AIMS Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 9(10): 28698–28721.
DOI:10.3934/math.20241392
Received: 16 July 2024
Revised: 17 September 2024
Accepted: 26 September 2024
Published: 10 October 2024

Research article

Change point detection for a skew normal distribution based on the Q-function

Yang Du[1,2] **and Weihu Cheng**[1,*]

[1] School of Mathematics, Statistics and Mechanics, Beijing University of Technology, Beijing 100124, China

[2] Department of Applied Mathematics, Harbin University of Science and Technology, Harbin 150080, China

* **Correspondence:** Email: chengweihu@bjut.edu.cn.

Abstract: In this paper, we enhanced change point detection in skew normal distribution models by integrating the EM algorithm's Q-function with the modified information criterion (MIC). The new QMIC framework improves sensitivity and accuracy in detecting changes, outperforming the modified information criterion (MIC) and the traditional Bayesian information criterion (BIC). Due to the complexity of deriving analytic asymptotic distributions, bootstrap simulations were used to determine critical values at various significance levels. Extensive simulations demonstrate that QMIC offers superior detection capabilities. We applied the QMIC method to two stock market datasets, successfully identifying multiple change points, and highlighting its effectiveness for real-world financial data analysis.

Keywords: skew normal distribution; change point detection; expectation maximization; Q-function; binary segmentation method
Mathematics Subject Classification: 62F03, 62P20

1. Introduction

In statistical analysis, the change point problem pertains to the phenomenon wherein observations demonstrate distinct distributions before and after an unspecified temporal threshold. This issue is crucial across various fields, including financial analysis, economics, and medical research. Tracing back to 1954, Page first introduced the concept of a change point and identified the presence of a singular change point using the well-established cumulative sum procedure [1]. Over subsequent decades, a plethora of methods for detecting change points has been developed, yielding substantial advancements in research, particularly within parametric methodologies. Arellano-Valle discussed the identification of change points in parameters of the location-scale skew normal distribution via Bayesian inference [2]. Moreover, Chernoff and Zacks explored the occurrence of mean shifts in normal distributions from a Bayesian viewpoint [3]. Hsu conducted tests for variance shifts in sequences of independent normal random variables, addressing scenarios where the initial variance level is unknown, employing two test statistics: one locally most powerful and the other based on the cumulative sum of chi-square values [4]. Hinkley inferred the change points of binomial parameter shifts in sequences of zero-one variables through maximum likelihood estimation [5]. Further details on the parametric approach are discussed in [6,7], while studies on non-parametric methods for change point detection are detailed in [8–10].

The likelihood ratio test is a method extensively employed for addressing the problem of detecting change points. Cai [11] introduced a test statistic of the likelihood ratio type to detect potential bathtub-shaped changes in the scale parameter of independent exponential distributions. Wang [12] utilized the likelihood ratio test statistic to identify changes in the parameters of a skew slash distribution. Ferreira [13] enhanced the use of the likelihood ratio test in their development of the mean-shift method for detecting outliers in regression models under skew scale-mixtures of normal distributions. Simulation studies have demonstrated the efficiency of the proposed method in detecting outliers. From a model selection perspective, determining whether a process has undergone change involves choosing between a model with a single set of parameters and one with multiple sets of parameters and change points. In this regard, information criteria serve as useful tools for change point detection. Chen [14] applied the Schwarz information criterion (SIC) to identify and locate all possible variance change points in sequences of independent Gaussian random variables. Ngunkeng [15] proposed a test statistic based on SIC to detect parameter changes in a skew normal distribution. Chen [16] developed a modified information criterion (MIC) by refining the measure of model complexity, which consistently selects the correct model and effectively detects changes at both early and late observation stages. Pan [17] applied the modified information criterion to detect and locate multiple change points in sequences of independent random variables. Said [18] proposed a change point detection procedure based on MIC for skew normal distributions. Kilai [19] conducted change point detection on the exponentiated generalized Gull alpha power exponential distribution using MIC and applied this method to COVID-19 data. It is noteworthy that both the likelihood ratio test and information criterion methods rely on the log-likelihood function. However, for complex models associated with intractable densities, deriving the log-likelihood function can be highly intricate, resulting in substantial computational burdens.

The expectation-maximization (EM) algorithm, introduced by Dempster [20], is a versatile technique for iteratively computing maximum likelihood estimates in models with incomplete data. This approach efficiently handles complex statistical models by treating latent variables as hypothetical missing data and utilizing EM algorithms. Numerous studies have discussed the convergence of parameter estimate sequences obtained through the EM algorithm [21–24]. The EM algorithm finds extensive applications in complex linear mixed models and finite mixed models, notably illustrated in [25–27]. Abe [28] introduced an overparameterized stochastic representation for the skew normal distribution, enhancing the EM algorithm's efficiency. This methodology has been extended to multivariate and mixture models, significantly improving computational performance. Moreover, Zhu developed a novel approach for assessing the local influence of incomplete data using the EM algorithm, employing a procedure similar to Cook's to replace the traditional observed data log-likelihood function with the conditional expectation of the complete data log-likelihood function [29]. Zhu's innovative work has significantly inspired our proposed method for detecting change points in complex models.

The rest of the paper is organized as follows: In Section 2, we introduce the Q-function modified information criterion (QMIC) method for change point detection and discuss the nice properties and motivations of the procedure. In Section 3, we adopt the QMIC method to detect the changes in parameters for the skew normal distribution and propose a procedure to detect change points. Simulations are performed in Section 4 to investigate the performance of the proposed method and to compare it with methods based on MIC and SIC statistics. In Section 5, we apply the QMIC procedure to three significant Latin American stock return markets to detect change points. Finally, conclusions are noted in Section 6.

2. Q-function modified information criterion

In this section, we introduce a novel method for change point detection in complex models, discussing its advantageous properties and underlying motivations.

2.1. A brief description of the EM algorithm

To elucidate our approach, we provide a concise overview of the expectation-maximization (EM) algorithm.

Let $Y_c = (Y_o, Y_m)$ be the complete dataset with a probability density function (PDF) $f(Y_c|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is an r-dimensional parameter vector, and Y_o and Y_m denote the observed data and missing data, respectively. In this setting, the complete data log-likelihood function can be given by

$$L_c(\boldsymbol{\theta}|Y_c) = \log f(Y_c|\boldsymbol{\theta}),$$

whereas the observed data log-likelihood function is given by

$$L_o(\boldsymbol{\theta}|Y_o) = \log f(Y_o|\boldsymbol{\theta}).$$

The complete data log-likelihood function is generally straightforward, in contrast to the observed data log-likelihood function, which is often complex in many statistical applications. This complexity forms the central focus of our study. The EM algorithm provides an alternative method for managing the observed data log-likelihood function, effectively simplifying the computational challenges associated with incomplete datasets.

The standard EM algorithm primarily comprises two steps: the expectation step (E-step) and the maximization step (M-step).

E-step: Given the complete dataset Y_c and parameter estimates $\hat{\boldsymbol{\theta}}^{(h)}$ of the h-th iteration, take the expectation with respect to the conditional distribution $f(Y_m|Y_o, \hat{\boldsymbol{\theta}}^{(h)})$, and then the Q-function is given by

$$Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(h)}) = \mathrm{E}\{L_c(\boldsymbol{\theta}|Y_c)|Y_o, \hat{\boldsymbol{\theta}}^{(h)}\}.$$

M-step: Determine $\hat{\boldsymbol{\theta}}^{(h+1)}$ by

$$\hat{\boldsymbol{\theta}}^{(h+1)} = \underset{\theta}{\mathrm{argmax}} Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(h)}).$$

Wu [24] illustrated that the sequence $\{\hat{\boldsymbol{\theta}}^{(h)}\}$ obtained by the EM algorithm converges to the maximum likelihood estimates $\hat{\boldsymbol{\theta}}$ under some mild conditions.

2.2. Method of QMIC

Let $Y_1, Y_2, \cdots, Y_n$ be a sequence of independent random variables and assume that Y_i with the PDF $f(y; \boldsymbol{\theta_i})$. We assume that all $f(y; \boldsymbol{\theta_i})$ come from the same parametric distribution family $\{f(y; \boldsymbol{\theta}) : \boldsymbol{\theta} \in R^d\}$, for $i = 1, 2, \cdots, n$. Our primary inquiry is whether a change point exists in this sequence and, if so, to identify its location. Consequently, we are interested in testing the null hypothesis

$$H_0 : \boldsymbol{\theta_1} = \boldsymbol{\theta_2} = \cdots = \boldsymbol{\theta_n} = \boldsymbol{\theta} \tag{2.1}$$

versus

$$H_1 : \underbrace{\boldsymbol{\theta_1} = \boldsymbol{\theta_2} = \cdots = \boldsymbol{\theta_k}}_{\theta_{1k}} \neq \underbrace{\boldsymbol{\theta_{k+1}} = \boldsymbol{\theta_{k+2}} = \cdots = \boldsymbol{\theta_n}}_{\theta_{2k}}. \tag{2.2}$$

The log-likelihood function under the null hypothesis can be expressed as

$$L(\boldsymbol{\theta}|Y) = \sum_{i=1}^{n} \log f(y_i; \boldsymbol{\theta}),$$

whereas under the alternative hypothesis, it has the form

$$L(\boldsymbol{\theta_{1k}}, \boldsymbol{\theta_{2k}}|Y) = \sum_{i=1}^{k} \log f(y_i; \boldsymbol{\theta_{1k}}) + \sum_{i=k+1}^{n} \log f(y_i; \boldsymbol{\theta_{2k}}).$$

Chen, Gupta, and Pank [16] considered the effect of change point location on model selection and proposed the modified information criterion. Under the null hypothesis, we define

$$\text{MIC}(n) = -2L(\hat{\boldsymbol{\theta}}|Y) + \dim(\hat{\boldsymbol{\theta}}) \log n.$$

The modified information criterion for the change point problem given by

$$\text{MIC}(k) = -2L(\hat{\boldsymbol{\theta}}_{\boldsymbol{1k}}, \hat{\boldsymbol{\theta}}_{\boldsymbol{2k}}|Y) + \left[2\dim(\hat{\boldsymbol{\theta}}) + \left(\frac{2k}{n} - 1\right)^2\right] \log n, \ for\ 1 \leq k < n,$$

where $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimator of $\boldsymbol{\theta}$, and $\hat{\boldsymbol{\theta}}_{\boldsymbol{1k}}$ and $\hat{\boldsymbol{\theta}}_{\boldsymbol{2k}}$ maximize $L(\boldsymbol{\theta_{1k}}, \boldsymbol{\theta_{2k}}|Y)$ for the given k.

Let

$$S_n = \text{MIC}(n) - \min_{1 \leq k < n} \text{MIC}(k) + \dim(\hat{\boldsymbol{\theta}}) \log n.$$

The addition of the term $\dim(\hat{\boldsymbol{\theta}}) \log n$ eliminates the constant term in the difference between MIC(k) and MIC(n). Chen, Gupta, and Pank [16] used the statistic S_n to test whether the null hypothesis (2.1) holds, and whether the null hypothesis is rejected when S_n takes a sufficiently large number.

Despite the utility of the MIC test, substantial challenges persist with complex models due to the intractability of the log-likelihood function. Consequently, it is imperative to explore alternative approaches to the log-likelihood function to address these difficulties effectively.

Motivated by the EM algorithm, we propose the following Q-function modified information criterion (QMIC) for change point detection. The idea is to apply the same procedure as MIC to the Q-function instead of the log-likelihood function. Under the null hypothesis, we define

$$\text{QMIC}(n) = -2Q(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}}) + \dim(\hat{\boldsymbol{\theta}}) \log n. \tag{2.3}$$

For $1 \leq k < n$, the QMIC in the presence of change points is

$$\text{QMIC}(k) = -2Q(\hat{\boldsymbol{\theta}}_{1k}, \hat{\boldsymbol{\theta}}_{2k}|\hat{\boldsymbol{\theta}}) + \left[2\dim(\hat{\boldsymbol{\theta}}) + \left(\frac{2k}{n} - 1\right)^2\right] \log n, \tag{2.4}$$

where $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimate of $\boldsymbol{\theta}$ obtained by the EM algorithm under the null hypothesis, and $\hat{\boldsymbol{\theta}}_{1k}, \hat{\boldsymbol{\theta}}_{2k}$ are the estimates of $\boldsymbol{\theta}_{1k}, \boldsymbol{\theta}_{2k}$ which satisfy

$$\hat{\boldsymbol{\theta}}_{1k}, \hat{\boldsymbol{\theta}}_{2k} = \underset{\theta_{1k},\theta_{2k}}{\text{argmax}} Q(\hat{\boldsymbol{\theta}}_{1k}, \hat{\boldsymbol{\theta}}_{2k}|\hat{\boldsymbol{\theta}}) = \underset{\theta_{1k},\theta_{2k}}{\text{argmax}} \text{E}\{L_c(\boldsymbol{\theta}_{1k}, \boldsymbol{\theta}_{2k}|Y_c)|Y_o, \hat{\boldsymbol{\theta}}\}.$$

The difference between QMIC(n) and $\min_{1\leq k<n}$ QMIC(k) plays a key role in the model selection. Based on this, we define a test statistic W_n to test whether the null hypothesis of no change holds.

Let

$$W_n = \text{QMIC}(n) - \min_{1\leq k<n} \text{QMIC}(k) + \dim(\boldsymbol{\theta}) \log n. \tag{2.5}$$

The statistic W_n can be considered as a measure of the difference between $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}_{1k}, \hat{\boldsymbol{\theta}}_{2k}$. The null hypothesis is rejected with a sufficiently large value of W_n, and then we suggest that there are change points in this sample and the location of the change point can be estimated by

$$\hat{k} = \underset{1\leq k<n}{\text{argmin}} \text{QMIC}(k). \tag{2.6}$$

Owing to the promising outcomes derived from the EM algorithm, the solutions $\hat{\boldsymbol{\theta}}_{1k}$ and $\hat{\boldsymbol{\theta}}_{2k}$ achieved in maximizing the Q-function lead to the formulation of the following lemma:

Lemma 2.1. *Suppose that $L_o(\boldsymbol{\theta}|Y_o)$ is unimodal in Ω with only a stationary point, where Ω is a subset in the d-dimensional space, and the first-order partial derivatives of $Q(\boldsymbol{\psi}|\boldsymbol{\theta})$ are continuous in $\boldsymbol{\psi}$ and $\boldsymbol{\theta}$. Then we obtain, under Wald conditions W1–W7 [16] and the null hypothesis,*

$$(\hat{\boldsymbol{\theta}}_{1k}, \hat{\boldsymbol{\theta}}_{2k}) \to \boldsymbol{\theta}$$

consistently in probability for all k, as $n \to \infty$.

Under Wald conditions W1–W7 and the alternative hypothesis with a change point at k satisfying the limit of k/n at (0, 1) as $n \to \infty$, then

$$(\hat{\boldsymbol{\theta}}_{1k}, \hat{\boldsymbol{\theta}}_{2k}) \to (\boldsymbol{\theta}_{1k}, \boldsymbol{\theta}_{2k})$$

consistently in probability for all k, as $n \to \infty$.

The above lemma is derived from the global convergence theorem [24] and the consistency lemma [16]; therefore, details of the proof are omitted here. Another motivation for our approach is its set of desirable properties, which are analogous to those of the MIC criterion.

Note that a sufficient number of sample observations are required to obtain the maximum likelihood estimates of the parameters. Perron and Vogelsang [30] considered a trimmed procedure for the method; the key idea is to trim k_0 samples at the beginning and k_1 samples at the end, which means detecting the change point in the location interval $[k_0 + 1, n - k_1]$. They pointed out that k_0 and k_1 can be arbitrarily selected according to individual needs.

Attention should be paid to multiple changes in the data. Vostrikova [31] proposed a binary segmentation method that assumes at most one change within each segmentation and repeats several consecutive steps to detect multiple changes. Based on the binary segmentation method, it is always possible to split the problem of multiple changes into several single change problems to deal with. Consequently, this paper focuses on developing a testing procedure for detecting a single change point.

3. QMIC for change point detection in skew normal distribution

In this section, we use the proposed QMIC approach to detect the change points in the skew normal distribution.

Let $Y = (Y_1, Y_2, \ldots, Y_n)$ be a random sample of size n from skew normal distribution $SN(\mu, \sigma^2, \lambda)$. We use the QMIC procedure to examine the possibility of changes in location parameter μ, scale parameter σ^2, and shape parameter λ. Here, the null hypothesis is interested in

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_n = \mu, \;\; {\sigma_1}^2 = {\sigma_2}^2 = \cdots = {\sigma_n}^2 = \sigma^2, \;\; \lambda_1 = \lambda_2 = \cdots = \lambda_n = \lambda \tag{3.1}$$

versus the alternative hypothesis

$$\begin{aligned} H_1 : &\underbrace{\mu_1 = \mu_2 = \cdots = \mu_k}_{\mu_{1k}} \neq \underbrace{\mu_{k+1} = \mu_{k+2} = \cdots = \mu_n}_{\mu_{2k}}, \\ &\underbrace{\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2}_{\sigma_{1k}^2} \neq \underbrace{\sigma_{k+1}^2 = \sigma_{k+2}^2 = \cdots = \sigma_n^2}_{\sigma_{1k}^2}, \\ &\underbrace{\lambda_1 = \lambda_2 = \cdots = \lambda_k}_{\lambda_{1k}} \neq \underbrace{\lambda_{k+1} = \lambda_{k+2} = \cdots = \lambda_n}_{\lambda_{2k}}. \end{aligned} \tag{3.2}$$

Before proceeding to the change point detection, we discuss the following lemma, which is indispensable for obtaining the Q-function.

Theorem 3.1. *A random variable $Y \sim SN(\mu, \sigma^2, \lambda)$ has a stochastic representation given by*

$$Y = \mu + \sigma\delta T + \sigma(1 - \delta^2)^{1/2} T_1,$$

where $\delta = \lambda / \sqrt{(1 + \lambda^2)}, T = |T_0|$, $|\cdot|$ denotes the absolute value, and T_0 and T_1 are independent standard normal random variables.

To develop an effective EM algorithm for skew normal distribution, identical to [32], we consider the following hierarchical representation for the SN model based on Lemma 3.1:

$$Y_i | T_i = t_i \sim N(\mu + \Delta \cdot t_i, \Gamma),$$

$$T_i \sim HN(0,1),$$

where $\Delta = \sigma\delta$, $\Gamma = (1-\delta^2)\sigma^2$, and $HN(0,1)$ is the half-normal distribution. We consider an important result that T_i given Y_i follows a truncated normal distribution on $(0,\infty)$ with parameters μ_{T_i} and M_T^2 before truncation, denoted by $T_i|Y_i \sim N(\mu_{T_i}, M_T^2)\mathbb{I}_{(0,\infty)}$, where

$$M_T^2 = \frac{\Gamma}{\Gamma+\Delta^2}, \qquad \mu_{T_i} = \frac{\Delta}{\Gamma+\Delta^2}(y_i-\mu). \tag{3.3}$$

Let $Y = (y_1, \cdots, y_n)^T$ be the observed data, $T = (t_1, \cdots, t_n)^T$ on behalf of the corresponding missing variables, and the complete data is the combination of Y and T, denoted by $W = (Y^T, T^T)^T$. To facilitate the implementation, let $\boldsymbol{\theta} = (\mu, \triangle, \Gamma)$ be the vector of parameters in focus.

Invoking the hierarchical representation of y_i, under the null hypothesis, the complete data log-likelihood function can be expressed as follows:

$$L_{H_0}(\boldsymbol{\theta}|W) = -n\log\pi - \frac{n}{2}\log\Gamma - \frac{1}{2\Gamma}\sum_{i=1}^{n}(y_i-\mu-\Delta t_i)^2 - \frac{1}{2}\sum_{i=1}^{n}t_i^2. \tag{3.4}$$

It is well known that the EM algorithm consists of two steps; namely, the expectation step (E-step) and the maximization step (M-step).

E-step: Given the observed dataset $Y = (y_1, \cdots, y_n)$ and current parameter estimates $\hat{\boldsymbol{\theta}}^{(h)}$ of the h-th iteration, take the expectation of the log-likelihood function for the condition distribution $f(T|Y, \hat{\boldsymbol{\theta}}^{(h)})$, and then the Q-function can be obtained:

$$\begin{aligned} Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(h)}) &= \mathrm{E}[L_{H_0}(\boldsymbol{\theta}|W)|Y, \hat{\boldsymbol{\theta}}^{(h)}] \\ &= -n\log\pi - \frac{n}{2}\log\Gamma - \frac{1}{2\Gamma}\sum_{i=1}^{n}[(y_i-\mu)^2 - 2\Delta(y_i-\mu)\hat{s}_{1i}^{(h)} + \Delta^2\hat{s}_{2i}^{(h)}] - \frac{1}{2}\sum_{i=1}^{n}\hat{s}_{2i}^{(h)}, \end{aligned}$$

where

$$\hat{s}_{1i}^{(h)} = \mathrm{E}[T_i|y_i, \hat{\boldsymbol{\theta}}^{(h)}] = \hat{\mu}_{T_i}^{(h)} + \hat{M}_T^{(h)}\frac{\phi(\mu_{T_i}^{(h)}/M_T^{(h)})}{\Phi(\mu_{T_i}^{(h)}/M_T^{(h)})}, \tag{3.5}$$

$$\hat{s}_{2i}^{(h)} = \mathrm{E}[T_i^2|y_i, \hat{\boldsymbol{\theta}}^{(h)}] = \hat{\mu}_{T_i}^{2(h)} + \hat{M}_T^{2(h)} + \hat{M}_T^{(h)}\mu_{T_i}^{(h)}\frac{\phi(\mu_{T_i}^{(h)}/M_T^{(h)})}{\Phi(\mu_{T_i}^{(h)}/M_T^{(h)})}, \quad i = 1, \cdots, n, \tag{3.6}$$

where $\mu_{T_i}^{(h)}$, $M_T^{2(h)}$ are defined as (3.3), $\phi(\cdot)$ denotes the PDF of standard normal distribution, and $\Phi(\cdot)$ is the cumulative density function (CDF) of standard normal distribution.

In certain cases, the complexity of the Q-function derived from the E-step can make the M-step analytically challenging. For skew normal distribution, this challenge is efficiently addressed by employing a series of conditional maximization (CM) steps.

CM-steps: Update $\hat{\boldsymbol{\theta}}^{(h)}$ by maximizing $Q(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}}^{(h)})$ over $\hat{\boldsymbol{\theta}}$, which leads to the following results:

$$\hat{\mu}^{(h+1)} = \sum_{i=1}^{n}(y_i - \Delta^{(h)}\hat{s}_{1i}^{(h)})/n\ ,$$

$$\hat{\Delta}^{(h+1)} = \sum_{i=1}^{n}(y_i - \hat{\mu}^{(h+1)})\hat{s}_{1i}^{(h)} / \sum_{i=0}^{n} \hat{s}_{2i}^{(h)} ,$$

$$\hat{\Gamma}^{(h+1)} = \frac{1}{n}\sum_{i=1}^{n}[(y_i - \hat{\mu}^{(h+1)})^2 - 2(y_i - \hat{\mu}^{(h+1)})\hat{\Delta}^{(h+1)}\hat{s}_{1i}^{(h)} + (\hat{\Delta}^{(h+1)})^2\hat{s}_{2i}^{(h)}].$$

For a prescribed value $\epsilon > 0$, if the value of the log-likelihood function of two consecutive iterations satisfies $|L_{H_0}(Y|\hat{\boldsymbol{\theta}}^{(h+1)})/L_{H_0}(Y|\hat{\boldsymbol{\theta}}^{(h)}) - 1| < \varepsilon$, the iterative process is broken up. In our study, we set $\varepsilon = 10^{-5}$, as recommended by Zeller et al. [33]. More details on the maximum likelihood estimates of the SN distribution can be found in [34].

The maximum likelihood estimates of the parameters can be obtained using the EM algorithm, denoted as $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\Delta}, \hat{\Gamma})$. Therefore, the expression of the $Q(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}})$ under the null hypothesis is

$$Q(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}}) = -n\log\pi - \frac{n}{2}\log\hat{\Gamma} - \frac{1}{2\hat{\Gamma}}\sum_{i=1}^{n}[(y_i - \hat{\mu})^2 - 2\hat{\Delta}(y_i - \hat{\mu})\hat{s}_{1i} + \hat{\Delta}^2\hat{s}_{2i}] - \frac{1}{2}\sum_{i=1}^{n}\hat{s}_{2i}, \tag{3.7}$$

where $\hat{s}_{1i}$ and $\hat{s}_{2i}$ can be obtained from (3.5) and (3.6), respectively, by replacing $\hat{\boldsymbol{\theta}}^{(h)}$ in them with $\hat{\boldsymbol{\theta}}$.

Under the alternative hypothesis, the complete data log-likelihood function is

$$\begin{aligned} L_{H_1}(\boldsymbol{\theta}|W) = &- n\log\pi - \frac{k}{2}\log\Gamma_{1k} - \frac{1}{2\Gamma_{1k}}\sum_{i=1}^{k}(y_i - \mu_{1k} - \Delta_{1k}t_i)^2 - \frac{1}{2}\sum_{i=1}^{k}t_i^2 \\ &- \frac{n-k}{2}\log\Gamma_{2k} - \frac{1}{2\Gamma_{2k}}\sum_{i=k+1}^{n}(y_i - \mu_{2k} - \Delta_{2k}t_i)^2 - \frac{1}{2}\sum_{i=k+1}^{n}t_i^2, \end{aligned} \tag{3.8}$$

where $\Delta_{ik} = \sigma_{ik}\lambda_{ik}/\sqrt{(1+\lambda_{ik}{}^2)}$, $\Gamma_{ik} = \sigma_{ik}^2/(1+\lambda_{ik}{}^2)$, $i = 1, 2$.

We consider the alternative hypothesis of the existence of a single change point as a combination of two skew normal distributed variables, one with sample size k and the other with $n-k$, such that maximizing the overall Q-function is equivalent to making these two components maximum. Therefore, with the current estimate of $\hat{\boldsymbol{\theta}}$, the EM algorithm is applied to each of these two components to obtain the maximum likelihood estimates of the parameters $\hat{\boldsymbol{\theta}}_{\mathbf{1k}} = (\hat{\mu}_{1k}, \hat{\Delta}_{1k}^2, \hat{\Gamma}_{1k})$ and $\hat{\boldsymbol{\theta}}_{\mathbf{2k}} = (\hat{\mu}_{2k}, \hat{\Delta}_{2k}^2, \hat{\Gamma}_{2k})$, respectively. Then we can obtain

$$\begin{aligned} Q(\hat{\boldsymbol{\theta}}_{\mathbf{1k}}, \hat{\boldsymbol{\theta}}_{\mathbf{2k}}|\hat{\boldsymbol{\theta}}) = &- n\log\pi - \frac{k}{2}\log\hat{\Gamma}_{1k} - \frac{1}{2\hat{\Gamma}_{1k}}\sum_{i=1}^{k}[(y_i - \hat{\mu}_{1k})^2 - 2\hat{\Delta}_{1k}(y_i - \hat{\mu}_{1k})\hat{s}_{1i} + \hat{\Delta}_{1k}^2\hat{s}_{2i}] - \frac{1}{2}\sum_{i=1}^{k}\hat{s}_{2i} \\ &- \frac{n-k}{2}\log\hat{\Gamma}_{2k} - \frac{1}{2\hat{\Gamma}_{2k}}\sum_{i=k+1}^{n}[(y_i - \hat{\mu}_{2k})^2 - 2\hat{\Delta}_{2k}(y_i - \hat{\mu}_{2k})\hat{s}_{2i} + \hat{\Delta}_{2k}^2\hat{s}_{1i}] - \frac{1}{2}\sum_{i=k+1}^{n}\hat{s}_{2i}. \end{aligned} \tag{3.9}$$

By substituting (3.7) and (3.9) into (2.5), we can obtain the W_n statistic under the null hypothesis (3.1) of the skew normal distribution

$$
\begin{aligned}
W_n =& n\log\hat{\Gamma} + \frac{1}{\hat{\Gamma}}\sum_{i=1}^{n}[(y_i-\hat{\mu})^2 - 2\hat{\Delta}(y_i-\hat{\mu})\hat{s}_{1i} + \hat{\Delta}^2\hat{s}_{2i}] - \min_{1\le k<n}\Big\{k\log\hat{\Gamma}_{1k} \\
&+ \frac{1}{\hat{\Gamma}_{1k}}\sum_{i=1}^{k}[(y_i-\hat{\mu}_{1k})^2 - 2\hat{\Delta}_{1k}(y_i-\hat{\mu}_{1k})\hat{s}_{1i} + \hat{\Delta}_{1k}^2\hat{s}_{2i}] + (n-k)\log\hat{\Gamma}_{2k} \\
&+ \frac{1}{\hat{\Gamma}_{2k}}\sum_{i=k+1}^{n}[(y_i-\hat{\mu}_{2k})^2 - 2\hat{\Delta}_{2k}(y_i-\hat{\mu}_{2k})\hat{s}_{2i} + \hat{\Delta}_{2k}^2\hat{s}_{2i}] + (\frac{2k}{n}-1)^2\log n\Big\}.
\end{aligned}
\tag{3.10}
$$

The null hypothesis (3.1) is rejected for a sufficiently large number of W_n. Here, we use the bootstrap method to simulate the distribution of the statistic W_n, and a suggested framework to obtain the critical value c_α of the W_n statistic for a given significance level α and the P-value is shown in Algorithm 1.

Algorithm 1. Bootstrap procedure for the W_n statistic and critical value c_α.

Input: Data: $Y_1, Y_2, \ldots, Y_n$, Significance level: α, Trimming parameters: k_0, k_1, Number of bootstrap samples: B.

Output: Critical value c_α, P-value.

(1) Estimate $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\Delta}, \hat{\Gamma})$ under H_0 using the EM algorithm.

(2) Compute the following parameter estimates:

$$
\hat{\sigma}^2 = \hat{\Delta}^2 + \hat{\Gamma}, \quad \hat{\lambda} = \frac{\hat{\Delta}}{\sqrt{\hat{\Gamma}}}.
$$

(3) For each $k \in [k_0+1, n-k_0]$, re-estimate $\hat{\boldsymbol{\theta}}_{\mathbf{1k}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{2k}}$ using the EM algorithm based on the current $\hat{\boldsymbol{\theta}}$.

(4) Calculate the test statistic W_n^* using Eq (3.10).

(5) Perform B bootstrap iterations:

  (a) Generate a bootstrap sample $Y_1^*, Y_2^*, \ldots, Y_n^*$ from the fitted distribution $SN(\hat{\mu}, \hat{\sigma}^2, \hat{\lambda})$.

  (b) Recalculate $W_n^{(b)(i)}$ for each bootstrap sample by repeating Steps 1–4, $i = 1, 2, \ldots, B$.

(6) Calculate the critical value c_α as the $100(1-\alpha)\%$ quantile of the $W_n^{(b)}$ values.

(7) Compute the P-value as:

$$
\text{P-value} = \frac{1}{B}\sum_{i=1}^{B} I(W_n^{(b)(i)} \ge W_n^*).
$$

(8) Decision rule:

  (a) Reject H_0 if $W_n^* > c_\alpha$ or P-value $< \alpha$.

  (b) If rejected, estimate the location of the change point k using Eq (2.6).

The QMIC method combined with the binary segmentation method can be utilized to address the detection of multiple change points of skew normal distribution.

4. Simulation results

In this section, we investigate the critical values of W_n and use simulations to compare the performance of several test methods in terms of powers. In addition to our proposed QMIC information criterion, we also consider the modified information criterion [18] and the Bayesian information criterion [15].

4.1. Critical values

We performed simulations of critical values under different combinations of parameters and sample sizes using Algorithm 1 in Section 3. The critical values were estimated with $B = 5000$ bootstrap samples of $n = 50, 100, 200, 300, 400$ at nominal test sizes $\alpha = 0.1, 0.05, 0.01$, respectively. The results, summarized in Table 1, indicate that the empirical critical values exhibit noticeable variation with the sample size at each significance level.

Table 1. Critical values under different significance levels α for the W_n statistic obtained by bootstrap simulation with $B = 5000$ samples of size n.

		α					α		
Parameters of SN	n	0.1	0.05	0.01	Parameters of SN	n	0.1	0.05	0.01
(2, 1, 2)	50	0.7445	1.7259	4.6582	(3, 3, 2)	50	0.6299	1.6657	5.0009
	100	0.5875	1.4271	4.1944		100	0.6831	1.6151	4.3323
	200	0.3580	0.8400	2.2657		200	0.8815	1.6615	3.5394
	300	0.7588	1.3115	3.0247		300	0.7653	1.3798	2.7957
	400	1.6036	2.3545	4.2669		400	1.0052	1.9998	2.7865
(1.5, 1.5, 2)	50	-0.1034	0.6576	3.9767	(2, 2, 1)	50	1.4553	2.6029	5.0203
	100	2.1997	3.4963	4.8935		100	3.1374	4.5598	8.0089
	200	3.2705	4.4386	7.5911		200	1.8309	2.9691	5.9207
	300	1.7664	2.6250	3.4809		300	3.5893	4.9501	7.6586
	400	2.2664	3.2613	5.8549		400	4.1812	5.6056	5.0678
(4, 4, 0)	50	0.5744	1.4334	4.0106	(3, 4, −2)	50	-0.3062	0.0981	1.5738
	100	3.3062	4.7456	8.3763		100	-0.1741	0.0698	1.4783
	200	5.1096	6.4657	10.3637		200	2.4018	3.5573	6.5670
	300	5.4263	7.0787	12.5718		300	1.4431	2.3089	4.0675
	400	5.4516	6.8953	11.0537		400	2.5494	3.4353	6.3534
(1.5, 1.5, 2)	50	-0.1034	0.6577	3.9768	(3, 1, 1)	50	1.0202	2.2220	5.6560
	100	2.1997	3.4963	7.2404		100	2.1467	3.5923	7.4245
	200	3.2704	4.4386	4.4915		200	1.8109	2.8508	5.8633
	300	1.7664	2.6250	3.4809		300	4.5889	5.9613	9.6741
	400	2.2664	3.2613	5.8549		400	5.2007	6.6797	11.0312

4.2. Power comparison

In this section, we conducted simulations under different scenarios to investigate the performance of the test procedures in terms of power. Additionally, we constructed the test statistic T_n based on the classical Bayesian information criterion (BIC), defined as follows, to facilitate comparison.

$$T_n = \text{BIC}(n) - \min_{1\le k<n} \text{BIC}(k) + 3\log n,$$

where BIC(n) under H_0 and BIC(k) under H_1 are defined as follows:

$$\text{BIC}(n) = -2\log L_{H_0}(\hat{\mu}, \hat{\sigma}, \hat{\lambda}) + 3\log(n),$$
$$\text{BIC}(k) = -2\log L_{H_1}(\hat{\mu}_1, \hat{\sigma}_1, \hat{\lambda}_1, \hat{\mu}_n, \hat{\sigma}_n, \hat{\lambda}_n) + 6\log n.$$

The pre-change distribution was set to follow $SN(\mu_1, \sigma_1, \lambda_1)$ with parameters $(\mu_1, \sigma_1, \lambda_1) = (2, 2, 1)$. Post-change, the distribution was set to $SN(\mu_n, \sigma_n, \lambda_n)$ with parameters $\theta_n = (\mu_n, \sigma_n, \lambda_n)$ taking values $(3, 3, 0)$, $(2.5, 2.5, 2)$, $(3, 3, 2)$, and $(1.5, 1.5, 1.5)$. The sample sizes considered were $n = 50, 100$, and 150. Changes were introduced approximately at the beginning ($\frac{1}{4}$), at the center ($\frac{1}{2}$), and at the end ($\frac{3}{4}$) of the sample size n. The results are listed in Tables 2 and 3.

Table 2. Power comparison between QMIC, MIC, and BIC for $\alpha = 0.1$.

n	k	$(\mu_1, \sigma_1, \lambda_1)$	Model	$(\mu_n, \sigma_n, \lambda_n)$					
				(3,3,0)	(2.5,2.5,2)	(3,3,2)	(1.5,1.5,1.5)	(0,2,1)	(1,2,1.5)
50	10	(2,2,1)	QMIC	0.218	0.190	0.562	0.278	0.683	0.385
			MIC	0.206	0.176	0.503	0.208	0.660	0.382
			BIC	0.182	0.097	0.247	0.276	0.695	0.364
	25	(2,2,1)	QMIC	0.437	0.443	0.814	0.463	0.805	0.526
			MIC	0.428	0.353	0.692	0.382	0.761	0.487
			BIC	0.277	0.296	0.660	0.352	0.752	0.496
	40	(2,2,1)	QMIC	0.328	0.237	0.416	0.138	0.776	0.355
			MIC	0.322	0.215	0.300	0.090	0.724	0.327
			BIC	0.245	0.152	0.412	0.063	0.718	0.319
100	25	(2,2,1)	QMIC	0.685	0.523	0.884	0.428	0.912	0.618
			MIC	0.697	0.501	0.880	0.411	0.897	0.602
			BIC	0.612	0.488	0.873	0.302	0.865	0.585
	50	(2,2,1)	QMIC	0.851	0.532	0.970	0.633	0.981	0.833
			MIC	0.871	0.530	0.992	0.656	0.975	0.756
			BIC	0.701	0.424	0.980	0.562	0.956	0.760
	75	(2,2,1)	QMIC	0.688	0.403	0.911	0.479	0.924	0.678
			MIC	0.665	0.379	0.903	0.513	0.917	0.650
			BIC	0.505	0.368	0.907	0.408	0.899	0.631
150	35	(2,2,1)	QMIC	0.812	0.690	0.989	0.663	0.986	0.841
			MIC	0.800	0.722	0.989	0.652	0.972	0.763
			BIC	0.798	0.604	0.977	0.607	0.983	0.788
	75	(2,2,1)	QMIC	0.975	0.880	0.970	0.883	0.984	0.867
			MIC	0.983	0.905	0.998	0.889	0.989	0.863
			BIC	0.919	0.876	0.999	0.849	0.956	0.842
	110	(2,2,1)	QMIC	0.940	0.756	0.998	0.702	0.998	0.868
			MIC	0.915	0.747	0.997	0.685	0.997	0.855
			BIC	0.923	0.720	0.992	0.658	0.995	0.847

Table 3. Power comparison between QMIC, MIC, and BIC for $\alpha = 0.05$.

n	k	$(\mu_1, \sigma_1, \lambda_1)$	Model	$(\mu_n, \sigma_n, \lambda_n)$					
				(3,3,0)	(2.5,2.5,2)	(3,3,2)	(1.5,1.5,1.5)	(0,2,1)	(1,2,1.5)
50	10	(2,2,1)	QMIC	0.156	0.121	0.402	0.185	0.580	0.292
			MIC	0.133	0.092	0.315	0.185	0.560	0.285
			BIC	0.096	0.045	0.191	0.178	0.552	0.280
	25	(2,2,1)	QMIC	0.328	0.237	0.634	0.235	0.668	0.323
			MIC	0.315	0.220	0.610	0.202	0.631	0.304
			BIC	0.165	0.163	0.498	0.197	0.609	0.315
	40	(2,2,1)	QMIC	0.227	0.115	0.325	0.043	0.697	0.237
			MIC	0.190	0.081	0.124	0.035	0.705	0.287
			BIC	0.180	0.084	0.309	0.020	0.708	0.260
100	25	(2,2,1)	QMIC	0.526	0.402	0.812	0.308	0.835	0.534
			MIC	0.531	0.373	0.760	0.287	0.824	0.526
			BIC	0.411	0.356	0.805	0.189	0.806	0.524
	50	(2,2,1)	QMIC	0.778	0.523	0.932	0.515	0.965	0.681
			MIC	0.760	0.396	0.913	0.501	0.933	0.692
			BIC	0.594	0.518	0.933	0.417	0.859	0.650
	75	(2,2,1)	QMIC	0.518	0.332	0.864	0.384	0.814	0.565
			MIC	0.523	0.290	0.831	0.382	0.808	0.453
			BIC	0.389	0.258	0.862	0.257	0.827	0.502
150	35	(2,2,1)	QMIC	0.738	0.639	0.950	0.587	0.965	0.764
			MIC	0.695	0.633	0.980	0.561	0.983	0.758
			BIC	0.702	0.504	0.969	0.575	0.942	0.699
	75	(2,2,1)	QMIC	0.927	0.852	0.997	0.780	0.976	0.821
			MIC	0.938	0.840	0.999	0.787	0.958	0.805
			BIC	0.886	0.755	0.995	0.769	0.925	0.800
	110	(2,2,1)	QMIC	0.907	0.595	0.998	0.563	0.917	0.811
			MIC	0.867	0.589	0.990	0.558	0.889	0.823
			BIC	0.870	0.545	0.979	0.533	0.903	0.762

From the simulation results, it is evident that the power of the proposed test increases with both the sample size and the magnitude of parameter changes. Notably, the power tends to be higher when the change occurs toward the middle of the data sequence, compared to changes occurring near the beginning or end.

Furthermore, our comparative analysis indicates that the proposed QMIC procedure is highly competitive relative to the MIC and BIC methods. In general, the QMIC procedure outperforms the BIC and MIC methods across a range of sample sizes and change locations. The power of all three tests increases with sample size, further underscoring the robustness of the QMIC procedure in detecting changes under various sample sizes and parameter shifts. Additionally, the QMIC procedure maintains strong performance even in scenarios involving changes in one or two parameters, further highlighting its effectiveness in a broader range of conditions.

5. Application

In this section, we draw a conclusion that the method proposed in this paper presents analytical results that are similar to those obtained from the modified information criterion [18] based on the classical log-likelihood function. To illustrate this result, we focus our attention on three significant Latin American stock return markets: the Chilean, Brazilian, and Mexican markets. The stock returns for each of the three countries were recorded weekly from October 31, 1995, to October 31, 2000, resulting in two time series, each with 261 data points. Arellano-Valle [2] used different models to fit the three datasets, and almost all the criterion suggest that the skew normal distribution is a good fitting model for the stock returns of these three markets, compared with the normal distribution. The data sources and further details regarding these datasets can be found in Ngunkeng and Wei [15] and Said, Wei, and Tian [18].

Let S_t be the stock return value at week t. In order to ensure the independence in the times series data and study the change in stock returns, in general, we study the stock return rates instead of the stock returns directly, which is defined as

$$R_t = \frac{S_{t+1} - S_t}{S_t}, \quad t = 1, 2, \cdots, n-1.$$

Before proceeding with the actual analysis, we need to test the independence of the transformed data R_t. Hsu [35] proposed several methods that can be used for testing the independence of such a transformed dataset. Here, we choose the Portmanteau test defined as

$$Q_k = n \sum_{i=1}^{k} r_i^2, \tag{5.1}$$

where r_i is the auto-correlation coefficient (ACF) at lag i, and k is the lag up to which the auto-correlation coefficient function is considered. Under the condition that the null hypothesis of independence holds, Q_k has an asymptotic χ^2 distribution with the degree of freedom k, denoted as $Q_k \sim \chi^2(k)$.

5.1. Chilean stock market

We consider the stock return rate series R_t of the Chilean market. The ACF plot and the normal Q-Q plot of R_t are depicted in Figure 1. The left graph in Figure 1 presents the auto-correlation coefficient of R_t, from which it is clear that R_t is not severely auto-correlated. Furthermore, the test statistic Q_k of R_t for the Chilean market is obtained as follows:

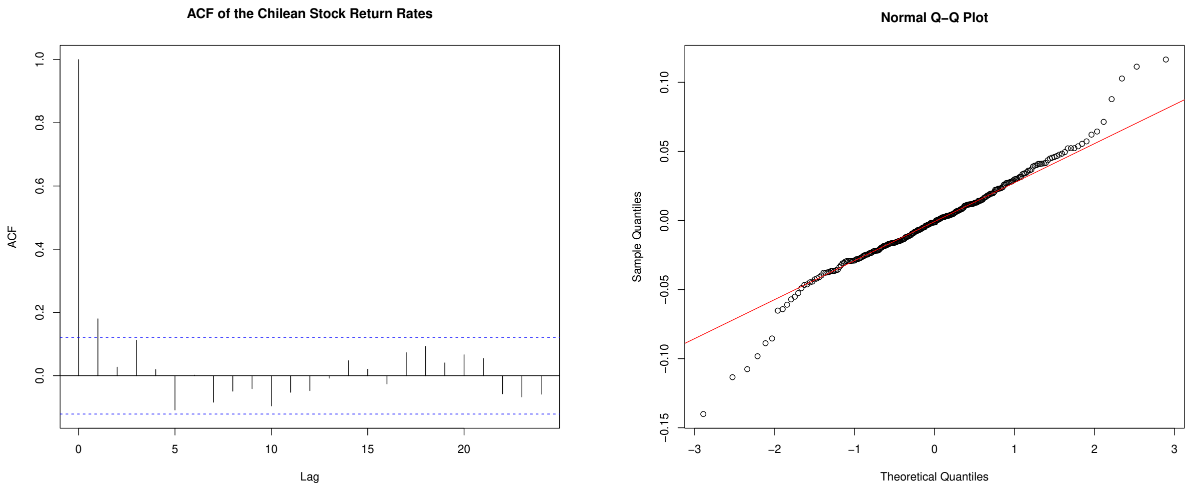
$$Q_{24} = 261 \times \sum_{i=1}^{24} r_i^2 = 261 \times 0.1210905 = 31.25 < \chi^2_{0.95}(24) = 36.415,$$

which indicates that the null hypothesis of independence of the transformed data failed to be rejected.

The right graph in Figure 1 illustrates that the R_t data deviates from the normal distribution, which indicates that the normality assumption fails. Therefore, as suggested by Arellano-Valle [2], we treat the skew normal distribution as a plausible candidate for fitting the R_t series for Chile. The estimated parameters for the fitted skew normal distribution are as follows:

$$\hat{\mu} = 0.0207\,(\text{SE} = 0.0076), \quad \hat{\sigma} = 0.0398\,(\text{SE} = 0.0045), \quad \hat{\lambda} = -0.9286\,(\text{SE} = 0.3841).$$

Here, "SE" denotes the standard error of the estimated parameters.



(a) ACF values of R_t. **(b)** The normal Q-Q plot of R_t.

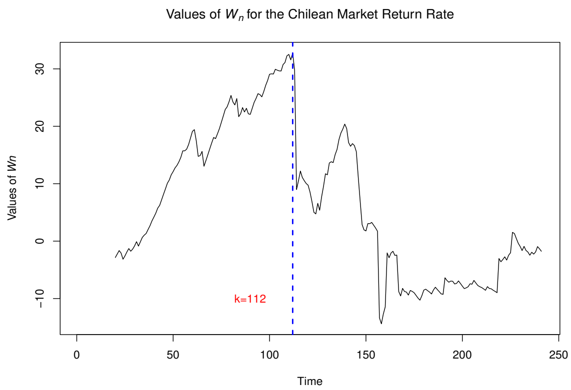
Figure 1. The behavior of the stock return rate R_t series of the Chilean market.

We apply the proposed QMIC procedure with the bootstrap method to test the hypothesis in Section 2, and implement the binary segmentation method to detect all possible change points in the R_t data. To enhance the robustness of our analysis, consistent with the approach described by Perron and Vogelsang [30], we trimmed 20 data points from both the beginning and the end of each series. Figure 2 shows the corresponding values of the W_n test statistics for each series of R_t, as well as the possible locations of the change points detected by the proposed method.

- Consider the sequence R_t for the Chilean stock market, $t = 1, 2, \cdots, 261$. Under the null hypothesis, QMIC(261) $= -707.3299$, $\min_{1\le k\le 261}$QMIC(k) = QMIC(112) $= -723.3759$, and the test statistic $W_n = 32.7396$. With the discussion in Section 2, we use the bootstrap method to simulate the distribution of the test statistic W_n and obtain the approximate P-value $= 0.000 < 0.05$ with $B = 2000$. Therefore, we reject the null hypothesis and conclude that there is a change in the R_t series and the first change is located in the 112th position. Moreover, it corresponds to the change point located at the 113th position (December 26, 1997) of the stock return data S_t with the associated stock return value of 736.133. The reason for the change point may be due to the Asian financial crisis of 1997, which reached its climax by a mini crash on October 27, 1997.
- With the binary segmentation method, we consider the series R_t for t from 1 to 112. Under the null hypothesis, we obtain QMIC(112) $= -500.0035$, $\min_{1\le k\le 112}$QMIC(k) = QMIC(24) $= -467.5512$, and the test statistic $W_n = -18.2969$. This indicates that the QMIC value under the original hypothesis is smaller than the QMIC value under the alternative hypothesis, and thus we accept the null hypothesis that there is no change point in this sequence.
- Then we consider the sequence R_t for t from 113 to 261, and we compute the QMIC value when the null hypothesis holds. QMIC(149) $= -362.0137$, $\min_{1\le k\le 149}$ QMIC(k) = QMIC(114) $= -356.9705$, the test statistic $W_n = 9.9687$, and the approximated P-value=0.007, which leads us to accept there is another change point in R_t, and the change occurs at the $112 + 114 = 226$th position in R_t, corresponding to the 227th position (March 3, 2000) of S_t with the stock return

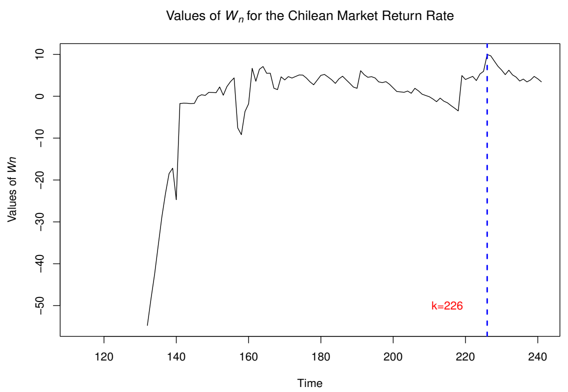
value 764.411. This change point may have been triggered by the Russian financial crisis of 1998, which devalued the ruble and suspended payments from its government to foreign creditors.

- In the following, we consider the sequence R_t for t from 113 to 226, and we have QMIC(114) = -269.8757, $\min_{1\le k\le 114}$ QMIC(k) = QMIC(57) = -262.1506, and the test statistic $W_n = 6.4836$ with the approximated P-value = 0.0145, and thus, we conclude that there is still a change point occurring at the 113 + 57 = 170th position in R_t corresponding to the 171st position (February 5, 1999) in S_t with the associated stock return 524.462. Argentina's financial crisis that emerged in December 1999 may have been responsible for this change point.

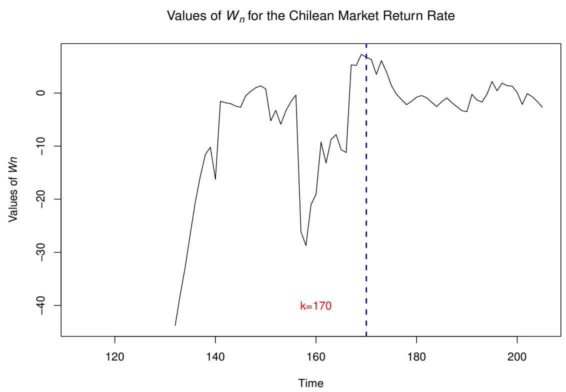


(a) Values of W_n for R_t, t from 1 to 261.

(b) Values of W_n for R_t, t from 1 to 112.



(c) Values of W_n for R_t, t from 113 to 261.

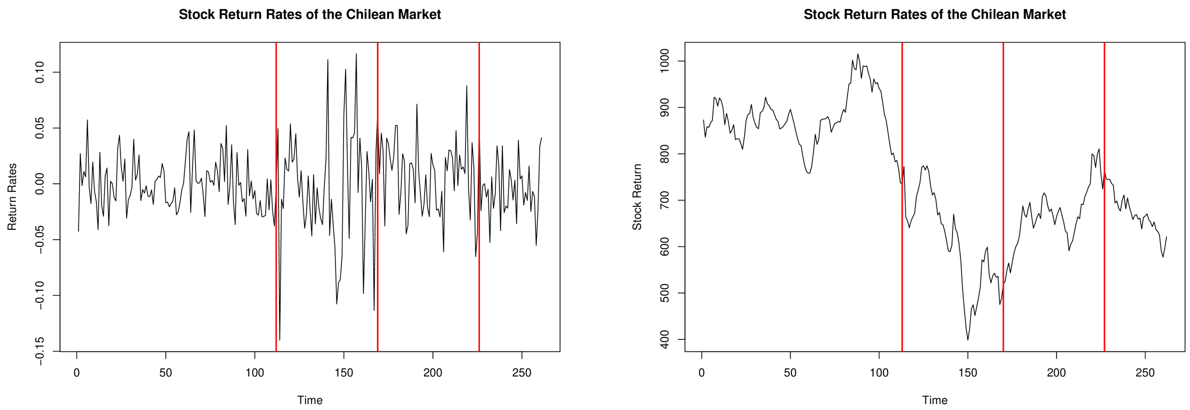


(d) Values of W_n for R_t, t from 113 to 226.

Figure 2. The corresponding values of the W_n test statistic for each sequence of Chilean R_t with the possible locations of the change points.

The graphs of the Chilean stock return S_t and return rate R_t with the locations of the corresponding change points are presented in Figure 3.

In our analysis, we identified three significant change points in the Chilean stock returns, each corresponding to major global financial events. As shown in Table 4, the first change point on December 26, 1997, aligns with the Asian financial crisis, reflecting its global impact on markets like Chile. The second, on February 5, 1999, corresponds with the Russian financial crisis, leading to a notable dip in returns due to global investor uncertainty. The final change point on March 3, 2000, is linked to the Argentine crisis, which also affected Chile through regional economic ties. These change points illustrate the sensitivity of the Chilean market to external financial shocks.



(a) Detected change points in R_t. **(b)** Detected change points in stock return S_t.

Figure 3. Detected change points in stock return rate R_t and stock return S_t.

Table 4. Change points of the Chilean stock return data.

Location	Date	Value	Possible reason
113th	December 26, 1997	736.133	The Asian financial crisis in 1997
171st	February 5, 1999	524.462	The Russian financial crisis in 1999
227th	March 3, 2000	764.411	The Argentina's crisis in 2000

Many results have been obtained for the variation point test of the Chilean return data. Arellano-Valle, Castro, and Loschi [2] detected only one change point in this dataset utilizing the Bayesian method and located the change point in the first week of February 1998. Ngunkeng and Wei [15] detected two change points in this dataset based on the Schwarz information criterion, which was at positions 113 and 170. The same result is also obtained by Said, Wei, and Tian [18] based on the modified information criterion. Using the method proposed in this paper, while detecting an approximate agreement with the change points mentioned in the latter two methods, we also detect another change point located at the 227th position. From Figure 3, we can see that the additional change point is also reasonable.

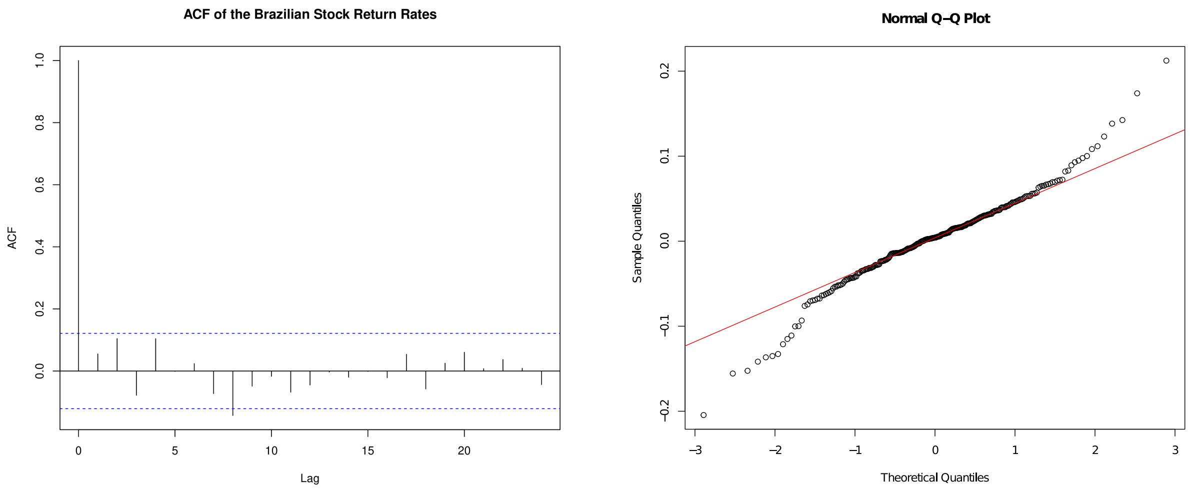
5.2. The Brazilian stock market

Subsequently, the Brazilian stock return data are examined using the identical analytical procedure employed for the Chilean stock returns. To test the independence in the Brazilian stock return data, we plot the ACF for the Brazilian stock return rate R_t, as shown in the left panel of Figure 4. The ACF plot exhibits no strong auto-correlation for the R_t data, and to be more convincing, we calculate the Portmanteau test statistic Q_k of R_t for the Brazilian market, and the result is given by

$$Q_{24} = 261 \times \sum_{i=1}^{24} r_i^2 = 261 \times 0.0910859 = 23.77342 < \chi^2_{0.95}(24) = 36.415.$$

Therefore, we fail to reject the null hypothesis of independence of R_t data for the Brazilian stock market. The right panel in Figure 4 shows that the normality of the R_t data is violated, so we use the skew normal distribution as a replacement for change point detectiion. The estimated parameters for the fitted skew normal distribution are as follows:

$$\hat{\mu} = 0.0406\,(\text{SE} = 0.0092), \quad \hat{\sigma} = 0.0650\,(\text{SE} = 0.0057), \quad \hat{\lambda} = -1.0803\,(\text{SE} = 0.3474).$$



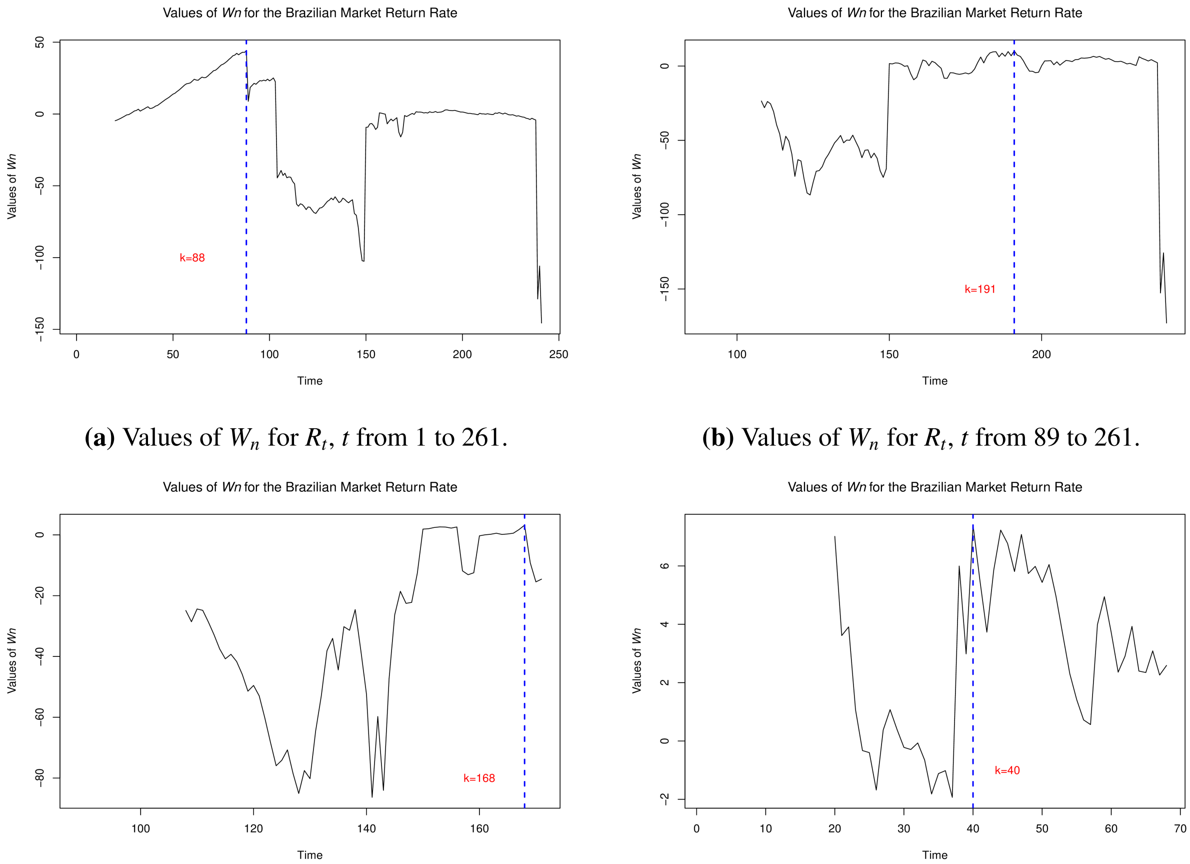
(a) ACF values of R_t. **(b)** The normal Q-Q plot of R_t.

Figure 4. The behavior of the stock return rate R_t series of the Brazilian market.

In the following, the suggested QMIC method and the binary segmentation procedure are applied to the R_t dataset to detect the potential change points in it. As with the previous analysis, we applied the same data trimming method. Figure 5 shows the corresponding values of the W_n test statistics across k for each sub-sequence of R_t and the possible locations of the change points. The procedures and findings of the change point detection for the R_t dataset of the Brazilian stock market are listed below.

- We first screen the whole R_t series from 1 to 261 for possible changes. Under the null hypothesis, we obtain that QMIC(261) = −491.5966, $\min_{1\leq k\leq 261}$ QMIC(k) = QMIC(88) = −518.8437, the test statistic $W_n = 43.9407$, and the approximated P-value = 0.000 < 0.05 with $B = 2000$ bootstrap sampling. Therefore, we reject the null hypothesis and conclude that there exists a change point in the R_t series and that the change occurs at the 88th position. This change point corresponds to the 89th position of the original series with a stock return value of 1278.702 on July 11, 1997, which may have been caused by the Asian financial crisis of 1997.
- Then we consider the sub-sequence R_t for t from 89 to 261, and the results are QMIC(173) = −250.2238, $\min_{1\leq k\leq 173}$ QMIC(k) = QMIC(103) = −244.9595, and W_n = 10.1956. Through B = 2000 bootstrap sampling, the P-value is approximately equal to 0.006, and hence, there is a change point in the sub-sequence of R_t. The change point occurs at position 103 of the sub-sequence, corresponding to the 103 + 88 = 191st position of the whole R_t, which is at the 192nd position of the original data with a stock return 675.134 on July 2, 1999. We believe that the Russian financial crisis of 1998 may have contributed to this change point.

- Next, we consider the sub-sequence R_t for t from 89 to 191, and we observe that QMIC(103) = −90.7087, $\min_{1\leq k\leq 103}$ QMIC(k) = QMIC(80) = −80.0023, W_n = 3.1978, and the approximated P-value = 0.1065 > 0.05. Therefore, there is no sufficient reason to assume that there is a change point in this sub-sequence.
- Furthermore, the R_t sequence for t from 1 to 88 is considered. The QMIC(88) = −261.4871, W_n = 7.39759, and $\min_{1\leq k\leq 88}$ QMIC(k) = QMIC(40) = −255.4527. The P-value obtained from B = 2000 bootstrap samples is 0.018, and thus there is a change point at position 40, which is equivalent to the 41st position in the Brazilian S_t series with a stock return of 728 on August 9, 1996. This change point may be due to the financial crisis of Mexico in 1995.

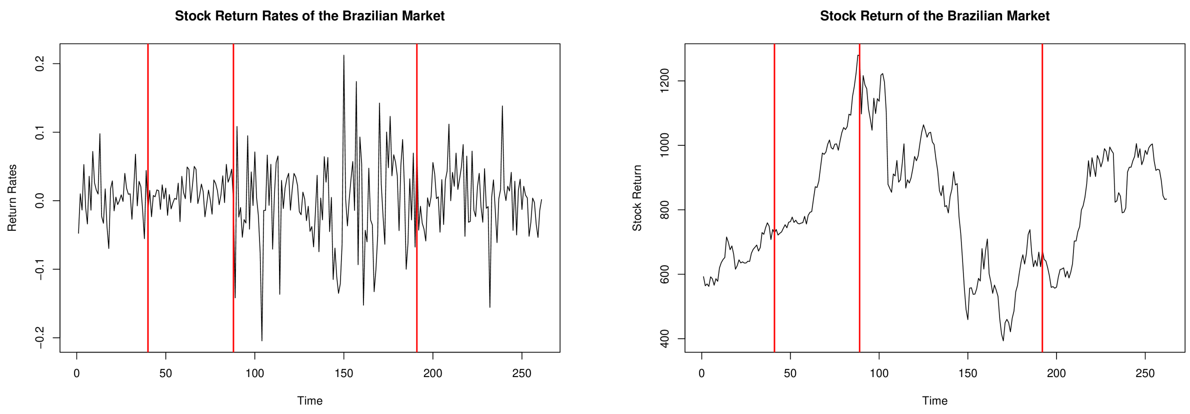


(a) Values of W_n for R_t, t from 1 to 261.
(b) Values of W_n for R_t, t from 89 to 261.
(c) Values of W_n for R_t, t from 89 to 191.
(d) Values of W_n for R_t, t from 1 to 88.

Figure 5. The corresponding values of the W_n test statistic for each sequence of Brazilian R_t with the possible locations of the change points.

The graphs of the Brazilian stock return S_t and return rate R_t with the locations of the corresponding change points are presented in Figure 6.

In summary, we can see that there are three change points in the Brazilian stock return, and the corresponding results can be found in Table 5. The first change point on August 9, 1996, coincides with the aftermath of the 1995 Mexican financial crisis, reflecting the broader regional economic instability. The second change point on July 11, 1997, corresponds to the onset of the 1997 Asian financial crisis, highlighting the global impact on emerging markets, including Brazil. The third change point on July 2, 1999, aligns with the aftermath of the 1998 Russian financial crisis, demonstrating the sensitivity of the Brazilian market to global financial disturbances.



(a) Detected change points in R_t. **(b)** Detected change points in stock return S_t.

Figure 6. Detected change points in stock return rate R_t and stock return S_t.

Table 5. Change points of the Brazilian stock return data.

Location	Date	Value	Possible reason
41st	August 9, 1996	728.000	The Mexican financial crisis in 1995
89th	July 11, 1997	1278.702	The Asian financial crisis in 1997
192nd	July 2, 1999	675.134	The Russian financial crisis in 1998

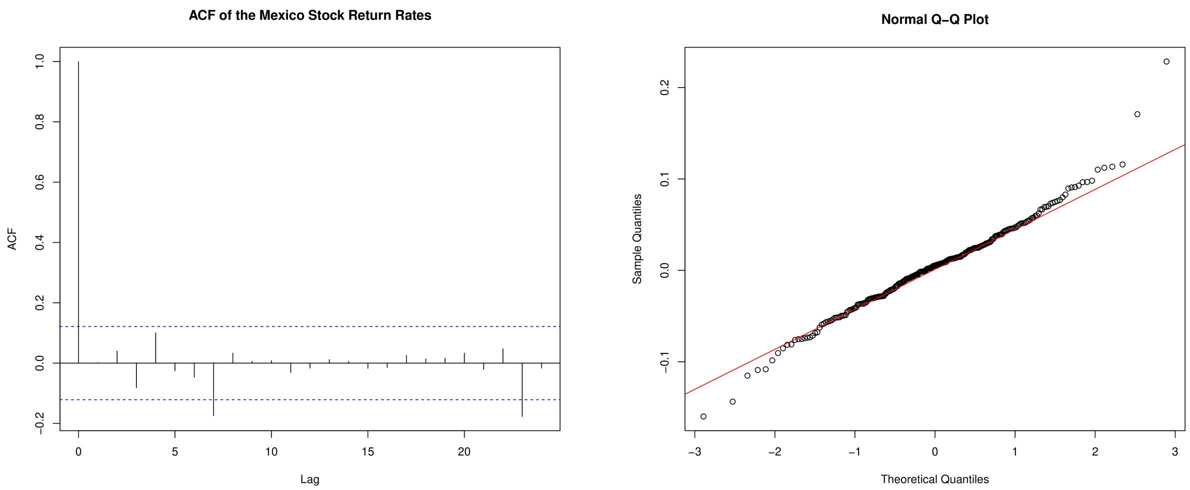
Arellano-Valle, Castro, and Loschi [2] showed that there was one change point in Brazilian stock returns and that the change point occurred in the first week of August 1997, while Ngunkeng and Wei [15] suggested that there were four change points in the data, occurring on July 11, 1997, March 8, 1996, July 31, 1998, and June 2, 2000, respectively. In this paper, using our proposed method, a total of three variation points are detected, including the most probable one (July 11, 1997), and the other two change points occur where there are significant changes in Brazilian stock return data. Our method ensures that the change points do not occur at the beginning or the end of the process, compared to Ngunkeng's results.

5.3. *The Mexican stock market*

Additionally, the Mexican stock return data are subjected to a similar analytical scrutiny, maintaining methodological consistency across all datasets.

To examine the independence of the Mexican stock market data, we calculated the statistic $Q_{24} = 261 \times \sum_{i=1}^{24} r_i^2 = 23.77342$, which is less than the critical value $\chi^2_{0.95}(24) = 36.415$. This leads us to not reject the null hypothesis, suggesting that the R_t series for the Mexican market are independent. The left graph in Figure 7 depicts the ACF values of the R_t series, and the right graph presents the normal Q-Q plot, which indicates a deviation from normality. Furthermore, the application of normality tests, such as the Shapiro-Wilk test, substantiates the nonconformance with the normality assumption. The estimated parameters for the fitted skew normal distribution are as follows:

$$\hat{\mu} = -0.0302\,(\text{SE} = 0.0091), \quad \hat{\sigma} = 0.0598\,(\text{SE} = 0.0055), \quad \hat{\lambda} = 1.0368\,(\text{SE} = 0.3554).$$



(a) ACF values of R_t. **(b)** The normal Q-Q plot of R_t.

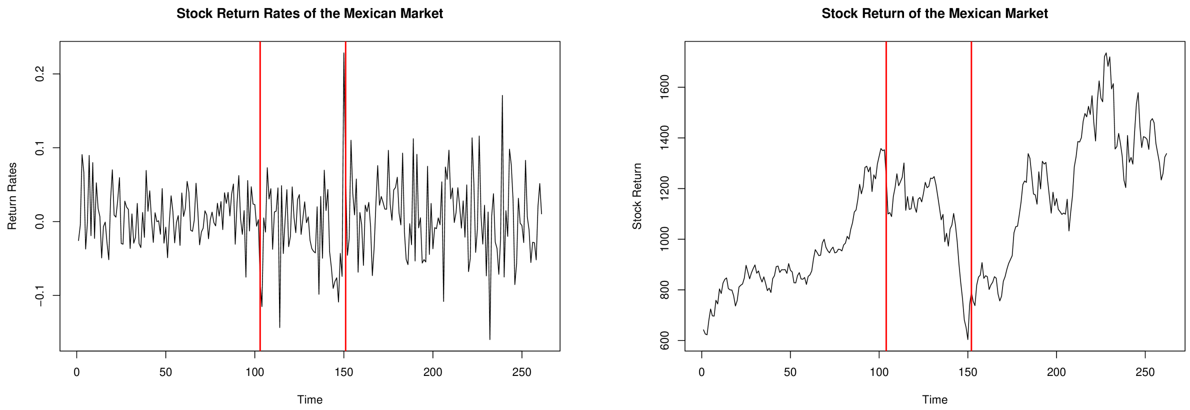
Figure 7. The behavior of the stock return rate R_t series of the Mexican market.

The procedures and results of change point detection in the R_t dataset of the Mexican stock market are presented below:

- We first screen the whole R_t series from 1 to 261 for possible changes. Under the null hypothesis, we obtain that QMIC(261) = −522.5629, $\min_{1\le k\le 261}$ QMIC(k) = QMIC(103) = −528.2583, the test statistic $W_n = 22.3890$, and the approximated P-value = 0.000 < 0.05 with $B = 2000$ bootstrap sampling. Therefore, we reject the null hypothesis and conclude that there exists a change point in the R_t series and that the change occurs at the 103rd position. This change point corresponds to the 104th position of the original series with a stock return value of 1242.851 on October 24, 1997.
- Similarly, we examine all possible subsequences using the binary segmentation method. We identified another change at the 152nd position in the R_t series, corresponding to a stock return value of 791.198. This change occurred on September 25, 1998.

In summary, we have identified changes at the 104th and 152nd positions, corresponding to October 24, 1997, and September 25, 1998, respectively. These changes may have been influenced by the Asian financial crisis of 1997 and the Russian financial crisis of 1998. Figure 8 displays the monthly stock return rate and the monthly stock return index for Mexico, including identified change points.

Arellano-Valle, Castro, and Loschi [2] showed that there was one change point in Mexican stock returns and that the change point occurred in the 1st week of September 1997. In contrast, the proposed QMIC approach identifies two distinct changes at the 104th and the 152nd positions, respectively, whereas another analytical method detects only a single change.



(a) Detected change points in R_t. **(b)** Detected change points in stock return S_t.

Figure 8. Detected change points in stock return rate R_t and stock return S_t.

By analyzing these three examples, it can be seen that our approach not only produces analytic results that are very similar to those obtained from a classical method based on the log-likelihood function but also provides programs for complex models with intractable likelihood functions.

6. Conclusions

In this study, we propose an enhancement to the modified information criterion (MIC) for detecting change points in skew normal distribution models by integrating the expectation-maximization (EM) algorithm's Q-function. This novel QMIC framework establishes a procedure for detecting simultaneous changes in all three parameters of a skew normal distribution, thereby improving the precision and robustness of change point detection across various data sequences. The complexity of deriving an analytic asymptotic distribution for the test statistic under QMIC is addressed through bootstrap simulations, allowing for flexible and accurate determination of critical values at different significance levels. Extensive simulations demonstrate that QMIC outperforms traditional methods, such as the MIC and the Bayesian information criterion (BIC), by more effectively capturing the nuances of skew normal distributions, particularly in datasets with non-normal characteristics.

The application of QMIC to stock market datasets highlights its practical utility, successfully identifying multiple change points that may be overlooked by conventional methods. This capability is especially valuable in financial data analysis, where detecting subtle shifts can significantly impact decision-making processes. The method's ability to uncover intricate structural changes underscores its potential for broader applications across various domains. In summary, the QMIC framework presents a valuable contribution to change point detection methodologies, offering a robust, sensitive, and versatile tool that complements existing state-of-the-art techniques, enhancing both the theoretical understanding and practical analysis of complex data.

Author contributions

Yang Du: conceptualization, methodology, writing-original draft and writingreview & editing; Weihu Cheng: validation and supervision. All authors have thoroughly reviewed and approved the final version of the manuscript for publication.

Conflict of interest

The authors declare no conflict of interest.

References

1. E. Page, Continuous inspection schemes, *Biometrika*, **41** (1954), 100–115. https://doi.org/10.2307/2333009
2. R. Arellano-Valle, L. Castro, R. Loschi, Change point detection in the skew-normal model parameters, *Commun. Stat.-Theory M.*, **42** (2013), 603–618. https://doi.org/10.1080/03610926.2011.611321
3. H. Chernoff, S. Zacks, Estimating the current mean of a normal distribution which is subjected to changes in time, *Ann. Math. Stat.*, **35** (1964), 999–1018. https://doi.org/10.1214/aoms/1177700517
4. D. Hsu, Tests for variance shift at an unknown time point, *J. R. Stat. Soc. C-Appl.*, **26** (1977), 279–284. https://doi.org/10.2307/2346968
5. D. Hinkley, E. Hinkley, Inference about the change-point in a sequence of binomial variables, *Biometrika*, **57** (1970), 477–488. https://doi.org/10.1093/biomet/57.3.477
6. K. Worsley, On the likelihood ratio test for a shift in location of normal populations, *J. Am. Stat. Assoc.*, **74** (1979), 365–367. https://doi.org/10.1080/01621459.1979.10482519
7. K. Worsley, Confidence regions and tests for a change-point in a sequence of exponential family random variables, *Biometrika*, **73** (1986), 91–104. https://doi.org/10.1093/biomet/73.1.91
8. B. Darkhovski, Nonparametric methods in change-point problems: a general approach and some concrete algorithms, *IMS Lecture Notes Monogr. Ser.*, **23** (1994), 99–107. https://doi.org/10.1214/lnms/1215463117
9. E. Gombay, L. Horvath, An application of U-statistics to change-point analysis, *Acta Sci. Math.*, **60** (1995), 345–357.
10. E. Gombay, M. Hušková, Rank based estimators of the change-point, *J. Stat. Plan. Infer.*, **67** (1998), 137–154. https://doi.org/10.1016/S0378-3758(97)00099-2
11. X. Cai, K. Said, W. Ning, Change-point analysis with bathtub shape for the exponential distribution, *J. Appl. Stat.*, **43** (2016), 2740–2750. https://doi.org/10.1080/02664763.2016.1143455
12. T. Wang, W. Tian, W. Ning, Likelihood ratio test change-point detection in the skew slash distribution, *Commun. Stat.-Simul. C.*, **51** (2022), 5068–5080. https://doi.org/10.1080/03610918.2020.1755869
13. C. Ferreira, T. Mattos, N. Balakrishnan, Mean-shift outliers model in skew scale-mixtures of normal distributions, *J. Stat. Comput. Sim.*, **86** (2016), 2346–2361. https://doi.org/10.1080/00949655.2015.1110819

14. J. Chen, A. Gupta, Testing and locating variance changepoints with application to stock prices, *J. Am. Stat. Assoc.*, **92** (1997), 739–747. https://doi.org/10.1080/01621459.1997.10474026

15. G. Ngunkeng, W. Ning, Information approach for the change-point detection in the skew normal distribution and its applications, *Sequential Anal.*, **33** (2014), 475–490. https://doi.org/10.1080/07474946.2014.961845

16. J. Chen, A. Gupta, J. Pan, Information criterion and change point problem for regular models, *Sankhyā: The Indian Journal of Statistics*, **68** (2006), 252–282.

17. J. Pan, J. Chen, Application of modified information criterion to multiple change point problems, *J. Multivariate Anal.*, **97** (2006), 2221–2241. https://doi.org/10.1016/j.jmva.2006.05.009

18. K. Said, W. Ning, Y. Tian, Modified information criterion for testing changes in skew normal model, *Braz. J. Probab. Stat.*, **33** (2019), 280–300. https://doi.org/10.1214/17-BJPS388

19. M. Kilai, G. Waititu, W. Kibira, R. Aldallal, M. Bakr, Y. Tashkandy, et al., Information approach for change point analysis of EGGAPE distribution and application to COVID-19 data, *Math. Probl. Eng.*, **2022** (2022), 9924902. https://doi.org/10.1155/2022/9924902

20. A. Dempster, N. Laird, D. Rubin, Maximum likelihood from incomplete data via EM algorithm, *J. R. Stat. Soc. B-Methol.*, **39** (1977), 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

21. R. Boyles, On the convergence of the EM algorithm, *J. R. Stat. Soc. B*, **45** (1983), 47–50. https://doi.org/10.1111/j.2517-6161.1983.tb01229.x

22. R. Redner, H. Walker, Mixture densities, maximum-likelihood and the EM algorithm, *SIAM Rev.*, **26** (1984), 195–237. https://doi.org/10.1137/1026034

23. X. Meng, D. Rubin, Maximum likelihood estimation via the ecm algorithm: a general framework, *Biometrika*, **80** (1993), 267–278. https://doi.org/10.2307/2337198

24. C. Wu, On the convergence properties of the EM algorithm, *Ann. Statist.*, **11** (1983), 95–103. https://doi.org/10.1214/aos/1176346060

25. T. Lin, J. Lee, S. Yen, Finite mixture modelling using the skew normal distribution, *Stat. Sinica*, **17** (2007), 909–927.

26. T. Lin, Maximum likelihood estimation for multivariate skew normal mixture models, *J. Multivariate Anal.*, **100** (2009), 257–265. https://doi.org/10.1016/j.jmva.2008.04.010

27. V. Lachos, P. Ghosh, R. Arellano-Valle, Likelihood based inference for skew-normal independent linear mixed models, *Stat. Sinica*, **20** (2010), 303–322.

28. T. Abe, H. Fujisawa, T. Kawashima, C. Ley, EM algorithm using overparameterization for multivariate skew-normal distribution, *Economet. Stat.*, **19** (2021), 151–168. https://doi.org/10.1016/j.ecosta.2021.03.003

29. H. Zhu, S. Lee, Local influence for incomplete-data models, *J. R. Stat. Soc. B*, **63** (2001), 111–126. https://doi.org/10.1111/1467-9868.00279

30. P. Perron, T. Vogelsang, Testing for a unit root in a time series with a changing mean: corrections and extensions, *J. Bus. Econ. Stat.*, **10** (1992), 467–470. https://doi.org/10.1080/07350015.1992.10509923

31. L. Vostrikova, Detecting "disorder" in multidimensional random processes, *Dokl. Akad. Nauk SSSR*, **259** (1981), 270–274.
32. R. Basso, V. Lachos, C. Cabral, P. Ghosh, Robust mixture modeling based on scale mixtures of skew-normal distributions, *Comput. Stat. Data Anal.*, **54** (2010), 2926–2941. https://doi.org/10.1016/j.csda.2009.09.031
33. C. Zeller, C. Barbosa Cabral, V. Lachos, L. Benites, Finite mixture of regression models for censored data based on scale mixtures of normal distributions, *Adv. Data Anal. Classif.*, **13** (2019), 89–116. https://doi.org/10.1007/s11634-018-0337-y
34. T. Mattos, A. Garay, V. Lachos, Likelihood-based inference for censored linear regression models with scale mixtures of skew-normal distributions, *J. Appl. Stat.*, **45** (2018), 2039–2066. https://doi.org/10.1080/02664763.2017.1408788
35. D. Hsu, Detecting shifts of parameter in gamma sequences with applications to stock-price and air-traffic flow analysis, *J. Am. Stat. Assoc.*, **74** (1979), 31–40. https://doi.org/10.2307/2286717