Forecasting the Gas demand by Room Type

Team Neo

EBAC, Institute of Systems Science, NUS

[Aastha Arora (A0178188L), Anusuya Manickavasagam (A0163300Y), Chetna Gupta (A0178260A), Kesavan Sridhar (A0163207M), Muni Ranjan (A0163382E), Pradeep Kumar (A0163453H), Zaira Hossain (A0178331E)]

**Abstract**

*This work is part of assignment to apply transfer function and/or cointegration (if applicable) on Consumer Price Index (CPI) and Exchange Rate (USD-INR). We aim to build a statistical model describing the relationship between an output variable Y and one input variable X. We have selected the data released by New Zealand government about Overseas Credit card spending and NZ resident traveler overseas. The result shows that the series are trending upwards and form an equilibrium that tend to return over time. We have evidence that Correlation plot indicates some two-way relation between both series and passing of Engle Grangers test confirms they are* ***cointegrated****. Further, in our error estimation mode, residuals(t) is found to be significant to forecast NZ traveler travelling overseas with Credit card spending taken as Independent variable. Hence, we could conclude that New Zealand traveler travelling overseas are not weakly exogenous. This means that travelling number moves to restore equilibrium when the system is out of bounce.*

# Abbreviation

**NZRT** New Zealand Resident Travelling overseas,

**OCS** Overseas Card Spending

# Executive Summary

Tourism is a very important area of any nation and it’s economy. Whether the people coming to any nation or its residents are going out, both contribute to GDP, employment. Using effective forecasting technique could help government, banks, Airlines and business to make strategies to cater to the needs of growing tourism industry.

In our analysis, we found that the relationship between NZRT and OCS is not one way. Hence transfer function couldn’t be applied. The Prewhiten correlation plot Figure [[06]](#_Figure_6_–) showed the pattern on both sides of lag order and indicated towards **Cointegration**. To confirm, we ran the Engle-Granger test and found evidence that two series are cointegrated. We then developed the two Error correction models (taking each variable as dependent) and found that e(t) is significant when NZRT is a dependent variable (Appendix. Figure [[12]](#_Figure_11_–)).

# Data Source, Partition, tools and Processing

We used the dataset from following source:

|  |  |
| --- | --- |
| Data Type | Source |
| NZ Credit Card Spending data | <https://rbnz.govt.nz/statistics/c13> |
| NZ Resident Travel | <http://www.stats.govt.nz/browse_for_stats/population/Migration/international-travel-and-migration-info-releases.aspx> |
| Time Period | Monthly Data, Jan 2010 – December 2016 |

There are different types of data variables available at above links and we used seasonally adjusted variables under “***Short-term NZ-resident travelers***” and “***Overseas billings on New Zealand issued cards***”.

## 3.1 Data Partition

Following data partition was performed:

|  |  |
| --- | --- |
| Train Set | Jan 2010 – Dec 2015 |
| Forecast Period | **Jan 2016 – Dec 2016** |

## 3.2 Tools

* Excel, Jmp, Gretl

## 3.3 Data Processing

We plotted the time series for both variables as shown in Figure [[1]](#_Figure_1_–).

We can observe above that both the series are trending upwards and don’t see to deviate too much from each other. It is possible that the series follow a long-term equilibrium relationship that they tend to return over time. The model building section below explains in detail about all the test runs.

**A note about the sharp dip in Sep 2011 –** As we can notice there is a sharp dip in September 2011 in terms of New Zealand resident traveling overseas and likewise Overseas Credit card spending. We investigated and [found that](http://www.stats.govt.nz/~/media/Statistics/Browse%20for%20stats/IntTravelAndMigration/HOTPSep11/IntTravelAndMigrationSep11HOTP.pdf) in 2011 School holidays which usually happens in September were moved to October to coincide with Rugby world cup finals. We imputed this data with past 6 months’ average as this is quite an irregular patternand not helpful in building a good model.

**NZRT –** No missing or garbage value was found.New Zealand traveling count was scaled down by dividing by 1000.

**OCS** – No missing or garbage value was found. Further, there were no changes made to this variable and used as it is.

# Model Building

### Transfer Function

We followed the procedure shown below (discussed in classroom) to try Transfer function on our data taking Overseas Card spend as output **Yt**and NZ resident traveling overseas as input variable **Xt**.

Our findings are:

1. Appendix, Figure [[2]](#_Figure_2_–). shows the 1st level differencing order to achieve stationarity. We can also observe the ACF and PACF plot. Since the data is already seasonal adjusted, we don’t see the sharp seasonal pattern as expected. ARIMA (1, 1, 1)(0,0,0) seems to be good fit to achieve the white Noise.
2. We created a model group in Jmp and found that ARIMA(1,1,2) is best performing model. Residual analysis showed White Noise, Figure[[3],[4],[5]](#_Figure_5_–).
3. Next, we Prewhitened the Input series Xt as shown in Figure[[6]](#_Appendix), and observed, it didn’t find one way relation from Xt -> Yt. However, we observed some signs of Cointegration, and we progressed our analysis of **Cointegration**

### Cointegration

We followed the steps as mentioned in the [tutorial](https://www.youtube.com/watch?v=bJgx3JLb7fI):

1. We ran the Engle-Granger test and it confirmed that two series are cointegrated.
   1. Determine ‘d’ in I(d) for the NZRT
   2. Determine ‘d’ in I(d) for the OCS
   3. Estimate cointegration regression: NZRTt =
   4. Calculate the residual ‘d’ in I(d)
      1. H0: Unit Root(i.e. not cointegrated)
      2. H1: No Unit Root(i.e., cointegrated)
2. Build the Error Correction Model
3. Model Checking
4. We figured out iteratively that NZRT as dependent variable shows exogenous behavior as compared to when OCS is taken as dependent variable.

# Model Performance

Tables below (Figure [[12]](#_Figure_11_–)[[13]](#_Figure_12_–)) shows the Model Performance parameter taking NZRT as the dependent variable for our **Error Estimation Model** created in Cointegration.

*OLS, using observations 2010:03-2015:12 (T = 70) Dependent variable: d\_NZRT*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| variable | coefficient | Std. error | t-ratio | p-value |
| d\_OCS\_1 | −0.0819673 | 0.0835990 | −0.9805 | 0.3304 |
| uhat1\_1 (e) | −0.969555 | 0.197726 | −4.904 | 6.29e-06 \*\*\* |
| d\_NZRT\_1 | 0.0794030 | 0.143081 | 0.5550 | 0.5808 |

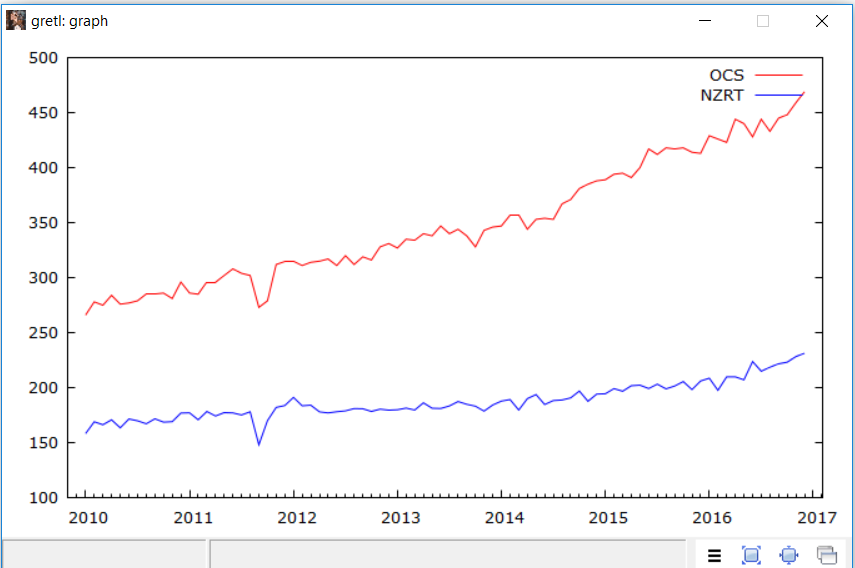
|  |  |  |  |
| --- | --- | --- | --- |
| AIC | Mean Error | RMSE | MAPE |
| 523.2726 | 13.677 | 21.563 | 6.9143 |

### Observations

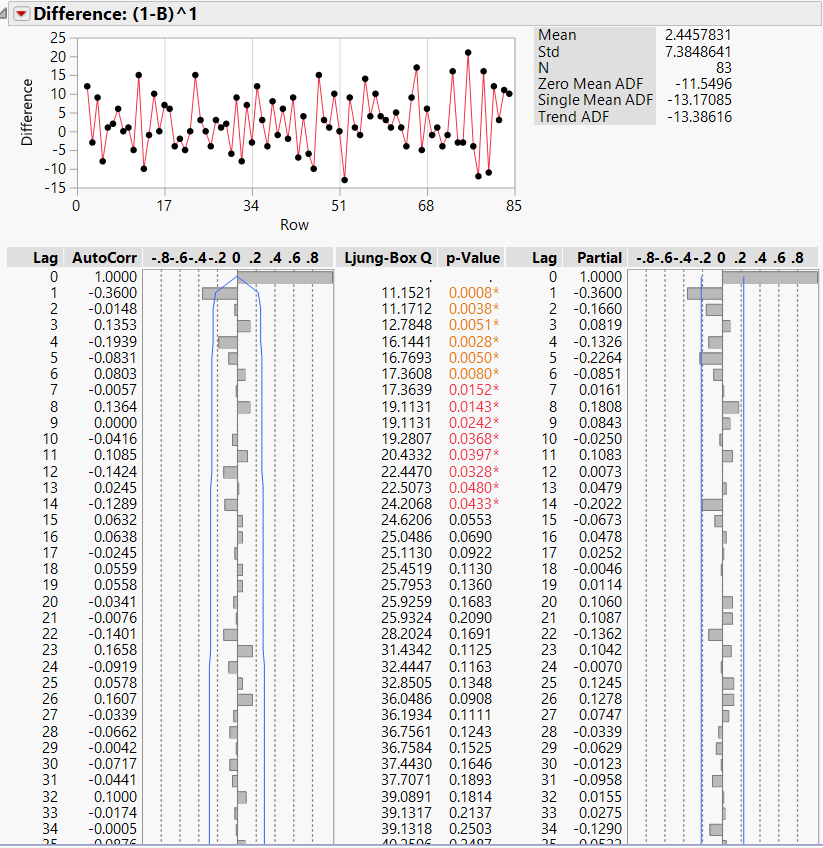
* Overall model performance is **Strong** with low RMSE, MAPE
* Residual (e) is significant in terms of P-value and hence shows that NZRT is not weakly exogenous
* Forecast Vs Actual NZRT plot Figure [[14]](#_Figure_13_–) showed that both series start to diverge after Sep 2016 in test set. This shows that model performance is decreasing over time and hence it should be upgraded continuously to maintain the stability of the model.

# Appendix

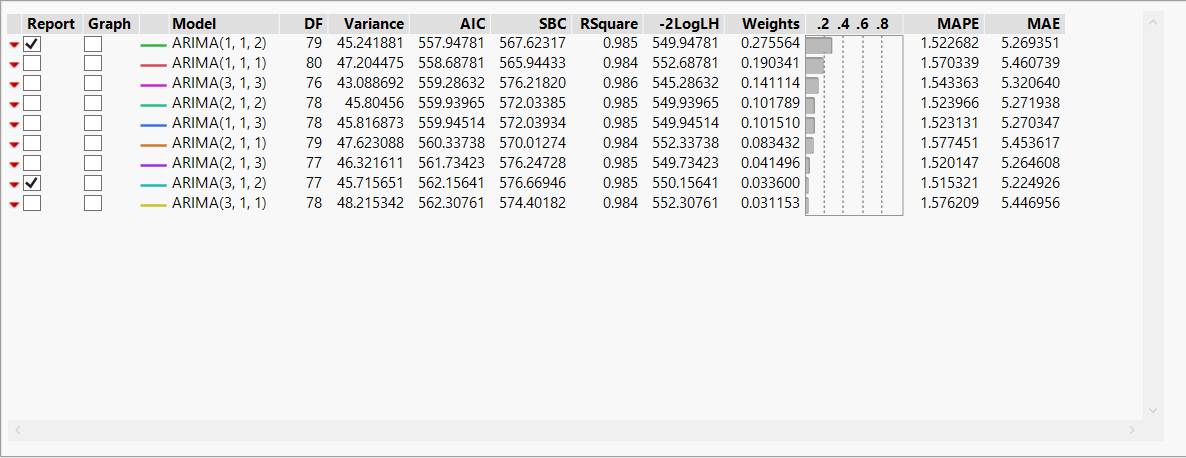
### Figure 1 – Time Series Plot



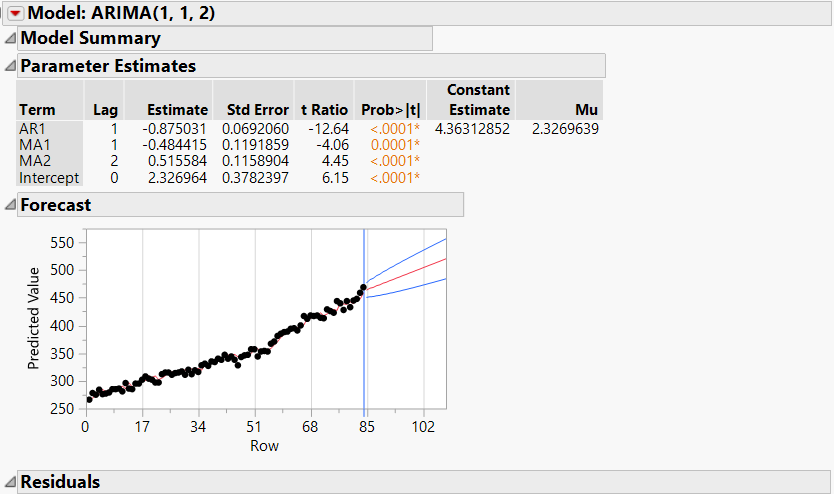
### Figure 2 – Difference (1-B)^1 Xt



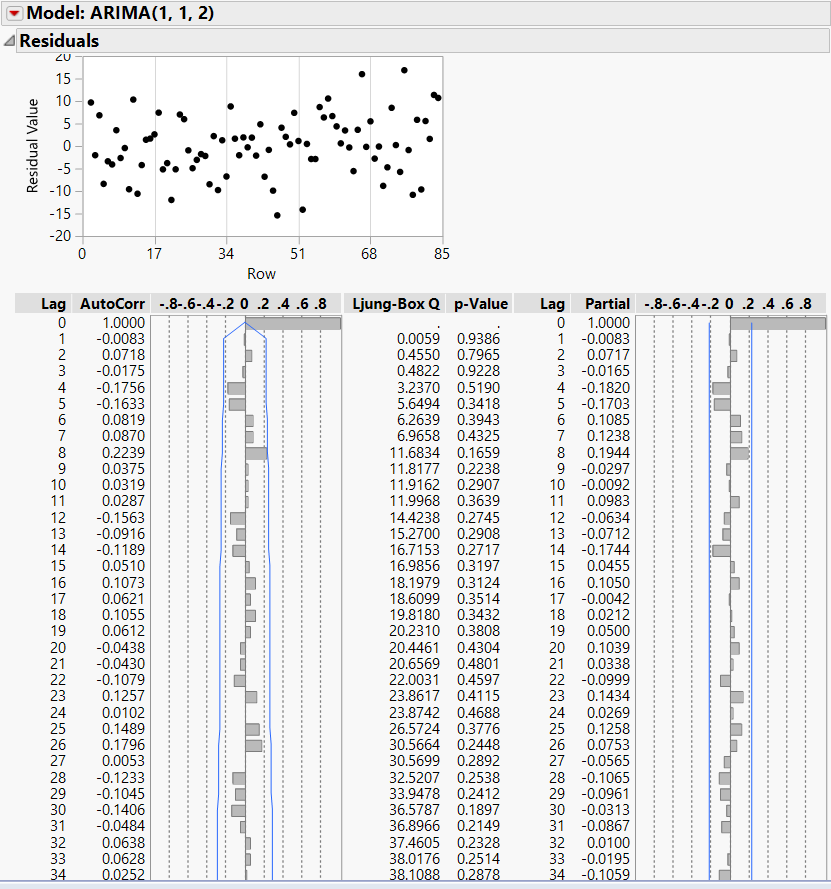
### Figure 3 – Model Comparison



### Figure 4 – Parameter Estimates

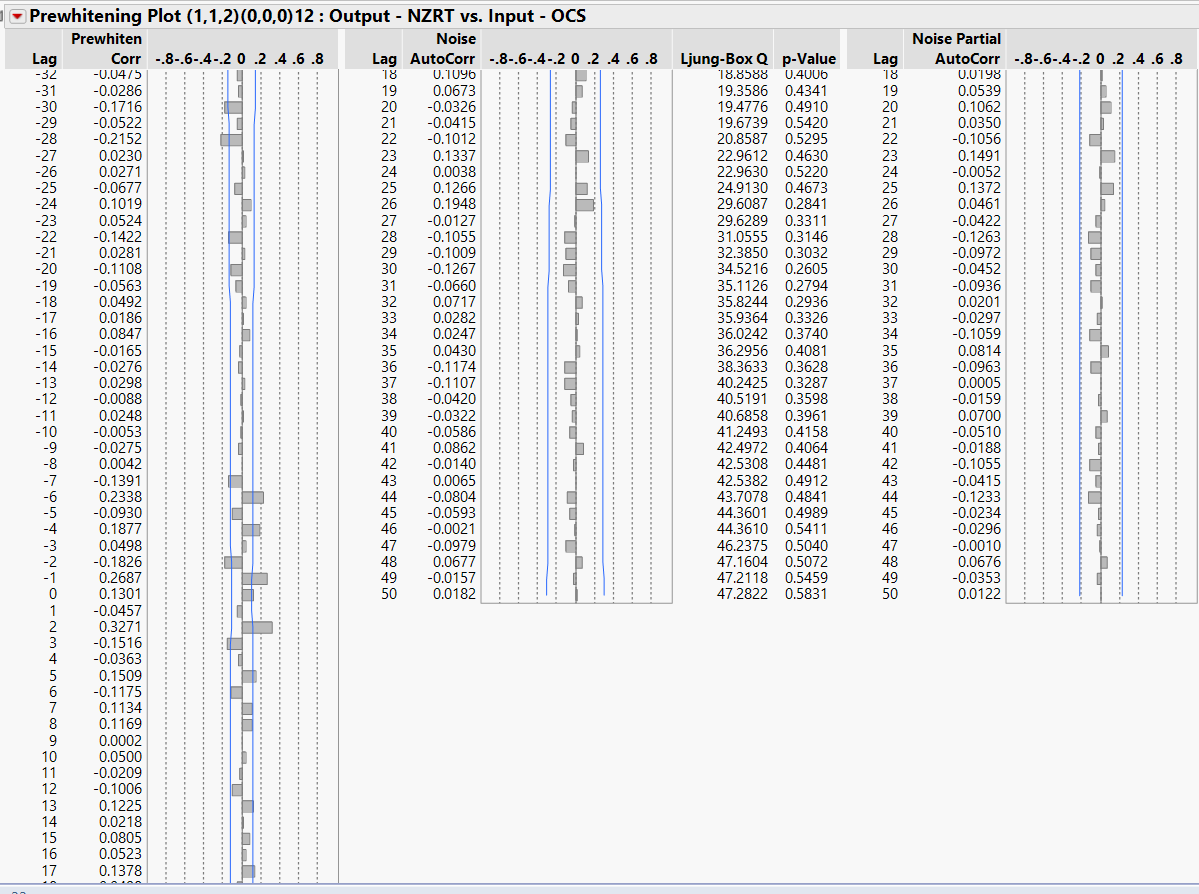


### Figure 5 – Residuals Analysis

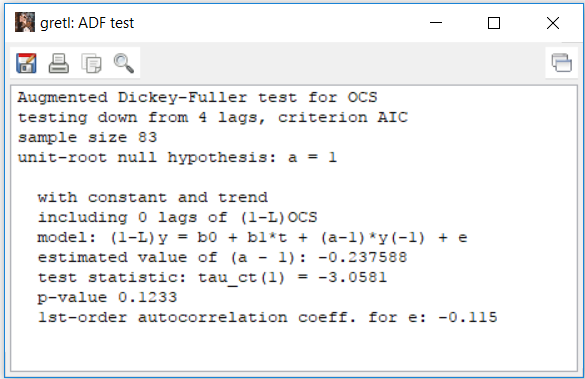




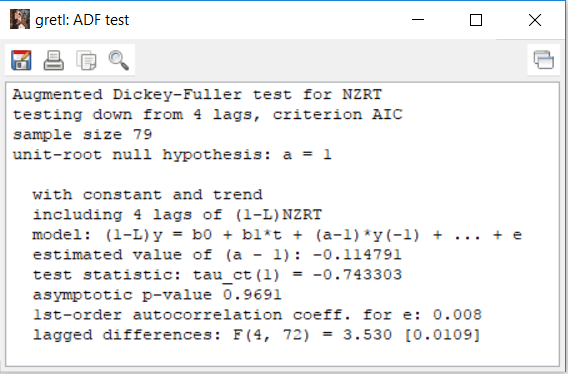
### Figure 6 – Prewhitening Plot NZRT vs OCS



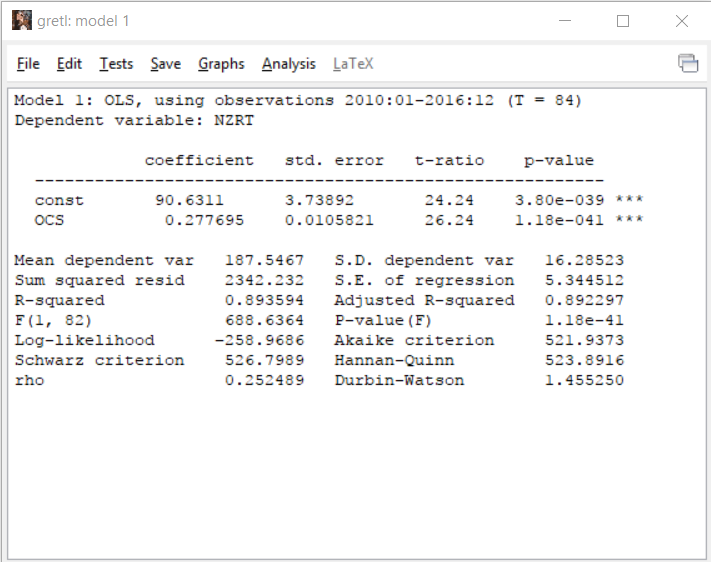
### Figure 7 – Unit Root Test OCS



### Figure 8 – Unit Root Test NZRT

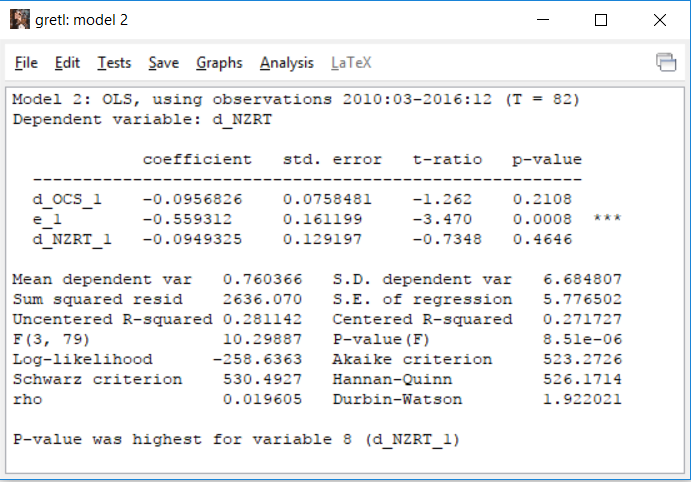


### Figure 9 – Residual Calculation using OLS

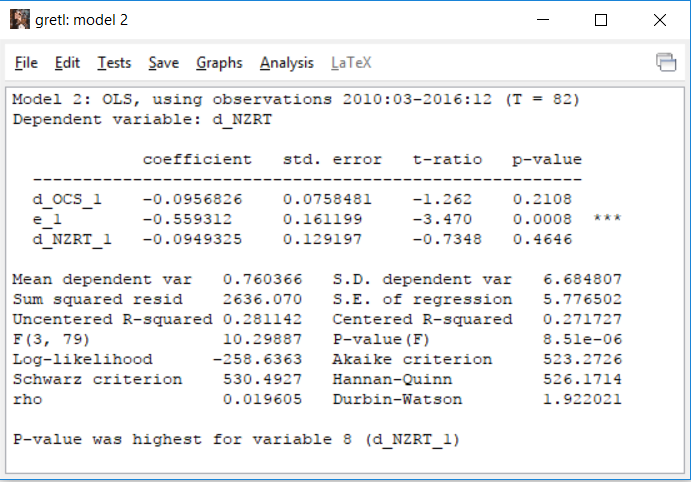


### Figure 10 – Unit Root Residual

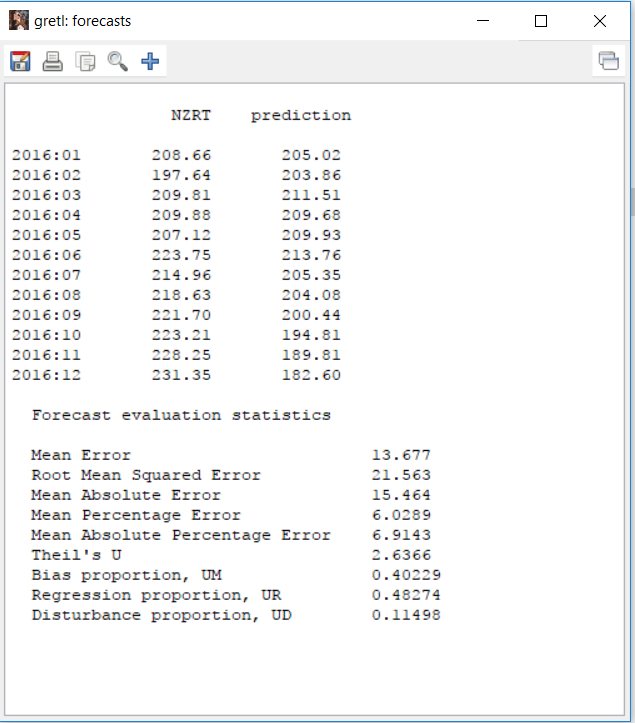
### Figure 12 – Unit Root Residual



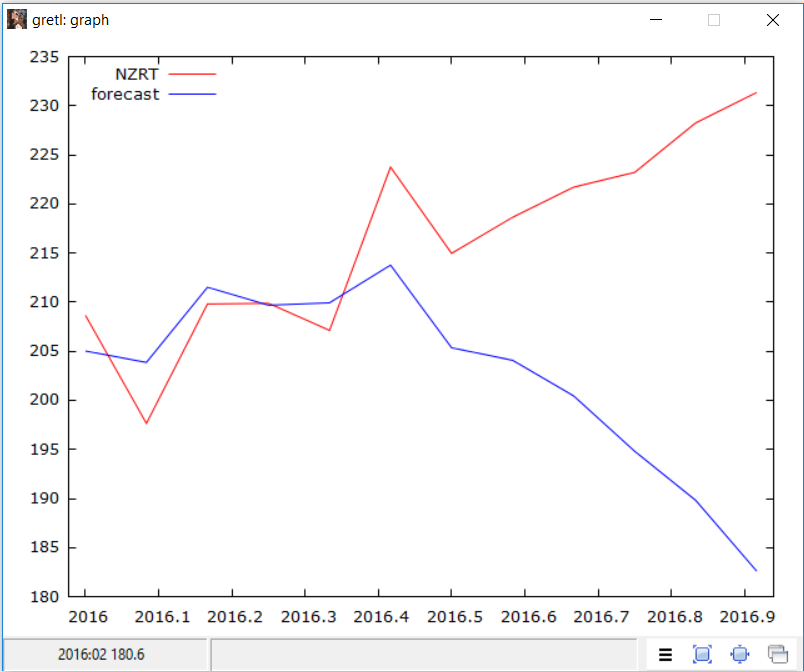
### Figure 12 – Error Estimation Model



### Figure 13 – Test Set Forecast



### Figure 14 – Test Set Forecast vs Actual plot



# References

* IVLE Notes
* Cointegration Tutorial - <https://www.youtube.com/watch?v=bJgx3JLb7fI>