Change Point Estimation in Volatility of a Time series using a Kolmogorov Smirnov Type Test Statistic

Josephine Njeri Ngure[1] and Anthony Gichuhi Waititu[2]

[1]Pan African University Institute for Basic Sciences, Technology, and Innovation, P.O BOX 62000-00200, ngurejosephine@gmail.com

[2]Department of Statistics and Actuarial Sciences, JKUAT, P.O BOX 62000-00200, awaititu@jkuat.ac.ke

Abstract

Detection of structural change in volatility of a time series is very important for understanding volatility dynamics and the stylized facts observed in financial time series. By applying the Nadaraya Watson kernel estimator of the mean function, estimated residuals are obtained. In this work, a Kolmogorov Smirnov type test statistic for change point estimation is developed and applied to conditional variances obtained from the squared residuals. The consistency of the change point estimator is shown through simulations. The developed estimator is then applied to KES/USD exchange rate data set to estimate a single change point.

Keywords: Volatility, Change point, Kolmogorov-Smirnov, ICSS, GARCH

Section 1: INTRODUCTION

When doing change point analysis, the major point of interest is to decide if the observations follow one model or if there is at least one time point when the model is believed to have changed. This therefore results in two sub-fields of change point analysis; change point detection and change point estimation. Change point analysis normally assumes that it is possible to segment the data into regimes and that the data structure is homogeneous within each regime. Of importance is that the change is assumed to be abrupt (as though each occurs completely between two observations) and not gradual or smooth. Although in many settings multiple change-points could be of paramount interest, we shall only seek to detect change point through the assumption of At Most One change point approach. Change point analysis can be performed in either the offline setting or online setting or by estimating single change point versus multiple change point among other scenarios. We define change point detection as the problem of finding abrupt changes in data when a property of the time series changes. For each observed time series, the instant where these structural changes occur are called change points. The time moment when the model has changed is called change point. Other synonyms for change point include but not limited to segmentation, structural breaks, break points, regime switching, and detecting disorder. Unlike change point detection, change point estimation tries to model and interpret known changes in the time series rather than identifying that a change has occurred. Change point estimation (estimates) focus on describing the nature and degree of the known change. Change points can be found in a wide range of literature including quality control, economics, medicine, environment, and even linguistics. Detection of large "homogeneous" segments of data enables one to identify "hidden" regularities in a time series behavior and to create a mathematical model for each segment of homogeneity. One choses change point with the intention of maximizing the separation between two segmentation. Thus, the goal of change point detection and estimation is to recover these segments as accurately as possible.

Subsection 1.1: Statement of the problem

A good change point estimation method in conditional variance (volatility) should be sensitive to the skewness of the observations (which does not happen when ranking is done) e.g. when dealing with financial data. With returns, there are times of uniquely high returns and others of uniquely low returns and thus rank bases test statistics makes the data robust to outliers. This is because, by ranking the observations, there is mitigation of the impact of extremely high or low returns (outliers). Reason being that regardless of how extreme an outlier is, it often receives the same rank as if it were slightly larger than the second largest observation. This is because all ranks are equally far apart from each other violating the stylized facts of returns. Hence, we adopt a non-ranking method of the observations. Thus, we aim to propose a theory of estimating the change point in volatility of a financial time series with USD/KES exchange rate dataset application in mind. In this work, the regression function (conditional mean function) and the conditional variance function are unknown hence we impose few or no restrictions to our data set. The inference based on non-parametric models is usually robust against misspecification of the underlying regression model and thus non-parametric models effectively avoid the problem of misspecification normally found in parametric approaches, which may yield inconsistent estimators. We thus take a non-parametric approach for our results to be robust with respect to model specifications.

Section 2: LITERATURE REVIEW

The first published article concerning change points analysis was done by (Page, 1954) who considered testing for a potential single change point for data from a common parametric distribution motivated by a quality control setting in manufacturing. Since then, change point analysis has developed rapidly with considerations on either multiple change point detection and estimation, different types of data and other assumptions being put into consideration.

(Chen G. a., 2005) proposed a procedure that was able to combine the least squares approach which does not require specific forms of the marginal or the transitional density functions (i.e the regression and conditional variance functions) to estimate the change points in the conditional variance {volatility} of a non-parametric model of time series in which the regression and the conditional variance were unknown. Further, the asymptotic properties of the estimators and test statistics were established. The location of the change point(s) was not been specified a priori like some other studies from previous scholars had assumed. Finally, the proposed test was consistent and more powerful than the non-parametric ones already existing tests in literature. Finally, the practicality of the methods was by application to the Hong Kong stock market index (HSI) series.

Change point analysis by (Gichuhi, 2008) in Bernoulli random variables based on neural networks motivated by a regression setting was the focus of the researcher. The parameters of the model were estimated using neural network method with the evidence that parameter estimates were only identifiable up to a given family of transformations and further derived the consistency and asymptotic normality of the network parameter estimates. A neural network based likelihood ratio test statistic to detect a change point in a given set of data was determined. The limiting distribution of the change point estimator was established. The results showed that the sample size, change point location and the size of change had an influence on change point detection. Through simulation, percentile bootstrap method showed superiority to profile log-likelihood ratio method in determining the change point confidence intervals.

Modeling of financial volatility in the presence of abrupt changes is a research done by (Ross, 2013) where the author incorporated the ICSS GARCH algorithm to detect changes in volatility of financial returns. Although the algorithm was simple to implement, its parameters were based on the assumption of the financial returns following a Gaussian distribution and thus can produce very many spurious jump points if this assumption is violated. By applying ICSS to heavy tailed series, poor results were found since extreme observations were misinterpreted as regime shifts. This indicated that the ICSS algorithm was only applicable to detect change points to weekly returns and using the algorithm to daily returns could generate too many spurious false positives for it to be useful because of the number of extreme values. Thus, due to this problem of ICSS GARCH algorithm, the author replaced the ICSS segmentation step with a technique which was purely based on non-parametric statistics which makes no assumptions on the true returns distribution and which in turn allows one to ignore the Gaussian assumptions and allows for its deployment on the daily returns which he coined NPCPM-GARCH. The author further analyzes several stock indices for change points in volatility mainly the Dow Jones Industrial Average, the German DAX, the VIX volatility index and the Japanese Nikkei 225. He further compared his results with those obtained from ICSS GARCH and found that his method gave a better fit to the data sets when measured using a standard criteria i.e at the same level of significance. The research showed that the ICSS algorithm was not able to ignore the assumption of Gaussian, which contradicted with the stylized facts of returns leading to the detection of very many change points some of which did not correspond to genuine long-term changes in volatility (false positives). This prompted the researcher to adopt the Mood test statistic in a sequential setting and which was found to work and detect true change points with non-Gaussian data.

Section 3: NON-PARAMETRIC TIME SERIES MODELLING

For a given time series $X_t; t = 1,2, \dots T$ non-parametric methods are used to analyze the features of interest. Conditional variances or conditional quartiles are required if interval forecasts or estimates of future volatility are desired as shall be necessary in this work. Suppose we let $X_t = \log(\frac{S_t}{S_{t-1}})$ be the return process in period t for $t = 1,2, \dots T$ and $X_{t-1}, X_{t-2}, \dots, X_{t-d}$ be the return processes at any time periods less than t, S_t is the price process of the stock in period t for $t = 1,2, \dots T$. We assume that there is a non-parametric and non-linear relationship between the current return values and the previous return values, modeled by a non-parametric autoregressive process of this form

$$X_t = m(X_{t-1}, X_{t-2}, \dots, X_{t-d}) + u_t \qquad t = 1,2, \dots, T \tag{1}$$

where u_t a series of innovations (random shocks) which is independent of $X_{t-1}, \dots, X_{t-d}$ satisfying

$$\mathbb{E}(u_t | X_{t-1}, X_{t-2}, \dots, X_{t-d}) = 0 \tag{2}$$

$m(.)$ is the conditional mean (smooth) function in period t given past time periods $X_{t-1}, X_{t-2}, \dots, X_{t-d}$ and it is the minimum mean squared error (MSE) 1-step predictor of X_t. The approximation precision of $m(.)$, increases with the sample size. Since in many situations point forecasting is too limited an objective, and the future volatility and higher order moments are of interest in addition to the conditional mean, we therefore let the following representation of the innovations u_t to hold

$$u_t = \sigma(X_{t-1}, X_{t-2}, \dots, X_{t-d}) z_t \tag{3}$$

and thus extend equation 1 above to a more general non-parametric conditional heteroskedastic model as in equation 4 below

$$X_t = m(X_{t-1}, X_{t-2}, \dots, X_{t-d}) + \sigma(X_{t-1}, X_{t-2}, \dots, X_{t-d}) z_t \tag{4}$$

$\mathbb{E}(X_t|X_{t-1} = x_1, X_{t-2} = x_2, \dots, X_{t-d} = x_d) = m(x_{t-1}, x_{t-2}, \dots, x_{t-d})$ is the non-linear autoregressive conditional mean (smooth) function of the returns.

Variance $(X_t|X_{t-1} = x_1, X_{t-2} = x_2, \dots, X_{t-d} = x_d) = \sigma^2(x_{t-1}, x_{t-2}, \dots, x_{t-d})$ is the non-linear autoregressive conditional variance (smooth) function of the returns,

z_t is an independent and identically distributed sequence of random variables with

$$\mathbb{E}(z_t|X_{t-1}, \dots, X_{t-d}) = 0 ,$$

$$\text{Variance } (z_t|X_{t-1}, \dots, X_{t-d}) = 1$$

and independent of $X_{t-1}, X_{t-2}, \dots, X_{t-d}$

Equation 4 above is a non-parametric autoregressive conditional heteroscedastic model and is the most flexible non-parametric time series model because it does not impose any (parametric) particular form on the conditional mean and conditional variance functions. Due to curse of dimensionality problem, where by as the dimension d grows, statistical and computational inefficiency comes in, we set $d = 1$ so that equation 4 above becomes

$$X_t = m(X_{t-1}) + \sigma(X_{t-1})z_t \tag{5}$$

Estimates of the functions $m(x)$ and $\sigma^2(x)$ are obtained by applying the Nadaraya Watson estimator of the unknown regression function (conditional mean at the evaluation points) and its properties such that

$$\widehat{m}_T(x) = \frac{\sum_{t=2}^{T} K\left(\frac{X_{t-1}-x}{b_x}\right) X_t}{\sum_{t=2}^{T} K\left(\frac{X_{t-1}-x}{b_x}\right)} \tag{6}$$

$$\hat{\sigma}_T^2(x) = \frac{\sum_2^T K\left(\frac{X_{t-1}-x}{b_x}\right)\left(X_t - \widehat{m}(X_{t-1})\right)^2}{\sum_{t=2}^{T} K\left(\frac{X_{t-1}-x}{b_x}\right)} \tag{7}$$

Under some assumptions, it can be shown that $\widehat{m}_T(x)$ is a consistent estimator of $m(x)$. $K(.): \mathbb{R} \to \mathbb{R}$ is a kernel function, which is continuous, symmetric, integrating to one with bounded support $[-1,1]$ in that the estimator only uses the observations in the interval $(x - b_x, x + b_x)$ and b_x is the bandwidth parameter or the tuning parameter. The bandwidth (smoothing parameter) controls the level of neighboring such that for a given kernel function K and a fixed x , observations (X_{t-1}, X_t) with X_{t-1} far from x are given more weights as b_x increases.This means that the larger the bandwidth is chosen, the less the mean function $\widehat{m}_T(x)$ is changing with x . Therefore, we can conclude that the degree of smoothness of the conditional mean function increases with the bandwidth. Therefore, it means that a weighted average of the observations is used as an estimator for the conditional mean function .

The Nadaraya Watson kernel regression estimator was first proposed independently by (Nadaraya, 1964) and (Watson, 1964)

The estimators of the mean function and conditional variance function have shown to be strongly consistent and asymptotically normal for α mixing observations. In this research, utilize the Epanechnikov kernel since it is the most efficient in minimizing the Mean Integrated Squared error putting in mind that the choice of the kernel is not as important as the choice of the bandwidth (this does not mean we disregard the choice of the kernel).

It is important to remember that when a kernel estimator is applied to dependent data, e.g. in financial time series returns data like in the case of this work, then it is affected only by the dependence among the observations in a small window and not by that between all data. This fact therefore reduces the dependence between the estimates so that most of the techniques developed for independent data are applicable as well. This is what we shall refer to as the **Whitening by window principle**. Also, the memory of the underlying process decreases with distance between events and that the rate of decay can be estimated by the mixing conditions some of which include the strong α mixing condition and the ϕ mixing condition as below

i. α (Strong) mixing condition. A sequence $\{X_t\}$ is said to be α mixing if

$\sup_{A \in \mathcal{F}_1^n B \in \mathcal{F}_{n+k}^\infty} |p(A \cap B) - p(A)p(B)| \leq \alpha_k \quad n = 1,2, \ldots$ and $\alpha_k \to 0$ as $k \to \infty$ and $\mathcal{F}_i^j$ is the σ −field generated by $X_i, \ldots, X_j$

ii. ϕ (Uniformly) mixing condition. This is a stronger condition which establishes that a sequence $\{X_t\}$ is said to be ϕ mixing if

$|p(A \cap B) - p(A)p(B)| \leq \phi_k \, p(A)$ for any $A \in \mathcal{F}_1^n$and $B \in \mathcal{F}_{n+k}^\infty$ and $\phi_k \to 0$ as $k \to \infty$.
For more on this, visit (Robinson, 1983).

The mixing conditions above control the dependency between X_i and $X_j$as the time distance $i - j$ increases. The rate at which α_k and ϕ_k goes to zero plays a noble role in showing the asymptotic properties of the non-parametric smoothing procedure. These conditions are usually difficult to check but if the process follows a stationary Markov chain, then geometric ergodicity imply absolute regularity, which in turn implies strong mixing conditions. Proposition 6 in subsection 2.4.2.3 of (Doukhan, 2012) give conditions on m, σ and the innovations that imply geometric ergodicity of $\{X_t\}$. This implies strong mixing properties with exponential mixing rates.

Equation 5 can generate heavy tailed distributions and we demonstrate this by considering a simple model

$$X_t = \sigma(X_{t-1}) z_t \tag{8}$$

with z_t having a standard normal distribution. By Jensen's inequality (Pishro-Nik, 2016),

$$\text{Kurtosis } (X_t) = \frac{\mathbb{E}(X_t^4)}{[\mathbb{E}(X_t^2)]^2} = \frac{\mathbb{E}[\sigma^4(X_{t-1}) z_t^4]}{[\mathbb{E}(\sigma^2(X_{t-1}) z_t^2)]^2} = 3\,\frac{\mathbb{E}[\sigma^4(X_{t-1})]}{[\mathbb{E}(\sigma^2(X_{t-1}))^2]} \geq 3 \tag{9}$$

This heavy tailed-ness feature implied by equation 5 makes it a successful mode for modelling data, which exhibit heavy tails e.g. financial time series data of returns. It is important to note that non-parametric time series approach has been highly appreciated by practitioners as a preliminary search method aimed at establishing the final parametric model.

Subsection 3.1: Single change point test statistic

In this section, we derive an estimator for change point in volatility of a non-parametric regression model for time series as shown in equation 5. Since we are in the off-line setting and only the conditional variance function that is changing with time, with the assumption that the conditional mean function is not changing with time, we shall formulate the change point problem as a hypothesis testing procedure of the following form

$$H_0: \sigma_t^2(X_{t-1}) = \sigma_{(1)}^2(X_{t-1}) \qquad t = 1,2, \dots, T \tag{10}$$

While the at most one change point alternative is

$$H_1: \begin{cases} \sigma_t^2(X_{t-1}) \ = \sigma_{(1)}^2(X_{t-1}) & \text{for } t = 1,2,\dots,\tau \\ \sigma_t^2(X_{t-1}) \ = \sigma_{(2)}^2(X_{t-1}) & \text{for } t = \tau+1,\dots,T \end{cases} \tag{11}$$

We want to locate (estimate) the change point position $\hat{\tau}$ and probably the number of change point (s) (Chen J. a., 2001). The regression function (conditional mean function) $m(x)$, the conditional variance function $\sigma(X)$, the distribution of the covariate as well as the distribution of the errors is completely unknown and we have not made any parametric form for them meaning our approach is fully non-parametric.

If we consider H_0 when we do not have a change point in volatility, we can rewrite equation 5

$$X_t = m(X_{t-1}) + \sigma(X_{t-1})z_i \quad \rightarrow u_t^2 = \{X_t - m(X_{t-1})\}^2 \ = \sigma^2(X_{t-1})z_t^2 \tag{12}$$

$t = 1,2,\dots,T$.

Which in turn implies that the conditional variance function of ε_t.

$$\mathbb{E}\{(X_t - m(X_{t-1})\}^2 = \sigma^2(X_{t-1}) \tag{13}$$

If we consider the alternative hypothesis, H_1 we can re-write the non-parametric model 3 with a single change point in volatility as

$$X_t = m(X_{t-1}) + \sigma_t(X_{t-1})z_t \rightarrow \ u_t^2 = \{X_t - m(X_{t-1})\}^2 = \sigma_t^2(X_{t-1})z_t^2 \tag{14}$$

This means that under the alternative hypothesis,

$$\mathbb{E}\{(X_t - m(X_{t-1})\}^2 = \sigma_t^2(X_{t-1}) \tag{15}$$

$t = 1,2,\dots,\tau,\tau+1,\dots,T$.

In this work, we assume that the conditional mean function (regression function) is stable and does not change with time but the conditional variance function is not stable and that $\mathbb{E}(z_t|X_{t-1}) = 0$, $\text{var}(z_t|X_{t-1}) = 1 \ \ \mathbb{E}(z_t)^4 < \infty$ with z_t being a sequence of random variables.

Suppose $\varepsilon_t = \frac{X_t - \widehat{m}(X_{t-1})}{\widehat{\sigma}(X_{t-1})}, t = 1,2,\dots,T$. We are concerned with testing non-parametrically our hypothesis above by defining the partial sum of the squared residuals across all possible sample segments as shown in equation 16 below

$$\varepsilon_T = \sum_{t=1}^{T} \varepsilon_t^2 , \qquad \varepsilon_\tau = \sum_{t=1}^{\tau} \varepsilon_t^2 , \quad \varepsilon_{\tau+1} = \sum_{t=\tau+1}^{T} \varepsilon_t^2 \tag{16}$$

Where $1 \leq \tau \leq T$ where $\sigma_{(1)}^2(X_{t-1})$ and $\sigma_{(2)}^2(X_{t-1})$ for $t = 1,2,\ldots,T$ denotes the conditional variance functions of the sequence $\{\varepsilon_t^2\}_{t=1}^{\tau}$ and $\{\varepsilon_t^2\}_{t=\tau+1}^{T}$ before change point instant and after the change point instant respectively. We propose a change point test statistic which is able to quantify the deviation between $\sigma_{(1)}^2(X_{t-1})$and $\sigma_{(2)}^2(X_{t-1})$ written as $l_p(\sigma_{(1)}^2(X_{t-1})$ - $\sigma_{(2)}^2(X_{t-1}))$ where for $p \geq 1$

$$l_p\left(\sigma_{(1)}^2(X_{t-1}) \; - \; \sigma_{(2)}^2(X_{t-1})\right) = \left(\sum_{t=1}^{T} w_\tau |\varepsilon_\tau - \varepsilon_{\tau+1}|^p\right)^{\frac{1}{p}} \tag{17}$$

Motivated by the l_p norm and properties of l_p space, the change point test statistic and change point estimator is constructed. We set $p = 2$ and work in l_2 norm. w_τ is a weight function, which is measurable and which depends on the sample size n and the change point position τ . It gives the sensitivity of the test statistic against different alternatives in the sense of the position of change. The weight function is arbitrary chosen so that it satisfies the condition that

$$\sum_{t=1}^{T} \varepsilon_t^2 = \frac{\tau}{T} \sum_{t=1}^{T} \varepsilon_t^2 \rightarrow \frac{1}{T}\left(\sum_{t=1}^{\tau} \varepsilon_t^2 - \frac{\tau}{T} \sum_{t=1}^{T} \varepsilon_t^2\right) = 0 \tag{18}$$

Simple Algebra will help derive the Kolmogorov Smirnov type statistic for change point detection.

The Kolmogorov Smirnov type statistic upon derivation becomes

$$D_\tau = \left(\frac{\tau}{T}\left(1 - \frac{\tau}{T}\right)\right)^{\frac{1}{2}} \left| \frac{1}{\tau} \sum_{t=1}^{\tau} \varepsilon_t^2 - \frac{1}{T-\tau} \sum_{t=\tau+1}^{T} \varepsilon_t^2 \right| \tag{19}$$

The KS type test statistic above gives more weight to the observations at the tails of a distribution and hence it is an appropriate statistic for change point detection especially when handling financial time series data, which exhibit heavy tails.

Subsection 3.2: Single change point estimator

The KS type estimator of change point will be the point where the KS type test statistic has its global maximum. This is because, the global maximum will often occur at the area of true change point (The point where we have maximum distance between the conditional variances of the residuals). Hence, a good choice of the estimator for the time of change is as given in equation 20 below as

$$\hat{\tau} = \text{Arg}\max_{\tau}|D_\tau| \tag{20}$$

The estimate $\hat{\tau}$ is the point at which there is maximal sample evidence for a break in the squared residual process. In the presence of a single break, we shall show, through simulation, that $\hat{\tau}$ is a consistent estimator of the unknown change point τ^* .

Section 4: APPLICATION OF THE KS TYPE ESTIMATOR

Consider the model below

$$X_t = m(X_{t-1}) + \sigma(X_{t-1})z_t \tag{21}$$

Suppose the model has a single change point in the volatility function defined as

$$\sigma(X_{t-1}) = \begin{cases} 2 + 0.7\varepsilon_{t-1}^2 & for\ t = 1,2,\dots,\tau \\ 1 + 0.035\varepsilon_{t-1}^2 & for\ t = \tau+1,\dots,T \end{cases} \tag{22}$$

$$m(x) \to X_t = 0.35X_{t-1} + \varepsilon_t + 0.4\varepsilon_{t-1}$$

Where $\varepsilon_t = \sigma(X_{t-1})z_t$ and z_t is a sequence of independent and identically distributed random variables with mean zero and variance 1 assumed from a normal distribution.

We thus create a table under different sample sizes with 1000 bootstrap samples in each. We fix the change point at $\frac{1}{3}T$, $\frac{1}{2}T$ and $\frac{2}{3}T$. In each simulation, the estimates of the change point highly depended on the locations of the change points and the sample size. The estimates were most accurate if τ^* was fixed around the middle of the sample.

To demonstrate consistency of the change point estimator, the distance between an estimation and the true change point index is obtained, then normalized by the size of sample (Truong, 2018).

This error was found to be decreasing to zero as the size of the samples grew unbounded which further verified the asymptotic consistency of the change point estimator. Consistency results only deal with change point fractions and not the time indexes themselves. We investigate the consistency of the estimator when the change point is fixed at $\frac{1}{3}T$, $\frac{1}{2}T$ and $\frac{2}{3}T$. Note that we are losing two observations in each simulation due to curse of dimensionality problem. The results from table 1 below showed that the change point estimator $\hat{\tau}$ was a consistent estimator of τ^*.

Table 1 : Table to demonstrate consistency of the change point estimator

Sample size T	True change point instant τ^*	Estimated change point instant $\hat{\tau}$	$\frac{\lvert\hat{\tau} - \tau^*\rvert}{T} \xrightarrow{p} 0$
50	$\frac{1}{3}T \to \tau^* =16$	14	0.04167
	$\frac{1}{2}T \to \tau^* = 24$	17	0.14583
	$\frac{2}{3}T \to \tau^* = 32$	20	0.27083
100	$\frac{1}{3}T \to \tau^*=32$	26	0.06122
	$\frac{1}{2}T \to \tau^* = 49$	36	0.13265
	$\frac{2}{3}T \to \tau^* = 66$	48	0.18367
200	$\frac{1}{3}T \to \tau^*=66$	53	0.06566
	$\frac{1}{2}T \to \tau^* = 99$	80	0.09596
	$\frac{2}{3}T \to \tau^* = 132$	107	0.12626
500	$\frac{1}{3}T \to \tau^*=166$	149	0.03414
	$\frac{1}{2}T \to \tau^* = 249$	226	0.04619
	$\frac{2}{3}T \to \tau^* = 332$	309	0.04619
1000	$\frac{1}{3}T \to \tau^*=332$	317	0.00751
	$\frac{1}{2}T \to \tau^* = 499$	484	0.01503
	$\frac{2}{3}T \to \tau^* = 668$	645	0.01151

Section 4.2: Data analysis and results

We apply the change point estimator to historical data set of USD/KES exchange rate data set from 2 January 2013 to 18 March 2019 to estimate change point in the conditional variance function (volatility) of exchange rate returns. The data set consisted of 2444 daily observations and the plot of the exchange rates is as shown in figure 1 below

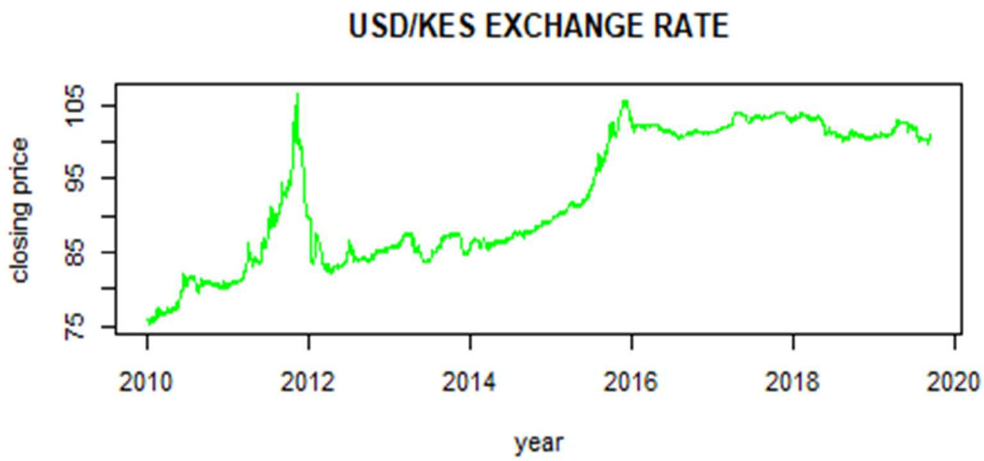


Figure 1: USD /KSH exchange rate

A plot of the returns is as shown in the figures 2 and 3 below

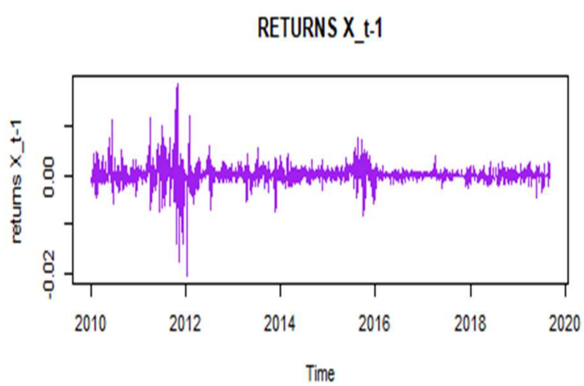


Figure 2: Returns X_{t-1}

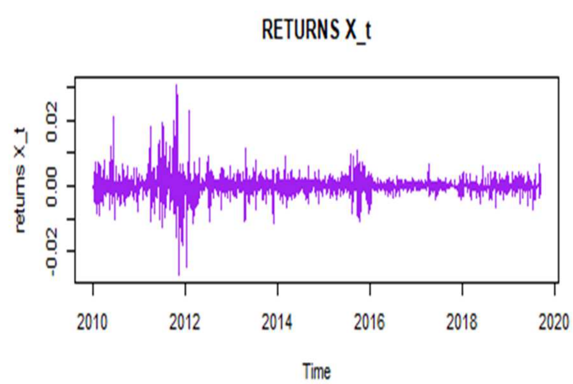


Figure 3: Returns X_t

A plot of the squared residuals is as shown in figure 4 below

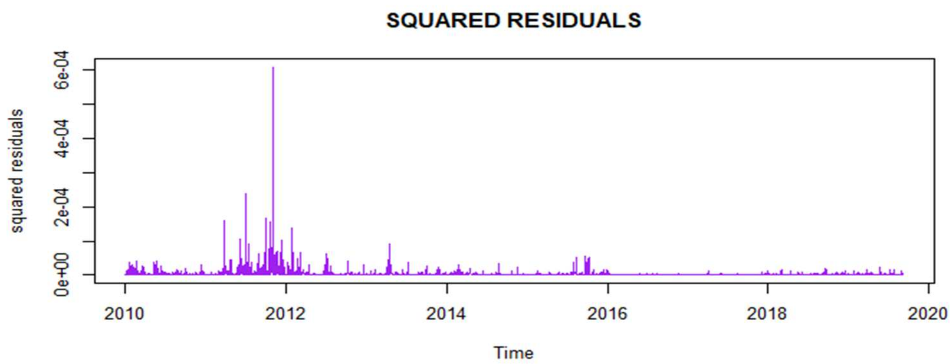
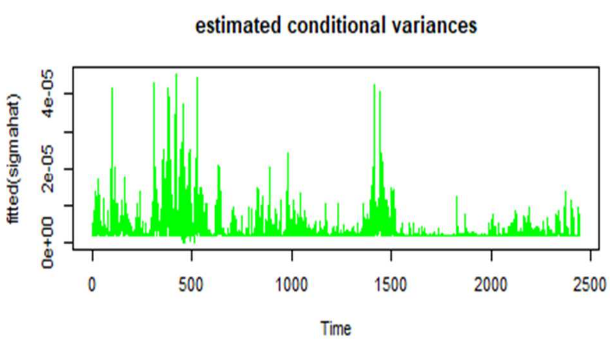


Figure 4: plot of the squared residuals

A plot of the estimated conditional variances and residuals is as shown below



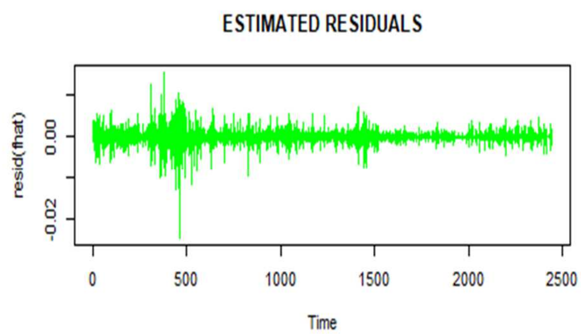
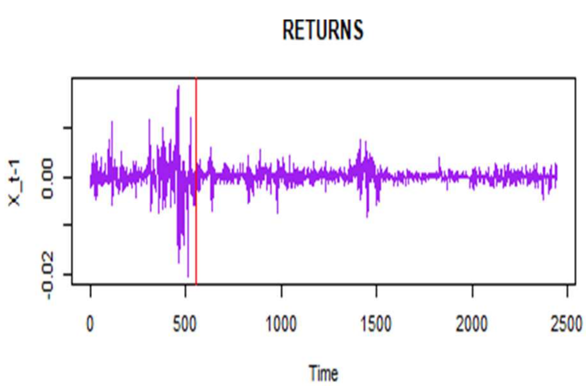


Figure 5: Estimated conditional variances.

Figure 6: Estimated residuals.



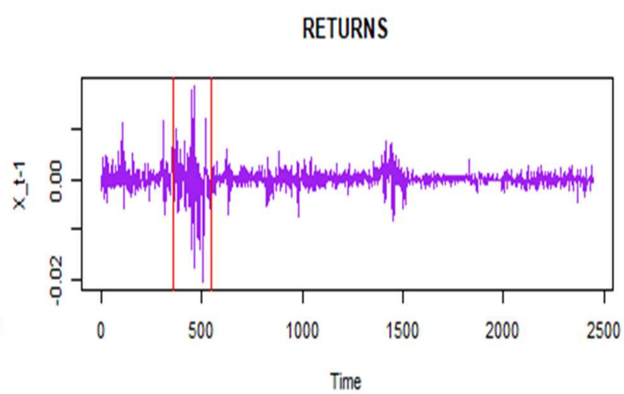


Figure 7: Returns with change point at 551

Figure 8: Returns with estimated change points at 551 and 358

From figure 2, we investigated the possibility of a single change point and were able to estimate it at point 551 corresponding to 13 February 2012. We also investigated the possibility of another change point by applying binary segmentation approach. Binary segmentation procedure allows for estimating the position of a single change point at each stage. The change point estimator is further applied to each sub-sequence of the returns. The next change point was estimated at point 358, which corresponded to 17 May 2011 and was as shown in figure 8 above.

The corresponding plots of the change point statistics at $\hat{\tau}$=551 and $\hat{\tau}$=358 were as shown in figures 9 and 10 below respectively

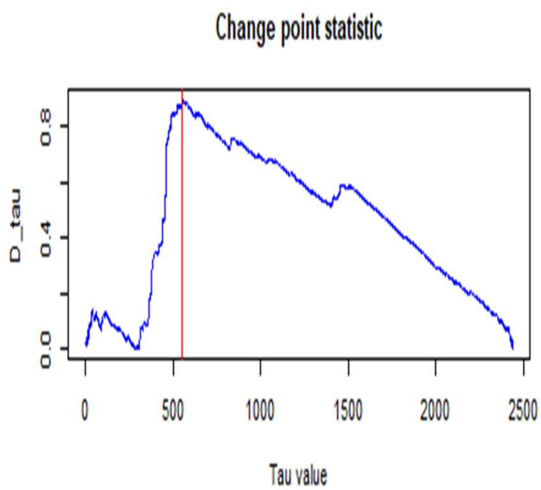


Figure 9: Change point statistic with estimated Change point at point 551

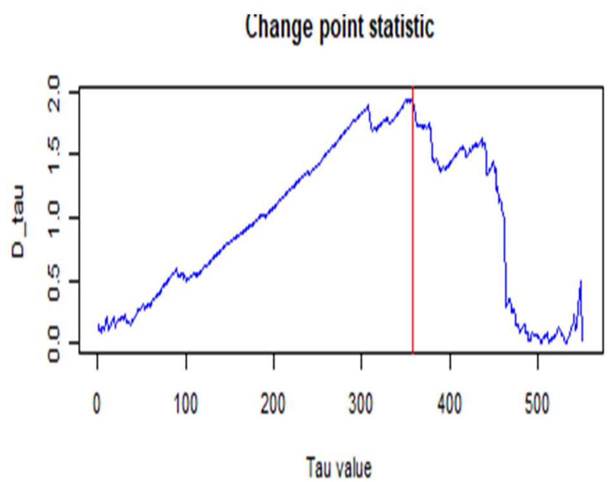


Figure 10: Change point statistic with estimated change point at point 358.

Subsection 4.3: Results and discussions

From figure 1, there is an increasing trend between January 2010 and October 2011 where the exchange rate prices were at the peak. Afterwards, one observes a decreasing trend after which the exchange rates started to rise again. The historical events associated with the behavior of the exchange rate plot were established. In 2011, there was the August 2011 stock market fall (Kibiy, 2016) because of price drop of stock prices in the stock exchanges across the major world markets in North America, Europe and Asia. The evidence from the seventh Bi-Annual Monetary Policy Committee Report issued by Central Bank of Kenya, in October 2011, can further support these results. A combination of both domestic and International economic developments during the six months period to October 2011 determined the conduct of monetary policy for Central Bank of Kenya overall mandate of price stability. These developments hence resulted in an escalation of inflationary pressure and exchange rate volatility, hence distorting the economy's recovery from the adverse effects of the global financial crisis. The USD/KSH exchange rate depreciated from 83.89 to 101.39 in the period April 2011 to October 2011. The depreciation of the shilling against US dollar could have been due to the buildup in the deficit in the current account. Current account is the gap between imports and exports of goods. This was because of a rise in imports of machinery and transport equipment, which are key inputs for the manufacturing sector required for the economic recovery process as well as the uncertainty in the global financial markets, caused by the debt crisis in the Eurozone.

From figure 7, to account for the change point at 17 May 2011, the foreign exchange market witnessed significant volatility between May 2011 and October 2011 reflecting the general volatility in the global financial markets as well as increase in demand for foreign exchange to finance imports. The result was that the Kenyan Shilling just like other currencies in the region and other global markets therefore weakened substantially e.g. the Kenyan shilling against the US dollar depreciated from an average of 84.2 in March 2011 to 101.39 in October 2011 (20.42% in percentage depreciation). Other reasons attributed to the change point in volatility on 17 May 2011 were high international food and fuel prices, the drought compounded by the conflict experienced in the Horn of Africa, the Euro crisis and major inefficiencies in Kenya’s agriculture sector.

To account for the change point in 13 February 2012, in both Kenya and Uganda, the economies reported slow growth at the beginning of 2012 following high inflation and high commercial bank interest rates.

Section 5: CONCLUSION AND RECOMMENDATIONS

In this paper, we have proposed a procedure to estimate change point in volatility of a time series modelled using a non-parametric approach. This non-parametric modeling is important in finance and non-parametric estimators are very powerful in distinguishing among many models like derivative pricing models. We have demonstrated the consistency of the change point estimator through simulations and seen that our estimator is consistent. One can easily extend the method to multidimensional non-parametric models (models of higher dimension) of the form

$X_t = m(X_{t-1}, X_{t-2}, \dots, X_{t-d}) + \sigma(X_{t-1}, X_{t-2}, \dots, X_{t-d})z_t$ where the regression function and the conditional variance functions should be estimated using multivariate kernel methods. $m(.)$, $\sigma(.)$ are multiple variable d functions, while at the same time being careful on how to deal with the curse of dimensionality problem which may lead to poor performance in higher dimensional regression problem since for $d > 2$, the subspace of $\mathbb{R}^{d+1}$ spanned by the data is rather empty.

BIBLIOGRAPHY

Chen, G. a. (2005). Nonparametric estimation of structural change points in volatility models for time series. *Journal of Econometrics, 126*(1), 79-114.

Chen, J. a. (2001). On change point detection and estimation. *Communications in statistics-simulation and computation, 30*(3), 665--697.

Doukhan, P. (2012). *Mixing: properties and examples* (Vol. 85). Springer Science \& Business Media.

Gichuhi, A. W. (2008). Nonparametric changepoint analysis for bernoulli random variables based on neural networks.

Kibiy, J. a. (2016). Determinants of Exchange Rate Volatility of the Kenyan Shilling Against World Major Currencies. *International Journal of Social Sciences and Information Technology, 2*, 1181-1198.

Mohr, M. (2018). Changepoint detection in a non parametric time series regression model.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability \& Its Applications, 9*(1), 141--142.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika, 41*(1/2), 100-115.

Pishro-Nik, H. (2016). *Introduction to probability,statistics, and random processes.*

Robinson, P. M. (1983). Nonparametric estimators for time series. *Journal of Time Series Analysis, 4*(3), 185--207.

Ross, G. J. (2013). Modelling financial volatility in the presence of abrupt changes. *Physica A: Statistical Mechanics and its Applications, 392*, 350-360.

Truong, C. a. (2018). A review of change point detection methods. *arXiv preprint arXiv:1801.00718*.

Watson, G. S. (1964). Smooth regression analysis. *Sankhy{\=a}: The Indian Journal of Statistics, Series A*, 359--372.