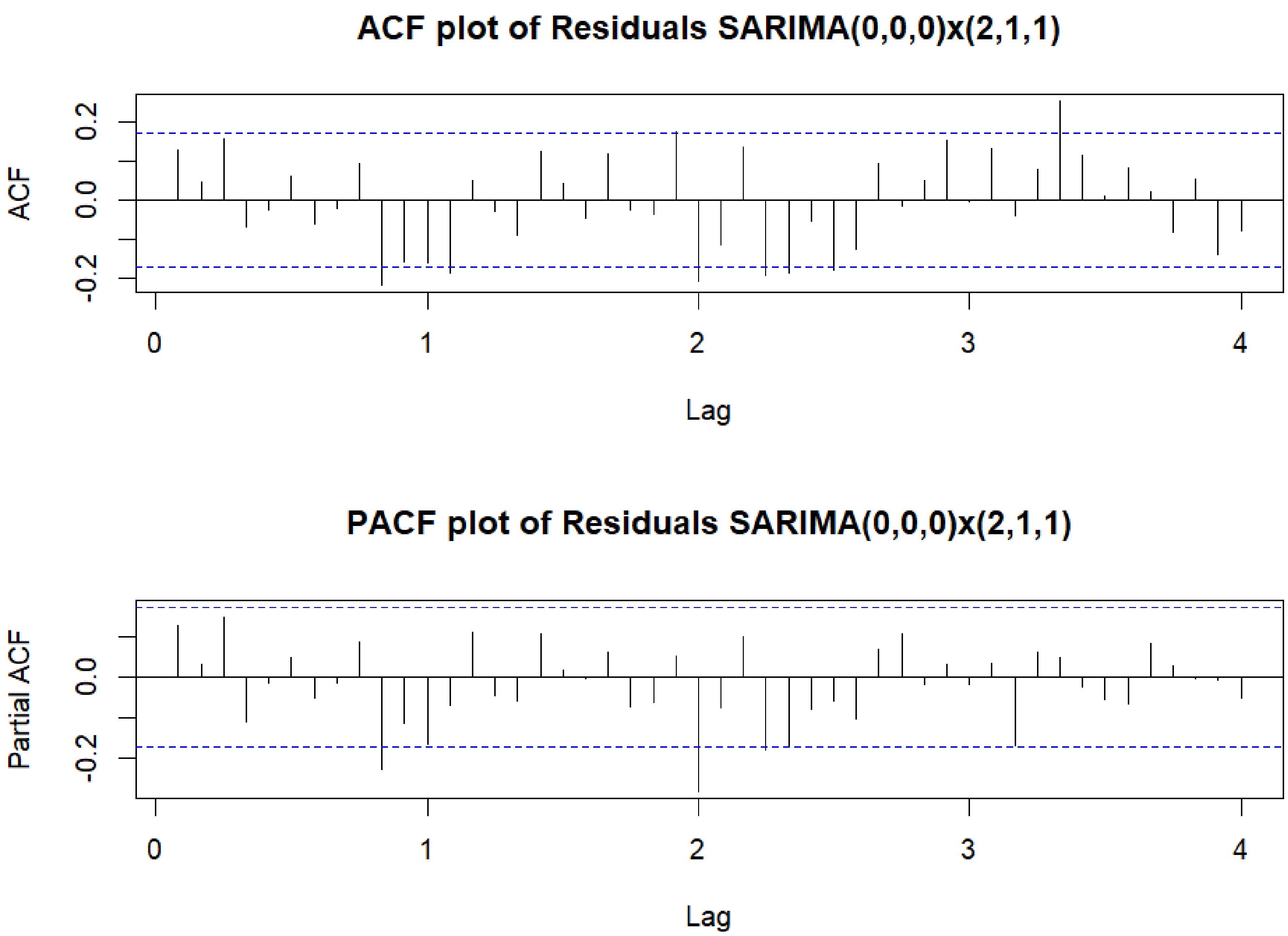


Figure 16- Residual auto-correlation plot of SARIMA(0,0,0) x (2,1,1)

ACF

PACF

SARIMA(2,0,2)X(2,1,1)

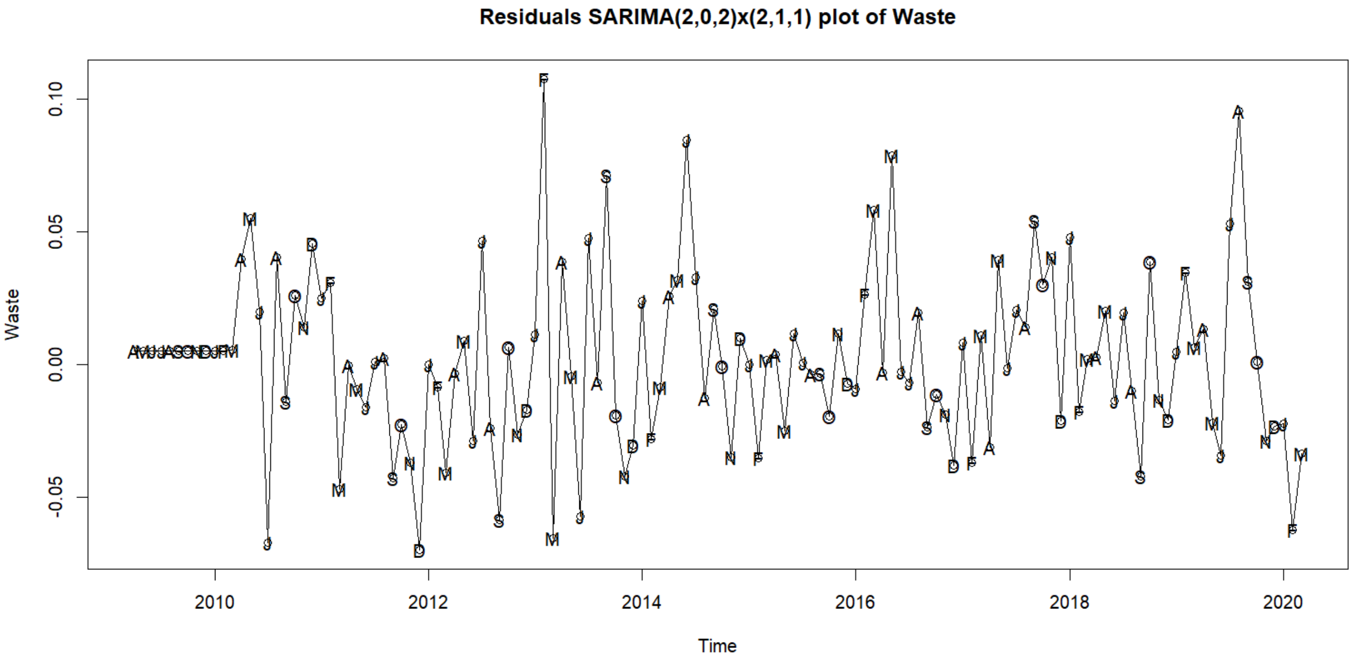
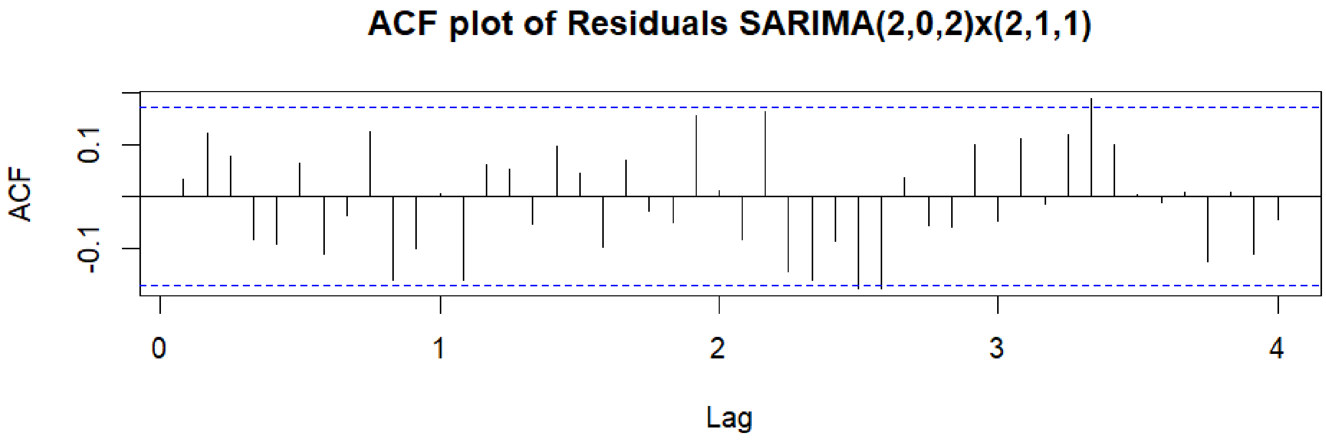


Figure 17- Residual time-series plot of SARIMA(2,0,2) x (2,1,1)



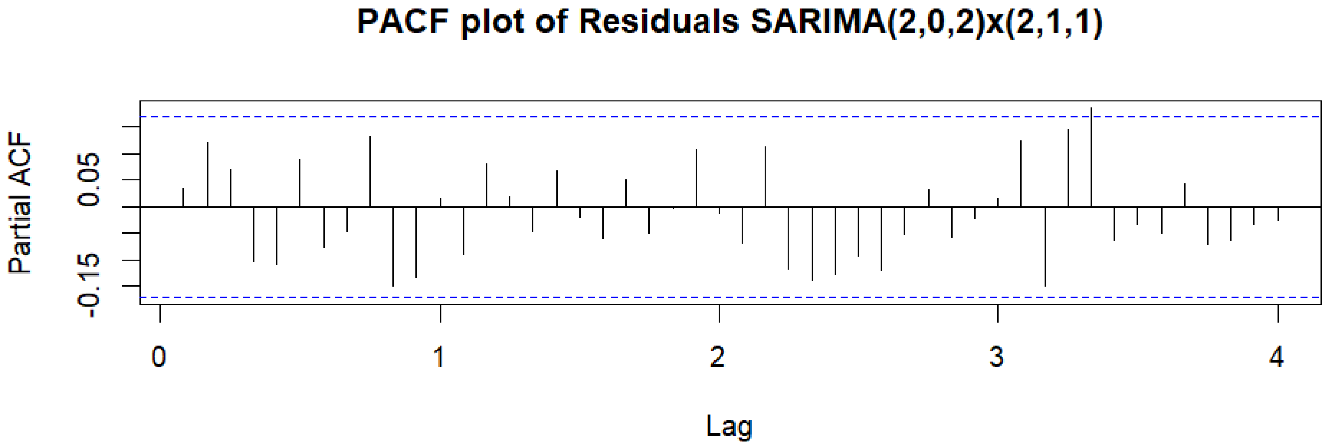


Figure 18- Residual auto-correlation plot of SARIMA(2,0,2) x (2,1,1)

ACF

PACF

DETERMINING SET OF SARIMA(p,d,q)x(P,D,Q) MODELS

ACF & PACF

EACF

```
> # ----------------------------------
> # EACF
> # ----------------------------------
> eacf(res.m2)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 o o o o o o o o o x o  o  x  o
1 x o o o o o o o o x o  o  x  o
2 x x o o o o o o o x o  o  x  o
3 x x o o o o o o o x o  o  o  o
4 o x o o o o o o o o o  o  o  o
5 x o x o o o o o o o o  o  o  o
6 x x o o o o o o o o o  o  o  o
7 x x o o o o o o o o o  o  o  o
```

Figure 19 - EACF Plot

BIC TABLE

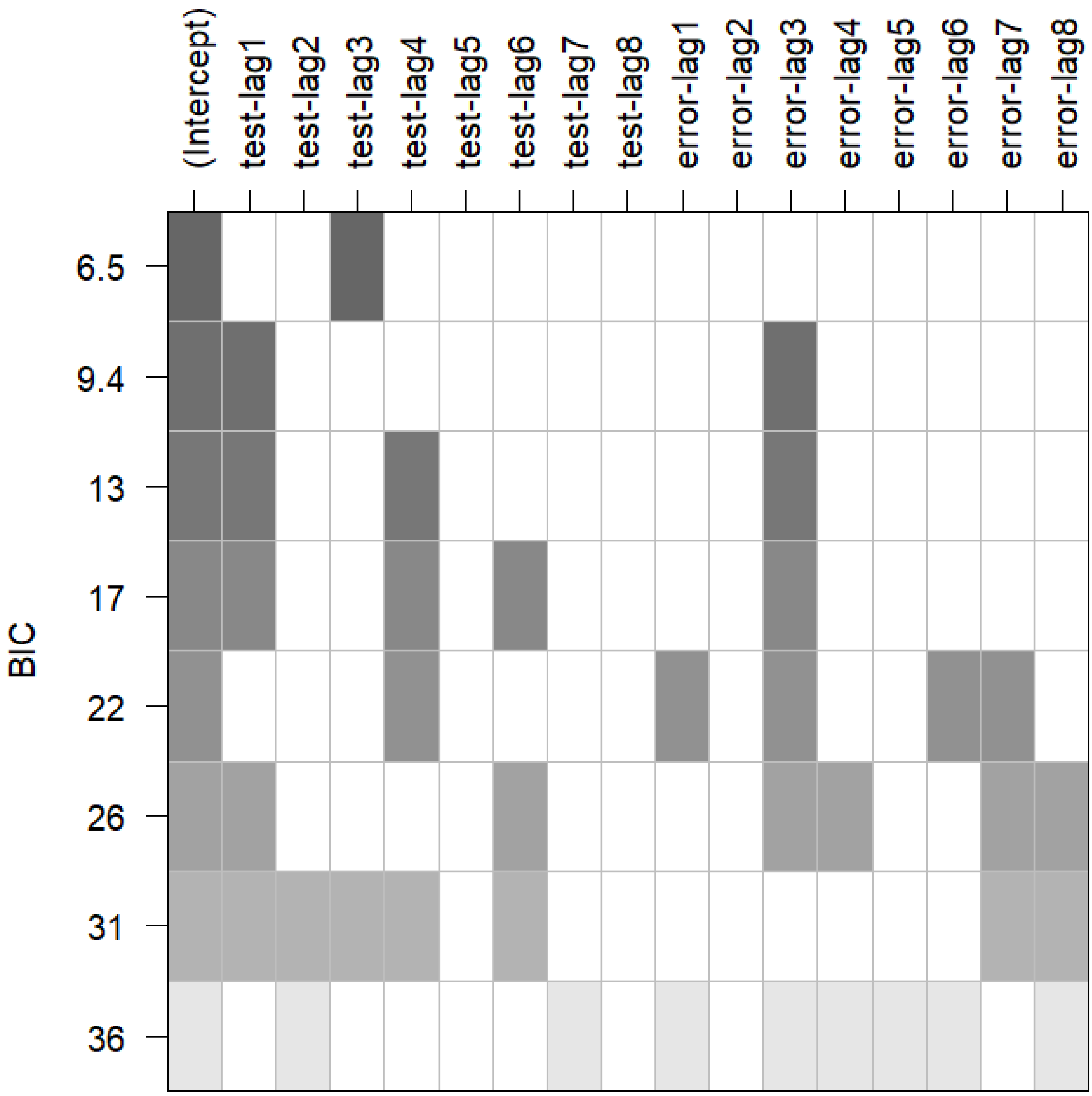


Figure 20 - BIC Table

DIAGNOSTIC & RESIDUAL ANALYSIS

SARIMA(0,0,0)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(0,0,0,"SARIMA(0,0,0) X (2,1,1)")

z test of coefficients:

      Estimate Std. Error  z value  Pr(>|z|)    
sar1  0.502024   0.088912   5.6463  1.64e-08 ***
sar2  0.497471   0.088906   5.5955  2.20e-08 ***
sma1 -0.974953   0.058083 -16.7854 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 21 - Co-efficient of SARIMA(0,0,0) x (2,1,1)

RESIDUAL CHECKING

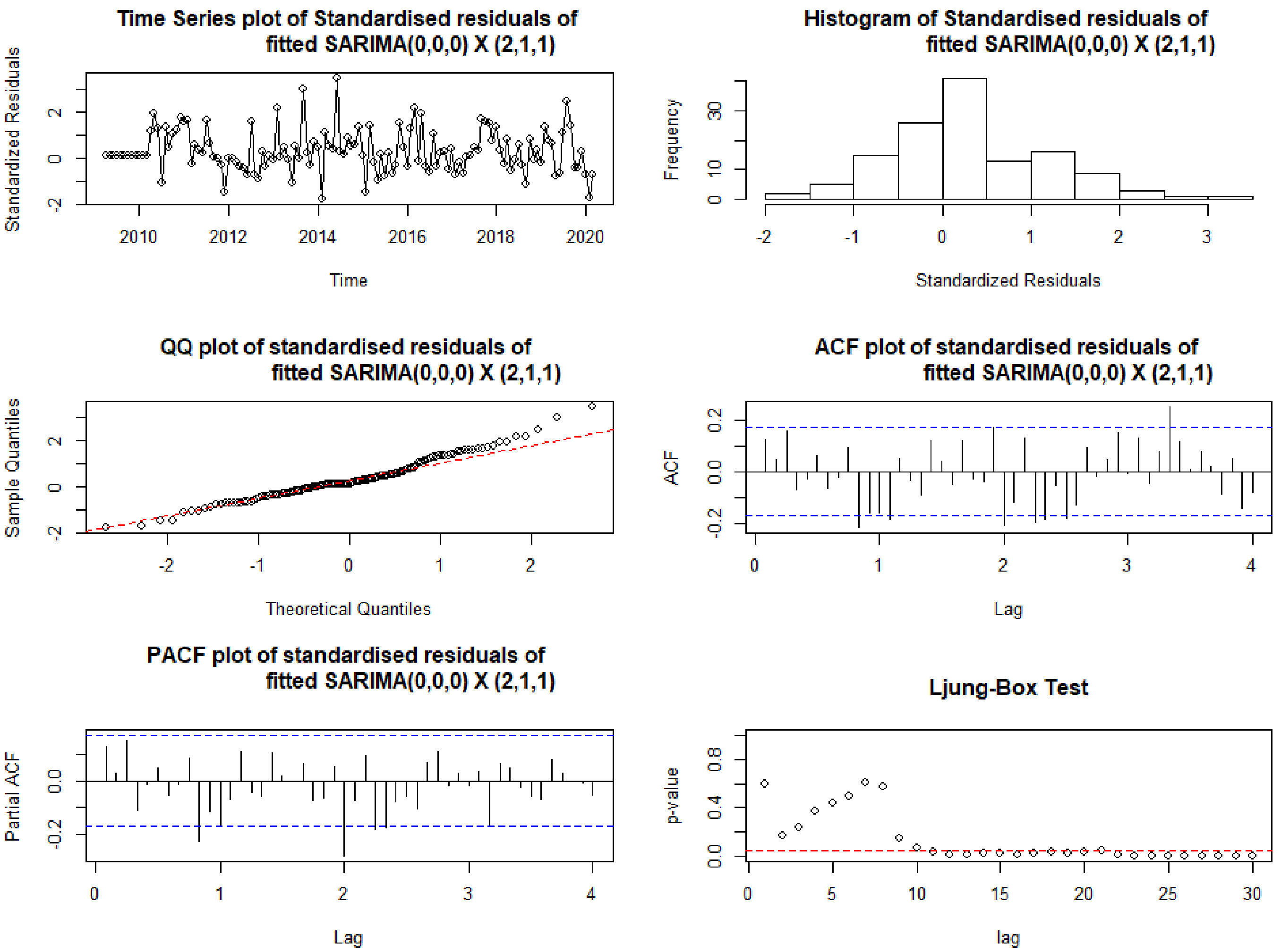


Figure 22 - Residual Diagnostics of SARIMA(0,0,0) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.97223, p-value = 0.008305
```

Figure 23 - Normality Check of SARIMA(0,0,0) X (2,1,1)

SARIMA(0,0,1)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(0,0,1,"SARIMA(0,0,1) X (2,1,1)")

z test of coefficients:

      Estimate Std. Error  z value  Pr(>|z|)
ma1   0.226463   0.092725   2.4423   0.01459 *
sar1  0.547743   0.092059   5.9499 2.683e-09 ***
sar2  0.451677   0.092074   4.9056 9.315e-07 ***
sma1 -0.978948   0.049256 -19.8746 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 24 - Co-efficient of SARIMA(0,0,1) x (2,1,1)

RESIDUAL CHECKING

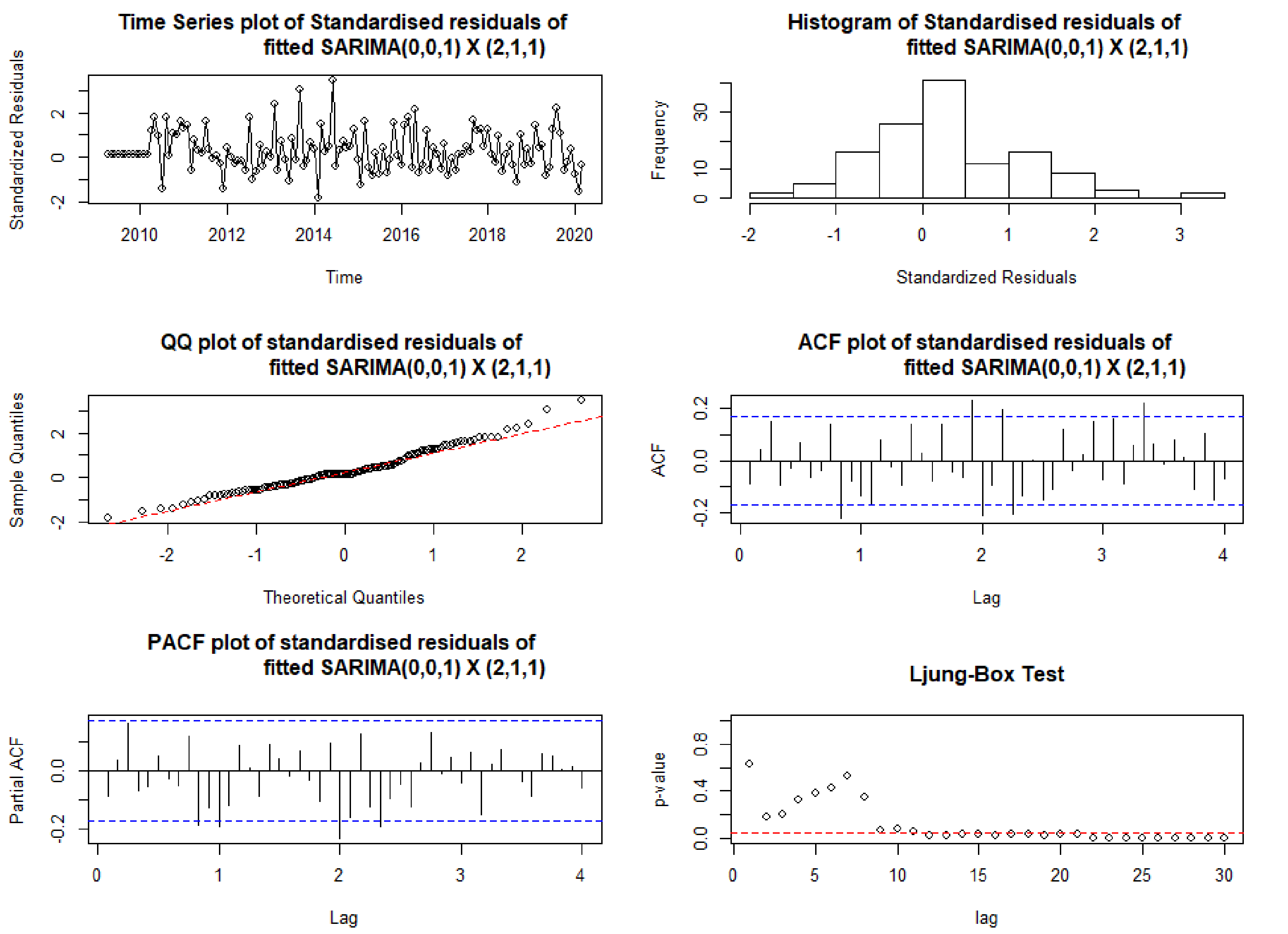


Figure 25 - Residual Diagnostics of SARIMA(0,0,1) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.97141, p-value = 0.006935
```

Figure 26 - Normality Check of SARIMA(0,0,1) X (2,1,1)

SARIMA(1,0,1)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(1,0,1,"SARIMA(1,0,1) X (2,1,1)")

z test of coefficients:

        Estimate   Std. Error   z value   Pr(>|z|)
ar1    0.99996391  0.00021718 4604.2063 < 2.2e-16 ***
ma1   -0.91506537  0.10068114   -9.0887 < 2.2e-16 ***
sar1  -0.05615592  0.13282076   -0.4228 0.6724450
sar2   0.01321980  0.12712548    0.1040 0.9171771
sma1  -0.91589623  0.24198758   -3.7849 0.0001538 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 27 - Co-efficient of SARIMA(1,0,1) x (2,1,1)

RESIDUAL CHECKING

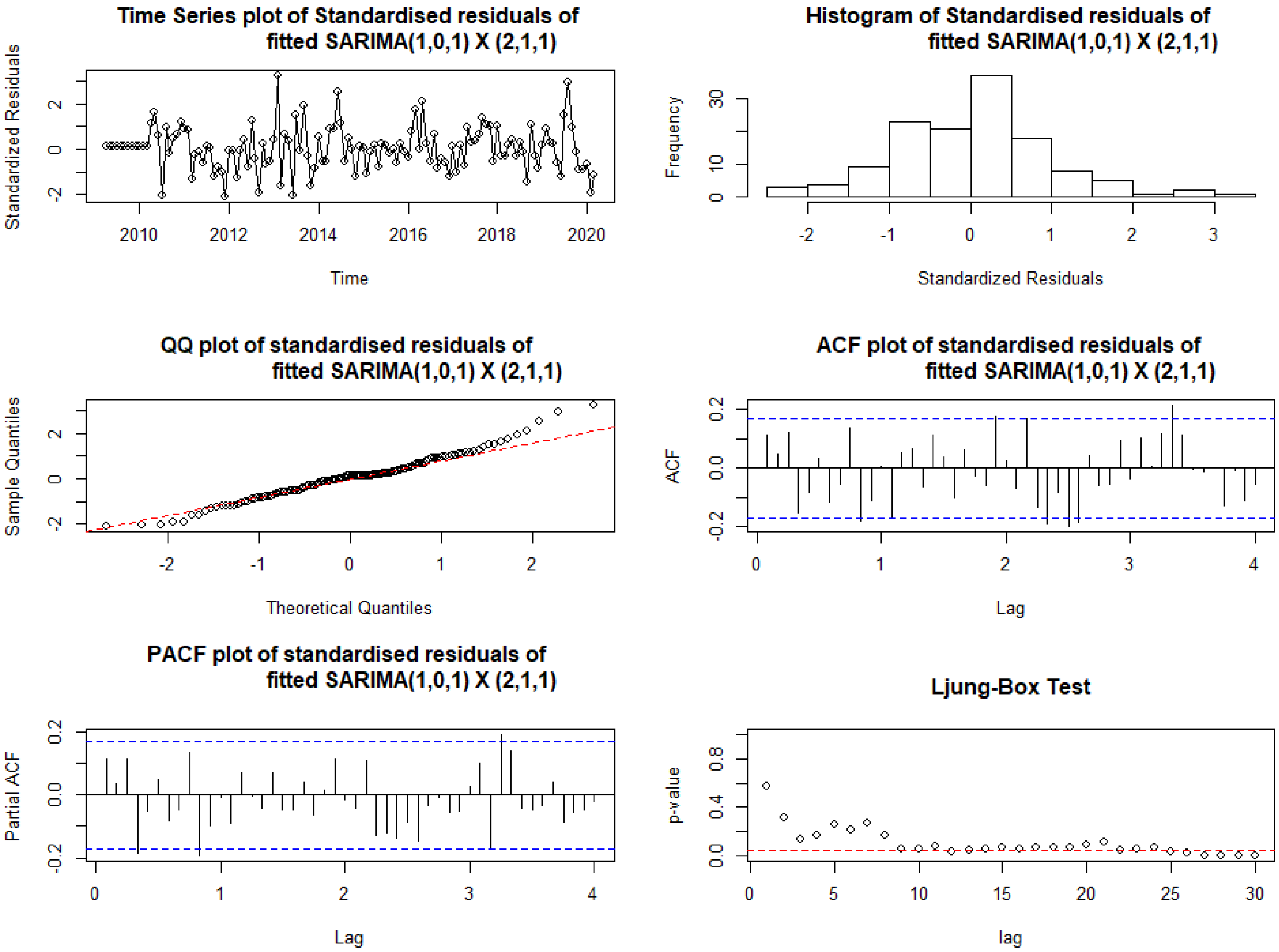


Figure 28 - Residual Diagnostics of SARIMA(1,0,1) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.97806, p-value = 0.0311
```

Figure 29 - Normality Check of SARIMA(1,0,1) X (2,1,1)

SARIMA(1,0,2)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(1,0,2,"SARIMA(1,0,2) X (2,1,1)")

z test of coefficients:

       Estimate   Std. Error    z value  Pr(>|z|)
ar1   1.0000e+00  1.6542e-05 60450.5097 < 2.2e-16 ***
ma1  -8.1168e-01  9.4433e-02    -8.5953 < 2.2e-16 ***
ma2  -1.4836e-01  9.4727e-02    -1.5662    0.1173
sar1  4.1322e-03  1.1788e-01     0.0351    0.9720
sar2  3.2363e-03  1.1952e-01     0.0271    0.9784
sma1 -9.6688e-01  2.0654e-01    -4.6812 2.852e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 30 - Co-efficient of SARIMA(1,0,2) x (2,1,1)

RESIDUAL CHECKING

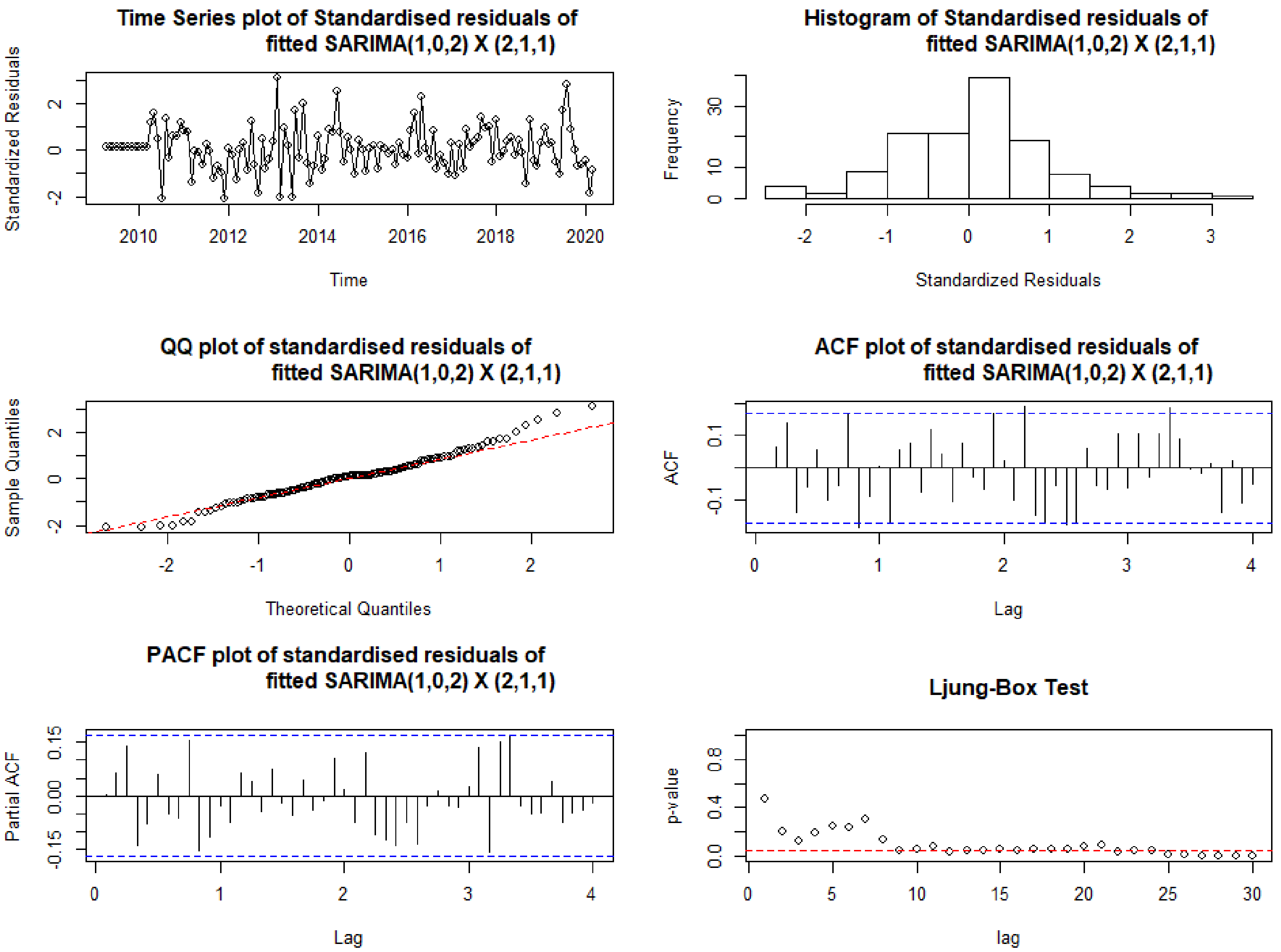


Figure 31 - Residual Diagnostics of SARIMA(1,0,2) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
	Shapiro-Wilk normality test

data:  res.model
W = 0.97942, p-value = 0.04267
```

Figure 32 - Normality Check of SARIMA(1,0,2) X (2,1,1)

SARIMA(1,0,3)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(1,0,3,"SARIMA(1,0,3) X (2,1,1)")

z test of coefficients:

        Estimate  Std. Error    z value  Pr(>|z|)    
ar1   1.0000e+00  1.2564e-05 79591.5918 < 2.2e-16 ***
ma1  -8.4548e-01  1.0805e-01    -7.8251 5.073e-15 ***
ma2   8.4623e-04  1.9124e-01     0.0044    0.9965    
ma3  -1.2734e-01  1.3724e-01    -0.9279    0.3535    
sar1  1.4141e-03  1.1994e-01     0.0118    0.9906    
sar2 -4.6263e-02  1.3137e-01    -0.3522    0.7247    
sma1 -9.4552e-01  1.9010e-01    -4.9738 6.566e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 33 - Co-efficient of SARIMA (1,0,3) x (2,1,1)

RESIDUAL CHECKING

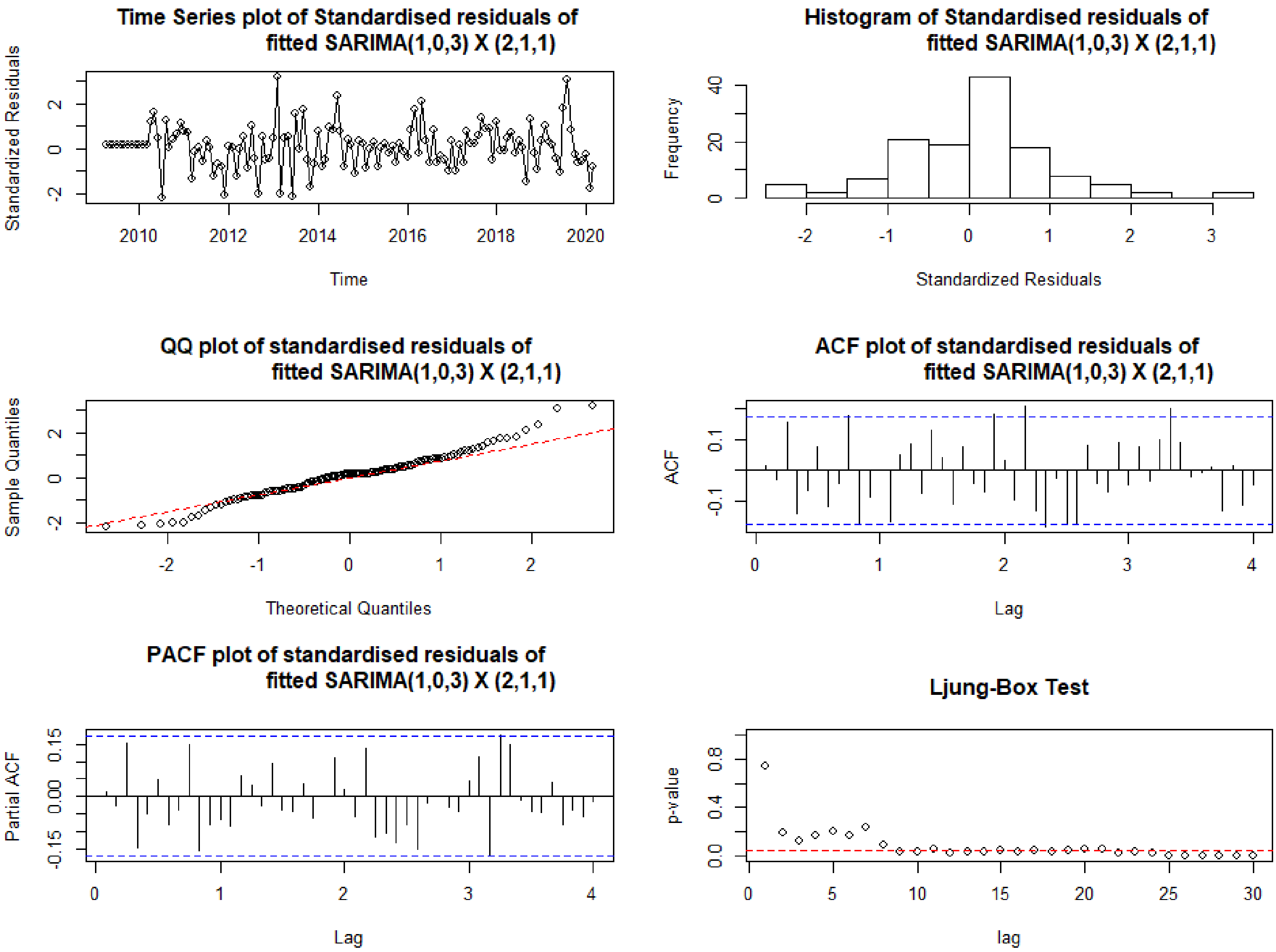


Figure 34 - Residual Diagnostics of SARIMA(1,0,3) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.97539, p-value = 0.01688
```

Figure 35 - Normality Check of SARIMA(1,0,3) X (2,1,1)

SARIMA(2,0,1)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(2,0,1,"SARIMA(2,0,1) X (2,1,1)")

z test of coefficients:

        Estimate  Std. Error z value  Pr(>|z|)
ar1   0.93632974          NA      NA        NA
ar2   0.05182068          NA      NA        NA
ma1  -0.68955077          NA      NA        NA
sar1 -0.06930528          NA      NA        NA
sar2  0.04488766  0.00061872   72.55 < 2.2e-16 ***
sma1 -0.76668420          NA      NA        NA
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 36 - Co-efficient of SARIMA(2,0,1) x (2,1,1)

RESIDUAL CHECKING

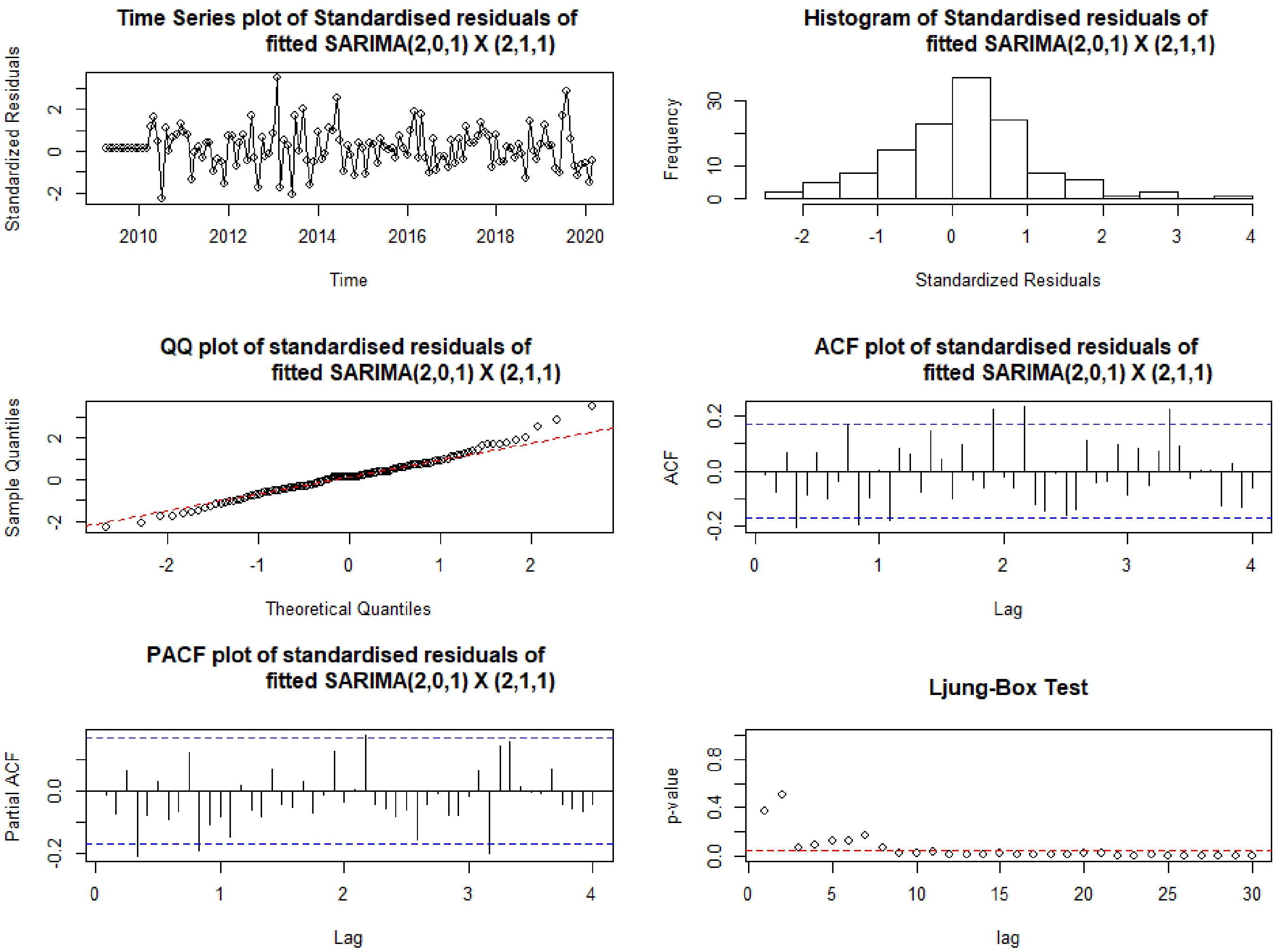


Figure 37 - Residual Diagnostics of SARIMA(2,0,1) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.98253, p-value = 0.08815
```

Figure 38 - Normality Check of SARIMA(2,0,1) X (2,1,1)

SARIMA(2,0,2)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(2,0,2,"SARIMA(2,0,2) X (2,1,1)")

z test of coefficients:

      Estimate Std. Error z value  Pr(>|z|)    
ar1   0.281124   0.176249  1.5950 0.1107044    
ar2   0.718853   0.176217  4.0794 4.516e-05 ***
ma1  -0.103874   0.156349 -0.6644 0.5064518    
ma2  -0.795486   0.136691 -5.8196 5.900e-09 ***
sar1 -0.027186   0.132147 -0.2057 0.8370070    
sar2  0.046374   0.127994  0.3623 0.7171180    
sma1 -0.927949   0.251675 -3.6871 0.0002268 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 39 - Co-efficient of SARIMA(2,0,2) x (2,1,1)

RESIDUAL CHECKING

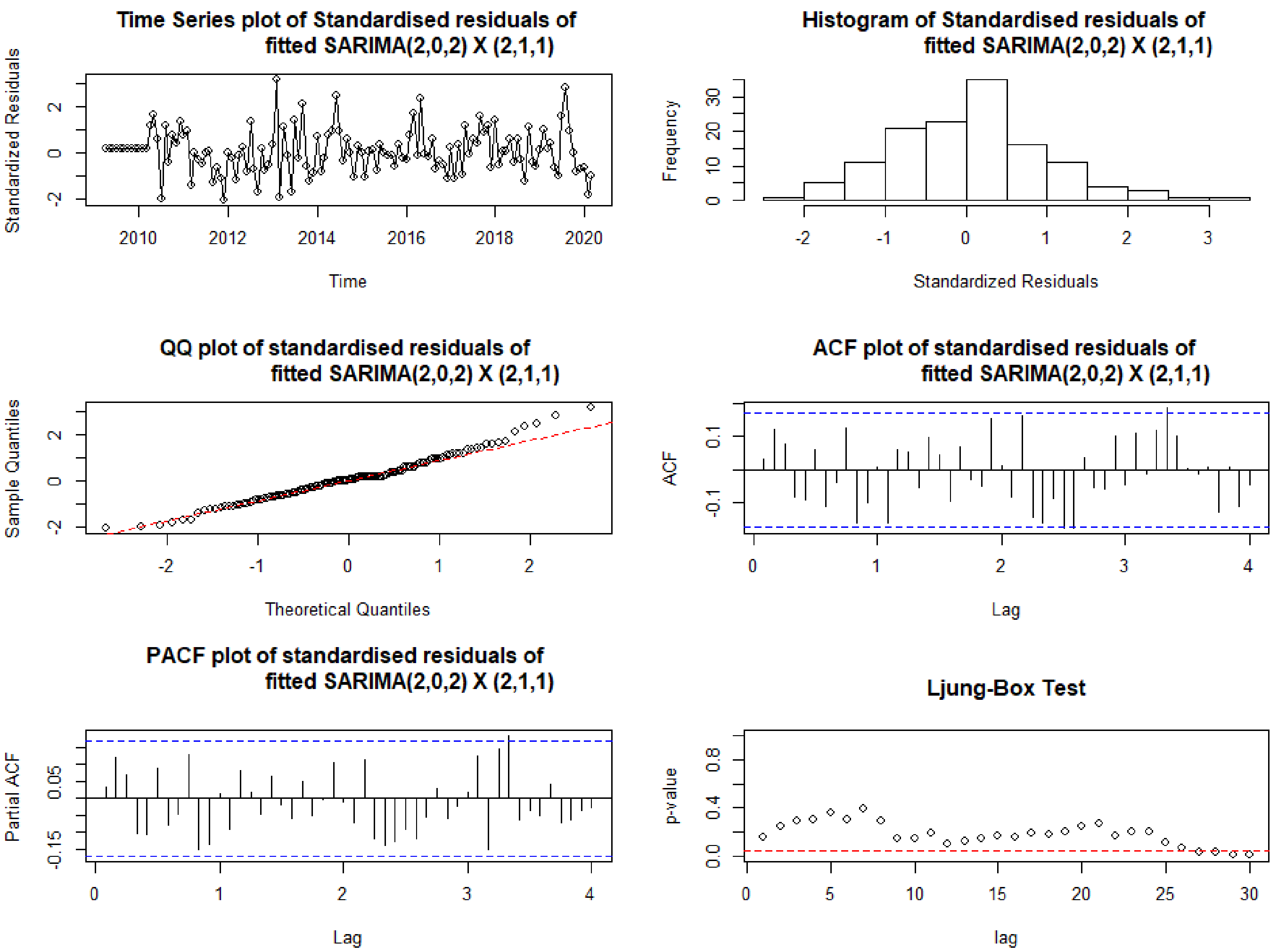


Figure 40 - Residual Diagnostics of SARIMA(2,0,2) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.98009, p-value = 0.04984
```

Figure 41 - Normality Check of SARIMA(2,0,2) X (2,1,1)

SARIMA(2,0,3)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(2,0,3,"SARIMA(2,0,3) X (2,1,1)")

z test of coefficients:

       Estimate Std. Error  z value  Pr(>|z|)
ar1   0.0027304         NA       NA        NA
ar2   0.9972674         NA       NA        NA
ma1   0.1881071  0.0989503   1.9010    0.0573 .
ma2  -0.9600614  0.0702775 -13.6610 < 2.2e-16 ***
ma3  -0.1483333  0.0949616  -1.5620    0.1183
sar1  0.0040930  0.1179262   0.0347    0.9723
sar2  0.0029454  0.1194263   0.0247    0.9803
sma1 -0.9686962  0.2222691  -4.3582 1.311e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 42 - Co-efficient of SARIMA(2,0,3) x (2,1,1)

RESIDUAL CHECKING

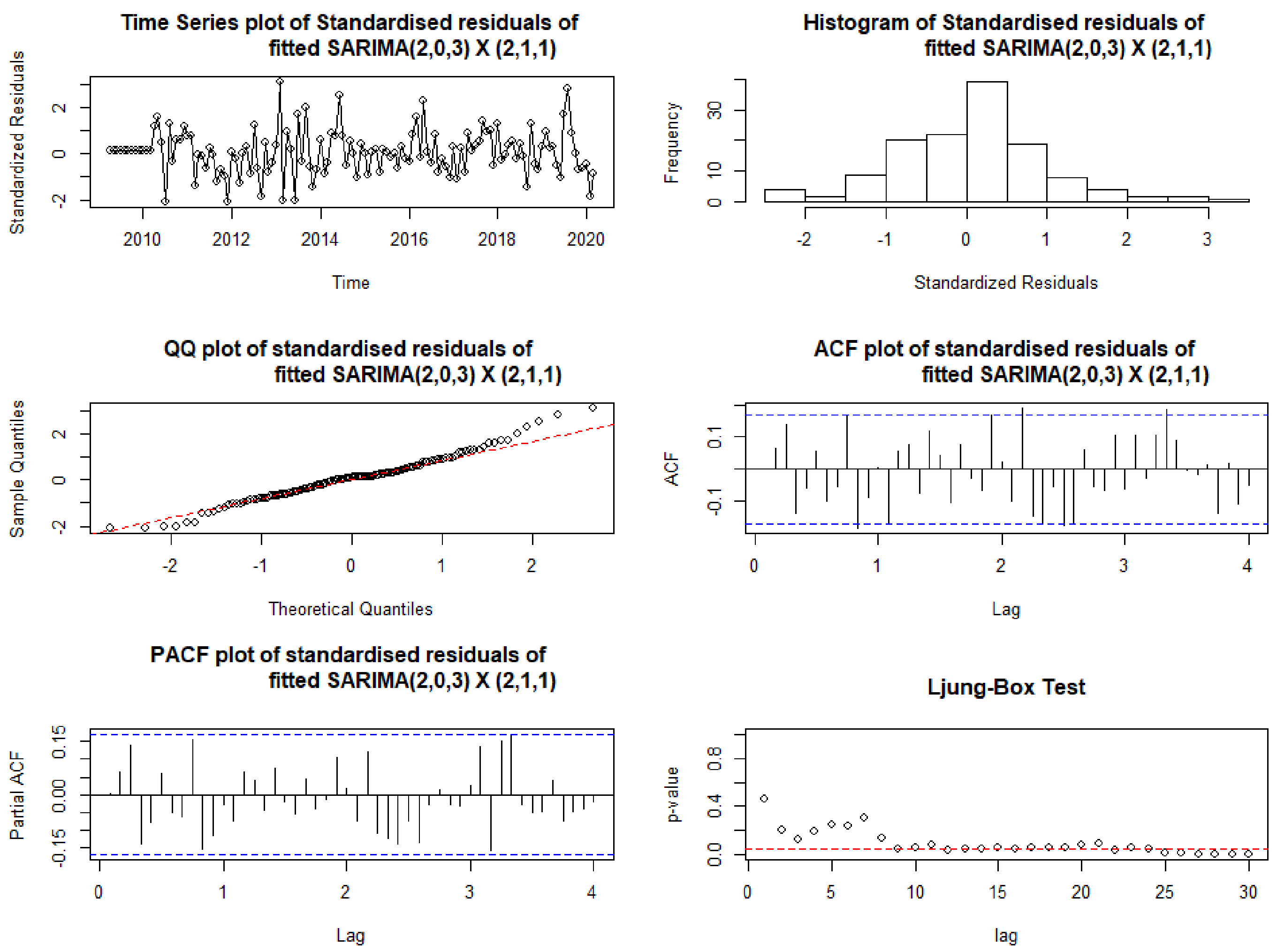


Figure 43 - Residual Diagnostics of SARIMA(2,0,3) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.97944, p-value = 0.04282
```

Figure 44 - Normality Check of SARIMA(2,0,3) X (2,1,1)

SARIMA(3,0,3)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(3,0,3,"SARIMA(3,0,3) X (2,1,1)")

z test of coefficients:

      Estimate Std. Error  z value  Pr(>|z|)    
ar1  -0.1517219  0.0129505 -11.7155 < 2.2e-16 ***
ar2   0.1523452  0.0089744  16.9756 < 2.2e-16 ***
ar3   0.9993669  0.0091746 108.9278 < 2.2e-16 ***
ma1   0.3596385  0.2229229   1.6133  0.106682    
ma2   0.1122161  0.4205575   0.2668  0.789602    
ma3  -0.7793285  0.2470279  -3.1548  0.001606 ** 
sar1 -0.1387942  0.1244101  -1.1156  0.264586    
sar2  0.0539036  0.1203240   0.4480  0.654163    
sma1 -0.9889562  0.2171524  -4.5542 5.258e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 45 - Co-efficient of SARIMA(3,0,3) x (2,1,1)

RESIDUAL CHECKING

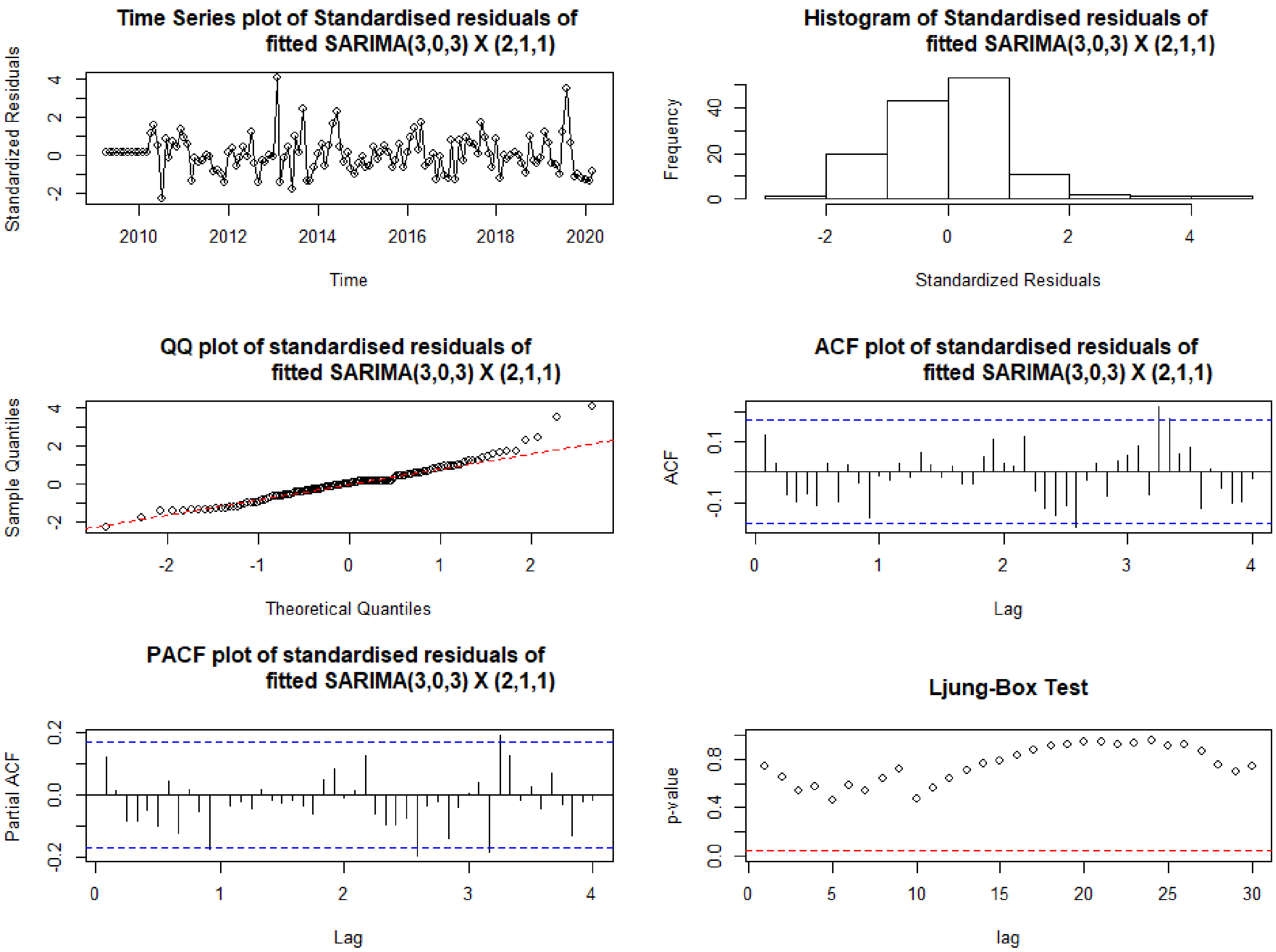


Figure 46 - Residual Diagnostics of SARIMA(3,0,3) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.95043, p-value = 0.0001078
```

Figure 47 - Normality Check of SARIMA(3,0,3) X (2,1,1)

DETERMINING BEST MODEL

AIC SCORE

```
> AIC_SCORE %>% unique() %>%  arrange(AIC)
                     Model       AIC
1 SARIMA(3,0,3) X (2,1,1) -439.5186
2 SARIMA(1,0,2) X (2,1,1) -432.5000
3 SARIMA(1,0,1) X (2,1,1) -432.0408
4 SARIMA(1,0,3) X (2,1,1) -431.4212
5 SARIMA(2,0,2) X (2,1,1) -431.3913
6 SARIMA(2,0,3) X (2,1,1) -428.5096
7 SARIMA(2,0,1) X (2,1,1) -424.8546
8 SARIMA(0,0,1) X (2,1,1) -390.2726
9 SARIMA(0,0,0) X (2,1,1) -386.3205
```

Figure 48 - AIC Sort

BIC SCORE

```
> BIC_SCORE %>% unique() %>%  arrange(BIC)
                     Model       BIC
1 SARIMA(1,0,1) X (2,1,1) -414.7440
2 SARIMA(1,0,2) X (2,1,1) -412.3204
3 SARIMA(3,0,3) X (2,1,1) -410.6905
4 SARIMA(1,0,3) X (2,1,1) -408.3588
5 SARIMA(2,0,2) X (2,1,1) -408.3289
6 SARIMA(2,0,1) X (2,1,1) -404.6750
7 SARIMA(2,0,3) X (2,1,1) -402.5644
8 SARIMA(0,0,1) X (2,1,1) -375.8586
9 SARIMA(0,0,0) X (2,1,1) -374.7893
```

Figure 49 - BIC Sort

Table 1 – Finding the Best SARIMA model

OVERFITTING THE BEST MODEL

SARIMA(3,0,4)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(3,0,4,"SARIMA(3,0,4) X (2,1,1)")

z test of coefficients:

       Estimate Std. Error  z value  Pr(>|z|)    
ar1  -0.1505894  0.0100010 -15.0575 < 2.2e-16 ***
ar2   0.1537674  0.0099960  15.3829 < 2.2e-16 ***
ar3   0.9968101  0.0061151 163.0088 < 2.2e-16 ***
ma1   0.4083850  0.0901214   4.5315 5.857e-06 ***
ma2  -0.0205230  0.0811787  -0.2528   0.80041    
ma3  -0.9615760  0.0811032 -11.8562 < 2.2e-16 ***
ma4  -0.2226861  0.0923130  -2.4123   0.01585 *  
sar1 -0.1067757  0.1193863  -0.8944   0.37112    
sar2  0.0414166  0.1197199   0.3459   0.72938    
sma1 -0.9580591  0.2026111  -4.7286 2.261e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 50 - Co-efficient of SARIMA(3,0,4) x (2,1,1)

RESIDUAL CHECKING

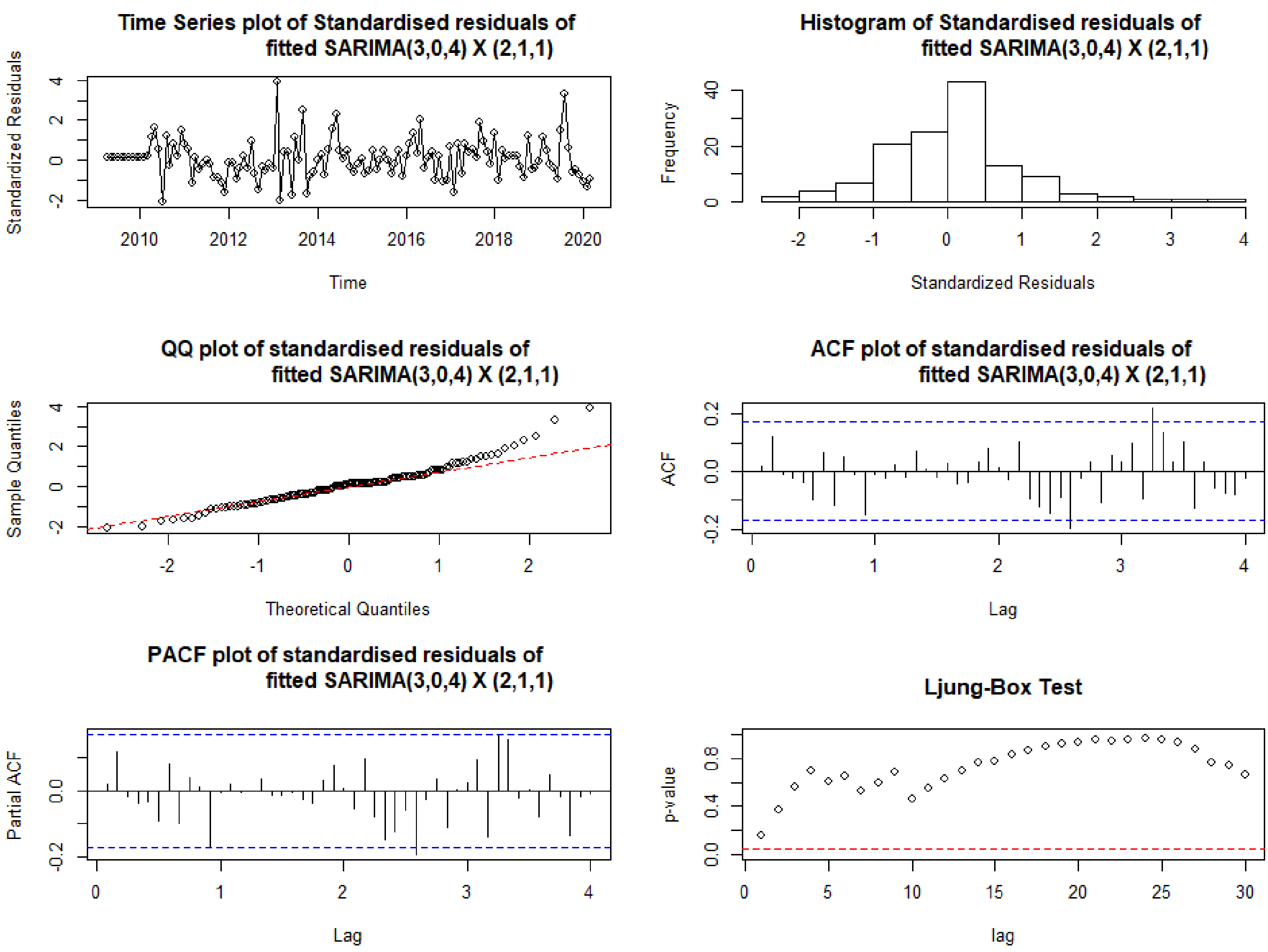


Figure 51 - Residual Diagnostics of SARIMA(3,0,4) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.95701, p-value = 0.0003632
```

Figure 52 - Normality Check of SARIMA(3,0,4) X (2,1,1)

SARIMA(4,0,3)x(2,1,1)

SARIMA(4,0,4)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(4,0,4,"SARIMA(4,0,4) X (2,1,1)")

z test of coefficients:

      Estimate Std. Error z value  Pr(>|z|)
ar1  -0.072412   0.487761 -0.1485 0.8819819
ar2  -0.050970   0.167672 -0.3040 0.7611372
ar3   0.752761   0.137208  5.4863 4.104e-08 ***
ar4   0.370602   0.459546  0.8065 0.4199817
ma1   0.263957   0.449746  0.5869 0.5572698
ma2   0.171707   0.142561  1.2045 0.2284147
ma3  -0.606289   0.203130 -2.9847 0.0028382 **
ma4  -0.619906   0.380230 -1.6303 0.1030286
sar1 -0.115555   0.130668 -0.8843 0.3765151
sar2  0.065919   0.156365  0.4216 0.6733370
sma1 -0.950250   0.246924 -3.8484 0.0001189 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 53 - Co-efficient of SARIMA(4,0,4) x (2,1,1)

RESIDUAL CHECKING

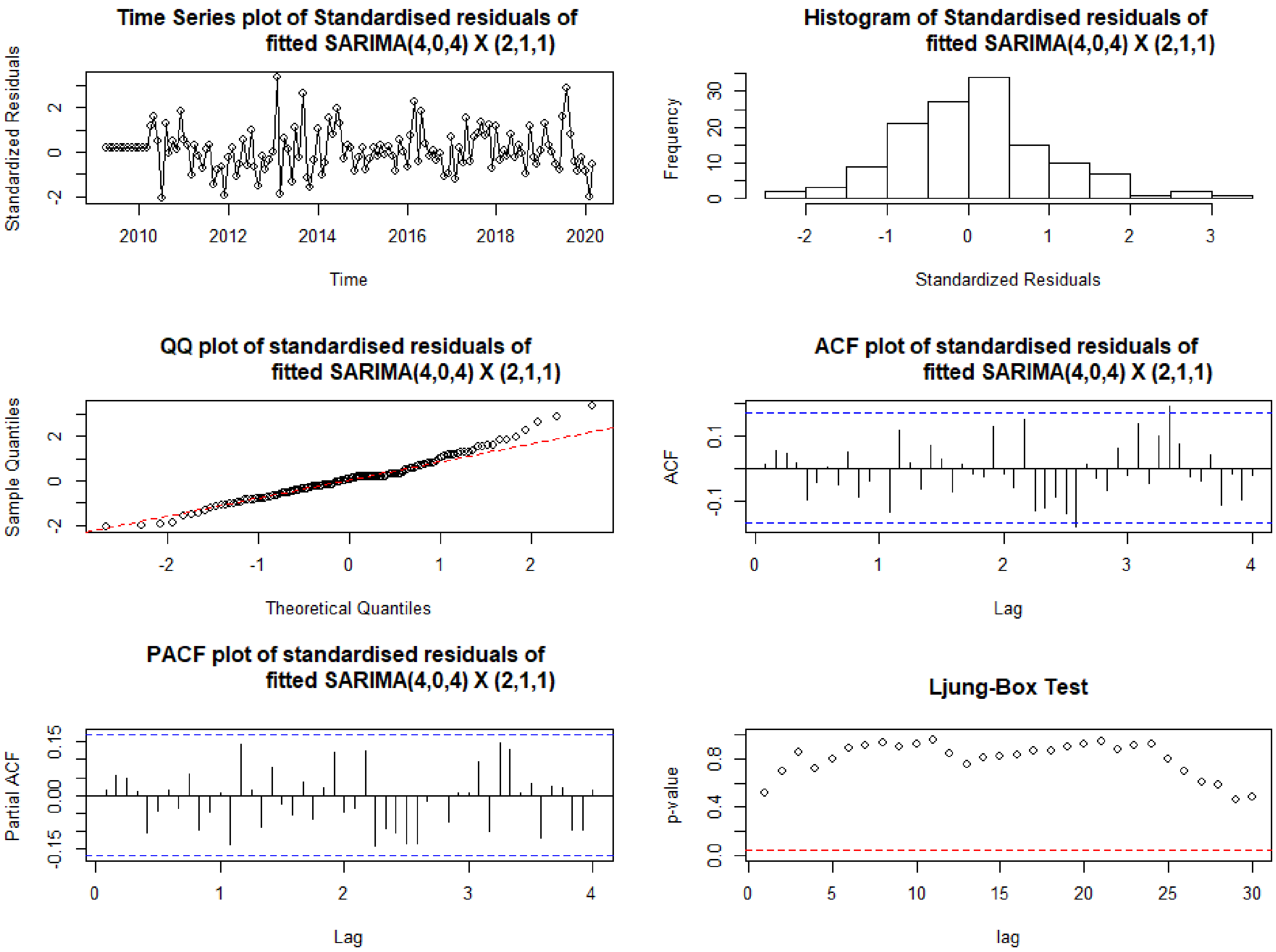


Figure 54 - Residual Diagnostics of SARIMA(4,0,4) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.97371, p-value = 0.01153
```

Figure 55 - Normality Check of SARIMA(4,0,4) X (2,1,1)

SARIMA(3,0,5)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(3,0,5,"SARIMA(3,0,5) X (2,1,1)")

z test of coefficients:

       Estimate Std. Error  z value  Pr(>|z|)    
ar1  -0.1527929  0.0078632 -19.4314 < 2.2e-16 ***
ar2   0.1537501  0.0078073  19.6931 < 2.2e-16 ***
ar3   0.9990400  0.0028823 346.6143 < 2.2e-16 ***
ma1   0.4237958  0.1026608   4.1281 3.657e-05 ***
ma2   0.1296202  0.1210022   1.0712   0.28407    
ma3  -0.9674829  0.0889546 -10.8761 < 2.2e-16 ***
ma4  -0.2731795  0.1075640  -2.5397   0.01109 *  
ma5  -0.1837491  0.1042266  -1.7630   0.07790 .  
sar1 -0.1272861  0.1185208  -1.0740   0.28284    
sar2 -0.0417562  0.1299078  -0.3214   0.74789    
sma1 -0.9632334  0.2073890  -4.6446 3.408e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 56 - Co-efficient of SARIMA(3,0,5) x (2,1,1)

RESIDUAL CHECKING

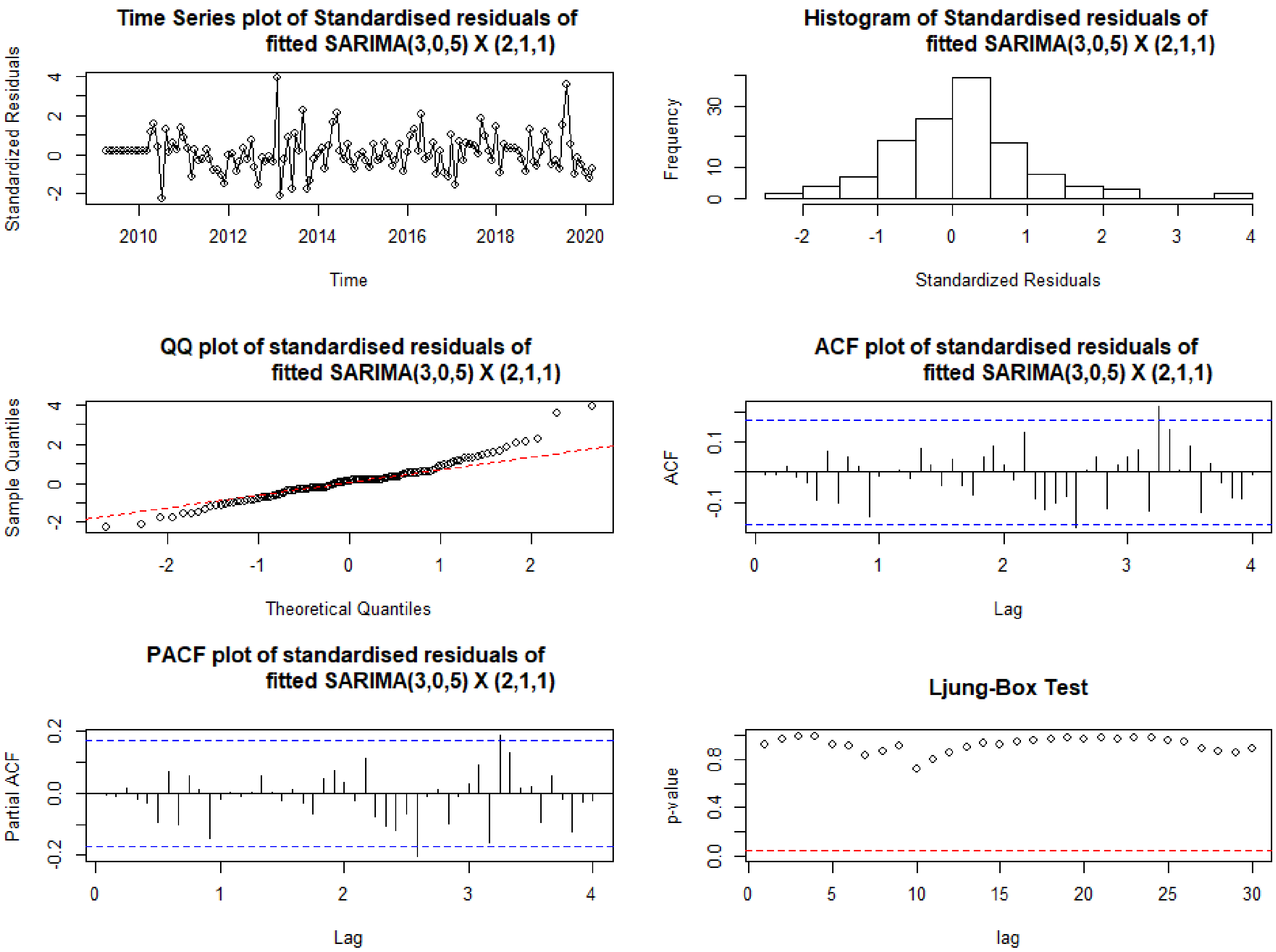


Figure 57 - Residual Diagnostics of SARIMA(3,0,5) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.95171, p-value = 0.0001357
```

Figure 58 - Normality Check of SARIMA(3,0,5) X (2,1,1)

SARIMA(3,0,6)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(3,0,6,"SARIMA(3,0,6) X (2,1,1)")

z test of coefficients:

      Estimate Std. Error  z value  Pr(>|z|)    
ar1  -0.1525919  0.0083372 -18.3025 < 2.2e-16 ***
ar2   0.1534140  0.0080786  18.9902 < 2.2e-16 ***
ar3   0.9991706  0.0028427 351.4880 < 2.2e-16 ***
ma1   0.4295072  0.1031968   4.1620 3.154e-05 ***
ma2   0.1390025  0.1236046   1.1246   0.26077    
ma3  -0.9179395  0.1195562  -7.6779 1.617e-14 ***
ma4  -0.2760209  0.1077427  -2.5619   0.01041 *  
ma5  -0.2051392  0.1132513  -1.8114   0.07008 .  
ma6  -0.0527767  0.1030472  -0.5122   0.60854    
sar1 -0.1269067  0.1227927  -1.0335   0.30137    
sar2 -0.0343617  0.1336352  -0.2571   0.79708    
sma1 -0.9372192  0.1982607  -4.7272 2.276e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 59 - Co-efficient of SARIMA(3,0,6) x (2,1,1)

RESIDUAL CHECKING

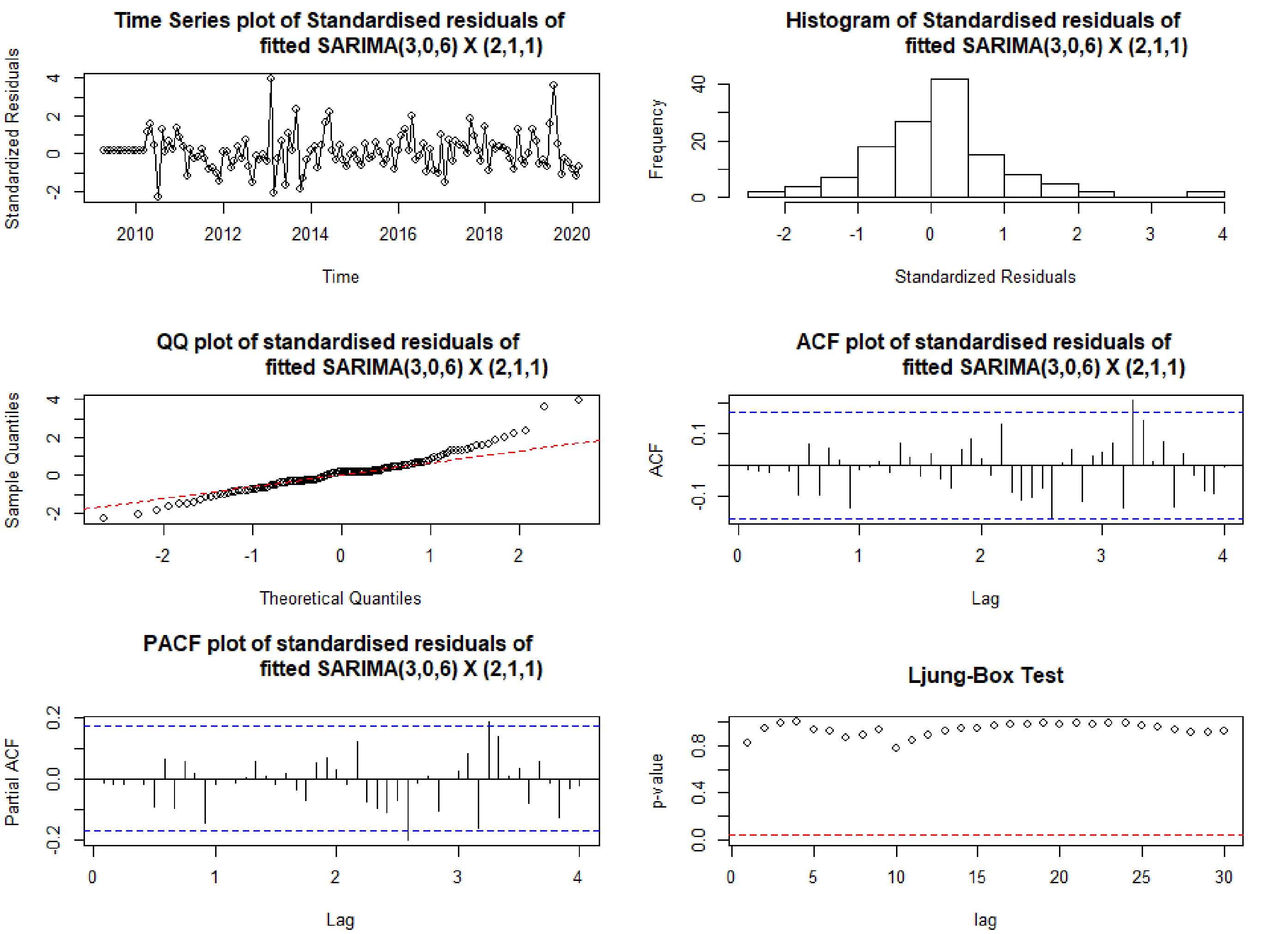


Figure 60 - Residual Diagnostics of SARIMA(3,0,6) x (2,1,1)

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.94998, p-value = 9.937e-05
```

Figure 61 - Normality Check of SARIMA(3,0,6) X (2,1,1)

SARIMA(4,0,5)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(4,0,5,"SARIMA(4,0,5) X (2,1,1)")

z test of coefficients:

      Estimate Std. Error z value  Pr(>|z|)    
ar1  -0.383847   0.295875 -1.2973 0.1945177    
ar2   0.038847   0.129270  0.3005 0.7637881    
ar3   0.706619   0.081018  8.7217 < 2.2e-16 ***
ar4   0.638359   0.260287  2.4525 0.0141860 *  
ma1   0.590208   0.312166  1.8907 0.0586665 .  
ma2   0.167342   0.105566  1.5852 0.1129244    
ma3  -0.540758   0.117378 -4.6070 4.085e-06 ***
ma4  -0.864293   0.240665 -3.5913 0.0003291 ***
ma5  -0.137408   0.161472 -0.8510 0.3947874    
sar1 -0.094113   0.137926 -0.6823 0.4950220    
sar2  0.056516   0.148205  0.3813 0.7029546    
sma1 -0.940883   0.230477 -4.0823 4.459e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 62 - Co-efficient of SARIMA(4,0,5) x (2,1,1)

RESIDUAL CHECKING

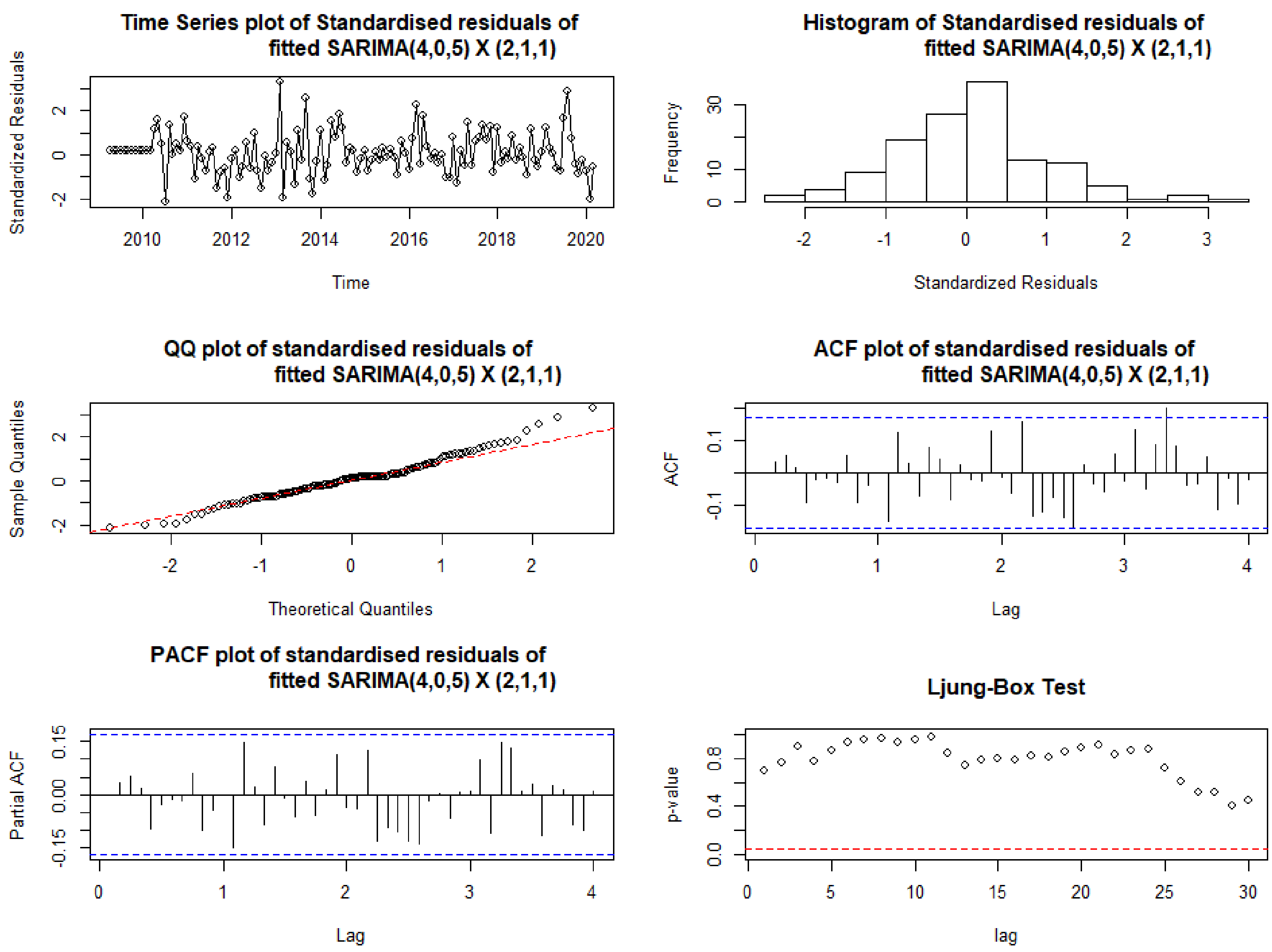


Figure 63 - Residual Diagnostics of SARIMA(4,0,5) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.97563, p-value = 0.0178
```

Figure 64 - Normality Check of SARIMA(4,0,5) X (2,1,1)

SARIMA(4,0,6)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(4,0,6,"SARIMA(4,0,6) X (2,1,1)")

z test of coefficients:

      Estimate Std. Error z value  Pr(>|z|)
ar1  -0.022957   0.551648 -0.0416  0.966805
ar2  -0.138545   0.178302 -0.7770  0.437143
ar3   0.800530   0.130619  6.1287 8.857e-10 ***
ar4   0.360971   0.497126  0.7261  0.467768
ma1   0.230534   0.572367  0.4028  0.687115
ma2   0.344548   0.148312  2.3231  0.020172 *
ma3  -0.589068   0.233243 -2.5256  0.011552 *
ma4  -0.630578   0.440343 -1.4320  0.152140
ma5  -0.087285   0.205955 -0.4238  0.671707
ma6  -0.183299   0.112767 -1.6255  0.104064
sar1 -0.073772   0.135437 -0.5447  0.585963
sar2  0.062513   0.130025  0.4808  0.630678
sma1 -0.958793   0.307074 -3.1224  0.001794 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 65 - Co-efficient of SARIMA(4,0,6) x (2,1,1)

RESIDUAL CHECKING

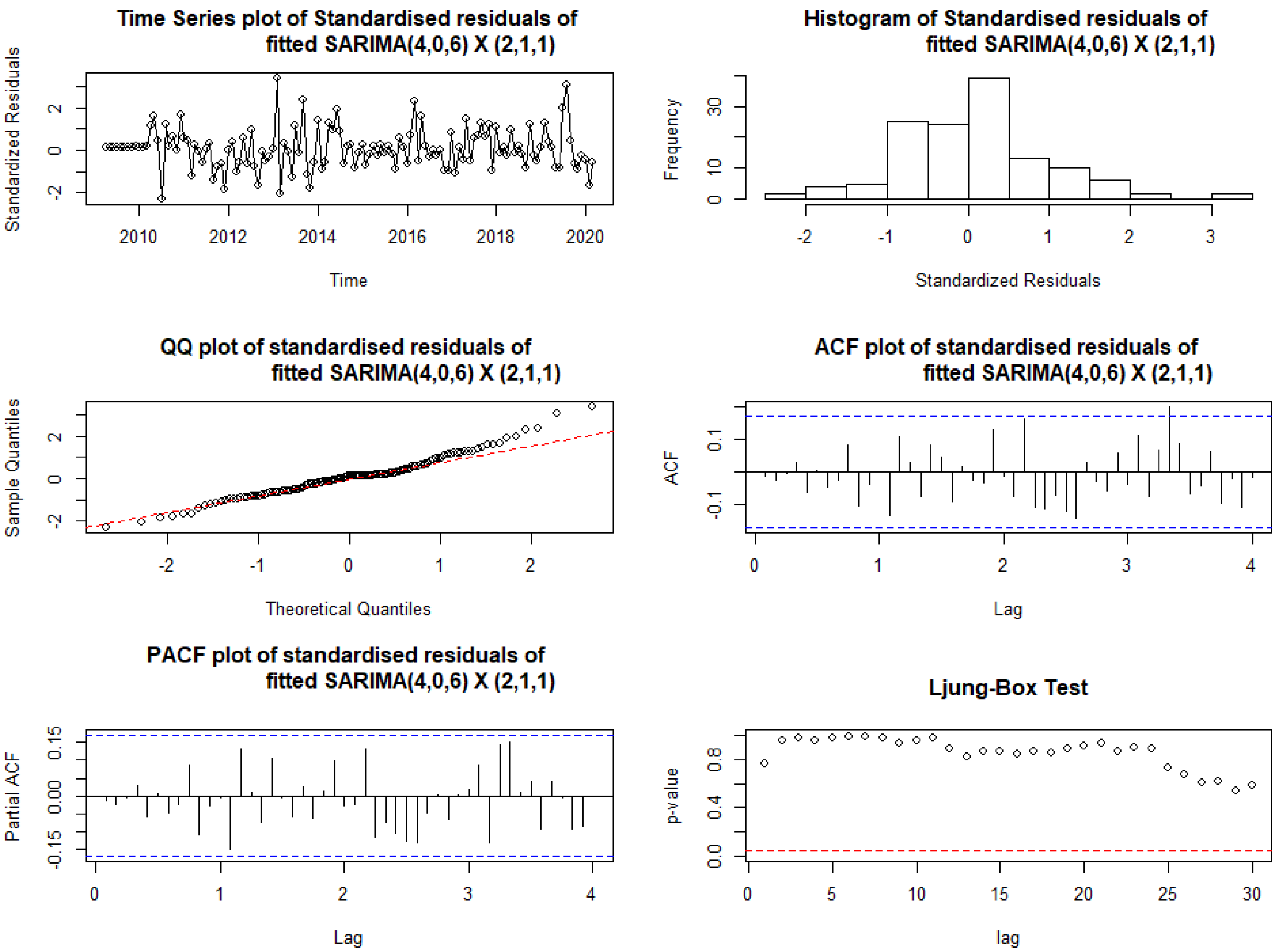


Figure 66 - Residual Diagnostics of SARIMA(4,0,6) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.97143, p-value = 0.00696
```

Figure 67 - Normality Check of SARIMA(4,0,6) X (2,1,1)

SARIMA(5,0,5)x(2,1,1)

CO-EFFICIENTS

```
> SARIMA(5,0,5,"SARIMA(5,0,5) X (2,1,1)")

z test of coefficients:

      Estimate Std. Error z value  Pr(>|z|)    
ar1  -0.954552   0.106139 -8.9934 < 2.2e-16 ***
ar2  -0.035807   0.079533 -0.4502  0.652549    
ar3   0.239283   0.077250  3.0975  0.001951 ** 
ar4   0.960748   0.079270 12.1199 < 2.2e-16 ***
ar5   0.790282   0.096189  8.2159 < 2.2e-16 ***
ma1   1.117702   0.104752 10.6700 < 2.2e-16 ***
ma2   0.265995   0.171310  1.5527  0.120492    
ma3  -0.044761   0.165440 -0.2706  0.786732    
ma4  -0.989144   0.158940 -6.2234 4.865e-10 ***
ma5  -0.909268   0.128291 -7.0875 1.365e-12 ***
sar1 -0.028621   0.046525 -0.6152  0.538432    
sar2 -0.012169   0.109768 -0.1109  0.911730    
sma1 -0.946262         NA      NA        NA    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 68 - Co-efficient of SARIMA(5,0,5) x (2,1,1)

RESIDUAL CHECKING

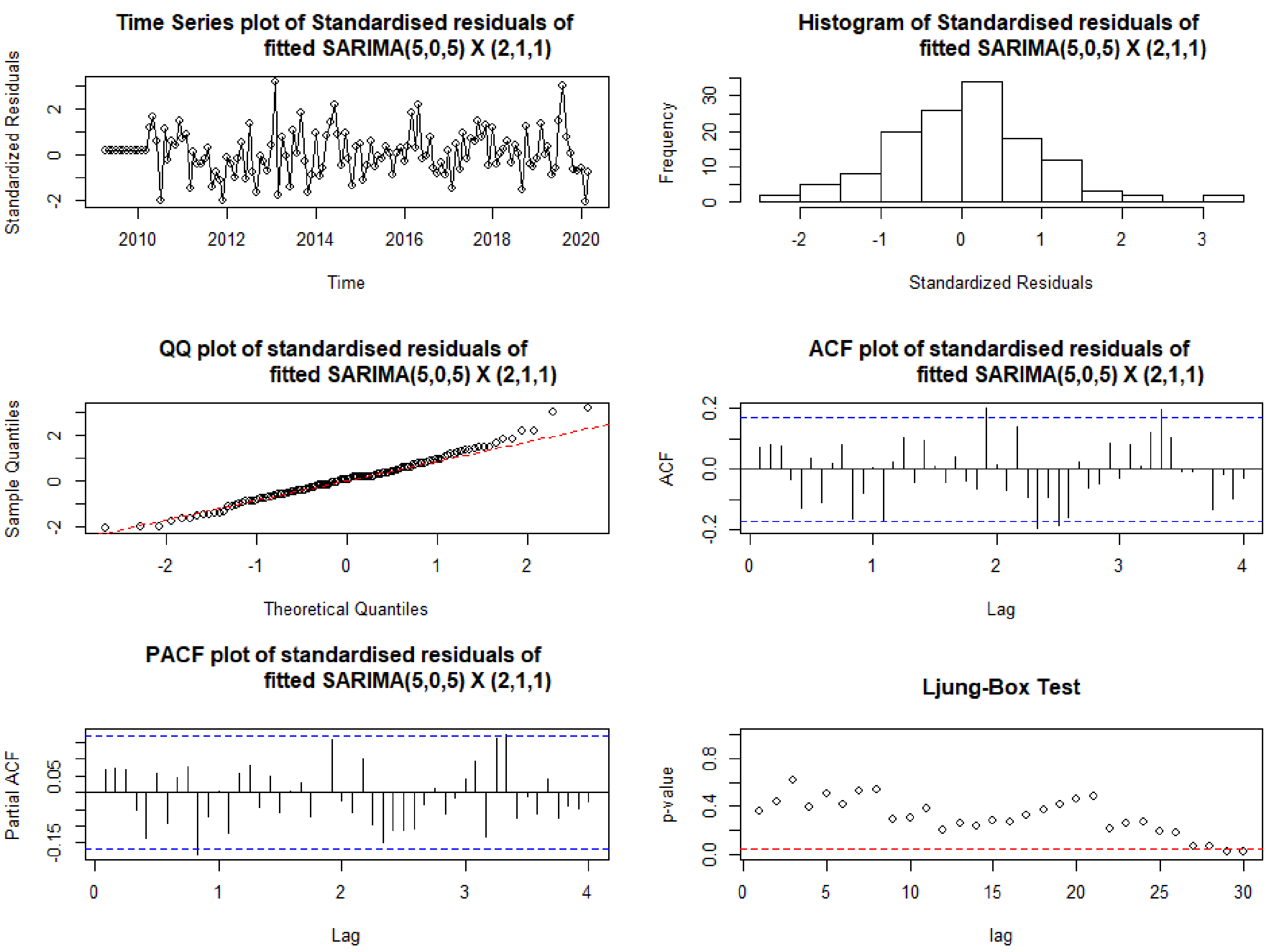


Figure 69 - Residual Diagnostics of SARIMA(5,0,5) x (2,1,1)

TIME – SERIES

HISTOGRAM

QQ-PLOT

ACF

PACF

Ljung- BOX

SHAPIRO-WILK TEST

```
        Shapiro-Wilk normality test

data:  res.model
W = 0.98472, p-value = 0.1467
```

Figure 70 - Normality Check of SARIMA(5,0,5) X (2,1,1)

FINAL SARIMA MODEL

PARAMETER ESTIMATION

FORECASTING

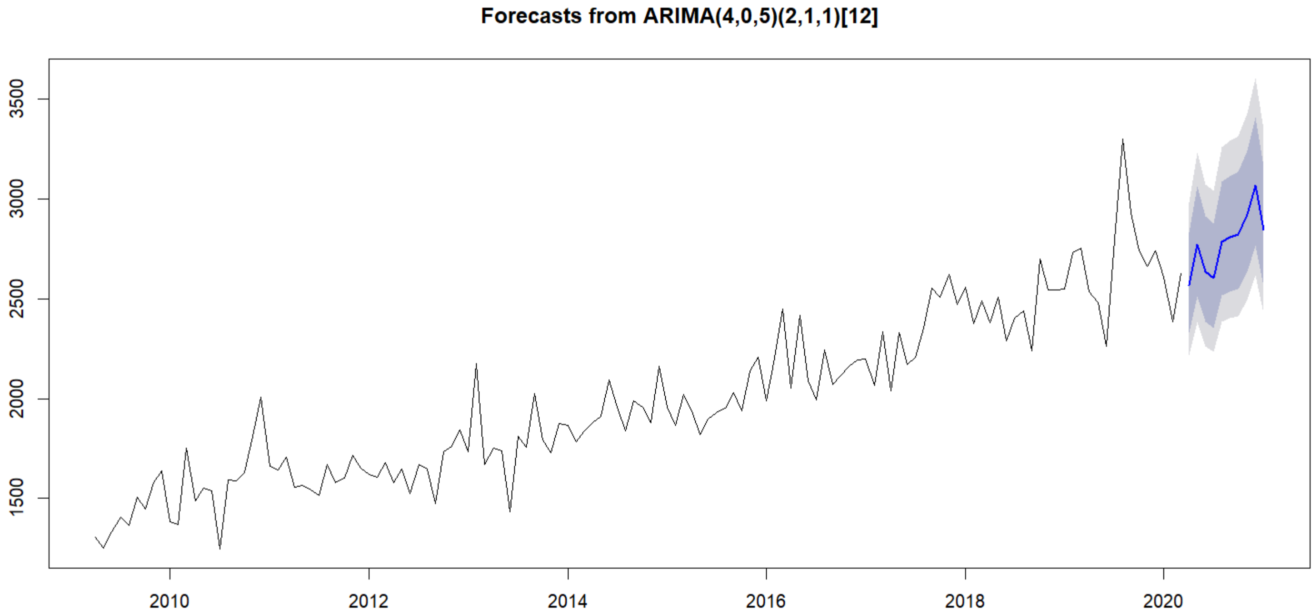


Figure 71 - Prediction Plot

```
> prediction
         Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
Apr 2020       2568.795 2332.496 2831.694 2217.168 2982.742
May 2020       2773.080 2510.978 3065.592 2383.356 3234.055
Jun 2020       2634.548 2384.835 2913.325 2263.276 3073.918
Jul 2020       2601.815 2353.675 2879.036 2232.948 3038.822
Aug 2020       2784.644 2515.860 3085.354 2385.230 3258.867
Sep 2020       2809.154 2535.528 3115.613 2402.653 3292.593
Oct 2020       2824.182 2547.365 3134.448 2413.018 3313.729
Nov 2020       2921.023 2633.689 3243.218 2494.283 3429.454
Dec 2020       3069.590 2766.116 3410.090 2618.949 3607.000
Jan 2021       2844.924 2562.907 3161.454 2426.178 3344.549
```

Figure 72 - Prediction Values for next 10 months

CONCLUSION

REFERENCES

Class Presentations (1-9)

Class Lecture - Notes (1-9)

Class Recordings (1-9)

APPENDIX

ABBREVIATION

ARIMA = Auto-Regressive Integrated Moving Average

AR = Auto-Regressive,

MA = Moving Average

BC = Box-Cox

adf = Augmented Dicky-Fuller Test

pp = Phillops-Perron Test

CI = Condifence Interval

AIC = Akaike Information Criterion

BIC = Bayesian Information Criterion

ACF = AutoCorrelation Function

PACF = Partial AutoCorrelation Function

EACF = Extended AutoCorrelation Function

BIC Table = Bayesian Information Criterion

TS = Time-Series

CODE

```
# Clearing the environment
rm(list=ls())

# libraries
library(readr)
library(dplyr)
library(tidyr)
library(TSA)
library(tseries)
library(forecast)
library(fUnitRoots)
library(lmtest)
library(fGarch)
library(CombMSC)
library(stats)
library(FitAR)

# data preparation
waste <- read_csv("C:/Users/Admin/Downloads/Waste_collected_per_month.csv")

waste <- waste %>% separate(date, into = c("month", "day","year"), sep = "/")  %>%
  separate(year, into = c("year", "time"), sep = " ") %>%
  arrange(year, month, day) %>% select(month, year, residential)

View(head(waste))
View(tail(waste))
```

```
# converting dataframe into time series
wasteTS = ts(waste$residential, start = c(2009,4), end = c(2020,3), frequency = 12)
wasteTS

#----------------------------------------------------------------------

# Functions for Time Series
plot_ts <- function(ts, transformation)
{
  win.graph(width = 45,
            height = 30,
            pointsize = 15)
  plot(ts,
       ylab = "Waste",
       main = c(paste0(toString(transformation),
                       " plot of Waste")),
       type="o")
  points(y=ts,x=time(ts), pch=as.vector(season(ts)))
}


# Function for ACF and PACF
autoCorrelation <- function(ts, time_series)
{
  win.graph(width = 20,
            height = 15,
            pointsize = 15)

  par(mfrow=c(2,1))
  acf(ts,
      lag.max = 48,
      main = c(paste0("ACF plot of ",toString(time_series))))

  pacf(ts,
       lag.max = 48,
       main = c(paste0("PACF plot of ",toString(time_series))))
  par(mfrow=c(1,1))
}


# Functions for its Residual Analysis
residual_model <- function(model, modelname)
{
  res.model = rstandard(model)
  win.graph(width = 20,
            height = 15,
            pointsize=15)

  par(mfrow=c(3,2))

  plot(y = res.model,
       x = as.vector(time(wasteTS)),
       main = c(paste0("Time Series plot of Standardised residuals of
                       fitted ", toString(modelname))),
       xlab = 'Time',
```

```
      ylab = 'Standardized Residuals',
      type = 'o')

  hist(res.model,
      main = c(paste0("Histogram of Standardised residuals of
                     fitted ", toString(modelname))),
      xlab='Standardized Residuals')

  qqnorm(y=res.model,
        main = c(paste0("QQ plot of standardised residuals of
                       fitted ", toString(modelname))))

  qqline(y=res.model,
        col = 2,
        lwd = 1,
        lty = 2)

  print(shapiro.test(res.model))

  acf(res.model,
     lag.max = 48,
     main = c(paste0("ACF plot of standardised residuals of
                    fitted ", toString(modelname))))

  pacf(res.model,
      lag.max = 48,
      main = c(paste0("PACF plot of standardised residuals of
                     fitted ", toString(modelname))))


  k=0
  LBQPlot(res.model,lag.max = 30, StartLag = k + 1, k = 0, SquaredQ = FALSE)
  par(mfrow=c(1,1))
}


# Function for SARIMA model
SARIMA <- function(p,d,q, modelname)
{
  ml = arima(wasteTSBC,order=c(p,d,q),
            seasonal = list(order = c(2,1,1), period = 12),
            method = "ML")

  print(coeftest(ml))
  residual_model(ml,toString(modelname))

  Model <- modelname

  AIC <- AIC(ml)
  aic <- cbind.data.frame(Model,AIC)
  AIC_SCORE <- rbind(AIC_SCORE, aic)
  AIC_SCORE <<- AIC_SCORE

  BIC <- AIC(ml,k=(log(length(wasteTS))))
  bic <- cbind.data.frame(Model,BIC)
```

```
  BIC_SCORE <- rbind(BIC_SCORE, bic)
  BIC_SCORE <<- BIC_SCORE
}

#----------------------------------------------------------------------
# Descriptive Analysis
#----------------------------------------------------------------------

# Original Time Series

par(mar=c(1,1,1,1))
plot_ts(wasteTS, "Time Series")


# ACF and PACF of original Time Series
autoCorrelation(wasteTS, "Time Series")


win.graph(width = 20,
          height = 15,
          pointsize=15)
# Scatter Plot of Consecutive Years
plot(y=wasteTS,
     x=zlag(wasteTS),
     ylab='Residential Waste (Tonnes)',
     xlab='Previous Month Residential Waste (Tonnes)',
     main = "Scatter plot of Residential Waste (Tonnes) in consequtive months")


# Correlation of Consecutive Years
y = wasteTS
x = zlag(wasteTS)
index = 2:length(x)
cor(y[index],x[index]) %>% round(3)

# ----------------------------------------------------------------------
# Box-cox Transformation
# ----------------------------------------------------------------------

win.graph(width = 20,
          height = 15,
          pointsize=15)
BC <- BoxCox.ar(wasteTS)

# box-cox confidence Interval
BC$ci

# maximum likelihood
lambda <- BC$lambda[which(max(BC$loglike) == BC$loglike)]
lambda

# box-cox transformation
wasteTSBC = ((wasteTS^lambda)-1)/lambda

plot_ts(wasteTSBC, "Box-Cox Transformation Time Series")
```

```
autoCorrelation(wasteTSBC, "Box-Cox Transformation Time Series")

# ------------------------------------------------------------------------
# Residual Approach SARIMA Models
# ------------------------------------------------------------------------

m1.temp = arima(wasteTSBC,order=c(0,0,0),
                seasonal = list(order = c(0,1,0), period = 12))
res.m1 = residuals(m1.temp)
plot_ts(res.m1,"Residuals SARIMA(0,0,0)x(0,1,0)")
autoCorrelation(res.m1,"Residuals SARIMA(0,0,0)x(0,1,0)")


m2.temp = arima(wasteTSBC,order=c(0,0,0),
                seasonal = list(order = c(2,1,1), period = 12))
res.m2 = residuals(m2.temp)
plot_ts(res.m2,"Residuals SARIMA(0,0,0)x(2,1,1)")
autoCorrelation(res.m2,"Residuals SARIMA(0,0,0)x(2,1,1)")

# P = 2, D = 1, Q = 1

# p = (1,2)
# q = (1,2,3)

# SARIMA(1,0,1)x(2,1,1),
# SARIMA(1,0,2)x(2,1,1),
# SARIMA(1,0,3)x(2,1,1),
# SARIMA(2,0,1)x(2,1,1),
# SARIMA(2,0,2)x(2,1,1),
# SARIMA(2,0,3)x(2,1,1),

m3.temp = arima(wasteTSBC,order=c(2,0,2),
                seasonal = list(order = c(2,1,1), period = 12))
res.m3 = residuals(m3.temp)
plot_ts(res.m3,"Residuals SARIMA(2,0,2)x(2,1,1)")
autoCorrelation(res.m3,"Residuals SARIMA(2,0,2)x(2,1,1)")

# ------------------------------------------------------------------------
# EACF
# ------------------------------------------------------------------------
eacf(res.m2)
# SARIMA(0,0,0)x(2,1,1),
# SARIMA(0,0,1)x(2,1,1),

# ------------------------------------------------------------------------
# BIC
# ------------------------------------------------------------------------
win.graph(width = 10,
          height = 10,
          pointsize = 15)
res = armasubsets(y=res.m2,
                  nar=8,
                  nma=8,
                  y.name='test',
```

```r
                  ar.method='ols')
plot(res)
# SARIMA(1,0,3)x(2,1,1),
# SARIMA(3,0,3)x(2,1,1),

# ----------------------------------------------------------------------
# set of possible SARIMA models;
# ----------------------------------------------------------------------

# SARIMA(0,0,0)x(2,1,1),
# SARIMA(0,0,1)x(2,1,1),
# SARIMA(1,0,1)x(2,1,1),
# SARIMA(1,0,2)x(2,1,1),
# SARIMA(1,0,3)x(2,1,1),
# SARIMA(2,0,1)x(2,1,1),
# SARIMA(2,0,2)x(2,1,1),
# SARIMA(2,0,3)x(2,1,1),
# SARIMA(3,0,3)x(2,1,1)

# ----------------------------------------------------------------------
# creating null data frame for AIC and BIC
# ----------------------------------------------------------------------

AIC_SCORE <- data.frame(Model = c(),
                        AIC = c())

BIC_SCORE <- data.frame(Model = c(),
                        BIC = c())

# ----------------------------------------------------------------------
# Model Fitting
# ----------------------------------------------------------------------

# SARIMA(0,0,0)x(2,1,1),
SARIMA(0,0,0,"SARIMA(0,0,0) X (2,1,1)")

# SARIMA(0,0,1)x(2,1,1)
SARIMA(0,0,1,"SARIMA(0,0,1) X (2,1,1)")

# SARIMA(1,0,1)x(2,1,1)
SARIMA(1,0,1,"SARIMA(1,0,1) X (2,1,1)")

# SARIMA(1,0,2)x(2,1,1)
SARIMA(1,0,2,"SARIMA(1,0,2) X (2,1,1)")

# SARIMA(1,0,3)x(2,1,1)
SARIMA(1,0,3,"SARIMA(1,0,3) X (2,1,1)")

# SARIMA(2,0,1)x(2,1,1)
SARIMA(2,0,1,"SARIMA(2,0,1) X (2,1,1)")

# SARIMA(2,0,2)x(2,1,1)
SARIMA(2,0,2,"SARIMA(2,0,2) X (2,1,1)")

# SARIMA(2,0,3)x(2,1,1)
```

```
SARIMA(2,0,3,"SARIMA(2,0,3) X (2,1,1)")

# SARIMA(3,0,3)x(2,1,1)
SARIMA(3,0,3,"SARIMA(3,0,3) X (2,1,1)")


# ----------------------------------------------------------------------
# Finding Best Model
# ----------------------------------------------------------------------
AIC_SCORE %>% unique() %>%  arrange(AIC)
BIC_SCORE %>% unique() %>%  arrange(BIC)

# ----------------------------------------------------------------------
# SARIMA(3,0,3)x(2,1,1) is best model, now, let's try overfit
# ----------------------------------------------------------------------

# SARIMA(4,0,3)x(2,1,1)
SARIMA(4,0,3,"SARIMA(4,0,3) X (2,1,1)")

# SARIMA(3,0,4)x(2,1,1)
SARIMA(3,0,4,"SARIMA(3,0,4) X (2,1,1)")

# ----------------------------------------------------------------------
# SARIMA(3,0,4)x(2,1,1) is best model, now, let's try overfit
# ----------------------------------------------------------------------

# SARIMA(3,0,5)x(2,1,1)
SARIMA(3,0,5,"SARIMA(3,0,5) X (2,1,1)")

# SARIMA(4,0,4)x(2,1,1)
SARIMA(4,0,4,"SARIMA(4,0,4) X (2,1,1)")


# ----------------------------------------------------------------------
# SARIMA(3,0,5)x(2,1,1) is best model, now, let's try overfit
# ----------------------------------------------------------------------

# SARIMA(3,0,6)x(2,1,1)
SARIMA(3,0,6,"SARIMA(3,0,6) X (2,1,1)")

# SARIMA(4,0,5)x(2,1,1)
SARIMA(4,0,5,"SARIMA(4,0,5) X (2,1,1)")


# ----------------------------------------------------------------------
# SARIMA(4,0,5)x(2,1,1) is best model, now, let's try overfit
# ----------------------------------------------------------------------

# SARIMA(4,0,6)x(2,1,1)
SARIMA(4,0,6,"SARIMA(4,0,6) X (2,1,1)")

# SARIMA(5,0,5)x(2,1,1)
SARIMA(5,0,5,"SARIMA(5,0,5) X (2,1,1)")

# ----------------------------------------------------------------------
```

```
# Forecasting
# -----------------------------------------------------------------------
forecasting = Arima(wasteTS,order=c(4,0,5),
                  seasonal=list(order=c(2,1,1), period=12),
                  lambda = -0.1,
                  method = "ML")
prediction = forecast(forecasting, h = 10)

win.graph(width = 10,
          height = 10,
          pointsize = 15)
plot(prediction)

prediction
```