mathematics

MDPI

Article

Change Point Analysis for Kumaraswamy Distribution

Weizhong Tian [1,*], Liyuan Pang [2], Chengliang Tian [3] and Wei Ning [4]

1 College of Big Data and Internet, Shenzhen Technology University, Shenzhen 518118, China
2 School of Science, Xi'an University of Technology, Xi'an 710048, China
3 College of Computer Science and Technology, Qingdao University, Qingdao 266071, China
4 Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403, USA
* Correspondence: tianweizhong@sztu.edu.cn

Abstract: The Kumaraswamy distribution is a common type of bounded distribution, which is widely used in agriculture, hydrology, and other fields. In this paper, we use the methods of the likelihood ratio test, modified information criterion, and Schwarz information criterion to analyze the change point of the Kumaraswamy distribution. Simulation experiments give the performance of the three methods. The application section illustrates the feasibility of the proposed method by applying it to a real dataset.

Keywords: Kumaraswamy distribution; change point; likelihood ratio test; modified information criterion; Schwarz information criterion; maximum likelihood estimate

MSC: 62C05; 62P30; 62E99; 62F03

check for updates

Citation: Tian, W.; Pang, L.; Tian, C.; Ning, W. Change Point Analysis for Kumaraswamy Distribution. *Mathematics* **2023**, *11*, 553. https://doi.org/10.3390/math11030553

Academic Editor: Vasile Preda

Received: 28 December 2022
Revised: 15 January 2023
Accepted: 17 January 2023
Published: 20 January 2023

1. Introduction

The change-point problem, introduced by Page [1,2], has become more important in many application fields, such as finance, hydrology, and genetics. In statistics, several theories and applications related to change-point analysis have been studied by scholars. Sen and Srivastava [3] deduced the exact and asymptotic distribution of the test statistics of a single change point in a normal random variable sequence. Cai et al. [4] considered the likelihood ratio test (LRT) and Schwarz information criterion (SIC) to detect the change-point problem of an exponential distribution. Chen and Ning [5] investigated the change point of an exponential-logarithmic distribution using the modified information criterion (MIC) method and applied it to biological and engineering aspects of the dataset. Said et al. [6] analyzed the change point of the skew-normal distribution by MIC, LRT, and the Bayesian information criterion (BIC). Wang et al. [7] extended the method of LRT into the skew-slash distribution. Tian and Yang [8] studied the change-point problem of weighted exponential distributions based on the LRT, MIC and SIC procedures.

In real life, we often encounter some measurements, such as the proportion of a certain feature, the scores of some ability tests, and different indicators and ratios, which are located in the $(0,1)$ interval. In such cases, bounded distributions are essential to model these phenomena. As we know, the Kumaraswamy (Kw) distribution plays an important role in bounded distributions. The Kw distribution was introduced by Kumaraswamy [9] to study the daily rainfall in hydrology. Its probability density function (pdf) was given by

$$f(x;\gamma,\beta) = \gamma\beta x^{\gamma-1}(1-x^{\gamma})^{\beta-1}, \quad 0<x<1,$$

where $\gamma>0$ and $\beta>0$ were shape parameters, and it was denoted by $X \sim Kw(\gamma,\beta)$. The density function is unimodal if $\gamma>1$ and $\beta>1$ and uniantimodal if $\gamma<1$ and $\beta<1$. The density function increases for $\gamma>1$ and $\beta \leq 1$, decreases for $\gamma \leq 1$ and $\beta>1$, and is constant for $\gamma=\beta=1$.

The *Kw* distribution was considered to be a substitutive model for the beta distribution in practical terms and has drawn much academic attention and concern. In fact, the *Kw* and beta distributions have the following properties in common: the shape types of their pdfs are the same, and the power function and the uniform distribution are similar in both their cases. Furthermore, the *Kw* distribution has some additional advantages over the beta distribution, such as its simple explicit formulas for the distribution functions and quantile function, which did not involve any special functions. Moreover, the simplicity of the quantile function provided a simple formula for random variable generation. See Jones [10] for a detailed description. Fletcher and Ponnambalam [11] used the *Kw* distribution to analyze reservoir storage capacity. Nadarajah [12] mentioned the *Kw* distribution as a special case of the beta distribution, and clarified that the *Kw* distribution was more effective than the beta distribution. Jones [10] systematically studied the basic statistical properties of the *Kw* distribution and estimated its parameters by the maximum likelihood estimation method. Nadar et al. [13] conducted a statistical correlation analysis of the *Kw* distribution for the recorded values. Meanwhile, some new families of distributions have been proposed based on the *Kw* distribution, such as Saulo et al. [14], who studied the *Kw* Birnbaum–Saunders distribution, which provided enormous flexibility in modeling heavy-tailed and skewed data. Lemonte et al. [15] established the exponentiated *Kw* distribution and used the model to effectively fit life data. Mameli [16] pointed out that the *Kw* skew-normal distribution was a valid alternative to the beta skew-normal distribution. Iqbal et al. [17] proposed the generalized inverted *Kw* distribution to model a dataset of prices of wooden toys for 31 children.

Based on our knowledge, there is little research on the change point of the *Kw* distribution. Therefore, it is of a certain significance to study the change-point detection of the *Kw* distribution. The remaining organizational parts of the paper are as follows. The related basic theoretical knowledge and three methods of change point detection based on the *Kw* distribution are introduced in detail in Section 2. Simulation studies are carried out for three different detection methods in Section 3. Real data applications are studied in Section 4. Some conclusions are given in Section 5.

2. Methodology

Let $X_1, X_2, \cdots, X_n$ be a sequence of independent *Kw* random variables with parameters $\gamma_1, \gamma_2, \cdots, \gamma_n$ and $\beta_1, \beta_2, \cdots, \beta_n$. We are interested in testing the null hypothesis,

$$H_0 : \gamma_1 = \gamma_2 = \cdots = \gamma_n = \gamma \; and \; \beta_1 = \beta_2 = \cdots = \beta_n = \beta,$$

against the alternative hypothesis

$$H_1 : \gamma_1 = \gamma_2 = \cdots = \gamma_k \neq \gamma_{k+1} = \gamma_{k+2} = \cdots = \gamma_n,$$

and

$$\beta_1 = \beta_2 = \cdots = \beta_k \neq \beta_{k+1} = \beta_{k+2} = \cdots = \beta_n.$$

Under H_0, the log-likelihood function is given by

$$\log L_0 = n \log(\gamma\beta) + (\beta - 1) \sum_{i=1}^{n} \log(x_i) + (\gamma - 1) \sum_{i=1}^{n} \log(1 - x_i^{\beta}). \tag{1}$$

We take the first derivatives of the Equation (1) with respect to γ and β. The MLEs $\hat{\gamma}$ and $\hat{\beta}$ of γ and β can be obtained by solving the following equations:

$$\frac{\partial \log L_0}{\partial \gamma} = \frac{n}{\gamma} + \sum_{i=1}^{n} \log(1 - x_i^{\beta}) = 0,$$
$$\frac{\partial \log L_0}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^{n} \log(x_i) - (\gamma - 1) \sum_{i=1}^{n} \frac{x_i^{\beta} \log(x_i)}{1 - x_i^{\beta}} = 0.$$

Under H_1, the log-likelihood function is given by

$$\begin{aligned}\log L_1 = {} & k \log(\gamma_1 \beta_1) + (\beta_1 - 1) \sum_{i=1}^{k} \log(x_i) + (\gamma_1 - 1) \sum_{i=1}^{k} \log(1 - x_i^{\beta_1}) \\ & + (n-k)(\log(\gamma_n \beta_n)) + (\beta_n - 1) \sum_{i=k+1}^{n} \log(x_i) + (\gamma_n - 1) \sum_{i=k+1}^{n} \log(1 - x_i^{\beta_n}).\end{aligned} \tag{2}$$

Similarly, we take the first derivatives of Equation (2) with respect to γ_1, β_1, γ_n and β_n. The MLEs $\hat{\gamma_1}$, $\hat{\beta_1}$, $\hat{\gamma_n}$ and $\hat{\beta_n}$ can be obtained by solving the following equations:

$$\frac{\partial \log L_1}{\partial \gamma_1} = \frac{k}{\gamma_1} + \sum_{i=1}^{k} \log(1 - x_i^{\beta_1}) = 0,$$
$$\frac{\partial \log L_1}{\partial \beta_1} = \frac{k}{\beta_1} + \sum_{i=1}^{k} \log(x_i) - (\gamma_1 - 1) \sum_{i=1}^{k} \frac{x_i^{\beta_1} \log(x_i)}{1 - x_i^{\beta_1}} = 0,$$
$$\frac{\partial \log L_1}{\partial \gamma_n} = \frac{n-k}{\gamma_n} + \sum_{i=k+1}^{n} \log(1 - x_i^{\beta_n}) = 0,$$
$$\frac{\partial \log L_1}{\partial \beta_n} = \frac{n-k}{\beta_n} + \sum_{i=k+1}^{n} \log(x_i) - (\gamma_n - 1) \sum_{i=k+1}^{n} \frac{x_i^{\beta_n} \log(x_i)}{1 - x_i^{\beta_n}} = 0.$$

2.1. Likelihood Ratio Test

The LRT is one of the most commonly used change point detection methods. The main idea of this method is to use the likelihood ratio idea to test the existence of some distribution parameter change point, that is, to estimate the relevant parameters by finding the maximum value of the likelihood function, where the change point itself is a parameter. The LRT method is a problem discussed earlier in change-point theory, which has been considered by many scholars. Said et al. [18] explained that the LRT procedure has considerable ability to detect the parameter changes of the skew-normal distribution model. Wang et al. [7] used LRT procedure to study the parameter changes of the skew-slash distribution. In the following, we describe the LRT test procedure in detail.

Assuming that k is an integer between 1 and n, if the change point occurs at k, we reject the null hypothesis H_0 for a sufficiently large value of the log-likelihood ratio $f_n(x;k)$, which is given by the following equation:

$$\begin{aligned} f_n(x;k) = -2\log\left(\frac{L_0}{L_1}\right) = -2\Bigg[& n \log(\hat{\gamma}\hat{\beta}) + (\hat{\beta} - 1) \sum_{i=1}^{n} \log(x_i) + (\hat{\gamma} - 1) \sum_{i=1}^{n} \log(1 - x_i^{\hat{\beta}}) \\ & - k \log(\hat{\gamma}_1 \hat{\beta}_1) - (\hat{\beta}_1 - 1) \sum_{i=1}^{k} \log(x_i) - (\hat{\gamma}_1 - 1) \sum_{i=1}^{k} \log(1 - x_i^{\hat{\beta}_1}) \\ & -(n-k)(\log(\hat{\gamma}_n \hat{\beta}_n)) - (\hat{\beta}_n - 1) \sum_{i=k+1}^{n} \log(x_i) - (\hat{\gamma}_n - 1) \sum_{i=k+1}^{n} \log(1 - x_i^{\hat{\beta}_n}) \Bigg]. \end{aligned}$$

We use $\hat{\gamma}, \hat{\beta}, \hat{\gamma}_1, \hat{\beta}_1, \hat{\gamma}_n$, and $\hat{\beta}_n$ to represent MLEs under the corresponding hypothesis of change point k. Since the change position k is unknown, the maximum value of the selected log-likelihood ratio test statistic is naturally defined as

$$Z_n = \max_{1<k<n} f_n(x;k).$$

Actually, if the change occurs at the very beginning or the very end of the data, we may not have enough observations to obtain the MLEs of the parameters, or the MLEs of the parameters may not be unique; see Said et al. [18]. Thus, we consider the trimmed version of the test statistics given by Zou et al. [19], as shown in the following formula:

$$Z'_n = \max_{k_0<k<n-k_0} f_n(x;k),$$

There are several choices for k_0. For example, Liu and Qian [20] suggested the choice of $k_0 = [\log n]^2$, Said et al. [18] chose $k_0 = 2[\log n]$, with $[x]$ representing the largest integer that is not greater than x. In this paper, we also choose $k_0 = 2[\log n]$. Thus, we reject H_0 if

$$\begin{aligned}
Z'_n &= \max_{k_0<k<n-k_0} f_n(x;k) \\
&= \max_{k_0<k<n-k_0} \Bigg\{ -2 \Bigg[n\log(\hat{\gamma}\hat{\beta}) + (\hat{\beta}-1)\sum_{i=1}^{n}\log(x_i) + (\hat{\gamma}-1)\sum_{i=1}^{n}\log(1-x_i^{\hat{\beta}}) \\
&- k\log(\hat{\gamma}_1\hat{\beta}_1) - (\hat{\beta}_1-1)\sum_{i=1}^{k}\log(x_i) - (\hat{\gamma}_1-1)\sum_{i=1}^{k}\log(1-x_i^{\hat{\beta}_1}) \\
&-(n-k)(\log(\hat{\gamma}_n\hat{\beta}_n)) - (\hat{\beta}_n-1)\sum_{i=k+1}^{n}\log(x_i) - (\hat{\gamma}_n-1)\sum_{i=k+1}^{n}\log(1-x_i^{\hat{\beta}_n}) \Bigg] \Bigg\}
\end{aligned}$$

is sufficiently large and the estimated change location $\hat{k} = \arg\max_{k_0<k<n-k_0} f_n(x;k)$. This means that for any given significance level α, we cannot reject H_0 if $Z'_n < c_{\alpha,n}$, where $c_{\alpha,n}$ is the critical value with respect to α for different sample size n. To obtain $c_{\alpha,n}$, we have to use the following theorem.

Theorem 1 (Csörgó and Horváth [21])**.** *Under H_0, as $n \to \infty$, for all $X \in \mathbb{R}$, we have*

$$\lim_{n\to\infty} P\left(A(\log u(n)) Z_n'^{\frac{1}{2}} - B(\log u(n)) \le x \right) = e^{-e^{-x}},$$

where

$$A(\log u(n)) = (2\log\log u(n))^{\frac{1}{2}},$$

$$B(\log u(n)) = 2\log\log u(n) + \log\log\log u(n) - \log\Gamma(1),$$

and

$$u(n) = \frac{n^2 - 2n[\log n] + (2[\log n])^2}{(2[\log n])^2}.$$

Proof. According to Theorem A1 in Csörgó and Horváth [21], which is given in Appendix A, let $t_1(n) = \frac{2[\log n]}{n}$, $t_2(n) = 1 - \frac{2[\log n]}{n}$. Then, we obtain

$$u(n) = \frac{1 - t_1(n)t_2(n)}{t_1(n)(1-t_2(n))} = \frac{n^2 - 2n[\log n] + (2[\log n])^2}{(2[\log n])^2}.$$

We consider the trimmed version of the test statistic Z'_n and use Theorem A1 instead of Corollary A1 in the proof. □

Using Theorem 1, the approximation of $c_{\alpha,n}$ is given by

$$
\begin{aligned}
1-\alpha =&P\big[Z'_n < c_{\alpha,n}|H_0\big] = P\big[0 < Z'_n < c_{\alpha,n}|H_0\big] = P\left[0 < Z'^{\frac{1}{2}}_n < (c_{\alpha,n})^{\frac{1}{2}}|H_0\right] \\
=&P\Big[-B(\log u(n)) < A(\log u(n))Z'^{\frac{1}{2}}_n - B(\log u(n)) < A(\log u(n))(c_{\alpha,n})^{\frac{1}{2}} \\
&- B(\log u(n))|H_0\Big] \\
=&P\left[A(\log u(n))Z'^{\frac{1}{2}}_n - B(\log u(n)) < A(\log u(n))(c_{\alpha,n})^{\frac{1}{2}} - B(\log u(n))\right] \\
&- P\left[A(\log u(n))Z'^{\frac{1}{2}}_n - B(\log u(n)) < -B(\log u(n))\right] \\
\cong& \exp\Big\{-\exp\Big\{B(\log u(n)) - A(\log u(n))(c_{\alpha,n})^{\frac{1}{2}}\Big\}\Big\} - \exp\{-\exp\{B(\log u(n))\}\}.
\end{aligned}
$$

Thus,

$$
c_{\alpha,n} \cong \left[\frac{\log[-\log(1-\alpha+\exp\{-\exp\{B(\log u(n))\}\})] - B(\log u(n))}{-A(\log u(n))}\right]^2. \tag{3}
$$

According to Equation (3), the empirical critical value $c_{\alpha,n}$ at different significance levels α and sample sizes n can be obtained, as shown in Table 1.

Table 1. Approximate critical values of LRT with different values of α and n.

n	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	n	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
15	27.9478	12.6744	8.8511	90	21.3668	13.6862	10.8365
20	21.6147	12.6386	9.4401	100	21.3745	13.7889	10.9661
35	21.4807	12.8847	9.7842	110	21.3845	12.6386	11.0768
40	21.4207	13.0768	10.0440	120	21.3957	13.9551	11.1734
50	21.3725	13.3602	10.4182	140	21.4191	14.0856	11.3344
60	21.3889	13.2315	10.2496	160	21.4050	14.0108	11.2422
70	21.3679	13.4170	10.4920	180	21.4238	14.1086	11.3626
80	21.3633	13.5646	10.6819	200	21.4425	14.1922	11.4649

2.2. Schwarz Information Criterion

The SIC was proposed by Schwarz [22] in order to remedy the inconsistency of estimators in the model based on the Akaike information criterion (AIC). The advantage of SIC is that it is unnecessary to derive the asymptotic distribution of complex test statistics. The SIC under H_0 is expressed as

$$
SIC(n) = -2\log L_0(\hat{\gamma}, \hat{\beta}) + 2\log n,
$$

and for a fixed change location $1 < k < n$ where k is an integer, we consider

$$
SIC(k) = -2\log L_1(\hat{\gamma}_1, \hat{\beta}_1, \hat{\gamma}_n, \hat{\beta}_n) + 4\log n,
$$

where $\log L_0(\cdot)$ and $\log L_1(\cdot)$ are the log-likelihood functions of the random sample under H_0 and H_1, respectively. The choice to accept H_0 or H_1 depends on the principle of the minimum information criteria, i.e., we fail to reject H_0 if

$$
SIC(n) < \min_{1<k<n} SIC(k),
$$

and we reject H_0 if

$$
SIC(n) > \min_{1<k<n} SIC(k),
$$

and the location of the change point can be estimated using $\hat{k}$ as follows:

$$SIC(\hat{k}) = \min_{1<k<n} SIC(k).$$

To make the conclusion more statistically convincing, we consider the following test statistic:

$$T_n = SIC(n) - \min_{1<k<n} SIC(k).$$

Thus, we fail to reject H_0 if $T_n < c_{\alpha,n}$ instead of $SIC(n) < \min_{1<k<n} SIC(k)$, where $c_{\alpha,n}$ is determined by

$$1 - \alpha = P\left[SIC(n) < \min_{1<k<n} SIC(k) + c_{\alpha,n}|H_0\right].$$

In fact,

$$\begin{aligned}
T_n &= SIC(n) - \min_{1<k<n} SIC(k) = \max_{1<k<n}[SIC(n) - SIC(k)] \\
&= \max_{1<k<n}\left[-2\log L_0(\hat{\gamma},\hat{\beta}) + 2\log n - \left(-2\log L_1(\hat{\gamma}_1,\hat{\beta}_1,\hat{\gamma}_n,\hat{\beta}_n) + 4\log n\right)\right] \\
&= \max_{1<k<n}\left[-2\left(\log L_0(\hat{\gamma},\hat{\beta}) - \log L_1(\hat{\gamma}_1,\hat{\beta}_1,\hat{\gamma}_n,\hat{\beta}_n)\right) - 2\log n\right] \\
&= Z'_n - 2\log n,
\end{aligned}$$

where Z'_n is the test statistic of the LRT. Therefore, we obtain that

$$Z'_n = T_n + 2\log n.$$

Theorem 2 (Csörgó and Horváth [21])**.** *Under H_0, as $n \to \infty$, for all $X \in \mathbb{R}$, we have*

$$\lim_{n\to\infty} P\left(A(\log n)Z_n'^{\frac{1}{2}} - B(\log n) \le x\right) = e^{-2e^{-x}},$$

where

$$A(\log n) = (2\log\log n)^{\frac{1}{2}},$$

and

$$B(\log n) = 2\log\log n + \log\log\log n - \log\Gamma(1).$$

Proof. In Csörgó and Horváth [21]'s $C1 - C9$ conditions, we use Theorem A2 from Csörgó and Horváth [21] to give the above conclusion; see Appendix A for Theorem A2. □

From Theorem 2 above, the approximate expression of $c_{\alpha,n}$ is derived as follows:

$$\begin{aligned}
1 - \alpha =& P\left[SIC(n) < \min_{1<k<n} SIC(k) + c_{\alpha,n}|H_0\right] \\
=& P[T_n < c_{\alpha,n}|H_0] = P\left[Z'_n - 2\log n < c_{\alpha,n}|H_0\right] = P\left[0 < Z_n'^{\frac{1}{2}} < (2\log n + c_{\alpha,n})^{\frac{1}{2}}\right] \\
=& P\left[-B(\log n) < A(\log n)Z_n'^{\frac{1}{2}} - B(\log n) < A(\log n)(2\log n + c_{\alpha,n})^{\frac{1}{2}} - B(\log n)\right] \\
=& P\left[A(\log n)Z_n'^{\frac{1}{2}} - B(\log n) < A(\log n)(2\log n + c_{\alpha,n})^{\frac{1}{2}} - B(\log n)\right] \\
&- P\left[A(\log n)Z_n'^{\frac{1}{2}} - B(\log n) < -B(\log n)\right] \\
\cong& \exp\left\{-2\exp\left\{B(\log n) - A(\log n)(2\log n + c_{\alpha,n})^{\frac{1}{2}}\right\}\right\} - \exp\{-2\exp\{B(\log n)\}\}.
\end{aligned}$$

Thus,

$$c_{\alpha,n} \cong \left[\frac{B(\log n)}{A(\log n)} - \frac{1}{A(\log n)}\log\log[1 - \alpha + \exp\{-2\exp\{B(\log n)\}\}]^{-\frac{1}{2}}\right]^2 - 2\log n. \quad (4)$$

According to Equation (4), the critical empirical value $c_{\alpha,n}$ based on the SIC method can be obtained under different significance levels α and sample sizes n, as shown in Table 2.

Table 2. Approximate critical values of SIC with different values of α and n.

n	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	n	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
15	21.1982	10.6171	6.7932	70	16.8038	8.0819	4.8147
20	20.1949	10.1444	6.4766	80	16.4902	7.8592	4.6200
25	19.5007	9.7790	6.2091	90	16.2178	7.6623	4.4463
30	18.9726	9.4804	5.9793	100	15.9772	7.4857	4.2894
35	18.5476	9.2275	5.7782	150	15.0740	6.8020	3.6737
40	18.1927	9.0080	5.5997	200	14.4507	6.3133	3.2268
50	17.6222	8.6400	5.2932	250	13.9751	5.9321	2.8751
60	17.1733	8.3381	5.0362	300	13.5907	5.6193	2.5847

2.3. Modified Information Criterion

The MIC approach was proposed by Chen et al. [23] to solve the issue of the redundancy of parameters caused by the SIC method. The MIC under the H_0 is expressed as

$$MIC(n) = -2\log L_0(\hat{\gamma}, \hat{\beta}) + 2\log n. \tag{5}$$

For a fixed change location $1 < k < n$,

$$MIC(k) = -2\log L_1(\hat{\gamma}_1, \hat{\beta}_1, \hat{\gamma}_n, \hat{\beta}_n) + \left[4 + \left(\frac{2k}{n} - 1\right)^2\right] \log n, \tag{6}$$

where $\log L_0(\cdot)$ and $\log L_1(\cdot)$ are the log-likelihood functions of the random sample under H_0 and H_1, respectively. Then, we fail to reject H_0 if

$$MIC(n) < \min_{1<k<n} MIC(k),$$

and we reject H_0 if

$$MIC(n) > \min_{1<k<n} MIC(k).$$

Therefore, we can estimate the position of the change point $\hat{k}$ by

$$MIC(\hat{k}) = \min_{1<k<n} MIC(k). \tag{7}$$

In addition, we give the critical empirical value of the MIC method by the test statistic S_n in order to detect the presence of a change point faster and more efficiently. In the case that the S_n value is large enough, we reject the null hypothesis H_0, and the S_n value is given by the following formula:

$$\begin{aligned} S_n &= MIC(n) - \min_{1<k<n} MIC(k) + 2\log n \\ &= -2\log L_0(\hat{\gamma}, \hat{\beta}) - \min_{1<k<n}\left\{-2\log L_1(\hat{\gamma}_1, \hat{\beta}_1, \hat{\gamma}_n, \hat{\beta}_n) + \left(\frac{2k}{n} - 1\right)^2 \log n\right\}. \end{aligned} \tag{8}$$

For a given significance level α, the critical value of the test statistic under the null hypothesis H_0 is simulated by the Bootstrap method. Namely, a certain number of Bootstrap samples are drawn from the generated random numbers by sampling with replacement, and then the values of the test statistics are obtained from the Bootstrap samples, which are sorted, and the percentage of the sorted test statistics is used as the critical value for a given significance level. Tables 3 and 4 are the critical value of the MIC detection method obtained by the bootstrap method with some specific Kw distributions.

Table 3. Approximate critical values for MIC under different parameters.

n	$Kw(\cdot)$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$Kw(\cdot)$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
15	$(4, 0.5)$	16.0214	12.1440	9.9910	$(0.5, 3.5)$	12.7805	8.4131	6.6169
20	$(4, 0.5)$	13.5348	9.4129	7.7097	$(0.5, 3.5)$	9.1797	4.8304	3.0177
30	$(4, 0.5)$	12.5291	9.3880	7.2124	$(0.5, 3.5)$	14.6126	9.0067	6.5348
40	$(4, 0.5)$	12.4641	9.4281	7.4635	$(0.5, 3.5)$	13.3377	8.6004	5.0923
50	$(4, 0.5)$	11.6223	7.6873	6.0628	$(0.5, 3.5)$	12.0217	7.2745	5.1606
55	$(4, 0.5)$	15.2560	12.1560	10.1627	$(0.5, 3.5)$	16.4481	13.0652	11.7789
60	$(4, 0.5)$	19.7908	13.6454	11.8123	$(0.5, 3.5)$	16.1109	12.2783	10.4150
80	$(4, 0.5)$	17.7720	13.5650	11.3690	$(0.5, 3.5)$	16.7704	12.3216	10.3872
100	$(4, 0.5)$	17.4697	12.9035	10.9909	$(0.5, 3.5)$	10.0437	6.7711	4.1757
150	$(4, 0.5)$	18.0801	13.0013	11.0911	$(0.5, 3.5)$	15.0437	11.7711	10.1757
200	$(4, 0.5)$	16.2452	12.5573	10.6762	$(0.5, 3.5)$	15.5848	11.9499	10.0195

Table 4. Approximate critical values for MIC under parameters.

n	$Kw(\cdot)$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
15	$(5, 2)$	19.4604	12.9669	10.3656
20	$(5, 2)$	18.4070	13.0958	10.5379
30	$(5, 2)$	21.8363	16.0958	14.8520
40	$(5, 2)$	17.5919	13.9549	11.5930
50	$(5, 2)$	17.2003	10.8230	8.3099
55	$(5, 2)$	18.2946	13.9232	11.8138
60	$(5, 2)$	17.9610	13.1129	11.1608
80	$(5, 2)$	16.4356	12.1177	10.8341
100	$(5, 2)$	11.3732	7.0083	6.4604
150	$(5, 2)$	15.2264	11.8086	9.2993
200	$(5, 2)$	15.6294	12.1649	10.5434

However, we do not know whether the real dataset satisfies H_0 or H_1, which would be a problem. Thus, we cannot re-sample the data directly. We first assume the data satisfying H_0, which indicates it should be fitted by a Kw distribution, say, $Kw_0 = Kw(\hat{\gamma}, \hat{\beta})$, where $\hat{\gamma}$ and $\hat{\beta}$ are obtained by the MLE method. Then. we generate a random sample based on Kw_0 denoted by $x_1, x_2, \ldots, x_n$. Then, B Bootstrap samples are drawn from this generated sample with replacement, denoted by $y_1^{(i)}, y_2^{(i)}, \ldots, y_n^{(i)}, i = 1, 2, \ldots, B$. For each Bootstrap sample, we calculate S_n denoted by $S_n^{(i)}, i = 1, 2, \ldots, B$. Thus, the p_value can be approximated as follows:

$$p\_value = \frac{1}{B} \sum_{i=1}^{B} I\left(S_n^{(*)} \leq S_n^{(i)}\right), \tag{9}$$

where $I(\cdot)$ is the indicator function and $S_n^{(*)}$ is the value of S_n calculated from the original real data.

3. Simulation

Power refers to the probability of accepting the correct alternative hypothesis after rejecting the null hypothesis in a hypothesis test. We did not consider whether the test procedures detected the correct change point because we only evaluated whether there was a change point. Then, we gave the performance of the test procedures based on the efficacy of Z_n', T_n, and S_n in different simulation scenarios. In the simulation study, the assessment of the robustness of the test relative to the underlying distribution was not the goal of the study; thus, all the data generated in the simulation part came from the Kw distribution. We conducted simulations 1000 times under $Kw(\gamma, \beta)$ with different values of the shape parameters γ and β. The test statistics Z_n', S_n and T_n were calculated and compared to the critical values corresponding to the significance levels of $\alpha = 0.01, 0.05$ and 0.1. After rejecting the null hypothesis, we calculated the powers of the SIC, the LRT, and the MIC with different sample sizes $n = 20$, 50, and 100 and assumed the change occurs at the position of approximately $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$ of the sample sizes n. The detailed results are displayed in Tables 5–13. We choose the parameter values $(\gamma_1, \beta_1) = (4, 0.5)$, (0.5, 3.5) and (5, 2) before the change, which was based on the increasing, decreasing, and unimodal types of the Kw distribution, respectively. The selection of parameter values after the change is based on the changing of one parameter or two parameters of the Kw distribution. In a word, the following three Kw distributions are considered:

I The distribution follows $Kw(4, 0.5)$ before the change and follows $Kw(\gamma_n, \beta_n)$ after the change, where (γ_n, β_n) are set to be $(4, 2.5), (0.2, 0.5), (2, 2), (4, 0.5)$.

II The distribution follows $Kw(0.5, 3.5)$ before the change and follows $Kw(\gamma_n, \beta_n)$ after the change, where (γ_n, β_n) are set to be $(0.5, 1.5), (1.2, 3.5), (0.8, 2.5), (0.5, 3.5)$.

III The distribution follows $Kw(5, 2)$ before the change and follows $Kw(\gamma_n, \beta_n)$ after the change, where (γ_n, β_n) are set to be $(5, 3.5), (0.5, 2), (1.5, 4.5), (5, 2)$.

Table 5. Powers of the LRT, SIC, and MIC procedures at $(\gamma_1, \beta_1) = (4, 0.5)$, $n = 20$.

				(γ_n, β_n)			
α	**k**	**Model**	(γ_1, β_1)	**(4, 2.5)**	**(0.2, 0.5)**	**(2, 2)**	**(4, 0.5)**
0.01	5	LRT	(4, 0.5)	0.116	0.037	0.146	0.000
		SIC	(4, 0.5)	0.048	0.015	0.075	0.000
		MIC	(4, 0.5)	0.483	0.594	0.610	0.016
	10	LRT	(4, 0.5)	0.107	0.248	0.187	0.001
		SIC	(4, 0.5)	0.024	0.076	0.086	0.000
		MIC	(4, 0.5)	0.567	0.809	0.707	0.018
	15	LRT	(4, 0.5)	0.036	0.225	0.079	0.004
		SIC	(4, 0.5)	0.008	0.159	0.018	0.000
		MIC	(4, 0.5)	0.592	0.804	0.583	0.015
0.05	5	LRT	(4, 0.5)	0.520	0.471	0.641	0.027
		SIC	(4, 0.5)	0.342	0.235	0.434	0.012
		MIC	(4, 0.5)	0.734	0.773	0.857	0.033
	10	LRT	(4, 0.5)	0.601	0.846	0.780	0.029
		SIC	(4, 0.5)	0.337	0.615	0.551	0.026
		MIC	(4, 0.5)	0.838	0.968	0.930	0.037
	15	LRT	(4, 0.5)	0.326	0.815	0.542	0.025
		SIC	(4, 0.5)	0.148	0.665	0.275	0.013
		MIC	(4, 0.5)	0.656	0.932	0.804	0.029
0.1	5	LRT	(4, 0.5)	0.726	0.782	0.831	0.049
		SIC	(4, 0.5)	0.558	0.510	0.673	0.030
		MIC	(4, 0.5)	0.853	0.928	0.912	0.065
	10	LRT	(4, 0.5)	0.813	0.962	0.923	0.051
		SIC	(4, 0.5)	0.583	0.840	0.794	0.037
		MIC	(4, 0.5)	0.912	0.988	0.975	0.065
	15	LRT	(4, 0.5)	0.610	0.893	0.778	0.030
		SIC	(4, 0.5)	0.354	0.835	0.549	0.017
		MIC	(4, 0.5)	0.782	0.970	0.908	0.037

Table 6. Powers of the LRT, SIC, and MIC procedures at $(\gamma_1, \beta_1) = (4, 0.5)$, $n = 50$.

				(γ_n, β_n)			
α	**k**	**Model**	(γ_1, β_1)	**(4, 2.5)**	**(0.2, 0.5)**	**(2, 2)**	**(4, 0.5)**
0.01	15	LRT	(4, 0.5)	0.795	0.869	0.930	0.000
		SIC	(4, 0.5)	0.617	0.672	0.837	0.000
		MIC	(4, 0.5)	0.979	0.998	0.998	0.002
	25	LRT	(4, 0.5)	0.828	0.995	0.955	0.001
		SIC	(4, 0.5)	0.639	0.960	0.904	0.000
		MIC	(4, 0.5)	0.996	1.000	1.000	0.006
	35	LRT	(4, 0.5)	0.530	0.988	0.977	0.002
		SIC	(4, 0.5)	0.300	0.976	0.718	0.000
		MIC	(4, 0.5)	0.964	0.999	0.994	0.002
0.05	15	LRT	(4, 0.5)	0.963	0.996	0.998	0.012
		SIC	(4, 0.5)	0.930	0.983	0.983	0.006
		MIC	(4, 0.5)	0.996	1.000	1.000	0.014
	25	LRT	(4, 0.5)	0.986	1.000	0.999	0.022
		SIC	(4, 0.5)	0.955	0.998	0.995	0.004
		MIC	(4, 0.5)	1.000	1.000	1.000	0.034
	35	LRT	(4, 0.5)	0.936	0.999	0.996	0.010
		SIC	(4, 0.5)	0.810	0.997	0.978	0.009
		MIC	(4, 0.5)	0.998	1.000	1.000	0.011
0.1	15	LRT	(4, 0.5)	0.991	1.000	0.999	0.026
		SIC	(4, 0.5)	0.974	0.996	0.993	0.024
		MIC	(4, 0.5)	0.999	1.000	1.000	0.032
	25	LRT	(4, 0.5)	0.997	1.000	1.000	0.029
		SIC	(4, 0.5)	0.992	1.000	0.999	0.029
		MIC	(4, 0.5)	1.000	1.000	1.000	0.034
	35	LRT	(4, 0.5)	0.989	0.999	0.999	0.017
		SIC	(4, 0.5)	0.931	0.998	0.996	0.011
		MIC	(4, 0.5)	1.000	1.000	1.000	0.028

Table 7. Powers of the LRT, SIC, and MIC procedures at $(\gamma_1, \beta_1) = (4, 0.5)$, $n = 100$.

				(γ_n, β_n)			
α	**k**	**Model**	(γ_1, β_1)	**(4, 2.5)**	**(0.2, 0.5)**	**(2, 2)**	**(4, 0.5)**
0.01	25	LRT	(4, 0.5)	0.995	1.000	0.998	0.002
		SIC	(4, 0.5)	0.989	1.000	0.997	0.001
		MIC	(4, 0.5)	0.999	1.000	1.000	0.008
	50	LRT	(4, 0.5)	1.000	1.000	1.000	0.000
		SIC	(4, 0.5)	0.998	1.000	1.000	0.000
		MIC	(4, 0.5)	1.000	1.000	1.000	0.005
	75	LRT	(4, 0.5)	0.972	1.000	1.000	0.002
		SIC	(4, 0.5)	0.913	1.000	0.996	0.000
		MIC	(4, 0.5)	0.991	1.000	1.000	0.006
0.05	25	LRT	(4, 0.5)	1.000	1.000	1.000	0.033
		SIC	(4, 0.5)	0.999	1.000	1.000	0.006
		MIC	(4, 0.5)	1.000	1.000	1.000	0.046
	50	LRT	(4, 0.5)	1.000	1.000	1.000	0.023
		SIC	(4, 0.5)	1.000	1.000	1.000	0.012
		MIC	(4, 0.5)	1.000	1.000	1.000	0.051
	75	LRT	(4, 0.5)	0.999	1.000	1.000	0.034
		SIC	(4, 0.5)	0.999	1.000	1.000	0.010
		MIC	(4, 0.5)	0.999	1.000	1.000	0.042

Table 7. *Cont.*

α	k	Model	(γ_1, β_1)	(γ_n, β_n) (4, 2.5)	(0.2, 0.5)	(2, 2)	(4, 0.5)
0.1	25	LRT	(4, 0.5)	1.000	1.000	1.000	0.075
		SIC	(4, 0.5)	1.000	1.000	1.000	0.033
		MIC	(4, 0.5)	1.000	1.000	1.000	0.093
	50	LRT	(4, 0.5)	1.000	1.000	1.000	0.072
		SIC	(4, 0.5)	1.000	1.000	1.000	0.040
		MIC	(4, 0.5)	1.000	1.000	1.000	0.099
	75	LRT	(4, 0.5)	1.000	1.000	1.000	0.094
		SIC	(4, 0.5)	1.000	1.000	1.000	0.039
		MIC	(4, 0.5)	1.000	1.000	1.000	0.096

Table 8. Powers of the LRT, SIC, and MIC procedures at $(\gamma_1, \beta_1) = (0.5, 3.5)$, $n = 20$.

α	k	Model	(γ_1, β_1)	(γ_n, β_n) (0.5, 1.5)	(1.2, 3.5)	(0.8, 2.5)	(0.5, 3.5)
0.01	5	LRT	(0.5, 3.5)	0.165	0.388	0.208	0.001
		SIC	(0.5, 3.5)	0.105	0.271	0.162	0.000
		MIC	(0.5, 3.5)	0.423	0.680	0.495	0.003
	10	LRT	(0.5, 3.5)	0.246	0.493	0.327	0.001
		SIC	(0.5, 3.5)	0.162	0.320	0.218	0.000
		MIC	(0.5, 3.5)	0.514	0.812	0.570	0.006
	15	LRT	(0.5, 3.5)	0.201	0.268	0.254	0.001
		SIC	(0.5, 3.5)	0.111	0.151	0.159	0.000
		MIC	(0.5, 3.5)	0.450	0.677	0.477	0.002
0.05	5	LRT	(0.5, 3.5)	0.446	0.694	0.518	0.028
		SIC	(0.5, 3.5)	0.334	0.586	0.411	0.004
		MIC	(0.5, 3.5)	0.879	0.953	0.892	0.036
	10	LRT	(0.5, 3.5)	0.543	0.794	0.616	0.045
		SIC	(0.5, 3.5)	0.394	0.736	0.469	0.017
		MIC	(0.5, 3.5)	0.903	0.987	0.923	0.059
	15	LRT	(0.5, 3.5)	0.476	0.599	0,517	0.024
		SIC	(0.5, 3.5)	0.300	0.542	0.377	0.006
		MIC	(0.5, 3.5)	0.866	0.959	0.898	0.039
0.1	5	LRT	(0.5, 3.5)	0.763	0.875	0.833	0.043
		SIC	(0.5, 3.5)	0.581	0.761	0.643	0.031
		MIC	(0.5, 3.5)	0.989	0.999	0.997	0.056
	10	LRT	(0.5, 3.5)	0.823	0.928	0.889	0.054
		SIC	(0.5, 3.5)	0.628	0.875	0.713	0.048
		MIC	(0.5, 3.5)	0.999	1.000	0.998	0.078
	15	LRT	(0.5, 3.5)	0.781	0.885	0.814	0.049
		SIC	(0.5, 3.5)	0.526	0.747	0.638	0.034
		MIC	(0.5, 3.5)	0.992	0.998	0.998	0.061

Table 9. Powers of the LRT, SIC, and MIC procedures at $(\gamma_1, \beta_1) = (0.5, 3.5)$, $n = 50$.

				(γ_n, β_n)			
α	k	Model	(γ_1, β_1)	(0.5, 1.5)	(1.2, 3.5)	(0.8, 2.5)	(0.5, 3.5)
0.01	15	LRT	(0.5, 3.5)	0.543	0.889	0.735	0.001
		SIC	(0.5, 3.5)	0.326	0.750	0.551	0.000
		MIC	(0.5, 3.5)	0.790	0.987	0.820	0.010
	25	LRT	(0.5, 3.5)	0.639	0.932	0.825	0.004
		SIC	(0.5, 3.5)	0.426	0.841	0.624	0.001
		MIC	(0.5, 3.5)	0.897	0.995	0.901	0.015
	35	LRT	(0.5, 3.5)	0.588	0.805	0.744	0.000
		SIC	(0.5, 3.5)	0.365	0.648	0.547	0.000
		MIC	(0.5, 3.5)	0.832	0.981	0.827	0.007
0.05	15	LRT	(0.5, 3.5)	0.724	0.975	0.844	0.029
		SIC	(0.5, 3.5)	0.608	0.929	0.810	0.010
		MIC	(0.5, 3.5)	0.929	0.999	0.965	0.033
	25	LRT	(0.5, 3.5)	0.803	0.989	0.906	0.033
		SIC	(0.5, 3.5)	0.742	0.956	0.858	0.010
		MIC	(0.5, 3.5)	0.954	1.000	0.988	0.048
	35	LRT	(0.5, 3.5)	0.732	0.966	0,850	0.034
		SIC	(0.5, 3.5)	0.635	0.886	0.789	0.012
		MIC	(0.5, 3.5)	0.926	0.999	0.971	0.048
0.1	15	LRT	(0.5, 3.5)	0.922	0.994	0.968	0.070
		SIC	(0.5, 3.5)	0.720	0.969	0.817	0.031
		MIC	(0.5, 3.5)	0.992	1.000	0.997	0.085
	25	LRT	(0.5, 3.5)	0.950	1.000	0.980	0.079
		SIC	(0.5, 3.5)	0.823	0.986	0.916	0.033
		MIC	(0.5, 3.5)	0.999	1.000	1.000	0.095
	35	LRT	(0.5, 3.5)	0.908	0.996	0.950	0.077
		SIC	(0.5, 3.5)	0.749	0.952	0.868	0.035
		MIC	(0.5, 3.5)	0.988	1.000	0.999	0.090

Table 10. Powers of the LRT, SIC, and MIC procedures at $(\gamma_1, \beta_1) = (0.5, 3.5)$, $n = 100$.

				(γ_n, β_n)			
α	k	Model	(γ_1, β_1)	(0.5, 1.5)	(1.2, 3.5)	(0.8, 2.5)	(0.5, 3.5)
0.01	25	LRT	(0.5, 3.5)	0.773	0.974	0.821	0.002
		SIC	(0.5, 3.5)	0.618	0.932	0.693	0.001
		MIC	(0.5, 3.5)	0.890	1.000	0.945	0.011
	50	LRT	(0.5, 3.5)	0.840	0.996	0.854	0.007
		SIC	(0.5, 3.5)	0.759	0.984	0.763	0.001
		MIC	(0.5, 3.5)	0.981	1.000	0.993	0.019
	75	LRT	(0.5, 3.5)	0.835	0.948	0.839	0.002
		SIC	(0.5, 3.5)	0.625	0.908	0.638	0.000
		MIC	(0.5, 3.5)	0.919	1.000	0.949	0.007
0.05	25	LRT	(0.5, 3.5)	0.809	0.997	0.880	0.032
		SIC	(0.5, 3.5)	0.673	0.993	0.857	0.007
		MIC	(0.5, 3.5)	0.939	1.000	0.994	0.060
	50	LRT	(0.5, 3.5)	0.895	1.000	0.957	0.038
		SIC	(0.5, 3.5)	0.780	1.000	0.911	0.012
		MIC	(0.5, 3.5)	0.992	1.000	0.998	0.074
	75	LRT	(0.5, 3.5)	0.819	0.996	0,870	0.025
		SIC	(0.5, 3.5)	0.674	0.996	0.813	0.007
		MIC	(0.5, 3.5)	0.955	1.000	0.992	0.039

Table 10. *Cont.*

α	k	Model	(γ_1, β_1)	(γ_n, β_n) (0.5, 1.5)	(1.2, 3.5)	(0.8, 2.5)	(0.5, 3.5)
0.1	25	LRT	(0.5, 3.5)	0.943	1.000	0.989	0.076
		SIC	(0.5, 3.5)	0.833	0.996	0.900	0.037
		MIC	(0.5, 3.5)	0.997	1.000	1.000	0.087
	50	LRT	(0.5, 3.5)	0.987	1.000	0.995	0.080
		SIC	(0.5, 3.5)	0.890	1.000	0.963	0.029
		MIC	(0.5, 3.5)	0.999	1.000	1.000	0.090
	75	LRT	(0.5, 3.5)	0.945	1.000	0.992	0.064
		SIC	(0.5, 3.5)	0.847	1.000	0.895	0.034
		MIC	(0.5, 3.5)	0.992	1.000	1.000	0.082

Table 11. Powers of the LRT, SIC, and MIC procedures at $(\gamma_1, \beta_1) = (5, 2)$, $n = 20$.

α	k	Model	(γ_1, β_1)	(γ_n, β_n) (5, 3.5)	(0.5, 2)	(1.5, 4.5)	(5, 2)
0.01	5	LRT	(5, 2)	0.274	0.263	0.295	0.000
		SIC	(5, 2)	0.099	0.114	0.118	0.000
		MIC	(5, 2)	0.483	0.600	0.513	0.007
	10	LRT	(5, 2)	0.381	0.494	0.333	0.002
		SIC	(5, 2)	0.144	0.218	0.119	0.000
		MIC	(5, 2)	0.686	0.749	0.808	0.007
	15	LRT	(5, 2)	0.186	0.358	0.152	0.001
		SIC	(5, 2)	0.069	0.144	0.036	0.000
		MIC	(5, 2)	0.445	0.617	0.611	0.005
0.05	5	LRT	(5, 2)	0.670	0.796	0.848	0.013
		SIC	(5, 2)	0.514	0.599	0.635	0.004
		MIC	(5, 2)	0.749	0.926	0.870	0.050
	10	LRT	(5, 2)	0.828	0.887	0.963	0.038
		SIC	(5, 2)	0.648	0.684	0.864	0.006
		MIC	(5, 2)	0.885	0.974	0.968	0.063
	15	LRT	(5, 2)	0.659	0.694	0,818	0.043
		SIC	(5, 2)	0.319	0.358	0.731	0.005
		MIC	(5, 2)	0.750	0.954	0.804	0.059
0.1	5	LRT	(5, 2)	0.743	0.922	0.884	0.053
		SIC	(5, 2)	0.629	0.784	0.861	0.049
		MIC	(5, 2)	0.823	0.983	0.957	0.074
	10	LRT	(5, 2)	0.782	0.975	0.894	0.062
		SIC	(5, 2)	0.635	0.888	0.881	0.051
		MIC	(5, 2)	0.857	0.998	0.996	0.080
	15	LRT	(5, 2)	0.623	0.871	0.883	0.049
		SIC	(5, 2)	0.569	0.658	0.820	0.043
		MIC	(5, 2)	0.835	0.996	0.968	0.080

Table 12. Powers of the LRT, SIC, and MIC procedures at $(\gamma_1, \beta_1) = (5, 2)$, $n = 50$.

				(γ_n, β_n)			
α	**k**	**Model**	(γ_1, β_1)	**(5, 3.5)**	**(0.5, 2)**	**(1.5, 4.5)**	**(5, 2)**
0.01	15	LRT	(5, 2)	0.512	0.886	0.999	0.001
		SIC	(5, 2)	0.403	0.719	0.998	0.001
		MIC	(5, 2)	0.600	1.000	1.000	0.004
	25	LRT	(5, 2)	0.519	0.954	1.000	0.002
		SIC	(5, 2)	0.415	0.761	1.000	0.001
		MIC	(5, 2)	0.649	1.000	1.000	0.013
	35	LRT	(5, 2)	0.507	0.932	1.000	0.000
		SIC	(5, 2)	0.404	0.754	0.998	0.000
		MIC	(5, 2)	0.612	1.000	1.000	0.005
0.05	15	LRT	(5, 2)	0.832	0.997	1.000	0.028
		SIC	(5, 2)	0.668	0.842	1.000	0.008
		MIC	(5, 2)	0.834	1.000	1.000	0.035
	25	LRT	(5, 2)	0.894	0.999	1.000	0.036
		SIC	(5, 2)	0.670	0.853	1.000	0.011
		MIC	(5, 2)	0.866	1.000	1.000	0.048
	35	LRT	(5, 2)	0.814	0.878	1.000	0.017
		SIC	(5, 2)	0.655	0.796	1.000	0.008
		MIC	(5, 2)	0.848	1.000	1.000	0.027
0.1	15	LRT	(5, 2)	0.909	1.000	1.000	0.064
		SIC	(5, 2)	0.777	0.998	1.000	0.028
		MIC	(5, 2)	0.948	1.000	1.000	0.078
	25	LRT	(5, 2)	0.923	1.000	1.000	0.069
		SIC	(5, 2)	0.796	0.999	1.000	0.038
		MIC	(5, 2)	0.962	1.000	1.000	0.083
	35	LRT	(5, 2)	0.881	1.000	1.000	0.050
		SIC	(5, 2)	0.737	0.979	1.000	0.031
		MIC	(5, 2)	0.958	1.000	1.000	0.081

Table 13. Powers of the LRT, SIC, and MIC procedures at $(\gamma_1, \beta_1) = (5, 2)$, $n = 100$.

				(γ_n, β_n)			
α	**k**	**Model**	(γ_1, β_1)	**(5, 3.5)**	**(0.5, 2)**	**(1.5, 4.5)**	**(5, 2)**
0.01	25	LRT	(5, 2)	0.755	0.917	1.000	0.002
		SIC	(5, 2)	0.628	0.767	1.000	0.000
		MIC	(5, 2)	0.799	1.000	1.000	0.009
	50	LRT	(5, 2)	0.869	0.993	1.000	0.003
		SIC	(5, 2)	0.722	0.869	1.000	0.001
		MIC	(5, 2)	0.890	1.000	1.000	0.013
	75	LRT	(5, 2)	0.832	0.962	1.000	0.002
		SIC	(5, 2)	0.710	0.814	1.000	0.001
		MIC	(5, 2)	0.765	1.000	1.000	0.010
0.05	25	LRT	(5, 2)	0.894	0.999	1.000	0.035
		SIC	(5, 2)	0.775	0.860	1.000	0.008
		MIC	(5, 2)	0.937	1.000	1.000	0.061
	50	LRT	(5, 2)	0.945	0.949	1.000	0.039
		SIC	(5, 2)	0.804	0.921	1.000	0.009
		MIC	(5, 2)	0.979	1.000	1.000	0.071
	75	LRT	(5, 2)	0.848	0.973	1.000	0.025
		SIC	(5, 2)	0.716	0.885	1.000	0.009
		MIC	(5, 2)	0.940	1.000	1.000	0.067

Table 13. *Cont.*

α	k	Model	(γ_1, β_1)	(γ_n, β_n) (5, 3.5)	(0.5, 2)	(1.5, 4.5)	(5, 2)
0.1	25	LRT	(5, 2)	0.979	1.000	1.000	0.065
		SIC	(5, 2)	0.898	0.998	1.000	0.028
		MIC	(5, 2)	0.989	1.000	1.000	0.086
	50	LRT	(5, 2)	0.984	1.000	1.000	0.081
		SIC	(5, 2)	0.918	1.000	1.000	0.037
		MIC	(5, 2)	0.997	1.000	1.000	0.091
	75	LRT	(5, 2)	0.943	1.000	1.000	0.075
		SIC	(5, 2)	0.852	0.999	1.000	0.030
		MIC	(5, 2)	0.992	1.000	1.000	0.087

From the simulation results in Tables 5–13, we observe that the power of the SIC procedure is generally the lowest for all situations, and the power of the MIC procedure is higher than the powers of the procedures based on SIC and LRT. At a small sample size of $n = 20$, the powers of the SIC and LRT procedures are relatively low compared to the MIC procedure; even the power of the MIC is not good enough. We also note that the generated data do not have variable points; in many cases, the rejection rate of the MIC test is greater than the nominal α level, probably because the MIC-based approach takes into account the effect of variable point location on model complexity. Next, we can also observe that as the significance level α and sample size n increase, the powers of the LRT, SIC and MIC procedures increase accordingly. The power values are higher when the change occurs around the middle of the data than the power values when the change occurs near the beginning or the end. Furthermore, we notice that the smaller the difference between (γ_1, β_1) and (γ_n, β_n), the smaller the power. In other words, when the parameter value of the null hypothesis and the alternative hypothesis are closer, the smaller the power is. Moreover, when sample sizes are large enough, the power approaches 1, which indicates that the three criteria are consistent. In the simulation results shown, if the statistics of the three criteria satisfy Pr (reject H_0 when H_0 is false) $\geq$ Pr (reject H_0 when H_0 is true), then the statistics of the three criteria are unbiased. From the comparison, the MIC test is usually anti-conservative and does not respect the nominal α, but it is the most powerful test in H_1 among the settings with good behavior under H_0. Therefore, we conclude that the MIC method has a significant ability to detect change points compared to the LRT and SIC methods.

4. Application

The *Kw* distribution is widely used in hydrology and related fields. Meanwhile, all the methods to detect the change point of the real dataset can be extended to the case where there may be a dependency between the observations, which is also common in the case of time series data, as in the literature, such as Chen and Ning [5] and Tian and Yang [8]. In this section, since the overall effect of the MIC is better, we consider applying the MIC testing procedure to detect possible change points in the following real datasets.

4.1. Shasta Reservoir

The first dataset describes the monthly water capacity from the Shasta reservoir in California, USA. The data are recorded for February from 1991 to 2010 (see for details the website http://cdec.water.ca.gov/reservoir_map.html (accessed on 15 December 2022), which can also be found in Sultana et al. [24]. The parameter estimates and the Kolmogorov–Smirnov (K-S) test correlation results are given in Table 14. The probability density fitting curves for the dataset are also shown in Figure 1, which means the dataset fits the *Kw* distribution reasonably well. In fact, Nadar et al. [13] used this dataset to conduct statistical analyses on the *Kw* distribution based on record data.

Table 14. The MLEs and the goodness-of-fit statistics for the Shasta reservoir dataset.

Model	n	$\hat{\gamma}$	$\hat{\beta}$	K–S (pval)
Kw	20	6.060	4.083	0.221 (0.245)

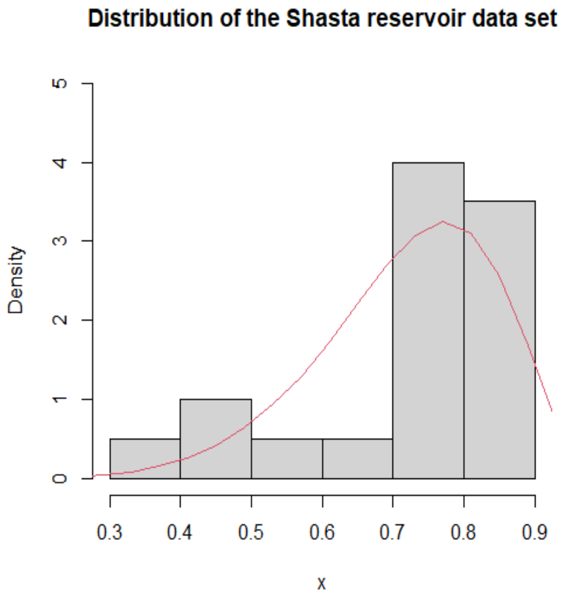

Figure 1. Histogram and PDF fitting of Shasta reservoir dataset.

We applied the MIC test criteria of Equations (5)–(7). Under the null hypothesis H_0, the $MIC(n)$ is calculated as -20.958. Under the alternative hypothesis H_1, $\min_{2 \leq k \leq 19} MIC(k)$ is calculated as -31.156, which corresponds to $k = 3$. The corresponding estimated values of the parameters are $\hat{\gamma}_1 = 4.131$, $\hat{\beta}_1 = 6.074$, $\hat{\gamma}_n = 10.801$, $\hat{\beta}_n = 9.253$ and $S_n = 16.189$ with $p\_value = 0.004$ when using Equations (8) and (9). Since p_value is less than 0.05, there is a change point occurring at position 3, which corresponds to the year 1993. According to Yates et al. [25], 1993 corresponded to a wet year in the Shasta reservoir, California. Figure 2 shows the dataset of monthly water capacity for the Shasta reservoir and the position of the change point.

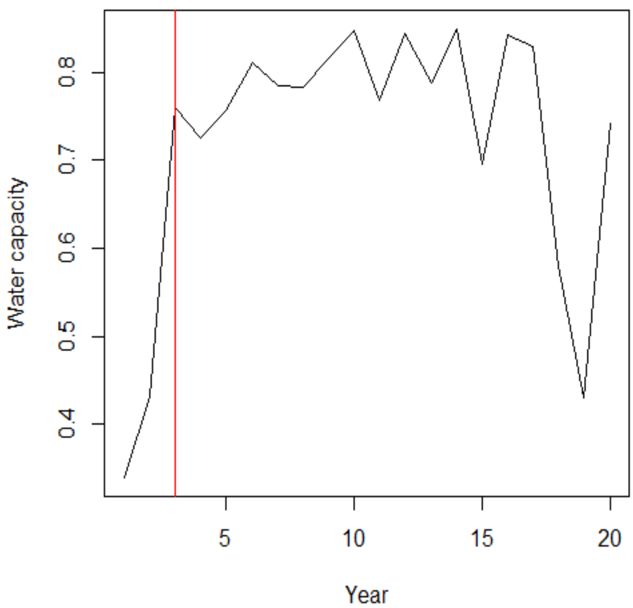

Figure 2. The Shasta reservoir dataset and position of change point.

Figure 3 shows the MIC values associated with different values of k. The estimated change location corresponds to the smallest MIC value.

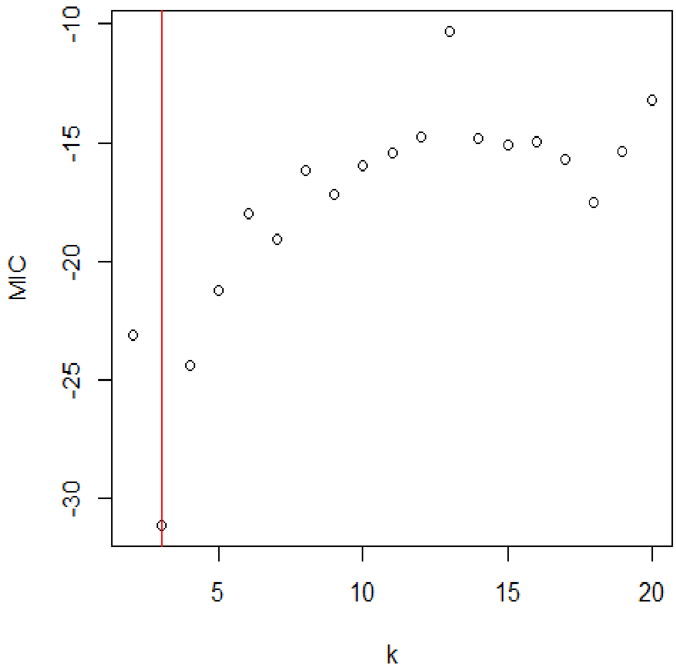

Figure 3. The distribution of MIC for the Shasta reservoir.

4.2. Susquehanna River

The second dataset describes the maximum flood level (in millions of cubic feet per second) for the Susquehanna River at Harrisburg, Pennsylvania, from 1890 to 1969. Each number is the maximum flood level for four years. Khan et al. [26] investigated these data with the Kw distribution and also considered fitting the flood data with the Kw distribution. Mazucheli et al. [27] used this dataset to verify the practicability of the unit Weibull distribution. Bantan et al. [28] applied the improved Kw model to this dataset, demonstrating the superiority of the distribution. Furthermore, the parameter estimates and the Kolmogorov–Smirnov (K-S) test correlation results are given in Table 15. The probability density fitting curves for the dataset are also shown in Figure 4.

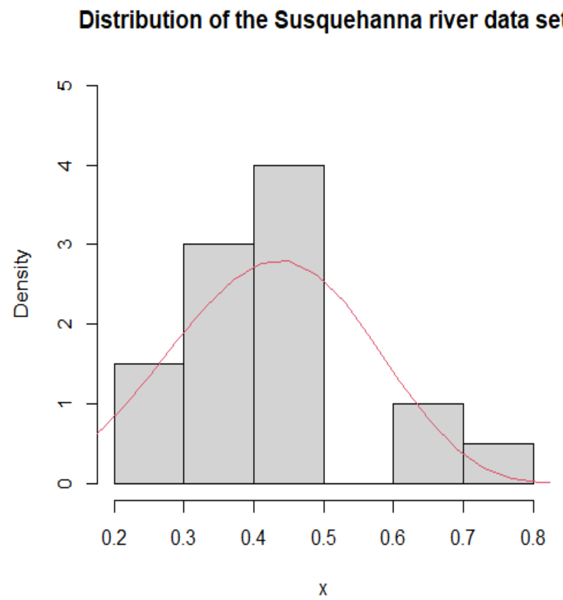

Figure 4. Histogram and PDF fitting of Susquehanna river dataset.

Table 15. The MLEs and the goodness-of-fit statistics for Susquehanna river dataset.

Model	n	$\hat{\gamma}$	$\hat{\beta}$	K–S (pval)
Kw	20	3.353	11.658	0.213 (0.284)

In order to detect the change point in the dataset, we obtained, under the null hypothesis H_0, the $MIC(n)$, which was calculated as -19.741. Under the alternative hypothesis H_1, $\min_{2\leq k\leq 19} MIC(k)$ was calculated as -34.055 which corresponds to $k = 13$. The corresponding parameters are $\hat{\gamma}_1 = 14.116$, $\hat{\beta}_1 = 3.444$, $\hat{\gamma}_n = 8.504$, $\hat{\beta}_n = 5.992$ and $S_n = 20.306$ with $p\_value = 0.018$. Since p_value is less than 0.05, we can say that the data have a change point, and the position of the change point is 13, which corresponds to the period 1934–1937. According to Roland et al. [29], a serious flood occurred in 1936. Figure 5 shows the dataset of the maximum flood level for the Susquehanna River and the position of the change point.

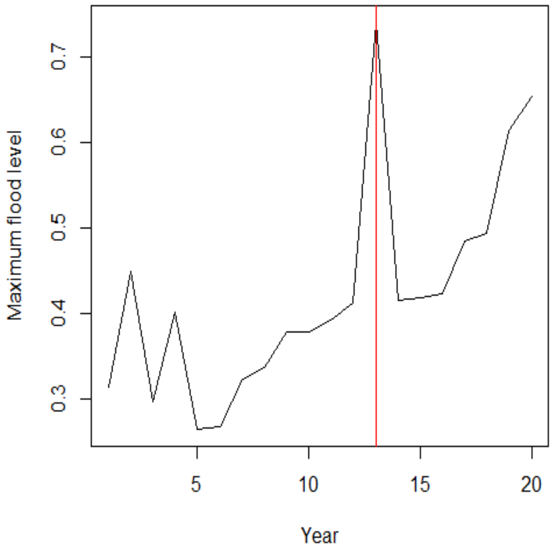

Figure 5. The Susquehanna river dataset and position of change point.

Figure 6 shows the MIC values for all possible values of k. The smallest value of the MIC corresponds to the estimated change location.

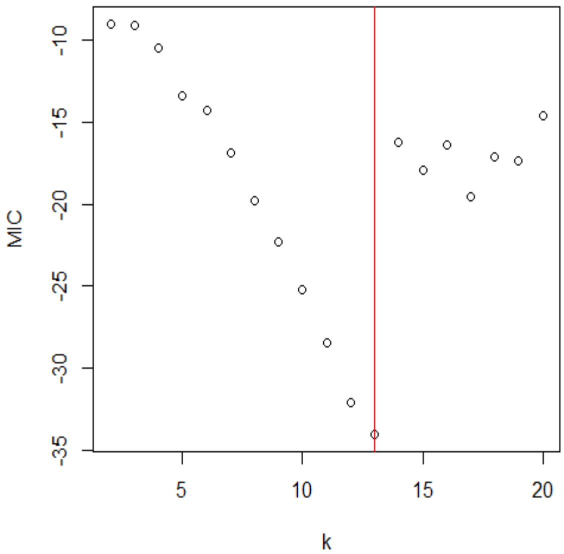

Figure 6. The distribution of MIC for the Susquehanna river.

4.3. Strengths of 1.5 cm Glass Fibres

The third dataset represents the strengths of 1.5 cm glass fibres, initially obtained by workers at the UK National Physical Laboratory. Glass fiber is used to make a variety of products. It is a good electrical insulator; therefore, it is used in the manufacture of many electrical and electronic products and circuit boards. It is also a heat-resistant material used to make products that heat up quickly, such as batteries and motors. The observations of the dataset are found in Elgarhy [30]. The parameter estimates and the Kolmogorov–Smirnov (K–S) test correlation results are given in Table 16. The probability density fitting curves for the dataset are also shown in Figure 7. Thus, the dataset fit the Kw distribution reasonably well.

Table 16. The MLEs and the goodness-of-fit statistics for 1.5cm glass fibre strengths dataset.

Model	n	$\hat{\gamma}$	$\hat{\beta}$	**K–S (pval)**
Kw	27	1.383	6.461	0.240 (0.074)

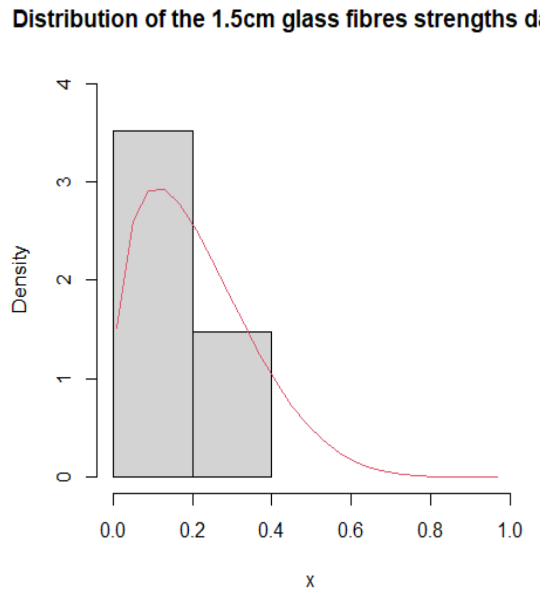

Figure 7. Histogram and PDF fitting of 1.5 cm glass fibre strengths dataset.

Under the null hypothesis H_0, the $MIC(n)$ was calculated as -24.539. Under the alternative hypothesis H_1, $\min_{2\leq k\leq 26} MIC(k)$ was calculated as -30.775, which corresponds to $k = 20$. The corresponding estimated value of the parameters are $\hat{\gamma}_1 = 2.414$, $\hat{\beta}_1 = 0.829$, $\hat{\gamma}_n = 2.844$, $\hat{\beta}_n = 1.404$, and $S_n = 12.828$, with $p\_value = 0.001$. Since p_value is less than 0.05, that is to say, there is a change point occurring at position 20, it can be seen that a change point is indicated at the strength of the twentieth, corresponding to the dataset of 0.10. This shows that for the dataset of twenty-seven strengths, the strength of the glass fibers changes at the twentieth strength, and this change remains until the twenty-seventh strength. Figure 8 shows the original data and change point position.

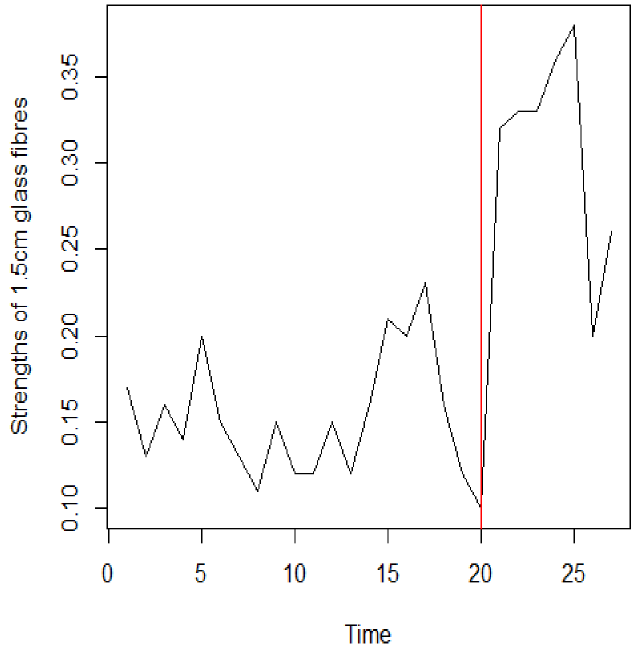

Figure 8. The strengths of 1.5cm glass fibres dataset and position of change point.

Figure 9 shows the values of the MIC associated with different values of k. The estimated change location corresponds to the smallest MIC value.

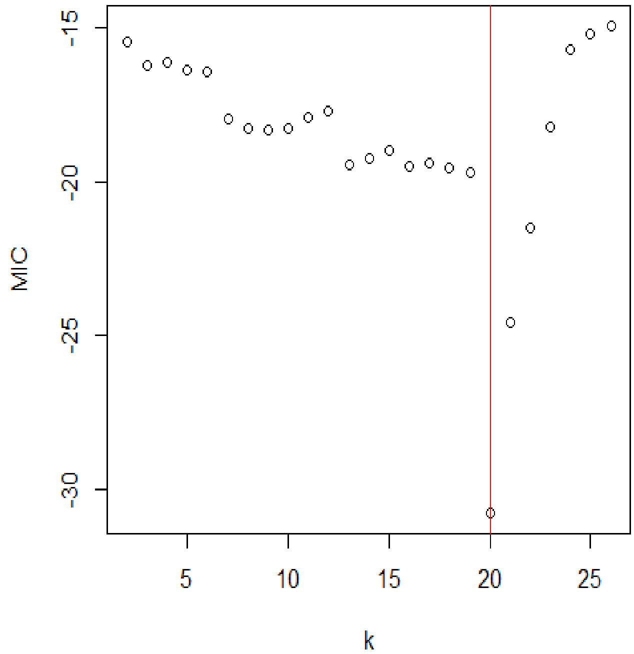

Figure 9. The distribution of the MIC for the strengths of 1.5 cm glass fibres.

5. Conclusions

In this paper, we use the LRT, SIC and MIC methods to perform a change point analysis of the *Kw* distribution, which is widely used in hydrology. Simulations are performed under different scenarios as a means to elucidate the performance of the three change point detection methods. The simulation results show that, in general, the MIC method has more advantages than the SIC and LRT methods in detecting the position of change points. Finally, the MIC method is used to detect the change point of real datasets, and significant change points can be detected. Although the MIC method can work well for the change point detection based on the Kw distribution, the power of it is not big enough for small sizes, and we will work on an alternative method to improve it.

Author Contributions: W.T.: Conceptualization, Methodology, Validation, Investigation, Resources, Supervision, Project Administration, Visualization, Writing—review and editing; L.P.: Software, Formal analysis, Data curation, Writing—original draft preparation, Visualization; C.T., W.N.: Software, Methodology, Visualization. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The source of the datasets are provided in the paper.

Acknowledgments: We would like to thank the reviewers for carefully and thoroughly reading this manuscript and for the thoughtful comments and constructive suggestions, which helped us improve the quality of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Theorem A1 (Csörgó and Horváth [21])**.** *If* $0 < t_1(n) < t_2(n) < 1$ *and*

$$u(n) = \frac{1 - t_1(n)t_2(n)}{t_1(n)(1 - t_2(n))} \to \infty, \text{ as } n \to \infty;$$

then we have

$$\lim_{n\to\infty} P(A(\log u(n)) \sup_{t_1(n)\le t\le t_2(n)} M_r(t) \le x + D_r(\log u(n))) = \exp(-e^{-x}),$$

for all x.

Corollary A1 (Csörgó and Horváth [21])**.** *We have for all* $0 < \lambda < \infty$

$$\lim_{n\to\infty} P(A(\log n) \sup_{\lambda/n\le t<1-\lambda/n} M_r(t) \le x + D_r(\log n)) = \exp(-2e^{-x}), \quad -\infty < x < \infty.$$

Theorem A2 (Csörgó and Horváth [21])**.** *If* H_0 *and* $C1 - C9$ *hold; then we have*

$$\lim_{n\to\infty} P\left\{ A(\log n) Z_n^{\frac{1}{2}} \le t + D_d(\log n) \right\} = \exp(-2e^{-t}),$$

for all t.

References

1. Page, E.S. Continuous inspection schemes. *Biometrika* **1954**, *41*, 100–115. [CrossRef]
2. Page, E.S. A test for a change in a parameter occurring at an unknown point. *Biometrika* **1955**, *42*, 523–527. [CrossRef]
3. Sen, A.; Srivastava, M.S. On tests for detecting change in mean. *Ann. Stat.* **1975**, *3*, 98–108. [CrossRef]
4. Cai, X.; Said, K.K.; Ning, W. Change-point analysis with bathtub shape for the exponential distribution. *J. Appl. Stat.* **2016**, *43*, 2740–2750. [CrossRef]
5. Chen, Y.J.; Ning, W. Information approach for a lifetime change-point model based on the exponential-logarithmic distribution. *Commun. Stat.-Simul. Comput.* **2019**, *48*, 1996–2003. [CrossRef]
6. Said, K.K.; Ning, W.; Tian, Y. Modified information criterion for testing changes in skew normal model. *Braz. J. Probab. Stat.* **2019**, *33*, 280–300. [CrossRef]
7. Wang, T.; Tian, W.; Ning, W. Likelihood ratio test change-point detection in the skew slash distribution. *Commun. Stat.-Simul. Comput.* **2020**, 1–13. [CrossRef]
8. Tian, W.; Yang, Y. Change point analysis for weighted exponential distribution. *Commun. Stat.-Simul. Comput.* **2021**, 1–13. [CrossRef]

9. Kumaraswamy, P. A generalized probability density function for double-bounded random processes. *J. Hydrol.* **1980**, *46*, 79–88. [CrossRef]
10. Jones, M.C. Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Stat. Methodol.* **2009**, *6*, 70–81. [CrossRef]
11. Fletcher, S.G.; Ponnambalam, K. Estimation of reservoir yield and storage distribution using moments analysis. *J. Hydrol.* **1996**, *182*, 259–275. [CrossRef]
12. Nadarajah, S. On the distribution of kumaraswamy. *J. Hydrol.* **2008**, *348*, 568–569. [CrossRef]
13. Nadar, M.; Papadopoulos, A.; Kızılaslan, F. Statistical analysis for Kumaraswamy's distribution based on record data. *Stat. Pap.* **2013**, *54*, 355–369. [CrossRef]
14. Saulo, H.; Leão, J.; Bourguignon, M. The kumaraswamy birnbaum-saunders distribution. *J. Stat. Theory Pract.* **2012**, *6*, 745–759. [CrossRef]
15. Lemonte, A.J.; Barreto-Souza, W.; Cordeiro, G.M. The exponentiated Kumaraswamy distribution and its log-transform. *Braz. J. Probab. Stat.* **2013**, *27*, 31–53. [CrossRef]
16. Mameli, V. The Kumaraswamy skew-normal distribution. *Stat. Probab. Lett.* **2015**, *104*, 75–81. [CrossRef]
17. Iqbal, Z.; Tahir, M.M.; Riaz, N.; Ali, S.A.; Ahmad, M. Generalized inverted Kumaraswamy distribution: Properties and application. *Open J. Stat.* **2017**, *7*, 645. [CrossRef]
18. Said, K.K.; Ning, W.; Tian, Y. Likelihood procedure for testing changes in skew normal model with applications to stock returns. *Commun. Stat.-Simul. Comput.* **2017**, *46*, 6790–6802. [CrossRef]
19. Zou, C.; Liu, Y.; Qin, P.; Wang, Z. Empirical likelihood ratio test for the change-point problem. *Stat. Probab. Lett.* **2007**, *77*, 374–382. [CrossRef]
20. Liu, Z.; Qian, L. Changepoint estimation in a segmented linear regression via empirical likelihood. *Commun. Stat.-Simul. Comput.* **2009**, *39*, 85–100. [CrossRef]
21. Csörgó, M.; Horváth, L. *Limit Theorems in Change-Point Analysis*; John Wiley and Sons Inc.: Hoboken, NJ, USA, 1997; Volume 18.
22. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
23. Chen, J.; Gupta, A.K.; Pan, J. Information criterion and change point problem for regular models. *Sankhya Indian J. Stat.* **2006**, *68*, 252–282.
24. Sultana, F.; Tripathi, Y.M.; Wu, S.J.; Sen, T. Inference for kumaraswamy distribution based on type I progressive hybrid censoring. *Ann. Data Sci.* **2022**, *9*, 1283–1307. [CrossRef]
25. Yates, D.; Galbraith, H.; Purkey, D.; Huber-Lee, A.; Sieber, J.; West, J.; Herrod-Julius, S.; Joyce, B. Climate warming, water storage, and Chinook salmon in California's Sacramento Valley. *Clim. Chang.* **2008**, *91*, 335–350. [CrossRef]
26. Khan, M.S.; King, R.; Hudson, I.L. Transmuted kumaraswamy distribution. *Stat. Transit. New Ser.* **2016**, *17*, 183–210. [CrossRef]
27. Mazucheli, J.; Menezes, A.F.B.; Ghitany, M.E. The unit-Weibull distribution and associated inference. *J. Appl. Probab. Stat.* **2018**, *13*, 1–22.
28. Bantan, R.A.; Chesneau, C.; Jamal, F.; Elgarhy, M.; Almutiry, W.; Alahmadi, A.A. Study of a Modified Kumaraswamy Distribution. *Mathematics* **2021**, *9*, 2836. [CrossRef]
29. Roland, M.A.; Underwood, S.M.; Thomas, C.M.; Miller, J.F.; Pratt, B.A.; Hogan, L.G.; Wnek, P.A. Flood-inundation maps for the Susquehanna River near Harrisburg, Pennsylvania, 2013. In *Scientific Investigations Report 2014–5046*; US Geological Survey: Reston, VA, USA, 2014; Volume 28.
30. Elgarhy, M. Exponentiated generalized Kumaraswamy distribution with applications. *Ann. Data Sci.* **2018**, *5*, 273–292. [CrossRef]