Province of KwaZulu-Natal

Provincial Treasury

IMES Unit

MODELING AND FORECASTING KWA-ZULU-NATAL'S POTENTIAL (AVERAGE) GDP DATA WITH A TIME SERIES ARIMA MODEL

2014

Clive Coetzee

General Manager: IMES Unit

Economist

clive.coetzee@kzntreasury.gov.za

033 897 4538

Working Paper 2a: dd January 2014

Introduction

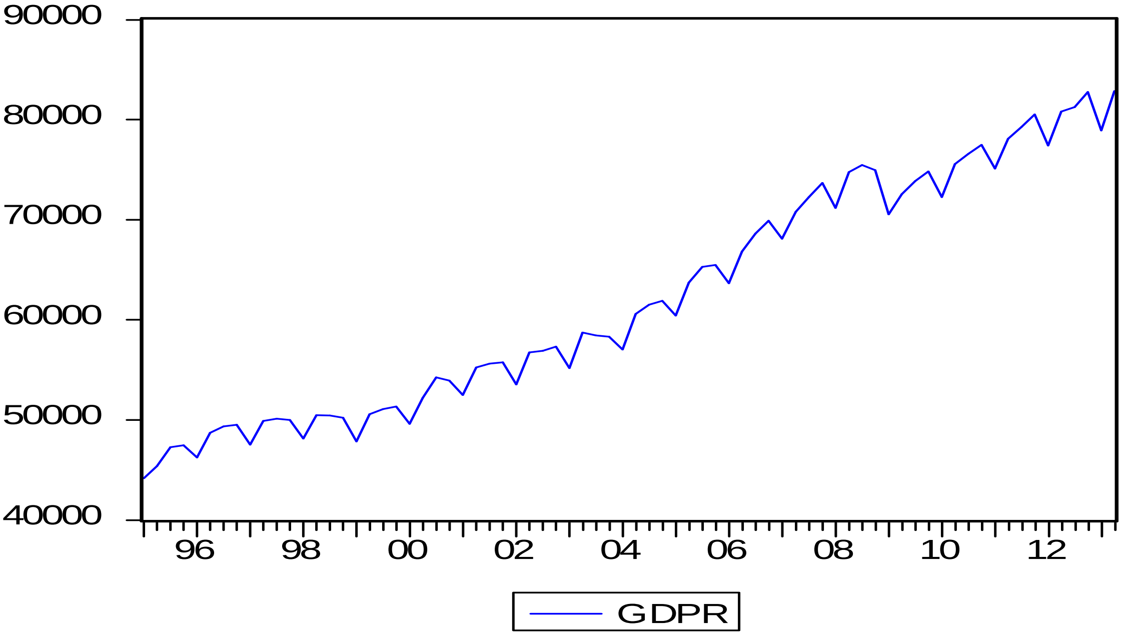
The gross domestic product (GDP), a basic measure of an economy's economic performance, is the market value of all final goods and services produced within the borders of a nation in a year. GDP can be defined in three ways, all of which are conceptually identical. First, it is equal to the total expenditures for all final goods and services produced within the country in a stipulated period of time (usually a 365-day year). Second, it is equal to the sum of the value added at every stage of production (the intermediate stages) by all the industries within a country, plus taxes less subsidies on products in the period. Third, it is equal to the sum of the income generated by production in the country in the period, which is compensation of employees, taxes on production and imports less subsidies, and gross operating surplus (or profits).

Potential Gross Domestic Product, or Potential GDP, refers to the highest level of real Gross Domestic Product output that can be sustained by a country over the long term. It is thus a measurement of what a country’s gross domestic product would be if it was operating at full employment and utilizing all of its resources. Generally, this amount is greater than the actual GDP of a country, thus, the difference between a country’s potential GDP and its actual (real) GDP is known as the GDP or output gap. The output gap may be caused by economic inefficiencies such as unemployment, inflation, and government regulations, which most economies encounter, thereby impeding production levels.

KZN Gross Domestic Product

The economy of KwaZulu-Natal is South Africa’s second largest after that of Gauteng contributing on average, 15.7% (2011) to the country’s GDP, which amounts to about R78 billion annually. The following paper is based on the Gross Domestic Product (GDPR) of KwaZulu-Natal, using Quarterly data for the period 1995q1 to 2013q2.

Figure 1.1: Provincial GDPR (R'm, constant 2005 prices)



The figure (figure 1.1) displays the quarterly data of GDPR for KZN, in constant 2005 prices, over the period 1995 to 2013. Since figure 1.1 exhibits seasonal effects, the GDPR time series is seasonally adjusted, as shown in the figure below (Figure 1.2).

Figure 1.2: Provincial Seasonal Adjusted GDPR (R'm, constant 2005 prices)

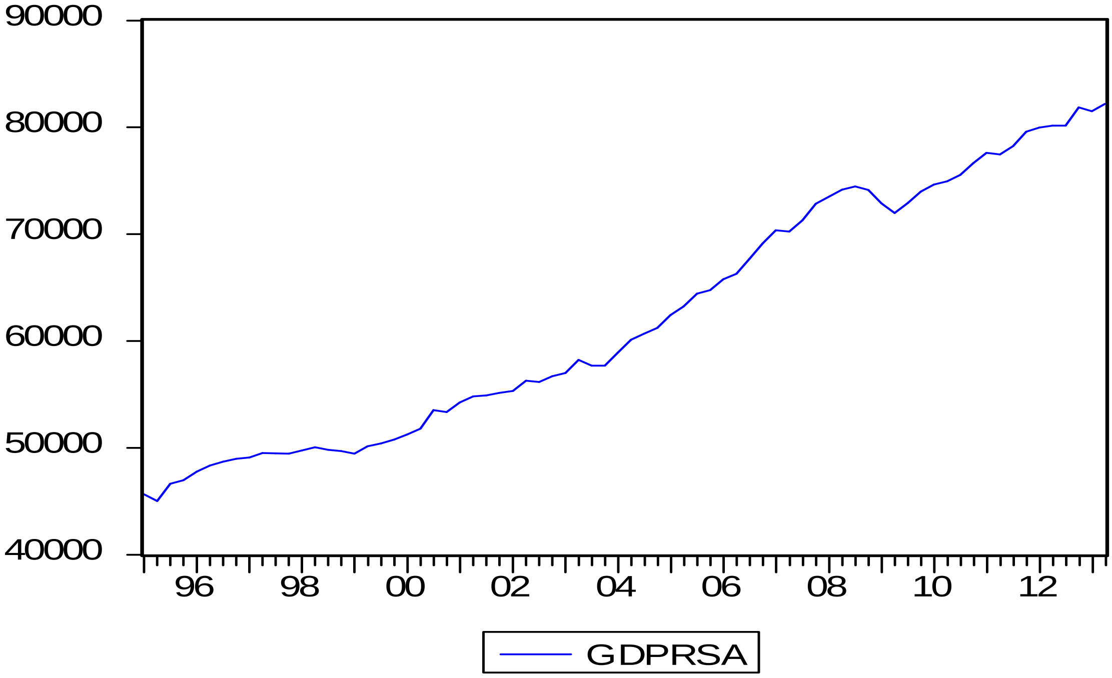
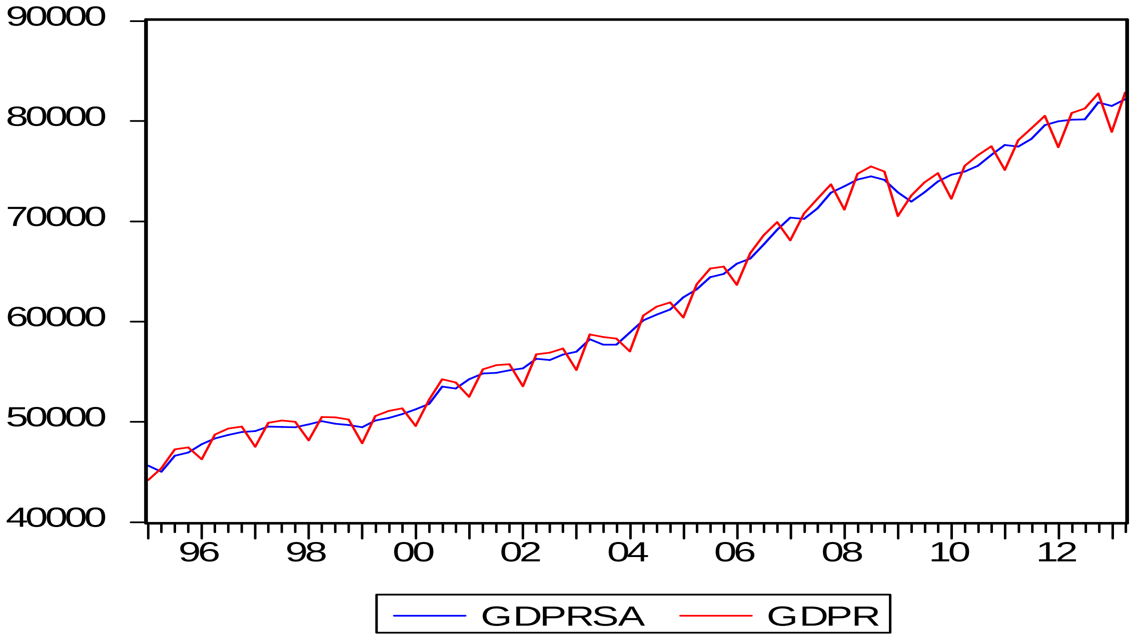
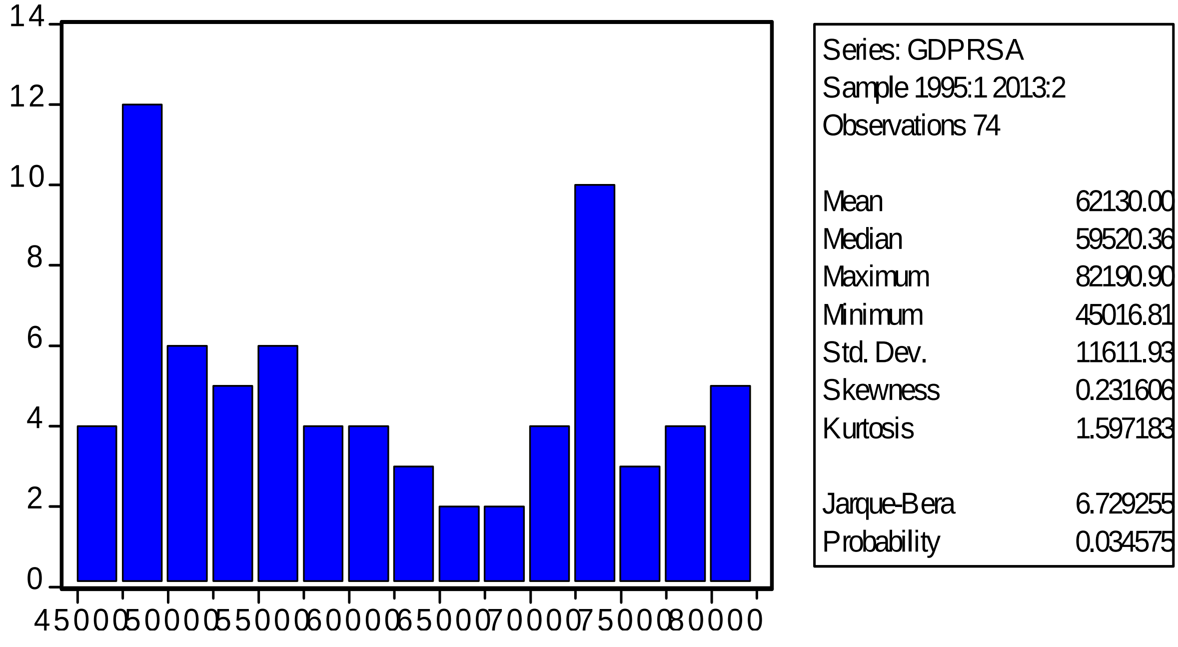


Figure 1.3: Provincial GDPR and Seasonal Adjusted Provincial GDPR (R'm, constant 2005 prices)



The Figure above (Figure 1.3) includes the seasonal adjusted GDPR which is generated using the Seasonal Adjustment: ratio to moving average method in EViews. In Figure 1.4, the p-value of 0.03 indicates that the series is not normally distributed.

Figure 1.4: Summary Statistics of KZN GDP Seasonal Adjusted



The GDPR is tested for non-stationarity (Figure 1.2) against the alternative that the variable is trend stationary. To perform the Unit Root Test on a AR(1) model, the following regression equation will be estimated:

$$y_t = \alpha + \delta t + \beta y_{t-1} + \mu_t$$

where:

y_t = variable to be tested (i.e. provincial seasonal adjusted GDPR)

α = constant

t = trend

u_t = white noise innovation

The Augmented Dickey-Fuller (ADF) Unit Root Test is based on the following three regression forms:

- with constant and trend (τ_τ)
- with constant (τ_μ)
- without constant and trend (τ)

and the testable hypothesis: H_0: $\beta = 0$ (i.e. y_t has a unit root).

The time series consists of 74 observations and includes 2 lags. The Table Below (Table 1.1) presents the results for each variable.

Table 1.1: Augmented Dickey-Fuller using Level and 1st Difference Data

Series	Model	ADF		
		Lags	$\tau_\tau \tau_\mu \tau$	$\varphi_3 \varphi_1$
Level	τ_τ	2	-2.11	1.81
	τ_μ	2	0.73	0.65
	τ	2	3.89	
		Lags	$\tau_\tau \tau_\mu \tau$	$\varphi_3 \varphi_1$

1st Difference	τ_τ	2	-4.53***	12.54***
	τ_μ	2	-4.41***	16.36***
	τ	2	-2.48***	

(** significant the 5 percent level, *** significant at the 1 percent level)

Comparing the ADF Test Statistics at the critical values of 1 percent, 5 percent, and 10 percent (tau values), and the F-Statistics at 1 percent, 5 percent, and 10 percent levels (phi values) both suggest that the time series is non-stationary in level format.

The ADF test suggests that the variable is stationary in 1st difference format and thus integrated to the order I(1). The stationary time series, 1st difference of provincial seasonal adjusted GDPR in R'millions at constant 2005 prices, is displayed in the figure below (Figure 1.5). The table (Table 1.2) shows that the correlogram of the seasonally adjusted variable, GDPR, is white noise as it shows no significant pattern, and further confirms that GDPRSA becomes stationary in 1st difference format.

Figure 1.5: The Provincial Seasonal Adjusted GDPR in 1st Difference Format (R'm, constant 2005 prices)

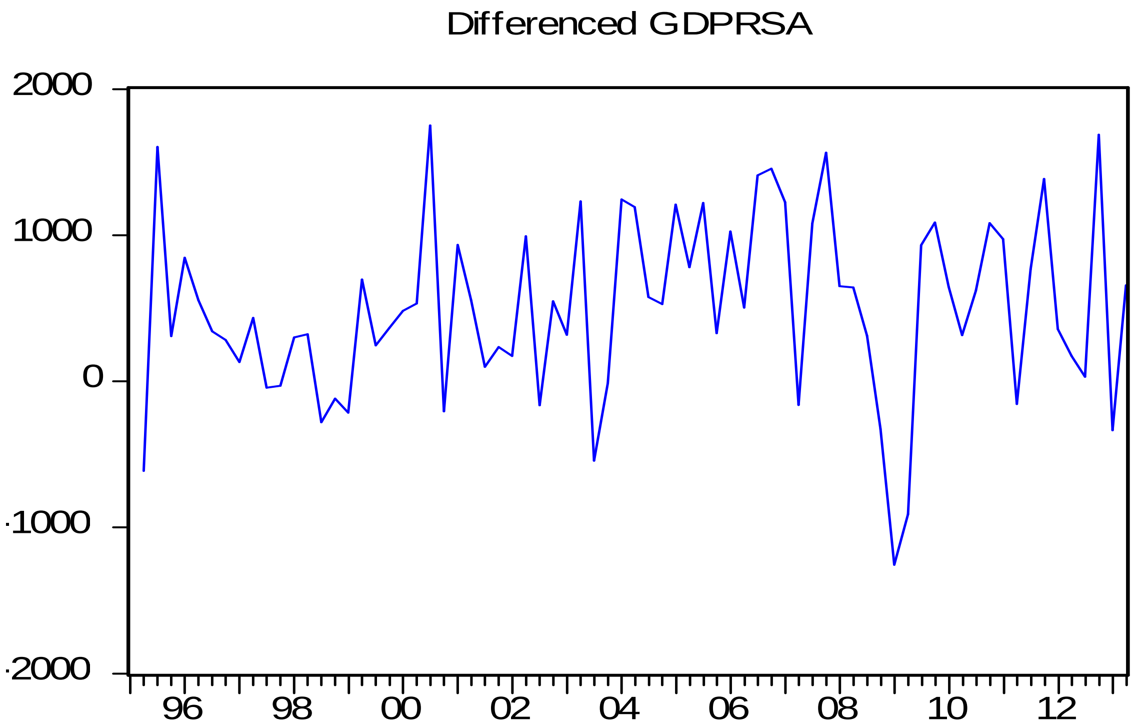


Table 1.2: Correlogram of GDPRSA in 1st difference format

Date: 11/25/13 Time: 13:57
Sample: 1995:1 2018:2
Included observations: 73

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |*. |	. |*. |	1	0.083	0.083	0.5232	0.469
. | . |	. | . |	2	0.053	0.047	0.7403	0.691
. | . |	. | . |	3	-0.002	-0.010	0.7406	0.864
. |** |	. |** |	4	0.214	0.214	4.3660	0.359
. | . |	. | . |	5	0.025	-0.009	4.4166	0.491
.*| . |	**| . |	6	-0.183	-0.215	7.1519	0.307
. |*. |	. |*. |	7	0.076	0.127	7.6364	0.366
. | . |	. | . |	8	-0.009	-0.054	7.6435	0.469
. | . |	. | . |	9	-0.027	-0.057	7.7074	0.564
**| . |	.*| . |	10	-0.207	-0.111	11.418	0.326
. | . |	. | . |	11	0.028	0.029	11.485	0.404
. | . |	.*| . |	12	-0.049	-0.075	11.697	0.470
. | . |	. |*. |	13	0.019	0.081	11.730	0.550
.*| . |	. | . |	14	-0.081	-0.035	12.340	0.579
. | . |	. | . |	15	0.057	0.049	12.651	0.629
. | . |	. | . |	16	0.002	-0.037	12.652	0.698
. | . |	. | . |	17	-0.043	-0.024	12.831	0.747
. | . |	. | . |	18	0.017	0.013	12.859	0.800
. | . |	. | . |	19	-0.048	-0.056	13.093	0.834
. |*. |	. |*. |	20	0.122	0.091	14.622	0.798

Arima Methodology and Application

ARMA models are used in time series analysis to describe stationary time series and to predict future values in this series. The ARMA model consists of an autoregressive (AR) model and a moving average (MA) model. The ARMA model is as follows:

$$\text{ARMA (p,q): } y_t = c + \sum_{i=1}^{p} \varphi y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \ , \varepsilon_t \sim WN\ (0, \sigma^2)$$

where:

- p is the order of the autoregressive model and
- q is the order of the moving average model.

The ARMA model is combined with two parts i.e. Autoregressive Model and the Moving Average Model.

$$\text{AR(p): } y_t = c + \sum_{i=1}^{p} \varphi y_{t-i} + \varepsilon_t \ , \varepsilon_t \sim WN\ (0, \sigma^2)$$

$$\text{MA(q)}: y_t = \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \ , \varepsilon_t \sim WN\ (0, \sigma^2)$$

The error terms ε_t are generally assumed to be independent identically-distributed random variables.

The Autoregressive Integrated Moving Average (ARIMA) will be used to determine the best model to fit the time series. ARIMA models form an important part of the Box-Jenkins approach to time-series modelling. A non-seasonal ARIMA model is classified as an ARIMA (p,d,q) model where:

- p is the number of autoregressive terms
- d is the number of non-seasonal differences
- q is the number of moving average lags

$\{y_t\}$ is said to be ARIMA (p,d,q) if:

$$(1 - L)^d \ \emptyset^* (L)\ y_t = c + \theta(L)\varepsilon_t$$

where:

$\emptyset^*$ (L) is defined in $\emptyset$(L)=(1 – L) $\emptyset^*$ (L) , $\emptyset^*$ (z) ≠ 0 for all $|z| \leq 1$. And $\emptyset$(L) is defined in $\emptyset$(z) ≠ 0 for all $|z| \leq 1$.

The process $\{y_t\}$ is stationary if and only if d=0, thus ARMA (p,q) process: $\emptyset(L) y_t = c + \theta(L)\varepsilon_t \ , \sim WN\ (0, \sigma^2)$.

The following criterions will be applied in order to determine the best model:

- Relatively small BIC/AIC
- Relatively small SEE
- And a relatively high Adjusted R^2

The ADF Test, line graph, and correlogram above strongly suggest that the ARIMA (0,1,0) is suitable for the time series. The table (Table 1.3) below displays the six best ARIMA Model combinations in 1st difference.

Table 1.3: ARIMA Models in 1st difference

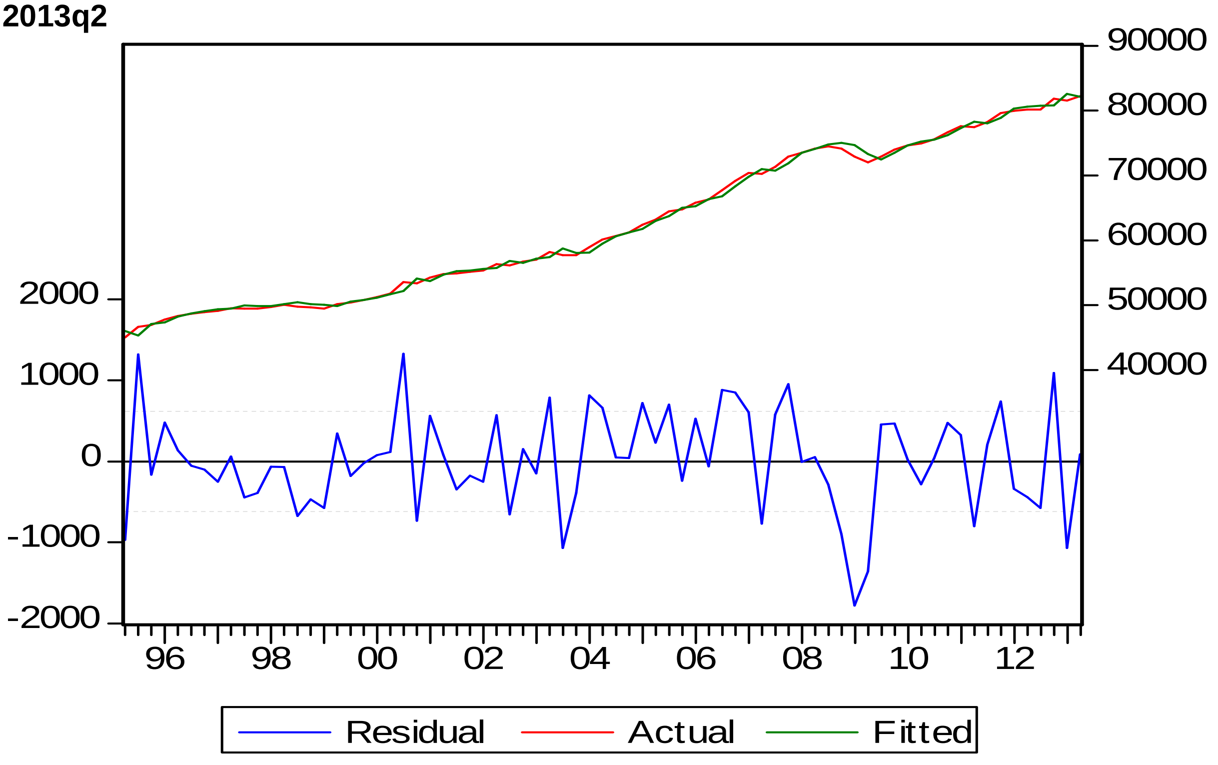
ARIMA Model	BIC	Adjusted R^2	SEE	Corrlogram

(1,1,1)	15.77	-0.019	616	0.124	
(1,1,4)	15.91	-0.162	658	-0.176	*
(2,1,2)	15.77	-0.048	615	0.133	
(3,1,3)	15.84	-0.118	639	0.119	
(4,1,4)	15.80	-0.060	625	0.142	
(5,1,5)	15.81	-0.050	627	0.146	
(6,1,6)	15.89	-0.121	652	0.118	

It is quite clear that ARIMA (1,1,1), ARIMA (1,1,4) and ARIMA (2,1,2) are the three better models from the table (Table 1.3) based on the criteria. Thus, these three models will be examined further, in the level format, in order to determine which model is the best fit for the time series.

The Figures (Figure 1.6, Figure 1.7 and Figure 1.8) below illustrates the actual, fitted values and residuals of the three models. The fitted values in figure 1.6 more closely fit the actual values of the time series, GDPRSA.

Figure 1.6: Actual and fitted values of GDPRSA of ARIMA (1,0,1), 1955q1 – 2013q2



* ARIMA (1,1,4) is included in Table 1.3 since both variables of the model are statistically significant i.e. AR(1): t_{stat}=3.171472 and MA(1): t_{stat}=4.435065.

Figure 1.7: Actual and fitted values of GDPRSA of ARIMA (1,0,4), 1955q1 – 2013q2

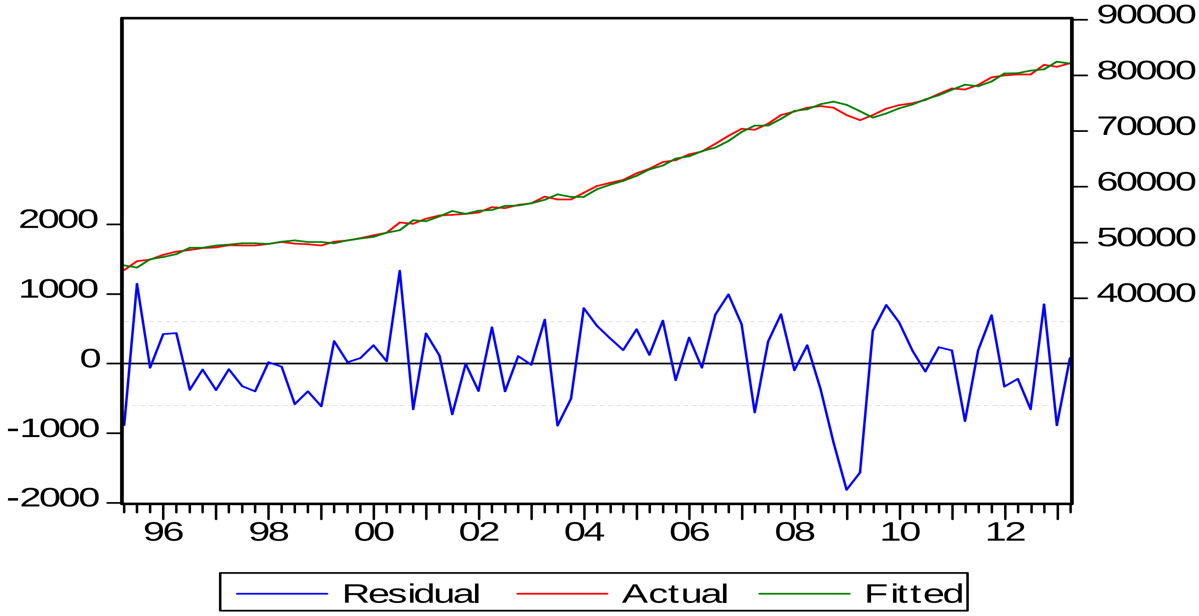
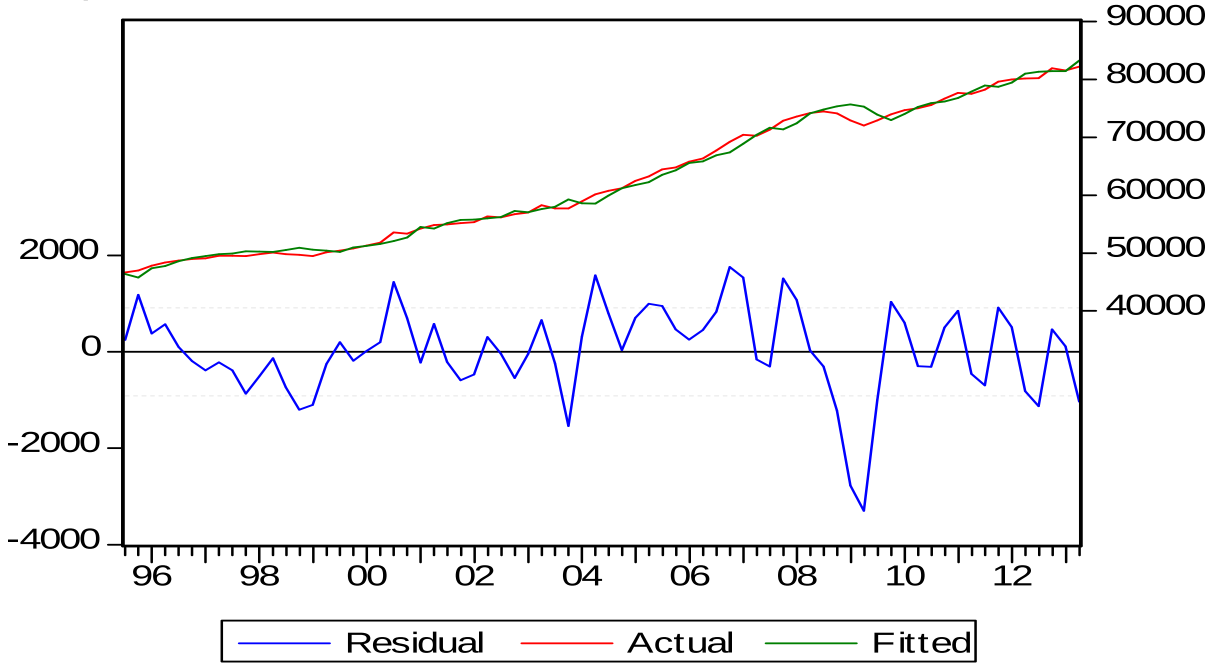
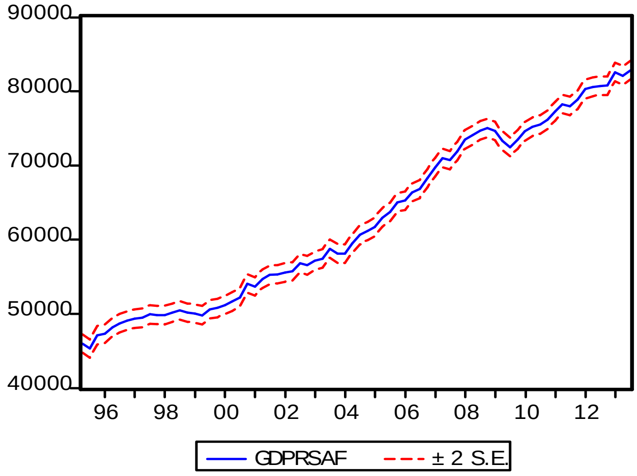


Figure 1.8: Actual and fitted values of GDPRSA of ARIMA (2,0,2), 1955q1 – 2013q2



In the figures (Figure 1.9, Figure 1.10 and Figure 1.11), the dependant variable, GDPRSA, is forecasted using the static method.

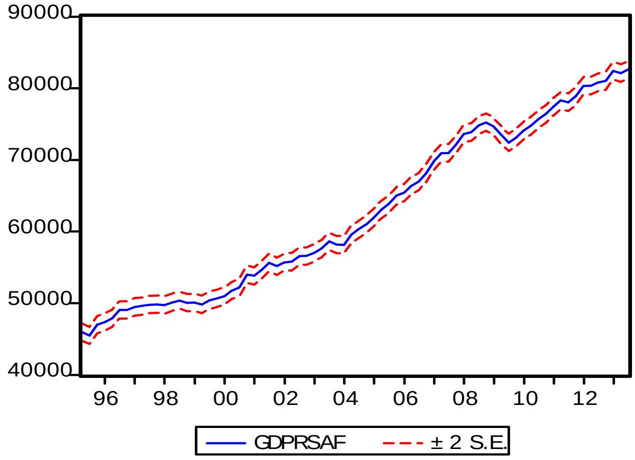
Figure 1.9: Forecasted values of GDPRSA of ARIMA (1,0,1), 1995q1 – 2013q2



Forecast: GDPRSAF
Actual: GDPRSA
Forecast sample: 1995:1 2018:2
Adjusted sample: 1995:2 2013:3
Included observations: 73

Root Mean Squared Error	611.5182
Mean Abs. Percent Error	475.4323
Mean Absolute Percentage Error	0.767726
Theil Inequality Coefficient	0.004823
Bias Proportion	0.000095
Variance Proportion	0.000389

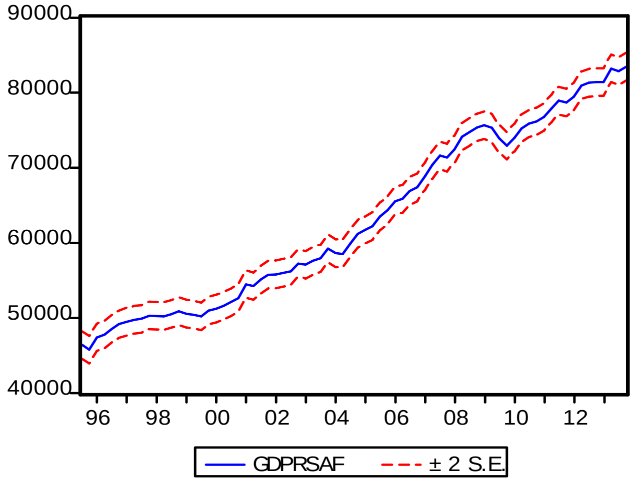
Figure 1.10: Forecasted values of GDPRSA of ARIMA (1,0,4), 1995q1 – 2013q2



Forecast: GDPRSAF
Actual: GDPRSA
Forecast sample: 1995:1 2018:2
Adjusted sample: 1995:2 2013:3
Included observations: 73

Root Mean Squared Error	594.5415
Mean Abs. Percent Error	466.3612
Mean Absolute Percentage Error	0.752527
Theil Inequality Coefficient	0.004689
Bias Proportion	0.000053
Variance Proportion	0.000800

Figure 1.11: Forecasted values of GDPRSA of ARIMA (2,0,2), 1995q1 – 2013q2



Forecast: GDPRSAF
Actual: GDPRSA
Forecast sample: 1995:1 2018:2
Adjusted sample: 1995:3 2013:4
Included observations: 72

Root Mean Squared Error	899.0406
Mean Abs. Percent Error	677.3740
Mean Absolute Percentage Error	1.070889
Theil Inequality Coefficient	0.007066
Bias Proportion	0.000253
Variance Proportion	0.002759

The Mean Absolute Percentage Error value of 0.75 (Figure 1.10) for ARIMA (1,0,4) is lower than the 0.76, 1.07 values for ARIMA (1,0,1) and (2,0,2) respectively. The Theil Inequality Coefficient lies between 0 and 1, with 0 indicating a perfect fit. Hence, the ARIMA (1,0,4) is relatively the most suitable model for the time series, GDPRSA.

Figure 1.12: Forecasted values of GDPRSA, 1995q1 – 2018q2

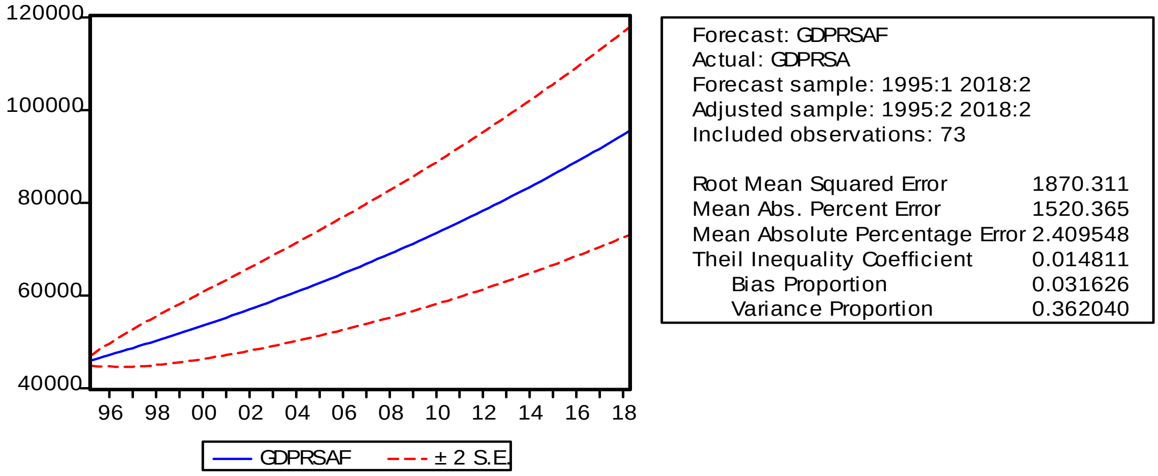
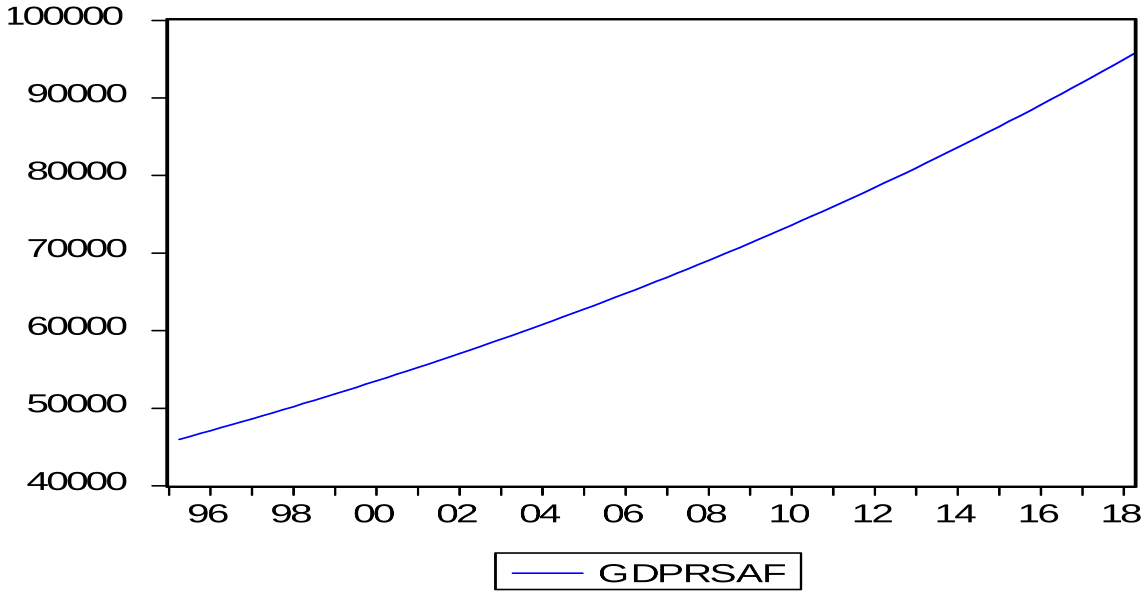


Figure 1.12 above displays the GDPRSA for the ARIMA (1,0,4) model using the dynamic method.

Figure 1.13: Forecasted values of GDPRSA



The figure (Figure 1.11) displays the behaviour of the forecasted values of the GDPRSA from the 2nd quarter of 1995 until the 2nd quarter of 2018.

KZN Potential GDP

The table (1.4) below displays the year-on-year growth rate of the forecasted values of the GDPRSA model from 2014 to 2018. The growth rate is initially 2.81% in 1996 but becomes stable at 3.22% from 2014 until 2018. The model, therefore, suggests that the long term potential growth rate for the KZN economy is estimated at 3.22 per cent.

Table 1.4: Year-on-Year Growth Rates (%), GDPRSA (1,0,4)

2014	3.22
2015	3.22
2016	3.22
2017	3.22
2018	3.22