

A Likelihood-Based Multiple Change Point Algorithm for Count Data with Allowance for Over-Dispersion

Shalyne Nyambura, Anthony Waititu, Antony Wanjoya, Herbert Imboga

School of Mathematics & Physical Sciences, Jomo Kenyatta University of Agriculture & Technology, Juja, Kenya
Email: nshalyne@gmail.com, agwaititu@gmail.com, awanjoya@gmail.com, imbogaherbert@jkuat.ac.ke

How to cite this paper: Nyambura, S., Waititu, A., Wanjoya, A. and Imboga, H. (2024) A Likelihood-Based Multiple Change Point Algorithm for Count Data with Allowance for Over-Dispersion. *Open Journal of Statistics*, 14, 518-545.
<https://doi.org/10.4236/ojs.2024.145023>

Received: September 21, 2024

Accepted: October 27, 2024

Published: October 30, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Count data is almost always over-dispersed where the variance exceeds the mean. Several count data models have been proposed by researchers but the problem of over-dispersion still remains unresolved, more so in the context of change point analysis. This study develops a likelihood-based algorithm that detects and estimates multiple change points in a set of count data assumed to follow the Negative Binomial distribution. Discrete change point procedures discussed in literature work well for equi-dispersed data. The new algorithm produces reliable estimates of change points in cases of both equi-dispersed and over-dispersed count data; hence its advantage over other count data change point techniques. The Negative Binomial Multiple Change Point Algorithm was tested using simulated data for different sample sizes and varying positions of change. Changes in the distribution parameters were detected and estimated by conducting a likelihood ratio test on several partitions of data obtained through step-wise recursive binary segmentation. Critical values for the likelihood ratio test were developed and used to check for significance of the maximum likelihood estimates of the change points. The change point algorithm was found to work best for large datasets, though it also works well for small and medium-sized datasets with little to no error in the location of change points. The algorithm correctly detects changes when present and fails to detect changes when change is absent in actual sense. Power analysis of the likelihood ratio test for change was performed through Monte-Carlo simulation in the single change point setting. Sensitivity analysis of the test power showed that likelihood ratio test is the most powerful when the simulated change points are located mid-way through the sample data as opposed to when changes were located in the periphery. Further, the test is more powerful when the change was located three-quarter-way through the sample data compared to when the change point is closer (quarter-way) to the first observation.

Keywords

Over-Dispersion, Multiple Changepoint, Binary Segmentation, Likelihood Ratio Test

1. Purpose and Objectives

1.1. Introduction

Count data arising from stochastic processes often exhibit over-dispersion, where the sample variance exceeds the sample mean. Several count data models have been proposed by researchers in available literature, but the problem of over-dispersion still remains unresolved, more so in the context of change point analysis. Discrete change-point procedures discussed so far work well for equi-dispersed data but produce biased estimates where data are over-dispersed. This study develops an algorithm for detecting and estimating multiple change points in the distribution of count data that exhibits either over-dispersion or equi-dispersion.

1.2. Objectives of the Study

1.2.1. General Objective

To develop a likelihood-based multiple change point algorithm for count data with allowance for over-dispersion.

1.2.2. Specific Objectives

- To develop a likelihood-based multiple change point algorithm for count data under the Negative Binomial distribution.
- To determine the critical values for the Likelihood Ratio Test for existence of change.
- To test the Negative Binomial multiple changepoint algorithm using simulated data.

2. Methodology

2.1. The Negative Binomial Distribution

In probability and statistics, the negative Binomial distribution is often used to model the number of successes in an infinite sequence of independent and identically distributed Bernoulli trials. The distribution has two parameters r and p , where r is a constant representing a fixed or predefined threshold for the number of successes required and $p \in [0,1]$ is the success probability, which is constant from one Bernoulli trial to the next. The probability distribution, $f(x; r, p)$ of a Negative Binomial random variable has two possible formulations which are contingent on the definition of measure of interest. The version used in this study counts the number of failures, x , before the r -th success. The probability mass function of X is defined by:

$$f(x; r, p) = \frac{\Gamma(x+r)}{x! \Gamma(r)} p^r q^x \quad (1)$$

where $p \in [0, 1]$, $q = 1 - p$, $r > 0$ and $x = 0, 1, 2, 3, \dots$.

The standard formulation of the Negative Binomial distribution is such that the mean and variance of random variable X are both derived quantities whose values are obtained from the primary parameters r and p as:

$$E[X] = \mu = \frac{rq}{p} \text{ and } \text{Var}(X) = \sigma^2 = \frac{rq}{p^2} = \mu + \frac{1}{r} \mu^2 \quad (2)$$

The parameter r is the measure for over-dispersion since the variance of X , σ^2 in Equation (2) exceeds the mean, μ by a function of r .

2.1.1. Parameterization of the Negative Binomial Distribution for the Proposed Change Point Algorithm

Under the formulation of the Negative Binomial distribution with the probability mass function defined as in Equation (1), consider the mean, μ and variance, σ^2 as the primary quantities and the parameters r and p as derived quantities such that the values of r and p can be derived from the mean and variance of the data distribution as:

$$r = \frac{\mu^2}{\sigma^2 - \mu} \text{ and } p = \frac{r}{r + \mu} \quad (3)$$

The factor $1/r$ in variance formula defined in Equation (2) is a sort of “clumping” parameter since as difference in variance and mean decreases, that is $\sigma^2 - \mu \rightarrow 0$ then $r \rightarrow \infty$ and $1/r \rightarrow 0$. In other words as $r \rightarrow \infty$, the variance, $\sigma^2 = \mu + \frac{1}{r} \mu^2$ approaches the mean, μ so that $NB(r, p)$ approaches the Poisson (μ) distribution with both mean and variance equal to the parameter μ . Therefore, under the parameterization given in Equation (3), Poisson is the limiting distribution for the Negative Binomial.

$$\lim_{r \rightarrow \infty} NB\left(r, p = \frac{r}{r + \mu}\right) = \text{Pois}(\mu) \quad (4)$$

A common scenario is where the factor $1/r$ is large. This happens in cases where the variance exceeds the mean such that the data are over-dispersed. It follows that $\sigma^2 - \mu \rightarrow \infty$ then $r \rightarrow 0$ and $1/r \rightarrow \infty$. The larger the factor $1/r$ the greater the amount of over-dispersion [1] and [2]. Equation (4) justifies the use of the negative binomial distribution to model count data that is both over-dispersed and equi-dispersed; hence its advantage over the standard Poisson model for count data.

2.1.2. The Likelihood Function of $NB(r, p)$ Distribution

Suppose we have a sample $X_1, X_2, X_3, \dots, X_N$ from a $NB(r, p)$ distribution with probability distribution $f(x; r, p)$ described as in Equation (1). The likelihood function is obtained as:

$$L(r, p) = \prod_{(i=1)}^N [f(x_i; r, p)] \quad (5)$$

The log-likelihood function is given by:

$$\begin{aligned} l(r, p) &= \sum_{i=1}^N \ln \left\{ \frac{\Gamma(x_i + r)}{x_i! \Gamma(r)} p^r q_i^{x_i} \right\} \\ &= \sum_{i=1}^N \ln(\Gamma(x_i + r)) - N \ln(\Gamma(r)) - \sum_{i=1}^N \ln(x_i!) + Nr \ln(p) + \sum_{i=1}^N x_i \ln(q) \end{aligned} \quad (6)$$

2.2. Dispersion Test

A count data distribution may exhibit any of three kinds of dispersion: under-dispersion, equi-dispersion or over-dispersion. The type of dispersion is dependent on the mean-variance relationship of the sample data. In this study, the variance-to-mean ratio (VMR), otherwise referred to as the dispersion index (D), is determined for the entire sample and for each segment, following which a dispersion test is conducted prior to applying the change point algorithm.

Starting with the assumption that sample count data are equi-dispersed so that the index of dispersion $D = 1$ and the sample are in agreement with a theoretical Poisson series, a dispersion test is conducted with the following hypotheses:

$$H_0 : D = 1 \quad (\text{Data are equi-dispersed})$$

versus

$$H_1 : D > 1 \quad (\text{Data are over-dispersed})$$

versus

$$H_1 : D < 1 \quad (\text{Data are under-dispersed})$$

The index of dispersion is computed from sample statistics as:

$$D = \frac{S^2}{\bar{X}} \quad (7)$$

where S^2 is the sample variance and \bar{X} is the sample mean.

The VMR of the sample data informs the choice of the distribution model to fit the data as summarized in **Table 1**.

Table 1. Mean-variance relationships.

Dispersion index (D)	Dispersion type	Proposed distribution
$D = 0$	Not dispersed	Constant variable
$0 < D < 1$	Under-dispersion	Binomial Distribution
$D = 1$	Equi-dispersion	Poisson Distribution
$D > 1$	Over-dispersion	Negative Binomial Distribution

Tests for significant departure of the index of dispersion from 1 are performed under the Chi-square or Normal distributions contingent on the size of sample data as described in [3]. For small ($n < 30$) samples, the dispersion test statistic is

defined by Equation (8)

$$d = D(n-1) \quad (8)$$

where d is approximated by a Chi-Square distribution with $(n-1)$ degrees of freedom. The decision criteria are such that agreement with the Poisson distribution is accepted if d falls between the chi-square table values at the probability levels $(1-\frac{\alpha}{2})$ and $(\frac{\alpha}{2})$. On the other hand, if d falls below the chi-square table value at the $(1-\frac{\alpha}{2})$ level, the alternative hypothesis of under-dispersion is accepted in favor of H_0 . In cases where d exceeds the chi-square table value at the $(\frac{\alpha}{2})$ level, the alternative hypothesis of over-dispersion is accepted in favor of H_0 .

For large ($n \geq 30$) samples, the test statistic is given by:

$$d = \sqrt{2D(n-1)} - \sqrt{2(n-1)-1} \quad (9)$$

where d is approximately normal. The null hypothesis is rejected based on a comparison of the absolute value of d against the two-sided critical values from the Standard Normal tables for a given level of the test. For a size $\alpha = 0.05$ test, H_0 is rejected if $|d| < -1.96$ or $|d| > 1.96$, in which case the distribution is said to be under-dispersed or over-dispersed, respectively. The change detection and estimation algorithm described in Section 2.3 is applied to over-dispersed and equi-dispersed samples, while under-dispersed samples are discarded.

2.3. The Negative Binomial Multiple Change Point Algorithm

The Negative Binomial Multiple Change Point Algorithm (NBMCPA) algorithm is developed in an iterative process involving recursive binary segmentation, hypothesis testing for existence of statistically significant change, and estimation of the change points, if any, using maximum likelihood approach. According to the parameterization described in Equation (3), the parameter r is defined such that its value depends only on the mean, μ and variance, σ^2 of the data distribution. On the other hand, the parameter p is dependent on parameter r and the mean μ of the distribution. As such, a change in parameter r automatically implies a change in parameter p . Therefore, for simplicity, the NBMCPA will explicitly consider only the hypotheses for a step change in the over-dispersion parameter r .

2.3.1. Assumptions for the Algorithm

The NBMCPA was built under the following data assumptions:

A₁: The sample data arise from a count process, and are therefore discrete.

A₂: There are no temporal dependencies in the sample data.

A₃: A step change in both parameters r and p of the Negative Binomial distribution occurs simultaneously.

2.3.2. The Step-Wise Recursive Binary Segmentation Procedure

The step-wise recursive binary segmentation (SWRBS) procedure, similar to the method discussed in [4] is an iterative process involving 5 elementary steps as illustrated by Figure 1.

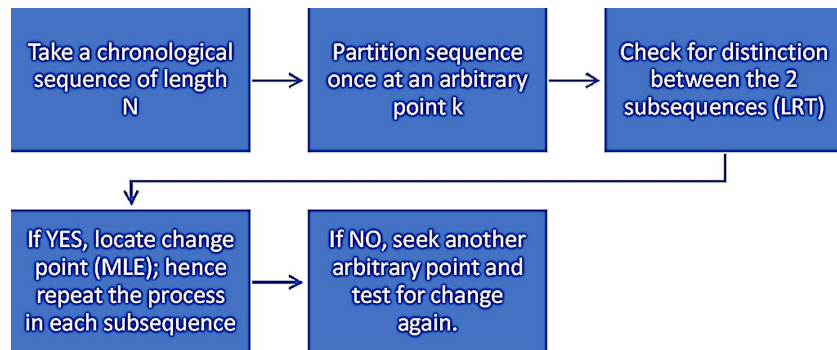


Figure 1. The SWRBS procedure.

Starting with a chronological sequence of length N from the Negative Binomial (r, p) distribution, partition the sequence into two parts at an arbitrary point k corresponding to the observation made at the random time point t_k . Check for the existence of a statistically significant distinction in the over-dispersion parameters r_1 and r_2 of the two sub-sequences by conducting a likelihood ratio test. If the two partitions are found to be significantly different with regard to the distributional parameters, then a change exists at or near the time point t_k . In this case, proceed to estimate the location of the said change using the maximum likelihood method.

Once the first change has been located, repeat the processes of partitioning, hypothesis testing and change point estimation in each of the two new sub-sequences formed. On the other hand, if no change is evident at the first arbitrary point k , seek an alternative arbitrary point and repeat the likelihood ratio test; hence obtain the MLE of the change point, if a change exists. The SWRBS procedure is repeated over and over until no more significant changes are identified.

2.3.3. Partitioning the Data Sequence

Consider a sequence of random calendar time points: $t_0, t_1, t_2, \dots, t_N$. Let x_1, x_2, \dots, x_N be a sequence of observed values or realizations of a stochastic process. Let $[t_0, t_1]$ be the length of the partition between consecutive time points 0 and 1, and $[t_1, t_2]$ be the length of the partition between consecutive time points 1 and 2, and so on. Assume that each partition $[t_i, t_j]$ for $j \in [1, N]$ and $i \in [0, N]$ and $i \neq j$ of the original sequence $\{x_i\}$ consists of a random sub-sequence $\{x_h, x_{h+1}, x_{h+2}, \dots\}$. Assume that the observations in each partition follow a Negative Binomial distribution with parameter r_j for $j \in [1, N]$ as shown in Figure 2.

2.3.4. The Likelihood Ratio Test for Existence of the First Change

An investigation as to whether a change exists at some random point $k \in [2, N - 2]$

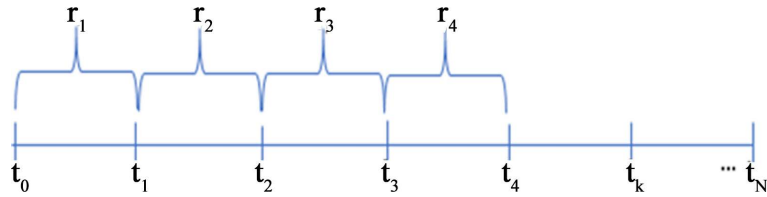


Figure 2. Timeline diagram showing sequence partitions.

in a sequence of N observations is done by conducting a two-tailed likelihood ratio test (LRT). The null hypothesis states that there is no change in the over-dispersion parameter r across the entire sample. The alternative hypothesis seeks a difference in the over-dispersion parameter r between the two data segments, such that the first partition of the sequence in the interval $[t_0, t_k]$ has the parameter r_α while the second partition in the interval $[t_{k+1}, t_N]$ has parameter r_β , where $r_\beta \neq r_\alpha$. The mathematical hypotheses are described as in Equation (10):

$$H_0 : r_1 = r_2 = \dots = r_N = r$$

versus

$$H_1 : r_1 = r_2 = \dots = r_k = r_\alpha \neq r_{k+1} = r_{k+2} = \dots = r_N = r_\beta \tag{10}$$

The log-likelihood function under H_0 is given by Equation (11) as:

$$l_0(\hat{r}, \hat{p} | x_i) = \sum_{i=1}^N \ln \left\{ \frac{\Gamma(x_i + \hat{r})}{x_i! \Gamma(\hat{r})} \hat{p}^{\hat{r}} \hat{q}^{x_i} \right\} \tag{11}$$

where the method of moments estimates (MME) of the model parameters are:

$$\hat{r} = \frac{\bar{X}^2}{S^2 - \bar{X}} \quad \text{and} \quad \hat{p} = \frac{\hat{r}}{\hat{r} + \bar{X}} \quad \text{and} \quad \hat{q} = 1 - \hat{p}$$

The statistics S^2 and \bar{X} are the unbiased estimates of the population variance and mean respectively obtained as:

$$\bar{X} = \frac{\sum_{i=1}^N x_i}{N} \quad \text{and} \quad S^2 = \frac{\sum_{i=1}^N (x_i - \bar{X})^2}{N - 1}$$

The log-likelihood function under H_1 for an arbitrary change point $k \in [2, N - 2]$ is defined in Equation (12).

$$l_1(\hat{r}_\alpha, \hat{r}_\beta, \hat{p}_\alpha, \hat{p}_\beta | x_i) = \sum_{i=1}^k \ln \left\{ \frac{\Gamma(x_i + \hat{r}_\alpha)}{x_i! \Gamma(\hat{r}_\alpha)} (\hat{p}_\alpha)^{\hat{r}_\alpha} (\hat{q}_\alpha)^{x_i} \right\} + \sum_{i=k+1}^N \ln \left\{ \frac{\Gamma(x_i + \hat{r}_\beta)}{x_i! \Gamma(\hat{r}_\beta)} (\hat{p}_\beta)^{\hat{r}_\beta} (\hat{q}_\beta)^{x_i} \right\} \tag{12}$$

where the method of moments parameter estimates for the lower partition of the data (r_α , p_α and q_α) are obtained as:

$$\hat{r}_\alpha = \frac{\bar{X}_k^2}{S_k^2 - \bar{X}_k}, \quad \hat{p}_\alpha = \frac{\hat{r}_\alpha}{\hat{r}_\alpha + \bar{X}_k} \quad \text{and} \quad \hat{q}_\alpha = 1 - \hat{p}_\alpha$$

whereas, for the upper segment of the sample data, the MME of model parameters r_β , p_β and q_β are given by:

$$\hat{r}_\beta = \frac{\bar{X}_{N-k}^2}{S_{N-k}^2 - \bar{X}_{N-k}}, \quad \hat{p}_\beta = \frac{\hat{r}_\beta}{\hat{r}_\beta + \bar{X}_{N-k}} \quad \text{and} \quad \hat{q}_\beta = 1 - \hat{p}_\beta$$

The statistics S_k^2 and \bar{X}_k are the unbiased estimates of the sub-population variance and mean for the first k observations respectively obtained as:

$$\bar{X}_k = \frac{\sum_{i=1}^k x_i}{k} \quad \text{and} \quad S_k^2 = \frac{\sum_{i=1}^k (x_i - \bar{X}_k)^2}{k-1}$$

Similarly, the statistics S_{N-k}^2 and \bar{X}_{N-k} are the unbiased estimates of the sub-population variance and mean for the next $N-k$ observations respectively obtained as:

$$\bar{X}_{N-k} = \frac{\sum_{i=k+1}^N x_i}{N-k} \quad \text{and} \quad S_{N-k}^2 = \frac{\sum_{i=k+1}^N (x_i - \bar{X}_{N-k})^2}{N-k-1}$$

The likelihood ratio statistic at an arbitrary point k takes the form:

$$\hat{\Lambda}_k = -2 \left(l_0(\hat{r}, \hat{p}) - l_1(\hat{r}_\alpha, \hat{r}_\beta, \hat{p}_\alpha, \hat{p}_\beta) \right) \quad (13)$$

The change point $\hat{k} \in [2, N-2]$ corresponding to the time point t_k is estimated such that the LRT statistic in Equation (13), or equivalently its square root, is maximized. Statistical significance of the estimated change point \hat{k} is determined by comparing the maximum value of $\hat{\Lambda}_k$ against the critical value of the LRT developed in Section 3.4. Of interest is to find the optimal value of the likelihood ratio as:

$$\hat{Z}_N = \max_{2 \leq k \leq N-2} \sqrt{\hat{\Lambda}_k} \quad (14)$$

The decision is made such that H_0 in Equation (10) is rejected if the LRT statistic is large so that $\hat{Z}_N > C$. The constant C is a critical value that is determined by the level of the test α , the sample size N , and the null distribution of the likelihood ratio test statistic in Equation (13) as in Gombay and Horvath (1990). Otherwise, small values of \hat{Z}_N such that $\hat{Z}_N \leq C$ indicate that a change exists at the time point t_k , but the change is not statistically significant.

2.3.5. Testing for the Second and Subsequent Change Points

Once the first change point is obtained, the likelihood ratio test is repeated in each of the two data segments formed. As a simple illustration, assume that there are $n = 100$ observations from the Negative Binomial distribution and that the first change point is estimated at time t_{20} . This results in two segments of the data set t_1, \dots, t_{20} and t_{21}, \dots, t_{100} having different parameters. A single change point algorithm would stop at this first change point. However, a multiple change point algorithm proceeds to investigate whether additional change points exist in the data set. This is done by conducting a likelihood ratio test for change in the lower segment (t_1, \dots, t_{20}) and in the upper segment (t_{21}, \dots, t_{100}) of the data, one at a time. To determine the possibility of change in the lower segment, the hypotheses in Equation (15) are tested:

$$H_0 : r_1 = r_2 = \dots = r_{20} = r$$

versus

$$H_1 : r_1 = r_2 = \dots = r_k = r_\alpha \neq r_{k+1} = r_{k+2} = \dots = r_{20} = r_\beta \quad (15)$$

The LRT procedure given in Section 2.3.4 is then followed. Similarly, to detect change in the upper data segment, the LRT procedure is applied with the test hypotheses defined in Equation (16).

$$H_0 : r_{21} = r_{22} = \dots = r_{100} = r$$

versus

$$H_1 : r_{21} = r_{22} = \dots = r_k = r_\alpha \neq r_{k+1} = r_{k+2} = \dots = r_{100} = r_\beta \quad (16)$$

Detection and location of the second and third change points, assuming they both exist, results in further sub-partitions of the data so that there are four segments in total for the entire sequence of size $n = 100$. Each of the four smaller partitions is then tested for change and the splitting process is repeated; hence the name step-wise recursive binary segmentation. The position of any viable change point, t_k must satisfy the inequality $t_2 \leq t_k \leq t_{n-2}$ so that a change point can neither occur at the first nor last two observations in the sequence. Instead the change point must be sandwiched between some two observations. In the problem of multiple change point analysis given a sample of size n , the maximum possible number of change points is $n - 3$, which excludes the first and last two observations.

2.4. Determining the Critical Values for the Likelihood Ratio Test for Existence of Change

The study makes use of the methods described by Gombay and Horvath on the asymptotics of maximum-likelihood ratio-type statistics for testing a sequence of observations for no change in parameters against a possible change while some nuisance parameters may remain constant over time [5]. In particular, Gombay and Horvath obtained extreme value approximations as well as Gaussian-type approximations for the square root of the likelihood ratio in Equation (13). They also approximated the maximum likelihood ratio using Ornstein-Uhlenbeck processes and obtained the upper bounds for the rate of approximation.

To derive critical values of the LRT statistic, this study makes use of the asymptotic distribution of $\sqrt{Z_n}$ as described in Equation (14). Different critical values are obtained for various small and medium sample sizes ($n = 12, 20, 60, 100, 200, \text{ and } 500$), various test sizes ($\alpha = 1\%, 5\% \text{ and } 10\%$).

The asymptotic critical values are derived as follows: Let $0 < \alpha < 1$ be the level of the double-sided Likelihood Ratio Test.

Define:

$$z_n = z_n(1 - \alpha) = \sup \left\{ x : P \left\{ \sqrt{Z_n} \leq x \right\} \leq 1 - \alpha \right\}$$

Define also:

$$\begin{aligned} r(h, l) &= r(h, l; 1 - \alpha) \\ &= \sup \left\{ x : P \left\{ \sup_{h \leq t \leq l} \left\{ B^{(d)}(t) / (t(1-t)) \right\}^{1/2} \leq x \right\} = 1 - \alpha \right\} \end{aligned} \quad (17)$$

Then, according to Gombay *et al.*, if the null hypothesis of no change holds and $h(n)$ and $l(n)$ are chosen such that both exceed $1/n$ then

$$\lim_{n \rightarrow \infty} P\{\sqrt{Z_n} > r(h(n), l(n))\} = \alpha$$

So that $r(h(n), l(n))$ is an asymptotically correct critical value of size α .

It can be shown that for all $0 < h < 1 - l < 1$.

$$\sup_{h \leq t \leq 1-t} \left(\frac{B^{(d)}(t)}{t(1-t)} \right)^{1/2} = \sup_{0 \leq t \leq \log((1-h)(1-l)/hl)} \Delta(t).$$

Such that for any $T > 0$

$$P\left\{ \sup_{0 \leq t \leq T} \Delta(t) > x \right\} = \frac{x^d \exp(-x^2/2)}{2^{d/2} \Gamma(d/2)} \left\{ T - \frac{d}{x^2} T + \frac{4}{x^2} + O\left(\frac{1}{x^4}\right) \right\} \quad (18)$$

The approximations for the distributions of $r(h(n), l(n))$ and $\sqrt{Z_n}$ proposed by Gombay *et al.* were applied to data assumed to follow the exponential, Poisson and Normal distributions [6]. Further, the approximations were based on a convenient choice of the parameters h and l such that

$$h(n) = l(n) = \frac{(\log n)^{3/2}}{n}$$

The value of parameter T was dependent on the choice of h and l as:

$$T = \log\left(\frac{(1-h)(1-l)}{hl}\right)$$

This study extends the method discussed in Gombay *et al.* to data obtained from the Negative Binomial distribution. Alternative values of $h(n)$ and $l(n)$ are tested in an attempt to get the best asymptotic critical values.

The assumptions made are that the limiting distribution of $\sqrt{Z_n}$ is the double exponential distribution, which is achieved through careful selection of the negative binomial parameters during simulations. Particularly, the parameters should be such that the amount of over-dispersion is neither extreme nor negligible. Further, moderate values for r and p should be chosen to ensure that the distribution exhibits properties conducive to convergence. Choosing r values that are not too small helps to avoid excessive variability and p is neither too high nor too low to maintain a balanced success rate.

3. Results and Discussion

3.1. The Multiple Change Point Algorithm

The negative binomial multiple change point algorithm was developed as a set of 9 steps based on the model equations discussed in the foregoing methodology. The sequential iterative steps are outlined here below:

Step 1

Input or load the sample data into a statistical software such as R. These data

could be either real observed count data or simulated data from a Negative Binomial Distribution.

Step 2

Compute the sample mean \bar{X} and variance S^2 for these data. Conduct a test for dispersion. If data are over-dispersed or equi-dispersed, proceed to estimate parameters, otherwise discard the sample. Using the sample mean and variance, find the method of moments estimates of the two parameters \hat{r} and \hat{p} of the negative binomial model according to Equation (11).

Step 3

Using the sample data and the parameter estimates in Step 2, compute the log-likelihood functions under the null and alternative hypotheses for various values of the arbitrary change point k as described in Equations (11) and (12) respectively. Next, compute the likelihood ratio statistic $\hat{\Lambda}_k$ as described in Equation (13) for each possible value of the change point k . Store all computed values of the likelihood ratio statistic in a single vector.

Step 4

Plot a graph of the likelihood ratio statistic $\hat{\Lambda}_k$ or equivalently, its square root, against the possible values of k . A line plot sufficiently shows at a glance the behavior of the likelihood ratio over sequential values of k .

Step 5

Investigate whether a change exists at some point k by visually inspecting the graph in Step 4. As a guide, when there is no change the likelihood ratio statistic does not have a unique maximum point. On the other hand, when a change exists at some point k , the graph of the likelihood ratio test statistic is such that it reaches a maximum value exactly or in the neighborhood of the point k . **Figure 3** and **Figure 4** represent illustrative graphs of a likelihood ratio test statistic in the absence of change and in the presence of a single change respectively.

Step 6

In case there is no change in model parameters for the given dataset, the change point algorithm comes to a stop and no further change points are sought. However, if a change exists, approximate the location of the change point using the maximum likelihood approach as the point k at which the statistic in Equation (14) attains a maximum.

Step 7

Once a change has been detected and its location estimated, determine whether the said change point is statistically significant by conducting a likelihood ratio test. The null hypothesis of no significant change is rejected if the test statistic in Equation (14) exceeds the critical value.

Step 8

If a change is found to be statistically insignificant, the algorithm comes to a stop and no further change points are sought. However, if a change is found to be significant at some point k , the current value of k is stored as a change point.

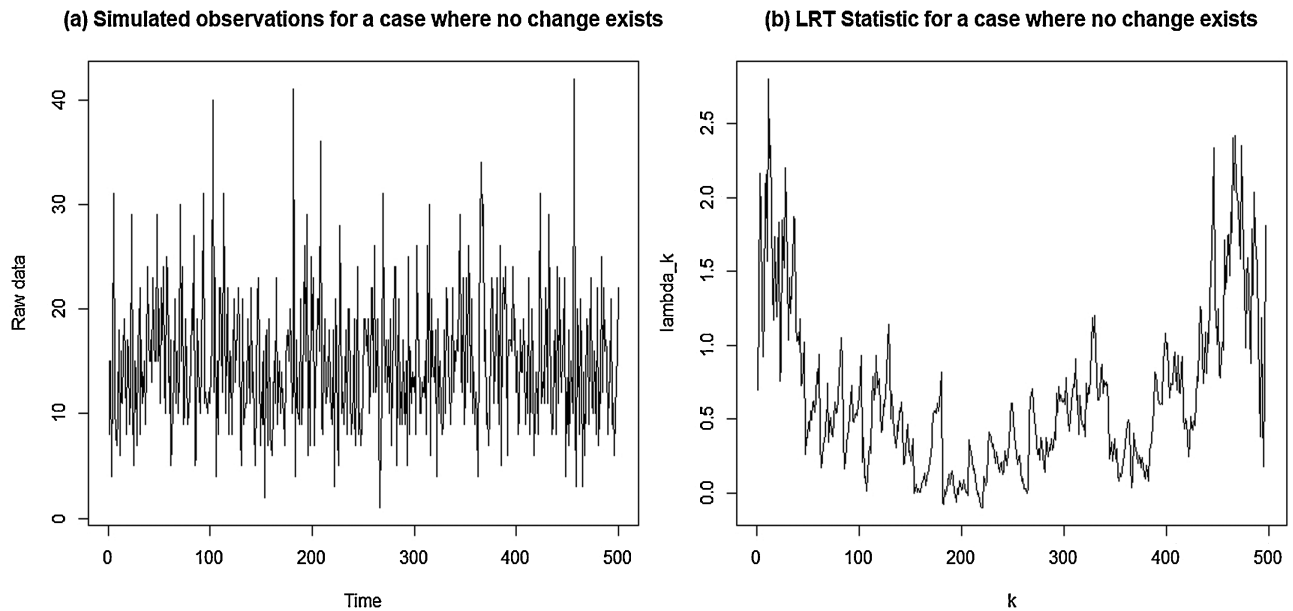


Figure 3. Sample graph showing a case of no change.

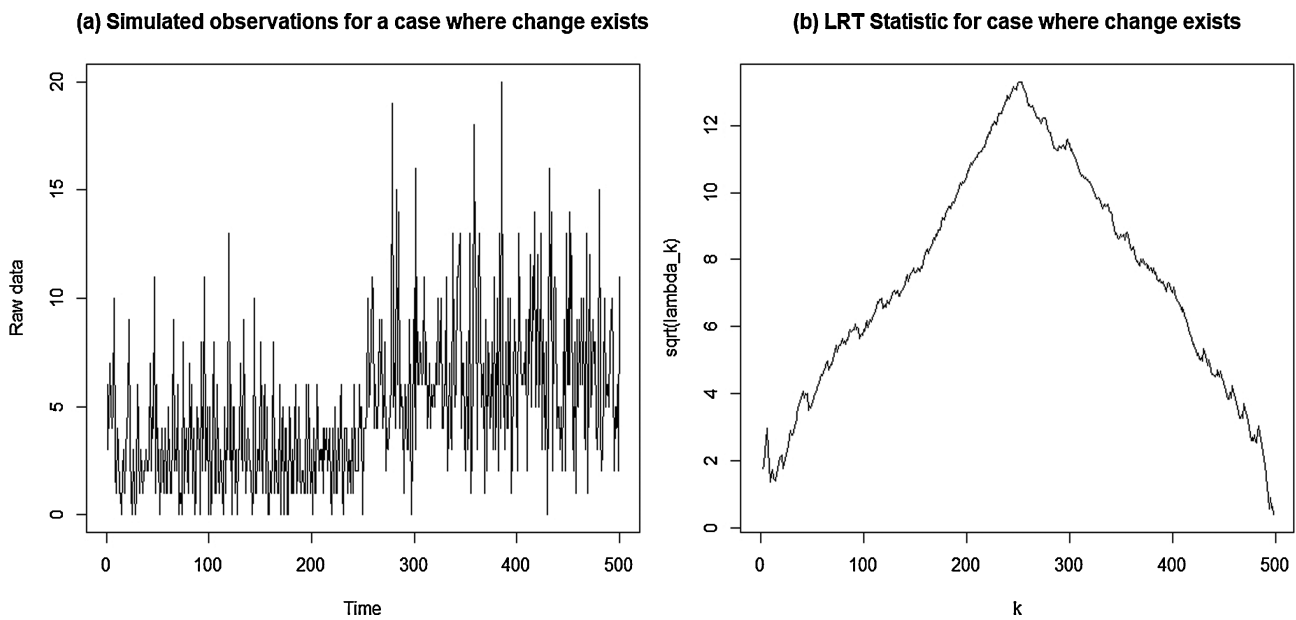


Figure 4. Sample graph showing a case of a single change.

Step 9

Partition the input data set into two segments with the boundary corresponding to the change point estimate k . Each segment is then taken, one at a time, and treated as the input data set in Step 1 then checked for the existence of change and the process repeats until no further change points are found.

Given this is an iterative process, the estimated values of k at each iteration must be recorded and stored. This process constitutes the step-wise recursive binary segmentation procedure discussed in Section 2. Figure 5 represents a model flow chart for the Negative Binomial Multiple Change Point Algorithm.

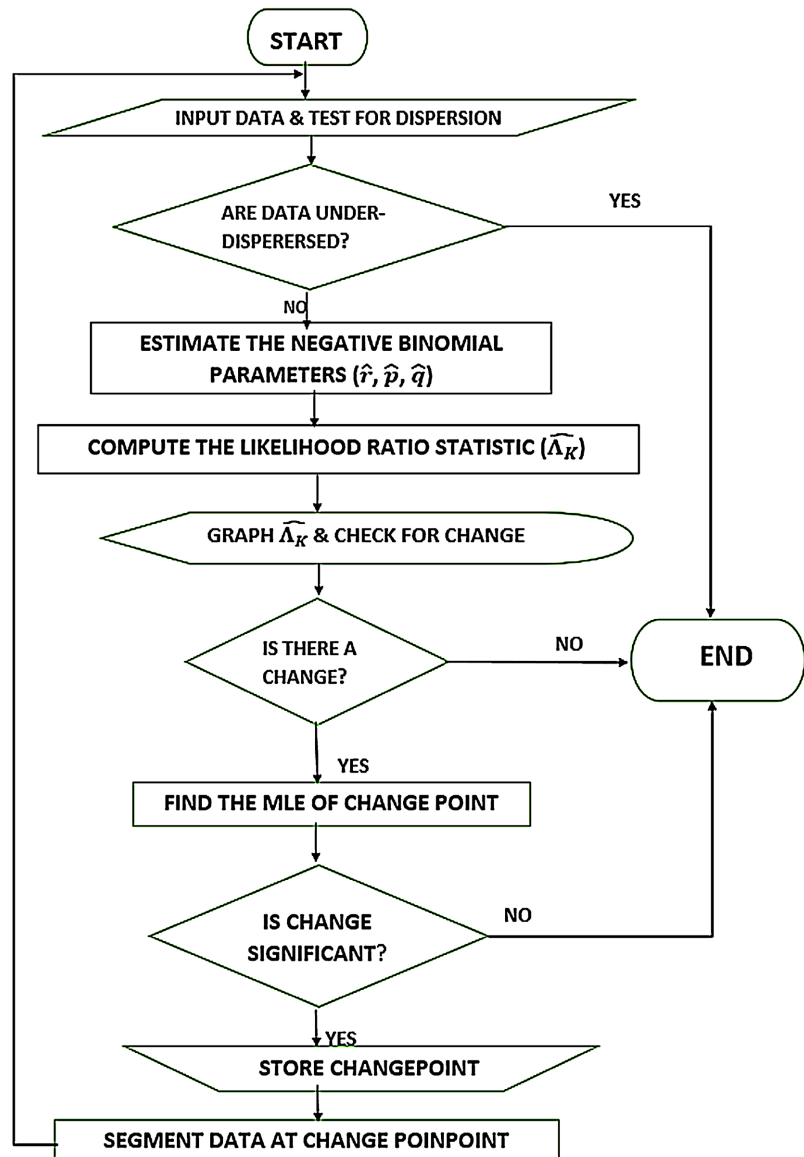


Figure 5. Schematic representation of the Negative Binomial Multiple Change Point Algorithm.

3.2. Critical Values of the Likelihood Ratio Test

This section refers back to the methods described in Subsection 2.4. The method proposed by Gombay and Horvath (1996) on the asymptotic distribution of the likelihood ratio test is applied in the determination of the critical values presented according to Equation (18) which states that for all $T > 0$

$$P\left\{\sup_{0 \leq t \leq T} \Delta(t) > x\right\} = \frac{x^d \exp(-x^2/2)}{2^{d/2} \Gamma(d/2)} \left\{T - \frac{d}{x^2} T + \frac{4}{x^2} + O\left(\frac{1}{x^4}\right)\right\} \quad (19)$$

Various combinations of $h(n)$ and $l(n)$ were tested for the current study, ensuring the condition $0 < h < 1 - l < 1$ was met. The choices of h and l selected were:

$$h(n) = l(n) = \frac{(\log n)^{1.5}}{n}$$

The parameter d in Equation (19) was taken as $d = 2$ to equal the number of unknown parameters in the Negative Binomial model. The value of parameter T was dependent on the choice of h and l . Since h and l were chosen to be equal, then T was computed as:

$$T = \log\left(\frac{(1-h)(1-l)}{hl}\right) = \log\left(\frac{(1-h)^2}{h^2}\right)$$

The asymptotic critical values for $\sqrt{Z_n}$ were determined in the R environment for 1000 iterations as the roots of Equation (19) for various sample sizes n and different levels of significance α . The results are summarized in **Table 2**.

Table 2. Critical values for the LRT at varying sample sizes and levels of the test.

Sample size (n)	Level of the test (α)	Critical value (C)
$n = 12$	0.10	2.900
	0.05	3.184
	0.01	3.735
$n = 20$	0.10	3.019
	0.05	3.294
	0.01	3.830
$n = 60$	0.10	3.209
	0.05	3.467
	0.01	3.978
$n = 100$	0.10	3.275
	0.05	3.527
	0.01	4.029
$n = 200$	0.10	3.349
	0.05	3.594
	0.01	4.086
$n = 500$	0.10	3.428
	0.05	3.666
	0.01	4.148

The critical values obtained were compared against the critical values obtained in [5]. Gombay and Horvath obtained critical values for data derived from the Poisson distribution with a mean of 10. For comparability, we chose the Negative Binomial parameters as $r = 10$ and $p = 0.5$ to achieve a mean of 10, similar to the Poisson distribution. **Table 3** shows a sample of critical values derived under

the Negative Binomial distribution against those presented by Gombay and Horvath for various sample sizes. The two sets of critical values are fairly consistent. An assessment of how sensitive the critical values are to changes in model parameters indicated no significant difference provided the double exponential limit distribution assumption holds.

Table 3. Critical values for the Negative Binomial LRT versus Gombay's critical values.

Sample size (n)	Level of the test (α)	Asymptotic critical values	Gombay's Critical values
$n = 20$	0.10	3.019	3.11
	0.05	3.294	3.60
	0.01	3.830	4.70
$n = 50$	0.10	3.183	3.18
	0.05	3.443	3.62
	0.01	3.958	4.69
$n = 100$	0.10	3.275	3.23
	0.05	3.527	3.64
	0.01	4.029	4.57
$n = 500$	0.10	3.428	3.31
	0.05	3.666	3.69
	0.01	4.148	4.54

3.3. Results of the Simulation Study

3.3.1. Particulars of the Simulation Study

The change point algorithm developed in Section 3 was tested using simulated data from a Negative Binomial distribution. Monte Carlo simulations were performed using the R software to showcase three scenarios: a case of no change, followed by a case of a single change, and finally a case of multiple changes in distribution parameters. In the case of single and multiple changes, synthetic datasets were generated such that the true change points were known. This allowed the comparison of detected points against the true values. The single change point study considered synthetic data for small and medium samples of size $n = 12$, $n = 20$, $n = 60$, $n = 100$, $n = 200$ and $n = 500$. The change points were set at each of the locations: $n/4$, $n/2$, and $3n/4$ where parameters r and p were assumed to change simultaneously, while the distributional form remained the same throughout the samples. For simplicity, the multiple change-point study considered only a case of two changes located at positions $n/4$ and $3n/4$ for a sample of size n .

The choices of Negative Binomial model parameters r and p for the different segments in the simulation were made to demonstrate the impact of different parameters of the Negative Binomial distribution on the generated data. Different

r and p combinations led to varying degrees of dispersion in the counts, providing an opportunity to analyze over-dispersion or changes in variability over the two segments. The rationale behind the specific parameter choices was to create diversity in data generation and illustrate change points while mimicking real-world scenarios in count data distributions.

In the no-change setting, a single set of parameters $r = 5$ and $p = 0.2$ were chosen for the entire data series. Setting $r = 5$ allows simulation of scenarios where 5 successful outcomes are expected (like 5 successful sales, recoveries) before stopping the trials. A moderate value of r led to a manageable amount of variability in the data, allowing the model to capture enough complexity without becoming overly simplistic or excessively complex. The choice of parameter $p = 0.2$ indicated a 20% chance of success in each individual trial. A lower probability of success leads to a larger number of trials being needed to achieve the desired amount of successes. This aligns with real-world scenarios where events may be rare, leading to larger counts of failures and capturing the essence of over-dispersed count data.

In the single change point setting, the parameters for the two segments were chosen as follows:

Segment 1: $r = 8$ and $p = 0.4$

The parameter choice for this segment indicates that you expect to see 8 successful outcomes (such as recoveries and disease incidences) before stopping the trials, with a success probability of 40%. A higher r value can produce more variability in the counts, reflecting scenarios where events are more frequent or where a greater number of successes are desirable before considering a stopping point. A moderate success probability reflects a reasonably likely event, which may be indicative of a favorable condition.

Segment 2: $r = 5$ and $p = 0.2$

The values of r and p were chosen such that 5 successes are expected with only a 20% chance of success for each trial. A lower r values in this segment may model a scenario where successes are less common, reflecting a different underlying process or condition. It can lead to less variability in counts, which might be appropriate if the success rate has significantly changed. A lower success probability in this segment means that more trials are needed to achieve the same number of successes, creating a higher variance in the outcomes. This can model a situation where conditions are critical, making successes more challenging to achieve.

In the multiple (two) change point setting, the parameter values for the three segments were chosen as follows:

Segment 1: $r = 8$ and $p = 0.4$

This segment simulates random observations where you expect to achieve 8 successes, with a 40% chance of success for each trial. A higher r value with a much lower p results in increased over-dispersion, which means greater variability in counts. This setting simulates a scenario where the stochastic process is more

unpredictable, leading to larger counts and more fluctuations.

Segment 2: $r = 3$ and $p = 0.3$

In this segment, data are simulated such that only 3 successes are expected with a 30% chance of success for each trial. The lower r suggests a shift to a less successful outcome scenario, perhaps indicating a less favorable condition. A p value of 0.3 suggests that while successes are still possible, they are less frequent than in the first segment, leading to a greater proportion of failures compared to successes.

Segment 3: $r = 5$ and $p = 0.2$

This segment is such that 5 successes are expected with a 20% chance of success per trial. A moderate r combined with a lower p reflects an even more challenging scenario for achieving successes, indicating a substantial decline in the likelihood of success compared to the previous segments. This can model conditions that have deteriorated significantly, resulting in fewer successful outcomes.

Overall, the parameter values were chosen to reflect a range of potential real-world scenarios where you might expect changes in counts due to different influencing factors (such as changes in policy, market conditions, or environmental factors). The chosen values for r and p led to a clear, yet not visually obvious, distinction between data segments, making it easier to demonstrate effectiveness of the change point detection algorithm developed.

This section gives a summary of results for the simulation study and power analysis showcasing performance of the NBMCPA given different sample sizes, model parameters, and location of change points. Throughout the simulation, the Likelihood ratio tests were conducted at the 5% level of significance. Visualizations of the likelihood ratio statistic are displayed for small and medium sample sizes and for each of three predefined change-point locations. The maximum values of the LRT statistics were compared to the critical values of the likelihood ratio test obtained in 3 to check for significance of change.

3.3.2. A Case Where No Change Exists

A random sample was generated under the Negative Binomial ($r = 5$, $p = 0.2$) distribution for a case of no change. The entire data set was such that both parameters r and p of the distribution remained constant throughout the series. **Figure 6** shows the graph of $\sqrt{\Lambda_k}$ for a sample of size $n = 500$.

The graph of the Likelihood Ratio Test statistic (as illustrated in **Figure 6** panel (b)) does not exhibit a unique or single maxima. Instead, the graph appears rugged, somewhat similar to the graph of raw data as displayed in panel (a). This indicates that no change is detected. In addition to the lack of a unique maxima on the graph, the absence of variation in the distribution parameters is emphasized by the fact that the highest point on the graph of the LRT statistic lies below the critical value, $C = 3.666$ at $\alpha = 0.05$ (see the dashed horizontal line). It is concluded that at the 5% level of significance, there is no statistically significant change in the distribution parameters.

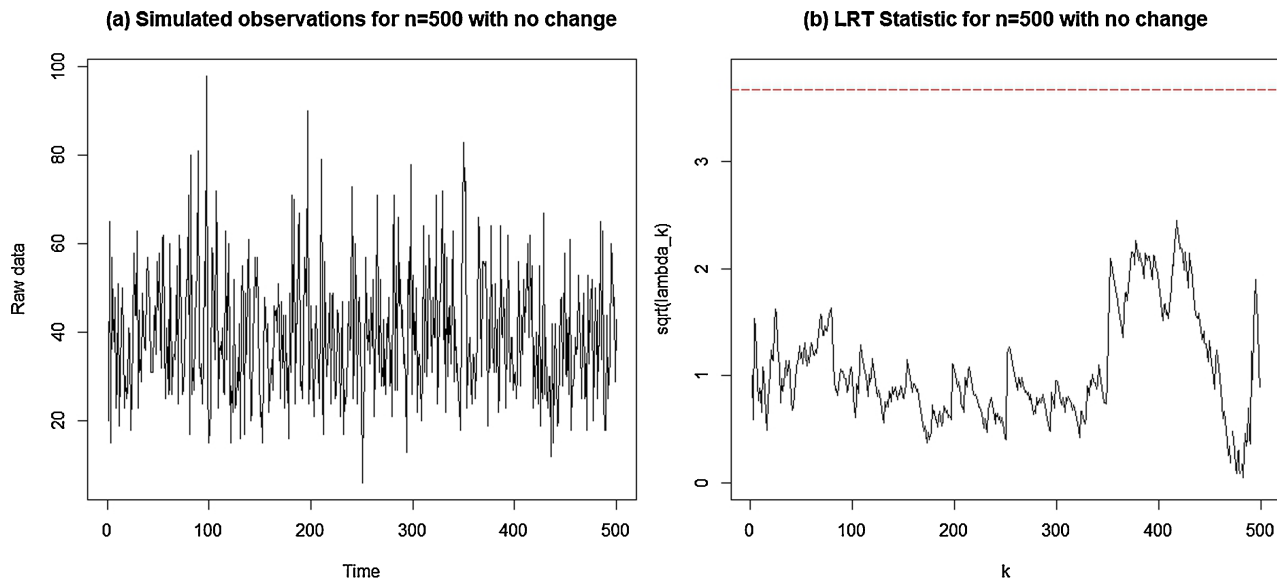


Figure 6. Graph showing a case of no change for a sample size $n = 500$.

3.3.3. A Case Where a Single Change Exists

Synthetic data were obtained randomly from the Negative Binomial distribution with three preset change points for various small and medium sample sizes, n . The change points were set such that both parameters r and p changed only once in the entire series at either of the locations $n/4$, $n/2$ or $3n/4$. However, the data segments formed both followed the Negative Binomial distribution. The change point estimates were obtained using the NBMCP algorithm and the results summarized in **Table 1**. The graphs of the square root of the LRT statistics, $\sqrt{\Lambda_k}$ with superimposed critical values are shown in **Figures 7-15**.

3.3.4. When the Sample Size Is $n = 50$

Figures 7-9 represent the raw simulated data (panel a)) and the results of the Likelihood Ratio Test (panel (b)) when the change point is set at $n/4$, $n/2$ and $3n/4$ respectively. The estimated change point in each case is indicated on the graph of the likelihood ratio statistic $\sqrt{\Lambda_k}$ using a vertical line. The horizontal dashed lines indicate the critical values of the test, used to determine whether or not a change, if present, is significant. The results displayed in **Figure 7** showed that when the change point was set at position $n/4$, the MLE of the change point was $\hat{k} = 13$. On the other hand, when the change was set at the middle of the series, $n/2 = 25$, (see **Figure 8**) the change point estimate was $\hat{k} = 25$. When the change point was set further away from the initial data point, at position $3n/4$ as shown in **Figure 9**, the change point estimate was $\hat{k} = 38$.

3.3.5. When the Sample Size Is $n = 200$

Figures 10-12 represent the raw simulated data and the results of the likelihood ratio test when the change point is set at $n/4$, $n/2$ and $3n/4$ respectively for $n = 200$. The estimated change point in each case is indicated on the graph of the likelihood ratios ($\sqrt{\Lambda_k}$) using vertical lines. The maximum points on the graph

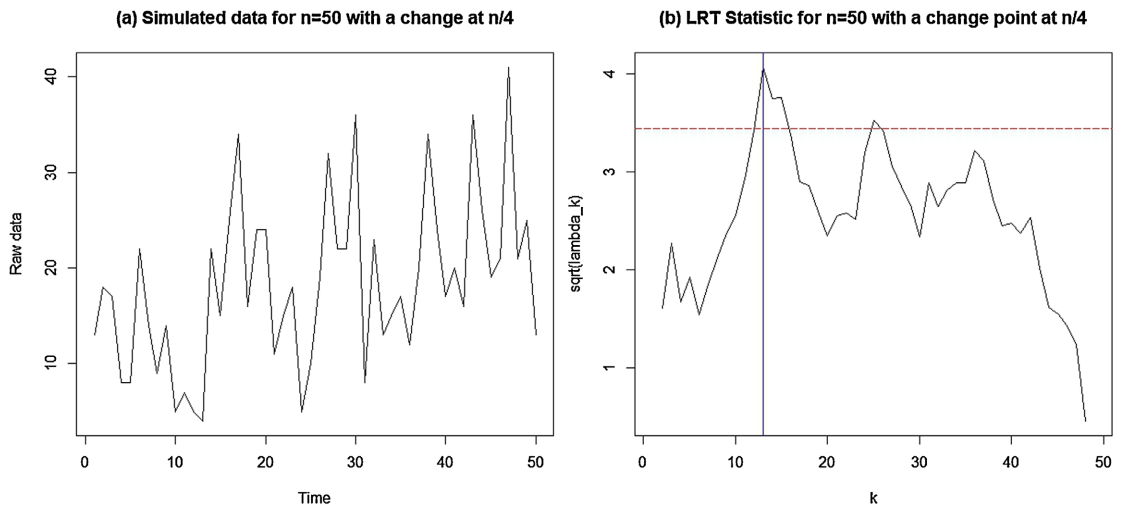


Figure 7. Graph showing a case of a single change at $n/4$ for a sample size $n = 50$.

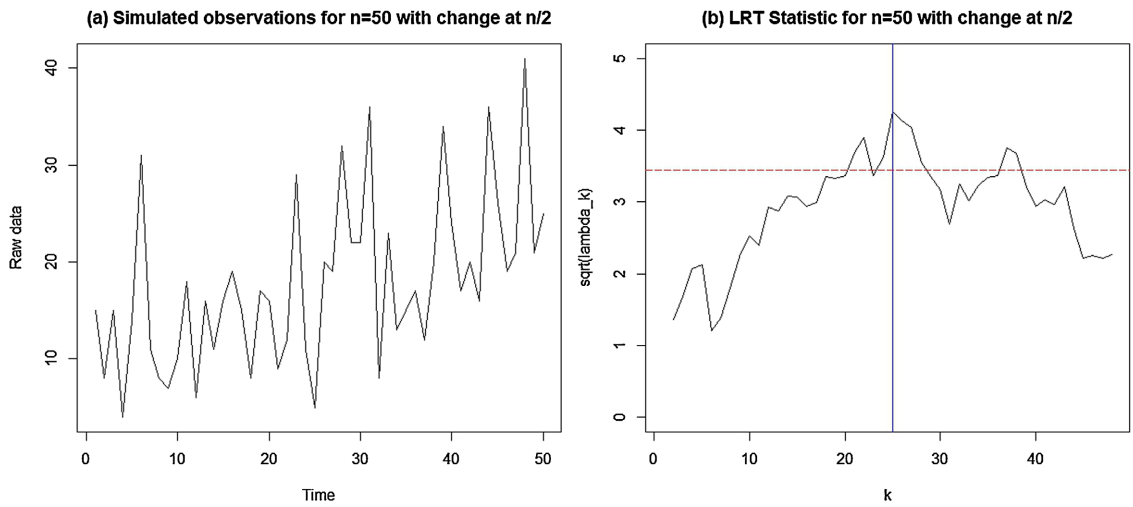


Figure 8. Graph showing a case of a single change at $n/2$ for a sample size $n = 50$.

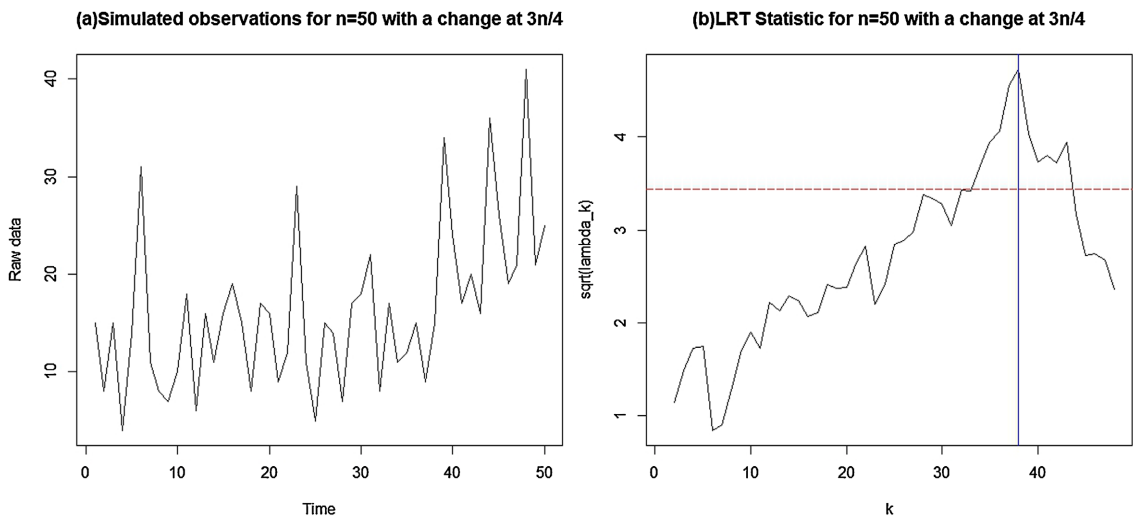


Figure 9. Graph showing a case of a single change at $3n/4$ for a sample size $n = 50$.

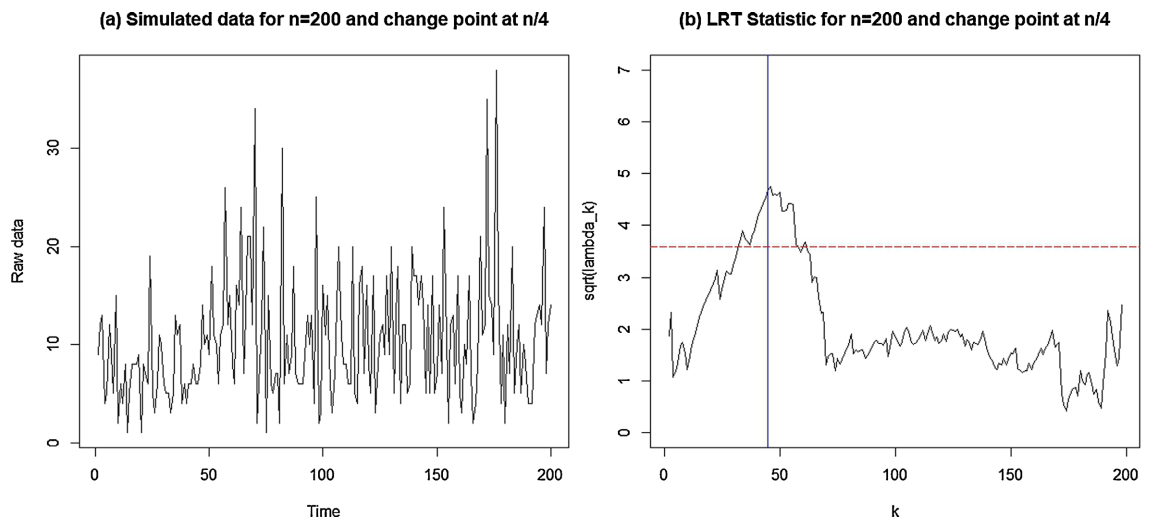


Figure 10. Graph showing a case of a single change at $n/4$ for a sample size $n = 200$.

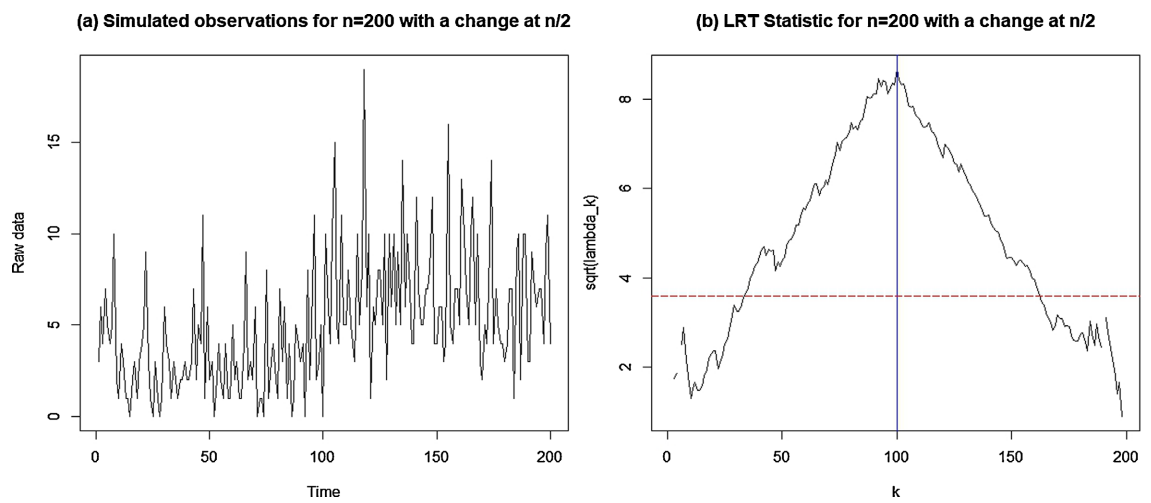


Figure 11. Graph showing a case of a single change at $n/2$ for a sample size $n = 200$.

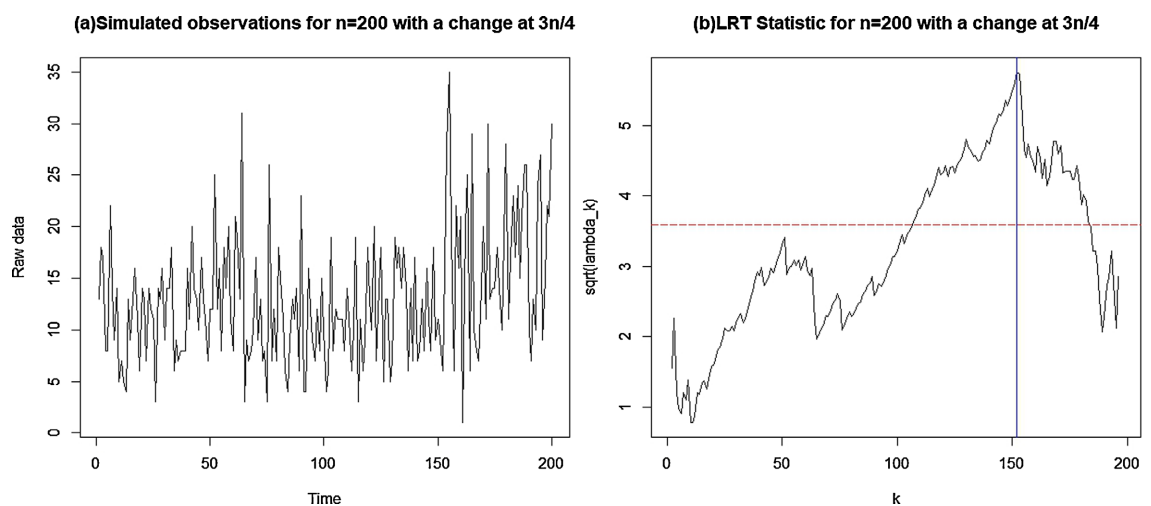


Figure 12. Graph showing a case of a single change at $3n/4$ for a sample size $n = 200$.

of the LRT statistic in each case exceed the critical value $C = 3.594$ indicating that the changes are significant at the 5% level.

3.3.6. When the Sample Size Is $n = 500$

Figures 13-15 represent the raw simulated data and the results of the likelihood ratio test when the change point is set at $n/4$, $n/2$ and $3n/4$ respectively for $n = 500$. The estimated change point in each case is indicated on the graph of the likelihood ratio statistic ($\sqrt{\Lambda_k}$) using vertical lines. The horizontal dashed lines indicate the critical values of the test statistic at the 0.05 level of significance.

Table 4 gives a summary of the estimated change points for different sample sizes ($n = 12, 20, 60, 100, 200, 500$) located at various points for the foregoing

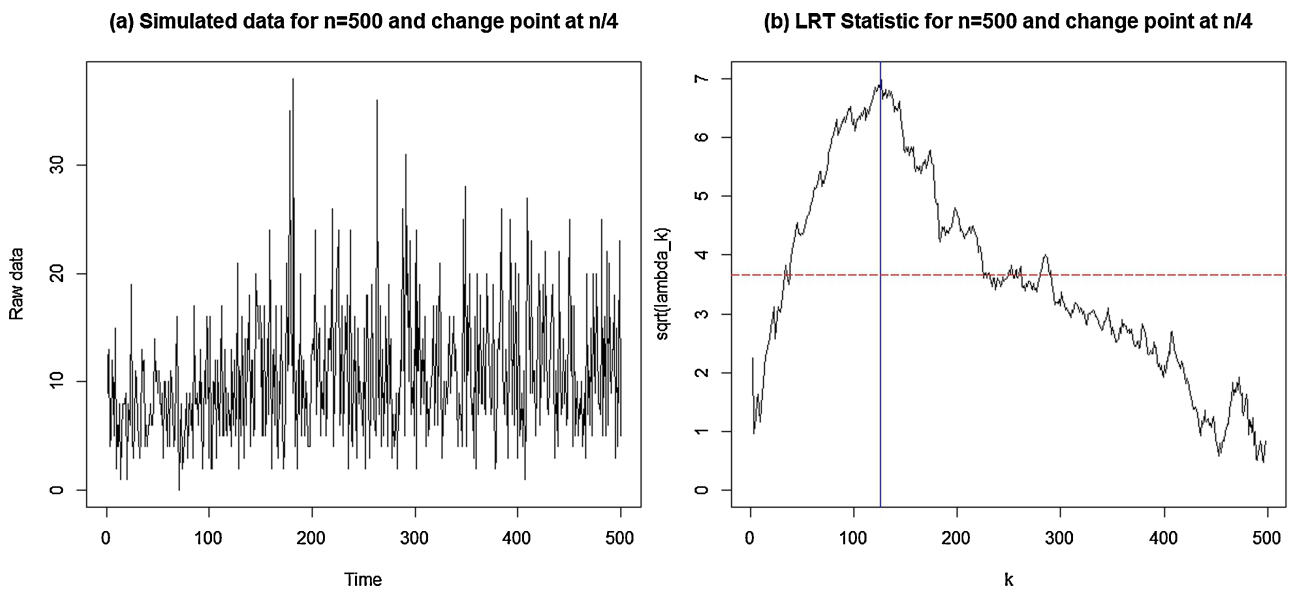


Figure 13. Graph showing a case of a single change at $n/4$ for a sample size $n = 500$.

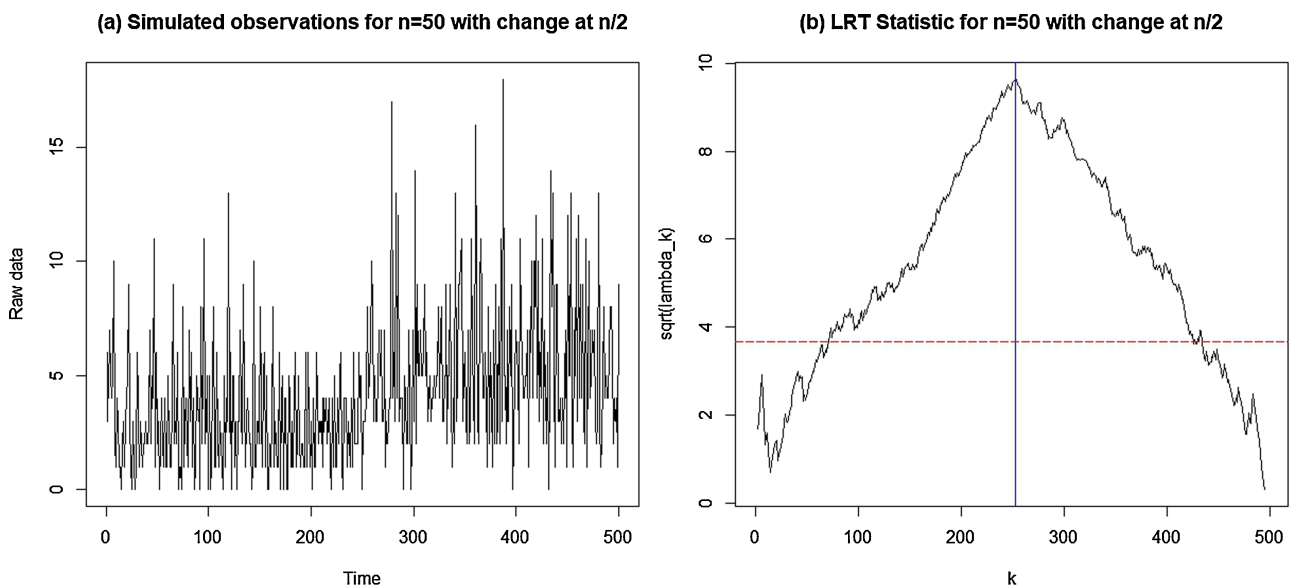


Figure 14. Graph showing a case of a single change at $n/2$ for a sample size $n = 500$.

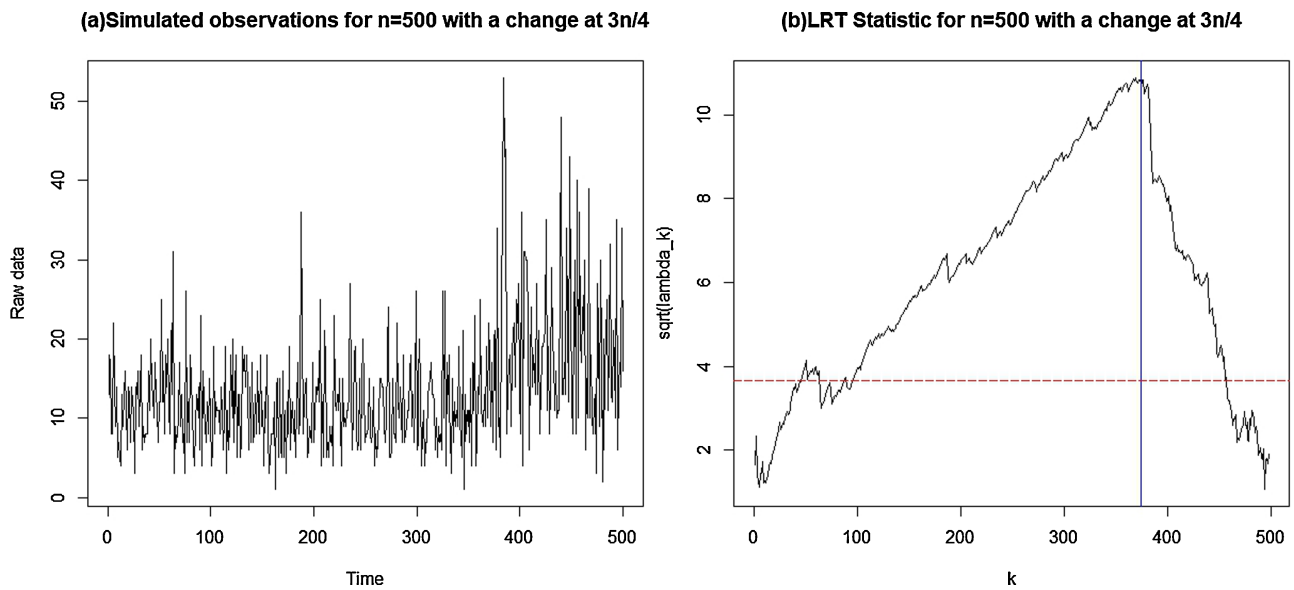


Figure 15. Graph showing a case of a single change at $3n/4$ for a sample size $n = 500$.

Table 4. Actual versus estimated changepoints across sample sizes and change locations.

Sample size (n)	Position of change	Actual cpt (k)	Estimated cpt (\hat{k})	Error (Δ)
$n = 12$	$n/4$	3	2	1
	$n/2$	6	5	1
	$3n/4$	9	9	0
$n = 20$	$n/4$	5	4	1
	$n/2$	10	9	1
	$3n/4$	15	15	0
$n = 60$	$n/4$	15	14	1
	$n/2$	30	30	0
	$3n/4$	45	45	0
$n = 100$	$n/4$	25	24	1
	$n/2$	50	50	0
	$3n/4$	75	75	0
$n = 200$	$n/4$	50	50	0
	$n/2$	100	100	0
	$3n/4$	150	150	0
$n = 500$	$n/4$	125	125	0
	$n/2$	250	250	0
	$3n/4$	375	375	0

single change-point scenario. The estimation error was calculated as the difference between the estimated change point and the true changepoint. The results showed that the algorithm detects and locates changes with very little ($\Delta = \pm 1$) or no ($\Delta = 0$) error. Changes located further away from the initial data point were more accurately estimated for small sample sizes.

3.3.7. A Case Where Multiple Changes Exist

The NBMCP algorithm was found to work well for the single change point case as indicated by the results in **Table 4**. The same algorithm was tested through a simulation study for a multiple change point case with a sample of size $n = 500$. For simplicity, two change points were specified quarter-way (at $n/4 = 125$) and three-quarter-way (at $3n/4 = 375$) respectively in the simulated data.

3.3.8. Location of First Change Point

The NBMCP algorithm detects changes by the order of their magnitude such that the first change point detected and estimated is the most pronounced among all changes present. The first change point lies at the point where the likelihood ratio test statistic first attains a maximum value. The vertical blocked line in **Figure 16** shows the estimate of the first change point, which corresponds to the $3n/4$ value. Potential additional change points in the data series only appear as peaks on the graph of the likelihood ratio statistic but are not marked as change points at this stage.

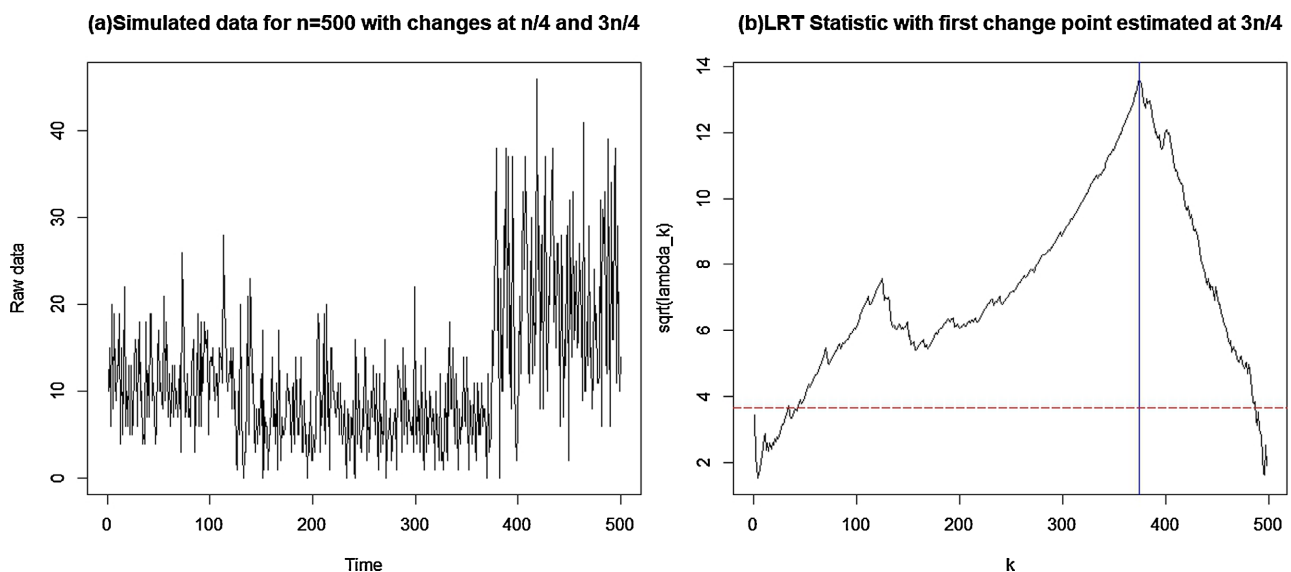


Figure 16. Simulated observations (a) and Likelihood ratio (b) showing the location of first change point at position $3n/4$ for $n = 500$.

3.3.9. Location of Second Change Point

Once the first change point is identified, the algorithm splits the time series into two parts at the first estimated change point. Change detection is then conducted in the lower partition and upper partitions provided the segments are not under-dispersed. The lower segment was discarded due to under-dispersion. **Figure 17**

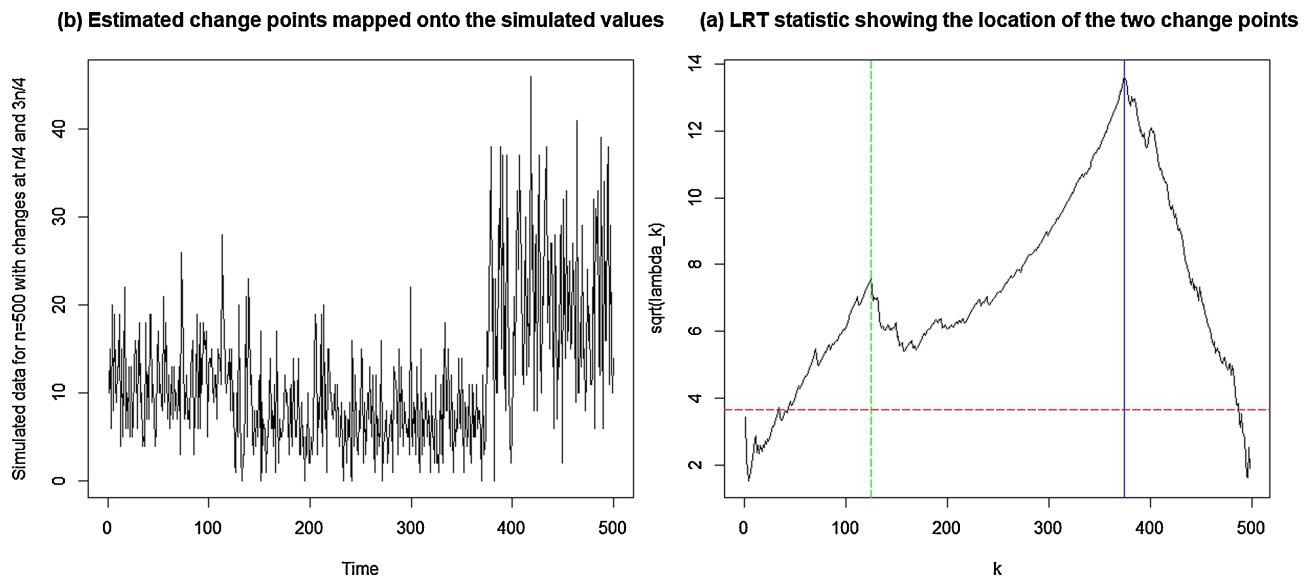


Figure 17. The location of the first change point at position $3n/4$ (vertical blocked line) and the second change point at $n/4$ (vertical dashed line) for $n = 500$.

shows estimates of the first change point (vertical blocked line) and the second change point (vertical dashed line).

A dispersion test of the resulting three segments, given the two change points located, showed that the samples were under-dispersed and hence discarded. No further changes were sought in the lower, middle and upper sub-partitions of the data. Following the dispersion test, the algorithm comes to a halt and only two change points are reported.

3.4. Power Analysis

Investigations into power of the Likelihood Ratio Test for existence of a change were conducted via a simulation study at the 5% level of significance. A null hypothesis of no change in the distributional parameters between any two data segments was tested against the alternative that a change in distributional parameters exists so that the data series has two segments, each following a negative binomial distribution but with varying dispersion parameters. Mathematically, the test hypotheses were given by Equation (20)

$$H_0 : r_1 = r_2 = \dots = r_n = r \text{ (No change)}$$

versus

$$H_1 : r_1 = r_2 = \dots = r_k = r_\alpha \neq r_{k+1} = r_{k+2} = \dots = r_n = r_\beta \text{ (change exists)} \quad (20)$$

Several datasets were generated from the negative binomial distribution under the alternative hypothesis and the LRT statistic in Equation (13) was computed for each dataset. The null hypothesis was rejected when the LRT statistic defined in Equation (14) fell short of the critical value for a given sample size (see **Table 2**). The proportion of times in the null hypotheses was correctly rejected constituted the power of the test, otherwise regarded as the probability of not making a

Type II error $(1 - \beta)$.

The simulations were performed over 1000 iterations for each sample size and change point location to investigate the algorithm's ability to detect and correctly estimate change points under different conditions. Power was calculated as the proportion of simulations where the change point was correctly detected.

$$\text{Power} = \frac{\text{Number of times the change point is correctly detected}}{\text{Total number of simulations with a change point}}$$

Sensitivity analysis of the test was performed in a single-change setting with change located at different positions ($n/4$, $n/2$ and $3n/4$) and for varying sample sizes ($n = 12$, $n = 20$, $n = 60$, $n = 100$, $n = 200$, and $n = 500$).

Power results for each combination of sample size and change point were stored and used to determine the optimal conditions for change detection through a comparative analysis. A tolerance level of 1 was considered in the analysis. **Table 5** summarizes the results of the power analysis for various sample sizes and locations of single change.

Table 5. Power of the likelihood ratio test for change.

Sample size (n)	Position of change	Test power
$n = 12$	$n/4$	0.084
	$n/2$	0.166
	$3n/4$	0.103
$n = 20$	$n/4$	0.311
	$n/2$	0.424
	$3n/4$	0.367
$n = 60$	$n/4$	0.404
	$n/2$	0.434
	$3n/4$	0.426
$n = 100$	$n/4$	0.477
	$n/2$	0.496
	$3n/4$	0.494
$n = 200$	$n/4$	0.498
	$n/2$	0.533
	$3n/4$	0.524
$n = 500$	$n/4$	0.513
	$n/2$	0.537
	$3n/4$	0.533

The results showed that the likelihood ratio test was most powerful in detecting and locating change when the changepoint was located midway through the data

set.

Change detection accuracy, in terms of the true positive rate, of the algorithm was higher for changes positioned three-quarter-way compared to when change was located closer to the first data point (quarter-way). These results were consistent with those presented in [7] and [8]. In addition, the power of the LRT was found to increase with the sample size with highest detection accuracy of 53.7% being recorded when the sample size was $n = 500$. Additional analysis showed that the test was most powerful when changes were larger, so that there was greater distinction between the segments. However, the test performed well even for subtle changes, especially larger sample sizes. Figure 18 visualizes the trend in test power as the sample size increases and change point location is varied.

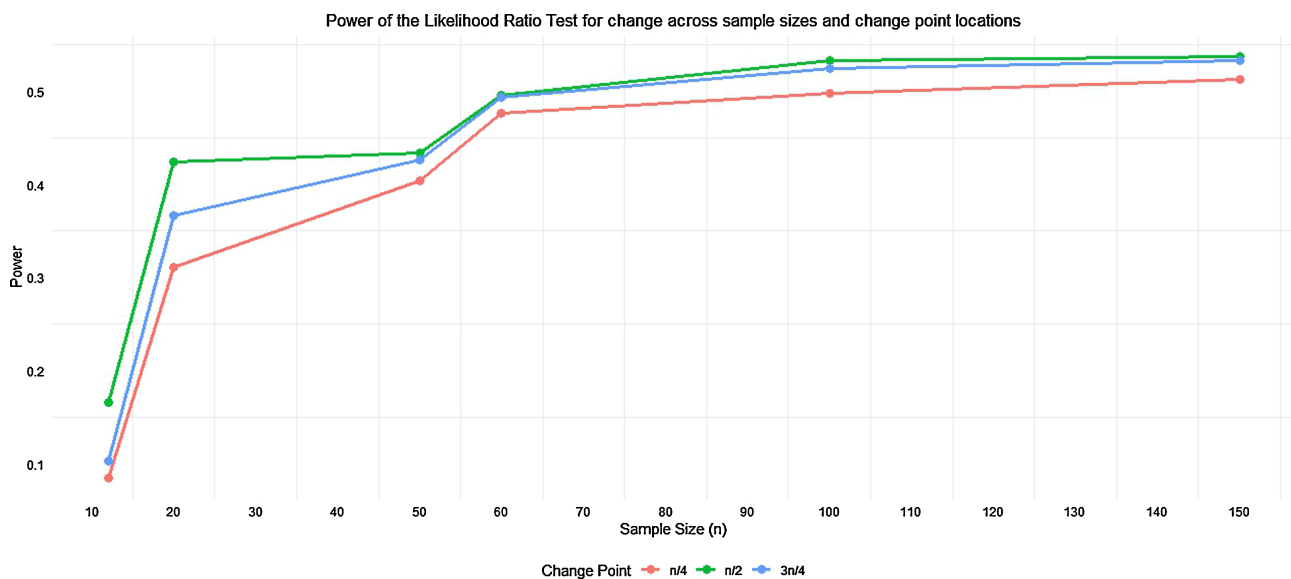


Figure 18. Sensitivity analysis of the power of the LRT for change detection.

4. Conclusions and Recommendations

This study considered change detection for a range of small and medium-sized datasets between $n = 12$ and $n = 500$. Given the foregoing results, the Negative Binomial multiple change point algorithm produces the expected results for different sample sizes and change point locations. Important to note, the algorithm does not erroneously detect changes when absent in actual sense for both large and small samples. This finding makes the NBMCP algorithm a robust and reliable method of detecting changes in a count data series. However, when change is present, the algorithm produces better results of the change point estimates for large and medium samples (n) compared to smaller datasets. The algorithm, when applied to larger data sets, detects multiple and subtle changes more accurately. While the accuracy of the algorithm is slightly lower in detecting small changes within small-sized datasets, there is an upside in reduced computation time compared to large datasets. In addition, the results showed that the NBMCPA produces better estimates of the change points when the actual points of variation are

further away rather than closer to the first data point. It was noted that when the change point was positioned three-quarter-way, the deviation of estimates from the actual changepoints was smaller, often 0, compared to when the change point was set at the $n/4^{\text{th}}$ position.

In a nutshell, comparative analysis of the power results revealed that various factors influence power including the sample size, so that larger samples generally increase power; effect size, with large differences between segments increasing power; significance level (α), such that higher levels of the test (e.g. $\alpha = 0.10$ instead of $\alpha = 0.05$) increases power but also raises the risk of Type I error; variability, so that less variation or noise in the data increases power; and changepoint location, so that power is higher when changes are located further away from the first observation. The algorithm correctly indicates that there is no change in synthetic data with no predefined changepoint indicating a high detection accuracy in this respect. However, further investigations can be done to investigate the False Positive Rate (FPR) of the algorithm where there are multiple subtle changes in close proximity to each other within small and medium datasets.

The NBMCPA is developed such that only a single change can be identified at a time. In the multiple change point setting, the algorithm starts by checking for change points in the entire series. In cases where only one change exists, then the algorithm stops when the first and only change point is located. However, where there are two or more change points, the most pronounced change is detected and located first, then the next most pronounced change is identified and so on.

This simulation study limited investigations to small and medium sized datasets between $n = 12$ and $n = 500$ for which critical values were obtained. Interested researchers are advised to look into the possibility of extending the methods of determining critical values for the likelihood ratio test with larger ($n > 500$) sample sizes. Finally, the algorithm developed in this study works well for over-dispersed and equi-dispersed datasets. Interested researchers may develop change point algorithms that work well for under-dispersed count data.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Lindén, A. and Mäntyniemi, S. (2011) Using the Negative Binomial Distribution to Model Overdispersion in Ecological Count Data. *Ecology*, **92**, 1414-1421. <https://doi.org/10.1890/10-1831.1>
- [2] Lord, D., Park, B.-J. and Model, P.-G. (2012) Negative Binomial Regression Models and Estimation Methods. Probability Density and Likelihood Functions. Texas A&M University, Korea Transport Institute, 1-15.
- [3] Ludwig, J.A. and Reynolds, J.F. (1988) Statistical Ecology: A Primer in Methods and Computing, Volume 1. John Wiley & Sons.
- [4] Korkas, K. and Fryzlewicz, P. (2017) Multiple Change-Point Detection for Non-Stationary Time Series Using Wild Binary Segmentation. *Statistica Sinica*, **27**, 287-311.

<https://doi.org/10.5705/ss.202015.0262>

- [5] Gombay, E. and Horváth, L. (1996) On the Rate of Approximations for Maximum Likelihood Tests in Change-Point Models. *Journal of Multivariate Analysis*, **56**, 120-152. <https://doi.org/10.1006/jmva.1996.0007>
- [6] Gombay, E. and Horvath, L. (1990) Asymptotic Distributions of Maximum Likelihood Tests for Change in the Mean. *Biometrika*, **77**, 411-414. <https://doi.org/10.1093/biomet/77.2.411>
- [7] Mundia, S., Gichuhi, A. and Kihoro, J. (2014) The Power of Likelihood Ratio Test for a Change-Point in Binomial Distribution. *Journal of Agriculture, Science and Technology*, **16**, 105-123.
- [8] Nyambura, S., Mundia, S. and Waititu, A. (2016) Estimation of Change Point in Poisson Random Variables Using the Maximum Likelihood Method. *American Journal of Theoretical and Applied Statistics*, **5**, 219-224. <https://doi.org/10.11648/j.ajtas.20160504.18>