# Modelling and Forecasting Volatility of Value Added Tax Revenue in Kenya

**Muthuri Evans Kithure, Anthony Waititu, Anthony Wanjoya**

Faculty of Science, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

**Email address:**

thurevans@gmail.com (M. E. Kithure), agwaititu@gmail.com (A. Waititu), awanjoya@gmail.com (A. Wanjoya)

**Abstract:** Taxation is one of the means by which governments finance their expenditure by imposing charges on citizens and corporate entities. Kenya Revenue Authority (KRA) is the agency responsible for the assessment, collection and accounting for of all revenues that are due to government. Volatile government revenue is a challenge for fiscal policy makers since it creates risks to government service provision and can make planning difficult, as revenue falls short of expenditure needs both frequently and unexpectedly. The main objective of this study was to model and forecast the volatility of VAT revenue collected in Kenya as well as computing its value at risk and the expected shortfall. Secondary data on daily VAT revenue collections for a period of 3 years was analyzed. The first step was to model the mean equation of the return series using the ARIMA model and ARIMA(3,0,3) was identified to be the most suitable since it had the least values of AIC and BIC. The Lagrange Multiplier test confirmed the presence of ARCH effects using the residuals of the mean equation. A number of heteroscedastic models were fitted and the TGARCH family (ARIMA(3,0,3)/TGARCH(1,2)) was preferred to fit the volatility of the returns. One step ahead forecasting of volatility of the returns was done using the model which gave a value of 7.212. Estimation of value at risk and expected shortfall involved use of POT method by fitting a GPD function to the return data series. The first step was determination of threshold by use of MRL plot and later fitting a GPD function to the return data series using the threshold. The shape, location and scale parameters were estimated using MLE and they were later used to compute the VaR loss and ES at 95% and 99% confidence intervals. The VaR at 95% and 99% was 1.45% and 1.49% respectively while the ES at both the intervals was 0.04% and 0.1% respectively. This study concluded that volatility is persistent in the daily VAT revenue collections and it can easily be modelled using conditional heteroscedastic models.

**Keywords:** Autoregressive Conditional Heteroscedasticity (ARCH), Expected Shortfall (ES),
Threshold Generalized Autoregressive Conditional Heteroscedasticity (TGARCH), Value at Risk (VaR)

## 1. Introduction

Revenue volatility can be defined as the extent of revenue fluctuation over the course of the business cycle. It is a serious concern particularly for state policymakers and fiscal administrators operating within the context of balanced budget requirements. It makes it hard to make accurate forecasts for future revenues and the establishment of long-term fiscal plans for the stable operation of public programs and services. With the global economy being liberalized and more tightly interwoven, the tax environment of governments has been increasingly volatile and uncertain over the past years, and as a result, fiscal planning and management have become more challenging particularly for state governments operating under the institutional constraints of balanced budget requirements.

The important implication of revenue volatility for state finance is that in the absence of adequate fiscal reserves, it is hard for states with volatile revenue bases to avoid massive spending cuts and tax hikes in times of economic crisis when governments counter cyclical fiscal actions are needed more than ever. Another implication is that such pro-cyclical austerity measures affect real economies, reducing households and businesses' propensity to consume and consequently creating the vicious circle of economic recession.

Capturing revenue volatility is very important. Looking at the perspective of a prospective planning, if we desire to make decisions about the structure our revenue sources on which we will rely on to provide sufficient revenue for all essential government services, we should consider both the mean rate of growth of the revenue sources and the volatility as well. Also, revenue volatility can be used as an independent variable in developing an understanding about past decisions such as revenue structure, the use of rainy day funds, or state borrowing trends among others.

A study was carried out on the importance of considering tax portfolios and economic conditions in determining the growth rate and volatility of state tax revenues. Using 1989 – 2009 state data and simple graphical constructs, various comparative analysis of the long-run growth rate and short-run volatility in percent changes of state economies and tax revenues and state tax portfolios were constructed. In the analysis, it was found that wide variations in these respects exist among states. Other findings suggested that in the short run, states cannot alter the underlying structure of the economy but can mitigate the impacts of the business cycle on their fiscal conditions by making changes to their tax portfolios [1].

Tax analysis and revenue forecasting are very essential in providing directions as to enhancing tax revenue, improving equity and efficiency of taxes, promoting investments and consequently economic growth with a view of increasing national income. In addition, it is important in monitoring processes and budget planning. It defines the resource envelope forming the basis for effective medium-term and long-term planning [2].

Revenue forecasting is an essential part of budgeting in the public sector. Therefore, it is necessary for a government to forecast the revenue it collects for planning purposes. Budgetary uncertainties have led to increased reliance on economic and revenue forecasting by state governments in recent years. Because of the magnitude of the fiscal problems facing many states, forecasting has assumed a more central role in the policy making process. As a result, revenue forecasts are closely examined and accuracy is essential for planning purposes. To improve accuracy, there is need for analysts to assemble as much information about their respective state economies as possible, including formal and informal consideration of alternative forecasts [2].

In Kenya, the tax revenue analysis and forecast is produced by a single agency, which is the Ministry of Finance in conjunction with the Kenya Institute for Public Policy Research and Analysis (KIPPRA). The ministry formulates the government budget by undertaking revenue projections and at the same time prepares the monthly/annual revenue targets for relevant revenue collection agencies in the country. VAT in Kenya is usually a tax on consumption, therefore its volatile depending on the economic seasons. The purpose of this project was to make use of heteroscedastic models to model and forecast the volatility of the VAT data series. The study also looked at the risks associated with the VAT by estimating its Value at Risk (VaR) and the Expected Shortfall (ES).

# 2. Methodology

Time series analysis is the procedure of fitting a time series to a proper model. It comprises methods that attempt to understand the nature of the data series and its characteristics which include trend, seasonality, outliers and cycle. When modeling the volatility of a return series, a mean equation is needed for testing the arch effects. ARMA or ARIMA models can be used to model the mean equation

## 2.1. Autoregressive Integrated Moving Average (ARIMA)

From application view point ARMA models are not the best to properly describe non-stationary time series, which are frequently encountered in practice. For this reason, the ARIMA model was proposed, which is an extension of an ARMA model to include the case of non-stationary data as well. In ARIMA models a non-stationary data is made stationary by applying finite differencing of the data points [3].

A time series model $y_t$ is an ARIMA model of order p, d, q, i.e. ARIMA(p, d, q), if

$$y_t' = \alpha_0 + \alpha_1 y_{t-1}' + \alpha_2 y_{t-2}' + \cdots + \alpha_p y_{t-p}' + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \varepsilon_t + \beta_2 \varepsilon_{t-2} + \cdots + \varepsilon_t + \beta_q \varepsilon_{t-q} \qquad (1)$$

Where
a) $y_t'$ is the differenced series (it may have been differenced more than once).
b) $\varepsilon_t$'s are random errors (or random shocks) at time period t.
c) $\alpha\,(i = 1,2, \dots p)$ and $\beta(i = 1,2, \dots q)$ are model parameters.
d) $\alpha_0$ is a constant.

## 2.2. Heteroscedastic Models

Volatility modeling provides an approach of calculating Value at Risk (VaR) of a financial position in risk management. Building a volatility model for a return series involves the following steps:
i   Specifying of a mean equation by testing for serial dependence in the data series. If need be, build an econometric model for the return series so as to be used to remove any linear dependence.
ii  Use of the squared residuals of the mean equation to test for ARCH effects.
iii Specifying of a volatility model if ARCH effects are statistically significant and performing a joint estimation of the mean and volatility equations.
iv  Checking the fitted model carefully and refining it if necessary.

## 2.2.1. Testing for ARCH Effects

Let $a_t = y_t - u_t$ be the residuals of the mean equation. The squared series $a_t^2$ is used to check for conditional heteroscedasticity (ARCH effect). This can be done using the

Lagrange multiplier test which is equivalent to the usual F statistic for testing $\alpha_i$ ($i = 1,2, \ldots m$) in the linear regression.

$$a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2 + e_t, t = m + 1, \cdots, T \quad (2)$$

Where;
i   $e_t$ denotes the error term.
ii   m is a pre-specified positive integer.
iii   T is the sample size.
This gives us

$$F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T-2m-1)} \quad (3)$$

Which is asymptotically distributed as chi-squared distribution with m degrees of freedom under the null hypothesis. This test is known as Lagrange multiplier test.

### 2.2.2. The ARCH Model

The first model that provides a systematic framework for volatility modeling is the ARCH model of [4]. The basic idea of ARCH models is that:
i   The mean corrected asset return $a_t$ is serially uncorrelated, but dependent.
ii   The dependence of $a_t$ can be described by a simple quadratic function of its lagged values.
Specifically, an ARCH (p) model assumes that

$$a_t = \sigma_t \epsilon_t \quad (4)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_p a_{t-p}^2 \quad (5)$$

Where
i   $\epsilon_t$ is a sequence of independent and identically distributed (iid) random variables with mean zero and variance 1.
ii   $\alpha_0 > 0$ and $\alpha_i \geq 0$ for i>0.

### 2.2.3. The GARCH Model

Although the ARCH model is simple, it requires many parameters to adequately describe the volatility process of an asset return. A useful extension known as the generalized ARCH (GARCH) model was therefore proposed [5]. For a log return series $y_t$, it's assumed that the mean equation of the process can be adequately described by an ARMA model. Let $a_t = y_t - u_t$ be the mean corrected log return. Then $a_t$ follows a GARCH (m,s) model if

$$a_t = \sigma_t \epsilon_t \quad (6)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \quad (7)$$

Where
i   $\epsilon_t$ is a sequence of independent and identically distributed (iid) random variables with mean zero and variance 1.
ii   $\alpha_0 > 0, \alpha_i \geq 0$ $\beta_j \geq 0$.
iii   $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_j) < 1$.
$\epsilon_t$ is often assumed to be a standard normal or standardized Student-t distribution.

### 2.2.4. The TGARCH Model

This is one of the models used to handle leverage effect [6]. A TGARC H (m, s) model assumes the form

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{s}(\alpha_i + \gamma_i N_{t-i})\, a_{t-i}^2 + \sum_{j=1}^{m} \beta_j \sigma_{t-j}^2 \quad (8)$$

Where
i   $N_{t-i} = 1$ if $a_{t-1} < 0$.
ii   $N_{t-i} = 0$ if $a_{t-1} \geq 0$.
iii   $\alpha_i, \beta_j$ and $\gamma_i$ are non-negative parameters satisfying conditions similar to those of GARCH models.
From the model, it is seen that a positive $a_{t-i}$ contributes $\alpha_i a_{t-i}^2$ to $\sigma_t^2$ whereas negative $a_{t-i}$ has a larger impact $(\alpha_i + \gamma_i)a_{t-i}^2$ with $\gamma_i > 0$. The model uses zero as its threshold to separate the impacts of past shocks. This model is preferred over the others because it allows the asymmetric effects between positive and negative asset returns [7].

*Model Selection Criteria*

Both AIC and BIC were used in the model selection. These can be expressed as;

$$AIC = -2\ln L + 2k \quad (9)$$

$$BIC = -2\ln L + 2\ln Nk \quad (10)$$

Where;
i   L is the maximized value of the likelihood function ii. N is the number of recorded measurements.
ii   k is the number of estimated parameters.

Existing studies shows that AIC is not consistent and has the tendency to choose models which are over-parameterized.

*Forecast Evaluation*

The mean absolute error (MAE) and the mean square error (MSE) were used as measures for evaluating forecasting performance. The MAE and MSE for n step ahead forecast are denoted as follows

$$MAE(n) = \frac{1}{N} \sum_{t=1}^{N} \left| \left(r_{\{t+n\}} - \bar{r}\right)^2 - \widehat{h}_t(n)\right| \quad (11)$$

$$MSE(n) = \frac{1}{N} \sum_{t=1}^{N} \left[\left(r_{\{t+n\}} - \bar{r}\right)^2 - \widehat{h}_t(n)\right]^2 \quad (12)$$

Where;
i   $r_{t+n}$ is the return over horizon n steps ahead at current time t.
ii   $\bar{r}$ is the mean of return.
iii   $\widehat{h}_t(n)$ is the forecasted conditional variance over horizon n steps ahead at current time t.

### 2.3. Value at Risk (VaR) and Expected Shortfall (ES)

One way of modeling VaR and ES by is applying extreme distribution based on methods such as Extreme Value Theory (EVT). EVT refers to branch of statistics which normally deals with the extreme deviations from the mean of a probability distribution. Results in EVT can be obtained either using Block Maxima model (BMM) or Peak Over Threshold model (POT). POT method models a distribution of excess over a given threshold. EVT shows that the limiting distribution of excess is a Generalized Pareto Distribution of GPD [8].

The POT method uses available data more efficiently compared to BMM. In POT, all the data which exceeds a particular threshold level is used while in BMM only the maximum from a block length is retained for distribution estimation.

### 2.3.1. GPD and EVT

The generalized Pareto distribution (GPD) is a family of continuous probability distributions. It is often used to model the tails of another distribution. It is specified by three parameters: location μ, scale β, and shape parameter ξ.

The two-parameter GPD (with scale parameter β and shape parameter ξ) has the following representation:

$$G_{\{\xi,\beta\}(y)} = \begin{cases} 1 - \left(1 + \frac{\xi}{\beta}\right)^{-\frac{1}{\xi}} & \xi \neq 0 \\ 1 - e^{-\frac{y}{\beta}} & \xi = 0 \end{cases} \qquad (13)$$

Where
i   $y > 0$ when $\xi \geq 0$.
ii   $0 \leq y \leq -\beta/\xi$ when $\xi < 0$.
iii   $\beta > 0$.

### 2.3.2. Peaks Over Threshold (POT)

To fit a GPD to our data set, adoption of the POT method that focuses on the distribution of excess above some high threshold is made. For $y - u \geq 0$ the excess distribution function can be rewritten as

$$F_{u(y-u)} = \frac{F(y) - F(u)}{1 - F(u)} \qquad (14)$$

and hence deduce the following reverse expression

$$F_{u(y)} = \left(1 - F(u)\right) F_u(y - u) F(u) \qquad (15)$$

which allows us to apply the POT method. There are two steps in applying the POT method. First is to choose an appropriate threshold and then fit the GPD function to data [9].

### 2.3.3. Mean Excess Function and Threshold Selection

For a random variable Y, the mean excess function is defined as

$$e(u) = E(Y - u | Y > u) \qquad (16)$$

i.e. the mean excess over a threshold u. If the underlying distribution $Y > u$ follows a GPD, then the corresponding mean excess is

$$e_n(u) = \frac{\beta + \xi u}{1 - \xi} \ \beta + \xi u > 0 \qquad (17)$$

provided $\xi < 1$. From the equation it's clearly seen that the mean excess function must be linear in u. More precisely, $Y > u$ follows a GPD if and only if the mean excess function is linear in u. This gives us a way of selecting an appropriate threshold.

### 2.3.4. Model Validation

Quantile plots can be used to assess the quality of a fitted

generalized Pareto model. If GPD is a reasonable fit for the excess above u, then the Q-Q plot should depict points that are approximately linear. Furthermore, the goodness-of-fit of GPD can be confirmed by utilizing the excess distribution plot and plot of the tail of underlying distribution. For a good fit, the excess should lie close to the theoretical curves. Lastly, a scatter plot of residuals should not depict any visible pattern to indicate independence of the excess.

### 2.3.5. VaR

For a random variable Y (usually the return in some risky financial instrument) with distribution function F over a specified time period, the VaR, for a given probability p, can be defined as the $p^{th}$ quantile of F, i.e.,

$$VaR_p = F^{-1}(1 - p) \qquad (18)$$

where $F^{-1}$ is the quantile function. VaR is a common measure of extreme risks and GPD is used to approximate this measure. In particular, using Equation of POT the following is obtained.

$$\widehat{VaR_p} = \begin{cases} \mu + \frac{\hat{\beta}}{\hat{\xi}} \{ \left(\frac{n}{N_\mu} P\right)^{-\hat{\xi}} - 1\} & \xi \neq 0 \\ \mu - \hat{\beta} \log\left(\frac{n}{N_\mu}(1 - P)\right) & \xi = 0 \end{cases} \qquad (19)$$

Where
i   β and ξ are the maximum likelihood estimates of the GPD parameters.
ii   n is the total sample size.
iii   Nu is the amount of observations above the chosen threshold.

### 2.3.6. ES

ES was proposed as a better measure of risk, which is sub-additive and also informs us about the likely magnitude of excess [10]. In contrast to VaR, ES is used measure the riskiness of an instrument by considering both the size and likelihood of losses above a particular threshold. ES gives the expected size of return that exceeds VaR.

$$ES_P = E(Y | Y > VaR_p \qquad (20)$$

And equivalently

$$ES_p = VaR_p + E(Y - VaR_p | Y > VaR_p) \qquad (21)$$

where the second term above represent the mean of the excess distribution. Proceeding as before, if the threshold $VaR_p$ is sufficiently large then the expected shortfall can be computed as follows.

$$\widehat{ES}_p = \frac{\widehat{VaR_p}}{1 - \hat{\xi}} + \frac{\hat{\beta} - \hat{\xi}\hat{\mu}}{1 - \hat{\xi}} \qquad (22)$$

## 3. Results and Discussion

Secondary data of the daily VAT revenue collections from January 2016 to March 2019 was used for this research.
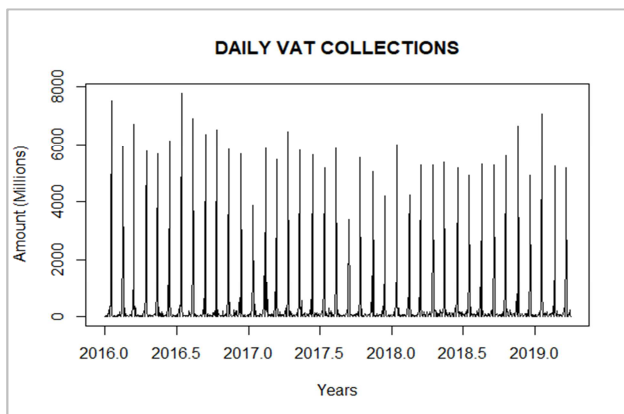
*Figure 1. Daily VAT Collection Plot.*

Figure 1 shows volatility clustering of the VAT revenue collections. It is noticeable that periods of low volatility are followed by periods of low volatility and periods of high volatility followed by periods of high volatility as well. The periods of high volatility are few days prior to the due date of payment. During these days, payments tend to increase with the highest payment being made on the due date of the month. After the due dates, the amounts of payments reduces drastically and the collections made in the proceeding days are usually low.

These VAT payments were converted in to their log returns and a time plot of the return series is as shown in Figure 2.
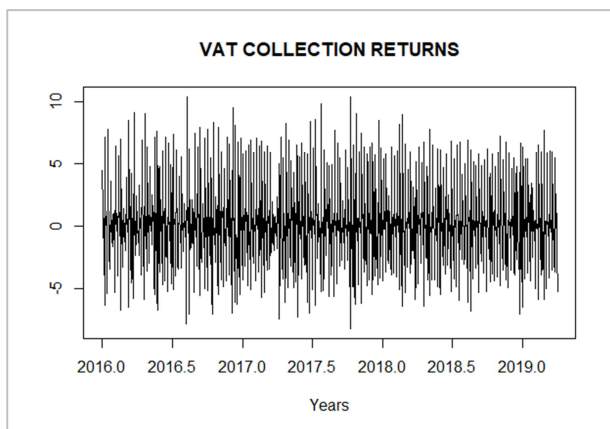


*Figure 2. A plot of daily VAT returns*

## 3.1. ARIMA Modeling

### 3.1.1. Stationary Test

The aim of this test was to check the presence or absence of unit root, where the null hypothesis is stated by default as the presence of unit root and the alternative hypothesis as absence of a unit root. Augmented Dickey Fuller (ADF) was used to test if data was stationary and produced the following results.

*Table 1. Stationary Test using ADF.*

| Test | p-value |
| --- | --- |
| ADF | 0.0000 |

Since the p-value is less than 0.05 significance level, we reject the null hypothesis and hence conclude that the returns data series is stationary. The series was later modeled using ARIMA models and checked for the arch effects.

### 3.1.2. Model Selection

Various ARIMA models were fitted to the differenced series and selection was made using the AIC and BIC values. The model with the lowest values of AIC and BIC is preferred.

*Table 2. Model Selection with AIC and BIC.*

| Model | BIC | AIC |
| --- | --- | --- |
| ARIMA(2,0,1) | 5583.38 | 5558.06 |
| ARIMA(2,0,2) | 5560.27 | 5529.88 |
| ARIMA(2,0,3) | 5567.33 | 5531.87 |
| ARIMA(3,0,1) | 5589.54 | 5559.16 |
| ARIMA(3,0,2) | 5567.32 | 5531.87 |
| ARIMA(3,0,3) | 5443.68 | 5403.16 |

From table 2, the ARIMA (3,0,3) is preferred to be the best model since it has the least values of AIC and BIC.

## 3.2. Heteroscedastic Modeling

### 3.2.1. Testing for ARCH Effects

This was done using the residuals of the mean equation. Using the squared residuals, Lagrange Multiplier (LM) test was used to test for the arch effects and the results obtained are as follows.

*Table 3. Lagrange Multiplier Test.*

| Test | df | p-value |
| --- | --- | --- |
| LM | 12 | 0.0000 |

The null hypothesis is that there is no arch effects. From table 3, the p-value obtained was less than 0.05. Therefore, the null hypothesis is rejected and a conclusion is made that there is presence of arch effects.

### 3.2.2. Model Selection

After confirming the presence of arch effects in our return data series, various ARIMA (3,0,3)/GARCH (p,q) as well as ARIMA(3,0,3)/EGARCH(p,q) and ARIMA(3,0,3)/TGARCH (p,q) models were considered for modelling the VAT returns and selection was made based on the values of the AIC and BIC.

*Table 4. TGARCH Model Selection using AIC and BIC.*

| Model | AIC | BIC |
| --- | --- | --- |
| ARIMA(3,0,3)/GARCH(1,1) | 4.630 | 4.677 |
| ARIMA(3,0,3)/GARCH(1,2) | 4.624 | 4.676 |
| ARIMA(3,0,3)/EGARCH(2,1) | 4.483 | 4.538 |
| ARIMA(3,0,3)/EGARCH(1,1) | 4.492 | 4.548 |
| ARIMA(3,0,3)/TGARCH(1,2) | 4.627 | 4.679 |
| ARIMA(3,0,3)/TGARCH(2,1) | 4.462 | 4.518 |

From table 4 above, it is evident that the best model is ARIMA(3,3)/TGARCH(1,2) since it had the lowest value of AIC and BIC using student-t distribution.

### 3.2.3. Parameter Estimation

Using maximum likelihood estimation, the parameters of our chosen model were estimated and the results obtained are as shown in table 5.

**Table 5.** *ARIMA(3,3)/TGARCH(1,2) Parameter Estimates..*

| Parameter | Estimate | Standard Error | p-value |
|---|---|---|---|
| $\varphi_1$ | 0.273657 | 0.016606 | 0.000000 |
| $\varphi_2$ | -0.678428 | 0.007762 | 0.000000 |
| $\varphi_3$ | 0.635825 | 0.016686 | 0.000000 |
| $\theta_1$ | -0.597181 | -0.597181 | 0.000000 |
| $\theta_2$ | 0.537376 | 0.000742 | 0.000000 |
| $\theta_3$ | -0.910576 | 0.000163 | 0.000000 |
| $\alpha_0$ | 1.579987 | 0.106565 | 0.000000 |
| $\alpha_1$ | 0.559060 | 0.060202 | 0.000000 |
| $\beta_1$ | 0.793704 | 0.029057 | 0.009586 |
| $\beta_2$ | 0.000002 | 0.031248 | 0.999994 |
| $\gamma$ | 1.000000 | 0.059067 | 0.000000 |

From table 5, the p-values of most of the parameters is less than 0.05 which indicates that they are statistically significant while few are greater than 0.05. The value of $\gamma$ is positive which implies that negative events will have more impact on the volatility of the returns than positive events.

### 3.2.4. Forecasting

Using the TGARCH model, one step ahead forecast of the VAT returns was done and the results are as shown in table 6

**Table 6.** *Volatility Forecasting.*

| Day | Series | Sigma |
|---|---|---|
| First | 2.314 | 7,212 |

## 3.3. Estimation of Value at Risk (VaR) and Expected Shortfall (ES)

This was done using the extreme value theory by using GPD function.

### 3.3.1. Determination of Threshold

Before fitting the GPD function to the data, it is first necessary to choose a threshold. Mean Residuals life (MRL) plot of the return data series was used to estimate the threshold as shown in figure 3
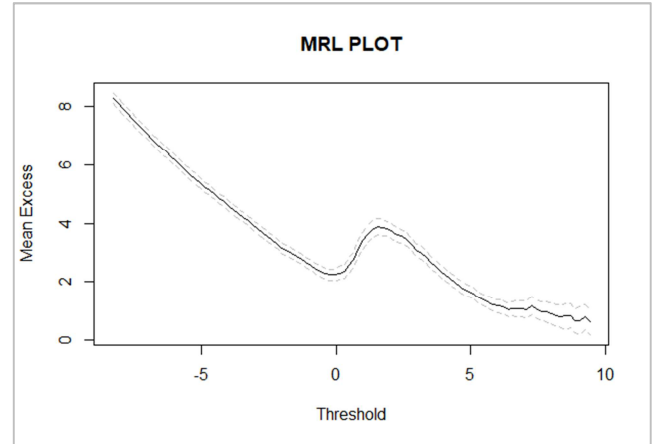


**Figure 3.** *Mean Residuals Life plot.*

Since the idea is to select a threshold whereby the graph is linear, a threshold of 0.1 appears to be a reasonable value since a reasonably straight line could be placed within the uncertainty bounds from this point up.

### 3.3.2. Estimation of GPD Parameters

The next step was to estimate the scale parameter $\sigma$ and the shape parameter $\xi$ by fitting a GPD function to the data series. Maximum likelihood estimate of the parameters was obtained and the results are as shown in table 7.

**Table 7.** *GPD Parameter Estimation.*

| Parameter | Estimate | SE |
|---|---|---|
| Scale ($\sigma$) | 1.0014385 | 0.14203701 |
| Shape ($\xi$) | 0.4558911 | 0.08111279 |
| Location ($\mu$) | -0.7932478 | 0.10900623 |

These values obtained were used to estimate the values of VaR and ES losses.

### 3.3.3. Model Adequacy

To assess the quality of the fitted model, various plots were used. These are excess distribution plot, tail of underlying distribution plot, scatter plot of residuals and Q-Q plot of residuals.
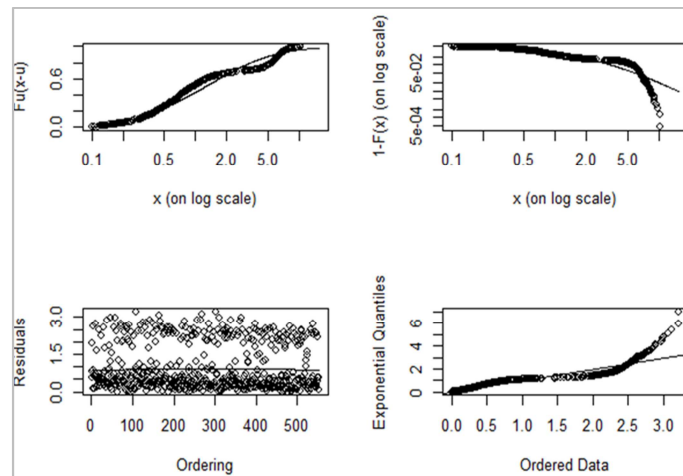


**Figure 4.** *GPD Model Validation Plots.*

From figure 4, the top plots from the left are excess distribution plot and plot of tail underlying distribution respectively. These plots show that the excess lie close to the theoretical curves. Scatter plot of the residuals does not depict any visible pattern in the residuals which indicate independence of the excess. Again the Q-Q plot depicts points which approximately linear. Therefore, all of the above plots confirmed the goodness of fit of our model.

### 3.3.4. VaR Computation

After confirming the goodness of fit of our model, the VaR was computed making use of the GPD parameters that were obtained through MLE. VaR was computed at 95% and 99% confidence interval and the results obtained were as shown in table 8.

***Table 8.** VaR Estimates.*

| CI | Estimate |
|---|---|
| 0.95 | -1.454972 |
| 0.99 | -1.488373 |

Table 8 shows that we are 95% and 99% confident that the worst loss of daily VAT collection will not exceed 1.45% and 1.49% respectively.

### 3.3.5. Estimation of Expected Shortfall

The ES was estimated in a similar way as the VaR by fitting GPD function to the data series. The scale parameter, shape parameter and the location parameter obtained by MLE from the GPD function together with the VaR estimated were used to compute the ES at different confidence intervals and the results obtained as in the table below

***Table 9.** ES Estimates.*

| CI | Estimate |
|---|---|
| 0.95 | -0.038158 |
| 0.99 | - 0.096084 |

Since ES is normally the expected loss of a return that exceeds VaR, we can conclude from Table 9 that if the VaR loss of our return becomes 1.45% at 95% confidence interval, then the worst loss on daily VAT collection will not exceed beyond 0.038%. Similarly if the VaR loss at 99% confidence interval is 1.49%, then the worst loss on daily VAT collection will not exceed beyond 0.096%.

## 4. Conclusions and Recommendation

### 4.1. Conclusions

This research study investigated the best model to capture the volatility of daily VAT revenue collected in Kenya from January 2016 to March 2019. Different statistical tests were carried out on the data so as to determine the best model to fit the returns. The analysis showed presence of components such as trend, seasonal variation and random changes.

Because of these, it was possible to model the data using time series ARIMA(3,0,3) model. The results showed that volatility was persistence in the VAT returns. ARIMA(3,0,3)/TGARCH(1,2) was identified as the best model to forecast the volatility of the revenue returns. Going further, the study estimated the market risks associated with the revenue returns. These were Value at Risk and the Expected Shortfall. Using Peak Over Threshold method of the Extreme Value Theory, various parameters that are required were computed by fitting a Generalised Pareto Distribution function in the data. These were then used to estimate the value of the risks (losses) at various confidence intervals.

### 4.2. Recommendations

The general recommendation goes to the future researchers to examine the effectiveness of other heteroscedastic models like EGARCH, IGARCH and GARCH-M in modeling and forecasting of revenue volatility. Also, other methods of risk measures like use of block maxima model (BMM) and GARCH model can also be looked at and a comparison made between the different methods. Future research can also look at other tax revenues and expenditure as well.

## References

[1]   G. C. &. N. R. D. Cornia, State tax revenue growth and volatility, Federal Reserve Bank of St. Louis Regional Economic Development, 2010.

[2]   ICPAK, "Kenya's revenue analysis 2010-2015," *A Historical Perspective to,* 2015.

[3]   R. A. R. Adhikari, An introductory study on time series modeling.

[4]   R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates," *Journal of the Econometric Society,* pp. 987-1007, 1982.

[5]   T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of econometrics,* pp. 307-327, 1986.

[6]   D. B. Nelson, "Conditional heteroskedasticity in asset returns," *Journal of the Econometric Society,* pp. 347-370, 1991.

[7]   G. e. a. Ali, "Egarch, gjr-garch, tgarch, avgarch, ngarch, igarch," *Journal of Statistical and,* pp. 57-73, 2013.

[8]   e. a. M. Gilli, "An application of extreme value theory for measuring financial rik," *Computational Economics ,* pp. 207-228, 2006.

[9]   E. p. R. Norberg, "Modelling extremal events for insurance and finance," *The Journal of the IAA,* pp. 285-286, 1998.

[10]  D. E. G. F. Ren, "Extreme value analysis of daily canadian crude oil prices," *Applied Financial Economics,* p. 941-954, 2010.