Journal of Modern Mathematics and Statistics
MEDWELL PUBLICATIONS SINCE 2000

# Box-Jerkins Approach to Fuel Prices and Currency Strength in the International Market

[1]Matthew Femi Olayiwola and [2]Samuel Babajide Atoyebi
[1]*Department of Mathematics, African Leadership Academy South Africa*
[2]*DSK Statistical Consult Nig. Ltd., Nigeria*

**Key words:** Box-Jerkins, fuel prices, exchange rate, stationarity test, AutoRegressive Integrated Moving Average, mathematical modeling, time series

**Corresponding Author:**
Matthew Femi Olayiwola
*Department of Mathematics, African Leadership Academy South Africa*

Page No.: 08-15
Volume: 15, Issue 1, 2021
ISSN: 1994-5388
Journal of Modern Mathematics and Statistics


**Abstract:** South African economy has been said to be fuel price dependent and that the fuel price keeps increasing on yearly basis. Thus, there is a need to develop a scientific model to predict the behaviour and future values of fuel prices and exchange rates. This study employs an Autoregressive Integrated Moving Average (ARIMA) in predicting the fuel prices and the strength of the South African rand in the international market based on the 35 years monthly data. The prices of fuel exhibit an upward trend variation. Among the competing models estimated per each variable, ARIMA (3, 1, 1), ARIMA (3, 1, 1), ARIMA (1, 1, 2) and ARIMA (1, 0, 1) are respectively found to be the best for modeling and predicting the behaviour and future values of diesel, paraffin, petrol and exchange rate (ZAD-USD). Their respective forecasting accuracy are 93.4, 91.7, 91.5 and 79.3%, respectively. Hence, predicting the future values of fuel prices and the strength of South African currency against the United States of America dollars using these models will help the policymakers and all stakeholders make well-informed decisions in planning. Further research can be carried out on the short and long-run relationship between the four variables. Also, the direction of causality among the variables should be looked into to determine which variable can be used to determine the future values of the other.

## INTRODUCTION

When the productive capacity of a country is regarded as increasing and it has a positive impact on the lives of the citizens, then the country's economy is said to be growing[1-3]. The real economic growth of South Africa per annum was around 2.83% during 1993-2017, during 1988-1992 there was about 1.4% which is a notable large growth[4]. However, according to Global Economic Outlook , it fell below the 3.1% international mean for developing countries. The growth of the economy of South Africa has been volatile and fluctuating, most especially because of its connection to the global economy. For example, the crisis of East Asia in 1998 affected the South African economy negatively, this happened when it was adjusting to some structural reforms and trade liberalisation global activities[5]. Government expenditure is another factor that has a huge capacity to influence economic growth and this largely relies on the amount borrowed or tax revenues collected.

It is consequently, imperative to examine the relationship between consumption and the price of fuel in South Africa.

Widespread concern had been created by the large and rapid increase experienced in the oil price over the last years and its impacts on poverty and economic growth in many developing countries. In the modern industrial economy, the quintessential commodity is petrol. Although initially, it was coal that was used to power the industrial revolution, Pennsylvania was where the commercial petrol well was drilled for the first time in 1859 and since then, when it comes to the share of petrol in the primary energy supply of the world, petrol has continued to gain prominence. It now has the largest share, it accounts for >35%. Petrol as an energy source is used for transportation as a liquid fuel, heating and most importantly for electricity generation. For some 90% of energy, the transport systems of the world (including road transport, airplanes, trains and ships) depend on petrol[6]. Consequently, many sectors are highly reliant on petrol in most developing countries. For the production of pesticides, herbicides and fertilizers, industrial agriculture heavily depends on natural gas and petrol as well as in transporting products to markets and power mechanized farm machinery.

One of the factors that influence price and consumption is the amount of fuel available. Consumptions can be very high when an item is abundant because of affordability, also the price will likely be cheaper when an item is abundant. As reported by South Africa Statistics[7] almost 95% of the crude oil requirements of South Africa are imported from Africa and the Middle East and that South Africa has limited oil reserves. Therefore, with South Africa producing just 5% locally, low consumption is expected from consumers and it is an indication of scarcity. The relationship between refinery products and the prices of oil had been investigated by several studies[8-10].

One of the major drivers of economic activities is the prices of oil and for global economic growth, the high prices have been perceived as not being favorable. The popular myth is that generally, high consumer prices are associated with high oil prices. For two reasons, the linkage between the Consumer Price Index (CPI) and oil prices is especially, important for the economy of South Africa. Firstly, in the world, one of the most unequal countries is South Africa in terms of income with a Gini coefficient of 63.1 in 2009[11]. What this means is that the larger part of the population is affected disproportionately by inflation because to keep up with rising prices they will need more income. Furthermore, South Africa is exposed to rising oil prices external shocks because it is an oil-importing country. For these reasons, determining the relationship between consumption and the price of fuel in South Africa is important.

The international crude oil prices are determined by the two leading international NYMEX WTI CRUDE Oil and Brent Oil (Trading Crude oil CFDs, 2020). The former determines the crude oil prices using the benchmark reported in oil future contracts as traded on CME's New York Mercantile Exchange while the latter instead determines the crude oil price using the European international benchmark as traded on the Intercontinental Exchange[12]. This is an indication of the vital role the exchange rate plays in the local price determination of the retail fuel price. A quite number of researches have reported that oil price is significantly related to the economic growth of South Africa. Olayiwola and Seeletse[13] submitted that there exists a unidirectional causality from the consumption of petrol in South Africa its retail price. It was reported by the duo in the same research that the South African economy is energy-dependent as the price of petrol affects the standard of living. Yang *et al*.[14] in a study on oil-price volatility study posited that recessions in oil-consuming nations are a function of higher oil prices. Consumption and price of oil are found to be the key to South African's energy security. The works of Stringer and Yang et al are found to be further sustained in a recent study on the statistical forecasting of petrol price in South Africa, the petrol price in South Africa showed an upward trend that continues into future and that an increase in the price of petrol reduces the disposable cash for consumers[1]. South Africa whose economy is driven by fuel as indicated in several studies imports 95% of its oil consumption from the Middle East and Africa[7]. The volumes of fuels available and their prices in South Africa yearly are strongly related to the strength of the South African rand (exchange rate) in the international market. Hence, this study is carried out to develop a mathematical model for predicting the future behaviour and values of fuel prices and exchange rate (ZAR-USD) in South Africa.

**Box-Jekins procedures:** Box-Jerkins's approach has been used by many researchers in modelling and forecasting the future values of different variables. ARIMA (4, 1, 1) and ARIMA (4, 1, 4) were found to be the best model for predicting the road traffic fatalites in South Africa[15] while ARIMA(1, 0, 2) was developed to model the maternal mortality ratio of the Okonfo Anokoye Teaching Hospital in Kumasi. The inflation rates in the Volta Region of Ghana can best be modelled by ARIMA (2, 1, 2) in studying the series behaviour[16]. Also, ARIMA (1, 0, 1) was found to be the best among the competing models to predict the lake Malawi water levels[17]. Box and Jerkins proposed a four-step iterative process in modelling and forecasting the time series data, namely; model identification, model estimation, model diagnosis and forecasting[18]. The presence of the time series components in the data can be detected graphically through the time series plot, a graphical representation of the series against time[19, 20].

Most time series variables are theoretically known to be nonstationary at a level due to the presence of components/variations in the series. Box and Jerkins[21] established that the stationarity condition for an autoregressive process of order p, AR, (p) requires that all roots of the characteristic equation must fall outside the unit circle. Several tests such as the Augmented Dickey-Fuller test[22]. Phillips-Perron test[23], Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test[24], etc. The ADF test is adopted in this study. The hypothesis o f no unit root is tested using the Augmented Dickey-Fuller (ADF) test is considered in testing for the null hypothesis that the series is nonstationary against the alternative hypothesis of staitionarity. This test is based on simple regression of actual series on its one-period lagged value:

$$Y_t = \theta Y_{t-1} + e_t \tag{1}$$

Where:
$Y_t$ = The actual series
$Y_{t-1}$ = One period lagged of the actual series
$e_t$ = A serially uncorrected white noise error term with a mean of zero and a constant variance

Reparametrizing (1) gives (2):

$$\Delta Y_t = \delta Y_{t-1} + e_t \tag{2}$$

where $\delta$ = (θ-1) and is the first difference operator.

The null hypothesis of $\delta = \theta$ from (2) has an estimated t-value. The Augmented Dickey-Fuller test was introduced to address the serial correlation of the error terms produced in the Dickey-Fuller test [20]. This test makes the error terms white noise by introducing the lagged first difference into the regression equation as in (3):

$$\Delta Y_t = \delta Y_{t-1} + \alpha_i \sum_{i=1}^{m} \Delta Y_{t-i} + e_t \tag{3}$$

Including intercept and time, t in Eq. 3 now gives:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \alpha_i \sum_{i=1}^{m} \Delta Y_{t-i} + e_t \tag{4}$$

The ADF unit root testing procedure uses this model:

$$\Delta y_t = \alpha + \beta_t + yYy_{t-1} + \sum_{j=1}^{\rho} \delta_j \Delta y_{t-j} + e_{it} \tag{5}$$

Where:
α = The intercept(a constant)
β = The coefficient of time, t
γ = The coefficient of $y_{t-1}$
P = The lag order of the autoregressive process
$\Delta_{yt}$ = The first difference of yt, $y_{t-1}$ the one time period lag values
$y_t$, $\Delta y_{t-j}$ = The changes in lagged values
$e_{it}$ = The white noise
γ = The parameter of interest in (5)
γ = 0 = Implies that the series is non-stationary

The identification of the Autoregressive and Moving Average (ARMA) order requires the careful study of the correlogram which is the plot of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) against lags. The ACF and PACF reveal the appropriateness of the autoregressive process and moving average process in fitting the time series data[21]. The autocorrelation values can be estimated using Eq. 6:

$$r_k = \frac{\sum\left(Y_t-\bar{Y}\right)\left(Y_{t+k}-\bar{Y}\right)/n-k}{\sum\left(Y_t-\bar{Y}\right)^2\backslash n} \tag{6}$$

For computational purposes, the autocorrelation at lag k is based on:

$$r_k = \frac{\sum\left(Y_t-\bar{Y}\right)\left(Y_{t+k}-\bar{Y}\right)}{\sum\left(Y_t-\bar{Y}\right)^2} \tag{7}$$

Where:
$r_k$ = The value of autcorrelation at lag k
$\bar{Y}, Y_t$ and $Y_{t+k}$ = Respectively the mean of the series, the value of the time series at
t = The value of time series at lag k
n = The number observations of the series

The Partial Autocorrelation Function (PACF) is usually denoted as which is mathematically represented as:

$$\phi_{kk} = \begin{cases} r_k & \text{at } k = 1 \\ \dfrac{r_k - \sum_{j=1}^{k-1} \phi_{j i_{k-1}} r_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{j i_{k-1}} r_k}, & \text{at } k = 2 \end{cases} \tag{8}$$

where, $r_k$ is the autocorrelation at lag k. The generalized AR(p) Model is represented in Eq. 9:

$$y_t = \alpha + \sum_{i=1}^{p} \vartheta_i y_{t-i} + u_t \tag{9}$$

p = the number of lagged values of the regressand included in the model (9) is the regression equation of the actual series on its past values that is the series is explained by its past values. Also, the MA is represented in Eq.10:

$$y_t = \pi + \sum_{j=1}^{q} d_j u_{t-i} + d_0 u_t \quad (10)$$

q = the number of lagged values of the error term in the model (10) is the regression equation of the actual series on its error past values that is the series is explained by its error past values (9) and (10) can be combined to produce an ARMA Model as represented in (11)[25]:

$$y_t = \varphi + \sum_{i=1}^{p} \vartheta_i y_{t-1} + \sum_{j=1}^{q} d_j u_{t-i} + d_0 u_t + u_t \quad (11)$$

Hence, when the time series data is nonstationary at level, then (11) becomes ARIMA (p, d, q) with d been the integration order.

The best ARMA/ARIMA Model can be chosen from the number of estimated tentative models based on the number of significant coefficients, the volatility of the model, adjusted R-square and the information criterion such as Akaike Information Criterion (AIC), Bayesian Information Xriterion (BIC), etc. AIC and BIC can be mathematically computed using the equations:

$$AIC = -2\ln L + 2k \quad (12)$$

$$BIC = -2\ln L + 2k\ln(n) \quad (13)$$

where, InL is the natural logarithm of the estimated likelihood function k = p+q is the number of parameters in the model and n is the total observations. The latter maximizes the log-likelihood function to penalize for over parameterized model while the former estimates the expected Kullback-Leibler divergence which the closeness of the candidate model to the actual. Generation of the future values of time series through the development of statistical models for stochastic simulation purposes can be referred to as forecasting[26]. Thus, predicting the future value of a time series variable is based on the use of the scientific method.

## MATERIALS AND METHODS

Monthly data on the basic fuel prices (in-land basic cost, IBLC) of diesel, paraffin and petrol and exchange rate (ZAR-USD) for the period of 1986-2020 were collected from the energy statistics of the Department of Mineral Resources and Energy South Africa[27]. The variables are graphically represented in a basic time series plot to examine the various components present. The appropriateness of the ARMA model is studied using the ACF and PACF and the order of integration is determined using the Augmented Dickey-Fuller test[22]. The best models are chosen from the estimated models using each model’s volatility, adjusted $R^2$, the number of the significance of coefficients and information criterion values. The selected ARIMA model is diagnosed as required by the theory. The forecast of the future values of each variable is done using the best model. Eviews 11 is used for all the analyses, recommendations and conclusions are based on the discussions of results (Fig. 1).

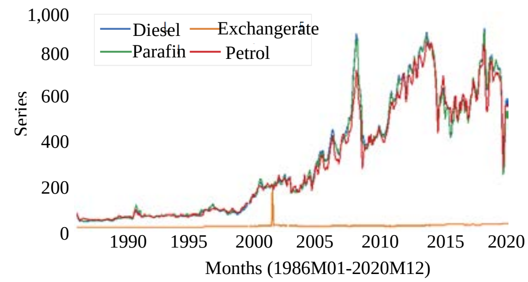



Fig. 1: Time series plots of the fuel prices and exchange rate

## RESULTS AND DISCUSSION

The time series plot of original road fatality data shown in Fig. 1 shows that there is trend variation in the series as the series mean changes systematically slowly with time. Though this is an indication of nonstationarity in the series but we cannot conclude based on the time series plot. Hence, the Augmented Dickey-Fuller test is required to check for the presence of the unit root in the series.

The results of the ADF unit root tests for both level and differenced series of fuel prices and exchange rate are shown in Table 1. The results show that the level series of fuel prices has a unit root while their first differenced has no unit root but the level series of the exchange rate has no unit root. However, South African annual exchange rate was found to be integrated of order in another study when annual data from 1960-2019 were used as repoted by the Augmented Dickey-Fuller and Phillips-Peron tests. The latter contradicts the   Hence, the exchangerate is integrated of order zero, I(0) while each fuel price variable is integrated of the first order I (1).

The correlogram plots of the Auto Correlation Functions (ACF) and Partial Autocorrelation Functions (PACF) for the first differenced diesel, paraffin and petrol and level of exchange rate series are plotted. These plots revealed that data can be modeled by an Auto Regressive Integrated Moving Average (ARIMA). Few lags are found to be significant for both the ACFs and PACFs, thus various models are estimated. Tables 2-5 show the results of tentative estimated ARIMA/ARMA Models.

The p-values for the intercept, autoregressive and moving average are reported in Table 2-5 for the nine tentative models estimated for each variable. In the diesel tentative models, 5 out of 9 have two coefficients significant at 1% level. Among these five models,

Table 1: Augmented Dickey-Fuller test results $H_0$ Series has a unit root

| Series | Model | Diesel | Exchange rate | Paraffin | Petrol |
|---|---|---|---|---|---|
| Level | Int.+trend | 0.1710 | 0.00000^* | 0.1373 | 0.3746 |
| Ist Difference | Int.+trend | 0.0000^* | NA | 0.0000^* | 0.0000^* |

Reseacher's computations from Eviews 11. * means significance at 1% level

Table 2: Results of tentative ARIMA (p, d, q) models for diesel

| Factors | ARIMA (1,1,1) | ARIMA (1,1,2) | ARIMA (1,1,3) | ARIMA (2,1,1) | ARIMA (2,1,2) | ARIMA (2,1,3) | ARIMA (3,1,1) | ARIMA (3,1,2) | ARIMA (3,1,3) |
|---|---|---|---|---|---|---|---|---|---|
| Cons. | 0.6547 | 0.6041 | 0.6107 | 0.6132 | 0.0133 | 0.4267 | 0.5831 | 0.4393 | 0.4644 |
| AR | 0.5658 | 0.0000 | 0.0000 | 0.0014 | 0.0000 | 0.0153 | 0.0000 | 0.0000 | 0.4867 |
| MA | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0107 | 0.6264 |
| Volatility | 887.61 | 882.94 | 924.71 | 881.18 | 1066.47 | 1058.15 | 857.13 | 1058.65 | 1065.21 |
| Adj. $R^2$ | 19.24% | 19.67% | 15.87% | 19.83% | 2.97% | 3.70% | 22.02% | 3.70% | 3.10% |
| AIC | 9.6469 | 9.6416 | 9.6875 | 9.6397 | 9.8314 | 9.8281 | 9.6123 | 9.8222 | 9.8284 |
| SC | 9.6856 | 9.6803 | 9.7262 | 9.6783 | 9.8707 | 9.8604 | 9.6510 | 9.8609 | 9.8671 |

Reseacher's computations using Eviews 11

Table 3: Results of tentative ARIMA (p, d, q) models for paraffin

| Factors | ARIMA (1,1,1) | ARIMA (1,1,2) | ARIMA (1,1,3) | ARIMA (2,1,1) | ARIMA (2,1,2) | ARIMA (2,1,3) | ARIMA (3,1,1) | ARIMA (3,1,2) | ARIMA (3,1,3) |
|---|---|---|---|---|---|---|---|---|---|
| Cons. | 0.6954 | 0.6525 | 0.6419 | 0.6556 | 0.5408 | 0.4845 | 0.6153 | 0.5003 | 0.5002 |
| AR | 0.2894 | 0.0000 | 0.0000 | 0.0000 | 0.6482 | 0.0137 | 0.0000 | 0.0000 | 0.8931 |
| MA | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.2416 | 0.0000 | 0.0000 | 0.0114 | 0.0910 |
| Volatility | 924.04 | 920.13 | 958.40 | 915.90 | 1142.20 | 1109.89 | 879.62 | 112.44 | 1118.22 |
| Adj. $R^2$ | 5.6% | 21.01% | 17.72% | 21.37% | 1.90% | 4.70% | 24.49% | 4.50% | 4.00% |
| AIC | 9.6872 | 9.6829 | 9.7236 | 9.6784 | 9.8980 | 9.8696 | 9.6385 | 9.8718 | 9.8772 |
| SC | 9.7259 | 9.7216 | 9.7622 | 9.7171 | 9.9367 | 9.9083 | 9.6771 | 9.9105 | 9.9158 |

Resarcher's computations using Eviews 11

Table 4: Results of tentative ARIMA (p, d, q) models for petrol

| Factors | ARIMA (1,1,1) | ARIMA (1,1,2) | ARIMA (1,1,3) | ARIMA (2,1,1) | ARIMA (2,1,2) | ARIMA (2,1,3) | ARIMA (3,1,1) | ARIMA (3,1,2) | ARIMA (3,1,3) |
|---|---|---|---|---|---|---|---|---|---|
| Cons. | 0.6879 | 0.5526 | 0.6508 | 0.6242 | 0.2974 | 0.3952 | 0.6497 | 0.4007 | 0.1571 |
| AR | 0.1736 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0000 |
| MA | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Volatility | 1010.24 | 973.74 | 1093.68 | 987.99 | 1138.98 | 1132.18 | 984.05 | 1133.92 | 1173.85 |
| Adj. $R^2$ | 19.24% | 22.15% | 12.57% | 21.81% | 8.90% | 9.45% | 21.33% | 9.30% | 6.20% |
| AIC | 9.7765 | 9.7401 | 9.8552 | 9.7443 | 9.8965 | 9.8901 | 9.7505 | 9.8915 | 9.9274 |
| SC | 9.8152 | 9.7788 | 9.8939 | 9.7830 | 9.9351 | 9.9288 | 9.7892 | 9.9302 | 9.9661 |

Researcher's computations using Eviews 11

Table 5: Results of tentative ARIMA (p, d, q) models for exchange rate

| Factors | ARIMA (1,0,1) | ARIMA (1,0,2) | ARIMA (1,0,3) | ARIMA (2,0,1) | ARIMA (2,0,2) | ARIMA (2,0,3) | ARIMA (3,0,1) | ARIMA (3,0,2) | ARIMA (3,0,3) |
|---|---|---|---|---|---|---|---|---|---|
| Cons. | 0.5143 | 0.0002 | 0.0003 | 0.0004 | 0.3076 | 0.0002 | 0.0001 | 0.0002 | 0.1780 |
| AR | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| MA | 0.0000 | 0.0004 | 0.0000 | 0.0000 | 0.0000 | 0.0007 | 0.0013 | 0.0005 | 0.0000 |
| Volatility | 90.68 | 99.11 | 99.16 | 98.73 | 92.71 | 98.99 | 99.19 | 99.01 | 94.06 |
| Adj. $R^2$ | 11.70% | 3.5% | 3.40% | 3.95% | 9.78% | 3.45% | 3.40% | 3.6% | 8.40% |
| AIC | 7.3665 | 7.4535 | 7.4541 | 7.4498 | 7.3892 | 7.4523 | 7.4544 | 7.4526 | 7.4039 |
| SC | 7.4051 | 7.4921 | 7.4923 | 7.4883 | 7.4278 | 7.4910 | 7.4930 | 7.4912 | 7.4425 |

Researcher's computations using Eviews 11

Table 6: Model adequacy results

| Factors | Diesel ARIMA (3,1,1) | Paraffin ARIMA (3,1,1) | Petrol ARIMA (1,1,2) | Exchange rate ARIMA (1,0,1) |
|---|---|---|---|---|
| RMSE | 54.9109 | 59.5511 | 63.9706 | 9.552 |
| MAE | 40.2241 | 42.6224 | 47.785 | 1.6848 |
| MAPE | 6.6112 | 8.3204 | 8.547 | 20.7097 |
| Symm. MAPE | 6.4909 | 8.0652 | 8.2421 | 17.9807 |

Appendices I-IV

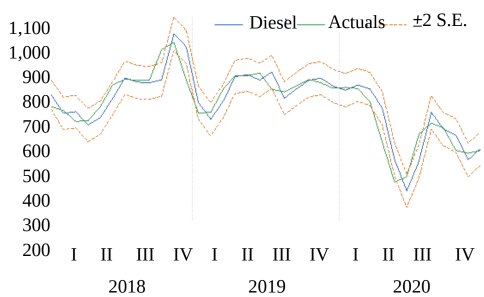
ARIMA (3, 1, 11) is found to be the best as it has the lowest volatility value of the highest adjusted coefficient of variation of and the least Akaike information criterion value of (Table 2). And in the paraffin tentative models, 4 out of 9 have two coefficients significant at one percent level. Among these four models, ARIMA (1, 1, 2) is found to be the best as it has the lowest volatility value of 90.67 highest adjusted coefficient of variation value of and the least Akaike information criterion value of (Table 3). However, in the petrol tentative models, 8 out of 9 have two coefficients significant at one percent level. Among these eight models, ARIMA (1, 0, 1) is found to be the best as it has the lowest volatility value of , highest adjusted coefficient of variation value of and the least

Akaike information criterion of  (Table 4). Also in the exchange rate tentative models 8 out of 9 have two coefficients significant at one percent level. Among these eight models, ARIMA (1, 0, 1) is found to be the best as it has the lowest volatility value of 90.67 highest adjusted coefficient of variation value of  and the least Akaike information criterion of  (Table 5). The four selected models are diagnosed through the correlogram of their residuals with each correlogram indicating that all information has been captured, hence the forecast will be based on these selected models, ARIMA (3, 1, 1), ARIMA (3, 1, 1), ARIMA (1, 1, 2) and ARIMA (1, 0, 1).

Table 6 shows the results of the forecast models for each variable over a period of 36 months (2018M01 to 2020M12). ARIMA (3, 1, 1) diesel forecast shows a slight deviation from the actual series, with the forecast accuracy of 93.4%. Also, ARIMA (3, 1, 1) paraffin forecast shows a slight deviation from the actual series, with a forecast accuracy of 9137%. ARIMA (3, 1, 1) petrol forecast shows a slight deviation from the actual series, with the forecast accuracy of ARIMA (3, 1, 1 exchangerate forecast shows a little or no deviation from the actual series with a forecast accuracy of . A very strong comovement is seen among the forecast and actual series or each variable (Appendices 1-5).

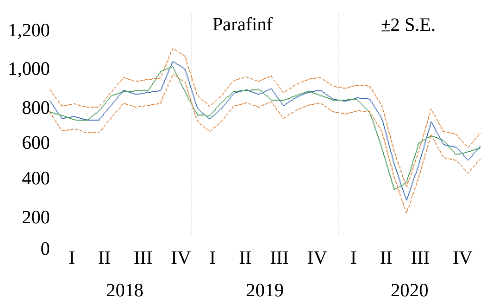
## CONCLUSION

The 35 years monthly series of diesel, paraffin and petrol in South Africa show an upward trend component with no seasonal component, thus these series are differenced to remove the trend and each is found to be integrated of order 1, . However, the exchange rate shows no trend variation and is stationary at level, . Many lags of the first differenced are found to be significant from the correlogram of both the ACF and PACF, however, the parsimonious models are estimated and checked for the best models based on the number of the significant coefficients; the lowest volatility; highest adjusted coefficient of determination and the lowest information criterion. ARIMA (3, 1, 1), ARIMA (3, 1, 1), ARIMA (1, 1, 2) and ARIMA (1, 0, 1) are respectively found to be the best predictive models for diesel, paraffin, petrol and exchange rate in South Africa. The residuals correlogram plots of the four models are found to be flat, indicating that all the information has been captured. Each of ARIMA (3, 1, 1), ARIMA (3, 1, 1), ARIMA (1, 1, 2) and ARIMA (1, 0, 1) respectively has  and 93.4%, 91.7%, 91.5% and 79.3%  forecast accuracy. These models can be used by the policymakers and all fuel and exchange rate stakeholders to predict the future values of the in-land basic cost of fuel per litre and the strength of South African rand against the United States of America dollars (ZAR-USD). This is necessary as fuel plays important role in growing the South African economy[13]. Further research can be carried out on the short and long-run relationship between the four variables. Also, the direction of causality among the variables should be looked into to determine which variable can be used to determine the future values of the other.



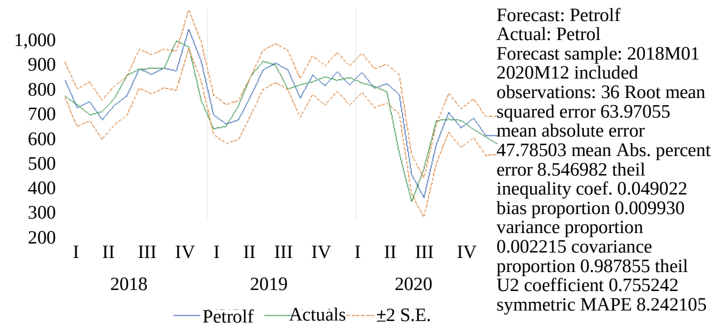

Forecast: Diesel Actual: Diesel
Forecast sample: 2018M01 2020M12 included observations: 36 Root mean squared error 54.91094 mean absolute error 40.22411 mean Abs. percent error 60611181 theil inequality coef. 0.039736 bias proportion 0.008619 variance proportion 0.001058 covariance proportion 0.990322 theil U2 coefficient 0.728635 symmetric MAPE 6.490942

Appendix 1: ARIMA (3, 1, 1) forecasted and actual diesel series; Researcher's computation from Eviews 11

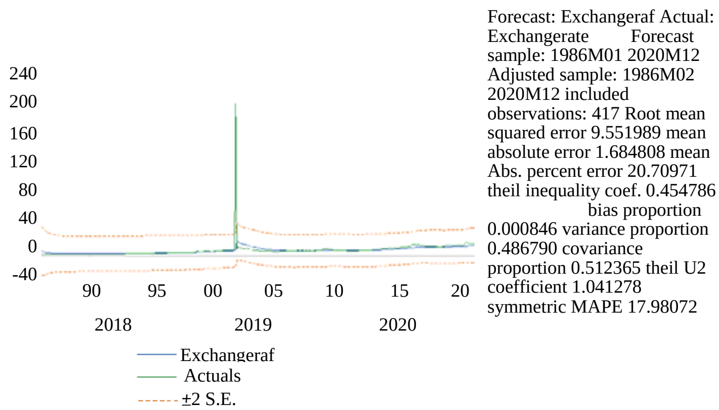



Forecast: Parafinf Actual: Parafinf Forecast sample: 2018M01 2020M12 included observations: 36 Root mean squared error 59.55110 mean absolute error 42.62244 mean Abs. percent error 8.320401 theil inequality coef. 0.043725 bias proportion 0.009599 variance proportion 0.001141covariance proportion 0.989260 theil U2 coefficient 0.737882 symmetric MAPE 8.065213

Appendix 2: ARIMA (3, 1, 1) forecasted and actual diesel series; Researcher's computation from Eviews 11




Appendix 3: ARIMA (1, 1, 2) forecasted and actual diesel series; Researcher's computation from Eviews 11




Appendix 4: ARIMA (1, 0, 1) forecasted and actual diesel series; Researcher's computation from Eviews 11

**ACKNOWLEDGMENTS**

The researchers acknowledge the support of the Department of Mineral Resources and Energy South Africa for providing the data used in this study.

**REFERENCES**

01. Olayiwola, M.F. and S.M. Seeletse, 2020. Statistical forecasting of petrol price in South Africa. J. Eng. Applied Sci., 15: 602-606.
02. Maia, J. and S. Hanival, 1994. An overview of the performance of the South African economy since 1994. The Presidency Office, Pretoria, South Africa.
03. Black, A. and S. Bhanisi, 2006. Globalisation, imports and local content in the South African automotive industry. Proceedings of the Development Policy Research Unit Conference, October 18-20, 2006, The Birchwood Hotel ands Conference Centre, Johannesburg, South Africa, pp: 18-20.
04. StatsSA., 2017. Statistics South Africa annual report. Pretoria Government Printers, Pretoria, South Africa.
05. Glenday, G., 2008. South African tax performance: Some perspectives and international comparisons. National Treasury of South Africa at the CSIR Convention Centre, Pretoria, South Africa.
06. Leggett, J.K., 2005. Half Gone: Oil, Gas, Hot Air and the Global Energy Crisis. Portobello Books, London, England, UK., ISBN: 9781846270048, Pages: 312.
07. Statistics South Africa, 2009. Energy statistics. Statistics South Africa, Pretoria, South Africa.
08. Serletis, A., 1994. A cointegration analysis of petroleum futures prices. Energy Econ., 16: 93-97.
09. Gjolberg, O. and T. Johnsen, 1999. Risk management in the oil industry: Can information on long-run equilibrium prices be utilized?. Energy Econ., 21: 517-527.
10. Girma, P.B. and A.S. Paulson, 1999. Risk arbitrage opportunities in petroleum futures spreads. J. Future. Markets, 19: 931-955.

11. Sukati, M., 2013. Cointegration analysis of oil prices and consumer price index in South Africa using STATA software. MPRA Paper No. 49797, University Library of Munich, Germany.
12. Trading Crude Oil CFDs Online, 2020. Plus 500TM crude oil trading. Plus500AU Pty Ltd, South African.
13. Olayiwola, M.F. and S.M. Seeletse, 2020. South African petrol price and consumption cointegration analysis. J. Eng. Applied Sci., 15: 1579-1588.
14. Yang, C.W., M.J. Hwang and B.N. Huang, 2002. An analysis of factors affecting price volatility of the US oil market. Energy Econ., 24: 107-119.
15. Olayiwola, M.F., 2020. Mathematical modeling of road traffic fatalities. Afr. J. Math. Stat. Stud., 3: 63-74.
16. Okyere, F., F. Mahama, S. Yemidi and E. Krampa, 2015. An econometric model for inflation rates in the volta region of Ghana. IOSR J. Econ. Finance, 6: 48-55.
17. Makwinja, R., T. Phiri, I.B. Kosamu and C.C. Kaonga, 2017. Application of stochastic models in predicting Lake Malawi water levels. Int. J. Water Resour. Environ. Eng., 9: 191-200.
18. Box, G. and Jenkins, 1976. Time Series Analysis and Control. 1st Edn., Holden-Day San Francisco, California, ISBN: 0-8162-1104-3.
19. Olayiwola, M.F. and S.M. Seeletse, 2014. Statistical modelling of petroleum price and consumption rate in South Africa. M.Sc. Thesis, University of Limpopo South Africa.
20. Kantz, H. and T. Schreiber, 2004. Nonlinear Time Series Analysis. 2nd Edn., Cambridge University Press, London, UK., ISBN: 9780521529020, Pages: 369.
21. Box, G.E.P. and D.A. Pierce, 1970. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. J. Am. Stat. Assoc., 65: 1509-1526.
22. Dickey, D.A. and W.A. Fuller, 1978. Distributional estimators for autoregressive time series with unit root. J. Am. Stat. Assoc., 74: 427-431.
23. Phillips, P.C.B. and P. Perron, 1988. Testing for a unit root in time series regression. Biometrika, 75: 335-346.
24. Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin, 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? J. Econometrics, 54: 159-178.
25. Cochrane, J.H., 1997. Time series for macroeconomics and finance. University of Chicago, Chicago, Illinois.
26. Woodward, W., H. Gray and A. Elliott, 2012. Applied Time Series Analysis. CRC Press, Boca Raton, Florida,.
27. DME., 2020. Energy statistics. Department of Minerals and Energy, South Africa.