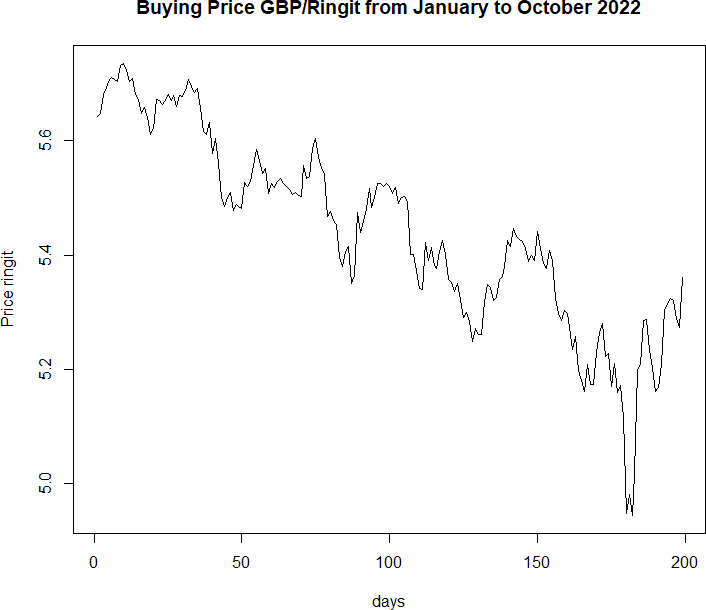
TIME SERIES ANALYSIS OF THE BUYING PRICE OF THE BRITISH POUND AGAINST MALAYSIAN RINGGITS FROM JAN 2022 TO OCTOBER 2022

In the context of our Time Series Analysis and Forecasting, we are going to perform

a study about a data set related to Foreign Exchange. Analysis of the buying price of the British Pound (GDP) against the Malaysian ringgit will be performed. The data will be extracted from Bank Negara Malaysia's official website

(<https://www.bnm.gov.my/web/guest/exchange-rates>) from January 2022 to October 2022.



## Figure 1.0 Buying price of GBP/ringgit from January 2022 to November 2022

Components of the series.

# Trend.

The time series plot reflects a downward movement of the British pound with time. As the graph plots the exchange rate of GDP to the Malaysian Ringit from January 2022 to November 2022. The graph starts with an exchange rate of 5.645 GBP, but fluctuations occur during the period the GBP decreases slowly over the period finally ending with an exchange rate of 5.306.

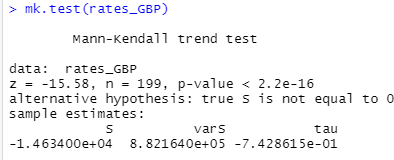
# Irregular variation (random variation)

Additionally, the time series reflects an irregular variation component as there is a sudden decrease in the exchange rate along the period over a short period. The graph shows that GBP hit an unforeseen downfall in value in September. According to Forbes, On September 26, the pound struck an all-time low versus the dollar by dropping to $1.03.

# Mann Kendell test

*Ho = No trend component*

*H1 = Existence of increasing or decreasing trend component*



## Figure 2.0: Output of Mann-Kendall trend test

As the p-value is less than a 5% level of significance we reject the Null hypothesis. Generally, the time series contains a trend component.

# Question (b)

From the previous analysis, the analysis and statistical test show the data has a trend component. Therefore, appropriate forecasting techniques are to be applied.

# Holt Method.

This is a forecasting technique which is used in a series with a trend component and has data that allows the anticipation of future upward movement. It is an exponential smoothing technique with an addition of a trend parameter to it.

# Equation

Ft + k = Lt + kTt

Lt = α At + (1 –α)(Lt-1+ Tt-1)

Tt = β(Lt - Lt-1) + (1- β) Tt-1

At = Actual value in period t

α = Smoothing constant to data β = Smoothing constant to trend K = Number of periods

# Simple Exponential model.

Although the method is not appropriate for data with the trend, the forecast reacts strongly to the immediate changes in data. Moreover, the exponential method is simpler and easier to use than other methods.

# Equation.

Ft + 1 = Ft + α(At - Ft) Or

Ft + 1 = α At + α(1 – α) At – 1 + (1 – α)2Ft – 1 + …….

# Linear and Exponential trend

The regression model permits a linear relationship to exist between a single predictor variable and the predicted variable y. The input properties of the linear regression method teach it how to create a weighted sum:

Linear regression will be useful since it can produce a formula that smoothly follows the rising trajectory identified in the data because the time series has a trend.

Yt = β0 + β1t+β2+β3+,,,,+ εt

Yt = Predicted values

β0 = Intercept

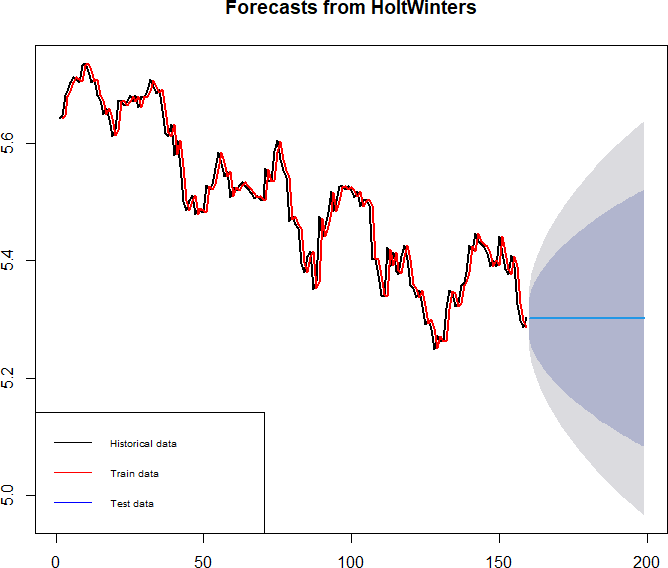
βt = Variable coefficient at time(t)

# Question (c)

The data is then split into training and testing sets to determine and forecast the data using the appropriate methods chosen above:

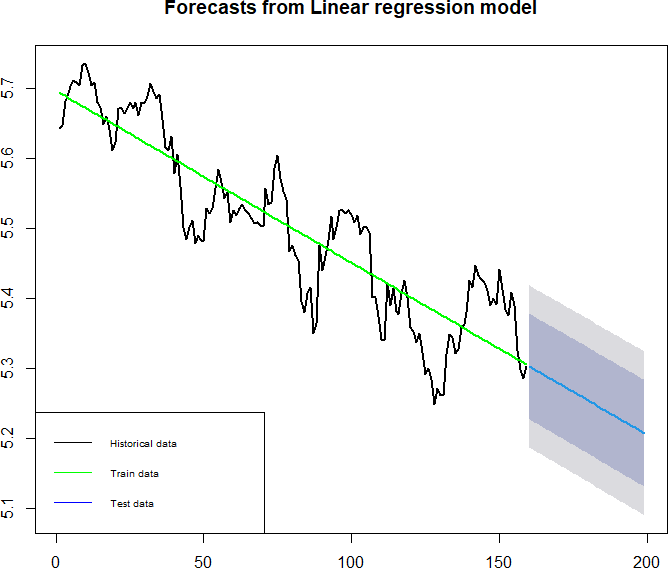
# Simple Exponential smoothing

Output



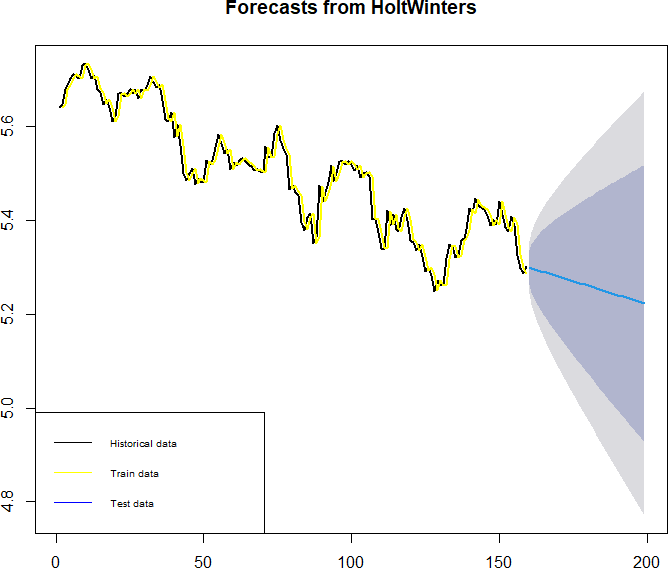
***Figure 3.0: Output Plot for Simple exponential smoothing***

# Linear trend model Output



***Figure 3.1: Output Plot for Linear Trend Model***

# Holts Method Output



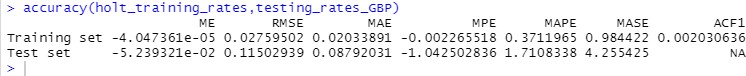
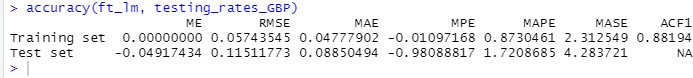
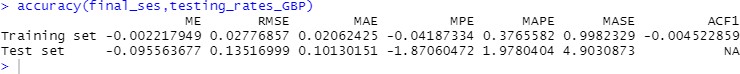
***Figure 3.2: Output Plot for Holt Winters smoothing***

# Question (d) Accuracy measures

From the analysis, two measures are used to test the performance of the forecast: RMSE and MAPE.

RMSE is the standard deviation of the prediction errors (residuals). The advantage of this accuracy measure is that this measure has the same units as the data series.

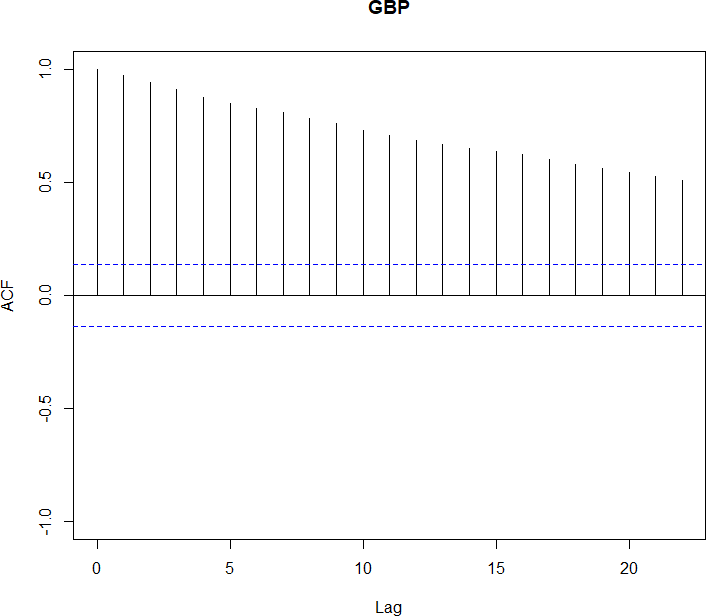
MAPE gives a percentage score of how the forecast deviates from the actual values. The advantage is that it’s useful for comparing performance across a series of data.



## Figure 4.0: Accuracy outputs for all models

From the analysis, Holt’s RMSE and MAPE values are 0.1150 and 1.7108. The Holt method has the least RMSE and MAPE values compared to other methods. Therefore, the Holt method performs better than other models.

# Question (e) ACF plot

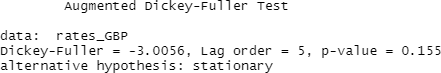


## Figure 5.0: ACF plot

The ACF plot shows the autocorrelation of the time series, from the plot it shows that the values are highly correlated to one another. From the graph, it is noted that there is a large significant ACF from the values. Moreover, there is a slow decrease in the size of the autocorrelations.

Generally, the ACF plot confirms there is a Trend component due to the high correlation of the values, Moreover ACF plots show the series is non-stationary but a test is needed to prove the hypothesis.

# Adf Test



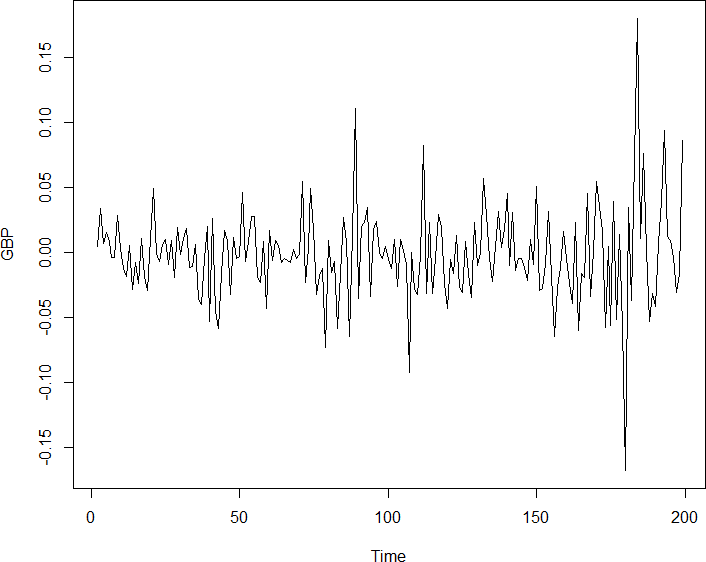
## Figure 5.1: Output for ADF test

**H0:** The series is not stationary

**H1**: The series is stationary

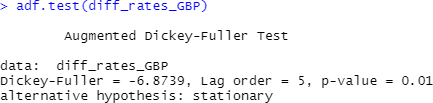
The p-value is 0.155 is more than 0.05 which means we accept the null hypothesis. As the ADF test shows the series as not stationary, differencing is then performed. **Differencing**.

The ADF test proves the time series is non-stationary, therefore the differencing is applied to the series to achieve stationarity in the series.



***Figure 5.2: Output plot after first differencing.***

# ADF test



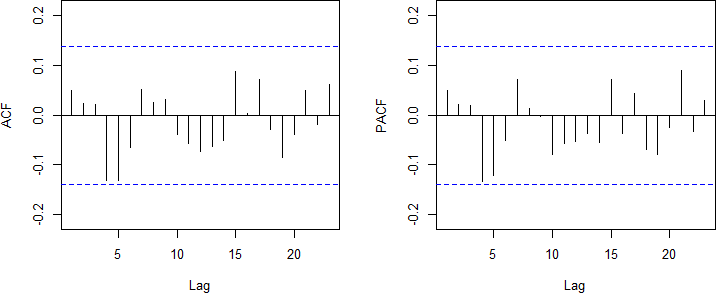
## Figure 5.3: Output for ADF test (ii)

The p-value is less than 5% therefore the series is stationary, after the first differencing.

# Question (f) (i)

From the previous analysis, the test proves the time series is not stationary. Therefore, in determining the suitable ARIMA models of the time series it is vital to use the differenced time series to determine the orders of the AR and MA models.

ACF and the PACF model of the differenced time series will be used to determine the order of the model.

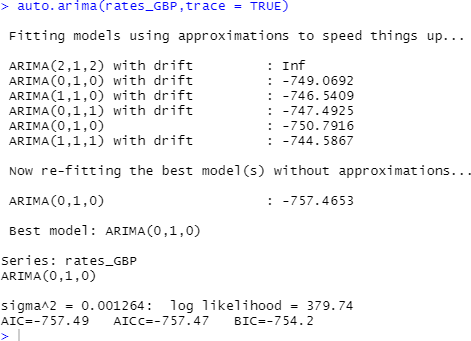


## Figure 6.0: Output plot for ACF and PACF

From the ACF and PACF plots, it can be observed that there are no significant spikes in both plot. The ACF plot which is used to determine the order of the Moving Average model (MA) shows there is no significant spike in all of the lags. Moreover, the PACF plot which determines the order of the Autoregressive model (AR) shows there is no significant spike in all of the lags.

Therefore, the observations from the ACF and PACF plots suggest that the first model is ARIMA (0,1,0).

Moreover, **auto.arima** function is used to determine the models which are available from the time series data. And the comparison of accuracy will be made to determine the viable second model for the time series data.



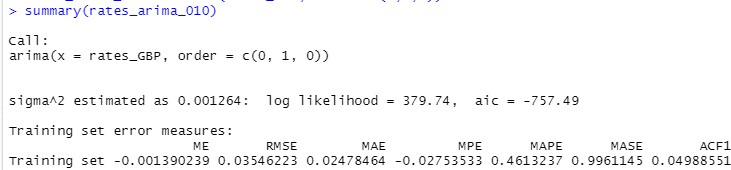
## Figure 6.1: Output auto. arima

From the figure 6.1 above, the **auto. arima** function suggests 6 new models. Firstly, we will choose the models with no drift which from the figure above is the first model we proposed. Secondly, we will compare the AIC values. The model with the least AIC value is the better fit for the model.

Therefore, from the figure above, the ARIMA(1,1,1) has the least AIC value and thus will be the second model.

# (ii)

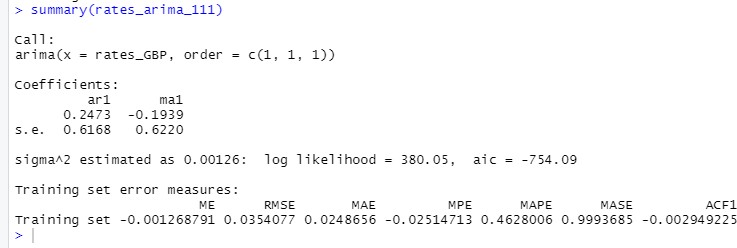
Summary output for ARIMA(0,1,0)



## Figure 7.0: Summary Output

Figure 7.0 above shows the model performs better than the naive model as its MASE value is less than 1.

Summary output for ARIMA(1,1,1)

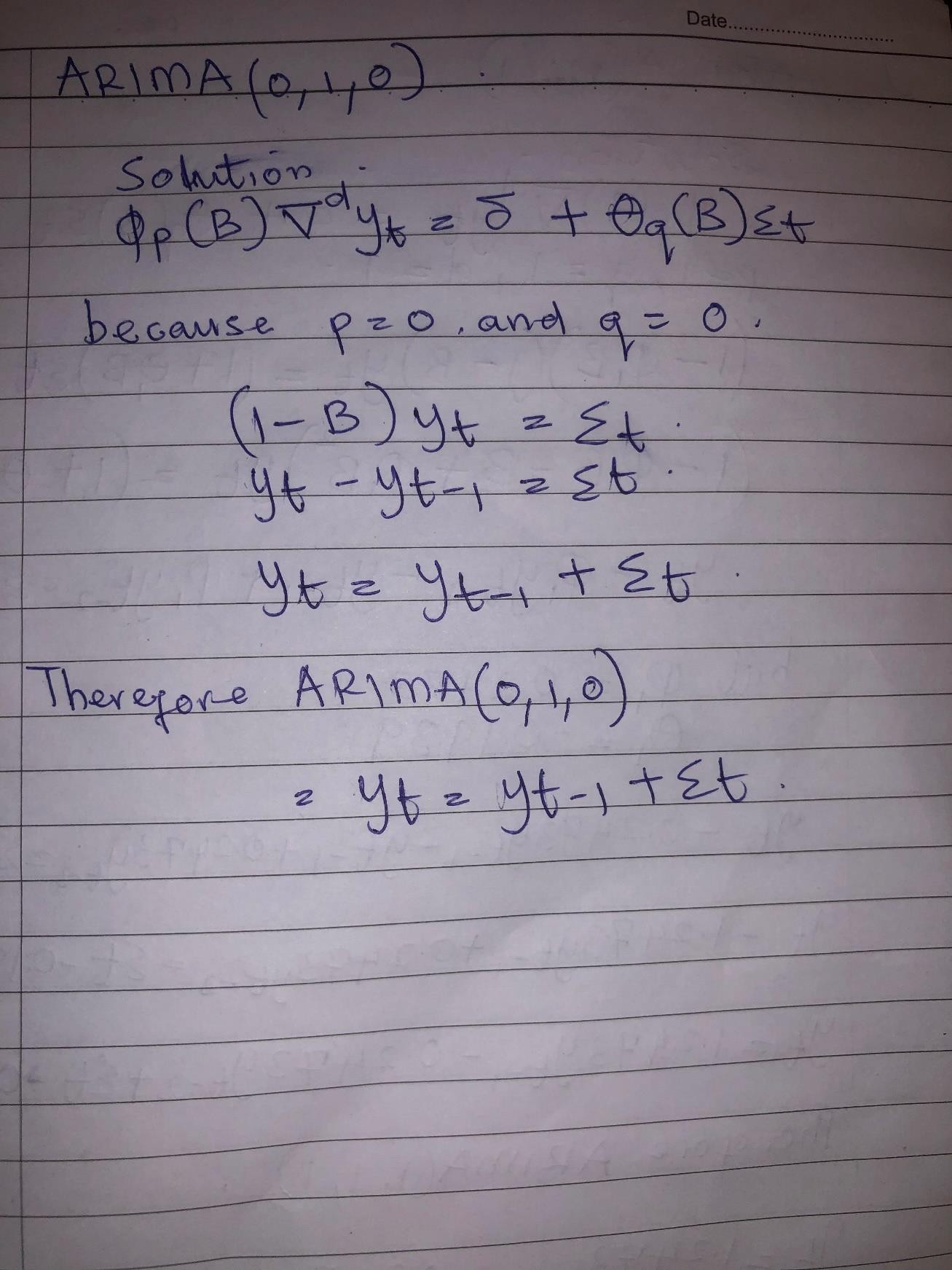


## Figure 7.1: Summary Output

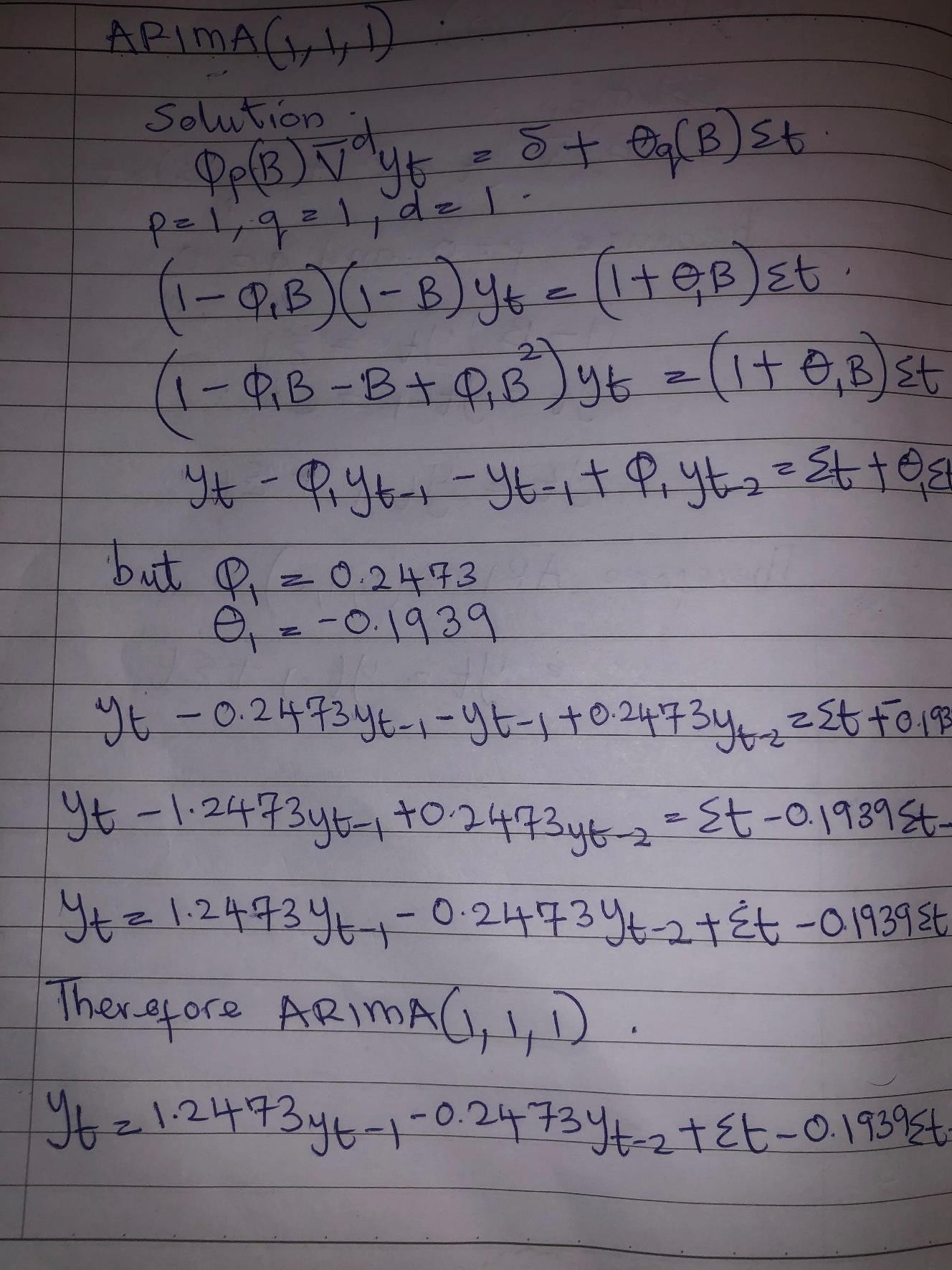
The figure above shows the model performs better than the naïve model MASE < 1.

# (iii)

Equation of the model ARIMA(0,1,0)



The equation of model ARIMA(1,1,1)



# (iv)

The overall model accuracy is provided by a chi-square test based on the Lyung Box statistic. According to the Lyung Box test, we test the null hypothesis to the Alternative hypothesis.

H0 = The model is adequate (errors are independent)

H1 = The model is not adequate (errors are not independent)

From the model ARIMA (0,1,0) the p-value is 0.4499 > 0.05 therefore we accept the null hypothesis, and the model is adequate.

From the model ARIMA (1,1,1) the p-value is 0.3754 > 0.05 therefore we accept the null hypothesis, and the model is adequate.

# (v)

To test for the significance of the parameters the coeftest function is used. According to this test, the parameters are retained if the p-value < 0.05.

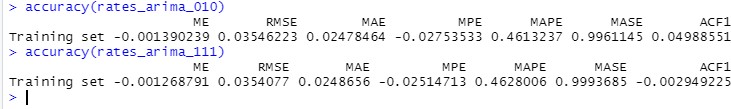
H0: θ = 0. H1: θ ≠ 0.

From the model ARIMA (1,1,1) the p-values of both coefficients are > 0.05, therefore the accepts Ho the coefficients are not significant.

The model ARIMA (0,1,0) cannot be tested as it does not have any coefficients.

# (vi)

From the analysis, two measures are used to test the performance of the forecast: RMSE and MAPE.



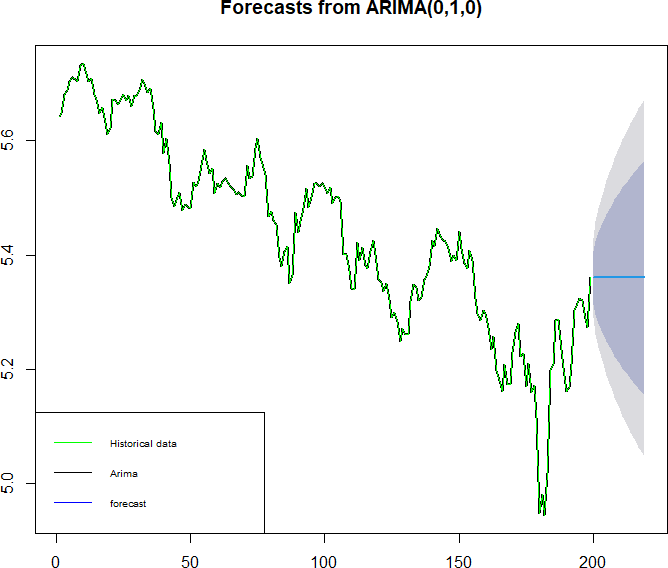
## Figure 8.0: Accuracy outputs

From the figure above, out of the two models, ARIMA (1,1,1) has the lowest RMSE value compared to ARIMA (0,1,0). But also, the ARIMA (0,1,0) has the least MAPE value compared to the other model.

Generally, the best model for this time series data is ARIMA (0,1,0) although it doesn’t have the least RMSE. Nevertheless, the coefficients of the ARIMA (1,1,1) are not significant. Therefore, the best model reduces to ARIMA (0,1,0).

# Question (g)

From the analysis, we have concluded that the best model for this data is ARIMA (0,1,0). Therefore, the model is then used to forecast the data for the next 20 periods.

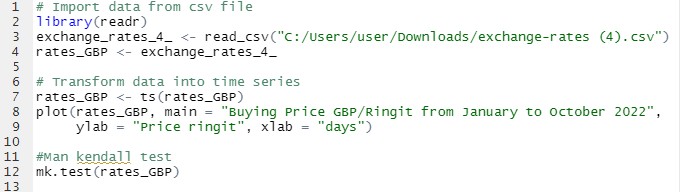


***Figure 9.0: Output plot for arima(0,1,0) model***

# APPENDIX

**Question (a)**

Code

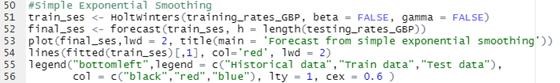


# Question (c)

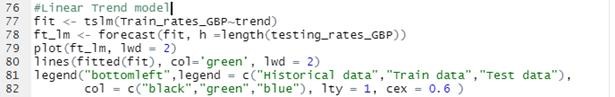
Splitting data



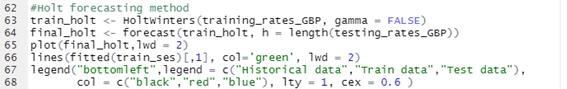
Simple Exponential smoothing



Linear Trend Model



Holt forecasting method

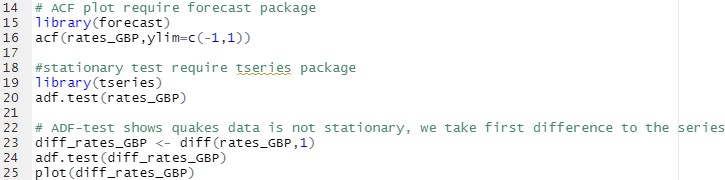


# Question (d)

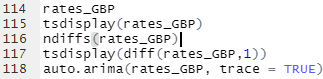
Accuracy Measures



# Question (e)



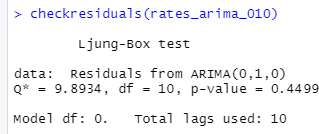
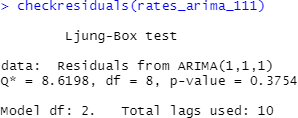
**Question (f) (i)**

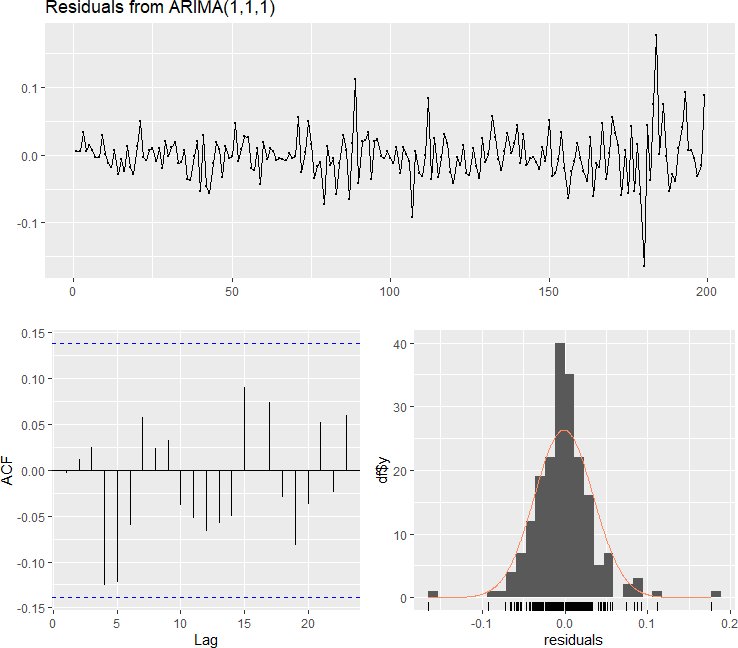


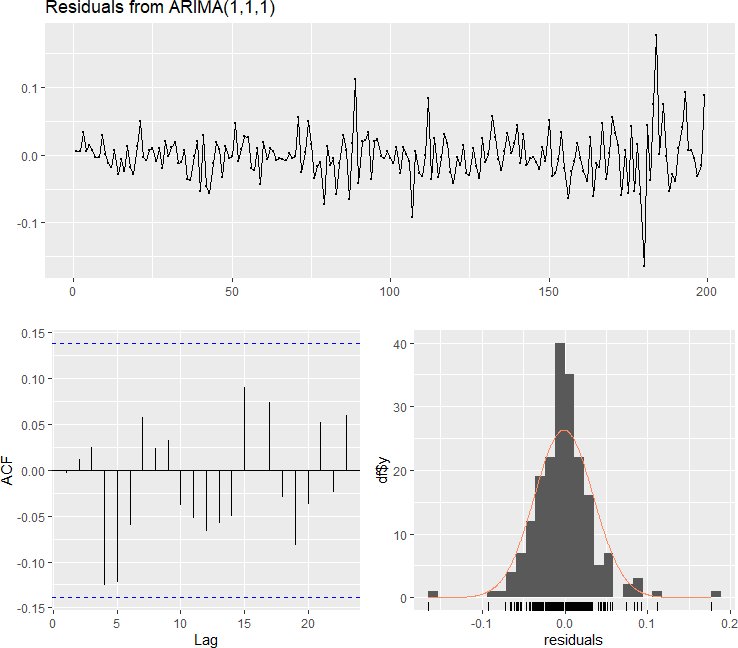
# (ii)



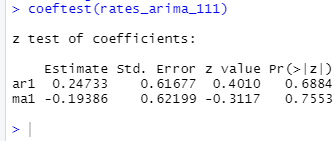
**Question (f) (iv)**







# Question (f) (iv)



**Question (g)**

