Allocation of short-term loans by the BRICS bank by forecasting industrial production of BRICS nations 2 years ahead

THE KEY QUESTION: How should the BRICS bank allocate its short-term (around 2 years) loan-portfolio to the 5 constituent countries?

DETAILS – The newly formed BRICS bank, now known as the New Development Bank, has recently issued its first loan in April 2016 to Brazil, China, India, and South Africa in the renewables energy space. As an ongoing engagement, the bank has reached out to our team of consultants to help decide what portion of its loan portfolio it should allocate to each country. Since the bank's loans are primarily in the manufacturing, heavy industries, and renewables space, both parties agree that the industrial output data published by the World Bank would be a good proxy to use. The loans will be allocated depending on how much the nations need it. The lower the industrial output, the bigger the need to infrastructural development. There is no dearth of demand for loans.

Forecasting Problem: Forecast 2 years industrial output for BRICS countries by quarter based on data from Jan 2002 to Nov 2016

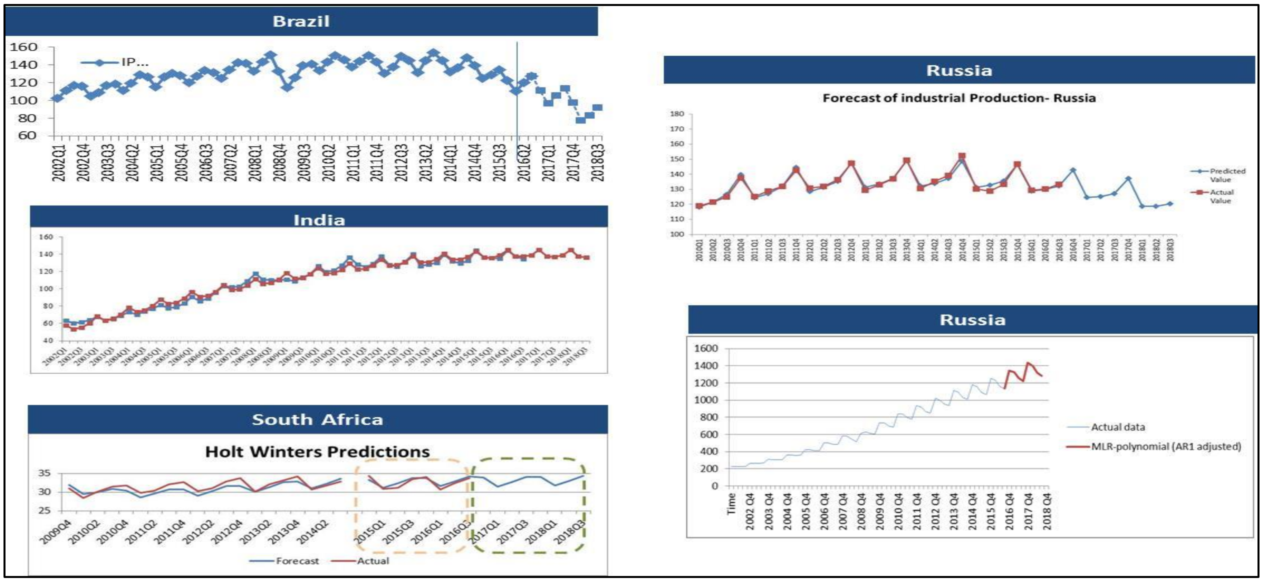
- Since BRICS use quarterly estimations of both financials and future planned investments, we have decided to aggregate the data into quarterly figures to help the Bank

RECOMMENDATION: BRICS bank should allocate a major part of the short term loans equally among Brazil and Russia because of 2 reasons:

- Brazil and Russia are the only 2 countries among all where industrial productions are forecasted to fall in the next 2 years. This necessitates fresh investment to give a boost.
- Both the countries have an industrial production of around $ 10 bn warranting an almost equal split

Also we would be better off by over-forecasting, because that way BRICS bank would truly be able to identify where the loan requirements are the most. And accordingly would be able to disburse the loan to most needy country.

These are the final forecasts for all 5 countries:



The data and the analysis method

<u>Source and aggregation:</u> The data consists of 5 time series sets – one each for Brazil, Russia, India, China, and South Africa - sourced from World Bank. Please refer Exhibit 1. World Bank provides industrial output data on a monthly basis in '00 of millions. However, since most companies in the BRICS use quarterly estimations of both financials and future planned investments, we have decided to aggregate the data into quarterly figures to help the Bank estimate quarterly loan forecasts. We have used a 2 year timeline for forecasting since heavy infrastructural projects may require significant investments and the bank can have sufficient time to generate such as a large asset base. Here is the summary of the data sets:

Countries	Trend	Seasonality	Comments
Brazil	Quadratic, decreasing	Annual, additive	Recession dip; dip after the Petrobras scandal
Russia	Quadratic, decreasing	Annual, additive	Recession dip; recent dip due to weak economy
India	Quadratic, increasing	Annual, additive	Weak recession effect; flattening growth
China	Quadratic, increasing	Annual, multiplicative	Weak recession effect; exponential growth
South Africa	Polynomial Degree 3, stable	Annual, slight multiplicative	Strong recession level change; stabilizing economy

Following are the details of the analysis for each of the series:

I. BRAZIL

1. **Data clean-up:** To avoid issues in prediction, the dip because of the 2008 recession was removed for a smoother trend. However the Petrobras scandal of 2014 has caused a steady downfall in the economy. Because the data is recent, the all the models over-predict as shown in *Exhibit 2*. Hence only data from Q1 of 2013 is used for the final model building.
2. **Model selection:** As shown in *Exhibit 2*, only Multiple Linear Regression overcomes the problem of over-prediction. An MLR model with a quadratic trend and additive seasonality (captured by 3 quarterly dummies) is used. However as shown in *Exhibit 2.2*, the residuals still exhibit a pattern. Hence a Holt's winter model is used to predict the residuals and adjusted in the MLR forecasts.
3. **Confidence in forecasts:** Inferences from the validation plots and the recentness of Petrobras scandal (2014) suggest that we cannot be highly confident about the future trend. Hence we express a **medium** confidence in the forecasts.
4. **Other comments:** Because the Petrobras scandal is very recent, we suggest a re-evaluation of the model every 2 quarters to take into account the recent IP data. Since the models have a tendency to over-forecast than under-forecast, we are on the safe side – since under-forecasts will have propensity to provide loans.

II. RUSSIA

1. **Data Clean-up:** In case of Russia, data changes in the post-recession period as compare to pre-recession. The seasonality is a lot more pronounced in the post-recession data (refer exhibit 3.1). Therefore for Russia we are using data from 2010Q1 onwards only for forecasting exercise.
2. **Model Selection:** We plotted and compared the residuals for the validation period (2014 Q4 till 2016 Q3) to determine the most suitable forecasting method (Exhibit 3.2). Residuals for MLR (Quadratic) with Additive seasonality fit the best. Moreover we plotted the ACF plot for residuals of MLR Quadratic with additive seasonality model, and verified that there was no correlation in errors.
3. **Confidence in forecast:** As shown in exhibit 3.3 p-values for coefficients are significant. Moreover the forecast captures the past data pretty well. Therefore we express **High** confidence in the forecasts.
4. **Other comments:** The data and forecasts are to be updated as we go along, this will enrichen our forecast for future periods.

III. INDIA

1. **Data clean-up:** India data was least impacted by the recession. Hence, the entire data from Q1 2002 to Q3 2016 has been used.
2. **Model selection:** As shown in *Exhibit 4.1*, the MLR model gave the least error plot compared to naïve-seasonal and Holt's Winter smoothing models. An ACF plot (Exhibit 4.2) however, displayed seasonality and an AR2 was applied to the residuals. Forecast errors were adjusted in validation to come up with a better plot (Exhibit 4.3)
3. **Confidence in forecasts:** As shown in Exhibit 4.4, MAPE is lowest for the AR2 adjusted MLR. However, if we are to forecast future values, we can only go up to 2 quarters and will need to keep revising forecasts with new data. In the absence of that, the MLR does a good job. Confidence in forecasts in **high**.
4. **Other comments:** External data such as FDI investment into India can be monitored to find correlation with industrial output and generate better forecasts.

CHINA

1. **Data clean-up:** The data for the Industrial production of China was provided in monthly intervals. Based on our forecasting objective, the data needed to be agglomerated into quarterly intervals. In doing so, we had to get rid of the data for the last 2 months of the data.
2. **Model selection:** To begin with, a naïve model (with lag 4) was run to forecast. Exhibit 5.1 shows the actual and forecast from the naïve model. As can be seen from the exhibit, this is not a particular indicator of the future. We ran Holts- Winter model with multiplicative model as the model showed both quadratic trend and seasonality. The parameters and error output are as shown in Exhibit 5.2. Although, MAPE for validation and training are low, the residual plot and ACF of the residual plots aren't acceptable for forecasting purposes. The plots are displayed in Exhibit 5.3. AR(2) model was run on the errors. An ACF plot of the model is displayed in Exhibit 5.4.

   We ran a MLR model to improve on the forecast by Holt- Winter method. The MLR model was run with a quadratic trend with multiplicative seasonality. The parameters of the model are shown in Exhibit 5.5. Although the model had a low MAPE, there was significant autocorrelation as shown in Exhibit 5.6. Hence, we adjusted the model by running AR(1) forecast.

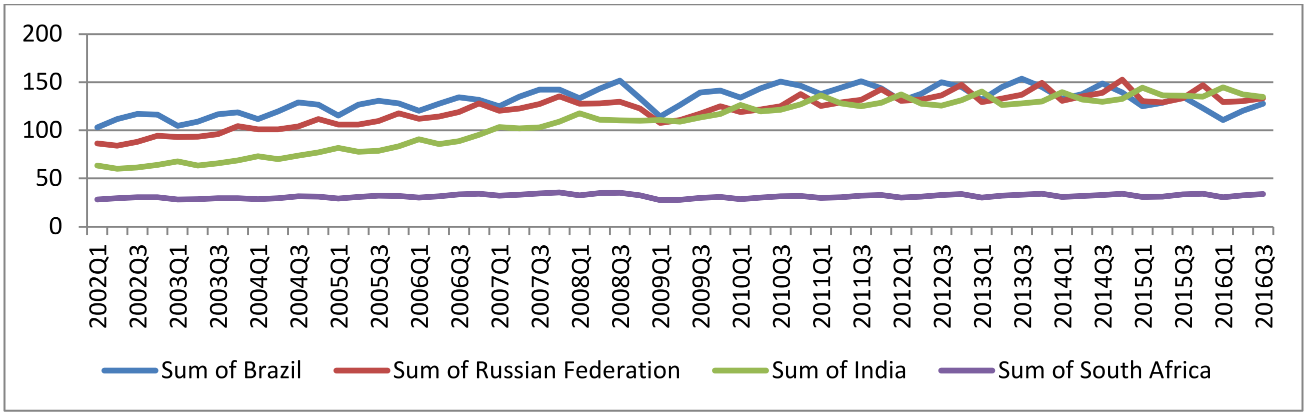
3. **Confidence in forecasts:** Very **high** confidence in the model-with MAPE of validation data approximatey 5% and the errors randomly distributed.

V. SOUTH AFRICA

1. **Data clean-up:** A stark fall in the level of data was observed post 2008 as can be seen in Exhibit 6.1 in training period and hence, the following measures had to be taken:
   - Exclusion of pre-recession data
   - Usage of a dummy variable for pre-recession versus post-recession data points
   - Fitting model over the entire series as is, expecting self-learning by the model
2. **Model selection:** We begin the benchmarking by taking Naïve predictors as the base. We see that the Naïve does a good job in terms of predicting the values; however, it is important to note that the series shows under-prediction to a certain extent, which is not a favorable case for the forecasting problem at hand. The results of Holt Winters over the entire dataset outperform all the other models and over the Holt Winters post-recession data only, implying that the effects of extra learning of seasonality from the data before recession is more than the bias created by its higher level. The results are as shown in Exhibit 6.1. A comparison across all the models is as shown in Exhibit 6.2 and Exhibit 6.4
3. **Confidence in forecasts:** As can be seen in Exhibit 6.3, we observe a fairly random distribution of residuals and no remaining seasonality. Also, comparing the MAPE as done in Exhibit 6.4 we express high **confidence** about the forecast.
4. **Other comments:** The evidence of polynomial trend (of power 3) has been analyzed by MLR with multiplicative seasonality. Also, the bias for pre-recession level has been analyzed by forecasting basis only post-recession data. Holt Winters Multiplicative over the complete dataset outperforms all these models as shown in Exhibit 6.2 and Exhibit 6.4 and is the most effective method in forecasting over validation and hence the final method of our choice for South Africa

Exhibits

Exhibit 1: Quarterly aggregated data for the BRIS countries and China



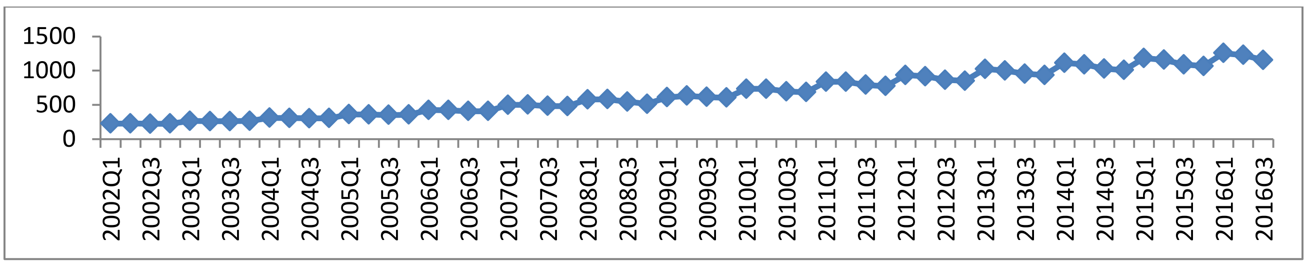
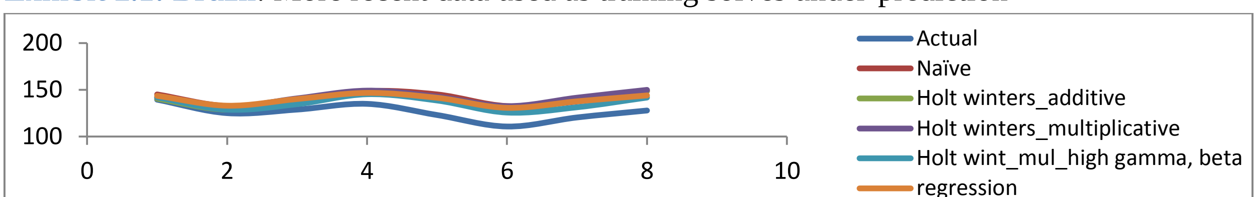
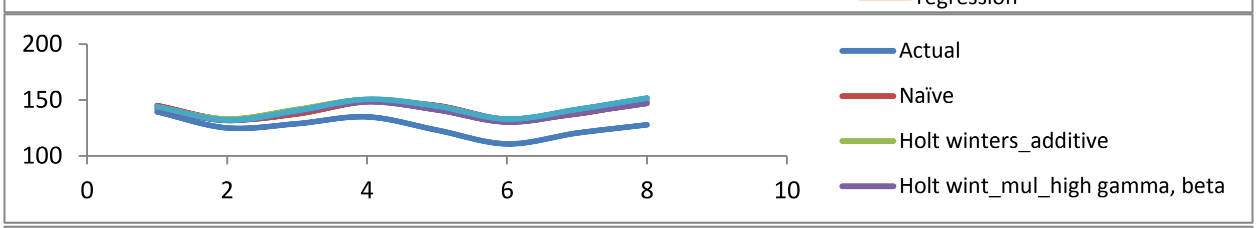


Exhibit 2.1: Brazil: More recent data used as training solves under-prediction





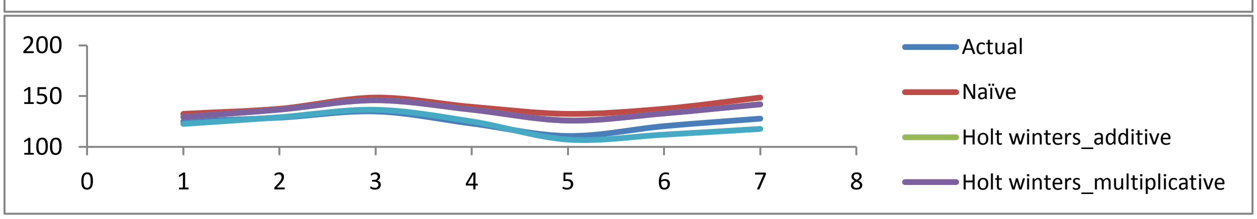


Exhibit 2.2:

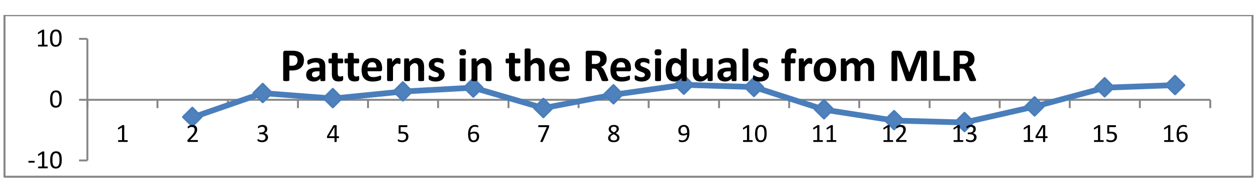
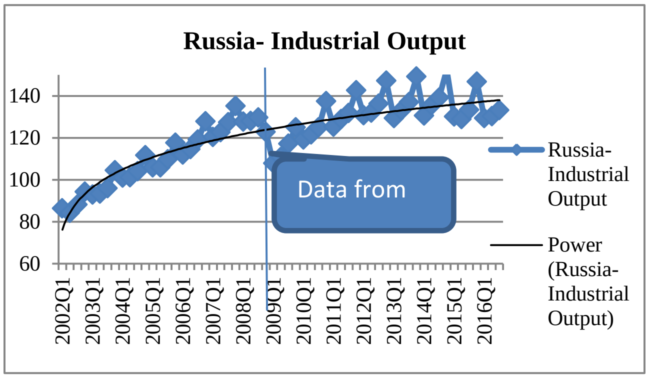


Exhibit 2.3

Input Variables	Coefficient	Std. Error	t-Statistic	P-Value	CI Lower	CI Upper	RSS Reduction
Intercept	146.8511	3.111308859	47.19912848	4.31E-12	139.8128	153.8893	267681
t	-0.35061	0.719621095	-0.48722042	0.637749	-1.97851	1.277282	932.575
quarter_Q1	-11.8014	2.138115725	-5.51955076	0.000371	-16.6382	-6.96468	605.9349
quarter_Q2	-1.86923	2.122853382	-0.88052749	0.401475	-6.67146	2.932997	130.9396
quarter_Q3	8.683856	2.138115725	4.061452685	0.002836	3.847102	13.52061	96.9175
t^2	-0.11044	0.043772499	-2.52303531	0.032606	-0.20946	-0.01142	47.35648

Residual DF	9
R^2	0.964399
Adjusted R^2	0.944621
Std. Error Estimate	2.72751
RSS	66.95381

Exhibit 3.1: Russia : Difference in nature of Data before and after recession



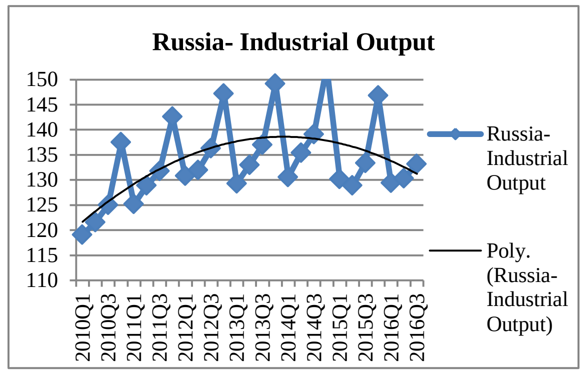
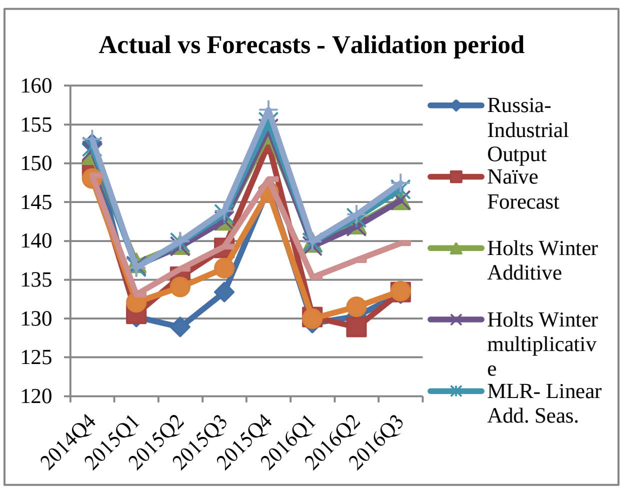


Exhibit 3.2 (A): Comparison of Actual vs Forecast and Residuals for Validation period



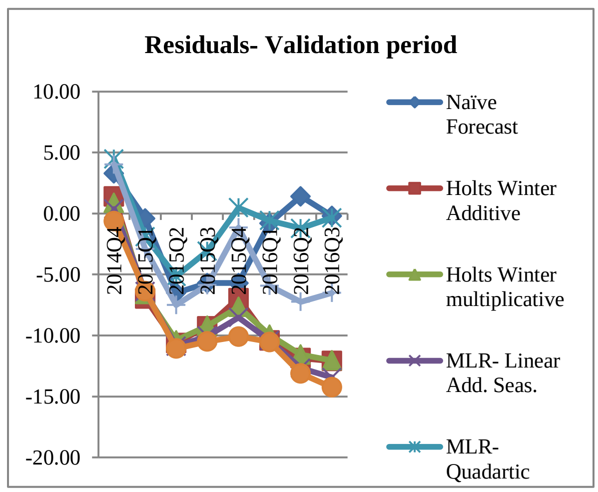
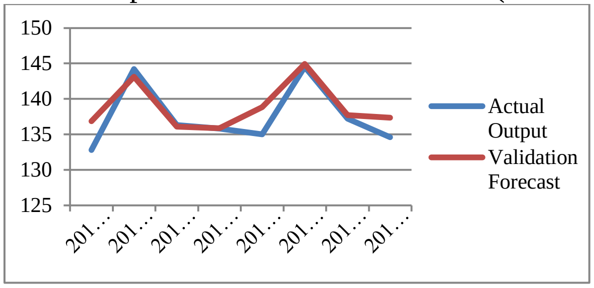
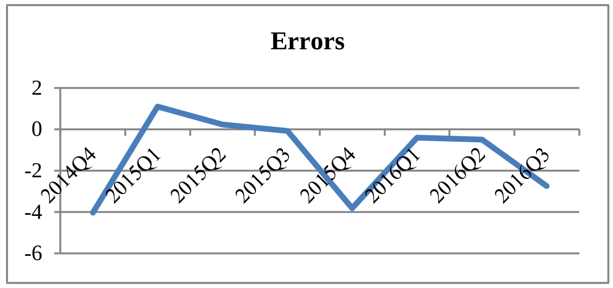


Exhibit 3.2 (B): ACF for residuals of MLR (Quardartic) with Additive seasonality model

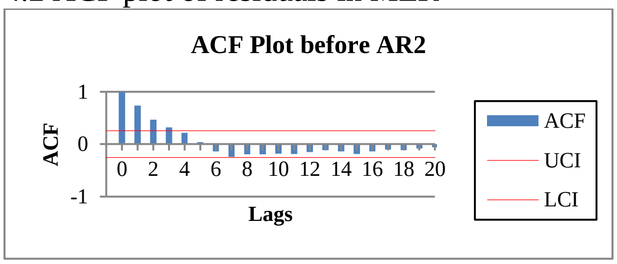
Exhibit 4 – India

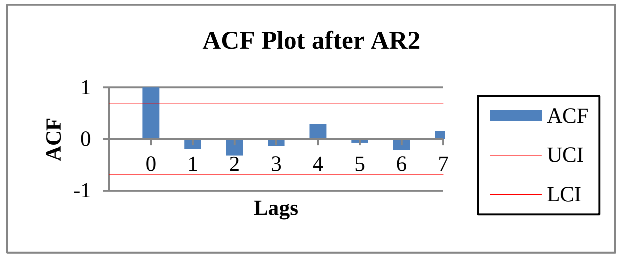
4.1 MLR plot of actuals vs forecasted (validation)



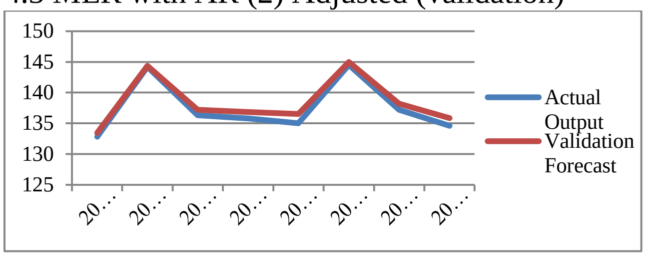


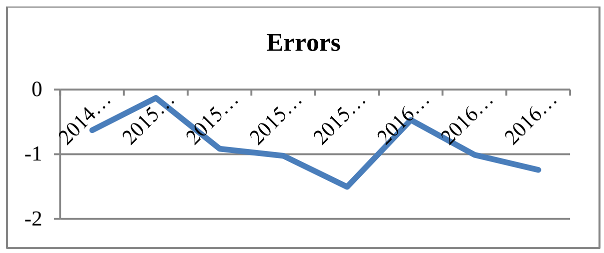
4.2 ACF plot of residuals in MLR





4.3 MLR with AR (2) Adjusted (validation)

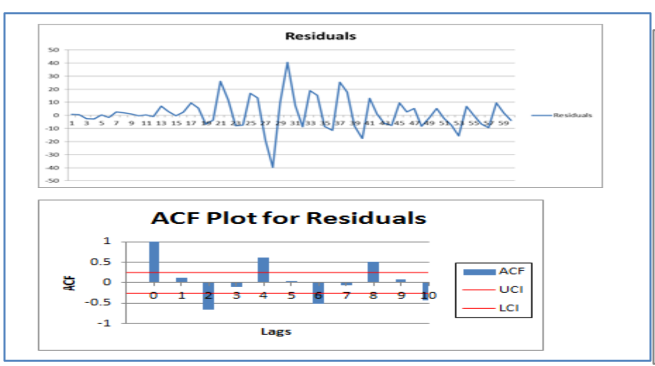


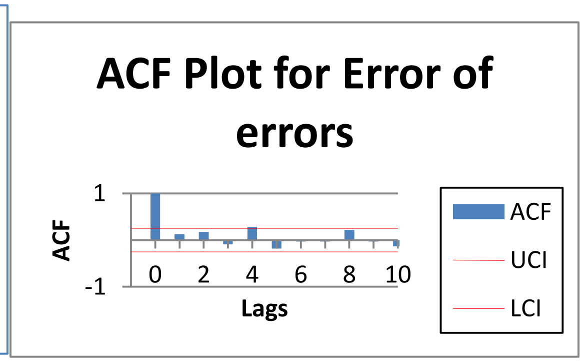


4.4 MAPE Comparison

MODEL	MAPE VALUE
Naïve-Seasonal	6%
MLR with additive trend	1.2%
MLR with AR(2) adjusted	0.63%
Holt's Winter	1.7%

Exhibit 5.1, 5.2, 5.3: China





Exhibits 5.4 & 5.5

Regression Model

Input Variables	Coefficient	Std. Error	t-Statistic	P-Value	CI Lower	CI Upper	RSS Reduction
Intercept	5.221923	0.009598348	544.0439643	2.95765E-89	5.202603	5.241244	2042.981
Time	0.044225	0.000720486	61.38169128	8.77312E-46	0.042774	0.045675	12.3126
Time sq	-0.000217	1.31762E-05	-16.491838	6.14931E-21	-0.00024	-0.00019	0.099587
Q1	0.148466	0.007522914	19.73512028	4.35168E-24	0.133323	0.163608	0.091026
Q2	0.111854	0.00751251	14.88903291	3.27663E-19	0.096732	0.126976	0.071014
Q3	0.042444	0.007506233	5.654527738	9.51062E-07	0.027335	0.057553	0.011703
Q4	0	0	N/A	N/A	0	0	0

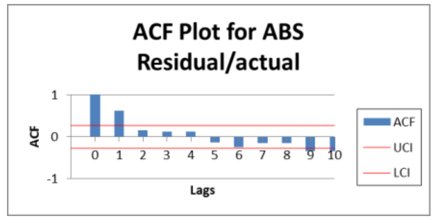


Exhibit 6.1: South Africa: Holt Winters Predictions

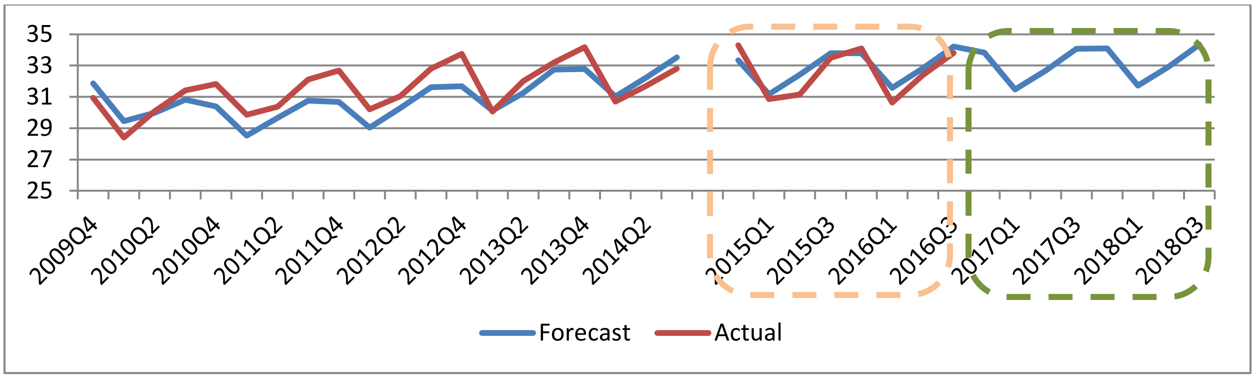


Exhibit 6.2: Performance Comparison: All methods

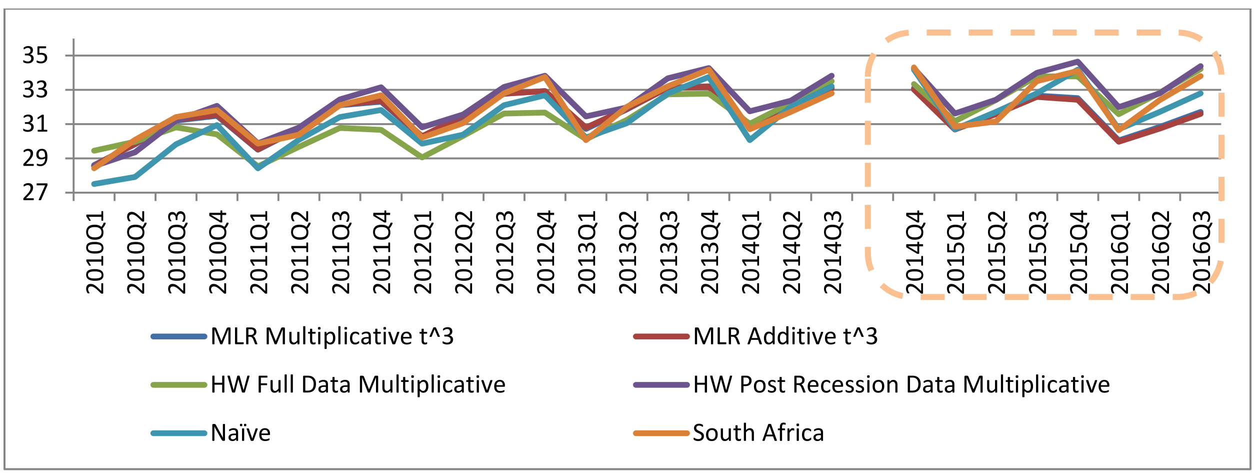
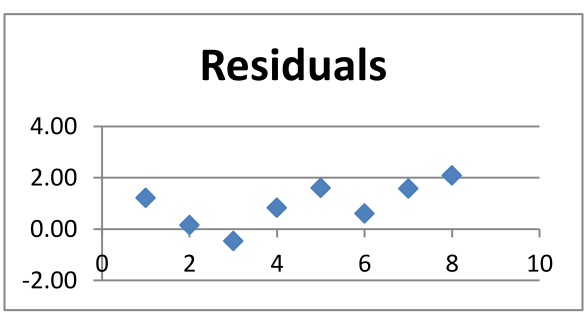


Exhibit 6.3 and Exhibit 6.4



Method	MAPE
MLR Multiplicative Power 3	3.21%
MLR Multiplicative Power 2	3.59%
MLR Additive Power 3	3.40%
MLR Additive Power 2	3.38%
HW Multiplicative Full Data	1.90%
HW Additive Full Data	2.00%
Naïve Prediction	1.28%
HW Multiplicative Partial Data	2.14%
HW Additive Partial Data	2.23%