TIME SERIES ARIMA MODEL DEVELOPMENT

UK Unemployment Rate

Author: Michael Dubem Igbomezie

Table of Content

1. Introduction

This project focuses on gathering time series data, transforming the data, building a suitable ARIMA model for the data, forecasting future values using the built model, and comparing this forecast with the real world outcome. The purpose of the project is to bridge the gap between the theories taught in the classroom setting with real world applications.

The Annual Unemployment Rate of the United Kingdom (UNRTUKA) was selected for this project. This series was constructed by the Bank of England as part of the Three Centuries of Macroeconomic Data project by combining data from a number of academic and official sources. The periodic range of the dataset is from the year 1760 to 2016. The Gretl Software was used for this project.

UNRTUKA

Figure 1. UNRTUKA Time Series Plot

2. Methodology and Intermittent Results

A. Model Selection

The autocorrelation of many different dataset (annual and seasonally adjusted monthly) were checked for white noise properties before our UNRTUKA dataset was chosen. Majority of these dataset were either not statistically different from a white noise or very close.

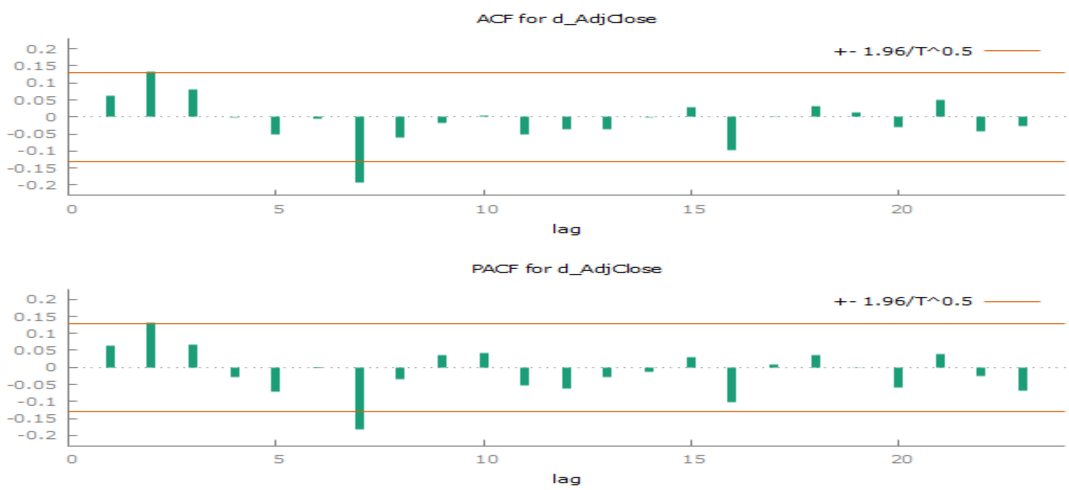

Figure 2. ACF and PACF Plot of GBPUSD Monthly Data

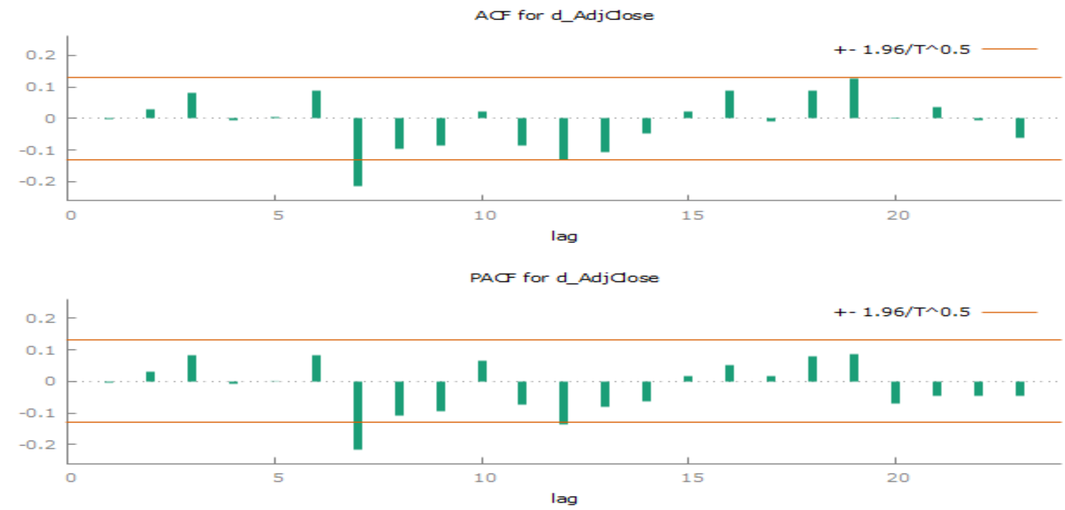

Figure 3. ACF and PACF Plot for EURUSD Monthly Data

This data gathering process showed how a lot of real world data can be likened to a white noise, unlike the classroom example where our data is artificially generated to fit our purpose. From the ACF plot for the GBPUSD dataset, it can be seen that the second lag is just at the significance band and the only other lag outside the significance band is the 7th lag. We cannot put high confidence in both these lags as the first is just at the significance band and the other is a relatively high lag. Similarly for the EURUSD ACF Plot, the only lag outside the significance band is the 7th lag which again is a relatively high lag, therefore we cannot put high confidence in this.

After searching for a while, the UNRTUKA dataset was eventually checked and selected. Due to the fact that the UNRTUKA dataset didn't appear to be stationary, the first difference of the dataset was taken. From the plot below, it can be observed that stationarity was achieved with just the first difference.

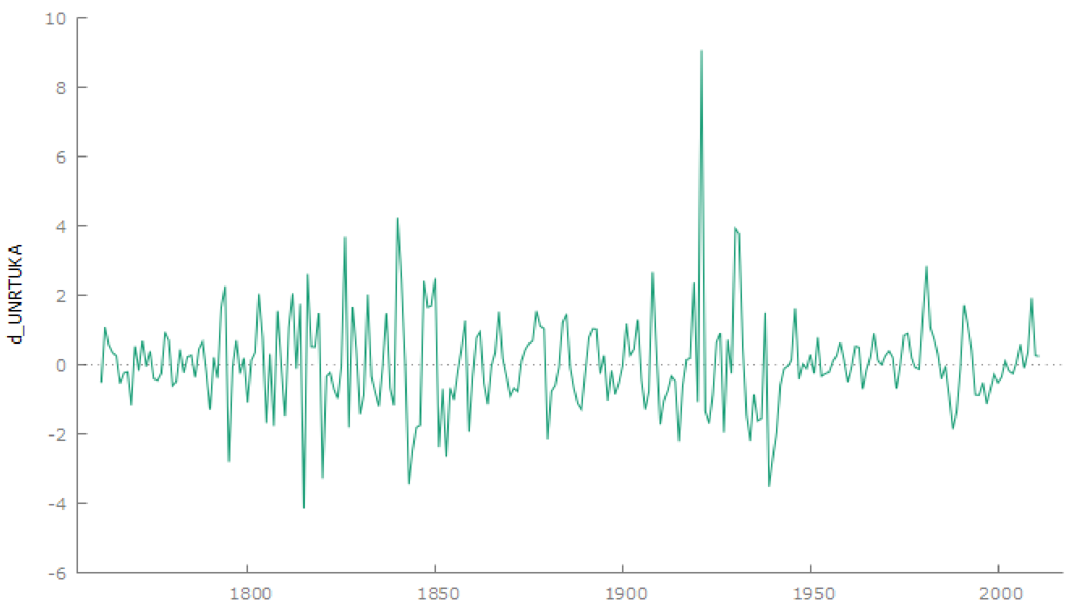

Figure 4. Time Series Plot of the First Difference of the UNRTUKA Dataset

After the first difference was taken, the ACF and PACF plot was also checked for attributes of a white noise.

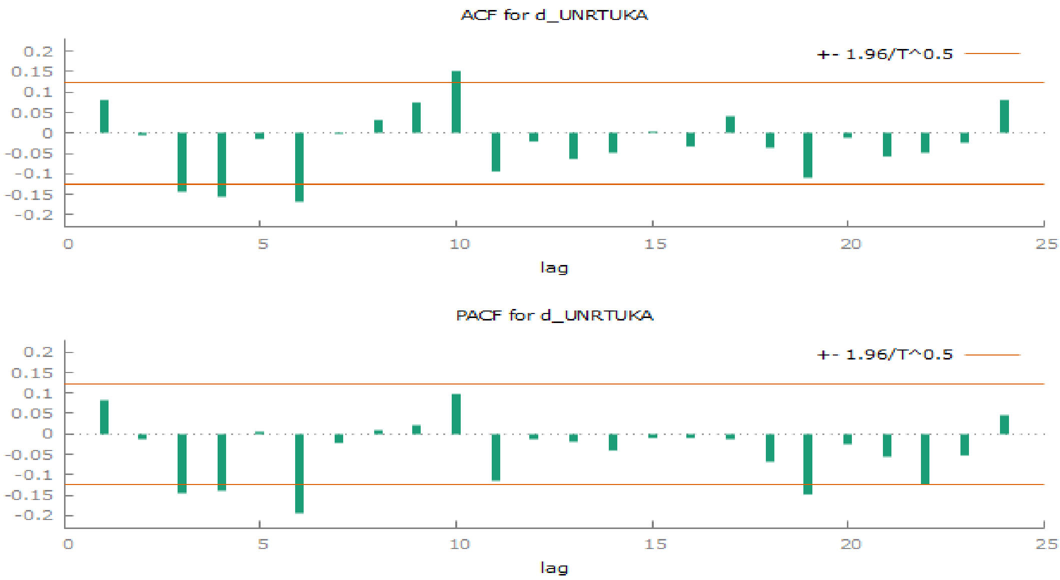

Figure 5. ACF and PACF Plot for the UNRTUKA Annual Data

Unlike the other two dataset, the UNRTUKA ACF and PACF Plot has it 3rd and 4th lag outside the significance band, which is sufficient enough to assume it is not generated by a white noise.

B. Determining the order of the ARIMA Model

For this section of the project, the UNRTUKA dataset was split in to training and testing set. The training period was set from 1760 to 2011, and the testing period was set from 2012 to 2016. After this was done, the AIC, BIC and HQC were used to determine the most likely autoregressive (AR) and moving average (MA) order for the chosen dataset. A maximum value of 3 was specified for both the AR and MA orders in the armax function in Gretl.

ARMAX_Criteria - Notepad

File Edit Format View Help

```
? armax(3, 3, d_UNRTUKA, null, 1, 1, 0, 1, 0)
================================================
 Information Criteria of ARMAX(p,q) for d_UNRTUKA
------------------------------------------------
 p, q          AIC           BIC           HQC
------------------------------------------------
 0, 0      875.4049      882.4558      878.2424
 0, 1      875.7164      886.2928      879.9726
 0, 2      877.6295      891.7313      883.3044
 0, 3      873.2970      890.9243      880.3907
 1, 0      875.7105      886.2869      879.9667
 1, 1      877.7018      891.8036      883.3767
 1, 2      863.4096      881.0369      870.5033
 1, 3      861.8920*     883.0447      870.4044
 2, 0      877.6721      891.7739      883.3470
 2, 1      862.4439      880.0712*     869.5376*
 2, 2      863.8881      885.0408      872.4005
 2, 3      863.8121      888.4903      873.7432
 3, 0      874.4304      892.0576      881.5240
 3, 1      863.2503      884.4030      871.7627
 3, 2      864.9314      889.6096      874.8625
 3, 3      862.7112      890.9148      874.0610
================================================
* indicates best models.
'9999.9999' suggests failures to estimate the models.
```

Figure 6. Order Selection using AIC, BIC, and HQC

The result above shows that the AIC selected ARIMA (1, 1, 3) as the optimal model, while the BIC and HQC selected ARIMA (2, 1, 1) as the optimal model. In some cases it could be logical to go for the model chosen by two out of three criteria used, however in this case both models were tested, and the best was selected thereof.

C. Estimating the coefficients of the two ARIMA Models, and Model Evaluation

To estimate the coefficients of the ARIMA model, the ARIMA function in Gretl was used. After this was done, different methods were used to check how well each model performed in representing the dataset.

- ARIMA (1, 1, 3)

Using the ARIMA function, the following coefficients were obtained for the first ARIMA model.

```
ARIMA_113 - Notepad
File  Edit  Format  View  Help
ARIMA_113:
ARIMA, using observations 1761-2011 (T = 251)
Estimated using AS 197 (exact ML)
Dependent variable: (1-L) UNRTUKA
Standard errors based on Hessian

             coefficient   std. error      z       p-value
  ---------------------------------------------------------
  const       0.00370842   0.00711555    0.5212   0.6022
  phi_1       0.791580     0.0549950    14.39     5.67e-047 ***
  theta_1    -0.785112     0.0809048    -9.704    2.90e-022 ***
  theta_2    -0.0553954    0.0792004    -0.6994   0.4843
  theta_3    -0.159493     0.0828617    -1.925    0.0543     *

Mean dependent var    0.017849   S.D. dependent var    1.375647
Mean of innovations   0.021204   S.D. of innovations   1.306591
R-squared             0.771035   Adjusted R-squared    0.768254
Log-likelihood       -424.9460   Akaike criterion      861.8920
Schwarz criterion     883.0447   Hannan-Quinn          870.4044

                        Real  Imaginary    Modulus  Frequency
  -----------------------------------------------------------
  AR
    Root  1           1.2633     0.0000     1.2633     0.0000
  MA
    Root  1           1.0000     0.0000     1.0000     0.0000
    Root  2          -0.6737    -2.4116     2.5040    -0.2934
    Root  3          -0.6737     2.4116     2.5040     0.2934
  -----------------------------------------------------------

Test for normality of residual -
  Null hypothesis: error is normally distributed
  Test statistic: Chi-square(2) = 65.067
  with p-value = 7.42833e-15
```

Figure 7. ARIMA (1, 1, 3) Coefficients

The complete estimated ARIMA (1, 1, 3) model is;

$$X_t = 0.7916X_{t-1} + 0.0037 + U_t - 0.7851U_{t-1} - 0.0554U_{t-2} - 0.1595U_{t-3}$$

The ACF and PACF plot of the residue from this model was checked to confirm if it is a white noise. This is so as to conclude that no extra information about the dataset is contained in its residue.

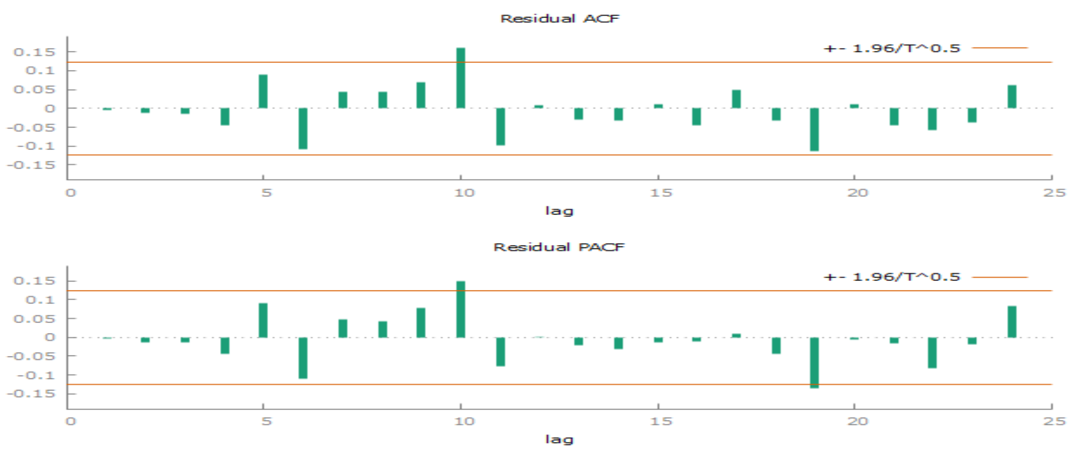

Figure 8. ACF and PACF plot for ARIMA (1, 1, 3) Residue

From the plot above, it can be observed that only lag 10 is outside our significance band, however this lag is a relatively high lag, so there's not much confidence in it. With this observation, it is safe to conclude that residue is simply a white noise hence there is no extra information about the dataset to be gained from it.

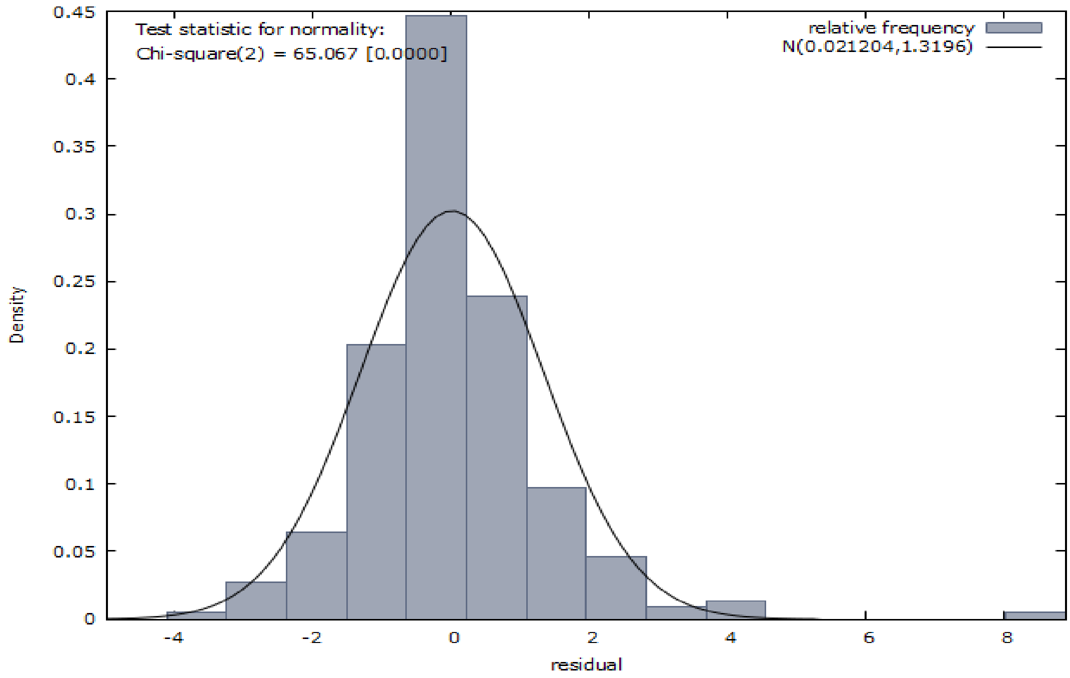

Figure 9. Normality Plot for ARIMA (1, 1, 3) Residue

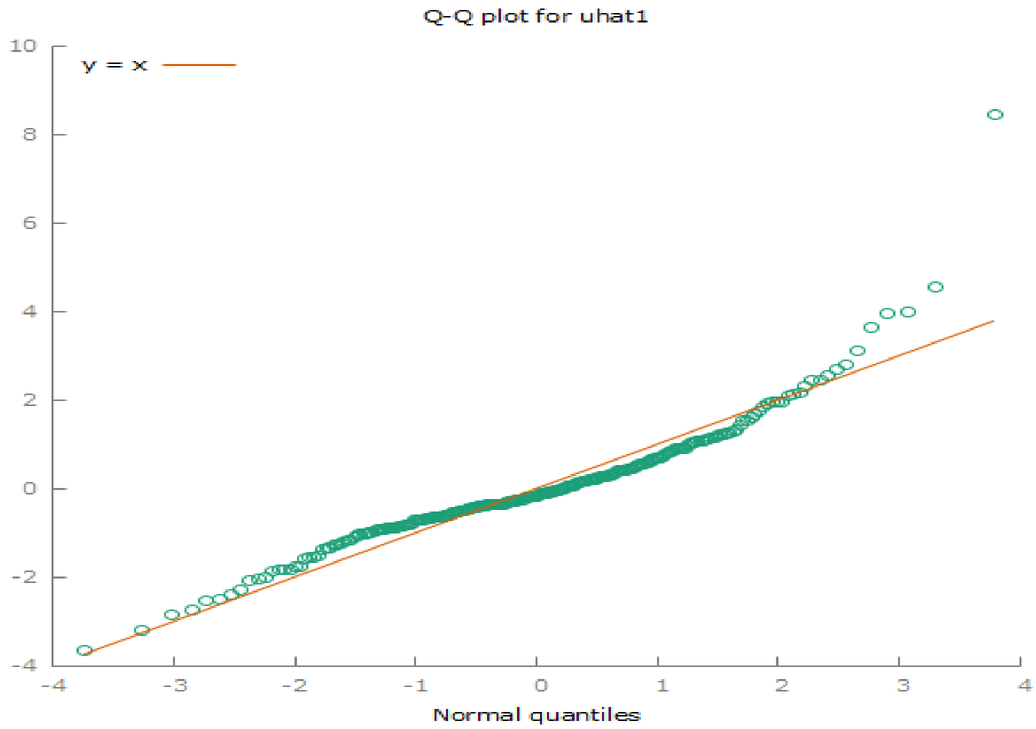

Figure 10. Q-Q Plot for ARIMA (1, 1, 3) Residue

Further analysis of the residue from the ARIMA (1, 1, 3) using the Normality Plot and the Q-Q Plot shows that the residue is a Gaussian white noise, so it is safe to assume that the residue is simple a result of non-avoidable stochasticity.

Furthermore, the ARIMA (1, 1, 3) is used to estimate that data in the testing set, so as to obtain the performance of the model on unseen data.

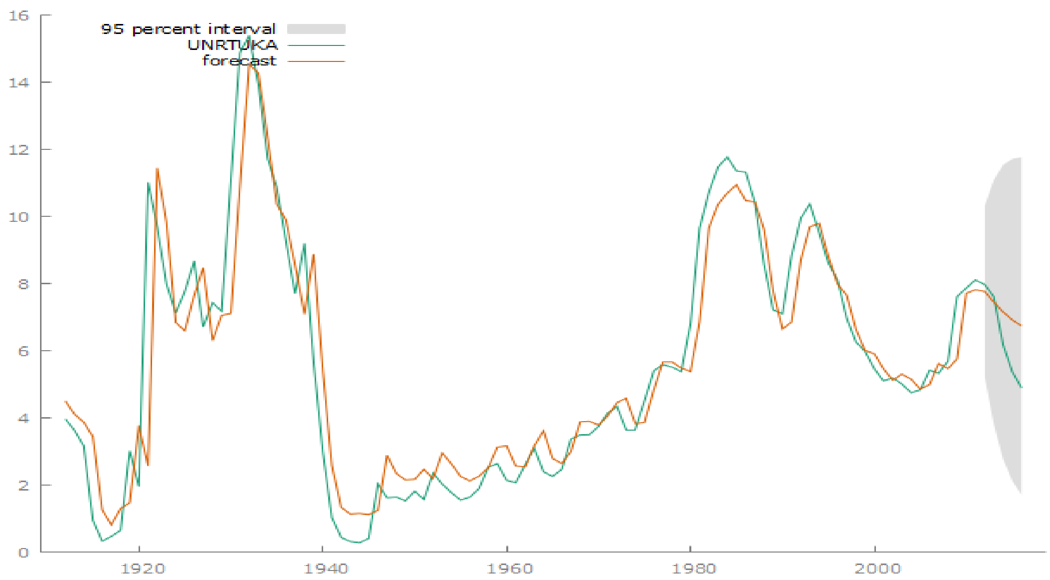

Figure 11. Forecast Plot on Test Data using ARIMA (1, 1, 3)

ARIMA_113_Forecast - Notepad

File Edit Format View Help

```
1997         6.97         7.65
1998         6.26         6.63
1999         5.98         6.01
2000         5.46         5.91
2001         5.10         5.46
2002         5.19         5.13
2003         5.01         5.30
2004         4.75         5.15
2005         4.83         4.87
2006         5.42         4.99
2007         5.33         5.62
2008         5.69         5.47
2009         7.61         5.75
2010         7.87         7.71
2011         8.11         7.82
2012         7.97         7.76         1.307         5.20 -     10.33
2013         7.61         7.45         1.854         3.82 -     11.08
2014         6.18         7.15         2.235         2.77 -     11.54
2015         5.38         6.92         2.444         2.13 -     11.71
2016         4.90         6.74         2.567         1.71 -     11.77

  Forecast evaluation statistics using 5 observations

  Mean Error                           -0.79793
  Root Mean Squared Error               1.1641
  Mean Absolute Error                   0.94416
  Mean Percentage Error               -15.454
  Mean Absolute Percentage Error       17.327
  Theil's U2                            1.7763
  Bias proportion, UM                   0.46984
  Regression proportion, UR             0.49385
  Disturbance proportion, UD            0.036309
```

Figure 12. RMSE value of the Forecast using ARIMA (1, 1, 3)

The RMSE value of the ARIMA (1, 1, 3) model is 1.1641.

- <u>ARIMA (2, 1, 1)</u>

Using the ARIMA function, the following coefficients were obtained for the second ARIMA model.

```
ARIMA_211 - Notepad
File  Edit  Format  View  Help
Function evaluations: 106
Evaluations of gradient: 23

Model 2: ARIMA, using observations 1761-2011 (T = 251)
Estimated using AS 197 (exact ML)
Dependent variable: (1-L) UNRTUKA
Standard errors based on Hessian

             coefficient   std. error      z        p-value
  ---------------------------------------------------------
  const        0.00400095   0.00760329     0.5262   0.5987
  phi_1        1.01015      0.0622491     16.23     3.22e-059 ***
  phi_2       -0.151574     0.0623561     -2.431    0.0151    **
  theta_1     -1.00000      0.0114177    -87.58     0.0000    ***

Mean dependent var   0.017849   S.D. dependent var   1.375647
Mean of innovations  0.019824   S.D. of innovations  1.313692
R-squared            0.768546   Adjusted R-squared   0.766680
Log-likelihood      -426.2220   Akaike criterion     862.4439
Schwarz criterion    880.0712   Hannan-Quinn         869.5376

                        Real  Imaginary    Modulus  Frequency
  -----------------------------------------------------------
  AR
    Root  1           1.2094     0.0000     1.2094     0.0000
    Root  2           5.4549     0.0000     5.4549     0.0000
  MA
    Root  1           1.0000     0.0000     1.0000     0.0000
  -----------------------------------------------------------
```

Figure 13. ARIMA (2, 1, 1) Coefficients

The complete estimated ARIMA (2, 1, 1) model is;

$$X_t = 1.0102X_{t-1} - 0.1516X_{t-2} + 0.004 - U_{t-1}$$

In a similar manner, the ACF and PACF plot of the model's residue was plotted to confirm if the residue can be assumed to be a white noise.

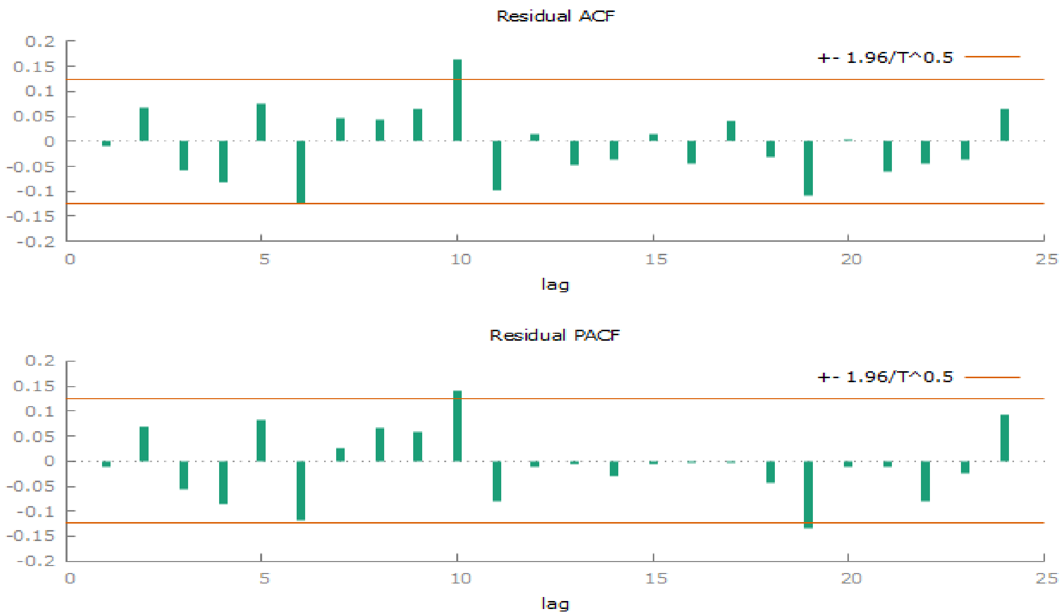

Figure 14. ACF and PACF plot for ARIMA (2, 1, 1) Residue

From the plot above, it can also be observed (as in the previous model) that only lag 10 is outside our significance band, however this lag is a relatively high lag, so there's not much confidence in it. With this observation, it is also safe to conclude that residue is simply a white noise hence there is no extra information about the dataset to be gained from it.

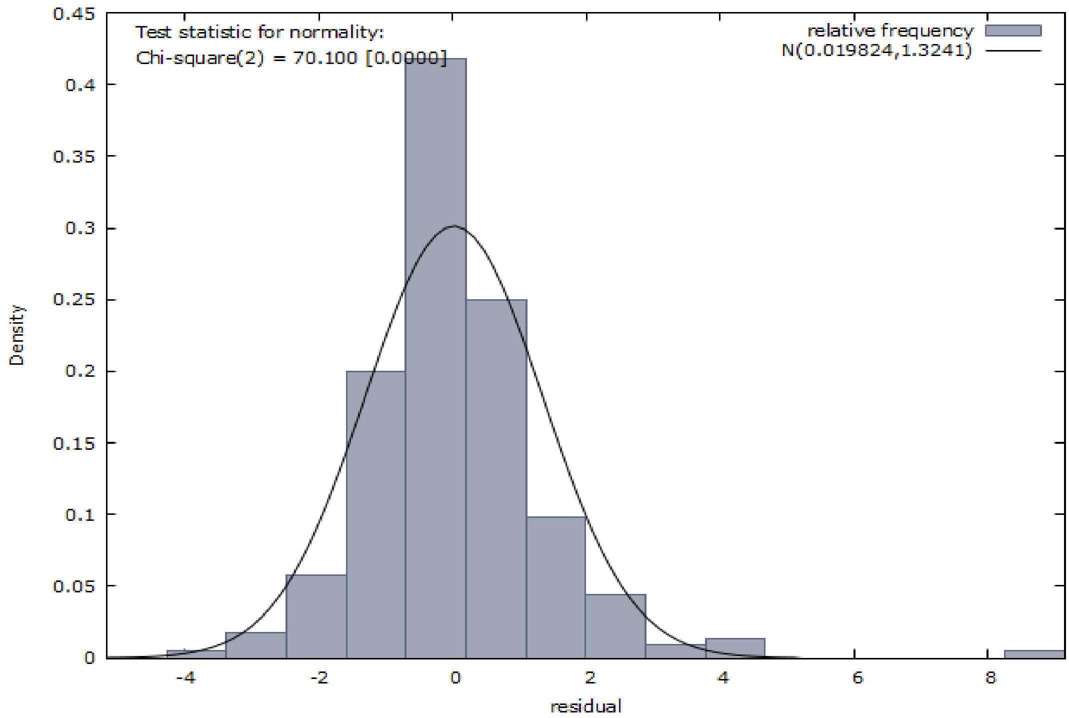

Figure 15. Normality Plot for ARIMA (2, 1, 1) Residue

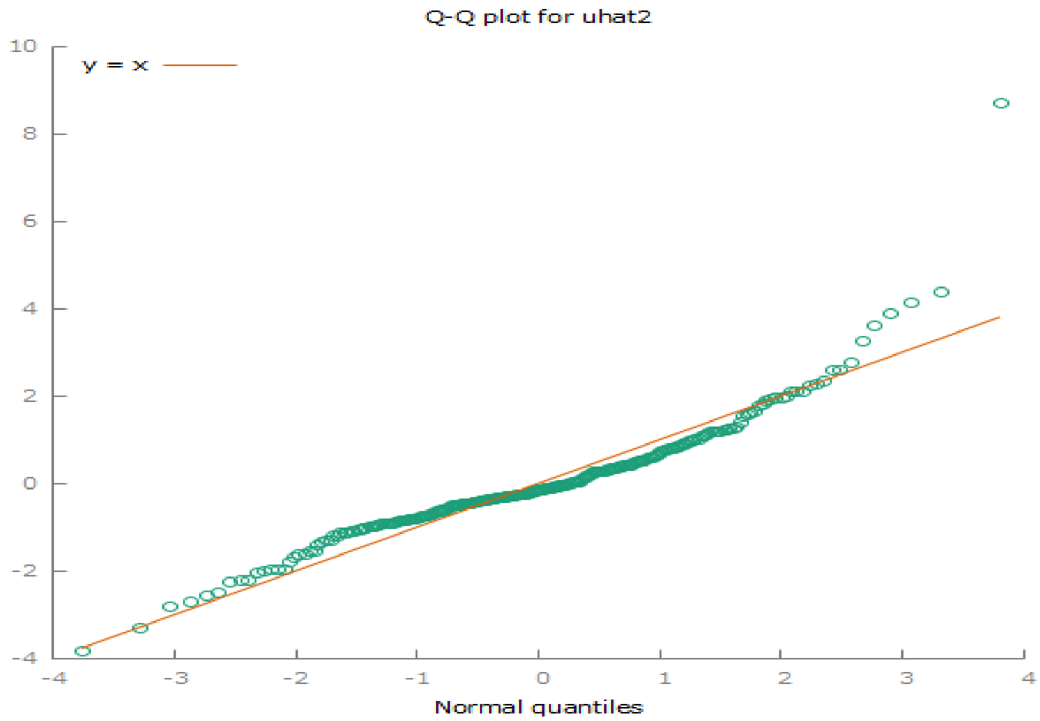

Figure 16. Q-Q Plot for ARIMA (2, 1, 1) Residue

Further analysis of the residue from the ARIMA (2, 1, 1) using the Normality Plot and the Q-Q Plot also shows that the residue is a Gaussian white noise, so it is safe to assume that the residue is simple a result of non-avoidable stochasticity

Using the ARIMA (2, 1, 1) to forecast on the test data, we get the following results.

Figure 17. Forecast Plot on Test Data using ARIMA (2, 1, 1)

```
ARIMA_211_Forecast - Notepad
File  Edit  Format  View  Help
1997          6.97          7.73
1998          6.26          6.66
1999          5.98          6.12
2000          5.46          5.94
2001          5.10          5.46
2002          5.19          5.17
2003          5.01          5.32
2004          4.75          5.12
2005          4.83          4.88
2006          5.42          5.01
2007          5.33          5.59
2008          5.69          5.41
2009          7.61          5.79
2010          7.87          7.68
2011          8.11          7.66
2012          7.97          7.86          1.314          5.28 -     10.43
2013          7.61          7.57          1.867          3.91 -     11.23
2014          6.18          7.31          2.188          3.03 -     11.60
2015          5.38          7.10          2.387          2.43 -     11.78
2016          4.90          6.93          2.513          2.00 -     11.85

  Forecast evaluation statistics using 5 observations

  Mean Error                          -0.94655
  Root Mean Squared Error              1.2947
  Mean Absolute Error                  1.0075
  Mean Percentage Error               -17.966
  Mean Absolute Percentage Error       18.741
  Theil's U2                           1.9746
  Bias proportion, UM                  0.53451
  Regression proportion, UR            0.43308
  Disturbance proportion, UD           0.032409
```

Figure 18. RMSE value of the Forecast using ARIMA (2, 1, 1)

The RMSE value of the ARIMA (2, 1, 1) model is 1.2947.

Comparing the two models, it can be observed that the RMSE of the ARIMA (1, 1, 3) model is lower than the RMSE of the ARIMA (2, 1, 1) model. As stated in the previous subsection of this report, both models were tested instead of just selecting the model with the highest vote from the AIC, BIC and HQC. It can be observed that in this case the model with the lesser vote eventually outperformed the model with the higher vote. This shows the importance of exhausting all available options, even though this isn't applicable in cases when resources are limited.

- <u>Ensemble Prediction and Performance Comparison</u>

To further exhaust all options, forecast of the testing set was carried out using an ensemble of both the ARIMA (1, 1, 3) and the ARIMA (2, 1, 1), and all three predictions were compared to see if the Ensemble model would outperform both models.

jupyter Ensemble_MSE (unsaved changes)

Dataset (Predictions) Importation

```
In [3]:  predictions = pd.read_excel('En.xlsx')

In [4]:  predictions.head()

Out[4]:
```

	Date	Forecast Leg (k)	Actual Value	ARIMA(1,1,3)	ARIMA(2,1,1)	Ensemble Prediction
0	2012	1	7.97	7.76	7.86	7.810
1	2013	2	7.61	7.45	7.57	7.510
2	2014	3	6.18	7.15	7.31	7.230
3	2015	4	5.38	6.92	7.10	7.010
4	2016	5	4.90	6.74	6.93	6.835

MSE Calculation and Comparison

```
In [6]:  # Calculation of Mean Squared Error (MSE)

         print('Mean Squared Error (ARIMA(1,1,3)): ',mean_squared_error(predictions['Actual Value'],predictions['ARIMA(1,1,3)']))
         print('Mean Squared Error (ARIMA(2,1,1)): ',mean_squared_error(predictions['Actual Value'],predictions['ARIMA(2,1,1)']))
         print('Mean Squared Error (Ensemble): ',mean_squared_error(predictions['Actual Value'],predictions['Ensemble Prediction']))

         Mean Squared Error (ARIMA(1,1,3)):  1.35356
         Mean Squared Error (ARIMA(2,1,1)):  1.6739799999999991
         Mean Squared Error (Ensemble):  1.507845
```

Figure 19. MSE Comparison of the ARIMA (1, 1, 3), ARIMA (2, 1, 1) and their Ensemble

The result of this comparison above shows that despite the fact that the ensemble prediction gave better results than the ARIMA (2, 1, 1), it still underperformed with respect to the ARIMA (1, 1, 3) model. As a result of this outcome, the ARIMA (1, 1, 3) model was selected from all considered models in this project.

D. Model Improvement and Comparison

As seen from Figure 7, the intercept in our ARIMA (1, 1, 3) model has no significant impact in our prediction. Coefficients without significant impact in the overall prediction can eventually act as noise or even increase cost of computation, which is the purpose of this subsection.

In this subsection the coefficients of the ARIMA (1, 1, 3) was re-estimated without adding an intercept, and the results were compared to its former version which had an intercept.

```
ARIMA_113_NC - Notepad
File  Edit  Format  View  Help
Function evaluations: 88
Evaluations of gradient: 23

Model 4: ARIMA, using observations 1761-2011 (T = 251)
Estimated using AS 197 (exact ML)
Dependent variable: (1-L) UNRTUKA
Standard errors based on Hessian

             coefficient   std. error      z        p-value
  ---------------------------------------------------------
  phi_1        0.792631    0.0549104     14.43     3.12e-047 ***
  theta_1     -0.785028    0.0809095     -9.703    2.94e-022 ***
  theta_2     -0.0553467   0.0791928     -0.6989   0.4846
  theta_3     -0.159625    0.0828334     -1.927    0.0540    *

Mean dependent var    0.017849   S.D. dependent var   1.375647
Mean of innovations   0.058917   S.D. of innovations  1.307310
R-squared             0.771192   Adjusted R-squared   0.768413
Log-likelihood       -425.0818   Akaike criterion     860.1636
Schwarz criterion     877.7908   Hannan-Quinn         867.2572

                       Real   Imaginary    Modulus   Frequency
  -----------------------------------------------------------
  AR
    Root  1           1.2616     0.0000     1.2616     0.0000
  MA
    Root  1           1.0000     0.0000     1.0000     0.0000
    Root  2          -0.6734    -2.4107     2.5029    -0.2934
    Root  3          -0.6734     2.4107     2.5029     0.2934
  -----------------------------------------------------------
```

Figure 20. ARIMA (1, 1, 3) Coefficients without the intercept

The complete estimated ARIMA (1, 1, 3) model without the intercept is;

$$X_t = 0.7926X_{t-1} + U_t - 0.7850U_{t-1} - 0.0553U_{t-2} - 0.1596U_{t-3}$$

The result of the comparison of the new version of the ARIMA (1, 1, 3) and its old version is as follows;

```
ARIMA_113_NC_Forecast - Notepad
File  Edit  Format  View  Help
1997          6.97          7.58
1998          6.26          6.56
1999          5.98          5.94
2000          5.46          5.84
2001          5.10          5.39
2002          5.19          5.06
2003          5.01          5.23
2004          4.75          5.08
2005          4.83          4.79
2006          5.42          4.92
2007          5.33          5.55
2008          5.69          5.40
2009          7.61          5.68
2010          7.87          7.64
2011          8.11          7.74
2012          7.97          7.69          1.307          5.13 -     10.25
2013          7.61          7.30          1.856          3.67 -     10.94
2014          6.18          6.94          2.239          2.55 -     11.33
2015          5.38          6.65          2.449          1.85 -     11.45
2016          4.90          6.42          2.573          1.38 -     11.46

  Forecast evaluation statistics using 5 observations

  Mean Error                         -0.59321
  Root Mean Squared Error             0.9662
  Mean Absolute Error                 0.82619
  Mean Percentage Error             -11.882
  Mean Absolute Percentage Error     14.877
  Theil's U2                          1.4661
  Bias proportion, UM                 0.37695
  Regression proportion, UR           0.571
  Disturbance proportion, UD          0.052048
```

Figure 21. RMSE value of the Forecast using ARIMA (1, 1, 3) without intercept

Comparing the RMSE in figure 21 (0.9662) to that of figure 12 (1.164), it can be observed that there was slight improvement in the model performance.

3. Final Result and Comparison with Real Outcome

A. Retraining the Final Chosen Model with the entire dataset for future Forecasting

After all experiment and improvements have been carried out on the chosen model (ARIMA 1, 1, 3), the model was then retrained on the entire dataset and used to predict 5 years into the future (2021) from the end of the duration covered in the dataset (2016), and the result is seen below;

```
ARIMA_113_NC_FullRange - Notepad
File  Edit  Format  View  Help
Function evaluations: 82
Evaluations of gradient: 21

Model 5: ARIMA, using observations 1761-2016 (T = 256)
Estimated using AS 197 (exact ML)
Dependent variable: (1-L) UNRTUKA
Standard errors based on Hessian

             coefficient   std. error      z        p-value
  ---------------------------------------------------------
  phi_1        0.790212    0.0541077    14.60     2.63e-048 ***
  theta_1     -0.781461    0.0796171    -9.815    9.68e-023 ***
  theta_2     -0.0565952   0.0783540    -0.7223   0.4701
  theta_3     -0.161944    0.0812144    -1.994    0.0461     **

Mean dependent var    0.004961   S.D. dependent var    1.366603
Mean of innovations   0.052388   S.D. of innovations   1.296615
R-squared             0.771680   Adjusted R-squared    0.768961
Log-likelihood       -431.4286   Akaike criterion      872.8572
Schwarz criterion     890.5831   Hannan-Quinn          879.9865

                        Real   Imaginary    Modulus  Frequency
  -----------------------------------------------------------
  AR
    Root  1            1.2655     0.0000     1.2655     0.0000
  MA
    Root  1            1.0000     0.0000     1.0000     0.0000
    Root  2           -0.6747    -2.3916     2.4849    -0.2938
    Root  3           -0.6747     2.3916     2.4849     0.2938
  -----------------------------------------------------------
```

Figure 22. ARIMA (1, 1, 3) Coefficients trained on the dataset's full range

The complete estimated ARIMA (1, 1, 3) model on the full-ranged dataset is;

$$X_t = 0.7902X_{t-1} + U_t - 0.7815U_{t-1} - 0.0566U_{t-2} - 0.1619U_{t-3}$$

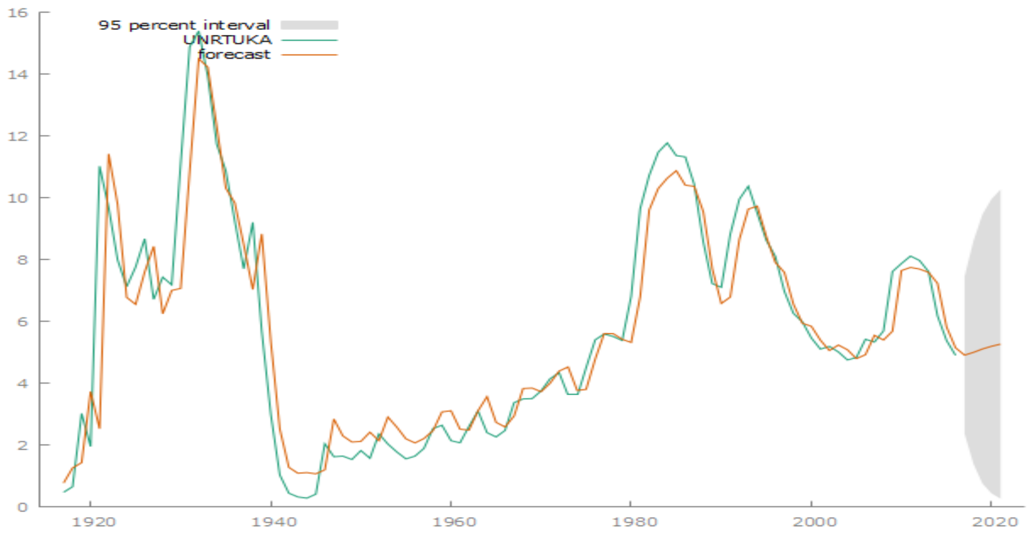

Figure 23. Future Forecast using the Final ARIMA (1, 1, 3) Model

After the final forecasting was done, further web search was carried out in order to obtain the actual values for the real world outcome for the UNRTUKA record, and fortunately there was a more recently updated dataset on the same index.

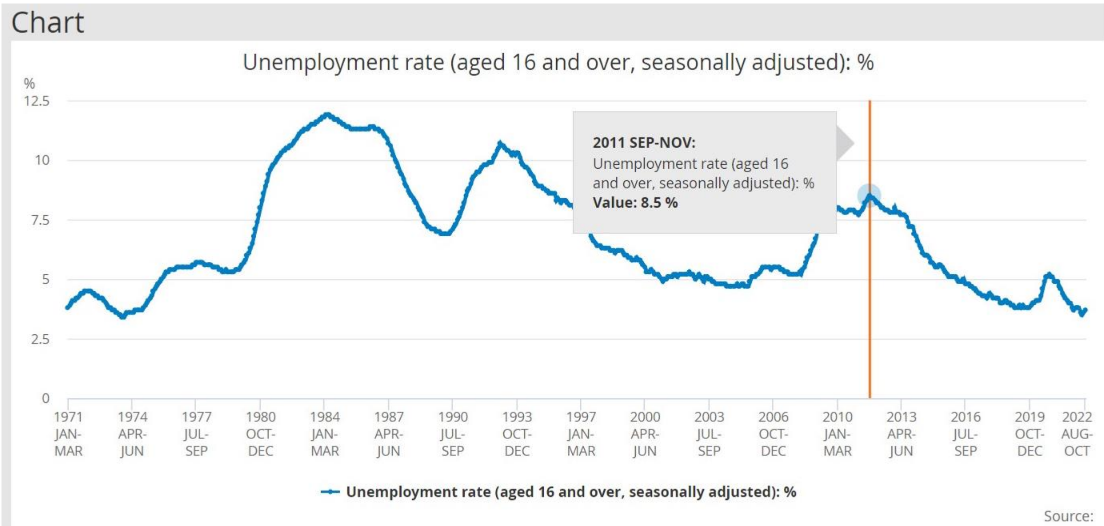

Figure 24. Real Word outcome of the UNRTUKA Index trend

Comparing the final model's future forecast with the real world outcome, it can be observed that the trained model did a pretty decent job in predicting the outcome of events.

4. Conclusion

In this project, the Unemployment Rate of the United Kingdom (UNRTUKA) was extracted from the web, transformed, analysed and modelled using the Gretl software. The three top ARIMA models (which was made up of the two top models according to the AIC, BIC and HQC, and their ensemble) were tested and compared, and the best was selected. The chosen model was then modified to improve performance, and this modified version was then trained on the entire dataset and was used to forecast the future values of the real world outcome. The result of this forecast was compared to the real world outcome, and was deemed a decent representation of the given time series.

The project was a very good learning experience and did a good job of bridging the gap between theory and practice. It was also a good way to encounter some of the hurdle real world scenarios pose that are not readily seen in classroom examples.

5. Reference

Unemployment Rate of the United Kingdom (UNRTUKA) used in the Project for modelling.

- https://fred.stlouisfed.org/series/UNRTUKA

Unemployment Rate of the United Kingdom (UNRTUKA) used for real world outcome comparison.

- https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/timeseries/mgsx/lms