

Research Article

A Weighted Error Distance Metrics (WEDM) for Performance Evaluation on Multiple Change-Point (MCP) Detection in Synthetic Time Series

Jin Peng. Qi ¹, Fang. Pu,² Ying. Zhu,³ and Ping. Zhang⁴

¹College of Information Science & Technology, Donghua University, Shanghai 201620, China

²Informationization Office, Donghua University, Shanghai 201620, China

³HNSW Health Pathology East Genetics Level 4, Campus Centre, Prince of Wales Hospital Randwick, Randwick, NSW, Australia

⁴Menzies Health Institute Queensland, Griffith University, Southport, Queensland, Australia

Correspondence should be addressed to Jin Peng. Qi; qipengkai@126.com

Received 27 October 2021; Revised 16 December 2021; Accepted 22 January 2022; Published 24 March 2022

Academic Editor: Radu-Emil Precup

Copyright © 2022 Jin Peng. Qi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Change-point detection (CPD) is to find abrupt changes in time-series data. Various computational algorithms have been developed for CPD applications. To compare the different CPD models, many performance metrics have been introduced to evaluate the algorithms. Each of the previous evaluation methods measures the different aspects of the methods. Based on the existing weighted error distance (WED) method on single change-point (CP) detection, a novel WED metrics (WEDM) was proposed to evaluate the overall performance of a CPD model across not only repetitive tests on single CP detection, but also successive tests on multiple change-point (MCP) detection on synthetic time series under the random slide window (RSW) and fixed slide window (FSW) frameworks. In the proposed WEDM method, a concept of normalized error distance was introduced that allows comparisons of the distance between the estimated change-point (eCP) position and the target change point (tCP) in the synthetic time series. In the successive MCPs detection, the proposed WEDM method first divides the original time-series sample into a series of data segments in terms of the assigned tCPs set and then calculates a normalized error distance (NED) value for each segment. Next, our WEDM presents the frequency and WED distribution of the resultant eCPs from all data segments in the normalized positive-error distance (NPED) and the normalized negative-error distance (NNED) intervals in the same coordinates. Last, the mean WED (MWED) and MWTD (1-MWED) were obtained and then dealt with as important performance evaluation indexes. Based on the synthetic datasets in the Matlab platform, repetitive tests on single CP detection were executed by using different CPD models, including ternary search tree (TST), binary search tree (BST), Kolmogorov–Smirnov (KS) tests, t -tests (T), and singular spectrum analysis (SSA) algorithms. Meanwhile, successive tests on MCPs detection were implemented under the fixed slide window (FSW) and random slide window (RSW) frameworks. These CPD models mentioned above were evaluated in terms of our WED metrics, together with supplementary indexes for evaluating the convergence of different CPD models, including rates of hit, miss, error, and computing time, respectively. The experimental results showed the value of this WEDM method.

1. Introduction

Change-point (CP) detection is the application of core techniques to detect abrupt changes in properties of time-series data. It has been widely studied in many real-world problems, such as atmospheric and financial analyses [1], fault detection in engineering systems [2, 3], changes

detection in a variance of oceanographic time series [4], genetic time-series analyses [5], and online detection of steady-state operation [6]. For example, the usage of this method to detect abnormal patterns in ECG and EEG signals may also be beneficial [4, 7–15]. This application would allow appropriate staff to be alerted of abrupt changes in a patient's medical situation and to provide on-time treatment [16, 17].

In addition, CPD models can be tightly combined with some nonlinear modeling approaches and their applications, such as classification of human hand movements [18], degradation signal for prognostic improvement [19], real-life hand prosthetic control [20], single-channel surface electromyography (sEMG)-based control [21]. CPD models utilize algorithms that cover the fields of data mining, statistics, and computer science, including parametric and nonparametric methods [8, 22–27]. Each CPD algorithm can be assessed from the aspect of detection accuracy, computational cost, or whether it can be a real-time detection.

Many performance metrics have been introduced to evaluate CPD algorithms based on the type of decisions they make [28]. Aminikhanghahi and Cook [29] reviewed the performance evaluation methods commonly used for CPD models. The evaluation can be based on a yes/no decision whether the resultant change point was detected within a certain distance from the actual change point. In this case, the CPD model can be treated as a binary classification model and can be evaluated with the usual measures, such as accuracy, sensitivity, specificity, or ROC curve [30, 31]. For real applications, for example, clinical decision-making, cut-offs applied to the model outcomes can be adjusted to achieve different sensitivity and specificity [32]. However, when the difference in time between the resultant eCP and the actual tCP represents the measure of CPD performance, then the evaluation of these algorithms is not as straightforward as for the binary classification. There is no single label against which the performance of the algorithm can be measured. A few useful metrics consider the distance between the eCP and the tCP to measure CPD method performance. These metrics include mean absolute error (MAE), mean squared error (MSE), mean signed difference (MSD), root mean squared error (RMSE), and normalized root mean squared error (NRMSE). Of these, except NRMSE normalizes the unit size of the predicted value and facilitates a more direct comparison of error between different datasets, the other methods measure only the absolute distances between the eCP and the tCP. However, even NRMSE does not count the difference between the situations when the eCP is before and after the actual tCP. It also fails to consider the relative position of the tCP within the total length of the time-series sample.

In our previous studies [33], a preliminary WED method was proposed for evaluating a CPD model for single change-point detection. In this existing method, a concept of weighted error distance (WED) is introduced for counting a normalized error distance between each pair of the resultant eCPs and the actual tCPs, and then the performance of different CPD models is ranked by the averaged WED accordingly [33]. In this study, a novel WEDM method is proposed to compare the overall performance of CPD models for MCPs detection on multiple data segments in a time series with different data features. Based on the previous WED measure, a concept of normalized error distance was introduced in this WEDM method, that allows comparisons of the distance between the estimated change-point (eCP) position and the target change point (tCP). During the successive MCPs detection, the proposed WEDM method

first divides the original sample into a series of data segments in terms of assigned tCPs, and then counts a normalized error distance (NED) value for each segment. Then, our WEDM presents the frequency and WED distribution of the resultant eCPs from all data segments in the normalized positive-error distance (NPED) and the normalized negative-error distance (NNED) intervals in the same coordinates. Last, the mean WED (MWED) and MWTD (1-MWED) were calculated and dealt with as important performance indexes. Based on the synthetic datasets in the Matlab platform, both repetitive tests on single CP detection and successive test on MCPs detection were executed by using different CPD models, including ternary search tree (TST) [8, 34], binary search tree (BST) [15, 24], Kolmogorov–Smirnov (KS) tests [22, 25], t -tests (T) [23, 35], and singular spectrum analysis (SSA) algorithms [36] recorded in our previous studies [22, 37]. Meanwhile, these CPD models above were evaluated under the random slide window (RSW) [8, 38, 39] and fixed slide window (FSW) frameworks [40–44] in terms of our WEDM and supplementary indexes including the rates of hit, miss, error, and computing time, respectively. The experimental results showed the value of this WEDM method.

2. Methods

In this part, the proposed WEDM is theoretically illuminated in the following steps. First, the diagnosed sample is divided into a series of data segments according to the assigned target MCPs. Second, a normalized error distance (NED) is calculated by comparing the distance between the resultant eCP position and the actual tCP within each data segment. Third, the frequency and WED distribution of the resultant eCPs detected from all segments are presented across the normalized positive-error distance (NPED) and the normalized negative-error distance (NNED) intervals in the same coordinates. Last, the metrics of mean WED (MWED) and mean WTD (MWTD) are given to efficiently evaluate a CPD model for MCPs detection on a series of data fluctuations in an identical time series.

2.1. Data Segmentation. Suppose a time-series signal $X = \{X_1, \dots, X_i, \dots, X_N\}$ can be observed as a trajectory of a multiple data distribution process, in which the segment X_i is defined by the following equation:

$$X_i = f_i(t) + \varepsilon_i, \quad (1)$$

where $t \in \{t_{i-1}+1, \dots, t_i\}$, $0 < i \leq M$, and $f_i \in \{f_1, \dots, f_M\}$ is a deterministic and piece-wise function of one-dimensional signal with change points (satisfying $f_i \neq f_{i+1}$, and $i = 1, \dots, M-1$ for insuring that abrupt changes occur), and $M \in \{1, 2, \dots, n\}$ is the number of data segment regimes and therefore $M-1$ is the number of abrupt changes, $0 = t_0 < t_1 < \dots < t_1 < \dots < t_M = n$. The number $M-1$ and locations $\eta_1, \dots, \eta_{M-1}$ of change points in the process are supposed to be unknown. The sequence $(\varepsilon_i)_{i \in N}$ is assumed to be random white noise and such that $E(\varepsilon_i)$ is exactly or approximately zero. In the

simplest case, $(\varepsilon_i)_{i \in N}$ is modeled as *i.i.d.*, but can also follow more complex time-series distributions.

Consider an observed time-series signal $X = \{X_1, \dots, X_i, \dots, X_N\}$ with $M-1$ change points mentioned above, one-part time series $X' = \{X_s, \dots, X_j, \dots, X_e\}$ with a size of N' is selected from X , $1 \leq s < j < e \leq N$, and $1 < N' \leq N$. Suppose a set of target MCPs $tMCP$ set = $\{tCP_1, \dots, tCP_n\}$ is contained within X' , and $1 \leq n \leq M-1$. In the proposed WEDM method, the diagnosed data sample X' is first divided into a series of data segments according to different target CP positions in the $tMCP$ set. The process of data segmentation is described below (Figure 1):

- (1) For each tCP_i to be diagnosed in the $tMCP$ set, the data segment Seg_i can be denoted as follows:

$$Seg_i = \{mtCP_{i-1}, \dots, tCP_i, \dots, mtCP_i\}, \quad (2)$$

where $1 < i < n$ and $1 < n < N'$, and two endpoints mCP_{i-1} and mCP_i in Seg_i are formulated as follows:

$$mtCP_{i-1} = \frac{(tCP_{i-1} + tCP_i)}{2} \text{ and } mtCP_i = \frac{(tCP_i + tCP_{i+1})}{2}. \quad (3)$$

- (2) Especially, the first Seg_1 and the last Seg_n can be presented according to the tCP_1 and tCP_n as follows:

$$Seg_1 = \{X_s, \dots, tCP_1\} \text{ and } Seg_n = \{tCP_n, \dots, X_e\}, \quad (4)$$

where X_s and X_e are the two endpoints in X' , respectively.

- (3) Then, the time series $X' = \{X_s, \dots, X_j, \dots, X_e\}$ can be divided into a set of data segments SEG set $_{X'} = \{Seg_1, \dots, Seg_n\}$. That is, $X' = \{Seg_1, \dots, Seg_n\}$, and the following equation holds

$$N_{X'} = \sum_{i=1}^n N_Seg_i, \quad (5)$$

where $N_{X'}$ is the total length of X' , and N_Seg_i refers to the size of Seg_i .

2.2. NED Evaluation on Single CP Detection. In the scheme of error distance (ED) measurement on single CP detection (Figure 2), each segment $Seg_i = \{X_a \dots X_c \dots X_b\}$ in time series $X' = \{Seg_1, \dots, Seg_n\}$ is divided into the former (left) part $\{X_a, \dots, X_{c-1}\}$ and the latter (right) part $\{X_{c+1}, \dots, X_b\}$ by the actual tCP_i located at the data point X_c and $1 \leq i \leq n$.

From a statistical point of view, we refer to the former (left) part as a positive area and the latter (right) part as a negative one. When applying a CPD to detect the actual tCP_i in the data segment Seg_i , a resultant eCP_i might be estimated from either the positive area or the negative one. A few concepts are introduced here to measure CPD model performance: true-positive distance (tPD), positive-error distance (pED), true-negative distance (tND), and negative-

error distance (nED). If the resultant eCP_i is detected on the left side of the tCP_i (positive area), then pED_i and tPD_i can be calculated. That is, the distance from the eCP_i to the tCP_i and the start point, respectively. Meanwhile, nED_i and tND_i are not applicable. Conversely, when the eCP_j is estimated from the right side of the tCP_i (negative area), nED_i equals the distance from eCP_j to tCP_i , and tND_i is the distance from the eCP_j to the end of the data segment Seg_i . At the same time, pED_i and tPD_i do not exist (Figure 2). These definitions can be represented in formulas (6)–(9) as follows:

$$\begin{aligned} tPD_i &= eCP_i - mCP_i \\ &= X_d - X_a, \end{aligned} \quad (6)$$

$$\begin{aligned} pED_i &= tCP_i - eCP_i \\ &= X_c - X_d, \end{aligned} \quad (7)$$

$$\begin{aligned} tND_i &= mCP_{i+1} - eCP_j \\ &= X_b - X_e, \end{aligned} \quad (8)$$

$$\begin{aligned} nED_i &= eCP_j - tCP_i \\ &= X_e - X_c. \end{aligned} \quad (9)$$

In which, X_a and X_b represent the start and endpoints of the time-series segment Seg_i , respectively, X_c is the position of actual tCP_i in the Seg_i , X_d and X_e refer to the positions of resultant eCP on the left or right side of the tCP_i respectively.

Basically, for a current data segment Seg_i in the scheme of NED evaluation on single CP detection (Figure 3), the distance between the start point and the tCP_i and the distance from the tCP_i to the end of each segment are both normalized to 1, and the normalized tCP position for each segment will match to the same point. In formulas (10)–(13), $tPDR_i$, $pEDR_i$, $tNDR_i$, and $nEDR_i$ can be interpreted as the normalized true-positive distance ($NtPD_i$), normalized positive-error distance ($NpED_i$), normalized true-negative distance ($NtND_i$), and normalized negative-error distance ($NnED_i$), respectively.

$$tPDR_i(NtPD_i) = \frac{tPD_i}{tPD_i + pED_i}, \quad (10)$$

$$pEDR_i(NpED_i) = \frac{pED_i}{tPD_i + pED_i}, \quad (11)$$

$$tNDR_i(NtND_i) = \frac{tND_i}{tND_i + nED_i}, \quad (12)$$

$$nEDR_i(NnED_i) = \frac{nED_i}{tND_i + nED_i}. \quad (13)$$

Thereafter, a normalized error distance NED^i in formula (14) is presented by a piecewise function of $NpED_i$ and $NnED_i$, according to the resultant eCP_i located at the positive or negative area.

$$NED^i = \begin{cases} NpED_i, & e - CP_i \text{ on positive area,} \\ NnED_i, & e - CP_i \text{ on negative area.} \end{cases} \quad (14)$$

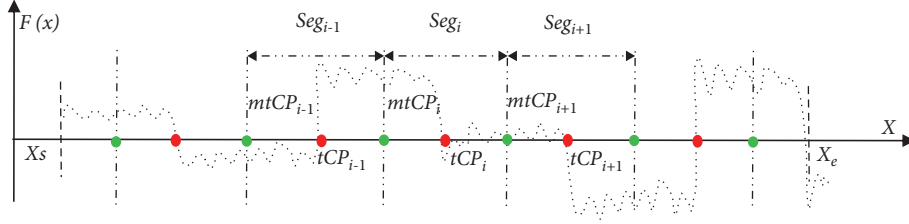


FIGURE 1: The scheme of WEDM evaluation on the target MCPs detection in the diagnosed X' .

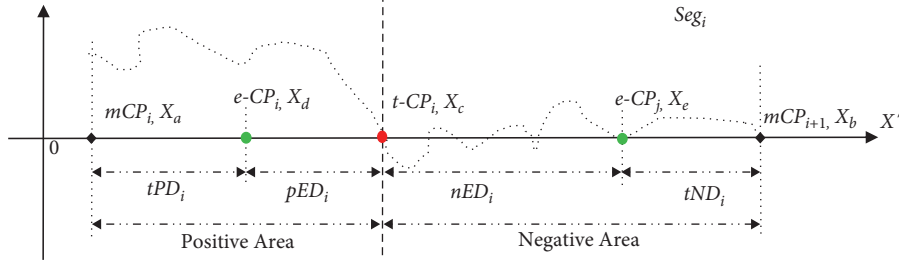


FIGURE 2: The scheme of error distance (ED) measurement on single CP detection in the data segment Seg_i . In the positive area, X_a represents the start point of Seg_i , and X_d is the position of resultant eCP_i within the positive area before the actual tCP_i . On the other hand, X_b represents the endpoint of Seg_i , and X_e stands for the eCP_j located within the negative area after the tCP_i .

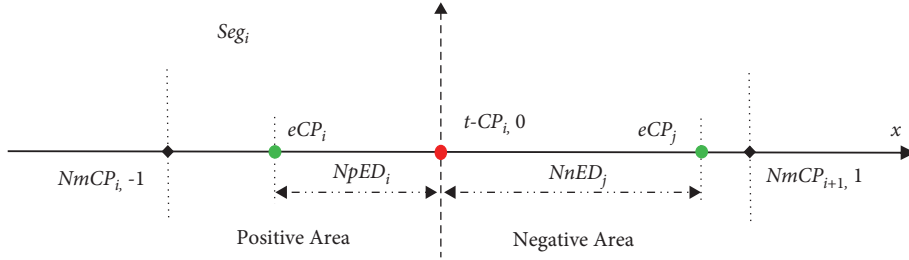


FIGURE 3: The scheme of NED evaluation on single CP detection in the data segment Seg_i . In which, “-1” and “1” represent the start and endpoints of Seg_i , and “0” refers to the position of actual tCP_i in the x -axis, respectively.

2.2.1. WED Evaluation on MCPs Detection. Given a series of data segments $SEG_{set_{X'}} = \{Seg_1, \dots, Seg_n\}$ in a diagnosed time series X' above, we can assemble all the resultant eCPs into an identical coordinate and present their NED values ranging from the positive area $[-1, 0]$ to the negative area $[0, 1]$ in the x -axis (Figure 4). Then, the frequencies of NED^i can be defined in the all resultant eCPs as follows:

$$Freq(NED^i) = \frac{Num(NED^i)}{Nt} \quad (15)$$

In which, $Num(NED^i)$ is the number of the resultant eCPs that their NED values equal to NED^i , and Nt is the number of resultant eCPs in total, $1 \leq i \leq Nt$.

Then, the weighted error distance WED^i is introduced according to the NED^i and $Freq(NED^i)$ in the resultant eCPs (Figure 5). For each eCP_i in the scattered distribution of resultant eCPs, its corresponding WED^i is equal to $WpED_i$ or $WnED_i$ depending on whether the NED^i is located at the positive-NpED or negative-NnED area ranging from -1 to 1 in the x -axis. The definitions of $WpED_i$, $WnED_i$, and WED^i are formulated as follows:

$$\begin{aligned} WpED_i &= Freq(NpED_i) * NpED_i, \\ WnED_i &= Freq(NnED_i) * NnED_i, \\ WED^i &= \begin{cases} WpED_i, & NED^i \text{ on } NpED \text{ area,} \\ WnED_i, & NED^i \text{ on } NnED \text{ area.} \end{cases} \end{aligned} \quad (16)$$

Thereafter, a mean weighted error distance ($MWED$) is defined as follows:

$$MWED = \frac{\sum_{i=1}^l WpED_i + \sum_{j=1}^r WnED_j}{l + r}, \quad (17)$$

where l and r refer to the numbers of the eCPs located before and after the actual tCPs (positive-NpED area and negative-NnED area), respectively. In most of the CPD models, when the search algorithm reaches the start or end of the time series, if no change point is found, then the resultant eCP can be set as either the start or the end. Therefore, the sum of l and r will be equal to N (the total number of actual tCPs to be diagnosed in a time series X'). Formula (17) can be simplified as follows:

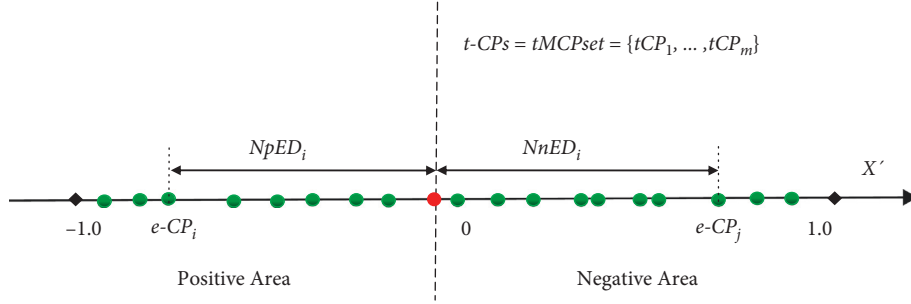


FIGURE 4: The scheme of NED evaluation for MCPs detection on $tMCPset = \{tCP_1, \dots, tCP_m\}$ in a time series X' . For each resultant eCP_i , the value of NED^i equals to $NpED_i$ or $NnED_i$ depending on that the eCP_i is located at the positive or negative area ranging from -1 to 1 in the x -axis.

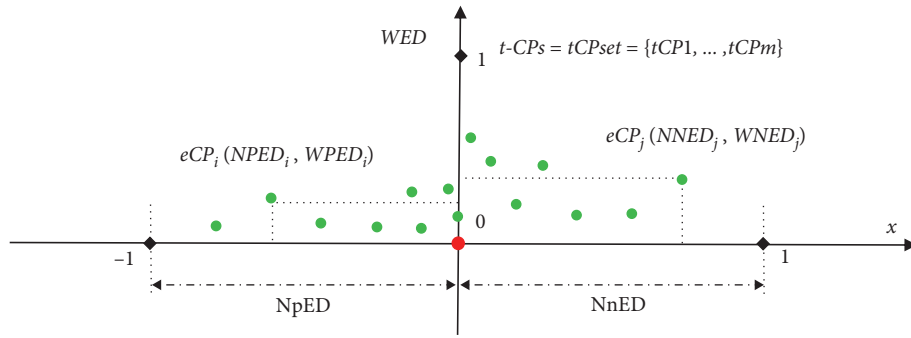


FIGURE 5: The scheme of WED metrics for MCPs detection on a set of target MCPs $tCPset = \{tCP_1, \dots, tCP_m\}$ in a time series X' . For each eCP_i in the scattered distribution of resultant $eCPs$, the value of WED^i refers to $WPED_i$ or $WNED_i$ according to whether the NED^i is located at the positive- $NpED$ or negative- $NnED$ area ranging from -1 to 1 in the x -axis.

$$MWED = \frac{\sum_{i=1}^N WED_i}{N} \quad (18)$$

Furthermore, following $MWED$, $1-MWED$ can be referred to as mean weighted true distance ($MWTD$) and used as a measure of the overall performance of a CPD model for MCPs detection on time series with a series of data fluctuations.

3. Results and Discussion

To accurately evaluate different CPD models, other related indexes were introduced besides our $WEDM$. In the synthetic experiments, time-series datasets were generated and assembled by using the Gaussian distribution function in the Matlab platform, and then repetitive tests on single CP detection were executed by using different TST, BST, KS, and SSA models. Meanwhile, the performance of CPD models was evaluated by using successive tests on MCPs detection that were implemented under different RSW and FSW frameworks, respectively.

3.1. Related Evaluation Indexes. In the synthetic tests, some other indexes are used for evaluating the convergence of different CPD models, including the hit, miss, and error rates, and computing time. Given a data segment Seg_i in the

time series X' mentioned above, the related definitions are introduced in terms of the error distance between the resultant $eCPs$ and the actual tCP_i as follows (Figure 6):

- (1) *Error distance:* Given an actual tCP_i assigned in the current data segment Seg_i , the error distance ED_{tCP_i} between each pair of the estimated eCP_j and the tCP_i is defined by $ED_{tCP_i} = |eCP_j - tCP_i|$.
- (2) *Hit area:* For the actual tCP_i , the hit area named HA_{tCP_i} is formulated by $HA_{tCP_i} = [tCP_i - hd_i, tCP_i + hd_i]$, where hd_i is the threshold value of error distance between tCP_i and eCP_j .
- (3) *Hit:* Given an error distance ED_{tCP_i} mentioned above, if $0 \leq ED_{tCP_i} \leq hd_i$ holds, then the tCP_i is hit by eCP_j and recorded by $Hit(tCP_i) = 1$. Therefore, the value of WED^i defined in formula (18) equals 0.
- (4) *Error:* On the other hand, if $ED_{tCP_i} > hd_i$ holds, then eCP_j is dealt as an error result labeled by $Error(eCP_j) = 1$. In this circumstance, the value of WED^i is within the range $(0, 1)$.
- (5) *Miss:* In addition, if no change point is detected from the Seg_i , then the target tCP_i is missed, and identified by $Miss(tCP_i) = 1$. Accordingly, the value of WED^i is set to be 1 because of the missing tCP_i .

Thereafter, the hit rate, miss rate, and error rate are formulated as follows:

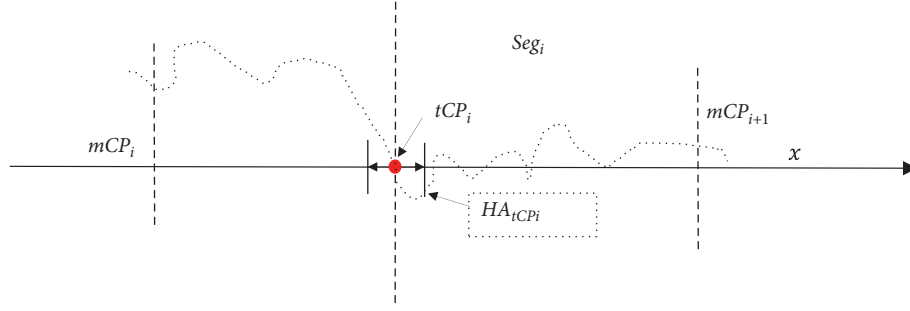


FIGURE 6: The scheme of single CP detection on the data segment Seg_i within a sliding window W_i . The definitions of hit, error, miss, and redundant are introduced according to the distance between tCP and eCP , respectively.

$$\begin{aligned} \text{Hit rate} &= \left(\frac{N_{hit}}{N_{eCPs}} \right) * 100\%, \\ \text{Miss rate} &= \left(\frac{N_{miss}}{N_{eCPs}} \right) * 100\%, \\ \text{Error rate} &= \left(\frac{N_{error}}{N_{eCPs}} \right) * 100\%. \end{aligned} \quad (19)$$

In which, $N_{hit} = \sum_{i=1}^{N_{tCPs}} \text{Hit}(tCP_i)$ is the number of actual $tCPs$ hit by the resultant $eCPs$, $N_{Miss} = \sum_{i=1}^{N_{tCPs}} \text{Miss}(tCP_i)$ is the part of actual $tCPs$ that are missed, and $N_{Error} = \sum_{i=1}^{N_{tCPs}} \text{Error}(eCP_i)$ stand for the number of the resultant $MCPs$ in which $D_{tCP_i} > hd_i$ holds. N_{eCPs} is the number of resultant $MCPs$ in total, and it is usually larger than N_{tCPs} , that is, the number of the actual $tCPs$ within the time series X' . Generally, it holds true that hit rate + miss rate + error rate = 1 for all the resultant $eCPs$.

- (6) *Computing time*: In addition, for a certain CPD model k , the computing time is mainly used for $tCPs$ detecting from the multiple data segments in X' , and it can be denoted as follows:

$$ST^k = \sum_{i=1}^{N_s} ST_i, \quad (20)$$

where ST_i refers to the computing time cost in the Seg_i , and N_s is the total data segments. Then, the normalized time is defined as follows:

$$NST^k = \frac{ST^k}{\sum_{k=1}^n ST^k}. \quad (21)$$

In which, ST^k stands for the computing time of the model k , and n is the total model to be compared. The NST^k represents the time ratio of model k to all methods, and then it can reflect the searching efficiency against others. Generally, both TST and BST models in our previous studies have a time complexity of nearly $O(\log N)$ [8, 10, 13]; therefore, they should be faster and more efficient than some traditional algorithms with time complexity about $O(N^2)$, such as KS, CUSUM, t -test, or SSA methods.

3.2. Repetitive Tests on Single CP Detection. In the first experiment, repetitive tests on single CP detection were executed on the synthetic dataset, that is, Dataset1 = $\{X^1, \dots, X^i, \dots, X^K\}$ that was generated by the Gaussian function in the Matlab R2016 platform. For each time series $X^i = \{x_1, \dots, x_i, \dots, x_N\}$ with single target CP, it is composed of both the positive area $X^{iL} = \{x_1, \dots, x_m\}$ and the negative area $X^{iR} = \{x_{m+1}, \dots, x_N\}$ before and after the assigned target $tCP_i = x_m$. The former X^{iL} and latter X^{iR} were generated by the normal distribution $N(\mu = 0, \sigma = 1)$ of size m (m time points included in the positive area), and $N(\mu = V, \sigma = 1)$ of size $N-m$ ($N-m$ time points in the negative area), respectively, where V is a constant mean value, and N is the total length of X^i .

Here, we first present the results from Dataset1 that was composed of multiple 20 data groups with different length N , variance V , and tCP , and each group contains 100 time-series samples. Therefore, Dataset1 included 2000 time series in total, and this experiment named Exp1 is performed by using TST, BST, KS, T, and SSA models, respectively. In our simulations, the time-series samples in each group were generated by selecting the random values of sample length N from 210 to 215, variance V from 1.0 to 3.7, and the position of actual tCP from 1 to N .

In the 20 groups of Exp1, the repetitive tests are executed by using different CPD models including the TST, BST, KS, T, and SSA, respectively (Figure 7). With the total 2000 time-series samples in Dataset1, the frequency and WED distribution of resultant $MCPs$ are illustrated from the positive-NpED range of $[-1, 0]$ to the negative-NnED range of $[0, 1]$ in the x -axis. From these results, we can see that if the resultant eCP is much closer to the central axis of $x = 0$, then the WED value generally gets smaller and tends to be 0, and vice versa. In all five models, TST and KS obtain the $eCPs$ that are mostly located near the central field of $x = 0$, and then have narrower WED distributions and smaller WED values than other models, except that TST has a few $eCPs$ fallen into the positive-NpED field. As for other BST, T, and SSA models, the $eCPs$ are mainly scattered with a wide range from the NpED to the NnED areas, therefore their WED distributions are wider and bigger, especially for T and SSA.

Meanwhile, these simulation results also illustrate that both TST and KS have better convergency than others, especially, the TST has the highest hit level and takes the

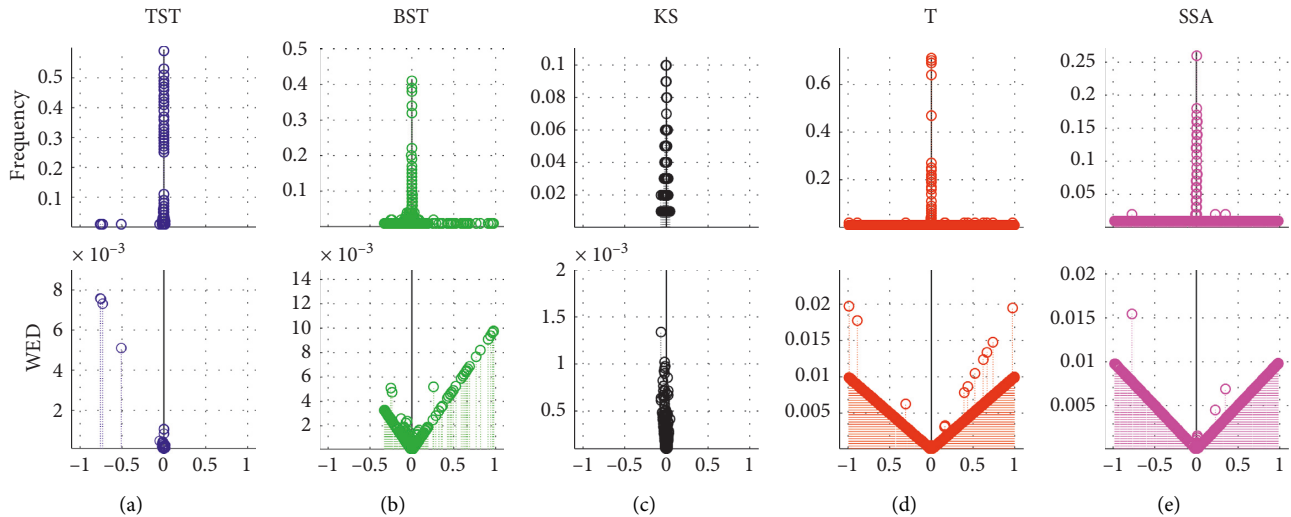


FIGURE 7: The frequency and WED distribution of resultant MCPs from the 20 groups in Dataset1. For the different models of (a) TST, (b) BST, (c) KS, (d) T, and (e) SSA, the frequency and WED distribution of the resultant MCPs are demonstrated from the NpED range of $[-1, 0]$ to the NnED range of $[0, 1]$ in the x -axis, respectively.

shortest convergent time in all five models. For the rest models, BST seems much better than others, and T has the worst convergency, because of the lowest hit, the biggest error, and convergent time in all five models. Furthermore, the mean analyses (Table 1) indicate that the TST takes the shortest computing time, has the highest hit rate, the smallest MWED, and the biggest MWTD out of the other four models. For T and SSA models, a lot of eCPs are scattered the whole field from NPED to NNED, especially, T has the biggest values of error rate and MWED and needs the longest time in all five models.

In addition, the efficiencies of five models are evaluated using random parameter values in a total of 20 tests. The dynamic tracks including hit rate, miss rate, error rate, and MWED are illustrated versus the test number from 1 to 20 (Figure 8). Also, the mean analyses on hit rate, miss rate, error rate, and MWED are presented in the histograms, in which, “1,” “2,” “3,” “4,” and “5” in x -axis refer to the TST, BST, KS, T, and SSA models, respectively. In the whole process of simulation tests, the TST model has a relatively higher hit rate with some fluctuations and keeps more stable and lower levels of miss rate, error rate, and MWED than others. Although KS has a smaller hit rate than TST and BST, it keeps lower tracks of miss and error rates than BST, T, and SSA. To some extent, BST has a bigger hit rate, and lower values of error rate and MWED than T and SSA, it seems unstable due to the drastic oscillations in the tracks of hit and miss rates. For T and SSA, both models have smaller hit rates and keep dramatic fluctuations in the tracks of error rate and MWED value, despite a lower miss rate than BST.

Furthermore, taking one representative test as an example, the simulations of single CP detection are repetitively executed by using 100 time-series samples with random values of parameters $N = 214$, $t_{CP} = 12267$, and $V = 1.9$. For different TST, BST, KS, T, and SSA models, the resultant eCPs are illustrated using the locations, distributions, frequency, and WED, in line with the test number, time-series

TABLE 1: The mean analyses of single CP detection in Exp1 by using TST, BST, KS, T, and SSA models.

| Items | Methods | | | | |
|------------|---------|--------|--------|--------|--------|
| | TST | BST | KS | T | SSA |
| MWTD | 0.9972 | 0.9633 | 0.9947 | 0.7030 | 0.8349 |
| Hit rate | 0.4040 | 0.1540 | 0.0430 | 0.0340 | 0.0585 |
| Miss rate | 0.0005 | 0.0035 | 0.0000 | 0.0005 | 0.0005 |
| Error rate | 0.0012 | 0.0038 | 0.0002 | 0.1202 | 0.0601 |
| MWED | 0.0028 | 0.0367 | 0.0053 | 0.2970 | 0.1651 |
| Time | 0.0032 | 0.0039 | 0.3126 | 0.5239 | 0.1566 |

positions, NPED, and NNED in the x -axis, respectively (Figure 9). For both TST and KS models, it is easy to see that most of the eCPs are located within the small range near the actual $t_{CP} = 12267$, and similar results can be found in the distribution, frequency, and WED analyses on the resultant eCPs. On the contrary, similar results for the rest of BST, T, and SSA models are that lots of the eCPs are randomly scattered across the fields from NPED to NNED, and small parts of the eCPs are gathered near the actual t_{CP} .

Then, the mean analyses for this representative test are summarized in terms of WMTD, hit rate, miss rate, error rate, MWED, and time (Table 2). The results show that the TST model has much smaller values of MWED, miss and error rates, and computing time, as well as the biggest values of hit rate and MWTD than others. Despite a long time and smaller hit rate than TST, KS kept similar levels of MWTD, hit, miss, and error rates with it. As for the rest BST, T, and SSA, although the three models had similar performance, BST had the biggest miss rate, T had the smallest MWTD and hit rate, and the biggest values of time, error rate, and MWED.

3.2.1. Successive MCPs Detection under the RSW Framework. In the second experiment, successive tests on MCPs detection were implemented by using other synthetic datasets

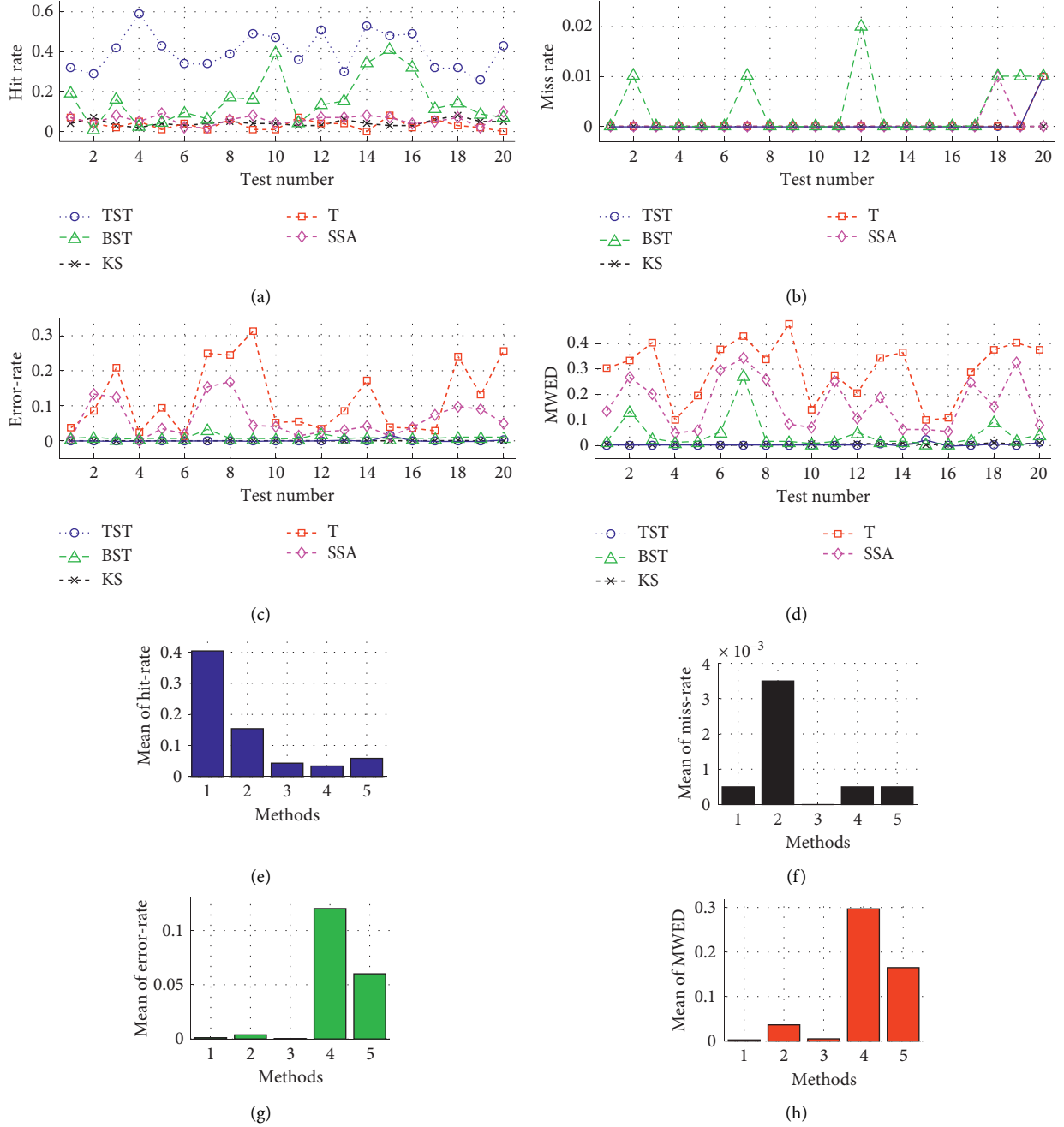


FIGURE 8: The results of multiple 20 tests on single CP detection by using 2000 synthetic time series in Dataset1 of Exp1, with random parameters of sample size (N) from 210 to 215, actual tCP from start to end of sample length (N), and variance (V) from 1.0 to 3.7. For TST, BST, KS, T, and SSA models, the dynamic tracks of (a) hit rate, (b) miss rate, (c) error rate, and (d) MWED versus simulation tests range from 1 to 20. In addition, the mean analyses on (e) hit rate, (f) miss rate, (g) error rate, and (h) MWED, in which, “1” “2”, “3”, “4”, and “5” in x -axis refer to the TST, BST, KS, T, and SSA models, respectively.

such as Dataset2 = $\{X^1, \dots, X^i, \dots, X^W\}$ that was composed of W time-series samples, and each sample $X^i = \{Seg^1, \dots, Seg^j, \dots, Seg^n\}$ was assembled by n data segments with different features and distributions. For a given $tMCP$ set = $\{tCP_1, \dots, tCP_n\}$, each tCP_i is assigned between two adjacent segments Seg^i and Seg^{i+1} , $1 \leq i \leq n-1$. Then, the sample X^i can be denoted as $X^i = \{x_1^{s1}, \dots, x_{Ns1-1}^{s1}, tCP_1, \dots, x_1^{sj}, \dots, x_{Nsj-1}^{sj}, tCP_j, x_1^{sn}, \dots, x_{Nsn-1}^{sn}, tCP_n\}$, where Nsj is the size of segment Seg^j in X^i . In the successive tests

on MCPs detection, two experiments named Exp2 and Exp3 were implemented based on Dataset2 under the RSW and FSW frameworks, respectively. For each experiment, a series of tests for MCPs detection was executed by using TST, BST, KS, T, and SSA models, respectively.

In Exp2, the number of segments n within each sample X_i was stochastically chosen from 15 to 30, and each data segment $Seg_j = \{x_1^{sj}, \dots, x_{Nsj}^{sj}\}$ was randomly generated by the Gaussian distribution $N(U_j, V_j)$ of length Nsj from 212 to

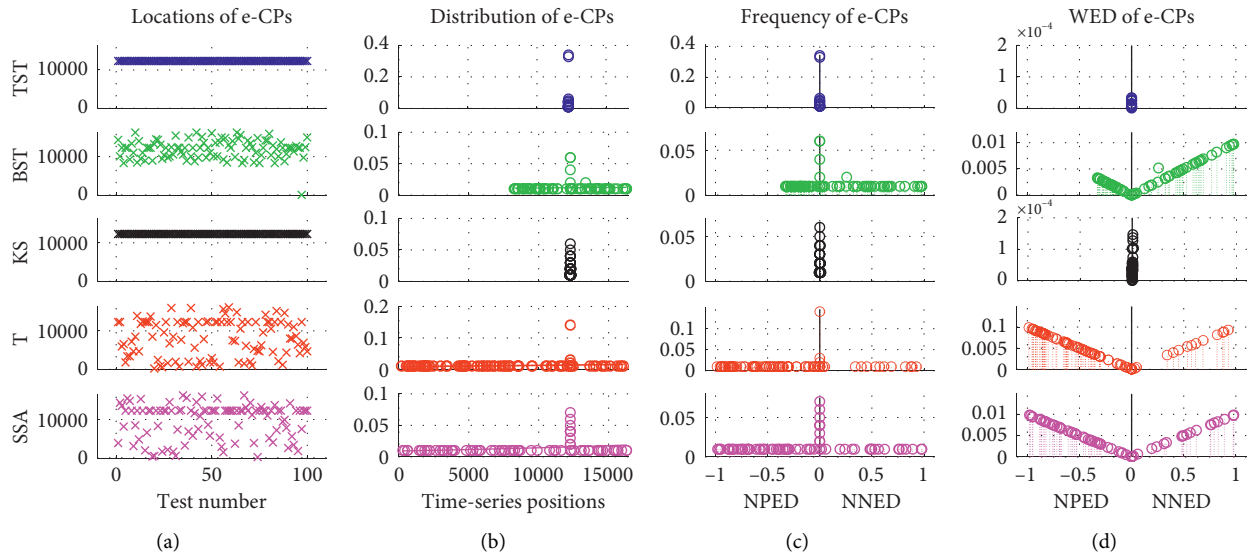


FIGURE 9: The repetitive simulations of single CP detection on 100 time series from one of 10 tests in Exp1, with random parameter values of sample size $N=214$, actual $tCP=12267$, and variance (V) = 1.9. By using different TST, BST, KS, T, and SSA models, the simulation results including (a) locations, (b) distributions, (c) frequency, and (d) WED of the resultant eCPs are represented in line with the test number, time-series positions, NPED, and NNED in the x -axis, respectively.

TABLE 2: The analyses of one representative test on repetitive single CP detection by different CPD models.

| items | Methods | | | | |
|------------|---------------|--------------|---------------|---------------|---------------|
| | TST | BST | KS | T | SSA |
| MWTD | 0.9998 | .7315 | 0.9980 | 0.5708 | 0.6567 |
| Hit rate | 0.3400 | 0.0600 | 0.0400 | 0.0010 | 0.0020 |
| Miss rate | 0.0000 | 0.0001 | 0.0000 | 0.0000 | 0.0000 |
| Error rate | 0.0001 | 0.0259 | 0.0001 | 0.2496 | 0.1536 |
| MWED | 0.0002 | 0.2685 | 0.0020 | 0.4292 | 0.3433 |
| Time | 0.0012 | 0.0012 | 0.3276 | 0.5861 | 0.0840 |

215, with mean U_j from 1.0 to $0.1 \times N_{MCPs}$, and variance V_j from 1 to $2.0 \times N_{MCPs}$, respectively. Here, we present the results of successive tests on MCPs detection under the RSW framework. First, the frequency and WED distribution of resultant MCPs (Figure 10) are displayed within the whole range from the negative-NPED field to the positive-NNED field in the x -axis. Generally, for a certain CPD model, the resultant MCPs are closer to the central axis $x=0$, their values of MWED are much smaller. In contrast, the bigger MWTD has, the better efficiency is, and vice versa. In all five models, the results (Figure 10) and the mean analyses (Table 3) show that most of the resultant MCPs detected by TST are located near the central axis $x=0$, and TST has the biggest hit rate, the smallest values of miss and error rates, therefore it has the highest MWTD out of others. For the BST model, although a lot of the resultant MCPs are scattered away from the central axis $x=0$, it has a smaller error rate and MWED, as well as a bigger hit rate and MWTD than the rest models. For KS, T, and SSA, the common feature is that most of the resultant MCPs are spread through the whole field ranging from -1 to 1 in the x -axis. KS has a bigger MWTD than the other two, T has the smallest

MWTD, and SSA has the biggest values of error rate and computing time in all five models.

Meanwhile, these simulations illustrate that the TST has the best convergence because it has the highest hit level, the lowest error, and takes the shortest convergent time in all five models. For the others, the BST model has much better convergence due to the higher hit, lower error, and shorter time than others. SSA seems the worst one in all five models, because of the lowest hit, the biggest error, and convergent time.

Second, the performance of five CPD models is demonstrated by a series of 10 tests in total, in which the respective parameters of the sample size N , the number of MCPs N_{MCPs} , the mean μ , and variance δ are randomly taken from 212–215, 15~30, $1 \sim 0.1 \times N_{MCPs}$, and $1 \sim 2 \times N_{MCPs}$, respectively. The results of dynamic tracks and mean analyses (Figure 11) indicate that the TST model still keeps a better grade with a higher and more stable level of hit rate, as well as the lower levels of error rate and MWED than the other four models. Although BST looks more efficient than KS, T, and SSA, the dynamic tracks in all four items present stronger fluctuations, especially for the miss rate.

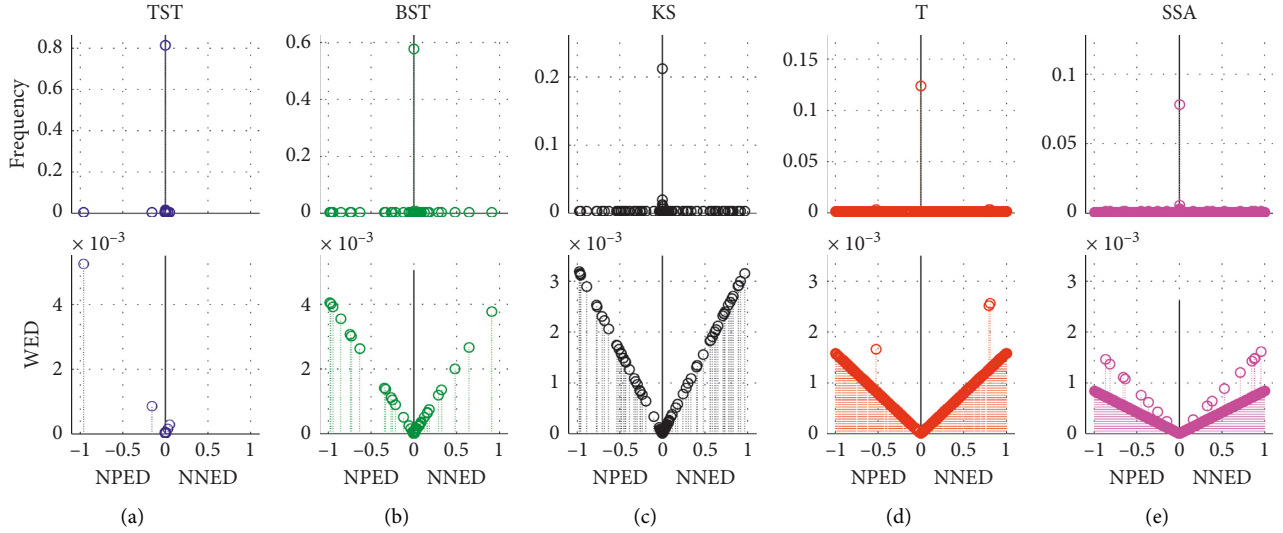


FIGURE 10: The analyses on the frequency and WED distribution of resultant MCPs in the total 10 tests of Exp2. For the different MCP models of (a) TST, (b) BST, (c) KS, (d) T, and (e) SSA under the RSW framework, the frequency and WED distribution of resultant MCPs are illustrated within the NPED ranging from -1 to 0 , and the NNED ranging from 0 to 1 in the x -axis.

TABLE 3: The performance analyses on MCPs detection by five CPD models in Exp2 under the RSW framework.

| items | Methods | | | | |
|------------|---------------|---------------|---------------|---------------|---------------|
| | TST | BST | KS | T | SSA |
| MWTD | 0.9264 | 0.8856 | 0.7850 | 0.5141 | 0.5299 |
| Hit rate | 0.8629 | 0.6255 | 0.2256 | 0.2029 | 0.0820 |
| Miss rate | 0.0398 | 0.0585 | 0.0430 | 0.0471 | 0.0018 |
| Error rate | 0.1006 | 0.3421 | 0.7682 | 0.7661 | 0.8851 |
| MWED | 0.0736 | 0.1144 | 0.2150 | 0.4859 | 0.4701 |
| Time | 0.0004 | 0.0003 | 0.1483 | 0.1134 | 0.7376 |

This probably means that BST has unstable performance during the process of MCPs detection. As for the rest models, they all have similar tracks of lower hit rate and bigger error rates. KS presents instability due to the fluctuant tracks of miss rate and MWED, and so does the T model because of the fluctuant miss rate in the total of random 10 tests. Also, the model's performance can be intuitively evaluated and distinguished from each other in terms of the mean analyses in the histograms (Figure 11(e)–11(h)).

Last, one representative test is selected from Exp2 above, and the simulations of MCPs detection are demonstrated by using a time series with $nMCPs = 25$ (Figure 12). For the diagnosed data sample (Figure 12(f)), the distributions of resultant MCPs are illustrated by using different CPD models of TST, BST, KS, T, and SSA models, respectively (Figure 12(a)–12(e)). The results of frequency and WED distribution of resultant MCPs (Figure 13) and mean analyses (Table 4) reveal that the TST is a superior one in all five models because most of the resultant MCPs hit the target MCP positions, and few of them are dealt with as miss or error states. The BST model takes second place due to a smaller hit rate and bigger error rate than TST. For the rest models, KS, T, and SSA get worse one by one because more

numbers of resultant MCPs are in the error state. As a result, the hit rate gets lower, and MWED takes bigger as well.

Successive MCPs detection under the FSW framework.

In the Exp3 under the FSW framework, the total of 30 data segments was arranged within each sample X_p , and each data segment $Seg_j = \{x_1^j, \dots, x_{N_{sj}}^j\}$ was randomly generated by the Gaussian distribution $N(U_j, V_j)$ of length N_{sj} from 212 to 215, with mean U_j from 1.0 to 0.1×30 and variance V_j from 1 to 2.0×30 , as well as with the size of fixed slide window N_{fsw} ranging from 26 to 215, respectively.

In our simulations, we execute a total of 10 successive tests on MCPs detection under the FSW framework. First, the frequency and WED distribution of resultant MCPs (Figure 14) are displayed from the negative-NPED field to the positive-NNED field in the x -axis. Generally, for a certain CPD model, the resultant MCPs are much closer to the central axis $x=0$, and their WED values are much smaller. The results (Figure 14 and Table 5) indicate that for the TST model, most of the resultant MCPs detected are located near the central axis $x=0$, and it has the biggest hit rate, the smallest values of error rate, MWED, and computing time; therefore, it has the highest MWTD in all five CPD models. As for BST, KS, T, and SSA models, the

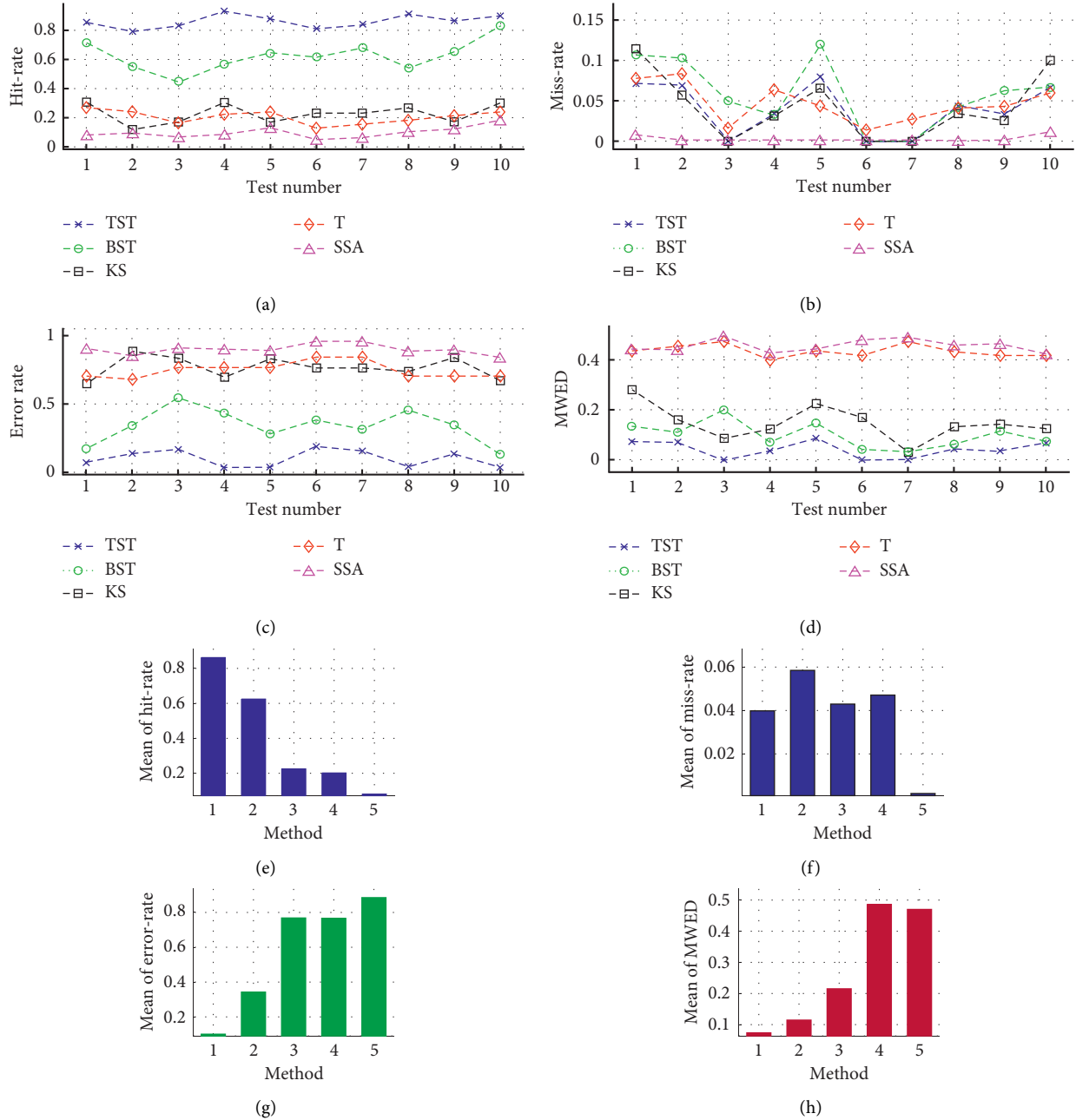


FIGURE 11: The simulations of MCPs detection on the total of 10 tests in Exp2 under the RSW framework, with random parameters of sample size N_{sj} from 212 to 215, the number of tCPs N_{MCPs} from 15 to 30, mean U_j from 1 to $0.1 \times N_{MCPs}$, and variance V_j from 1 to $2 \times N_{MCPs}$, respectively. For the different TST, BST, KS, T, and SSA models, the performance analyses are denoted in (a) hit rate, (b) miss rate, (c) error rate, and (d) MWED, respectively. Furthermore, the mean analyses are illustrated in histograms of (e) hit rate, (f) miss rate, (g) error rate, and (h) MWED, in which, “1,” “2,” “3,” “4,” and “5” in x -axis refer to TST, BST, KS, T, and SSA, respectively.

common feature is that most of the resultant MCPs are randomly scattered through the whole field ranging from -1 to 1 in the x -axis. For KS, it has a smaller miss rate and MWED and a bigger MWTD than the others. Although BST has a bigger hit rate and shorter time, it has a bigger MWED and smaller MWTD than TST and KS. T and SSA have much bigger values of MWED, error rate, and smaller MWTD, especially SSA has the smallest MWTD and the biggest values of error rate and time in all five models.

Meanwhile, these simulations illustrate that the TST has the best convergency, in terms of the highest hit, the lowest error, and the shortest time in all five models. For the other four models, the BST model is much better than the rest ones, because it has a relatively higher hit level, lower error rate, and much shorter time than others. Unfortunately, SSA has the worst convergency in all five models, due to the lowest hit level, the biggest error rate, and the longest convergent time out of the other four models.

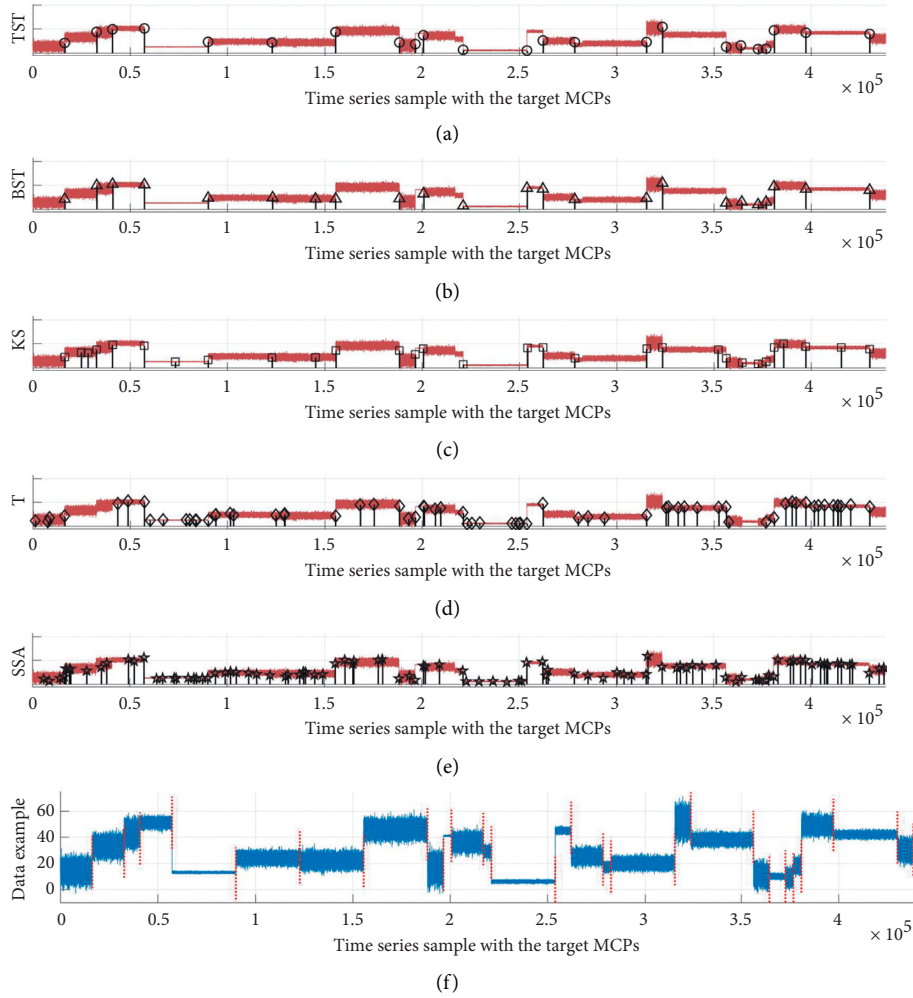


FIGURE 12: The simulations of MCPs detection on the representative test with $N_{MCPs} = 25$. For one selected sample in (f), the resultant MCPs are illustrated by using different models of (a) TST, (b) BST, (c) KS, (d) T, and (e) SSA, respectively.

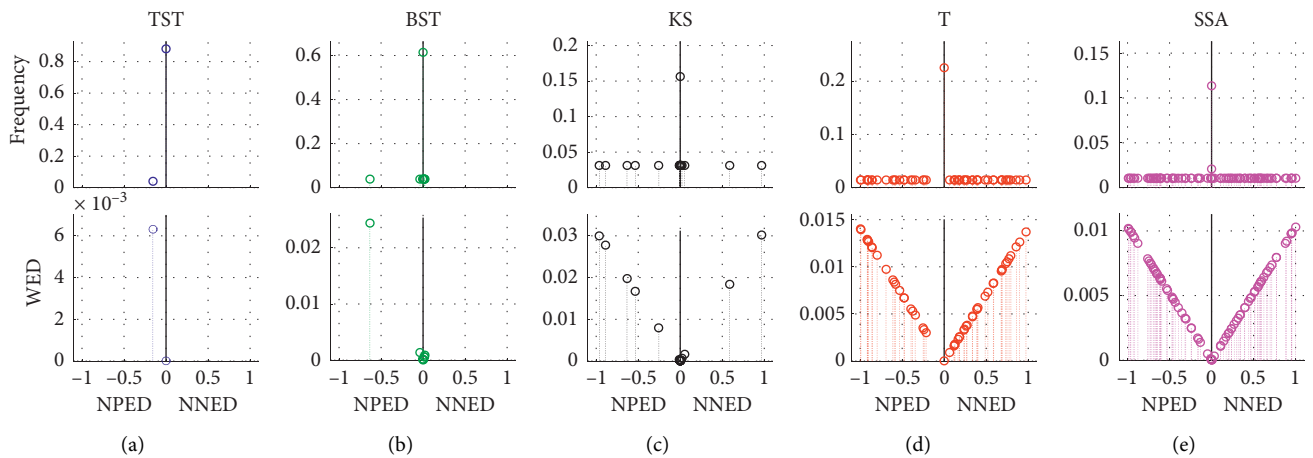


FIGURE 13: The results of WED evaluation on the 100 samples with $N_{MCPs} = 25$ in Exp2. For the different MCP models of (a) TST, (b) BST, (c) KS, (d) T, and (e) SSA, the frequency and WED distribution of resultant MCPs are illustrated within the NPED ranging from -1 to 0 , and the NNED ranging from 0 to 1 in the x -axis, respectively.

TABLE 4: The mean analyses on five CPD models in one representative MCPs detection test with $N_{MCPs} = 25$.

| items | Methods | | | | |
|------------|---------------|---------------|---------------|---------------|---------------|
| | TST | BST | KS | T | SSA |
| MWTD | 0.9655 | 0.9319 | 0.8836 | 0.6042 | 0.5779 |
| Hit rate | 0.9310 | 0.5667 | 0.3030 | 0.2222 | 0.0727 |
| Miss rate | 0.0345 | 0.0333 | 0.0303 | 0.0635 | 0.0000 |
| Error rate | 0.0345 | 0.4333 | 0.6970 | 0.7619 | 0.8909 |
| MWED | 0.0345 | 0.0681 | 0.1164 | 0.3958 | 0.4221 |
| Time | 0.0004 | 0.0003 | 0.1889 | 0.1209 | 0.6896 |

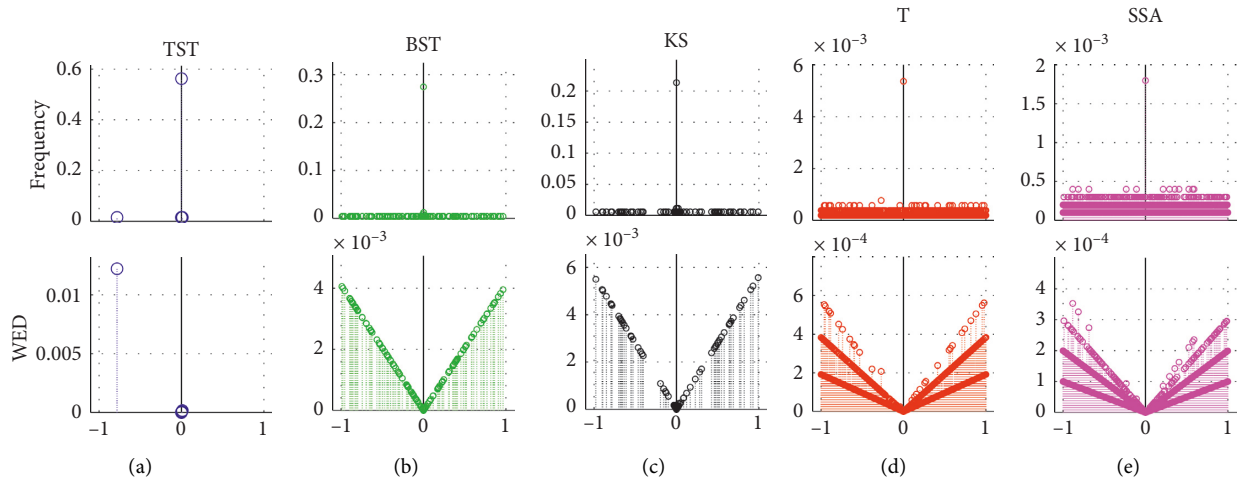
FIGURE 14: The analyses on the frequency and WED distribution of resultant MCPs in the total 10 tests of Exp3. For the different MCP models of (a) TST, (b) BST, (c) KS, (d) T, and (e) SSA under the FSW framework, the frequency and WED distribution of resultant MCPs are illustrated within the NPED field ranging from -1 to 0 and the NNED field ranging from 0 to 1 in the x -axis.

TABLE 5: The mean analyses on MCPs detection in Exp3 by five CPD models under the FSW framework.

| Items | Methods | | | | |
|------------|---------------|---------------|---------------|---------------|---------------|
| | TST | BST | KS | T | SSA |
| MWTD | 0.9875 | 0.7758 | 0.8268 | 0.5063 | 0.5009 |
| Hit rate | 0.7867 | 0.5167 | 0.3106 | 0.1862 | 0.0525 |
| Miss rate | 0.1900 | 0.0930 | 0.0894 | 0.0977 | 0.0633 |
| Error rate | 0.0200 | 0.4186 | 0.6194 | 0.7271 | 0.9004 |
| MWED | 0.0125 | 0.2242 | 0.1732 | 0.4937 | 0.4991 |
| Time | 0.0006 | 0.0008 | 0.1419 | 0.0766 | 0.7802 |

Second, the performance evaluation on five CPD models is demonstrated respectively by a series of successive MCPs detection tests in Exp3. Generally, the dynamic tracks and histogram analyses (Figure 15) show that all five CPD models present respective instability in response to the size of the fixed slide window, N_{fsw} ranging from 26 to 215, especially for the TST, BST, and KS models. Despite the TST model having the biggest miss rate with drastic fluctuations, it still keeps a better efficiency due to the highest hit rate and the lowest levels of error rate and MWED out of the other four models. As for the rest ones, BST seems better than KS, T, and SSA, because of the higher hit rate and the slightly decreasing level of error rate. Although KS reversely keeps decreasing hit rate and increasing error rate with big

fluctuation, it seems better than T and SSA, on account of lower levels of miss rate and MWED. Both T and SSA present inefficiency and insensitivity in response to the increasing N_{fsw} , especially for the SSA model, with the lowest hit rate and the highest levels of error rate and MWED out of other ones.

Last, taking the TST model as an example, five representative simulations are selected from the total 10 tests in the FSW framework of Exp3 (Figure 16(a)–16(e)), and then the performance evaluation is listed under the values of $N_{fsw} = 26, 28, 212, 214,$ and 215 , respectively (Table 6). Given one data sample with $N_{MCPs} = 30$ (Figure 16(f)), the results of MCPs detection show that the TST model presents the best performance as $N_{fsw} = 212$, in terms of the biggest values of

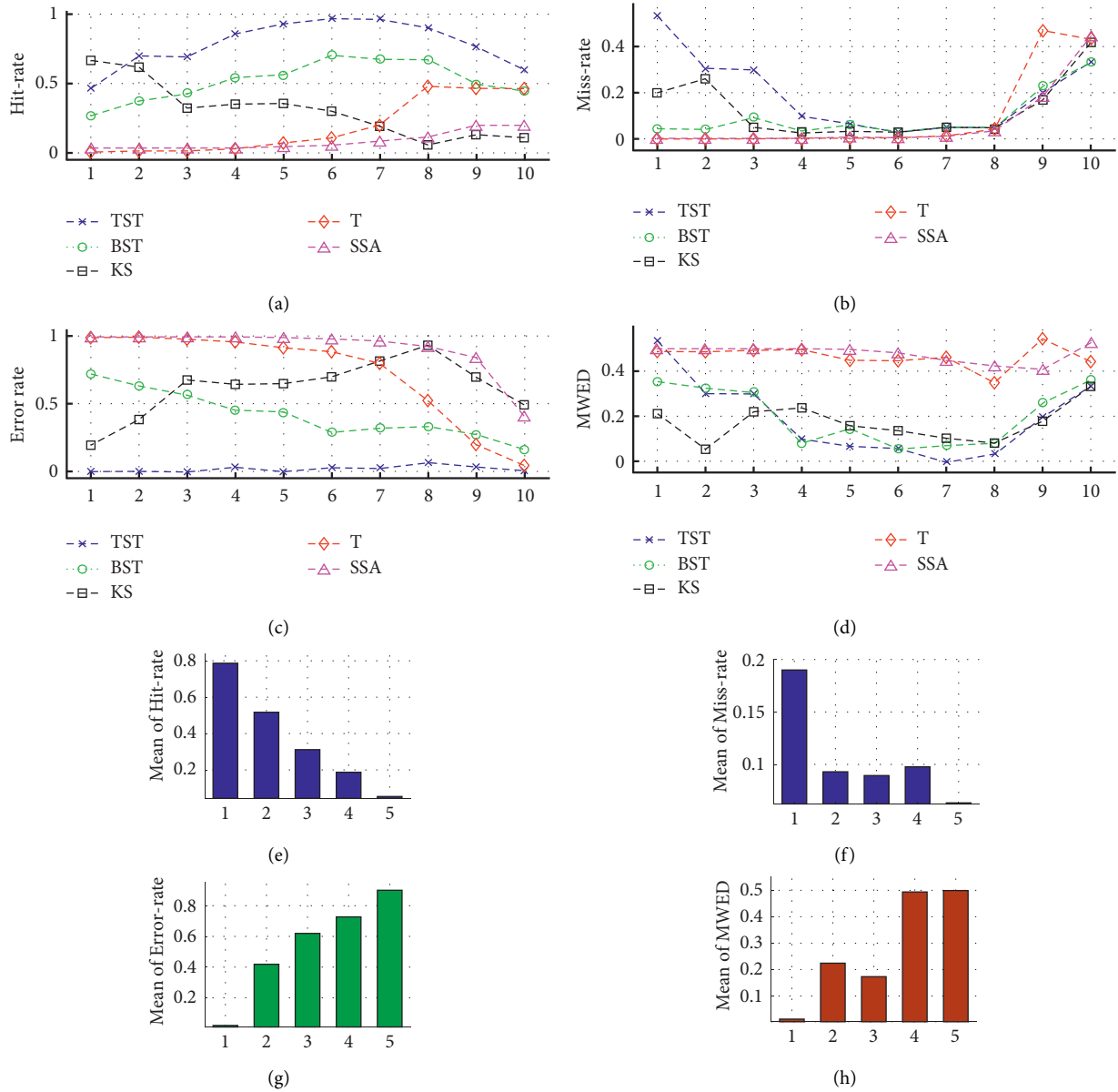


FIGURE 15: The simulations of MCPs detection on the total of 10 tests in Exp3 under the FSW framework, with random parameters of sample size N_j from 212 to 215, the fixed number of tCPs $N_{MCPs} = 30$, mean U_j from 1 to 0.1×30 , and variance V_j from 1 to 2×30 , respectively. For the different CPD models of TST, BST, KS, T, and SSA, the performance analyses are denoted in (a) hit rate, (b) miss rate, (c) error rate, and (d) mwed, respectively. Furthermore, the mean analyses are illustrated in (e) hit rate, (f) miss rate, (g) error rate, and (h) MWED, in which, “1,” “2,” “3,” “4,” and “5” in x -axis refer to TST, BST, KS, T, and SSA, respectively.

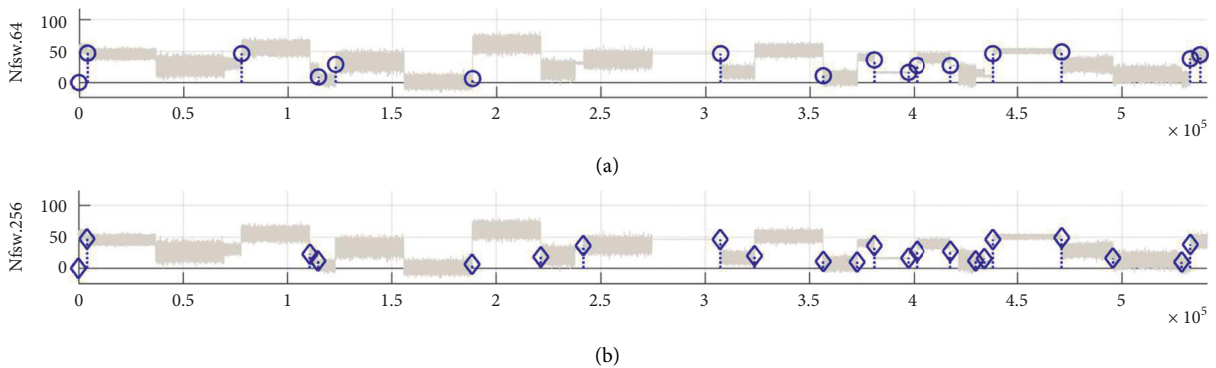


FIGURE 16: Continued.

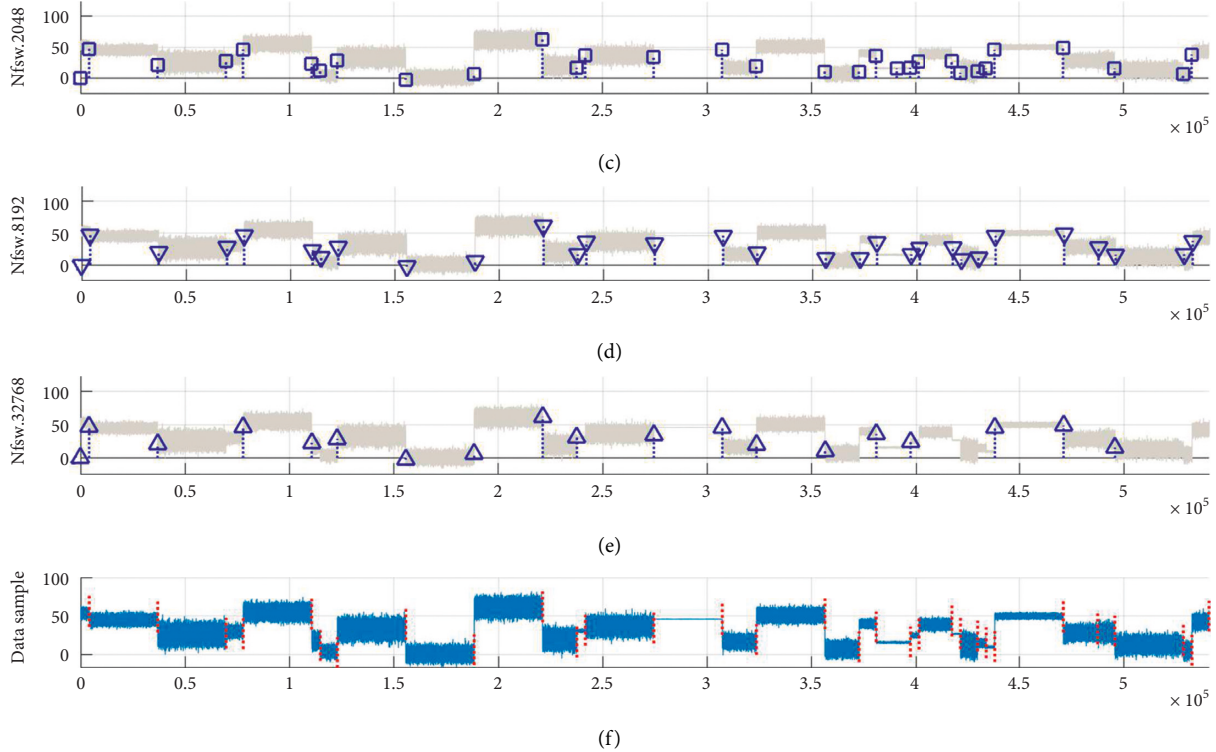


FIGURE 16: The simulations of MCPs detection by TST model under different sizes of fixed slide window N_{fsw} in the FSW framework. Given the diagnosed data sample with $N_{MCPs} = 30$ in (f), the resultant MCPs detection is illustrated under different N_{fsw} values of (a) 26, (b) 28, (c) 212, (d) 214, and (e) 215, respectively.

TABLE 6: The performance evaluations on the TST model with different N_{fsw} under the FSW framework in Exp3.

| Items $N_{fsw} =$ | Hit rate | Miss rate | Error rate | MWED | MWTD |
|-------------------|---------------|---------------|---------------|---------------|---------------|
| 26 | 0.4667 | 0.5333 | 0.0000 | 0.5333 | 0.4667 |
| 28 | 0.7000 | 0.3000 | 0.0000 | 0.3000 | 0.7000 |
| 212 | 0.9667 | 0.0303 | 0.0333 | 0.0586 | 0.9414 |
| 214 | 0.7667 | 0.2000 | 0.0333 | 0.2002 | 0.7998 |
| 215 | 0.6000 | 0.3333 | 0.0000 | 0.3333 | 0.6667 |

hit rate and MWTD, and the smallest values of miss and error rates and MWED in all five tests. However, the efficiency of TST tends to be worse as the value of N_{fsw} takes too bigger or too smaller. Therefore, the size of the fixed slide window is a key factor for the FSW framework during the MCPs detection.

In all, these results in the two experiments above suggest that the proposed WED method can visually present the distribution of resultant eCPs in the error state and the normalized distance from the target position of zero in the x -axis. The simulation results suggest that the mean analyses of MWED can generally count the mean value of error ratio against total tests and then measure the efficiency of a certain model in the successive MCPs detection. The performances of different CPD models can be evaluated, and the better ones can be discerned from the others.

4. Conclusions and Discussion

In this study, a novel WEDM method is proposed for evaluating the overall performance of a CPD model across not only repetitive tests on single CP detection, but also successive tests on multiple change-point (MCP) detection on synthetic time series under different RSW and FSW frameworks. In this WEDM method, a concept of normalized error distance was introduced that allows comparisons of the distance between the estimated change-point (eCP) position and the target change-point (tCP) in the synthetic time series. Especially, both positive- and negative-error distances between resultant eCPs and actual tCPs are weighted or normalized for creating WED metrics.

As opposed to previous methods, our WEDM allows comparison when CPD is used across multiple time-series

samples with different lengths and variances, especially cross multiple data segments in an identical time series, with different patterns, such as data distributions, segment sizes, and number and positions of targets tCPs. In the successive MCPs detection, our WEDM method first divides the original sample into a series of data segments in terms of assigned target change points and then calculates a normalized error distance (NED) value for each segment. Next, WEDM presents the frequency and WED distribution of the resultant eCPs from all data segments in the normalized positive-error distance (NPED) and the normalized negative-error distance (NNED) intervals in the same coordinates. Last, the mean WED (MWED) and MWTD (1-MWED) were obtained and dealt with as important performance indexes.

In our simulations, a series of MCPs detection tests were executed by using synthetic time-series datasets in the Matlab platform, and the proposed method was applied to the evaluation of the CPD utilizing TST, BST, KS, T, and SSA models under repetitive single CP detection in Exp1, successive MCPs detection under the RSW in Exp2, and FSW framework in Exp3, respectively. The results of the study showed its ability to compare the results from the CPD models working with a series of synthetic tests on multiple time-series samples. The WED metrics offer a new way of evaluating CPD performance. It allows better visualization of the distribution of the resultant eCPs when the CPD models work on multiple time series with different data features, as well as multiple data segments of a time-series sample with different data patterns. Meanwhile, the convergence of different CPD models was analyzed in terms of the dynamic tracks and mean analyses on the value of WED, as well as other measurements, including the rates of hit, error, and miss, and the computational cost. Our WEDM method can not only offer a visualizable and overall measure but also give better advice for users as to what CPD models to use based on the application.

Data Availability

Some synthetic time-series datasets were generated in the Matlab simulation platform, and no real datasets are used specially for the experimental validations in this study.

Conflicts of Interest

All authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Prof. Qing Zhang and Prof. Mohan Karunanithi of the Australia e-Health Research Centre, CSIRO Computation Informatics, for their assistance, support, and advice for this paper. Also, the authors appreciate the editors and referees for their very helpful comments that led to a substantial improvement of this manuscript. This paper is supported by the National Natural Science Foundation of China (no. 61104154) and the

Specialized Research Fund for Natural Science Foundation of Shanghai (nos. 16ZR1401300 and 16ZR1401200).

References

- [1] R. J. Bolton and D. J. Hand, "Statistical fraud detection: a review," *Statistical Science*, vol. 17, no. 3, pp. 235–249, 2002.
- [2] K. Yamanishi, J.-I. Takeuchi, G. Williams, and P. Milne, "Online unsupervised outlier detection using finite mixtures with discounting learning algorithms," *Data Mining and Knowledge Discovery*, vol. 8, no. 3, pp. 275–300, 2004.
- [3] U. Murad and G. Pinkas, "Unsupervised profiling for identifying superimposed fraud," in *Principles of Data Mining and Knowledge Discovery*, pp. 251–261, Springer, Manhattan, NY, USA, 1999.
- [4] R. Killick, I. A. Eckley, K. Ewans, and P. Jonathan, "Detection of changes in variance of oceanographic time-series using changepoint analysis," *Ocean Engineering*, vol. 37, no. 13, pp. 1120–1126, 2010.
- [5] Y. Wang, C. Wu, Z. Ji, B. Wang, and Y. Liang, "Non-parametric change-point method for differential gene expression detection," *PLoS one*, vol. 6, no. 5, Article ID e20060, 2011.
- [6] J. Wu, Y. Chen, and S. Zhou, "Online detection of steady-state operation using a multiple-change-point model and exact Bayesian inference," *Iise Transactions*, vol. 48, no. 7, 2016.
- [7] D. Matteo, M. L. C. Simone, M. P. Sara, and F. Matteo, "A comparison between power spectral density and network metrics: an EEG study," *Biomedical Signal Processing and Control*, vol. 57, 2019.
- [8] J. P. Qi, Y. Zhu, F. Pu, and P. Zhang, "A novel RSW&TST framework of MCPs detection for abnormal pattern recognition on large-scale time series and pathological signals in epilepsy," *PLoS One*, 2021.
- [9] M. Jalayer, C. Orsenigo, and C. Vercellis, "Fault detection and diagnosis for rotating machinery: a model based on convolutional LSTM, Fast Fourier and continuous wavelet transforms," *Computers in Industry*, vol. 125, 2020.
- [10] M. G. Terzano and L. Parrino, "Chapter 8 the cyclic alternating pattern (CAP) in human sleep," *Handbook of Clinical Neurophysiology*, vol. 6, pp. 79–93, 2005.
- [11] D. U. Elif, "Eigenvector methods for automated detection of electrocardiographic changes in partial epileptic patients," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, 2009.
- [12] T. S. Kumar and V. Kanhangad, "Detection of electrocardiographic changes in partial epileptic patients using local binary pattern based composite feature," *Australasian Physical & Engineering Sciences in Medicine*, 2017.
- [13] A. L. Goldberger, L. A. Amaral, L. Glass et al., "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiological signals," *Circulation*, vol. 101, no. 23, pp. E215–20, 2000.
- [14] M. G. Terzano, L. Parrino, A. Sherieri et al., "Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep," *Sleep Medicine*, vol. 2, no. 6, p. 537, 2002.
- [15] J. P. Qi, Q. Zhang, Y. Zhu, and J. Qi, "A novel method for fast change-point detection on simulated time series and electrocardiogram data," *PLoS One*, vol. 9, no. 4, in English, Article ID e93365, 2014.
- [16] V. Pillai, D. A. Kalmbach, and J. A. Ciesla, "A meta-analysis of electroencephalographic sleep in depression: evidence for genetic biomarkers," *Biological Psychiatry*, vol. 70, no. 10, pp. 912–919, 2011.

- [17] A. B. de Luna, I. Cygankiewicz, A. Baranchuk et al., "Prinzmetal angina: ECG changes and clinical considerations: a consensus paper," *Annals of Noninvasive Electrocardiology*, vol. 19, no. 5, pp. 442–453, 2014.
- [18] N. Rabin, M. Kahlon, S. Malayev, and A. Ratnovsky, "Classification of human hand movements based on EMG signals using nonlinear dimensionality reduction and data fusion techniques," *Expert Systems with Applications*, vol. 149, Article ID 113281, 2020.
- [19] Y. Wen, J. Wu, Q. Zhou, and T.-L. Tseng, "Multiple-change-point modeling and exact bayesian inference of degradation signal for prognostic improvement," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 2, pp. 613–628, 2019.
- [20] G. Patel, M. Nowak, and C. Castellini, "Exploiting knowledge composition to improve real-life hand prosthetic control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, p. 1, 2017.
- [21] A. Prakash and S. Sharma, "Single-channel Surface Electromyography (sEMG) Based Control of a Multi-Functional Prosthetic Hand," *Instrumentation Science & Technology*, vol. 49, pp. 1–18, 2021.
- [22] J.-P. Qi, Q. Zhang, Y. Zhu, and J. Qi, "A novel method for fast Change-Point detection on simulated time series and electrocardiogram data," *Plos One*, vol. 9, no. 4, Article ID e93365, 2014.
- [23] V. Moskvina and A. A. Zhigljavsky, "An algorithm based on singular spectrum analysis for change-point detection," *Communications in Statistics-Simulation and Computation*, vol. 32, no. 2, pp. 319–352, 2003.
- [24] J.-P. Qi, J. Qi, and Q. Zhang, "A fast framework for abrupt change detection based on binary search trees and Kolmogorov statistic," *Computational Intelligence and Neuroscience*, vol. 2016, pp. 1–16, 2016.
- [25] B. S. Darkhovski, "Nonparametric methods in change-point problems: a general approach and some concrete algorithms," *Institute of Mathematical Statistics Lecture Notes-Monograph Series*, vol. 23, pp. 99–107, 1994.
- [26] F. Yang, Y. Cui, F. Wu, and R. Zhang, "Fault monitoring of chemical process based on sliding window wavelet DenoisingGLPP," *Processes*, vol. 9, no. 1, p. 86, 2021.
- [27] S. Ariens, E. Ceulemans, and J. K. Adolf, "Time series analysis of intensive longitudinal data in psychosomatic research: a methodological overview," *Journal of Psychosomatic Research*, vol. 137, 2020.
- [28] D. J. Cook and N. C. Krishnan, *Activity Learning: Discovering, Recognizing, and Predicting Human Behavior from Sensor Data*, Wiley, Hoboken, NJ, USA, 2015.
- [29] S. Aminikhanghahi and D. J. Cook, "A survey of methods for time series change point detection," *Knowledge and Information Systems*, vol. 51, no. 2, pp. 339–367, 2017.
- [30] M. S. Pepe, "Receiver operating characteristic methodology," *Journal of the American Statistical Association*, vol. 95, no. 449, pp. 308–311, 2000.
- [31] J. A. Hanley, "Receiver operating characteristic (ROC) methodology: the state of the art," *Critical Reviews In Diagnostic Imaging*, vol. 29, pp. 307–335, 1989.
- [32] M. Grzybowski and J. G. Younger, "Statistical methodology: III. Receiver operating characteristic (ROC) curves," *Academic Emergency Medicine*, vol. 4, no. 8, pp. 818–826, 1997.
- [33] J. P. Qi, Y. Zhu, and P. Zhang, "A WED Method for Evaluating the Performance of Change-Point Detection Algorithms," in *Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1406–1410 in English, Madrid, Spain, December 2018.
- [34] J. P. Qi, J. Qi, F. Pu, and T. Gong, "Multi-channel detection for abrupt change based on the Ternary Search Tree and Kolmogorov statistic method," in *Proceedings of the Chinese Control Conference (CCC)*, Hangzhou, China, July 2015.
- [35] V. Moskvina and A. Zhigljavsky, "An algorithm based on singular Spectrum analysis for change-point detection," *Communications in Statistics - Simulation and Computation*, vol. 32, no. 2, pp. 319–352, 2003.
- [36] Z. Yang, K.-T. Fang, and S. Kotz, "On the Student's t-distribution and the t-statistic," *Journal of Multivariate Analysis*, vol. 98, no. 6, pp. 1293–1304, 2007.
- [37] J. P. Qi, Q. Gu, Y. Zhu, and P. Zhang, "A KST framework for correlation network construction from time series signals," in *Proceedings of the Ninth International Conference on Graphic and Image Processing*, Qing Dao, China, April 2017.
- [38] J. Qi, F. Pu, J. Liu, J. Zhu, and H. Gong, "A preliminary RSWKS framework for multiple change points (MCPs) detection on pathological signal analysis in partial epilepsy," in *Proceedings of the 2020 Chinese Automation Congress (CAC)*, Shanghai, China, November 2020.
- [39] M. H. M. Noor, Z. Salci, and K. I.-K. Wang, "Adaptive sliding window segmentation for physical activity recognition using a single tri-axial accelerometer," *Pervasive and Mobile Computing*, vol. 38, pp. 41–59, 2017.
- [40] W. Li, W. Guo, X. Luo, and X. Li, "On sliding window based change point detection for hybrid SIP DoS attack," in *Proceedings of the Services Computing Conference (APSCC)*, December 2010.
- [41] U. Yun, G. Lee, and E. Yoon, "Advanced approach of sliding window based erasable pattern mining with list structure of industrial fields," *Information Sciences*, vol. 494, pp. 37–59, 2019.
- [42] T. J. Leng and L. Q. Tang, "Based on the fixed length sliding window Spectrum curve of gene identification method study," *Applied Mechanics and Materials*, vol. 644–650, pp. 5351–5355, 2014.
- [43] A. Villalba and D. Carrera, "Constant- time approximate sliding window framework with error control," in *Proceedings of the 2019 IEEE 22nd International Symposium on Real-Time Distributed Computing (ISORC)*, pp. 99–107 in English, Valencia, Spain, May 2019.
- [44] A. Villalba, J. L. Berral, and D. Carrera, "Constant-time sliding window framework with reduced memory footprint and efficient bulk evictions," *IEEE Transactions on Parallel and Distributed Systems*, vol. 30, no. 3, pp. 486–500, 2019, in English.