ANALYSIS OF THE VOLATILITY OF THE ELECTRICITY COSTS IN KENYA USING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL

Mohammed Mustapha Wasseja and Samwel N Mwenda

Kenya National Bureau of Statistics

ABSTRACT: *Electricity has proved to be a vital input to most developing economies. As the Kenyan government aims at transforming Kenya into a newly-industrialized and globally competitive, more energy is expected to be used in the commercial sector on the road to 2030. Therefore, modelling and forecasting of electricity costs in Kenya is of vital concern. In this study, the monthly costs of electricity using Autoregressive Integrated Moving Average models (ARIMA) were used so as to determine the most efficient and adequate model for analysing the volatility of the electricity cost in Kenya. Finally, the fitted ARIMA model was used to do an out-off-sample forecasting for electricity cost for September 2013 to August 2016. The forecasting values obtained indicated that the costs will rise initially but later adapt a decreasing trend. A better understanding of electricity cost trend in the small commercial sector will enhance the producers make informed decisions about their products as electricity is a major input in the sector. Also it will assist the government in making appropriate policy measures to maintain or even lowers the electricity cost.*

KEYWORDS: Electricity, ARIMA, forecast, Kenya

INTRODUCTION

Inadequate electricity generation capacity and an unreliable power supply have been perennial problems in Kenya for over a decade. As in other African countries, lack of integration between planning and implementation has plagued the industry. For example, a 328MW expansion in capacity planned for the period 1996 to 2000 saw only 205 MW actually being installed, leaving a significant shortfall [1].

As Kenya aspire to be a middle income economy as envisaged in Vision 2030, it faces an enormous task of meeting energy needs due to the high expectation in growth, to power the economy. The country therefore needs to come up with strategies and investment plans to secure sustainable supply of energy especially electricity to meet the growing demand, since it is considered a key enabler to achieving vision 2030.Since electricity is a necessary input for many sectors in the economy, a proper and timely forecasting of electricity costs can affect the base for many business entities.

Electricity plays a major role in the socio-economic development of the nation. It's the cheapest source of energy compared with the other sources and has also proved to be extremely efficient and cost effective. In Kenya, electricity is mostly used in commercial sector, manufacturing and processing, construction, agricultural and residential sectors among many other purposes which are sourced from hydro, thermal oil and geothermal sources. The installed capacity by 2007 were 677.3, 389.3 and 128 Megawatts for hydro, thermal oil and geothermal, respectively.

Hydropower has long dominated Kenya's generating capacity, in 2010; it supplied almost 57% of the country's electricity. However, severe droughts in the 1990s had virtually paralysed the industry; with the 1999 drought (the worst since 1949), leading to a 79% decrease in hydro

capacity between July and December 2000 [1].Power cuts were widespread and commerce and industry suffered significant losses. When the 1999 drought exacerbated existing problems, an additional 105 MW of emergency thermal generation was deployed in 2000.These two interventions proved costly however, the price of electricity rose considerably and has remained high ever since. This has led to a more fundamental reform of the electricity sector which has led to establishment of an independent regulator, and unbundling and liberalisation of the electricity sector. Also the ever fluctuating fuel prices and poor performance of Kenyan currency in forex markets has led to high cost volatility. Exploitation of small scale hydro electricity is impeded by high installation cost, inadequate hydrological data, effects of climate change and limited local capacity to manufacture small hydro power

Problem Statement

The electric power industry in many countries all over the world is moving from a centralized operational approach to a competitive one. The understanding of electric power supply as a public service is being replaced by the notion that a competitive market is more appropriate mechanism to supply energy to consumers with high reliability and low cost. Financial contracts to hedge against risks of price volatility are possible and advisable but they do not affect the physical operation of the system. The power firm usually passes on costs accrued due to rise on fuel prices as well as instability in the local currency to consumers. Kenya Power passes the two charges to its clients without any attempts to shield them from adverse effects. They change frequently due to the volatile nature of oil prices and the Kenyan currency, making the power bills unpredictable. Kenya heavily relies on power produced using hydro dams and during dry spells the country is forced to increase the share of electricity generated using diesel. In the framework of competitive electricity markets, power producers and consumers need accurate price forecasting tools. Cost forecasts embody crucial information for producers and consumers when planning bidding strategies in order to maximize their benefits and utilities, respectively. This study therefore aimed at coming up with an effective and efficient model for forecasting the volatile electricity cost to enhance producers and consumers make appropriate decisions.

General Objective of the study

The general objective of the study was to forecast electricity cost in Kenya using the Autoregressive Integrated Moving Average (ARIMA) model.

Specific Objectives

i. To fit an ARIMA model to the data on electricity cost in Kenya.

ii. To evaluate the efficiency of the ARIMA model on fitting the electricity cost in Kenya.

iii. To forecast electricity cost in Kenya using the ARIMA model.

LITERATURE REVIEW

Introduction

This chapter outlines the theoretical literature review on electricity costs, Kenyan electricity industry, Kenya Power (KP), Kenya electricity Generating Company (KenGen), Independent

Power Producers (IPPs), electricity regulation, Energy Regulatory Commission (ERC). It also focus on the effect of electricity cost to the economy, electricity cost and inflation, the ARIMA model, the Box-Jenkins methodology, conceptual framework and model application.

Electricity Costs Forecast

Forecasting is the process of making predictions about an event whose outcome is unknown at present. It's potentially useful in decision aids and economists have devoted considerable efforts to develop assess forecasting methods[2]. In today's business world, forecasting is of significant importance and interest for most of the companies since an accurate forecast can provide any business entity in a given market a strategic advantage over its competitors. It is the main tool for decision making. With the structural changes in the power industry over the past 10 years, the market price of electricity has become the key determinant of resource value[3].Very little though has been done to forecast short time electricity cost.

Kenya's electricity industry

The structure of Kenya's electricity-supply industry can be traced back to the reforms that swept the industry in the mid 1990s. A policy paper on economic reforms [4]set out the government's intention to separate the regulatory and commercial functions of the sector, facilitate restructuring and promote private-sector investment. Consequently the Electric Power Act of 1997 was passed. The government's primary function, through the Ministry of Energy, became policy formulation, and its regulatory authority was devolved to the newly established Electricity Regulatory Board (ERB) which became functional in 1998[5]. At industry level, rationalisation and unbundling translated into a redefinition of the scope of KP's activities such that it now focuses only on the transmission and distribution (including supply) of electricity, while KenGen, established in 1997, has taken over all the generation activities that KP had previously managed. While a number of IPPs have subsequently entered the industry, KenGen and KPLC (both loosely referred to as state-owned entities) remain the dominant players.

In addition, in late 2008, the government incorporated the Kenya Electricity Transmission Company (KETRACO) to construct, operate and maintain new, publicly funded, high-voltage electricity transmission networks.

Electricity cost and inflation

Inflation is the persistent increase in the prices of commodities with time. Inflation in Kenya has resulted in an increase the prices of commodities and these coupled with the high cost of electricity is simply beyond the reach of many Kenyans. Many companies had resulted to hiking the prices of their products in order to meet the rising electricity costs, a cost that is being passed on directly to the consumers. The high electricity costs had greatly increased the cost of doing business in the Kenya making its products quite uncompetitive both at the regional and international level

Effects of Electricity Cost to the Economy

Electricity has proved to be a vital input to economy development and growth. Electric power in Kenya is used in the agricultural sector, industrial sector, residential purposes and many others to run the machines. High electricity cost has detrimental effects on the country's manufacturing sector which is considered to be the main engine behind Kenya's economic growth. It is a fact that over the years, Kenya has become very uncompetitive with regard to

the cost and provision of power. This is the reason why the recent president's directive was well over due. The President took it upon himself to issue a directive to the Energy and Finance ministries to review taxes and levies on electricity in a bid to reduce the electricity costs whose disastrous effects were already being felt on the country's economy[6].The current costs are simply outrageously high and unaffordable by all. Given the extent of the problem, action should be taken immediately to save Kenya's economy from collapsing.

RESULTS AND DISCUSSIONS

Analysis was done using GRETL programme for Econometric and Time-series software Version 1.9.13

The ARIMA Model

ARIMA model is one of the time series models used in forecasting. This model unlike others does not assume knowledge of any underlying economic model or structural relationships. It is assumed that values of the past series and the previous disturbance terms have information for the purposes of forecasting. The major advantage of ARIMA is that it requires data on the time series in question only. The disadvantages of the model are;

     i.   It is necessary to have a great number of basic data.

     ii.   There is no simple method of adjustment of parameters of ARIMA model.

     iii. Also for estimation one or another model is used and it means the risk in calculations and therefore it is necessary to do periodic verification of adequacy of the applied model.

     iv. The model depends on historic information.

Prerna (2012) provides that the ARIMA model is a modification of the ARMA model. ARMA models represent a combination of the Auto Regressive (AR) and Moving Average models[7]. The application of the ARIMA model to time series analysis is due to Box and Jenkins (1976).The ARIMA model will be applied in this study to forecast the electricity cost using monthly costs data from Kenya Power (KP) in conjunction with Kenya National Bureau of Statistics (KNBS)

The study employed the Box-Jenkins (1976) methodology that involves stages of identification, estimation, diagnostic checking and forecasting of a univariate time series[8]. An exploratory research design was adopted for a sample of 58 observations. Purposive sampling method was used to obtain this sample. The monthly data was obtained from Kenya power, Energy Regulatory Commission and the Kenya National Bureau of Statistics (KNBS) for the period November 2008 to August 2013. The Results indicated that the Autoregressive Integrated Moving Average models (ARIMA 2,1,2) was the best model to fit the electricity data since it yielded the least values for the Schwartz Bayesian Criterion (SBC), Akaike Information Criterion (AIC), Mean Square Error (MSE), Log Likelihood and Hannan-Quinn Tests criterion and therefore outperformed other models in forecasting

Model Identification

This stage involves the use and observing the behaviours of autocorrelation function (AFC) and partial autocorrelation function (PAFC) to more precisely determine if the time series is stationary or not and tentatively identify a Box-Jenkins model resulting from correlograms (Al-Zeaud, 2011). Correlograms are the plots of the ACFs and PACFs against lag lengths. The autocorrelation coefficient (ACF) measures the correlation between a set of observations and a lagged set of observation in a time series. The sample autocorrelation r_k coefficient is given by,

$$r_k = \frac{\sum(Y_t - \hat{Y})(Y_{t+k} - \hat{Y})}{\sum(Y_t - \hat{Y})(Y_t - \hat{Y})}$$

Where, Y_t = The data from the stationary time series

Y_{t+k} = The data from k time period ahead of t

$\hat{Y}$ = The mean of the stationary time series

The estimated partial autocorrelation function (PACF) is used as a guide together with the estimated autocorrelation function in choosing one or more ARIMA models that fit the data available. The idea of Partial autocorrelation is that we want to measure how Y_t and Y_{t+k} are related. If a graph of the ACF of the time series values either cuts off fairly quickly or dies down fairly quickly, then the time series should be considered stationary. If the ACF graph dies down extremely slowly, then the time series values should be considered non-stationary. If the series is not stationary, it needs to a logarithmic transformation and then converting it to a stationary process by differencing. The behaviour of the various models on the ACF and PACF can be shown as follow

Table 1: Behaviour of ACF and PACF

Model	ACF	PACF
AR(p)	Dies down	Cuts off after lag q
MA(q)	Cuts off after lag p	Dies down
ARMA(p,q)	Dies down	Dies down

Estimation of parameter

Not only does the Box-Jenkins model have to be stationary, it also has to be invertible. Invertible means recent observations are more heavily weighted than more remote observations; the parameters $(\varphi_1, \varphi_2..., \varphi_p, \theta_1, \theta_2...\theta_p)$ used in the model decline from the most recent observations down to the further past observations. Having identified the appropriate p and q values, the next stage is to estimate the parameters of the auto regressive and moving average terms included in the model. To estimates the parameters, we can use the least squares method and the maximum likelihood.

Diagnostic Checks

This is a Jenkins method of checking whether or not the residuals ε_t were white noise. This is done by checking the Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) coefficients. The assumption made here is that the error term is independent and identically distributed according to a normal distribution. The various checks include; Akaike Information Criteria (AIC), Mean Squared Error (MSE) and Schwartz Bayesian Criteria (SBC). Akaike Information Criteria (AIC) can be generalised by the form

$AIC = 2k - 2ln(L)$

where k is the number of parameters in the model and L is the maximised value of the likelihood function for the estimated model. Given a set of models, the best model to choose is one with the minimum AIC value .The SBC which is also known as the Bayesian Information Criteria (BIC) formula is given as;

$-2.lnp(^x/_k) = BIC = -2.lnL + k\ ln(n)$

where x is the observed data, n is the number of observations, k is the number of parameters to be estimated, $p(x/k)$ is the probability of the observed data given the number of parameters and L is the maximised likelihood function for the estimated model. The Mean Squared Error (MSE) of an estimator $\acute{\emptyset}$ with respect to the estimated parameter $\emptyset$ is defined as;

$$MSE = \frac{\sum_{t=1}^{n}(\in_t)(\in_t)}{n} = \frac{\sum_{t=1}^{n}(y_{t+1} - \hat{y}_t)(y_{t+1} - \hat{y}_t)}{n}$$

Where n is the size of the time series data.

Forecasting using the Model

This stage involves getting the precise estimates of the coefficients of the ARIMA model in the identification stage. These estimates are obtained by fitting the time series data to the ARIMA model. This stage also provides some warning signs on the adequacy of the model whereby if the model does not satisfy the adequacy conditions, it is rejected.

Model Application

Various forecasts on electricity prices have been done but not much has been done in Kenya. Javier, Rosario, Francisco and Antonio (2003) used the ARIMA model to predicts next- day electricity prices in Spain[9]. Beatrice (2004) used the ARIMA model to forecast the sale of electricity in USA while Wu Min and Cao (2008) used ARIMA model to forecast the consumption of electricity in China, Xier (2009) used the ARIMA model to analyse monthly temperature of Stockholm. Smart (2010) used ARIMA models for modelling and forecasting maternal mortality in Ghana. Researchers have also participated in the forecasting of other energy models. These studies include; Prerna (2012) forecasted natural gas price in London using a time series and non-parametric approach, Syed *et al.*,(2012) investigated on the impact of oil prices on food inflation in Pakistan[10]. It is clear that little has been done in Kenya and thus this paper aims at exploring the under researched sector in the country.

Data Analysis

Gretl, an econometrics package, was used to run and analyse the time series data in order to obtain estimates that can be used to forecast future small commercial electricity cost. The raw data obtained from Kenya Power, the Energy Regulatory commission and the Kenya National Bureau of Statistics was entered into the software and run for the ARIMA process. For the data to be run, it had to undergo log transformation to make it easier to transform the data to stationarity. A comparison of various combinations of different ARIMA Models was then carried out to establish the best model to fit. The values of the parameter estimates for the model were obtained from the output of the run data. The analysis also involved drawing time plots, the ACF and descriptive statistics. The diagnostic test values; Akaike Information Criteria (AIC), Schwartz Bayesian Criteria (SBC), Hannan-Quinn and the Mean Squared Error (MSE) were used where the best model to fit was established. The final model was estimated by the Gaussian Likelihood estimation method, including the estimation of coefficients (φ_1, $\varphi2$, $\varphi3$, $\varphi4$.....). Using the model obtained, at 95% confidence interval, I forecasted the electricity cost for small commercial for the preceded 36-month period, i.e. September 2013 to August 2016.

Model Specification

This part focuses on the overall outlook of the Moving Average (MA) model, the Autoregressive (AR) Process and the Autoregressive Integrated Moving Average (ARIMA) process.

Moving Average model (MA)

Moving Average models were as a result of an invention by Slutsky in 1937[11]. MA process is one in which the current value of a time series depends upon current and past random error variables. In time series, the moving-average (MA) model is a common approach for modelling univariate time series models. The notation MA (q) refers to the moving average model of order q:

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

Where μ is the mean of the series, the $\theta_1 ... \theta_q$ are the parameters of the model and the ε_t, ε_{t-1}... are the white noise error terms. The value of q is called the order of the MA model. This can be equivalently written in terms of the backward shift operator B as

$$X_t = \mu + (1 + \theta_1 B + \cdots + \theta_q B^q)\varepsilon_t.$$

Thus, a moving-average model is conceptually a linear regression of the current value of the series against current and previous (unobserved) white noise error terms or random shocks. The random shocks at each point are assumed to be mutually independent and to come from the same distribution, typically a normal distribution, with location at zero and constant scale. Fitting the MA estimates is more complicated than with autoregressive models (AR models) because the lagged error terms are not observable. This means that iterative non-linear fitting procedures need to be used in place of linear least squares.

The moving-average model is essentially a finite impulse response filter applied to white noise, with some additional interpretation placed on it. The role of the random shocks in the MA model differs from their role in the AR model in two ways. First, they are propagated to future

values of the time series directly: for example, ε_{t-1} appears directly on the right side of the equation for X_t. In contrast, in an AR model ε_{t-1} does not appear on the right side of the X_t equation, but it does appear on the right side of the X_{t-1} equation, and X_{t-1} appears on the right side of the X_t equation, giving only an indirect effect of ε_{t-1} on X_t. Second, in the MA model a shock affects X values only for the current period and q periods into the future; in contrast, in the AR model a shock affects X values infinitely far into the future, because ε_t affects X_t, which affects X_{t+1}, which affects X_{t+2}, and so on forever.

An Autoregressive (AR) Process

Autoregressive (AR) models were first introduced by Yule in 1926 [11].An AR process is one in which the current value of a time series depends upon the past values of a time series and the random error. The notation AR (p) indicates an autoregressive model of order p. The AR (p) model is defined as

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t$$

Where $\varphi_1, \cdots, \varphi_p$ the parameters of the model are, c is a constant, and ε_t is white noise. This can be equivalently written using the backshift operator$B$ as

$$X_t = c + \sum_{i=1}^{p} \varphi_i B^i X_t + \varepsilon_t$$

So that, moving the summation term to the left side and using polynomial notation, we have

$$\phi(B)X_t = c + \varepsilon_t .$$

Auto-Regressive Integrated Moving Average model: ARIMA (p, d, q)

The ARIMA model is a modification of the ARMA models arrived at by transforming the time series data to Stationarity. ARIMA models became highly popular in the 1979s among academics, in particular when it was shown through empirical studies (Cooper, 1972) that they could outperform the large and complex econometric models, popular at that time, in a variety of situations. If we differentiate the time series data d, then the model becomes stationary (Ransang and Titida 2006). By applying ARMA (p, q) to it, we say the original time series has been transformed to stationarity and is referred to as an Auto Regressive Integrated Moving Average model (p, d, q) where p is the number of autoregressive terms, d is the number of non-seasonal differences and q is the number of lagged forecast errors in the prediction equation. Stationarity in the model can be detected by the use of the Auto Correlation Function graph (ACF).If a graph of the time series values either cuts off or dies down fairly quickly, then the times series should be considered stationary. If however the graph of the ACF dies down extremely slowly, then the time series values should be considered non-stationary. If the original time series is stationary, $d = 0$ and the ARIMA model reduces to ARMA model.

The difference linear operator (Δ) is given by;

$\Delta Y_t = Y_t - Y_{t-1} = Y_t - B Y_t = (1 - B) Y_t$

The stationary series W_t is obtained as the dth difference (Δ^d) of Y_t

$W_t = \Delta^d Y_t = (1 - B)^d Y_t$

The ARIMA model denoted as ARIMA (p, q, d) thus has the general form given by;

$\emptyset_p(B)(1-B)^d Y_t = \mu + \theta_q B \varepsilon_t$ or

$\emptyset_p(B) W_t = \mu + \theta_q B \varepsilon_t$

Preliminary Analysis

The preliminary analysis was done by use of time plots for the various series presented in Figures 1 and 2.
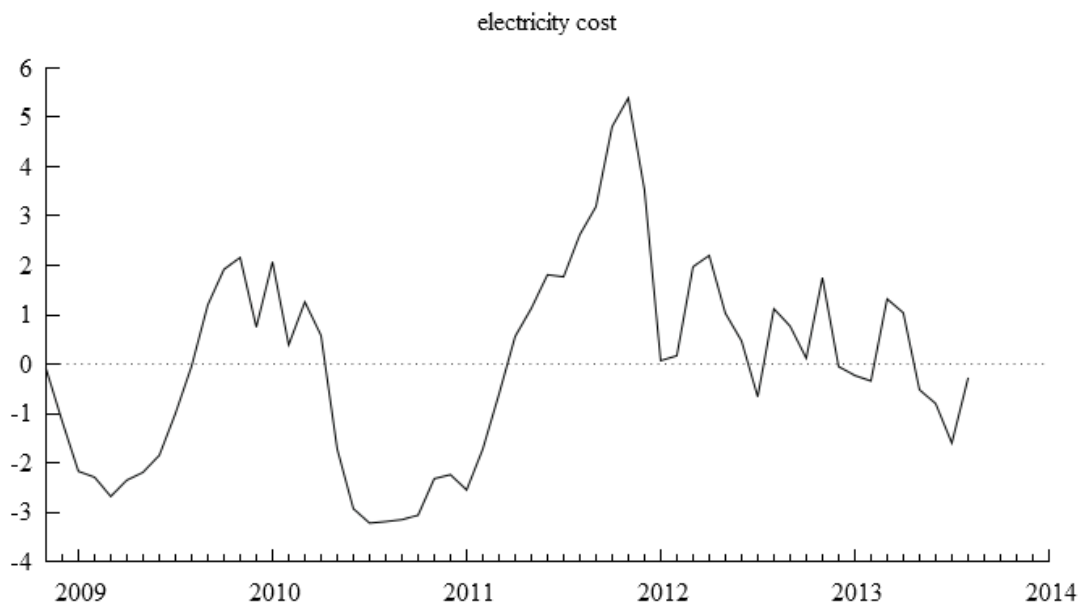

Figure 1: Time plot for the monthly average electricity cost for small commercial

By inspecting the time plot visually, the mean and variance are non-constant implying that the data is non-stationary. The non-constant mean and variance provides a suggestion of the utilization of a non-linear model.

The series was transformed to attain Stationarity by taking the first differences of the natural logarithms of the values in each series. The equation representing the transformation is given by $Y_t = ln(P_t) - ln(P_{t-1})$ where P_t represents the monthly value for each series. The resulting plots for the returns are presented below:

Figure 2: Time plots for log differenced cost

The time plot of the series now shows stationarity. The mean and the variance show the property of being constant.

Summary Statistics

The analysis involved used a total of 58 observations. The data displayed a sample mean of 18.0325861 and a standard deviation of 2.02269486.Standard deviation of 2.02269486 indicates that the electricity costs are fluctuating from the estimated sample mean by 2.02269486.

Summary statistics for the log-difference electricity cost data display features of a stationary time series.

The results are shown in Table 2.

Table 2: Summary Statistics for Log –Differenced Electricity cost

Statistics	Value
Minimum	-0.17493
Maximum	0.097558
Mean	-0.00019668
Standard Deviation	0.059285
Skewness	-0.53959
Kurtosis	0.05745

JarqueBera Statistics 2.77381

KPPS test   0.0494222

ADF test   -7.5894

The summary statistics for the Log-Differenced Electricity Cost series can be interpreted as follows. The mean value of -0.00019668 tends to zero which is an important characteristics of stationarity time series data. The standard deviation is 0.059285. The KPSS and Augmented Dickey Fuller Tests are used to verify whether or not the differenced series is stationary and whether or not there exists a unit root. The p-value (0.0494222) of the KPSS is smaller than 0.1 hence it rejects the null hypothesis and therefore the series is stationary. The p-value (-7.5894) of the ADF is smaller than 0.1 so it rejects the null hypothesis that there exists a unit root. From the above results, we find that the differenced series is stationary and there is no unit root. The differenced data had a kurtosis of 0.057245 and JarqueBera test of 2.77381, which suggested that the data was not normally distributed, it was negatively skewed with a skewness of -0.53959.This means that more observations were on the left hand side.

Auto-correlation Function (ACF)

The autocorrelation function also provides very useful information that is typical of a non-stationary process, where the autocorrelation declines slowly as the number of lags increases as indicated in the autocorrelation function (ACF) of electricity cost in figure 3.Random walk behaviour of time series is expected.
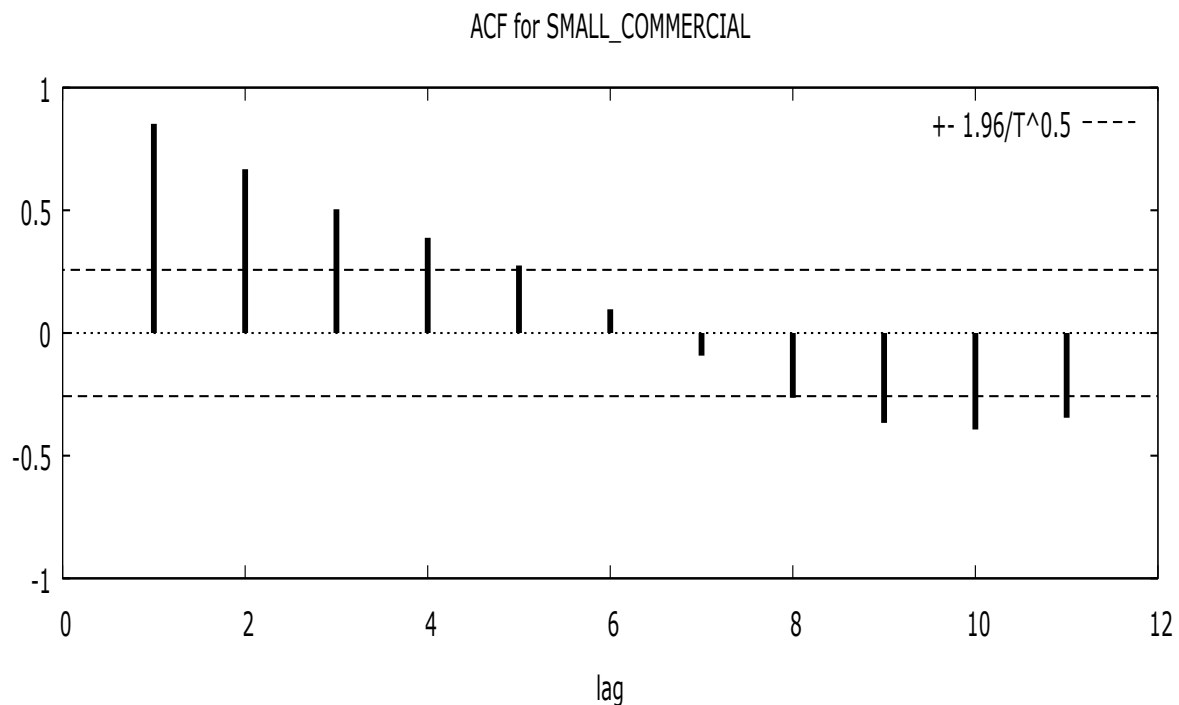

Figure 3: ACF for electricity cost

The correlograms in Figure 3 indicates the cycles gradually dying away, thus the series is non-stationary. To make the data stationary, we obtain the logarithm of the first order difference of the dependent variable (electricity cost). The plot for log-differenced cost time series data is shown in Figure 4.
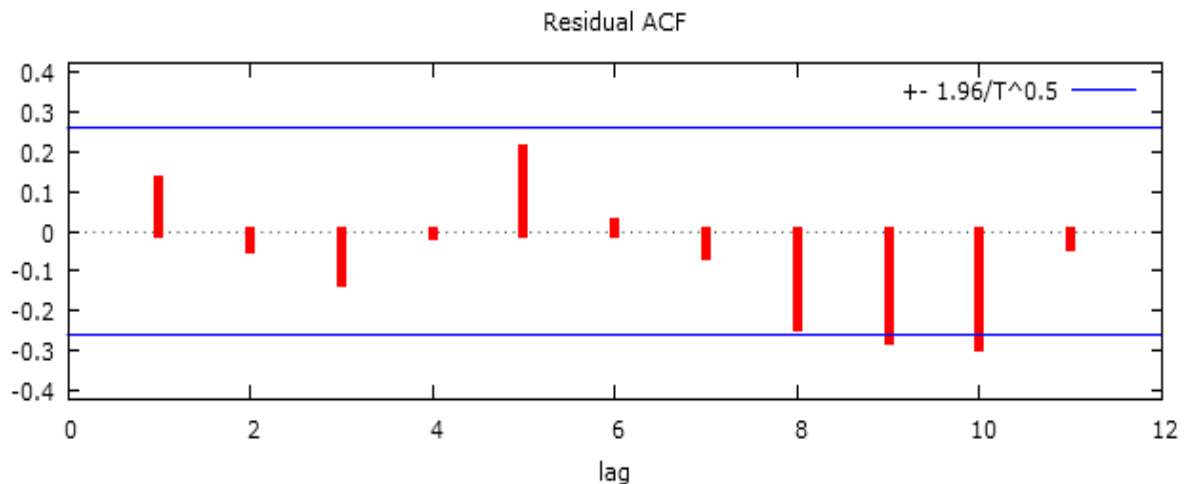

Figure 4: ACF of Log- Differenced costs

Figure 4 shows the ACF of the log differenced data. The autocorrelations here resemble white noise and hence depicts the property of a stationary series. By log differencing the electricity costs, stationarity of the cost was achieved.

Estimation Results

The results on Table 3 show the coefficient estimates of Autoregressive and Moving Averages schemes of electricity cost for small commercial sector in Kenya that resulted from modelling using Ordinary Least Square (OLS) method. All coefficients are not statistically significant at 5% level of significance.

Table 3: Results of Estimation

Dependent variable: Electricity Cost

Variable	Estimate	Std.Err	z Ratio	p-Value
Intercept	0.02171	0.0005455	0.2727	0.9529
AR1	-0.7876	0.1388	-5.676	0.00047
AR2	0.1044	0.138	0.7565	0.4493
MA1	-0.00701	0.07058	-0.027	0.9993
.MA2	-0.999997	0.07058	-14.17	0.0000041

Comparison with Other ARIMA models

The above model was compared with other models and using a model selection criterion that includes Log Likelihood, Schwarz criterion, Hannan-Quinn criterion and Akaike criterion. The best model is the one that displays lesser figures of the above named criterion. ARIMA (2, 1, 2) outperforms the other competing models as presented in Table 4.

Table 4: Comparison of various models

	ARIMA ARIMA	ARIMA	ARIMA	ARIMA	
ARIMA	(0,1,1)	(1,1,0)	(1,1,1)	(2,1,1)	(2,1,2)
Log Likelihood	76.741	68.7028	77.3672	77.4168	78.16747
Schwarz Criterion	-141.41	-139.33	-138.63	-134.71	-132.18
Hannan-Quinn Criterion	-145.13	-144.05	-143.59	-140.91	-139.62
Akaike Criterion	-147.48	-147.41	-146.73	-144.83	-144.33
MSE	0.2069	0.2810	0.2028	0.2024	0.1969
Ljung-Box Q(12)	20.9281	25.173	15.061	14.651	14.50

The Suitable Model

Using the lowest information criteria, the best model was chosen. ARIMA (2, 1, 2) model displays lesser results compared to other models and hence the best model to forecast electricity cost for small commercial in Kenya. Basing on the estimation result on Table 3, the model can be presented as:

$X_t = 0.02171 - 0.7876 X_{t-1} + 0.1044 X_{t-2} - 0.00701 e_{t-1} - 0.999997 e_{t-2} + e_t$

The model has an intercept value of 0.02171. Coefficients are not statistically significant at 5% significance level. The coefficient of X_{t-2} is positive, implying that they are positively correlated to their immediate formers, while the coefficients of $X_{t-1}, e_{t-1}, e_{t-2}$ are negative indicating negative correlation with their immediate formers.

Diagnostic Checking

The model was tested for normality .The statistics used to test this are JarqueBera statistic (1987), residual Skewness, residual kurtosis and the normal QQ plot. The JarqueBera statistic (2.77381) rejected the null hypothesis of normality in the residuals in the series. Residual kurtosis and Skewness should be zero for the normality condition series; however this is not the case in the results obtained. The series has a negative Skewness implying that most of the observations are to the left of the normal distribution.

For the normal QQ curve, the observation should be along the normal curve for the normality condition. According to the normal QQ plotted, the observations does not lie along the normal curve, hence the result obtained in this statistic are not normally distributed. The normal QQ curve is represented in Appendix 1.

Model independence was assessed by inspecting the sample autocorrelations of the residuals to see if they would resemble the white noise. Figure 6 shows that, the residuals approximate white noise and hence the model was independent.

The Durbin Watson Statistic is the most universally accepted testing criterion for measuring the magnitude of autocorrelation in the residuals from a model. The Durbin Watson Statistic should be approximate 2. A statistic of 2.37381 with p-value 0.2498 indicates that there was absence of both positive and negative autocorrelation.

Model efficiency was evaluated using MSE .The model's MSE was 0.1969 which was the smallest as compared to other models fitted hence efficient in modelling volatility of the electricity cost based on MSE.

Model adequacy was tested using the Ljung-Box Q (12) for residuals. The Box Pierce Q Statistics for the residuals is 14.505 with p-value of 0.06952 indicating that the model is adequate since the p-values exceed 5 percent.

Forecasting

Using ARIMA (2, 1, 2) model, a thirty six month forecasting of the electricity cost for small commercial was performed. The Table 5 below shows the forecasts result and their respective standard errors.

Table 5: Forecasted costs

: Forecasted costs

Month	Year	Cost	Std.errors
September	2013	18.37	0.987
October	2013	18.61	1.362
November	2013	18.22	1.630
December	2013	17.60	1.862
January	2014	17.37	2.090
February	2014	17.7545	2.31327
March	2014	18.3782	2.51512
April	2014	18.6283	2.68506
May	2014	18.2650	2.83094
June	2014	17.6559	2.97075
July	2014	17.4086	3.11803
August	2014	17.7699	3.27143
September	2014	18.3840	3.41718
October	2014	18.6479	3.54460
November	2014	18.3080	3.65675
December	2014	17.7086	3.76601
January	2015	17.4481	3.88281
February	2015	17.7862	4.00663
March	2015	18.3903	4.12651
April	2015	18.6670	4.23301
May	2015	18.3500	4.32770
June	2015	17.7610	4.42040

July 2015	17.4880	4.51998
August 2015	17.8034	4.62642
September 2015	18.3969	4.73062
October 2015	18.6856	4.82412
November 2015	18.3912	4.90773
December 2015	17.8130	4.98967
January 2016	17.5284	5.07779
February 2016	17.8214	5.17246
March 2016	18.4039	5.26585
April 2016	18.7038	5.35028
May 2016	18.4315	5.42606
June 2016	17.8645	5.50030
July 2016	17.5692	5.58011
August 2016	17.8403	5.66610

Table 5 shows the general trend of the forecasted electricity costs for small commercial in Kenya. In the month of September, the forecasted cost was sh.18.37. There was an increment in the cost from August which was sh.17.74 .The forecasted cost increased further in October to 18.61, before dropping in November to 18.22 and the decreasing trend of the forecasted cost has continued for the next two months until January 2014 when the price will be sh.17.37.The cost will then adapt increasing trend and then decreasing trend. The interchanging trend will continue for the next thirty six months. The forecasted costs as shown in table 5, depicts the fluctuating nature of the electricity cost in Kenya.

SUMMARY CONCLUSION AND RECOMMENDATION

Summary

The study sought to forecast the electricity cost for small commercial in Kenya for the next five months. It used the Gretl programme for Econometric and Time-series software Version 1.9.13 in data analysis. After electricity cost was input to the computer, Stationarity was tested and achieved by using the Log-differencing technique where the first log-difference exhibited a stationary series.The ARIMA (p, d, q) model was applied to the data. Different orders for differencing (d) were tested in all cases and the first difference (d=1) was found to be the most adequate. Four different orders p and q values were tested for ARIMA (p, 1, q): (1,1,0), (0,1,1), (1,1,1) and (2,1,2).In the four series, the order (2,1,2) is the best choice. This is determined by comparing the models; selecting using the model selection criterion and the diagnostic tests whereby ARIMA (2, 1, 2) outperformed all the other models.

The model was tested for normality using the Skewness, kurtosis, JarqueBera statistics and the Q-Q normal plot. The model was found not to be normally distributed. Model adequacy was tested using the Ljung-Box statistics and found to be adequate. Model efficiency was tested using the MSE statistics and found to be efficient.

The model was tested for independence by inspecting the sample autocorrelations of the residuals. It was found to be adequate as the residuals resembled white noise. Autocorrelation was also tested using the Durbin-Watson statistic and there was no negative or positive

correlation. Forecasting was done for a thirty-six month ahead and the series described increasing trend and then decreasing trend.

Conclusion

According to the findings of the research, the ARIMA model (2, 1, 2) was ascertained for forecasting electricity costs for small commercial in Kenya as presented below;

$X_t = 0.02171 - 0.7876 X_{t-1} + 0.1044 X_{t-2} - 0.00701 e_{t-1} - 0.999997 e_{t-2} + e_t$

This is the recommended electricity costs forecasting model for small commercial in Kenya. Some of the coefficients were statistically insignificant at 5% level. The forecasting result for electricity costs depicts increasing trend and then decreasing trend. This interchanging trend will be expected over the preceding thirty-six months.

Recommendation

From the findings of this study, it is recommended the Energy Regulatory Commission and the Kenya power to adopt a stable form for electricity costs in Kenya especially in small commercial sector. Since the electricity costs in Kenya are highly influenced by diesel prices which are characterised by unexpected fluctuations, thereby fluctuating the electricity costs. The government through ERC should take the appropriate measures to stabilise the costs as high electricity costs affects the economy negatively.

Electricity in Kenya is a major input in many sectors especially for production. High electricity cost may lead to higher operating costs compared to the neighbouring countries. This may lead to relocation of industries to countries with lower operation cost leading to slow economic growth. With the government maintaining low costs, more investments from both local and foreign investors will be experienced in the country thereby leading to significant economic growth and development.

The purpose of this study was to tell the likely future events. The forecasted results in this case showed that there are expectations of decreasing electricity costs in the next first five months. This requires the government to take the appropriate measures to ensure that this trend of decreasing cost is maintained.

Suggestions for further research

I would recommend that more forecasts for electricity be done using other Models like the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) other than the ARIMA and establishing which model is the best to fit and analyse the data with more precision.

REFERENCE

[1]   P. Nyoike, "'Is the Kenyan Electricity Regulatory Board Autonomous?'," Energy Policy, no. 30, p. 987–997, 2002.
[2]   P. G. Allen, "Economic Forecasting in Agriculture," International Journal of Forecasting, vol. 4, no. 10, pp. 81-135, 1994.

[3]   gregory and Adam, "The Role of Vibrant Retail Electricity Markets in Assuring that Wholesale Power Markets Operate Effectively," electricity journal, vol. 12, no. 10, pp. 61-73, 1999.

[4]   Government of Kenya, "Economic Recovery Strategy for Wealth and Employment Creation 2003–2007," The Policy Framework Paper, Nairobi, 1996.

[5]   Eberhard A. & K Gratwick, "The Kenyan IPP Experience," Working Paper No. 49, Management Programme in Infrastructure Reform and Regulation, Graduate School of Business, University of Cape Town, South Africa, 2005.

[6]   Government of kenya, "Economic Recovery Strategy for Wealth and Employment Creation 2003–2007," Nairobi, 2003.

[7]   Prerna, M., "Forecasting Natural Gas Price-Time Series and Nonparametric Approach," 2012.

[8]   Box, G.E.P. and Jenkins, G.M., "Time Series Analysis," San Francisco: Holden-Day, San Francisco, 1976.

[9]   Javier, C., Rosario, E., Francisco, J.N. and Antonio, J.C., "ARIMA Models to Predict Next Day Electricity Prices," IEEE Transaction on Power Systems, vol. 18, no. 3, pp. 1014-1015, 2003.

[10]   Syed, A.A., Muhammad, R., Amir, R. and Ammar, G.B, "Impact of Oil prices on Food Inflation in Pakistan," Interdisciplinary Journal of Contemporary Research in Business, vol. 13, no. 11, pp. 123-124, 2012.

[11]   Makridakis, S. and Hibon, M., "ARMA Models and the Box Jenkins Methodology," Journal of Forecasting, vol. 1, no. 2, pp. 111-153, 1995.