Currency rates in the Time of Coronavirus

AUSTRALIAN DOLLAR EXCHANGE RATE IN US DOLLARS
DECEMBER 2019 – MAY 2020

Thomas Robertson

Abstract

The exchange rate of the Australian dollar (AUD), in United States dollars (USD), over the period from 1 December 2019 to 28 May 2020, was considered. During this time, the response to COVID-19 had major economic impacts. Using time series techniques, a model was created based on this entire time series. However, the dramatic fall and swift recovery of the AUD in terms of USD in March and early April prompted the creation of another model to be used for predicting the exchange rate, using only the time series since 17 April. This model was used to predict the next 10 instances of the exchange rate time series.

Table of Contents

Introduction

The Australian dollar (AUD) changes in value over time. It is often measured in terms of how many United States dollars (USD) it will purchase. To be sure, the United States currency changes in value also, so this measurement is a relative measurement comparing AUD to USD. This report looks at the value of AUD in USD over the past six months, from the start of December 2019 to 28 May 2020, using data from exchange-rates.org. At the start of this period, the AUD was trading at 0.676 USD. During this time, COVID-19 and the various responses to it by the civil authorities, in both Australia and the United States as well as other nations, affected economies in both nations. This economic impact is likely to affect exchange rates. At one stage during in March 2020, the Australian dollar dropped dramatically (in USD) before recovering quite quickly. The AUD dropped 7 cents from 64.9 US cents to 57.9 US cents in one week between 11 March and 18 March, but by 14 April it was trading at 64.4 US cents. This is likely to have been connected to the concern over the economic impact of the response to the disease. To be sure, the currency exchange rate is changing all the time but in the context of the whole six-month period, the drop in particular was quite sudden. This presents challenges to modelling the exchange rate in this period, but nevertheless this project sets out to do so.

Methodology

The R programming language was used throughout this project. The data was downloaded as a csv file from exchange-rates.org and read into an R data frame. This data contained the exchange rate for each day in the period except all Saturdays. This would be because foreign exchange markets do not see – relatively – much activity over the weekend. The Sunday exchange rate is presumably the rate at which markets opened on Monday. A time series of the exchange rate was created and plotted. Two approaches to modelling the data were taken.

First approach

The first approach was to create a model based on the whole time series. As it was not stationary, the series was differenced. The differenced series was plotted. Checks for normality and stationarity were performed. Note that the 5% significance level was chosen to be used for statistical tests throughout this project. To inform the selection of candidate models for the time series, the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the differenced series were plotted. Then the extended autocorrelation function (EACF) was checked. Finally, a Bayesian Information Criterion (BIC) table was created to further inform the decision as to which candidate Autoregressive Integrated Moving Average (ARIMA) models would be suitable.

Parameters were estimated for each candidate model. The residuals of each model were analysed. To further check each model, a generalisation of each was created and the parameters checked to see if there were any significant differences between the candidate model and its generalisation.

Finally, a model was chosen. Using this model, predictions were made for the next 10 days.

Second Approach

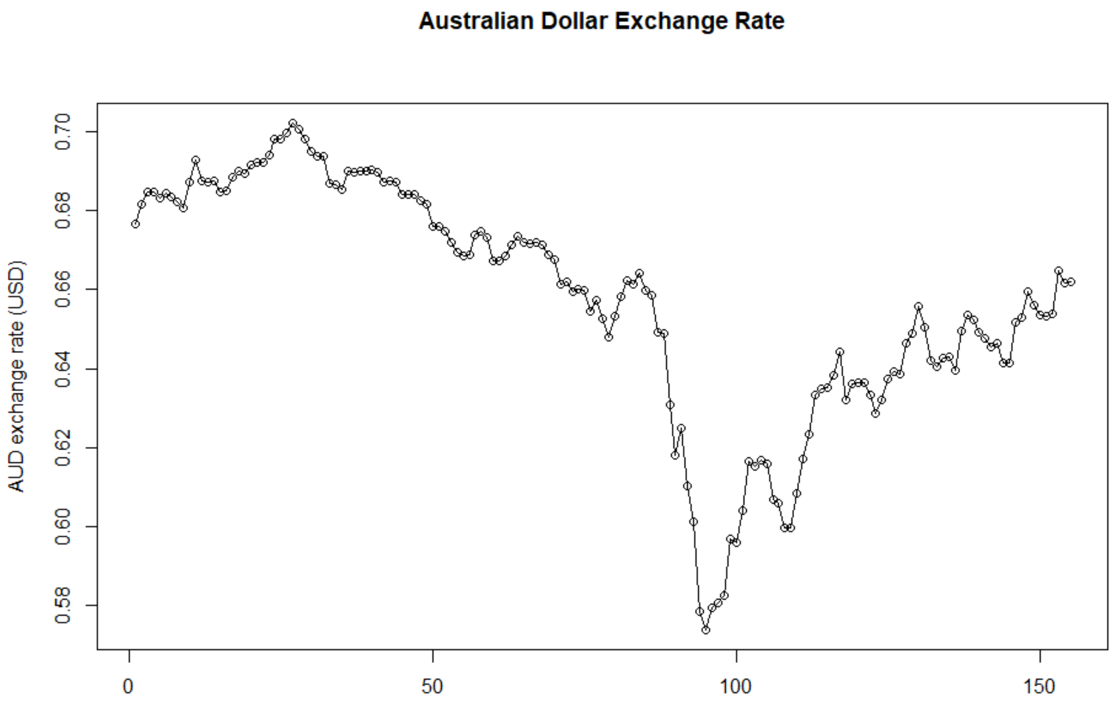
The second approach was to consider the time series after the sudden fall and recovery. This smaller time series was analysed and modelled in a similar method to the larger series, with two exceptions: firstly, that as it was found that this smaller time series was normally distributed, the Akaike

Information Criterion (AIC) was also used, and secondly, it was found that there was an appreciable seasonal aspect to the time series, a seasonal component was introduced.

Results

First approach

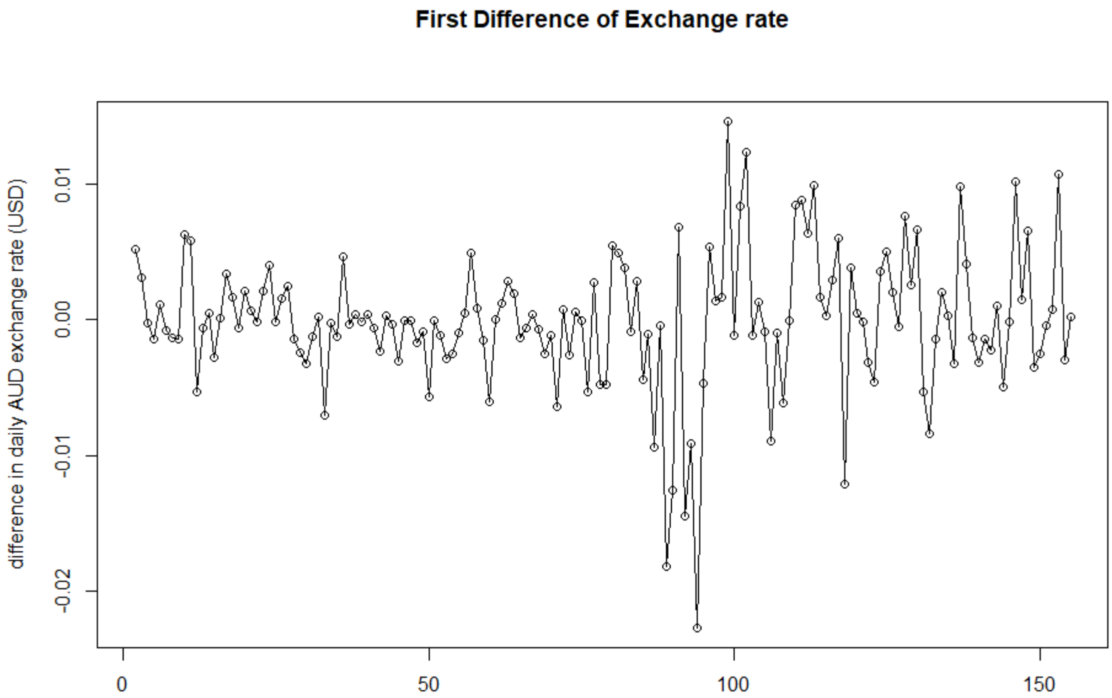
The time series of the exchange rate between 1 December 2019 and 28 May 2020 was plotted. The plot showed that it was clearly not stationary, and a Dickey-Fuller unit root test confirmed the same. The results are shown below.



Data from EXCHANGE-RATES.ORG

```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 2
  STATISTIC:
    Dickey-Fuller: -0.3606
  P VALUE:
    0.4993
```

The null hypothesis that the series is non-stationary cannot be rejected. The series was differenced. The plot of the differenced time series, on the next page, looks more stationary, and indeed the Dickey-Fuller unit root test rejected the non-stationary hypothesis.




```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 10
  STATISTIC:
    Dickey-Fuller: -3.296
  P VALUE:
    0.01
```

The p-value was in fact less than 0.01 as per a warning message given by R.

The time series of the exchange rate and its first difference were both found to not be normally distributed.

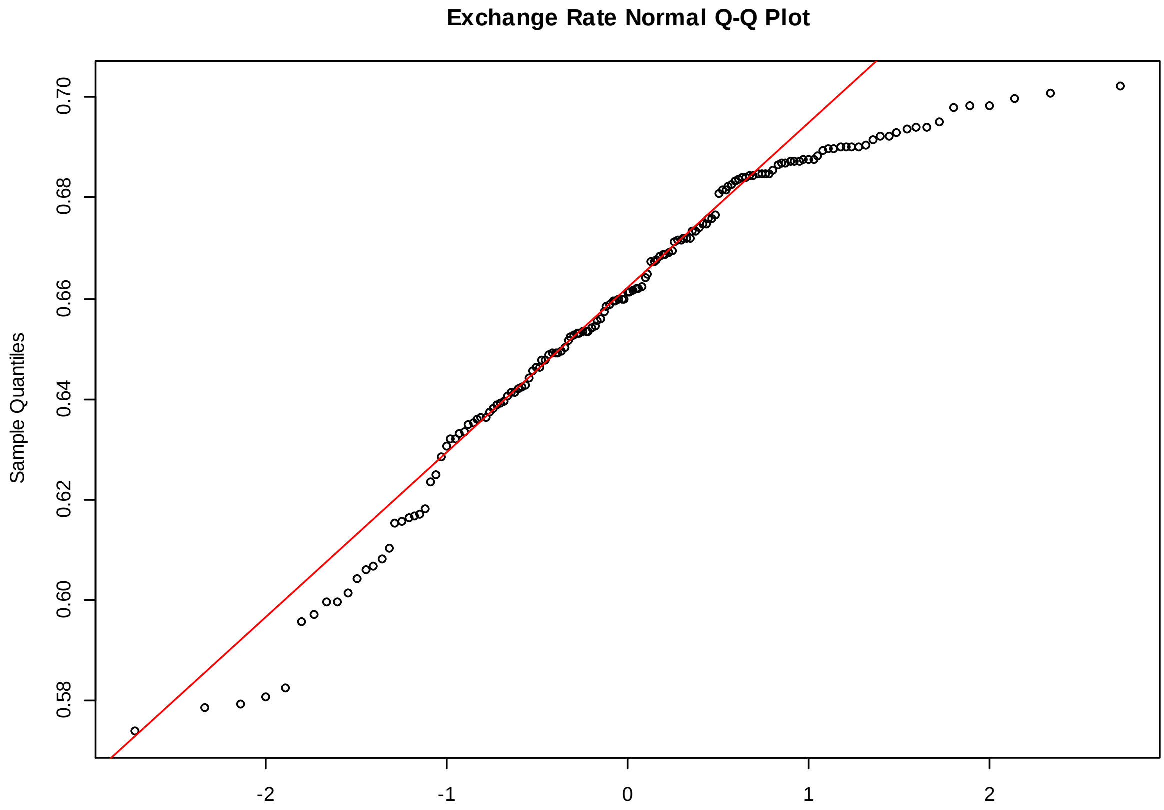
The Shapiro-Wilk test, when applied to the original time series, yielded the following results:

```
Shapiro-Wilk normality test

data:  time_series
W = 0.93918, p-value = 3.279e-06
```

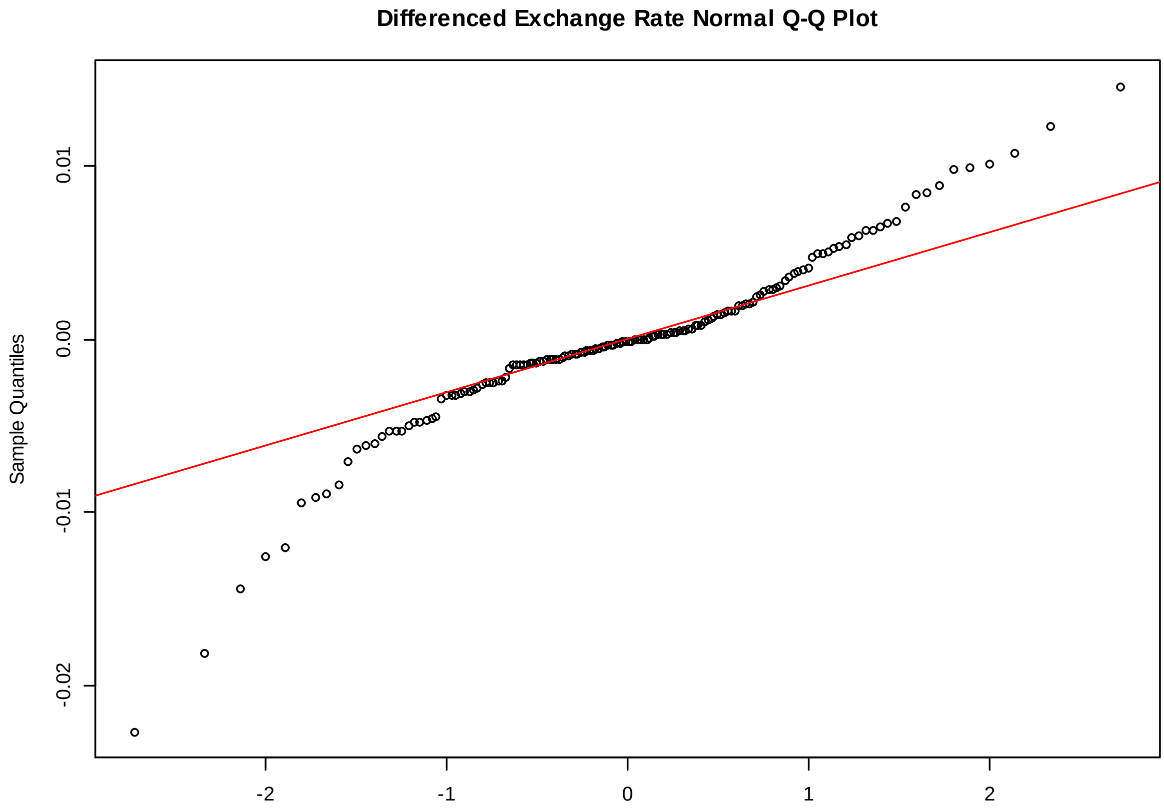
And the same test applied to the first difference gave the following results:

```
Shapiro-Wilk normality test

data:  time_series
W = 0.93263, p-value = 1.147e-06
```

The respective Q-Q plots of the undifferenced and differenced time series are on the next page.


Exchange Rate Normal Q-Q Plot
Sample Quantiles
0.58
0.60
0.62
0.64
0.66
0.68
0.70
-2
-1
0
1
2

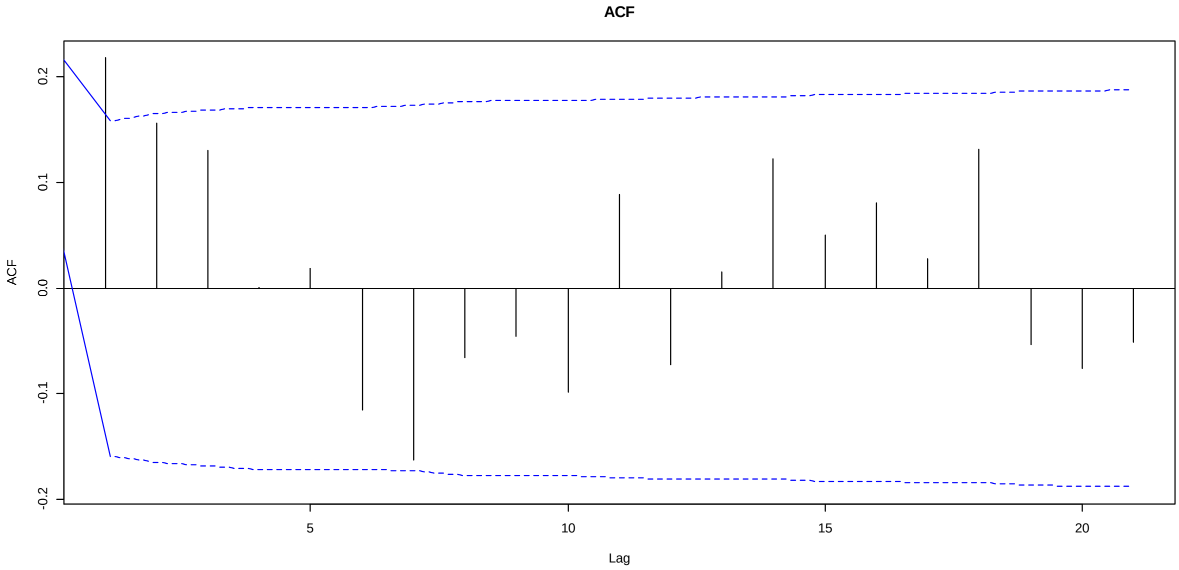
1
2

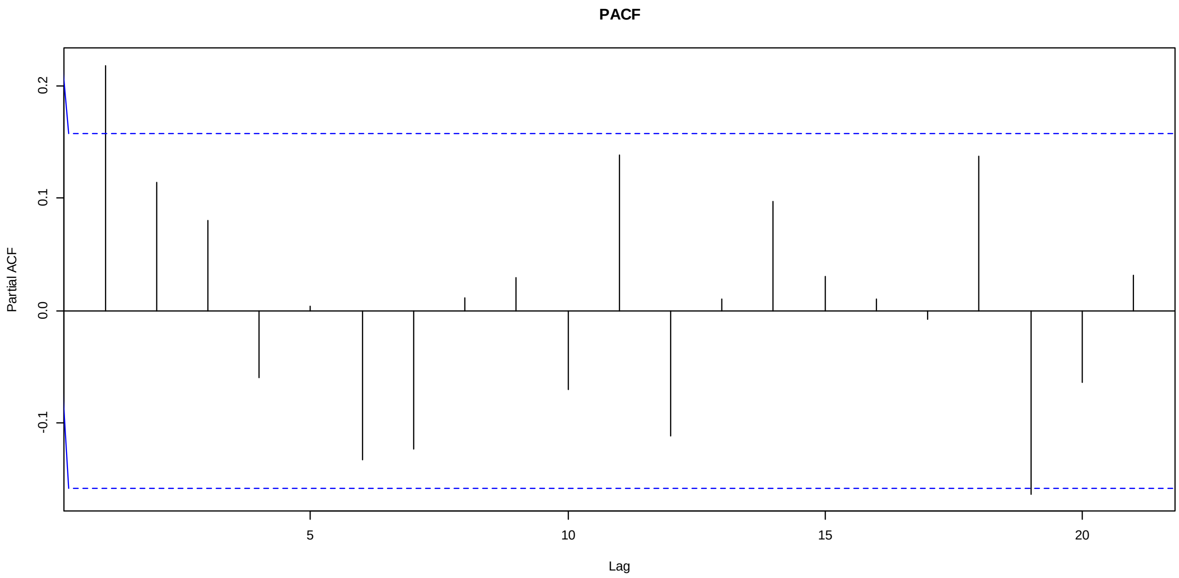

Differenced Exchange Rate Normal Q-Q Plot
Sample Quantiles
-0.02
-0.01
0.00
0.01
-2
-1
0
1
2

2
3

Neither taking the logarithm nor applying the Box-Cox transformation resulted in a normally distributed time series. The divergence from normality presents some difficulty as it means tests such as evaluating the AIC that require normality cannot be employed. Nevertheless, as the time series of the first difference is a stationary series, an attempt to fit an ARIMA model for the undifferenced series was still undertaken.

The ACF and PACF of the first difference are shown below.





Lag 1 can be seen to have significant autocorrelation as well as partial autocorrelation. This suggests an ARMA(1, 1) model may be suitable for the first difference, and consequently an ARIMA(1, 1, 1) model for the undifferenced series.

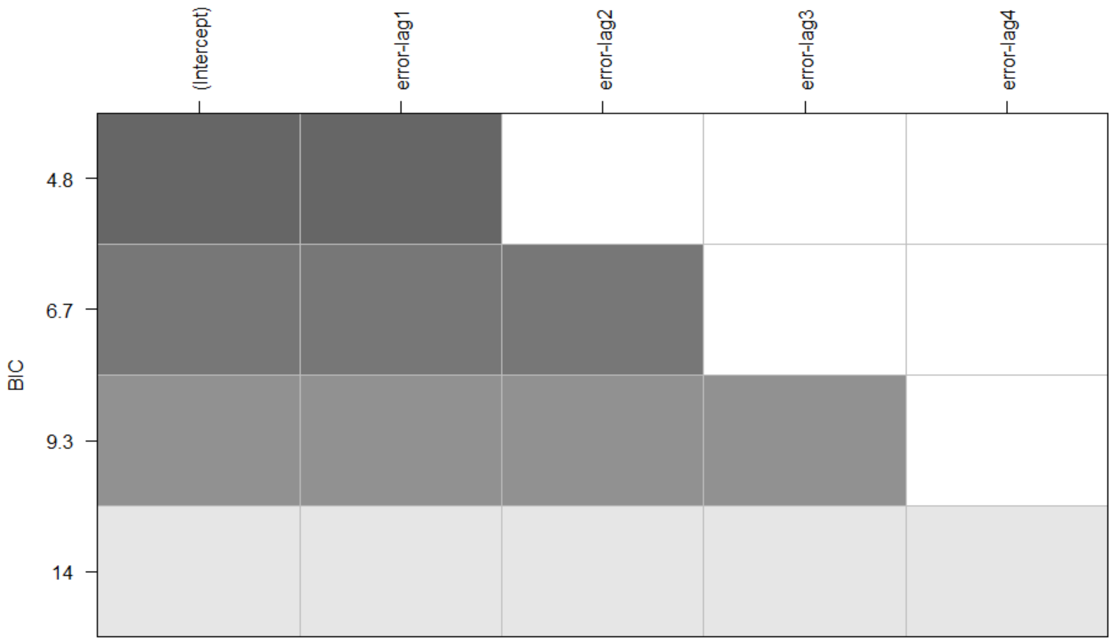
Next the EACF should be considered.

EACF of first difference

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

The EACF suggests that an ARMA(0, 1) may be the correct model of the first difference.

Finally, BIC was also considered.



The BIC table also supports an ARMA(0, 1) model for the first difference and hence an ARIMA(0, 1, 1) model for the undifferenced exchange rate time series.

There are thus two candidate ARIMA models to consider: ARIMA(0, 1, 1) and ARIMA(1, 1, 1).

Fitting parameters yielded the following coefficients:

ARIMA(0, 1, 1)

```
z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ma1 0.181620   0.073148  2.4829  0.01303 *
```

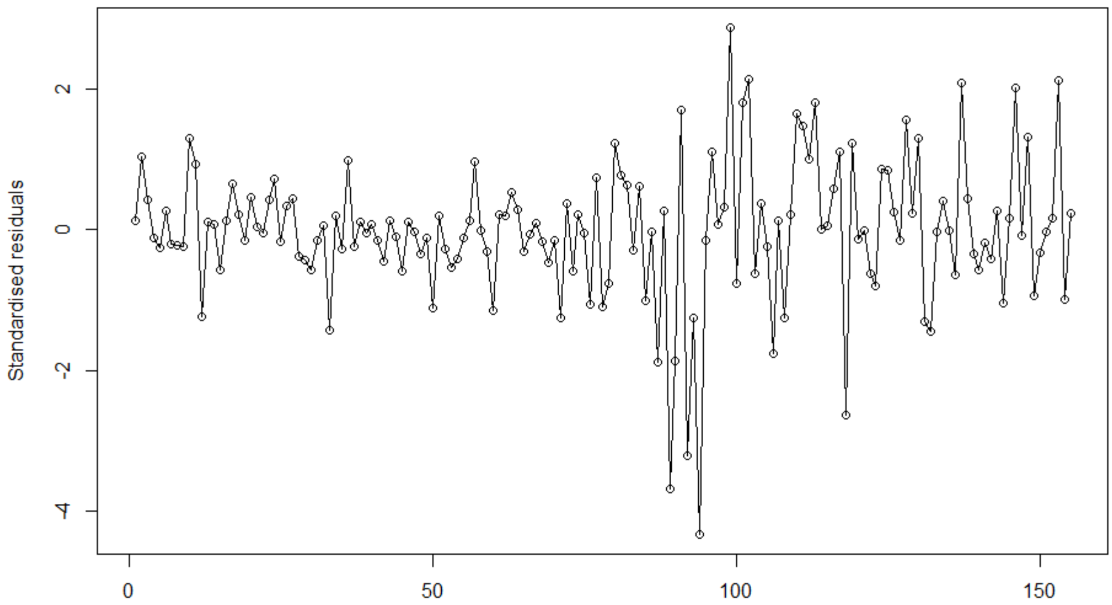
ARIMA(1, 1, 1)

```
z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ar1  0.60390    0.17798  3.3932 0.0006909 ***
```

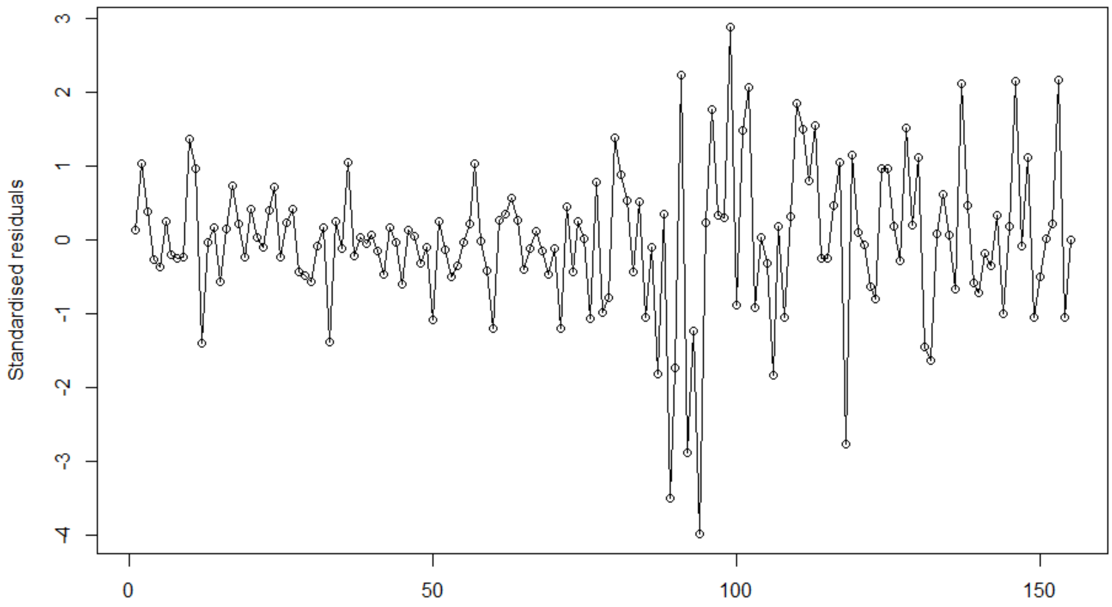
```
ma1 -0.40025    0.19862 -2.0152 0.0438883 *
```

In spite of the differences in models, the standardised residuals looked quite similar.

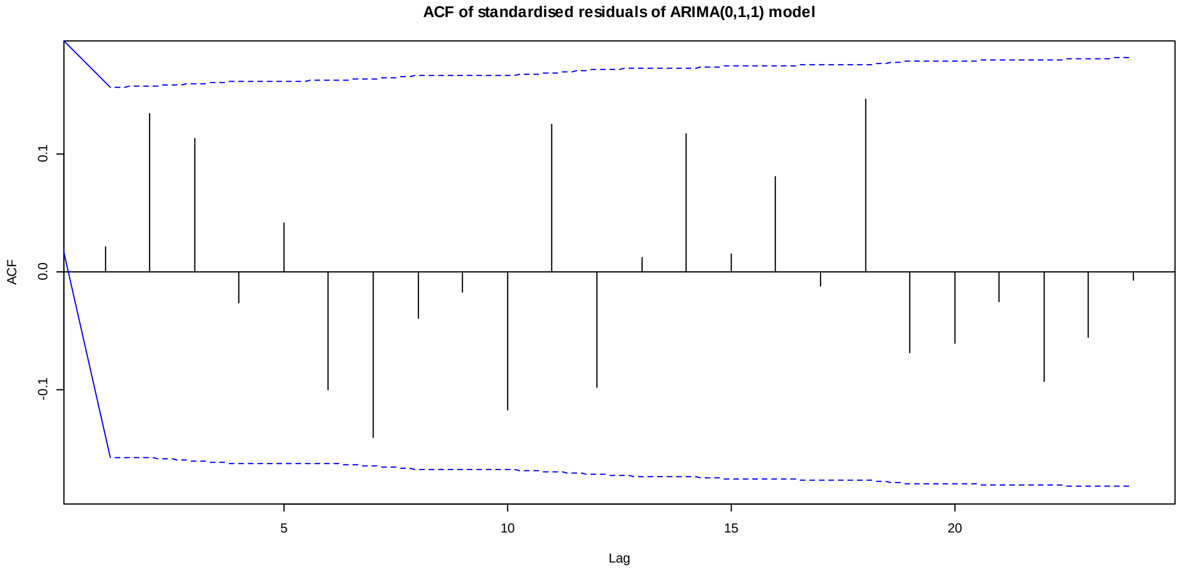
Standardised residuals of ARIMA(0,1,1) model

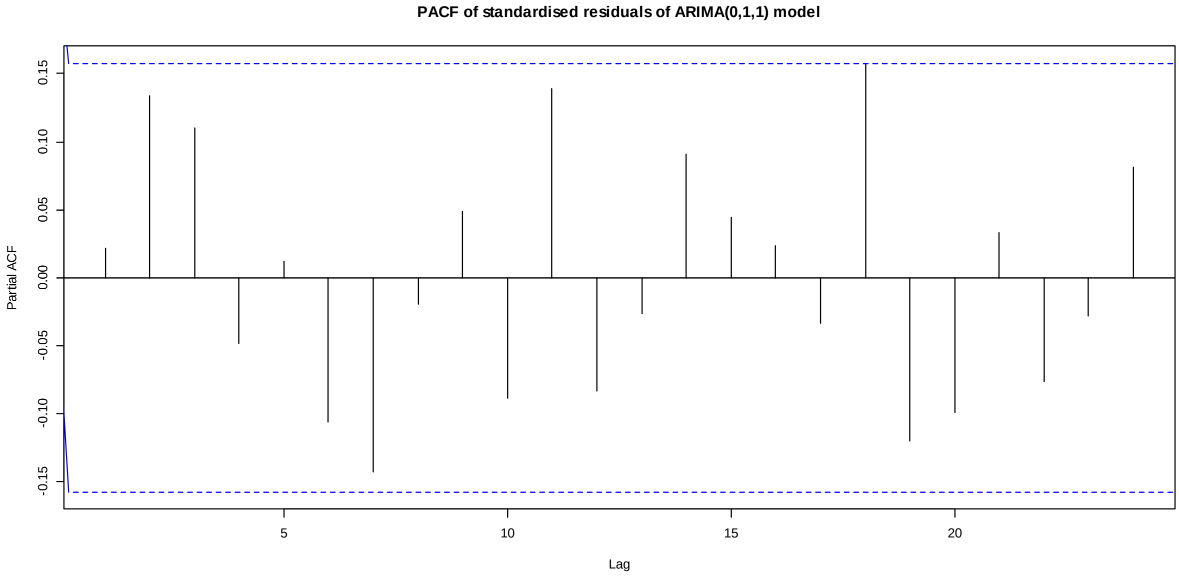


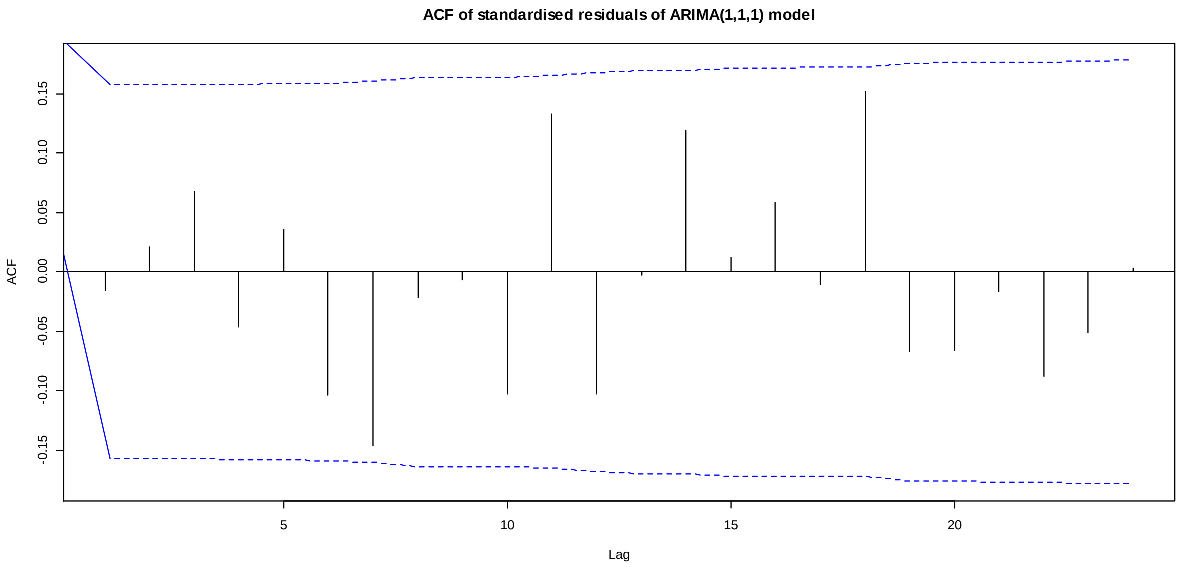
Standardised residuals of ARIMA(1,1,1) model

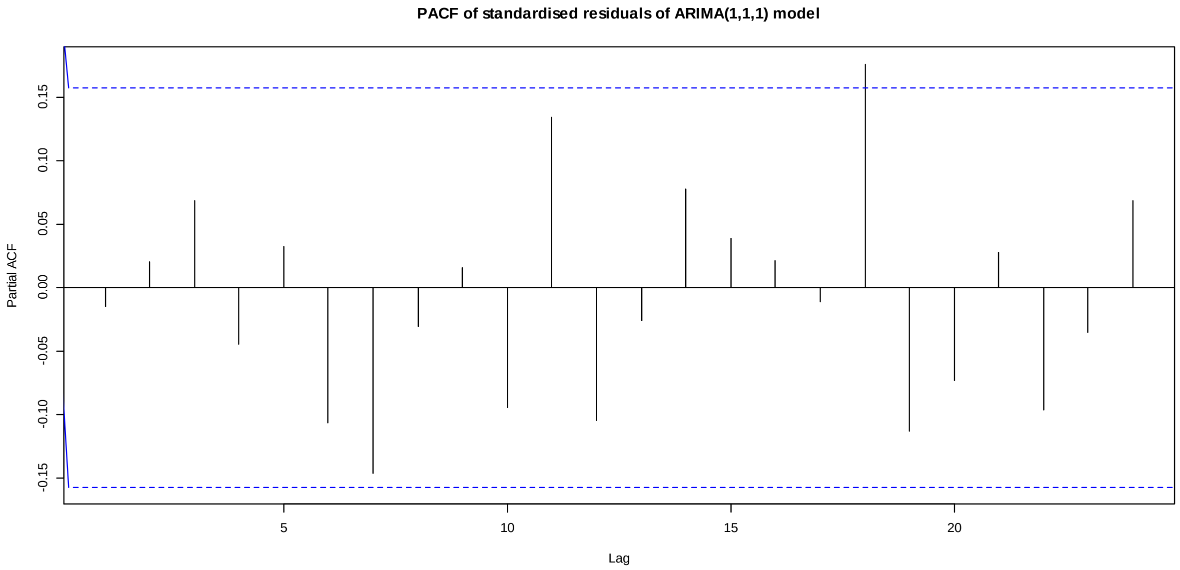


The similarity between the above two plots is interesting given the differences in the coefficients. The volatility around data points 90-100 corresponds to the sudden decline and recovery in the exchange rate in March. The ACF and PACF of the standardised residuals were plotted and are shown on the following pages.


ACF of standardised residuals of ARIMA(0,1,1) model
ACF
Lag


PACF of standardised residuals of ARIMA(0,1,1) model
Partial ACF
Lag





In the residuals of both models lag 18 has high partial autocorrelation. It is touching the dotted line in the case of the ARIMA(0, 1, 1) model and extends beyond it in the case of the ARIMA(1, 1, 1) model. But overall the residuals of the latter model actually look less correlated, with the first five lags having small ACF and PACF.

Ljung-Box tests also help to understand if the residuals are significantly correlated. The results for each are below.

<u>ARIMA(0, 1, 1)</u>

```
        Box-Ljung test

data:  rstandard(m1)
X-squared = 24.566, df = 18, p-value = 0.1374
```

<u>ARIMA(1, 1, 1)</u>

```
Box-Ljung test

data:  rstandard(m2)
X-squared = 20.523, df = 18, p-value = 0.3042
```

Generalising or overfitting a model can also provide support for that model. If the coefficients in the generalised model do not differ significantly from those in the original model, and the additional parameters are not significant, then it shows that these additional parameters likely lead to overfitting and should not be included.

Both models were extended by adding an extra moving average (MA) term. The parameter estimates of the candidate models and the overfitted extensions thereof, with standard errors, z value and p-values are listed below.

ARIMA(0, 1, 1)

```
z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ma1 0.181620   0.073148  2.4829  0.01303 *
```

ARIMA(0, 1, 2) – extension of ARIMA(0, 1, 1)

```
z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ma1 0.173383   0.080254  2.1604  0.03074 *
ma2 0.123160   0.082514  1.4926  0.13554
```

ARIMA(1, 1, 1)

```
z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ar1  0.60390    0.17798  3.3932 0.0006909 ***
```

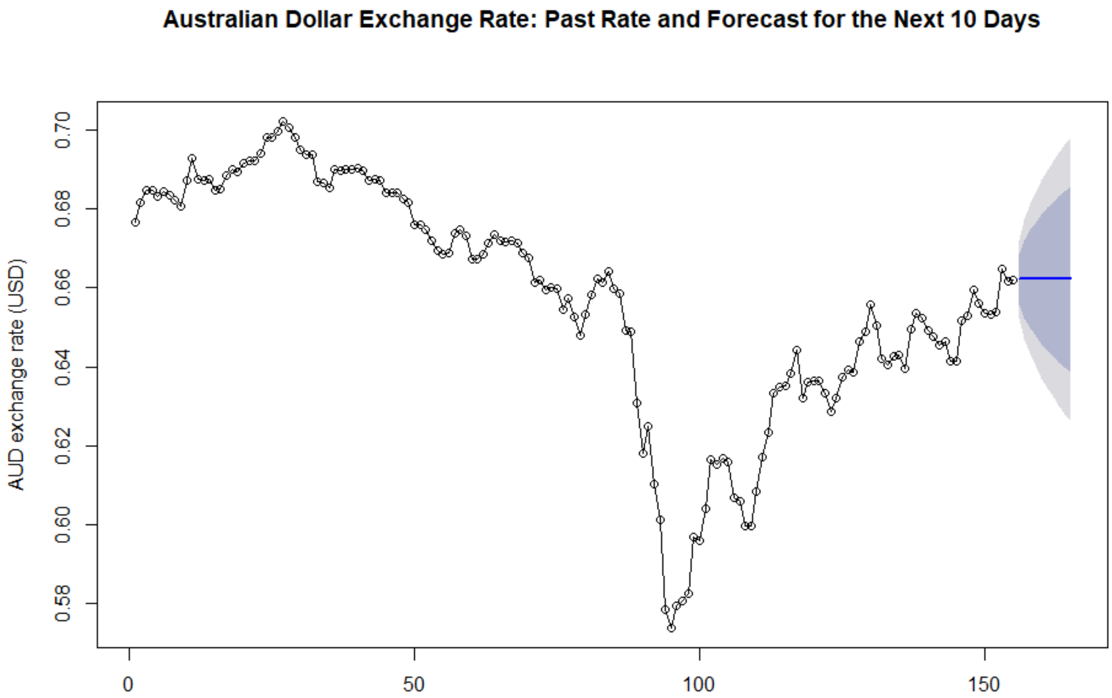
ARIMA(1, 1, 2) – extension of ARIMA(1, 1, 1)

```
z test of coefficients:

     Estimate Std. Error z value Pr(>|z|)
ar1  0.531348   0.235109  2.2600  0.02382 *
ma1 -0.349693   0.240811 -1.4521  0.14646
ma2  0.063888   0.096867  0.6595  0.50955
```

In both cases the additional MA term proved insignificantly different from zero. Also the difference between the common coefficients of the ARIMA(0, 1, 1) and the extension of it, ARIMA(0, 1, 2) as well as the difference between the common coefficients of the ARIMA(1, 1, 1) and the extension of it, ARIMA(1, 1, 2) were not significant at the 5% significance level, being well within one, let alone two, standard deviations. It was concluded that these generalised models were indeed overfittings of each model, and these overfitted models discarded.

As there is not much separating these two models, it would be appropriate to choose the simpler of the two – the ARIMA(0, 1, 1) model – due to the principle of parsimony.

Based on this model then, a prediction for the next 10 instances in the time series was made.

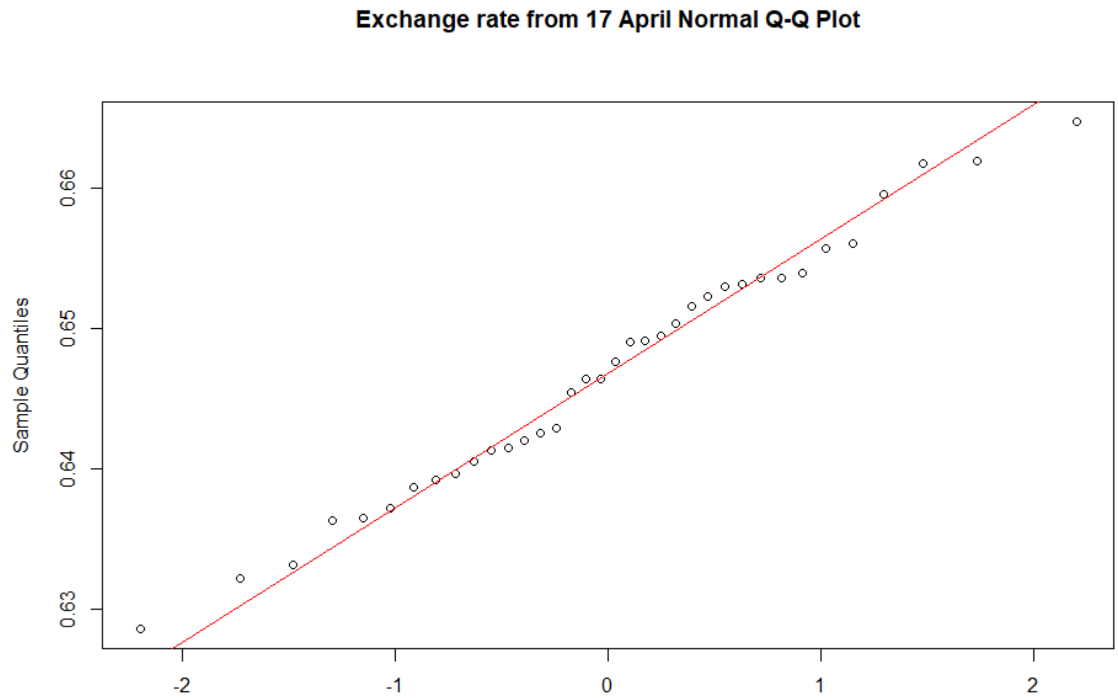


This model predicts the exchange rate to on average stay the same as at present.

Second approach

From about 9 April there appears to be a steady increase in the exchange rate, after the sudden fall and recovery from the middle of March to early April. This later section of the overall should be considered separately for modelling and predictions as it may represent the current situation better. To make sure that the time considered only included that after the sudden decline and recovery during March and early April, the time series was started at 17 April. By then the time series seems to be settling into a gradual increase with small fluctuations. Therefore it may be better for modelling the present situation. The plot of this smaller time series is shown on the next page.

Unlike the larger December to May time series, this shorter time series, with 36 data points rather than 155, is in fact normally distributed. A Q-Q plot and the results of the Shapiro-Wilk test are shown below.




```
        Shapiro-Wilk normality test

data:   time_series
```

```
W = 0.98438, p-value = 0.8803
```

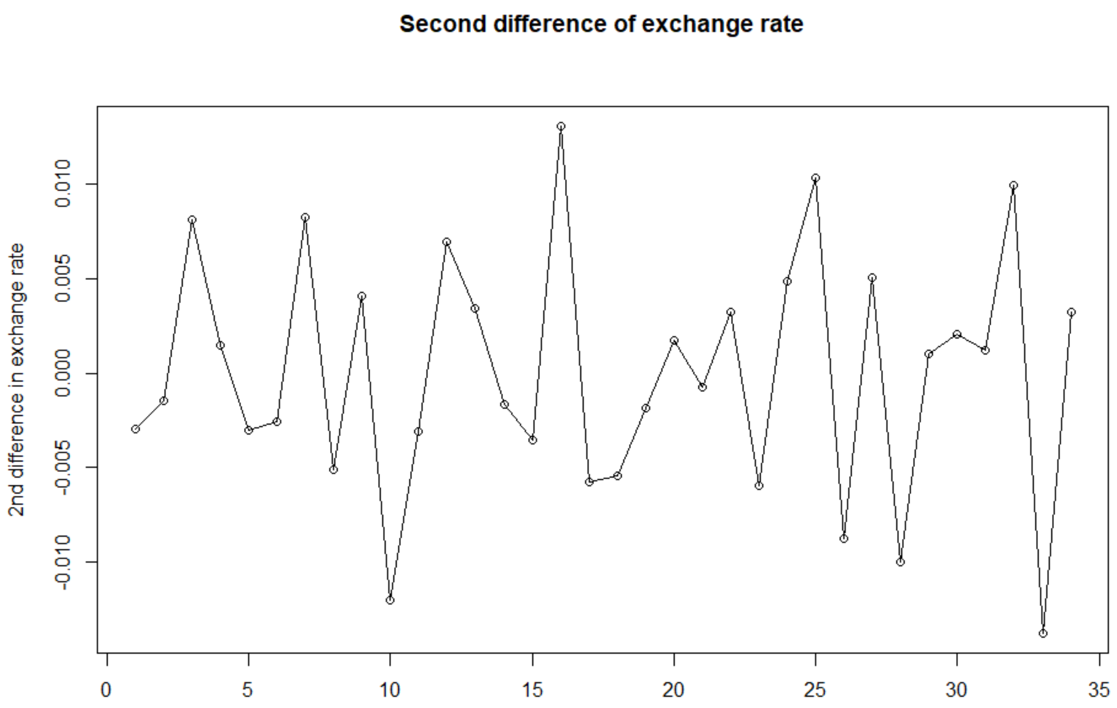
This time series is clearly not stationary and unsurprisingly the Dicky-Fuller unit root test did not reject the null hypothesis of non-stationarity.

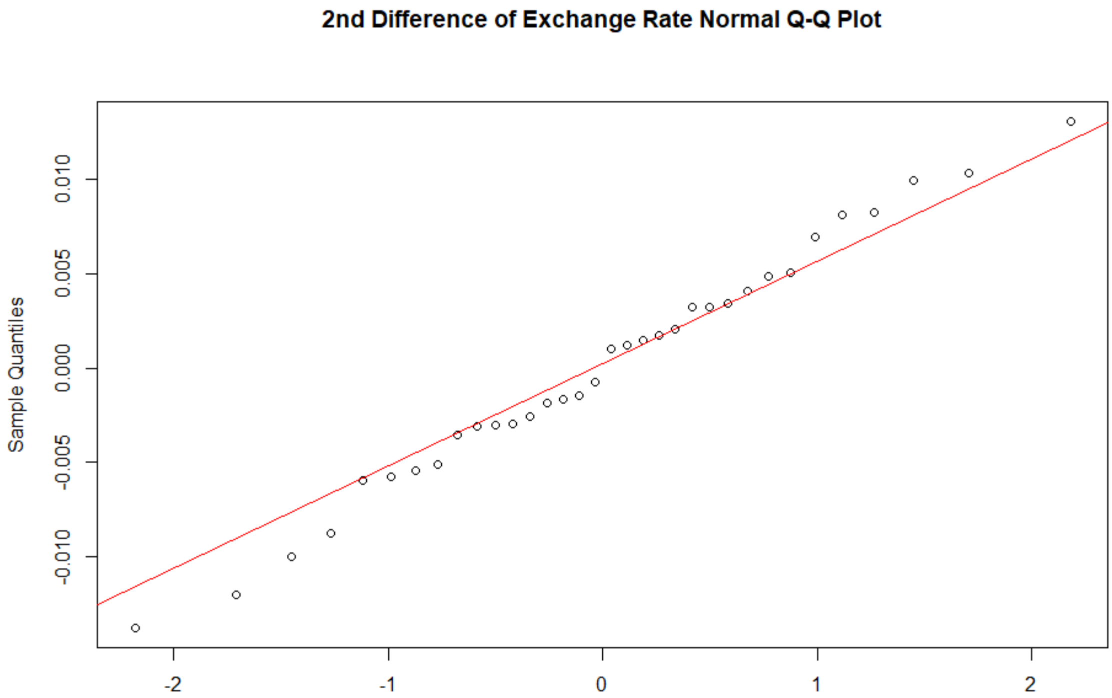
```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 0
  STATISTIC:
    Dickey-Fuller: 0.9124
  P VALUE:
    0.8989
```

The first difference was taken but the Dicky-Fuller unit root test did not reject the null hypothesis that this differenced time series was stationary either, with the p-value greater than 0.05 as seen in the results below.

```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 8
  STATISTIC:
    Dickey-Fuller: -1.5901
  P VALUE:
    0.1043
```

Consequently, the second difference was taken. Its plot and results of the tests for its normality are as follows.



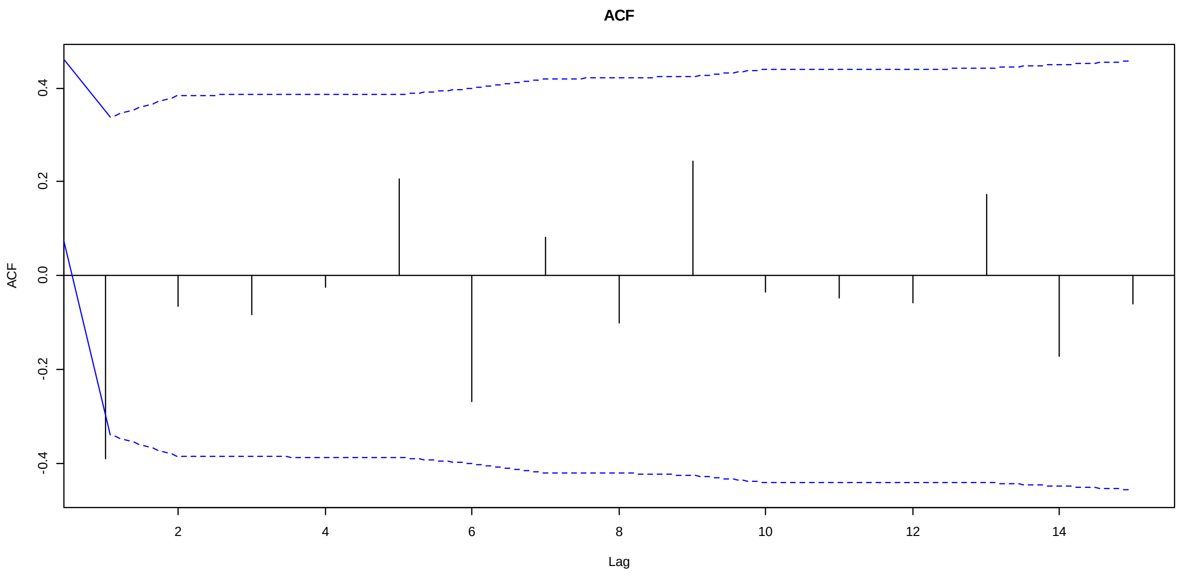


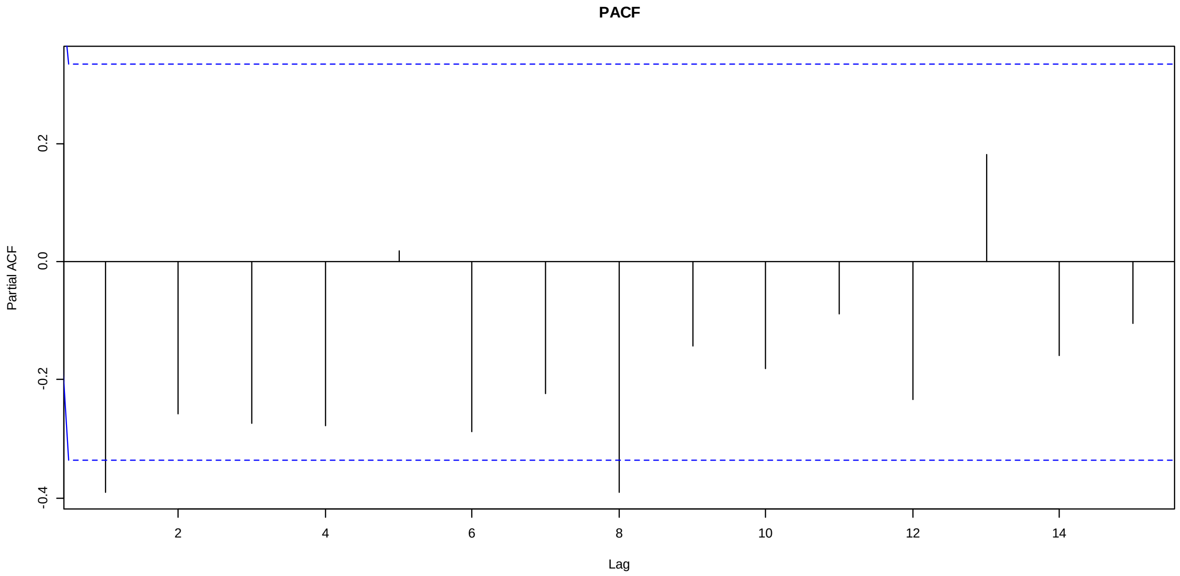

```
        Shapiro-Wilk normality test

data:  time_series
W = 0.98908, p-value = 0.9777
```

The Shapiro-Wilk test fails to reject the assumption of normality. Next the second difference was tested for stationarity. The Dicky-Fuller unit root test reject the assumption of non-stationarity, as shown below. Note that a warning message that "p-value smaller than printed p-value" was displayed in R, indicating that the p-value was even smaller than 0.01.

```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 4
  STATISTIC:
    Dickey-Fuller: -3.2017
  P VALUE:
    0.01
```

A stationary series being obtained, the next step was to select an appropriate model, identifying which autoregressive (AR) and moving average (MA) terms to include. To this end the ACF and PACF were plotted.



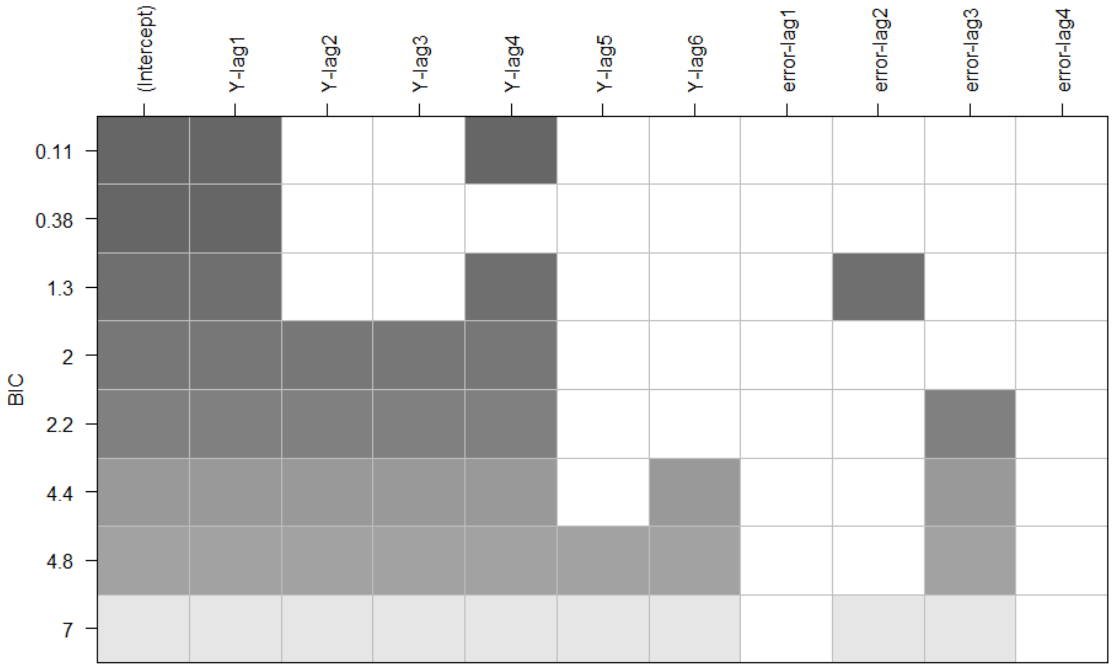


Lag 1 appears to have significant autocorrelation and partial autocorrelation. Hence an ARIMA(1,2,1) model may be suitable for the undifferenced time series.

Next EACF was considered.

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

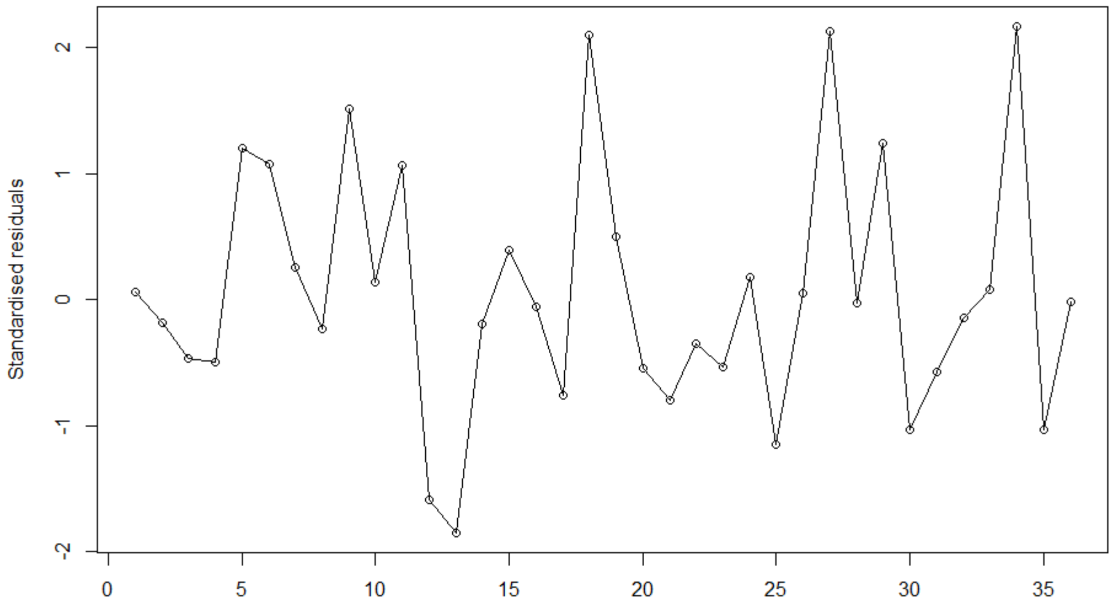
Based on the EACF, an ARIMA(0,2,1) model may be suitable. A BIC table was also created.



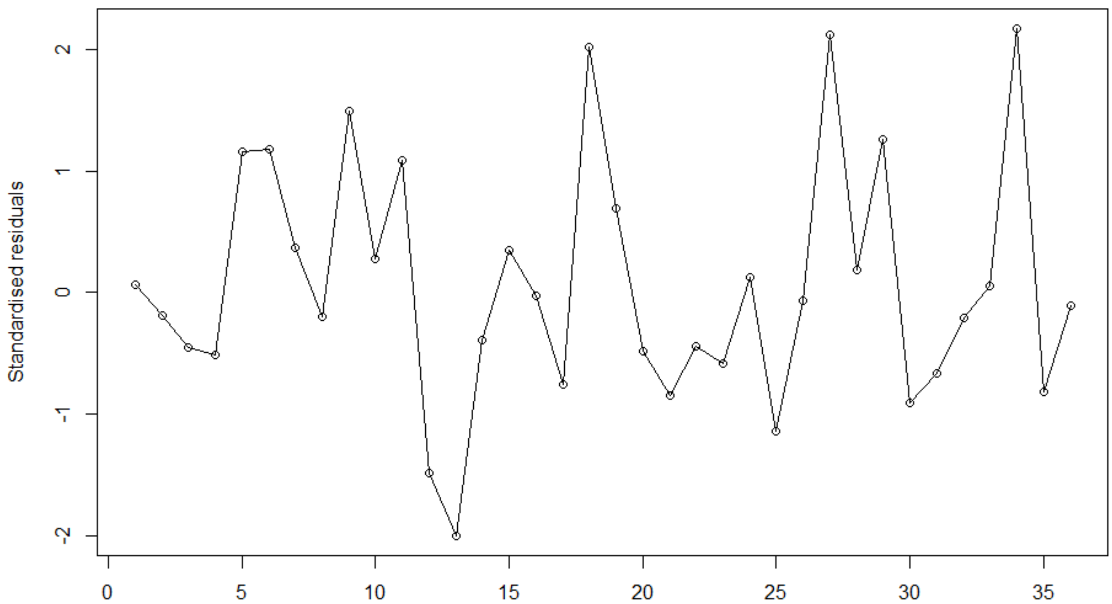
Based on the BIC table the ARIMA(4,2,0) or ARIMA(1,2,0) models were added to the set of candidate models.

Next parameters were estimated for each of these models. The residuals of each model were then analysed including the checking for normality as well as plotting ACF and PACF. The normality of the residuals was checked with the Shapiro-Wilk test and the null hypothesis of normality was not rejected for the residuals of any model. Plots of the standardised residuals, followed by the ACF and PACF of each model, are shown on the following pages.

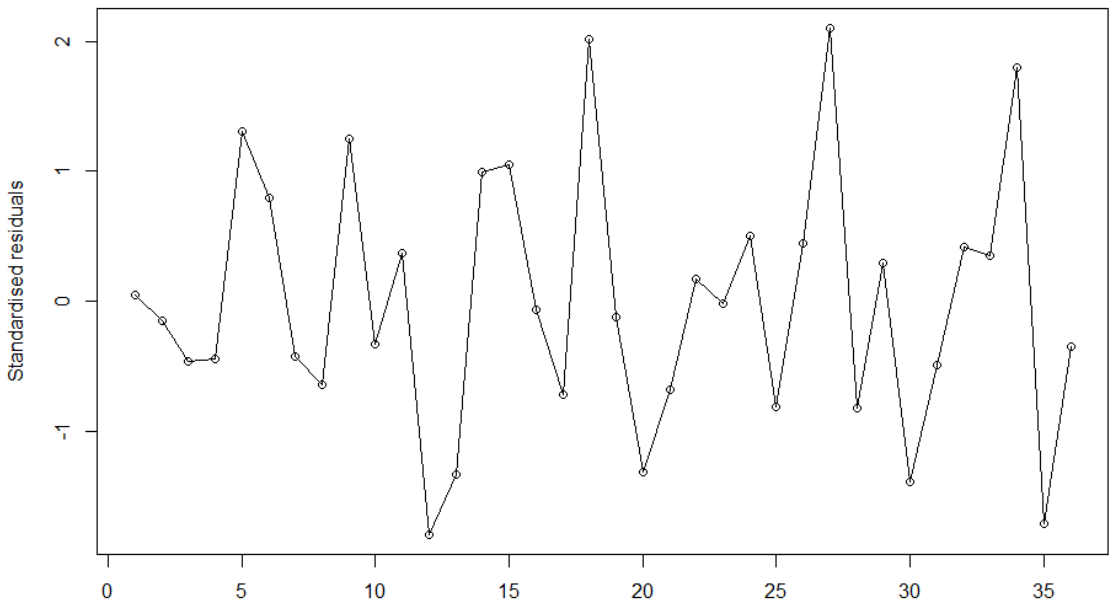
Standardised residuals of ARIMA(1,2,1) model



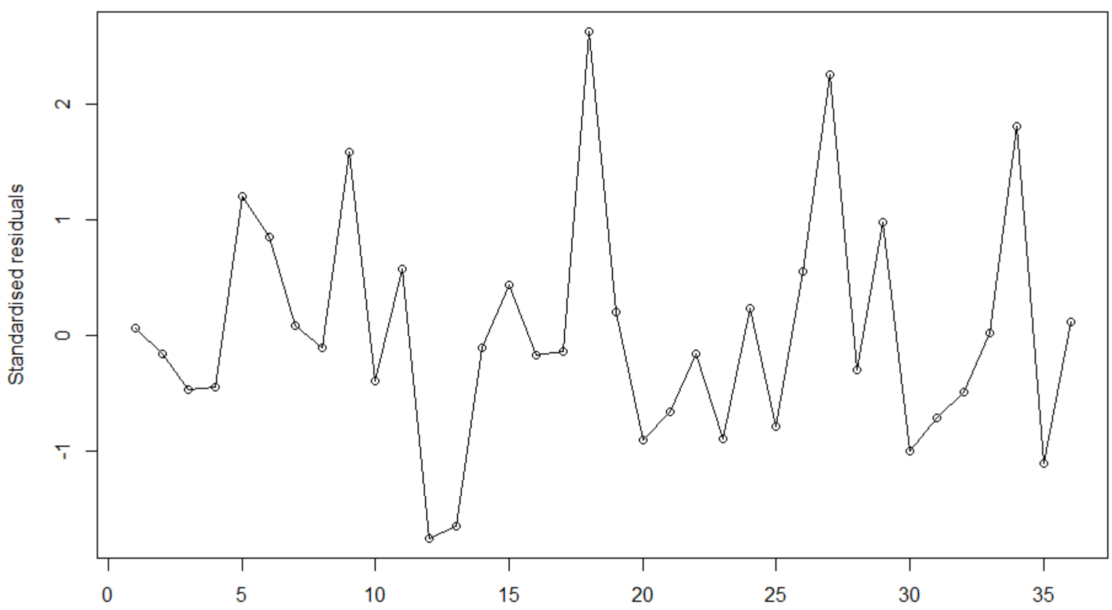
Standardised residuals of ARIMA(0,2,1) model



Standardised residuals of ARIMA(1,2,0) model

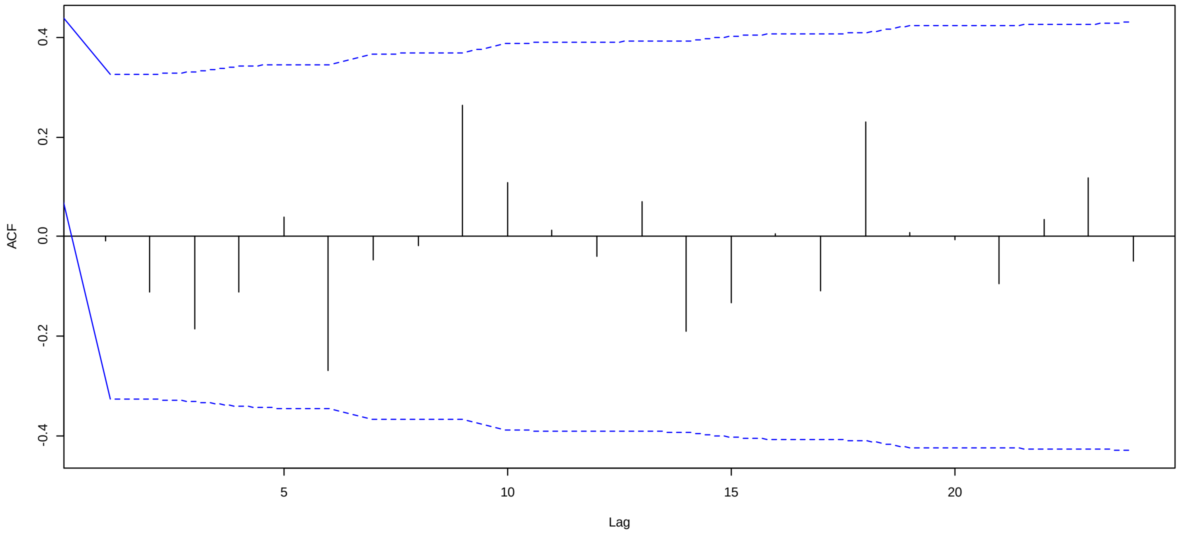


Standardised residuals of ARIMA(4,2,0) model

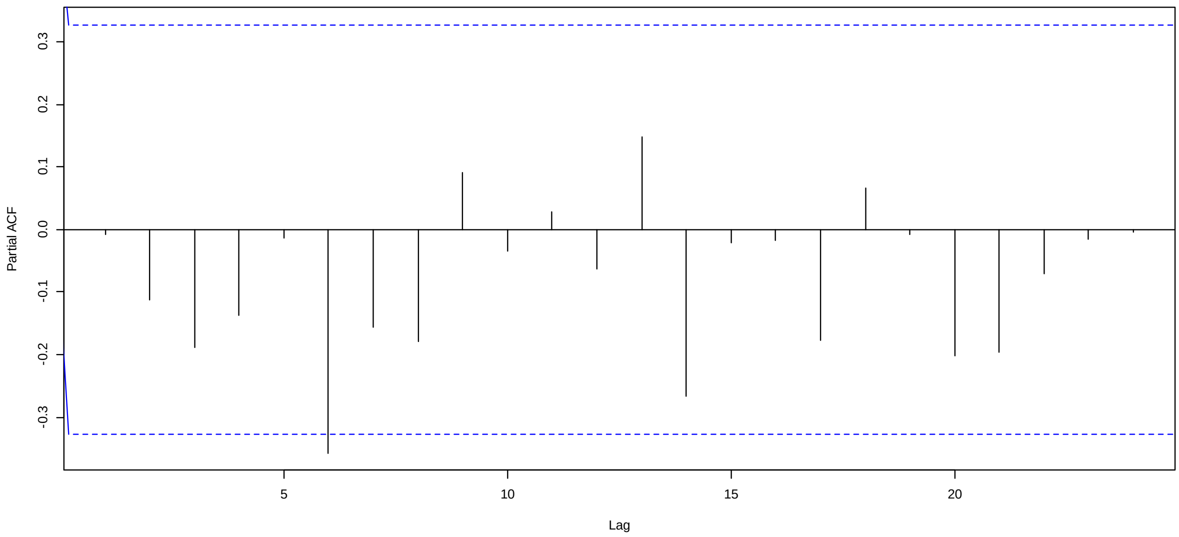


.

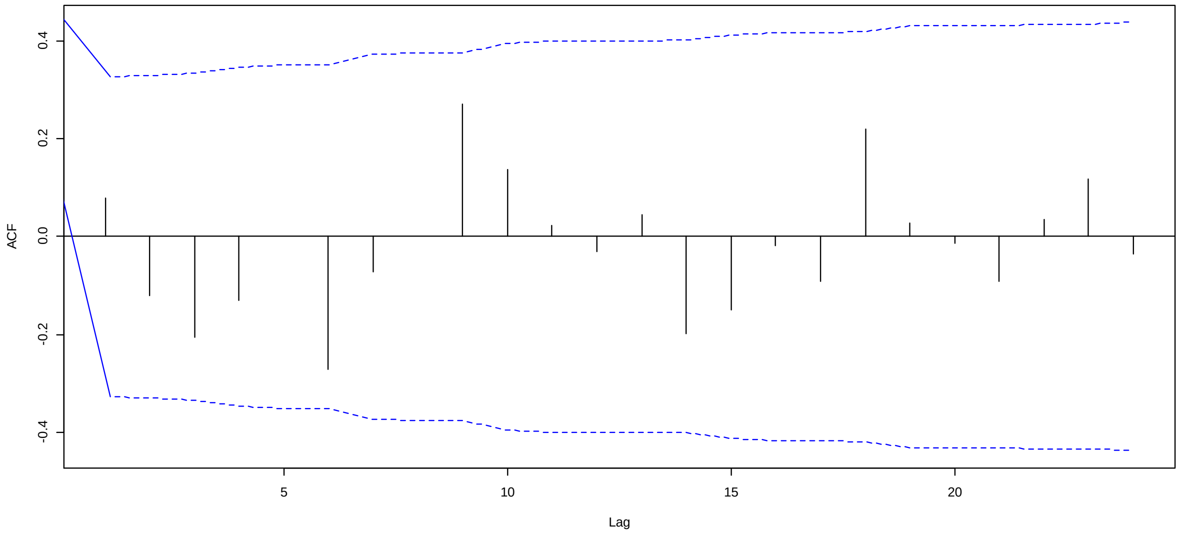
ACF of standardised residuals of ARIMA(1,2,1) model



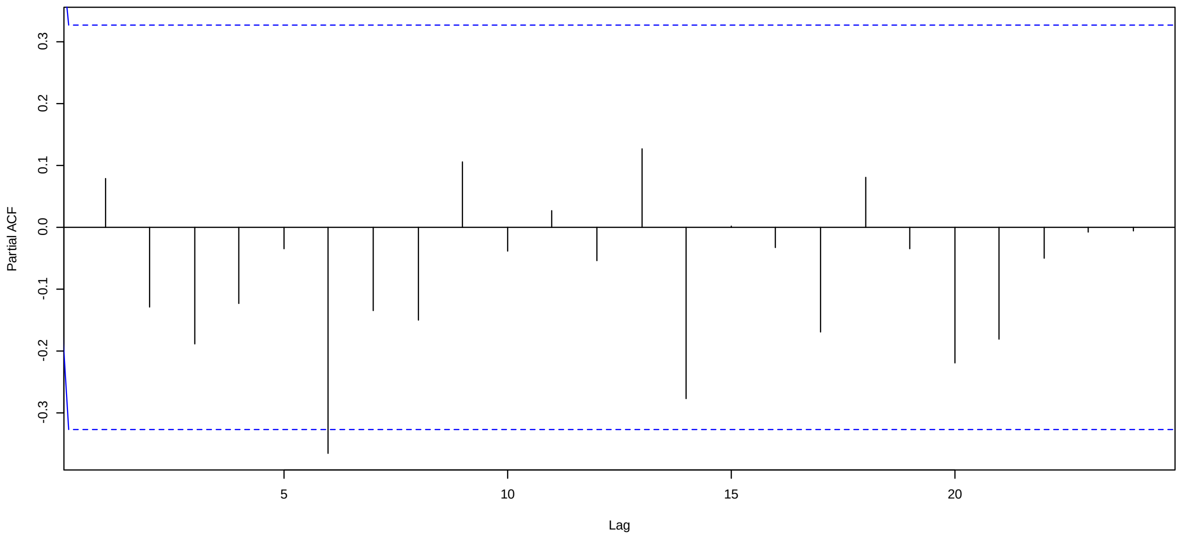
PACF of standardised residuals of ARIMA(1,2,1) model



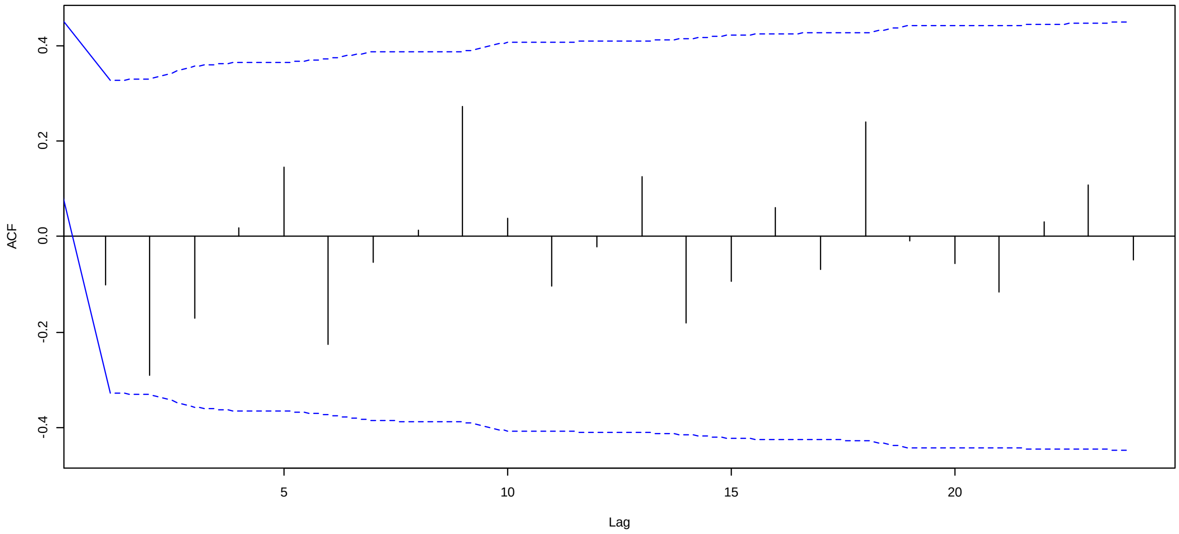
ACF of standardised residuals of ARIMA(0,2,1) model



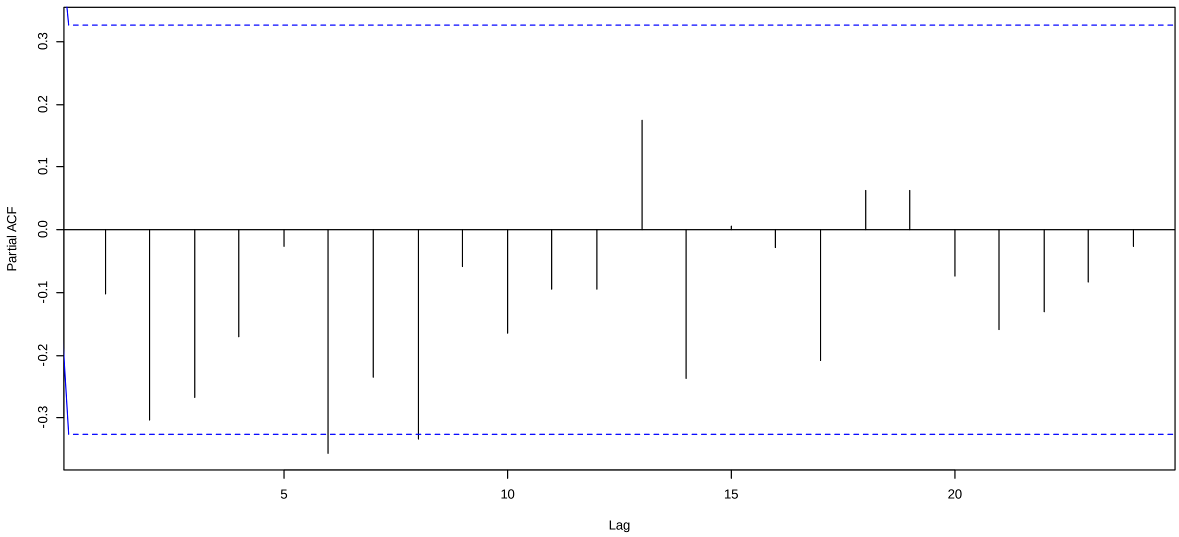
PACF of standardised residuals of ARIMA(0,2,1) model

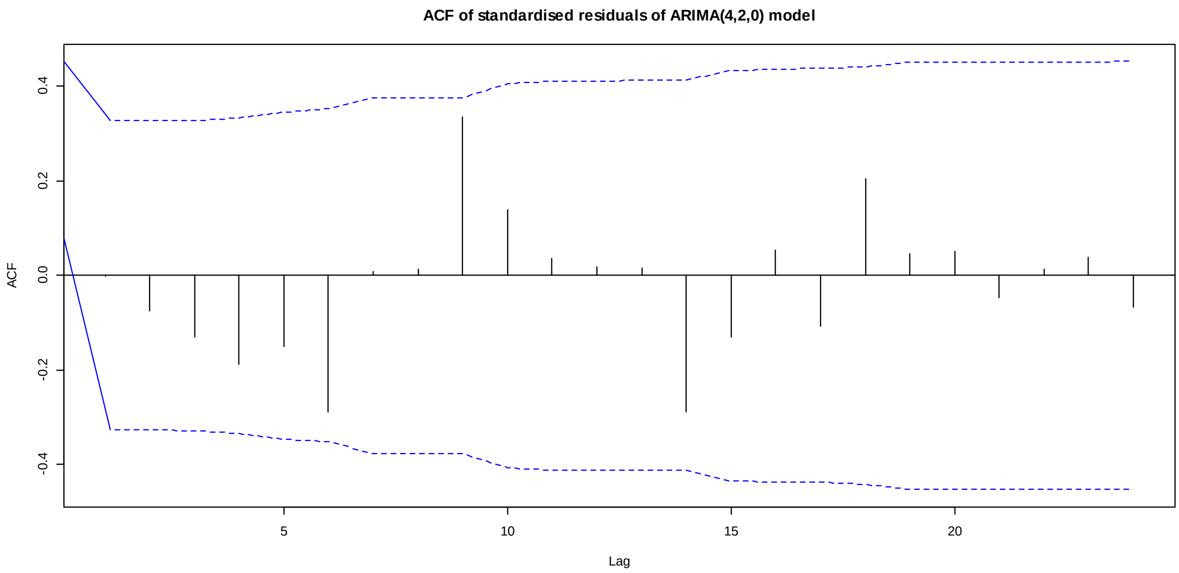


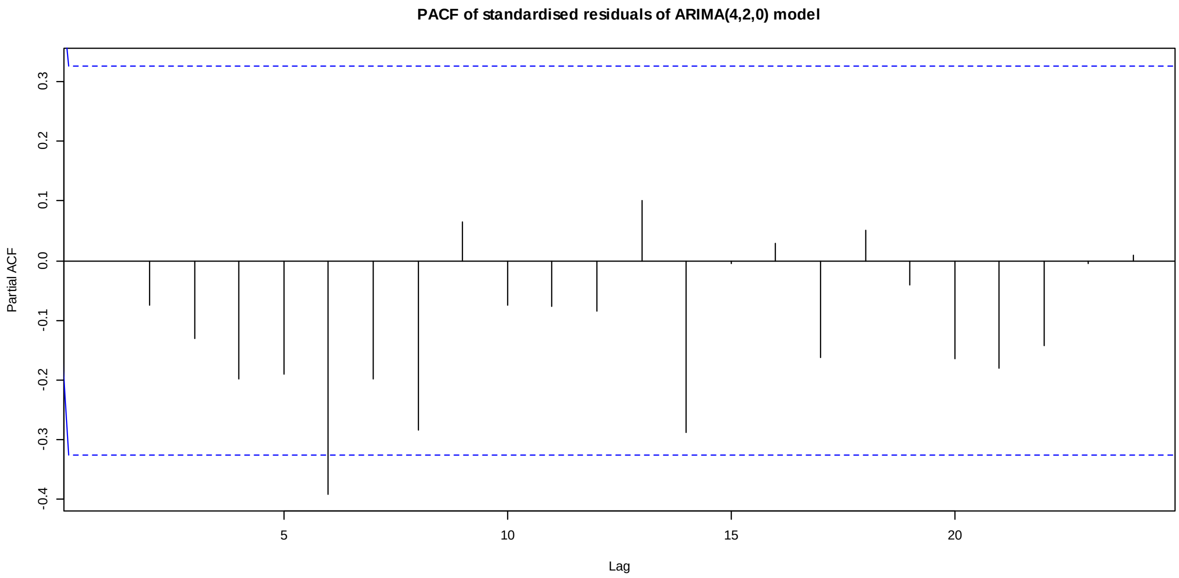
ACF of standardised residuals of ARIMA(1,2,0) model



PACF of standardised residuals of ARIMA(1,2,0) model



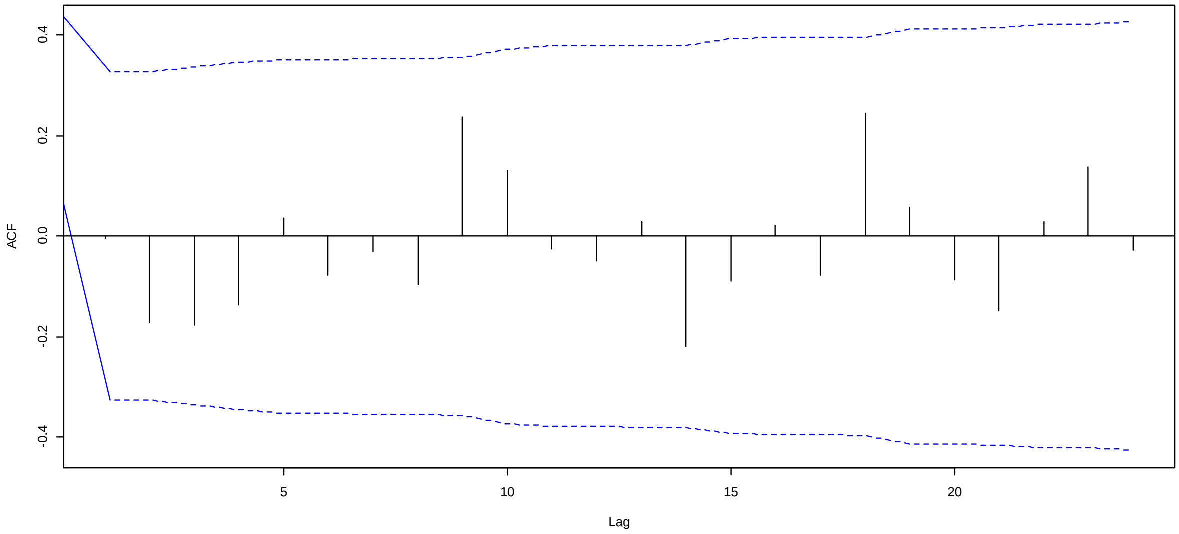




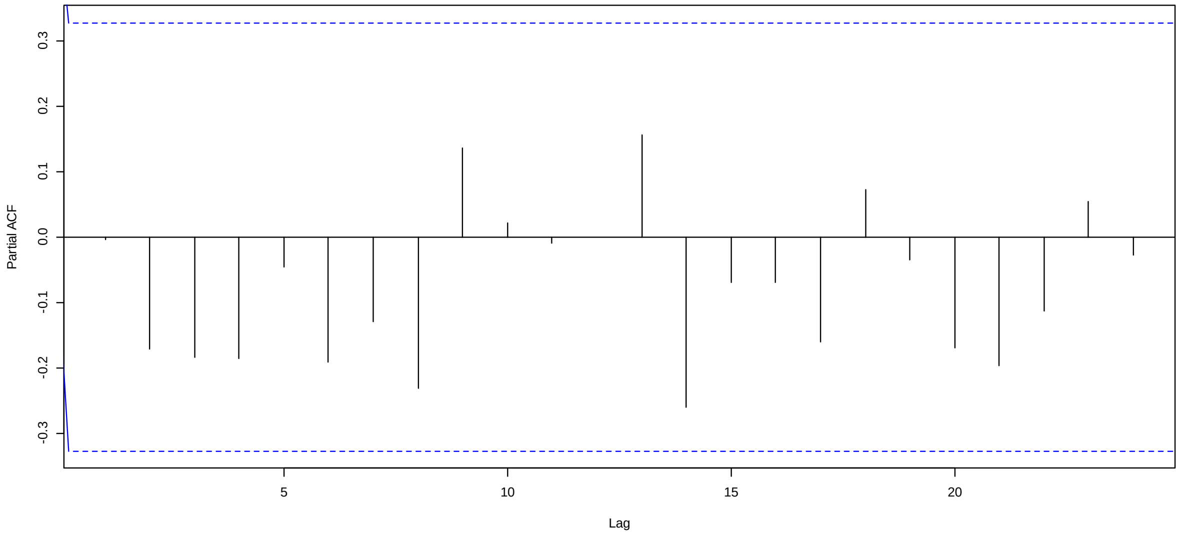
In all the above PACF plots, it can be seen that the sixth lag has significant partial autocorrelation. Now this is interesting because there were observations of the recorded for six days of each week. There may be a weekly seasonal aspect to the time series. This may explain the significant sixth lags in the PACF plots. Hence a seasonal first order AR component, with a period of 6, was added to each model.

Residual analysis was undertaken on the resulting SARIMA models, with the seasonal AR component added. Again there was no model for which the null hypothesis of normality was rejected. The ACF and PACF plots of the standardised residuals are shown below.

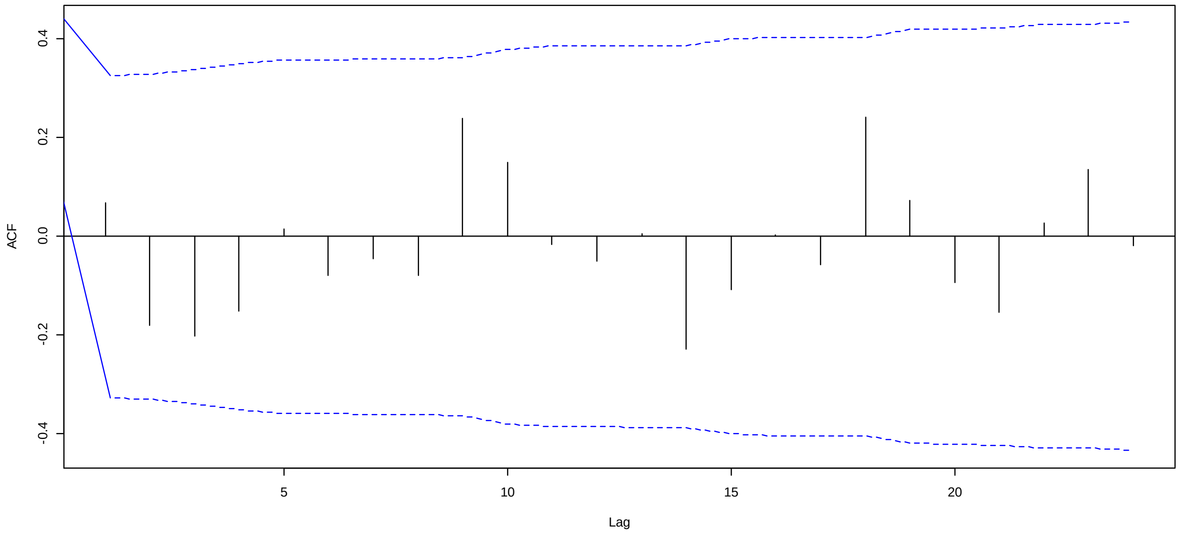
ACF of standardised residuals of SARIMA((1,2,1) * (1,0,0)) model



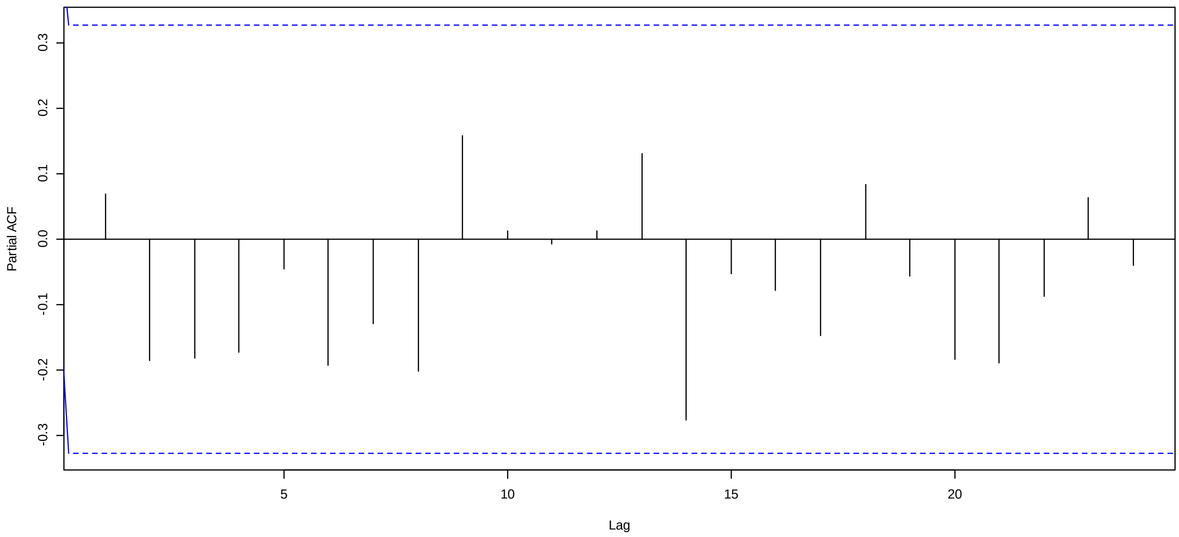
PACF of standardised residuals of SARIMA((1,2,1) * (1,0,0)) model

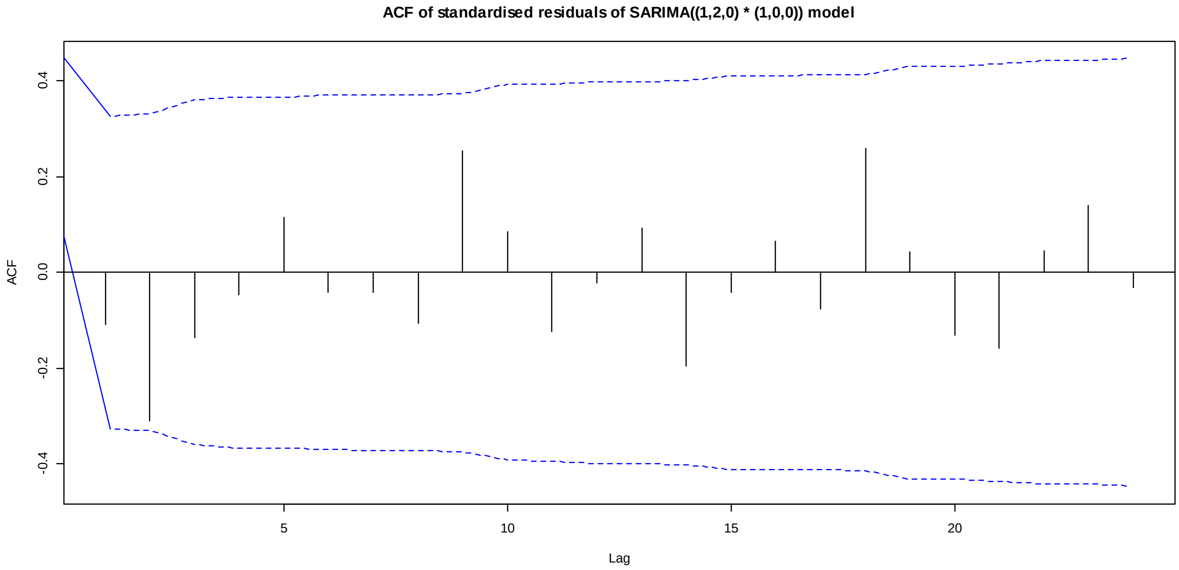


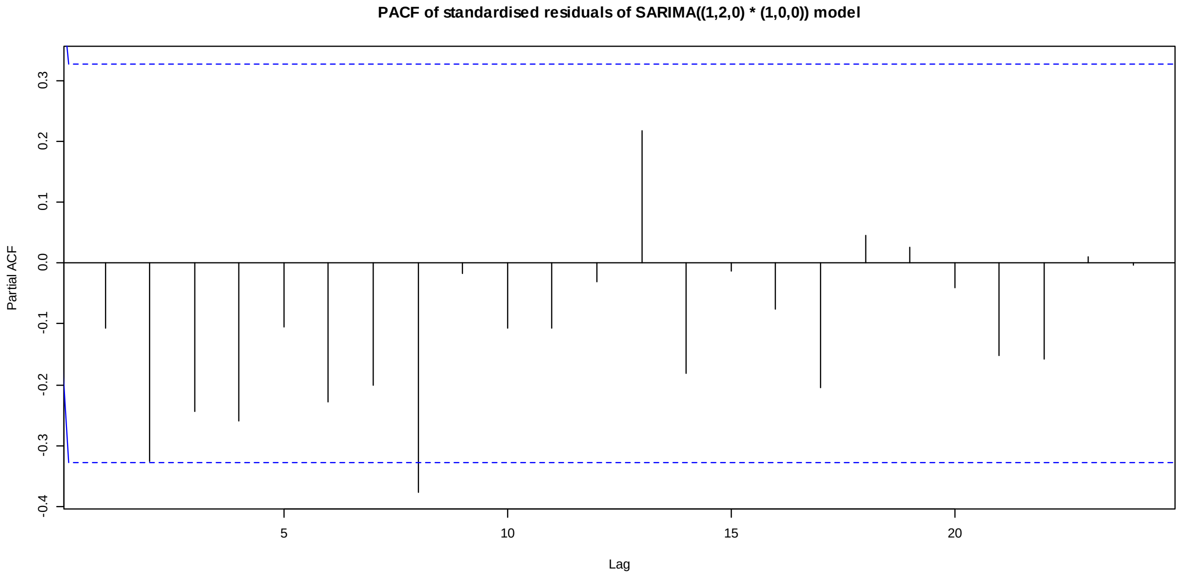
ACF of standardised residuals of SARIMA((0,2,1) * (1,0,0)) model

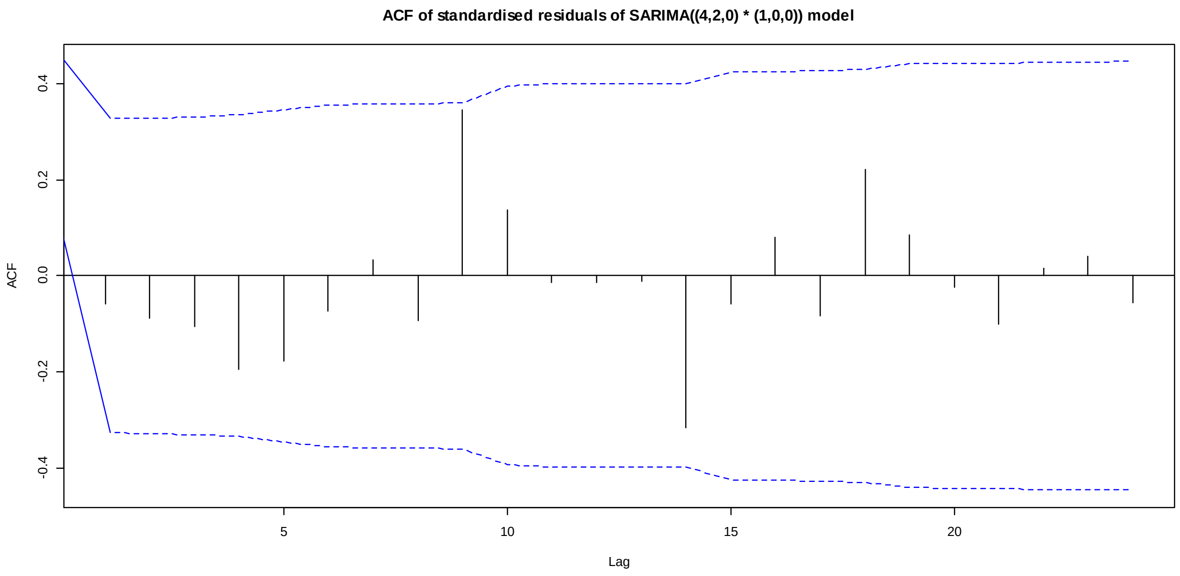


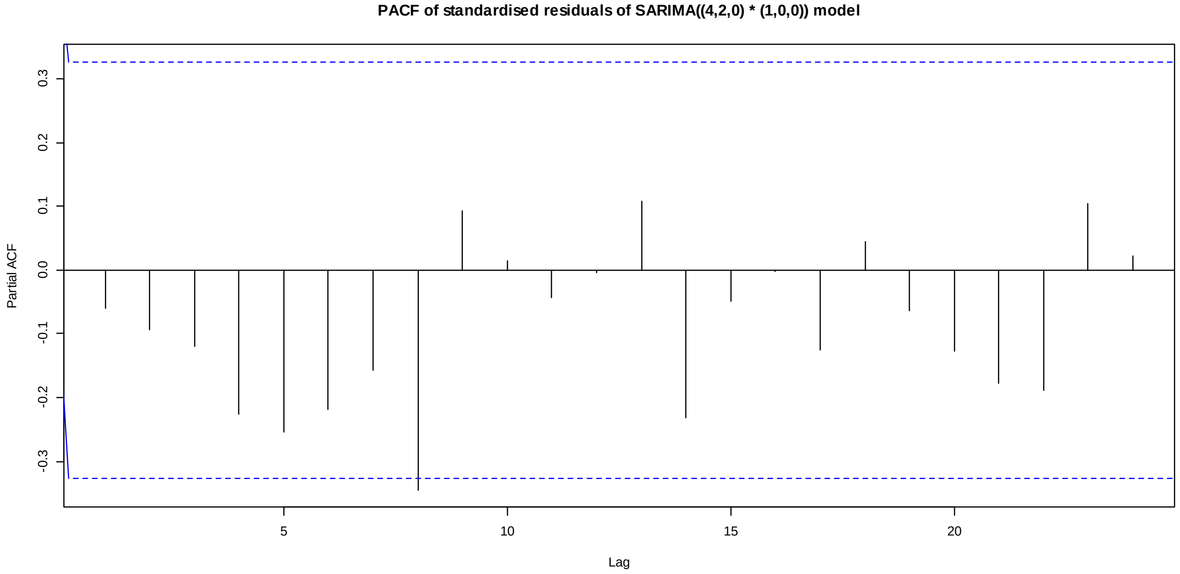
PACF of standardised residuals of SARIMA((0,2,1) * (1,0,0)) model




ACF of standardised residuals of SARIMA((1,2,0) * (1,0,0)) model
ACF
0.4
0.2
0.0
-0.2
-0.4
5
10
15
20
Lag


PACF of standardised residuals of SARIMA((1,2,0) * (1,0,0)) model
Partial ACF
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
5
10
15
20
Lag





The latter two models have significant partial autocorrelation at lag 8, and in the case of the SARIMA((1,2,0)*(1,0,0)) model, lag 2 is on the borderline of being significant. The other two models were analysed further. Firstly the correlations of the residuals were tested further using the Ljung-Box test. Neither result led to the rejection of the null hypothesis that they were not correlated, with p-values of 0.8446 and 0.7737 respectively for the SARIMA((1,2,1)*(1,0,0)) and SARIMA((0,2,1)*(1,0,0)).

Next the AIC was considered. The SARIMA((1,2,1)*(1,0,0)) and SARIMA((0,2,1)*(1,0,0)) models had AIC values of -260.4747 and -262.1972 respectively. As a lower AIC is preferable, this favours the SARIMA((0,2,1)*(1,0,0)) model.

Finally, it will be worthwhile too review the estimates for the coefficients.

SARIMA((1,2,1)*(1,0,0))

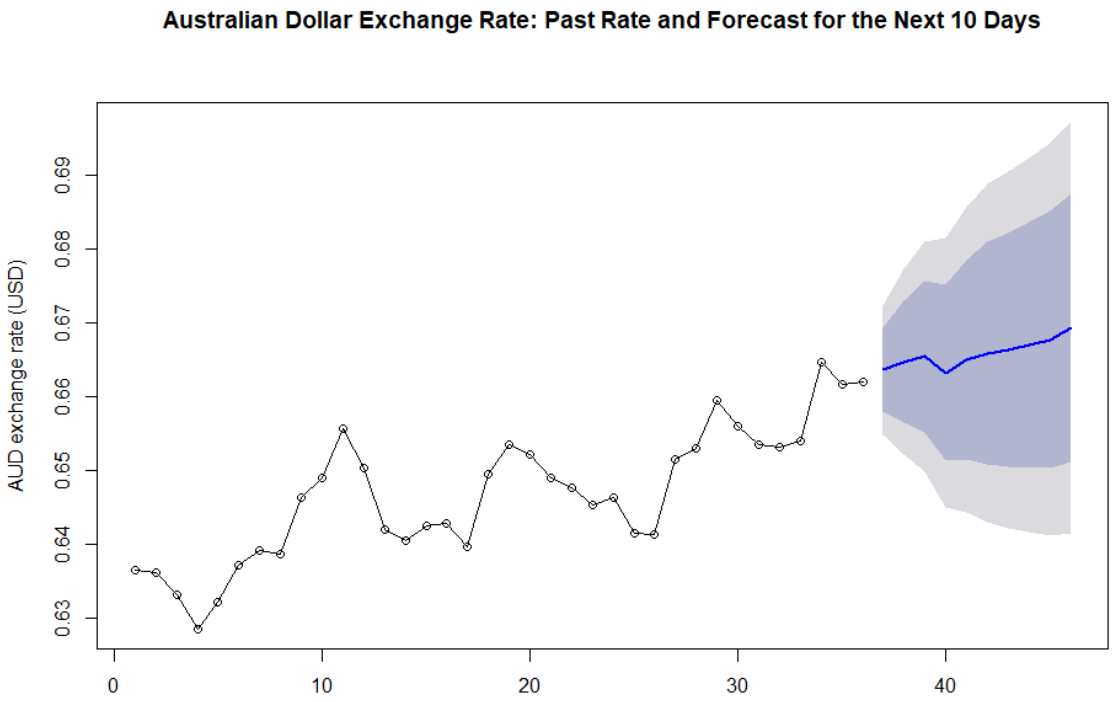
```
z test of coefficients:

      Estimate Std. Error  z value Pr(>|z|)    
ar1   0.090707   0.172351   0.5263  0.59868    
ma1  -0.999998   0.087050 -11.4877  < 2e-16 ***
sar1 -0.292080   0.170485  -1.7132  0.08667 .  
```

SARIMA((0,2,1)*(1,0,0))

```
z test of coefficients:

      Estimate Std. Error  z value Pr(>|z|)    
ma1  -1.000000   0.090262 -11.0789  < 2e-16 ***
sar1 -0.294901   0.169985  -1.7349  0.08276 .  
```

It can be seen that there is little difference in the estimates of the first-order ordinary MA coefficient, or in the first-order seasonal AR coefficient, which are common to both. Also, the additional term in the SARIMA((1,2,1)*(1,0,0)) model, the first-order ordinary AR coefficient, is not significantly different from zero, with a p-value of approximately 0.6. Consequently ,the principle of parsimony means that the SARIMA((0,2,1)*(1,0,0)) should be selected. This decision is supported by the fact that this model has superior AIC.

Having selected a model, it was used to forecast the next 10 instances (which in keeping with the collected data should be the next 10 days not including Saturdays) in the exchange rate time series. The forecast, with 80% and 95% confidence intervals, is shown below.



This model predicts some continued growth in the exchange rate.

Discussion

The time series of the exchange rate showed a steep decline and speedy recovery between days 84 and 113, which corresponded to the time between 6 March and 9 April. After the initial quick recovery there appeared to be a more gradual upward trend, just as before the steep decline there was an extended period of gradual decline. With this in mind, the second approach of modelling the last section separately to make a forecast for the future seems preferable. Of course, as the exchange rate approaches the rate it was trading at before the decline, it may start to flatten out. It would not be expected to grow indefinitely. Nevertheless, it seems appropriate in the short term to work with the model obtained by the second approach, namely the SARIMA((0,2,1)*(1,0,0)) model. This model based on the final sub-section may also be more informative, as it is less influenced by the sudden shocks that took place as traders reacted to the coronavirus and the response to it. This model also has a weekly seasonal component which may assist in understanding any weekly pattern affecting the exchange rate between the Australian dollar and the United States dollar.

As time goes by, different economic circumstances in Australia and the United States will affect the exchange rate. These models were developed based on data in one six-month period. As time goes by, new trends can be expected to emerge.

Conclusion

The time series of the Australian-United States currency exchange rate in the previous six months was analysed and modelled. In this time the coronavirus and the actions taken in response had a major economic impact in many, if not all countries. It was observed that during this time there was a sudden decrease followed by a swift recovery in the price of the Australian dollar in US dollars. Afterwards there appeared to be a more gradual increase in the exchange rate continuing to the end of the series. In light of this it seemed appropriate to make a predictive model based on this final sub-section of the time series, in addition to a model based on the whole series. Using techniques including differencing to obtain a stationary series, considering ACF, PACF, EACF and BIC to determine the number of autoregressive and moving average coefficients, and evaluating models with residual analysis, overfitting and the Akaike Information Criterion, two models with parameters fitted were used to make predictions of the next 10 instances in the exchange rate time series. It was concluded that the model created using only the final 36 instances of the time series was a more useful model for making predictions in the short-term.

Reference

EXCHANGE-RATES.ORG 2020, *US Dollars (USD) per Australian Dollar (AUD)*, Historical Exchange Rate Table, csv file, Exchange-Rates.org, MBH Media, Inc., viewed 29 May 2020, <https://www.exchange-rates.org/history/USD/AUD/T>.

Appendix: R Code

```
# This project will attempt to explore and model the daily Australian Dollar exchange

# rate to US Dollars from 1 December 2019 to 28 May 2020.


# Load packages

library(TSA)

library(fUnitRoots)

library(lmtest)


# Define useful functions

check.normality = function(time_series, title) {

  qqnorm(x = time(time_series), y = time_series, main = paste(title, "Normal Q-Q Plot"))

  qqline(time_series, col = 'red')

  test=shapiro.test(time_series)

  return(test)

}

check.for.stationarity <- function(time.series) {

  a=ar(diff(time.series))$order

  test=adfTest(time.series, lags = a)

  return(test)

}

# check.normality and check.for.stationarity defined in my assignment 2

# though I have improved check.normality by adding a title to the Q-Q plot.


check.acf.and.pacf = function(time.series) {

  win.graph(width = 20, height = 20, pointsize = 8)

  par(mfrow=c(2,1))

  acf(time.series, ci.type='ma', main = 'ACF')

  pacf(time.series, main = 'PACF')

  par(mfrow=c(1,1))

}
```

```
plot.residuals <- function(model) {
  p=as.character(model$call$order)[2]
  d=as.character(model$call$order)[3]
  q=as.character(model$call$order)[4]
  if (is.null(model$call$seasonal)==FALSE) {
    P=as.character(model$call$seasonal$order)[2]
    D=as.character(model$call$seasonal$order)[3]
    Q=as.character(model$call$seasonal$order)[4]
    model_label=paste0("SARIMA((",paste(p, d, q, sep = ","),") * (",paste(P,D,Q, sep = ","),"))")
    label=paste("Standardised residuals of", model_label, "model")
    plot(rstandard(model), type='o', main=label, ylab="Standardised residuals")
  }
  else {
    model_label=paste0("ARIMA(",paste(p, d, q, sep = ","),")")
    label=paste("Standardised residuals of", model_label, "model")
    plot(rstandard(model), type='o', main=label, ylab="Standardised residuals")
  }

}
plot.residuals.acf.and.pacf = function(model) {
  p=as.character(model$call$order)[2]
  d=as.character(model$call$order)[3]
  q=as.character(model$call$order)[4]

  if (is.null(model$call$seasonal)==FALSE) {
    P=as.character(model$call$seasonal$order)[2]
    D=as.character(model$call$seasonal$order)[3]
    Q=as.character(model$call$seasonal$order)[4]
    model_label=paste0("SARIMA((",paste(p, d, q, sep = ","),") * (",paste(P,D,Q, sep = ","),"))")
    label1=paste("ACF of standardised residuals of", model_label, "model")
    label2=paste("PACF of standardised residuals of", model_label, "model")
```

```
    win.graph(width = 20, height = 20, pointsize = 8)
    par(mfrow=c(2,1))
    acf(rstandard(model), ci.type='ma', main = label1, lag.max = 24)
    pacf(rstandard(model), main=label2, lag.max = 24)
  }
  else {
    model_label=paste0("ARIMA(",paste(p, d, q, sep = ","),")")
    label1=paste("ACF of standardised residuals of", model_label, "model")
    label2=paste("PACF of standardised residuals of", model_label, "model")

    win.graph(width = 20, height = 20, pointsize = 8)
    par(mfrow=c(2,1))
    acf(rstandard(model), ci.type='ma', main = label1, lag.max = 24)
    pacf(rstandard(model), main=label2, lag.max = 24)
  }
  par(mfrow=c(1,1))
}
# plot.residuals and plot.residuals.acf.and.pacf developed using code from plot.residuals function from my assignment 2

# Load data and convert to time series
AUS.US.DOLLAR.EXCHANGE.RATE=read.csv("HistoryExchangeReport.csv")
exchange.rate=ts(AUS.US.DOLLAR.EXCHANGE.RATE$Rate)

# Display plot of time series
plot(exchange.rate, type = 'o', ylab = 'AUD exchange rate (USD)', main ='Australian Dollar Exchange Rate')

# Check if normal
check.normality(exchange.rate,'Exchange Rate')
#data points not normally distributed
```

```
# Check if stationary
check.for.stationarity(exchange.rate)

# The time series clearly looks non-stationary. Taking the difference may solve this
diff.exchange.rate=diff(exchange.rate, differences = 1)

# Plot the differenced series
plot(diff.exchange.rate, type = 'o', ylab = 'difference in daily AUD exchange rate (USD)', main ='First Difference of Exchange rate')

# Check if differenced series is normal
check.normality(diff.exchange.rate,'Differenced Exchange Rate')
# data is not normally distributed

# The lack of normality is a problem.
# Check if there is a transformation that will normalise the data
# Try taking logarithm
log.exchange.rate=log(exchange.rate)
check.normality(log.exchange.rate,'Logarithm of exchange rate')
check.normality(diff(log.exchange.rate),'Difference of logarithm of exchange rate')

# Try BoxCox transform
lambda=BoxCox.ar(exchange.rate)$mle
bc.exchange.rate=((exchange.rate-1)^lambda)/lambda
check.normality(bc.exchange.rate, 'Power transform of exchange rate')
check.normality(diff(bc.exchange.rate),'Difference of power transform of exchange rate')

# Plot normality of original and differenced series together for report
win.graph(height = 18, width = 12, pointsize = 8)
par(mfrow=c(2,1))
```

```
check.normality(exchange.rate,'Exchange Rate')
check.normality(diff.exchange.rate,'Differenced Exchange Rate')

# Check if differenced series is stationary
check.for.stationarity(diff.exchange.rate)
# Apparently the differenced time series is stationary

# Check ACF and PACF of differenced series
check.acf.and.pacf(diff.exchange.rate)

# Check EACF
eacf(diff.exchange.rate)

# Check BIC
res=armasubsets(diff.exchange.rate, nar = 1, nma = 4)
plot(res)



# ARIMA(0,1,1) or ARIMA(1,1,1) are candidate models
# Now estimate parameters for both these models
m1=arima(exchange.rate, order = c(0,1,1))
coeftest(m1)

m2=arima(exchange.rate, order = c(1,1,1))
coeftest(m2)

# Residual analysis
plot.residuals(m1)
check.normality(resid(m1), 'ARIMA(0, 1, 1)')
plot.residuals.acf.and.pacf(m1)
```

```
plot.residuals(m2)

check.normality(resid(m2), 'ARIMA(1, 1, 1)')

plot.residuals.acf.and.pacf(m2)

# Neither model has normally distributed residuals.

# However the first difference of the exchange rate is not normally distributed so this is not surprising

# Overall the correlations of the ARIMA(1,1,1) lags appear smaller,

# though it does have a significant correlation at lag 18.

# But to get an additional assesment of the residuals, the Ljung-Box test will be employed

Box.test(rstandard(m1), lag = 18, type = "Ljung-Box")

Box.test(rstandard(m2), lag = 18, type = "Ljung-Box")

# In both cases, the p-value is larger than 0.05 so cannot reject the null hypothesis that the lags are uncorrelated


# Consider generalisations of above models by including an additional ma term


# Extending ARIMA(0,1,1)

# If the coefficients common to both ARIMA(0,1,1) and ARIMA(0,1,2) do not differ

# significantly, and the additional term is not significant, then this supports

# the original ARIMA(0,1,1) model

m3=arima(exchange.rate, order = c(0,1,2))

m3

coeftest(m3)

# compare to ARIMA(0,1,1)

coeftest(m1)

# The difference in the common coefficient is not statistically significant.

# Neither is the additional (ma2) coefficient significantly different from zero

# So the ARIMA(0,1,1) model seems good, and the ARIMA(0,1,2) can be discarded.


# Extending ARIMA(1,1,1)

# If the coefficients common to both ARIMA(1,1,1) and ARIMA(1,1,2) do not differ

# significantly, and the additional term is not significant, then this supports
```

the original ARIMA(1,1,1) model

m4=arima(exchange.rate, order = c(1,1,2))

m4

coeftest(m4)

compare to ARIMA(1,1,2)

coeftest(m2)

The difference in the common coefficient is not statistically significant.

Neither is the additional (ma2) coefficient significantly different from zero

So the ARIMA(1,1,1) model seems good, and the ARIMA(1,1,2) can be discarded.

#Cannot use AIC as data not normally distributed

#Choose ARIMA(0,1,1) due to the principle of parsimony

library(forecast)

preds=forecast(exchange.rate, model = m1, h = 10)

plot(preds, type='o', ylab = 'AUD exchange rate (USD)', main ='Australian Dollar Exchange Rate: Past Rate and Forecast for the Next 10 Days')

Another model that may be more accurate is a model considering only the days

since the dramatic fall and recovery in March.

Consider original time series to see when to start this time series

plot(exchange.rate, type = 'o', ylab = 'AUD exchange rate (USD)', main ='Australian Dollar Exchange Rate')

since about day 113, there has been an upward trend.

However it may be worth starting the series a little later to increase confidence that the time series is over its dramatic fluctuations.

Hence start a series from the 120th recorded day in the series.

create a time series starting from this point

steady.increase=exchange.rate[120:155]

This series will obviously have an upward trend and so will need to be differenced

But first view its plot

```
plot(steady.increase, type='o',ylab = 'AUD exchange rate (USD)', main ='Australian Dollar Exchange Rate')

# Check normality
check.normality(steady.increase, 'Exchange rate from 17 April')
# For this shorter time series, the null hypothesis of normality cannot be rejected

# Check stationarity
check.for.stationarity(steady.increase)
# As already obseved, time series not stationary

# Take first difference
diff.steady.increase=diff(steady.increase)

# Plot the differenced series
plot(diff.steady.increase, type='o', ylab = 'difference in exchange rate', main='First difference of exchange rate')
# There still appears to be a slight upward trend.

# Check normality
check.normality(diff.steady.increase, 'First difference of exchange rate')
# For the first difference also, the null hypothesis of normality cannot be rejected

# check stationarity
check.for.stationarity(diff.steady.increase)
# The null hypothesis of non-stationarity cannot be rejected.
# It seems the "steady increase" is not so steady after all!

# Take the second difference and plot it.
diff.steady.increase=diff(steady.increase, differences=2)
plot(diff.steady.increase, type='o', ylab = '2nd difference in exchange rate', main='Second difference of exchange rate')
```

```
# Check if is normally distributed
check.normality(diff.steady.increase, '2nd Difference of Exchange Rate')
# For the second difference, the null hypothesis of normality cannot be rejected.

# Check if it is stationary
check.for.stationarity(diff.steady.increase)
# The null hypothesis of non-stationarity is rejected. P-value < 0.01.

# Check its autocorrelation and partial autocorrelation
check.acf.and.pacf(diff.steady.increase)
# the first lag has significant autocorrelation and partial autocorrelation
# So ARIMA(1,2,1) may be a good model

# Check eacf
eacf(diff.steady.increase, ar.max = 8, ma.max = 6)
# The top left corner of 'o's is at (0,1). So ARIMA(0,2,1) may be a good model.

# Check BIC
res2=armasubsets(diff.steady.increase, nar = 6, nma = 4)
plot(res2)
# Based on BIC, ARIMA(1,2,0) and ARIMA(4,2,0) are candidate models.

# Set of candidate models: {ARIMA(1,2,1), ARIMA(0,2,1), ARIMA(1,2,0) and ARIMA(4,2,0)}
# Consider ARIMA(1,2,1) model
model1=arima(steady.increase, order = c(1,2,1))
plot.residuals(model1)
check.normality(resid(model1),'ARIMA(1,2,1)')
plot.residuals.acf.and.pacf(model1)

# ARIMA(0,2,1) model
```

```
model2=arima(steady.increase, order = c(0,2,1))

plot.residuals(model2)

check.normality(resid(model2),'ARIMA(0,2,1)')

plot.residuals.acf.and.pacf(model2)


# ARIMA(1,2,0) model

model3=arima(steady.increase, order = c(1,2,0))

plot.residuals(model3)

check.normality(resid(model3),'ARIMA(1,2,0)')

plot.residuals.acf.and.pacf(model3)


# ARIMA(4,2,0) model

model4=arima(steady.increase, order = c(4,2,0))

plot.residuals(model4)

check.normality(resid(model4),'ARIMA(4,2,0)')

plot.residuals.acf.and.pacf(model4)


# In all the above plots the PACF of the residuals was outside the error bounds at lag 8.



# There may be some seasonal (weekly) trend.

# As Saturdays were not included, there were only 6 days of the week given.

# Hence the period of a weekly seasonal model would be 6.

# This could help explain why the residuals each of the above models had a significant partial
autocorrelation in lag 6.

# So adding a seasonal first order AR component, with a period of 6, to the above models may better
reflect the time series being modelled.


# Consider SARIMA((1,2,1)*(1,0,0)) model

model1.1=arima(steady.increase, order = c(1,2,1), seasonal = list(order=c(1,0,0), period=6))

plot.residuals(model1.1)

check.normality(resid(model1.1),'SARIMA((1,2,1),(1,0,0))')
```

```
plot.residuals.acf.and.pacf(model1.1)

# SARIMA((0,2,1)*(1,0,0)) model
model2.1=arima(steady.increase, order = c(0,2,1), seasonal = list(order=c(1,0,0), period = 6))
plot.residuals(model2.1)
check.normality(resid(model2.1),'SARIMA((0,2,1),(1,0,0))')
plot.residuals.acf.and.pacf(model2.1)

# SARIMA((1,2,0)*(1,0,0)) model
model3.1=arima(steady.increase, order = c(1,2,0), seasonal = list(order=c(1,0,0), period = 6))
plot.residuals(model3.1)
check.normality(resid(model3.1),'SARIMA((1,2,0),(1,0,0))')
plot.residuals.acf.and.pacf(model3.1)
# lags 8 (and to a lesser extent lag 2) in the PACF are concerning

# SARIMA((4,2,0)*(1,0,0)) model
model4.1=arima(steady.increase, order = c(4,2,0), seasonal = list(order=c(1,0,0), period = 6))
plot.residuals(model4.1)
check.normality(resid(model4.1),'SARIMA((4,2,0),(1,0,0))')
plot.residuals.acf.and.pacf(model4.1)
# The PACF does not look quite as bad as for the previous model, but lag 8 is still outside the white noise bounds

# In the last two models there appears to be significant partial autocorrelation in the residuals in lag 8.

# These models shall be discarded for now.
# SARIMA((1,2,1)*(1,0,0)) and SARIMA((0,2,1)*(1,0,0)) will be compared to see which candidate model is better

# Consider the residuals autocrrelations further. Use Ljung-Box test.
Box.test(resid(model1.1), lag = 8, type = "Ljung-Box")
```

Box.test(resid(model2.1), lag = 8, type = "Ljung-Box")

The Ljung-Box text does not lead to rejection of the null hypothesis that lags are uncorrelated for either model.

AIC for each model should be considered

AIC(model1.1)

AIC(model2.1)

The SARIMA((0,2,1)*(1,0,0)) has better (i.e. lower) AIC

The parameters and the significance of each can be checked.

coeftest(model1.1)

coeftest(model2.1)

The two models both have an ordinary first order MA term and a seasonal first order AR term.

The parameters of these common terms are not significantly different.

The SARIMA((1,2,1)*(1,0,0)) model also has an ordinary first order AR term, but it is not significant with a p-value about 0.6.

The principle of parsimony favours the SARIMA((0,2,1)*(1,0,0)) model.

It seems best to select this model, especially given it has lower AIC.

Now predict the next 10 values for the exhange rate with this model

prediction=forecast(steady.increase, model=model2.1, h = 10)

plot(prediction, type='o', ylab = 'AUD exchange rate (USD)', main ='Australian Dollar Exchange Rate: Past Rate and Forecast for the Next 10 Days')