

---

# Moving Block Bootstrap Method for Determining Confidence Intervals for a Change Point in Time Series in the Presence of Conditional Heteroscedasticity

Josephine Njeri Ngunjiri<sup>1, \*</sup>, Anthony Gichuhi Waititu<sup>2</sup>, Simon Maina Mundia<sup>1, 3</sup>

<sup>1</sup>Pan African University Institute for Basic Sciences, Technology and Innovation (PAUSTI), Nairobi, Kenya

<sup>2</sup>Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi, Kenya

<sup>3</sup>Department of Statistics and Actuarial Sciences, Dedan Kimathi University of Technology (DEKUT), Nyeri, Kenya

## Email address:

[jngure@kyu.ac.ke](mailto:jngure@kyu.ac.ke) (Josephine Njeri Ngunjiri), [awaititu@jkuat.ac.ke](mailto:awaititu@jkuat.ac.ke) (Anthony Gichuhi Waititu),

[mundiamaina@yahoo.com](mailto:mundiamaina@yahoo.com) (Simon Maina Mundia)

\*Corresponding author

## To cite this article:

Josephine Njeri Ngunjiri, Anthony Gichuhi Waititu, Simon Maina Mundia. (2025). Moving Block Bootstrap Method for Determining Confidence Intervals for a Change Point in Time Series in the Presence of Conditional Heteroscedasticity. *Mathematics and Computer Science*, 10(2), 38-43. <https://doi.org/10.11648/j.mcs.20251002.12>

**Received:** 3 February 2025; **Accepted:** 19 March 2025; **Published:** 6 May 2025

---

**Abstract:** This paper seeks to use the Moving Block Bootstrap method to determine the confidence intervals for a change point in conditional variance function of data exhibiting conditional heteroscedasticity and heterogeneity. Confidence intervals for a change point normally provide or give a range within which the true change point location is likely to lie. This is usually based on a specified confidence level. This helps to in turn determine whether the change point is statistically significant especially after determining the critical values for the distribution of the change point test statistic. Confidence intervals are also called interval estimates as opposed to point estimates which provide a single estimate for a parameter.

**Keywords:** Change Point, Bootstrap, Moving Block, Confidence Interval

---

## 1. Introduction

A Non parametric Auto-Regressive Conditional Heteroscedastic model is considered in the simulations. The conditional mean and conditional variance functions are first estimated using Nadaraya Watson kernel estimator. A test statistic for unknown abrupt change point in volatility which takes into consideration conditional heteroskedasticity and the fourth moment of the time series (hence the residuals) is considered. The test statistic is based on  $L_2$  norm of the conditional variance functions of the squared residuals. The volatility change point estimator is there after obtained by maximizing the change point test statistic. Consistency of the change point estimator is clearly shown in the work of [8]. The limit distribution of the change point test statistic is obtained hence the critical values as in the work by [9] and by simulation, the test is shown to be consistent with asymptotic power one. A point estimate provides a single plausible value of

a certain parameter. However, the point estimate being a single value does not express the statistical variation, or random error, that the estimate has. Hence, instead of supplying a single estimate of a parameter, a logical step is to provide a plausible range of values constructed from sample data so that the population parameter occurs within that range at a specified probability. The plausible range of values is called a confidence interval and the specified probability before sampling is called confidence level. Confidence interval for the change point can be determined by either knowing or approximating the limit distribution of  $\hat{\tau}_n - \tau$  where  $\tau$  is the true change point and  $\hat{\tau}_n$  is its estimate for a sample of size  $n$ , [3]. One approximation is the bootstrap approach [7].

The bootstrap method introduced by [4] samples individual data points with replacement and is only applied to independent and identically distributed (i.i.d) data [5]. For a stationary time series, successive observations are usually correlated but observations separated far apart in time are

nearly uncorrelated. Hence one cannot re-sample individual data points but re-sample blocks of data. The idea behind block re-sampling for stationary series is that individual blocks of observations separated far enough in time will be nearly uncorrelated hence are treated as exchangeable. In this study, due to the nonparametric set-up, and since the data displayed dependence and conditional heteroskedasticity, then blocks of observations were re-sampled. Otherwise, statistics computed from the re-sampled data or from transformations of the re-sampled data without blocking are usually inconsistent.

The block bootstrap methods introduced by [6] divides the time series into blocks of  $b$  consecutive observations. The blocks which can be overlapping or non-overlapping in addition to being of either fixed or variable length are then re-sampled with replacement to create pseudo-time series on which the statistic of interest is computed. They include the Moving Block Bootstrap (MBB), Non-overlapping Block Bootstrap (NBB), the Circular Block Bootstrap (CBB) and the Stationary Block Bootstrap (SBB) [1].

The motivation behind block bootstrap is that the dependence structure of the original data is preserved within the block by sampling blocks of data so that one obtains asymptotically correct estimates if the length of the block converges to infinity with the size of the sample. Further, since the length of the blocks increases with the size of the sample, when data is generated by a weakly dependent process, the MBB introduced by [6] is able to reproduce the underlying dependence structure of the process in an asymptotic manner and thus was employed in this work. In the MBB, overlapping blocks of fixed length are applied. This allows for more blocks than when there is no overlapping and hence is more efficient. [7] provides theoretical and practical examples showing that fixed block lengths are better than varying ones and that overlapping blocks are much better than non-overlapping ones.

For analysis, the block length  $b$ , which acts as a tuning parameter, and which increases with the sample size should be long enough so as to preserve heteroscedasticity and the correlation structure in the time series which involves a trade-off. If the block size is too small, the time dependency of the data is destroyed whenever a new block is started while if the block size is too large, then one obtains few blocks, the bootstrap samples are not random enough, hence the pseudo-data may tend to look similar. In both cases, the consequence is the decline in the average accuracy of the MBB. This implies that there exist an optimal block size  $\hat{b}$  which is able to maximize this accuracy. Optimal block length depends on the sample size, the test statistic being considered and the data generating process (correlation structure in the data). Due to this reason, the block size which on average produced the best estimate of change point (minimizes the mean squared error) was utilized.

## 2. Moving Block Bootstrap Procedure

### 2.1. Step Wise Description

To perform the MBB procedure, given a time series  $\{X_t, t = 1, 2, \dots, n\}$ , estimate the change point  $\hat{\tau}_n$ .

Obtain  $N = n - b + 1$  overlapping blocks of data  $B_t = (X_t, \dots, X_{t+b-1})$  of length  $b < n$ . For simplicity, one supposes that  $b$  divides  $n$  and with the condition that for dependent data  $b \rightarrow \infty$  and  $\frac{b}{n} \rightarrow 0$  as  $n \rightarrow \infty$ . The first block contains observations 1 to  $b$ , the second contains observations 2 to  $b+1$  and the last contains observations  $n-b+1$  to  $n$  as

$$\begin{aligned} B_1 &= (X_1, \dots, X_b) \\ B_2 &= (X_2, \dots, X_{b+1}) \\ &\vdots \\ B_N &= (X_{n-b+1}, \dots, X_n) \end{aligned}$$

From the  $N$  overlapping blocks  $\{B_1, B_2, \dots, B_N\}$ , re-sample with replacement  $l = \frac{n}{b}$  blocks to form pseudo-data  $\{B_1^*, B_2^*, \dots, B_l^*\}$  of length  $n = lb$  where each of the  $N$  overlapping blocks has a  $\frac{1}{n-b+1}$  chance of being selected. Stitch all the  $l$  blocks together so as to form the bootstrap sample also called the pseudo-time series (the order in which the blocks are sampled is the same order followed when stitching the blocks in order to obtain the bootstrapped series). The re-sampling is done  $\mathbb{L}$  times. The estimate of change point  $\hat{\tau}_n^*$  is then obtained for each of the many sets ( $\mathbb{L}$  sets) of pseudo-data so as to obtain the bootstrap change point vector  $\hat{\tau}_n^{*1}, \dots, \hat{\tau}_n^{*\mathbb{L}}$ . As long as the block size increases with the sample size, the distribution of the bootstrap statistic approximates the asymptotic distribution. The bootstrap change point vector is then arranged in ascending order. From the replicates of change point, one is able to estimate the distribution function of  $\hat{\tau}_n^* - \hat{\tau}_n$  where  $\hat{\tau}_n^*$  is the estimate of the change point of the re-samples.

### 2.2. Percentile Bootstrap Confidence Intervals for the Change Point

To construct a confidence interval for the time of change, the distribution of  $\hat{\tau}_n - \tau$  where the value of  $\hat{\tau}_n$  gives an estimate for the time of change and  $\tau$  is the true time of change is approximated using the MBB approach (The bootstrap distribution of  $\hat{\tau}_n^* - \hat{\tau}_n$  approximates the distribution of  $\hat{\tau}_n - \tau$ ). The percentile method for construction of confidence intervals is then applied.

Theoretically, a  $100(1 - \alpha)\%$  confidence interval for the change point  $\tau$  is given by the form

$$100(1 - \alpha)\%ci = (\hat{\tau}_n - a_2, \hat{\tau}_n - a_1) \quad (1)$$

where  $a_1$  and  $a_2$  satisfy

$$P(\hat{\tau}_n - a_2 \leq \tau \leq \hat{\tau}_n - a_1) = 1 - \alpha \quad (2)$$

In addition, if one wants an equitailed interval, then the requirement is

$$P(\tau \leq \hat{\tau}_n - a_2) = \frac{\alpha}{2} = P(\tau > \hat{\tau}_n - a_1) \quad (3)$$

hence  $a_1$  and  $a_2$  are the quantiles of the random variable  $\hat{\tau}_n - \tau$ .

Suppose that  $\tau_{\frac{\alpha}{2}}^*$  and  $\tau_{1-\frac{\alpha}{2}}^*$  are the quantiles of  $\hat{\tau}_n^*$  such that

$$P\left(\hat{\tau}_n^* \leq \tau_{\frac{\alpha}{2}}^*\right) = P\left(\hat{\tau}_n^* > \tau_{1-\frac{\alpha}{2}}^*\right) = \frac{\alpha}{2} \quad (4)$$

Then one has that

$$P\left(\tau_{\frac{\alpha}{2}}^* \leq \hat{\tau}_n^* \leq \tau_{1-\frac{\alpha}{2}}^*\right) = 1 - \alpha \quad (5)$$

Equation (5) implies that

$$P\left(\tau_{\frac{\alpha}{2}}^* - \hat{\tau}_n \leq \hat{\tau}_n^* - \hat{\tau}_n \leq \tau_{1-\frac{\alpha}{2}}^* - \hat{\tau}_n\right) = 1 - \alpha \quad (6)$$

The bootstrap principle for confidence intervals assumes that one can approximate the quantiles of  $\hat{\tau}_n - \tau$  by the quantiles of  $\hat{\tau}_n^* - \hat{\tau}_n$ . If one assumes that this bootstrap approximation is possible, then from Equation (3) and Equation (6)

$$P\left(\tau_{\frac{\alpha}{2}}^* - \hat{\tau}_n \leq \hat{\tau}_n - \tau \leq \tau_{1-\frac{\alpha}{2}}^* - \hat{\tau}_n\right) \approx 1 - \alpha \quad (7)$$

Introducing  $-\hat{\tau}_n$  in Equation (7), one obtains

$$P\left(\tau_{\frac{\alpha}{2}}^* - \hat{\tau}_n - \hat{\tau}_n \leq \hat{\tau}_n - \tau - \hat{\tau}_n \leq \tau_{1-\frac{\alpha}{2}}^* - \hat{\tau}_n - \hat{\tau}_n\right) \approx 1 - \alpha \quad (8)$$

$$P\left(-\tau_{\frac{\alpha}{2}}^* + \hat{\tau}_n + \hat{\tau}_n \geq -\hat{\tau}_n + \tau + \hat{\tau}_n \geq -\tau_{1-\frac{\alpha}{2}}^* + \hat{\tau}_n + \hat{\tau}_n\right) \approx 1 - \alpha$$

$$P\left(\hat{\tau}_n - \left(\tau_{1-\frac{\alpha}{2}}^* - \hat{\tau}_n\right) \leq \tau \leq \hat{\tau}_n - \left(\tau_{\frac{\alpha}{2}}^* - \hat{\tau}_n\right)\right) \approx 1 - \alpha$$

where  $\left(\tau_{1-\frac{\alpha}{2}}^* - \hat{\tau}_n\right)$  and  $\left(\tau_{\frac{\alpha}{2}}^* - \hat{\tau}_n\right)$  are the bootstrap approximations of  $a_2$  and  $a_1$  respectively in Equation (1) in the equitailed scenario. As noted in [4], estimation of confidence intervals for a parameter transformation gives better confidence intervals than the direct estimation for the parameter. Hence, the random variable  $\hat{\tau}_n$  is transformed

using a symmetric function say  $h(\cdot)$  (which has a symmetric distribution) denoted as

$$\hat{\omega}_n = h(\hat{\tau}_n) \quad (9)$$

From Equation (9), one has

$$P\left[\hat{\omega}_n - \left(\omega_{1-\frac{\alpha}{2}}^* - \hat{\omega}_n\right) \leq h(\tau) \leq \hat{\omega}_n - \left(\omega_{\frac{\alpha}{2}}^* - \hat{\omega}_n\right)\right] \approx 1 - \alpha \quad (10)$$

where  $\omega_{\frac{\alpha}{2}}^*$  is the  $\frac{\alpha}{2}$  quantile of the transformed random variable  $h(\hat{\tau}_n^*)$ . Due to symmetry, it is known that  $a_1 = -a_2$  and hence  $\left(\omega_{1-\frac{\alpha}{2}}^* - \hat{\omega}_n\right) = -\left(\omega_{\frac{\alpha}{2}}^* - \hat{\omega}_n\right)$  which means that one can transform the approximated quantiles of  $\hat{\omega}_n - h(\tau)$  in Equation (10) such that

$$P\left[\hat{\omega}_n + \left(\omega_{\frac{\alpha}{2}}^* - \hat{\omega}_n\right) \leq h(\tau) \leq \hat{\omega}_n + \left(\omega_{1-\frac{\alpha}{2}}^* - \hat{\omega}_n\right)\right] \approx 1 - \alpha \quad (11)$$

which reduces to

$$P\left(\omega_{\frac{\alpha}{2}}^* \leq h(\tau) \leq \omega_{1-\frac{\alpha}{2}}^*\right) \approx 1 - \alpha \quad (12)$$

Simplifying Equation (12) and transforming it back to its initial (original) scale gives

$$P\left(\tau_{\frac{\alpha}{2}}^* \leq \tau \leq \tau_{1-\frac{\alpha}{2}}^*\right) \approx 1 - \alpha \quad (13)$$

so that

$$P\left(\tau_{n,((\mathbb{L}+1)\frac{\alpha}{2})}^* \leq \tau \leq \tau_{n,((\mathbb{L}+1)(1-\frac{\alpha}{2}))}^*\right) \approx 1 - \alpha \quad (14)$$

Hence, the  $100(1 - \alpha)\%$  bootstrap percentile confidence interval for the time of change  $\tau$  is given by

$$100(1 - \alpha)\% \text{ci} = \left(\tau_{n,[(\mathbb{L}+1)\frac{\alpha}{2}]}^*, \tau_{n,[(\mathbb{L}+1)(1-\frac{\alpha}{2})]}^*\right) \quad (15)$$

where the square brackets means that one rounds to the nearest integer. For example, when  $\alpha = 0.05$  and  $\mathbb{L} = 999$ , then lower =  $25^{th}$  and upper =  $975^{th}$  value of the ordered change point estimates from the bootstrap distribution forms a 95% confidence interval for the change point estimate. This implies that the limits of the  $100(1 - \alpha)\%$  confidence interval for  $\hat{\tau}_n$  are

given by the  $(100 \cdot \frac{\alpha}{2})^{th}$  and the  $(100 \cdot (1 - \frac{\alpha}{2}))^{th}$  percentiles of the approximated distribution of change point obtained from the  $\mathbb{L}$  bootstrap samples.

### 3. Simulation Study

#### 3.1. Confidence Intervals for a Single Change Point

To investigate the confidence interval for the change point, the case of a single break in volatility when the data generating process is an ARMA(1, 1) – ARCH(1)

$$X_t = 0.35X_{t-1} + \epsilon_t + 0.4\epsilon_{t-1}, \quad \epsilon_t = \sigma(X_{t-1})z_t, \quad z_t \sim N(0, 1)$$

$$\sigma_t^2(X_{t-1}) = \begin{cases} (0.1 + \Delta) + 0.1\epsilon_{t-1}^2 & 1 \leq t \leq \tau \\ 0.1 + 0.1\epsilon_{t-1}^2 & \tau < t \leq n \end{cases} \quad (16)$$

with  $\Delta = 0.4$  was considered. The sample size is fixed at  $n = 500$  and the change point at  $\tau = 249$ .  $\hat{\tau}_n$  is estimated from the simulated data.  $\mathbb{L} = 999$  bootstrap replications of  $\hat{\tau}_n$  are done in line with the moving block bootstrap procedure. The change point estimated from the simulated data using Equation (16) with  $\Delta = 0.4$  is  $\hat{\tau}_{500} = 246$  against the true one  $\tau = 249$ . Figure 1 displays the results.

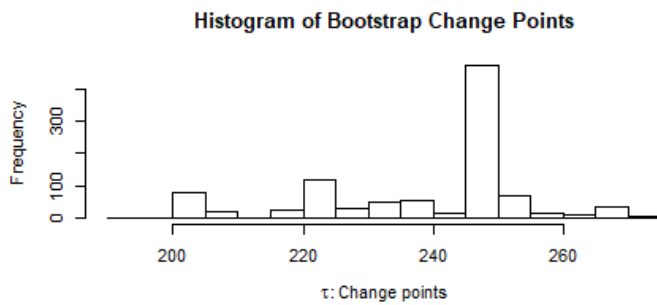


Figure 1. A Histogram of  $\mathbb{L} = 999$  bootstrap replications of  $\hat{\tau}_n$ .

One can deduce from the histogram of change points in Figure 1 that the distribution of the change point estimates is not symmetric.

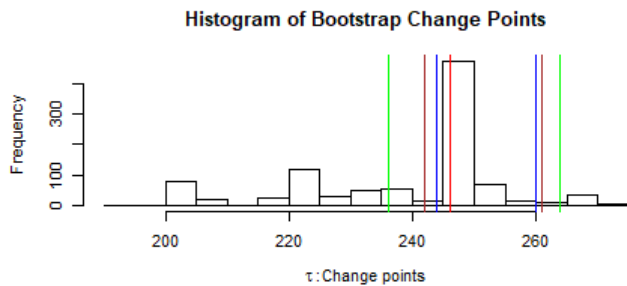


Figure 2. A Histogram of  $\mathbb{L} = 999$  bootstrap replications of  $\hat{\tau}_n$ . The red vertical line represents  $\hat{\tau}_{500} = 246$ . The two blue vertical lines represent the 90% confidence interval. The two brown vertical lines represent the 95% confidence interval. The two green vertical lines represent the 99% confidence interval.

Table 1. Confidence Interval results for  $\mathbb{L} = 999$  bootstrap replications of the change point  $\hat{\tau}_{500} = 246$  from a sample size  $n = 500$  generated as in Equation (16) with  $\Delta = 0.4$ .

Confidence level	True change point $\tau$	Confidence Interval
90%	249	244-260
95%	249	242-261
99%	249	236-264

From Figure 2, one can deduce that the percentile bootstrap confidence intervals for the change point are not symmetric around the change point estimate  $\hat{\tau}_n$ . This is especially so because the bootstrap distribution of the change point estimates is not symmetric. The results are similar to those obtained by [2].

#### 3.2. Coverage Performance

Due to variation in sampling which brings about sampling error, a confidence interval can fail to cover the true change point  $\tau$ . A 95% confidence interval means that on collecting a large number of samples and constructing the corresponding confidence interval for each sample, about 95% of the intervals will contain the change point. Hence 0.95 is the coverage probability. The coverage performance of the Percentile Bootstrap confidence intervals for the change point  $\tau$  was assessed by using the empirical coverage rate  $c$  given by Equation (17) below as

$$c = \frac{\text{Number of confidence intervals including } \tau}{\text{Total number of trials}} \times 100 \quad (17)$$

In cases where the lower end of the confidence interval is larger than  $\tau$  then this is referred to as “miss left”, and in cases where the upper end of the confidence interval is less than  $\tau$ , then it is referred to as “miss right”. The “miss” rate is the sum of the “miss left” rate and “miss right” rate. When determining confidence intervals by bootstrapping, a  $100(1 - 2\alpha)\%$  confidence interval  $(\hat{\tau}_{Lower}, \hat{\tau}_{Upper})$  is usually expected to have a probability  $\alpha$  of miss-coverage of the true change point  $\tau$  from above or below. This means,

$$\text{Prob}(\tau < \hat{\tau}_{Lower}) = \alpha \text{ or } \text{Prob}(\tau > \hat{\tau}_{Upper}) = \alpha \quad (18)$$

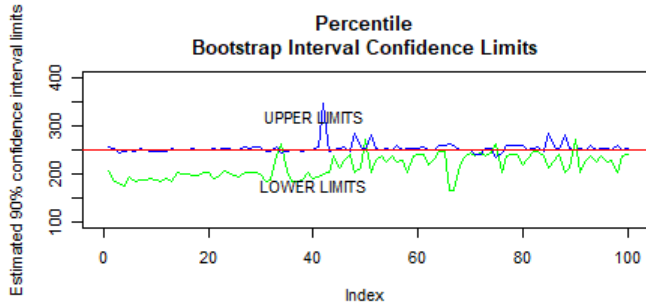
Table 2. Results for 100 Percentile Bootstrap Confidence interval realizations for the volatility change point  $\tau$  from a sample size  $n = 500$  generated as in Equation 16 with  $\Delta = 0.4$ . In each realization,  $\mathbb{L} = 999$  bootstrap replications of  $\hat{\tau}_{100} = 246$  are done.

Confidence level (%)	% Miss Left	% Miss Right
90%	12	4
95%	8	2
99%	3	1

Approximate confidence intervals are usually graded on how accurately they are able to match Equation (18) as noted in [4]. Table 2 displays the number of times in percentage

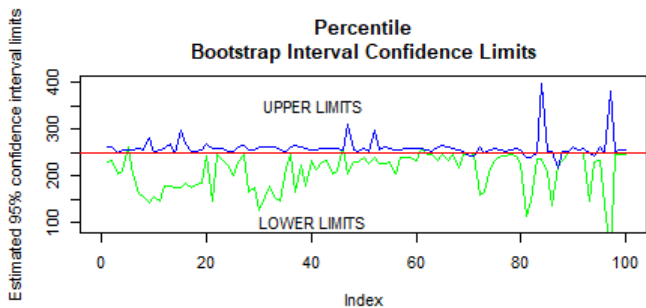
that the percentile bootstrap confidence intervals missed the true change point on the left and on the right hand side in 100 realizations.

From Table 2, the percentile bootstrap confidence intervals do not have the expected coverage probably because the bootstrap distribution of the change point estimates is skewed (asymmetric) and not centered on the true change point. The method over covers on the right and under covers on the left. Remedies to this exist which could be looked into in the future.



**Figure 3.** A graph of 100, 90% confidence interval realizations of the change point  $\tau = 249$  from a sample of size  $n = 500$  simulated as per Equation (16) with  $\Delta = 0.4$ . For each realization,  $\mathbb{L} = 999$  bootstrap replications of  $\hat{\tau}_n$  were done.

In Figure 3, the red horizontal line indicates the true change point  $\tau = 249$ , the green line represents the lower 90% Confidence Interval Limits while the blue line represents the upper 90% Confidence Interval Limits.



**Figure 4.** A graph of 100, 95% confidence interval realizations of the change point  $\tau = 249$  from a sample of size  $n = 500$  simulated as per Equation (16) with  $\Delta = 0.4$ . For each realization,  $\mathbb{L} = 999$  bootstrap replications of  $\hat{\tau}_n$  were done.

In Figure 4, the red horizontal line indicates the true change point  $\tau = 249$ , the green line represents the lower 95% Confidence Interval Limits while the blue line represents the upper 95% Confidence Interval Limits.

### 3.3. Summary on Confidence Intervals for the Change Point

This research aimed at determining confidence intervals for the change point  $\tau$ . The 90%, 95% and 99% percentile bootstrap confidence intervals for  $\tau$  were determined using Moving Block Bootstrap approach. The distribution of the bootstrap change point estimates was not symmetric and

not centered around the true change point. The percentile bootstrap method was found to over cover on the right and under cover on the left. probably due to asymmetric distribution of the bootstrap change point estimates.

## Acknowledgments

The authors greatly acknowledge Pan African University Institute for Basic Sciences, Technology, and Innovation (PAUSTI) in collaboration with African Development Bank and African Union for funding this research. The USD/KSH exchange rate data set used to support the findings of this study is available from The Central Bank of Kenya <https://www.centralbank.go.ke>.

## References

- [1] Michael R Chernick. *Bootstrap methods: A guide for practitioners and researchers*, volume 619. John Wiley & Sons, 2011.
- [2] R Dennis Cook and Sanford Weisberg. Confidence curves in nonlinear regression. *Journal of the American Statistical Association*, 85(410): 544–551, 1990.
- [3] Alexandra Da Costa Dias and P Embrechts. *Change-point analysis for dependence structures in finance and insurance*. University of Leicester, 2004.
- [4] Bradley Efron and Robert J Tibshirani. An introduction to the bootstrap. *Monographs on statistics and applied probability*, 57: 1–436, 1993.
- [5] AW Gichuhi, J Franke, and JM Kihoro. Parametric change point estimation, testing and confidence interval applied in business. *Journal of Agriculture, Science and Technology*, 14(2): 136–148, 2012.
- [6] Hans R Künsch. The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17 (3): 1217–1241, 1989.
- [7] Soumendra Nath Lahiri. *Resampling methods for dependent data*. Springer Science & Business Media, 2013.
- [8] Josephine Njeri Ngure, Anthony Gichuhi Waititu, and Simon Maina Mundia. Consistency of an estimator for change point in volatility of financial returns. *Journal of Mathematics Research*, 13 (1), 2021.
- [9] Josephine Njeri Ngure, Anthony Gichuhi Waititu, and Simon Maina Mundia. Detection and estimation of change point in volatility function of foreign exchange rate returns. *International Journal of Data Science and Analysis*, 9(1): 1–12, 2023. <https://doi.org/10.11648/j.ijdsa.20230901.11>