

# Trend analysis and change point techniques: a survey

Shilpy Sharma<sup>1</sup> · David A. Swayne<sup>1</sup> · Charlie Obimbo<sup>1</sup>

<sup>1</sup> School of Computer Science, University of Guelph, Guelph, Canada

Received: 29 September 2015/Revised: 13 January 2016/Accepted: 14 January 2016/Published online: 8 February 2016  
© Joint Center on Global Change and Earth System Science of the University of Maryland and Beijing Normal University and Springer-Verlag Berlin Heidelberg 2016

**Abstract** Trend analysis and change point detection in a time series are frequent analysis tools. Change point detection is the identification of abrupt variation in the process behavior due to distributional or structural changes, whereas trend can be defined as estimation of gradual departure from past norms. We examine four different change point detection methods which, by virtue of current literature, appear to be the most widely used and the newest algorithms. They are Wild Binary Segmentation, E-Agglomerative algorithm for change point, Iterative Robust Detection method and Bayesian Analysis of Change Points. We measure the power and accuracy of these current methods using simulated data. We draw comparisons on the functionality and usefulness of each method. We also analyze the data in the presence of trend, using Mann–Kendall and Cox–Stuart methods together with the change point algorithms, in order to evaluate whether presence of trend affects change point or vice versa.

**Keywords** Change point analysis · Trend analysis · LO(W)ESS

## 1 Introduction

A change point can be defined as unexpected, structural, changes in time series data properties such as the mean or variance. Many change point detection algorithms have

been proposed for single and multiple changes. Several parametric and nonparametric methods are available depending on the change in mean or variance (Adams and MacKay 2007; Fryzlewicz 2013; Gerard-Merchant et al. 2008; Matteson and James 2013; Reeves et al. 2007). In water quality data, a change point can occur due to floods, droughts or other changes occurring in the environment. Figure 1 shows an example of a change point. Our motive is to study, categorize and examine change point detection methods: Bayesian Analysis of Change Points (BCP), Wild Binary Segmentation (WBS), E-Agglomerative algorithm (E-Agglo.), and Iterative Robust Detection (IR), and to draw inference on their functionality and effectiveness. We found that different methods often give different numbers and locations of change points. WBS and E-Agglo. are the latest techniques for change point detection, whereas IR is the improved version of iterative two-phase regression algorithm and BCP overcomes the problem of change point detection in “change point” package in R. Our intention is to analyze which of these methods are most capable of finding exact number and locations of change points, among above-mentioned methods.

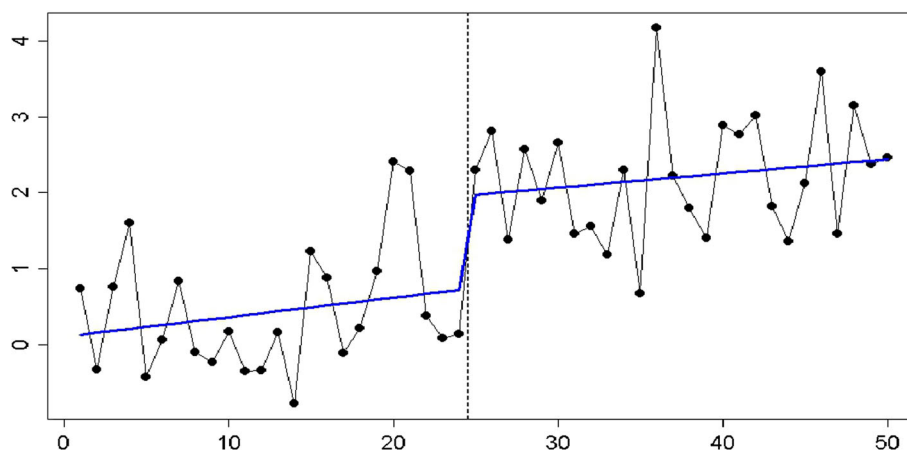
Trend analysis is defined as a process of estimating gradual change in future events from past data. Different parametric and nonparametric techniques are used to estimate trends. The Mann–Kendall, Spearman rho, seasonal Kendall, the Sen’s  $t$ , and Cox–Stuart are the most widely used methods (Hamed 2008; Hessa et al. 2001; Patal and Kahya 2006; Yue et al. 2002). We have focused on the Cox–Stuart and Mann–Kendall to estimate trend. Abrupt or gradual changes exist together. For example, in hydrological data, if there is a sudden release of water from an impoundment followed by a return to a prior regime, it

---

✉ Shilpy Sharma  
Shilpy@uoguelph.ca

David A. Swayne  
dswayne@cis.uoguelph.ca

Charlie Obimbo  
cobimbo@uoguelph.ca

**Fig. 1** Change point example

appears as if the sharp downward trend may be misinterpreted as a series of change points. We found only a few references to analysis of combined change point and trend (Xiong and Guo 2004). Analyzing only change point and overlooking trend or vice versa may mislead the calculation. It could be possible that the trend masks abrupt change. For example, estimating trend alone could imply that the trend is either increasing or decreasing. This concern is answered when both change point and trend are analyzed together. Analyzing both possibilities can make improved prediction for future events.

The remainder of this paper is organized as follows. Section 2 reviews existing change point methods, used for our experiments. Section 3 focuses methods used, Sect. 4 shows the simulation results of each method, Sect. 5 emphasizes on trend and change point interaction, and Sect. 6 provides conclusions.

## 2 Review of existing methods

### 2.1 E-Agglomerative

This algorithm is used to detect multiple change points in data set. It is able to detect numbers and location of change points, and it is considered as computationally efficient (Matteson and James 2013). The number of change points is estimated based on “goodness of fit” of clusters in the data set. The R package, “*ecp*,” is available to calculate number of change points (James and Matteson 2013). In *ecp*, two different methods, E-Divisive and E-Agglomerative algorithms, are provided for univariate as well as multivariate data. Divisive methods estimate change points using a bisection algorithm, whereas agglomerative methods detect abrupt change using agglomeration (James and Matteson 2013). According to Matteson, the agglomerative approach is preferable to the divisive approach (Matteson and James 2013).

### 2.2 Wild Binary Segmentation (WBS)

This change point detection method claims to detect the exact number and potential locations of change points with no prior assumptions. The R package, *WBS*, is available with two methods, Standard Binary Segmentation (SBS) and Wild Binary Segmentation (WBS). According to Fryzlewicz (Baranowski and Fryzlewicz 2014), WBS overcomes the problems of the binary segmentation technique and is an improvement in terms of computation. It uses the idea of computing Cumulative Sum (CUSUM) from randomly drawn intervals considering the largest CUSUMs to be the first change point candidate to test against the stopping criteria (Fryzlewicz 2013). This process is repeated for all the samples.

### 2.3 Bayesian analysis of change points (BCP)

This algorithm works on a product partition model and computes the change point based on data and current partition. Tuning parameter values are provided for accurate analysis of change points (Erdman and Emerson 2007). The Markov chain Monte Carlo (MCMC) implementation of Bayesian analysis can be completed in  $O(n^2)$  time. We used the R package, “*bcp*,” for our experiments (Erdman and Emerson 2007).

### 2.4 Robust detection of change point

This methods works well when outliers are present, but it prompts possible false alarms as it works on “hierarchical binary splitting” techniques. Least square regression does not perform well in the presence of outliers. However, robust detection overcomes that problem (Gerard-Merchant et al. 2008). There is no R package available; so, we ran Iterative Robust Detection method using a Python implementation, provided by Stooksbury (Gerard-Merchant et al. 2008).

### 3 Materials and methods

To demonstrate the performance and effectiveness of change point detection methods, we generated synthetic data with 1000 data points. The reason for generating artificial data is that we can position in advance the number and locations of change points, which greatly simplifies our experiment. For combined trend and change detection, we generated 500 data points. The change points are generated from Eq. (1).

$$X_t = \mu + \delta_t + \varepsilon_t \quad (1)$$

where  $\mu$  = mean of series,  $\varepsilon_t$  = random error, and  $\delta_t$  = the point in time where the shift occurs

Trend is generated from Eq. (2).

$$X_t = \mu + \delta_t + \varepsilon_t + \beta_t \quad (2)$$

here  $\beta_t$  is the added trend.

### 4 Experimental results

An artificially generated data set of 1000 data points was generated with Eq. (3) (Liu et al. 2013).

- Simulation 1: Data points are generated using

$$y(t) = 0.7y(t-1) - 0.4y(t-2) + \varepsilon \quad (3)$$

where  $\varepsilon$  is a noise with varying  $\mu$  value and constant  $\sigma^2$ . A change point is inserted at steps 100 ( $\mu = 1.5$ ,  $\sigma^2 = 1.5$ ), 200 ( $\mu = 2.5$  and  $\sigma^2 = 1.5$ ), 300 ( $\mu = 1$ ,  $\sigma^2 = 1.5$ ), 200 ( $\mu = 2.5$ ,  $\sigma^2 = 1.5$ ), and 200 ( $\mu = 1.5$ ,  $\sigma^2 = 1.5$ ), and the locations of change points are 100, 300, 600, 800, and 1000.

Analysis: The results from Simulation 1 in Table 1a show that all the methods are capable of detecting change in the mean of data. In order to test the computational efficiency, we investigated WBS, BCP, and E-Agglomerative methods using 100 random iterations, as shown in Table 1b.

In Table 1b, 10 and 25 % represent variation from data points considered to be near the actual change points, “missed” represents the number of times actual change points locations are missed, and “extra” represents the number of times phantom change points to be detected. From our experiment, we observed that WBS appears to have the best combination of accuracy and speed. BCP, because of its sensitivity, generates unnecessary locations and numbers. E-Agglomerative produces accurate locations and numbers, but it is computationally cumbersome.

- Simulation 2: A change point is added at every 250 steps between 1 and 1000. Finally, noise with mean 0 and variance 4 is added. The change points are at steps 250, 500, and 750.

**Table 1** Change in mean

Method	Locations				
(a) Change point locations					
WBS	99, 299, 601, 808				
Divisive					
$\alpha = 1$	100, 300, 602, 809				
$\alpha = 2$	100, 300, 602, 809				
BCP	99, 298, 601, 808				
Robust detection	100, 300, 602, 808				
Method	10 %	25 %	Missed	Extra	Time (s)
(b) Results from 100 iterations					
WBS	67	97	8	1	32.53
Divisive					
$\alpha = 1$	70	91	11	10	
$\alpha = 2$	72	92	7	5	36,769.08
BCP	80	100	10	80	270.53

**Table 2** Step data with constant mean and variance

Method	Locations				
(a) Change point locations					
WBS	250, 501, 758				
Divisive					
$\alpha = 1$	251, 502, 759				
$\alpha = 2$	251, 502, 759				
BCP	240, 501, 752				
Robust detection	250, 501, 758				
Method	10 %	25 %	Missed	Extra	Time(s)
(b) Results from 100 iterations					
WBS	86	98	22	10	35.78
Divisive					
$\alpha = 1$	84	97	4	5	27,350.66
$\alpha = 2$	85	98	3	4	
BCP	80	87	33	27	267.2

Analysis: The results in Table 2a from Simulation 2 show that all the methods are capable of detecting step change in the data, while Table 2b shows the results from 100 iterations of WBS, BCP and E-Agglomerative methods. In Simulation 2, WBS appears again to be computationally faster over BCP and E-Agglomerative.

- Simulation 3: In this test, data points are generated, from Eq. (3). A change point is added at 200, 300, 400, and 100 steps. The noise is added with mean 0 and

different variance at steps 200 ( $\mu = 1.5$  and  $\sigma^2 = 2.5$ ), 300 ( $\mu = 1.5, \sigma^2 = 3.5$ ), 400 ( $\mu = 1.5, \sigma^2 = 4.5$ ), and 100 ( $\mu = 1.5, \sigma^2 = 5.5$ ), and locations are 200, 500, 900, and 1000.

Analysis: The results, in Table 3a, from Simulation 3 show that none of the methods are able to detect change in the variance of data. While performing this experiment, we tried to change the variance value, but we encountered no method has the ability to detect change in the variance. In order to reduce the computation time for E-Agglomerative method, we parallelized the code using “doParallel” package in R.

- Simulation 4: In this test, data points are generated using Eq. 4 where  $w$  is randomly generated data and  $\varepsilon$  is noise with  $\sigma^2$  0.8.

$$y(t) = \sin(w) + \varepsilon \tag{4}$$

Analysis: The change points are added at every 250 steps, i.e., at location 250, 500, and 750. The results, in Table 4a, from simulation 4 show that robust detection is not able to detect change points. The results, in Table 4b, show that WBS again outperforms.

### 5 Trend and change point interaction

#### 5.1 Trend masked change point

In order to assess the interaction between trend and change point, we generated data points from Eq. (2). Only 500 data points are generated for simplicity, and they clearly show

**Table 3** Change in variance

Method	Locations				
(a) Change point locations					
WBS	No change point detected				
Divisive					
$\alpha = 1$	574				
$\alpha = 2$	No change point detected				
BCP	No change point detected				
Robust detection	No change point detected				
Method	10 %	25 %	Missed	Extra	Time (s)
(b) Results from 100 iterations					
WBS	8	20	81	59	32.71
Divisive					
$\alpha = 1$	4	10	99	4	14,009.17
$\alpha = 2$	1	2	11	6	
BCP	NA	NA	NA	NA	221.1

**Table 4** Sine wave data

Method	Locations				
(a) Change point locations					
WBS	261, 501, 749				
Divisive					
$\alpha = 1$	262, 502, 750				
$\alpha = 2$	262, 500, 749				
BCP	262, 500, 749				
Robust detection	No change point detected				
Method	10 %	25 %	Missed	Extra	Time (s)
(b) Results from 100 iterations					
WBS	83	96	19	4	36.24
Divisive					
$\alpha = 1$	85	94	9	12	20,768.36
$\alpha = 2$	88	95	9	5	
BCP	60	70	70	90	221.1

the relations between change point and trend. Figure 2 shows the presence of change point and trend in the series.

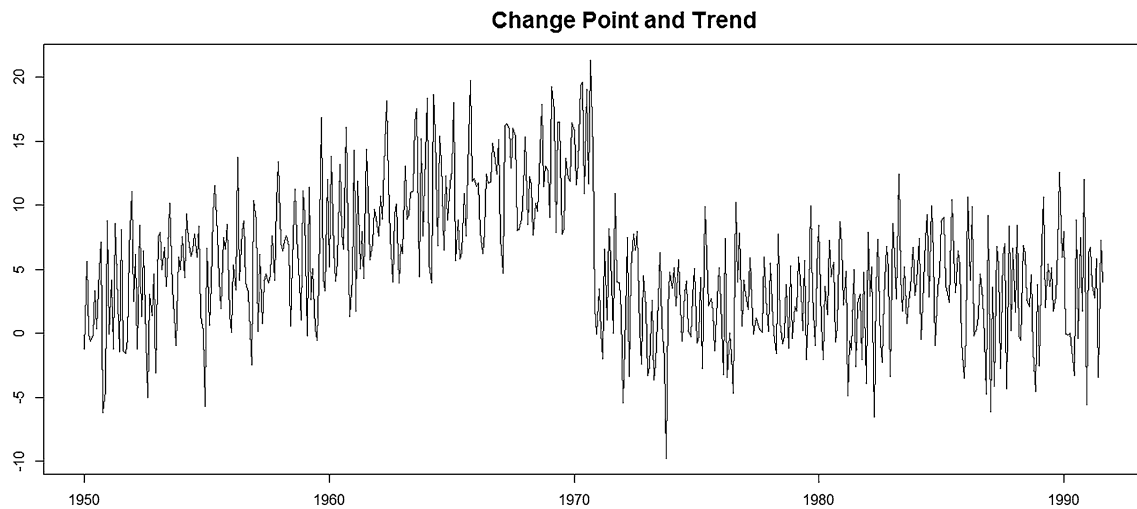
For assessing the relationship, following steps have been taken:

1. detect the presence of change points
2. analyze the data set for trend
3. divide the data into segments based on change points
4. analyze those segment for the presence of trend

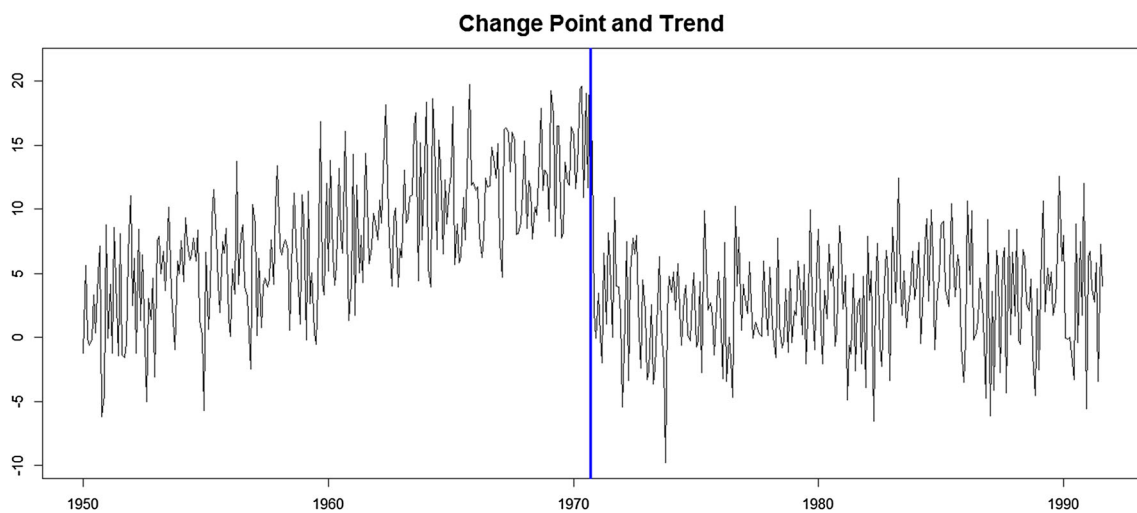
We used the WBS and BCP methods for detection of change points in our artificial data set. According to these methods, a change point appeared at location 250 as shown in Fig. 3.

Similarly, we analyze the same data set for trend as shown in Fig. 4. For both Mann–Kendall and Cox–Stuart method, a downward trend is predicted when the complete data set is analyzed. If we consider this as the final conclusion, it is a misleading result. When we divide the data set into segments and test them before and after the change point for trend analysis, we get the actual picture.

Since there is one change point at 250, we see that it divides the series into two parts at that point. When we analyze the first part, with both Mann–Kendall and Cox–Stuart, an upward trend is predicted, as shown in Fig. 5a, whereas the second half shows no trend in the data set, in the Fig. 5b. Using repeated Monte Carlo simulations, we found that 72 % of times the trend has affected the location of change points. We used piecewise LOWESS curve as shown in Fig. 6. From this, it is clear that there was an upward trend before the change point. Due to some abrupt change, where there could be a change point, the trend changes.



**Fig. 2** Change point and trend



**Fig. 3** Change point location

## 5.2 Change point masked trend

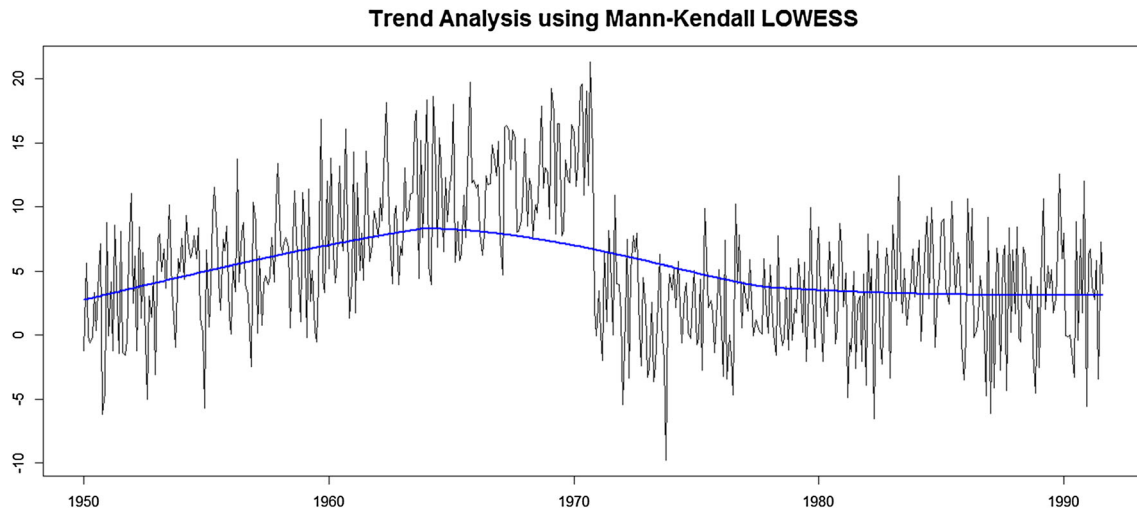
In order to test whether change point can mask trend in the data set, we simulated data points with a steep rise, followed by decay to prior consideration. This situation could occur in hydrological time series when a large amount of water is suddenly released and the system returns to previous flow level. Such a situation initially shows an upward trend and then downward trend. We used the steps as in Sect. 5.1. Initially, we examined the number of change points in the time series (Fig. 7). Analysis shows a series of artificial change points. We also tested our data for trend.

By itself, Mann–Kendall could not detect trend in the time series. According to the step 4 in Sect. 5.1, we tested the data before and after each change point. Figure 8 shows a typical result trend, change points, and piecewise trend

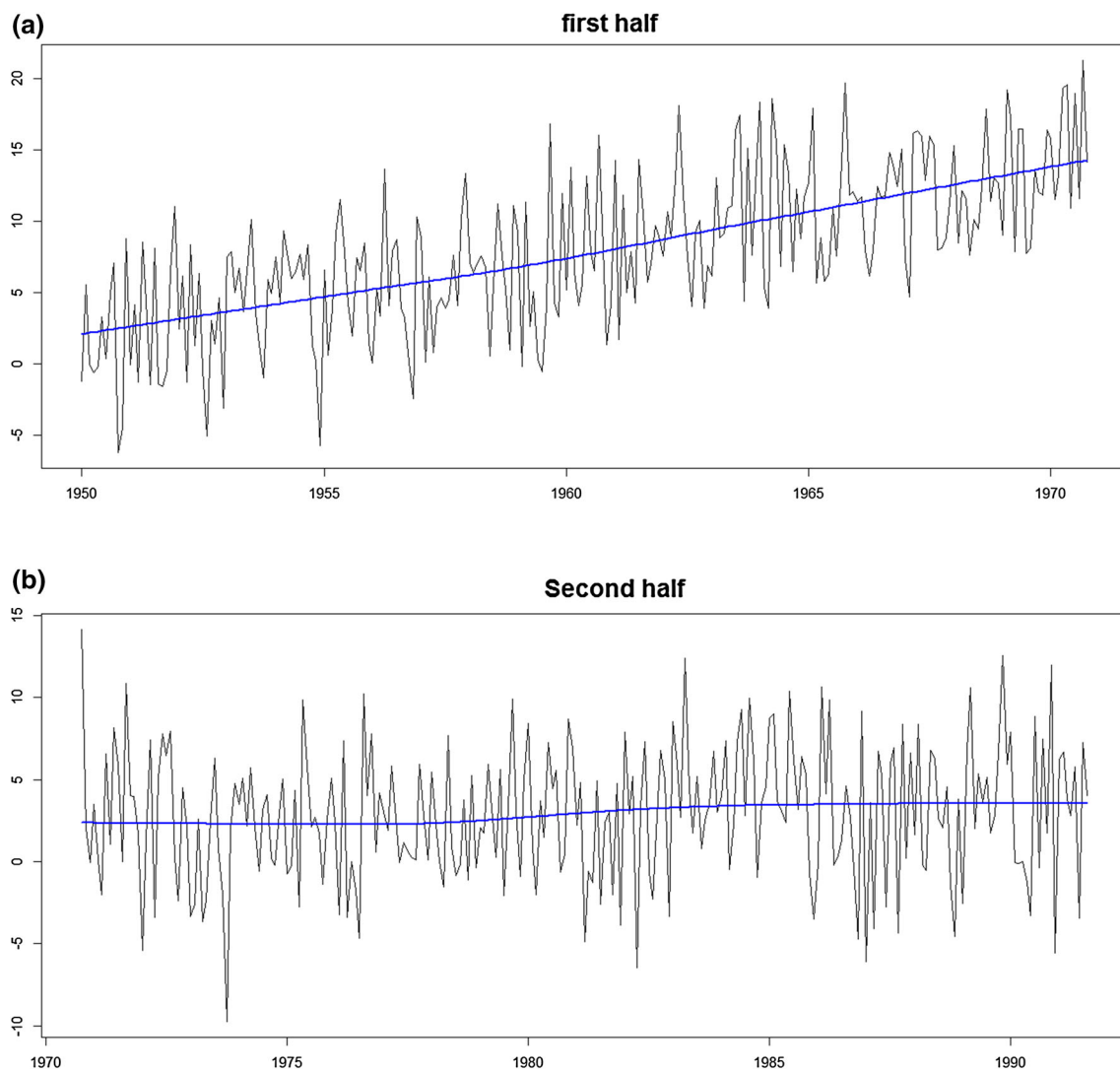
with LOWESS and piecewise LOWESS. Using the Monte Carlo simulations, we found that in 78 % of cases the change points have marked the change in the direction of the trend.

## 6 Conclusions

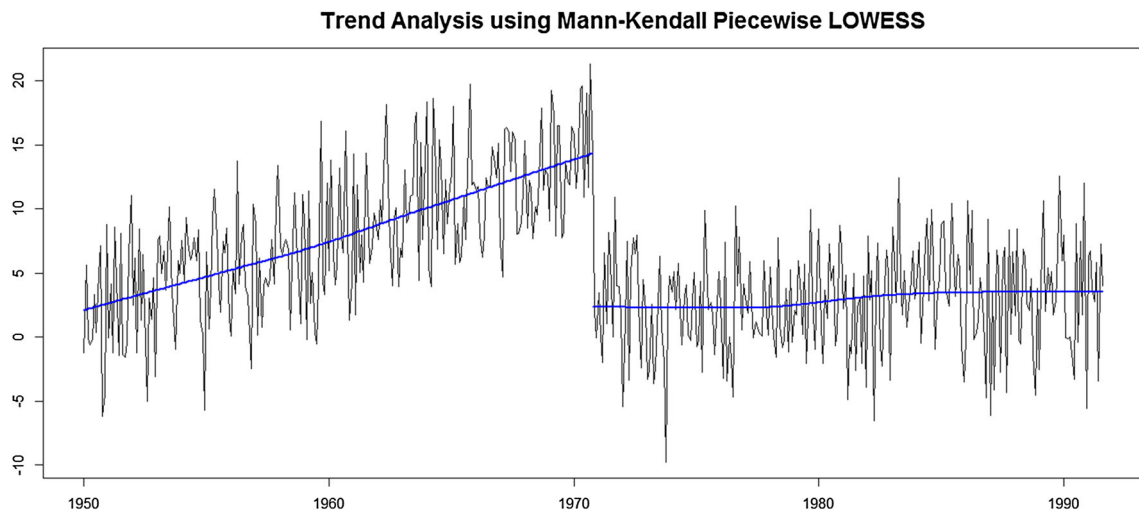
In this paper, we compared four different change point detection methods using an artificial data set. All of the methods fail to detect the change point location for the change in variance alone. From our experiments, we found that BCP is very sensitive and detects more change points other than the correct ones. However, E-Agglomerative accurately identifies correct locations and numbers of change points, but is computationally expensive. Our



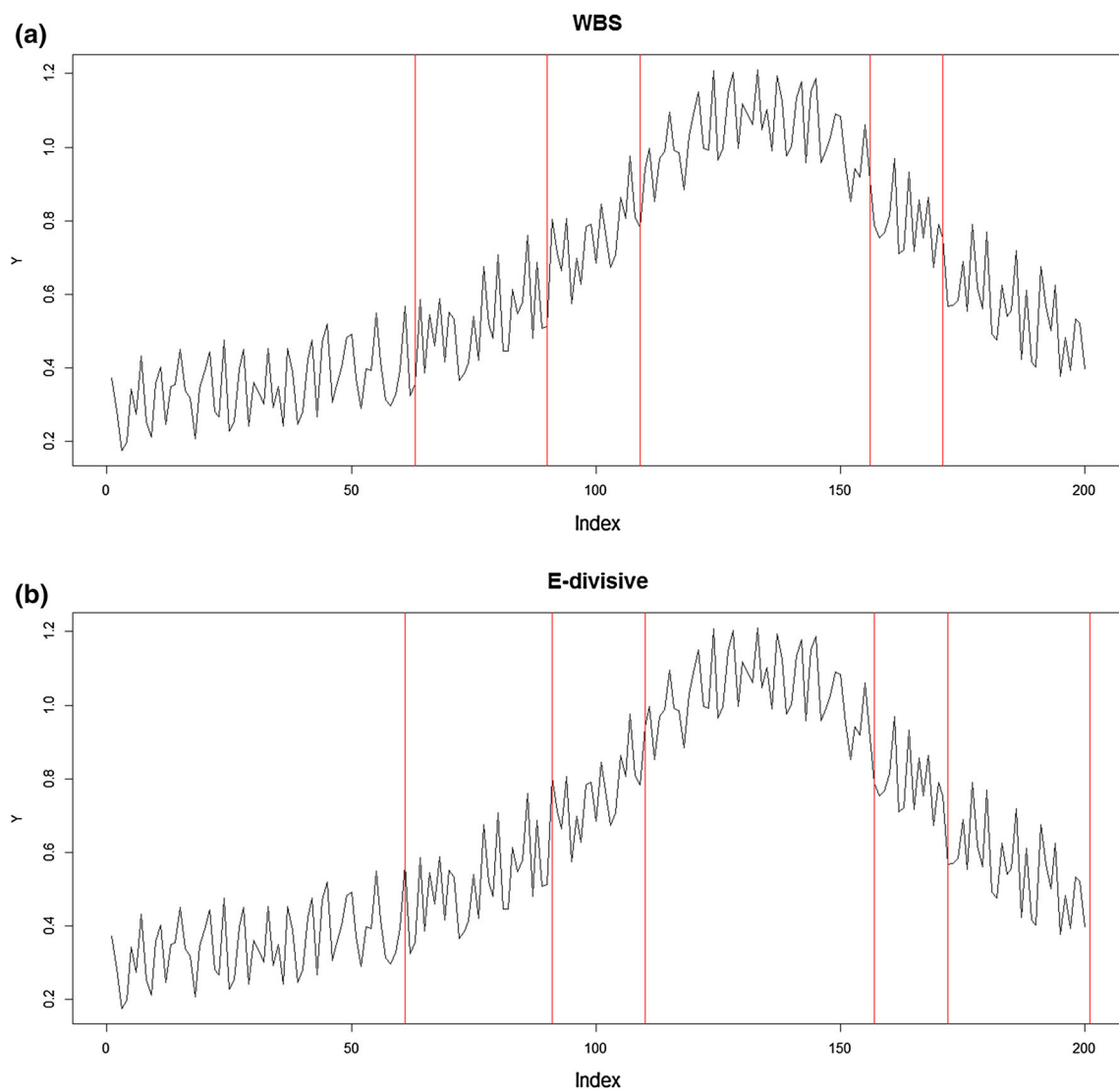
**Fig. 4** Trend for complete series



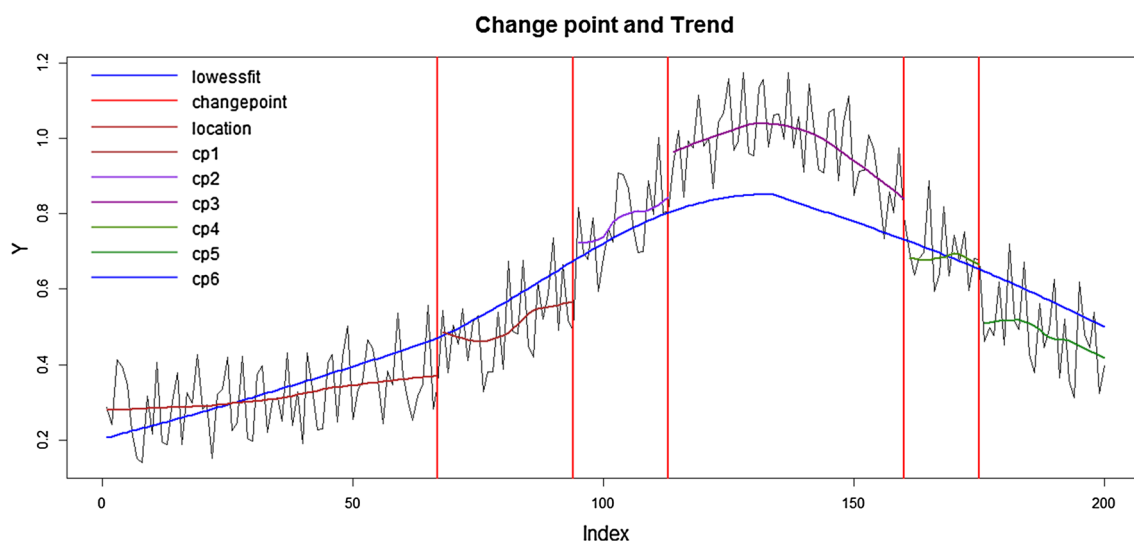
**Fig. 5** Trend analysis before (a) and after (b) change point



**Fig. 6** Trend analysis before and after change point



**Fig. 7** Change point using *WBS* (a) and *E-Divisive* (b)



**Fig. 8** Change point, trend, and piecewise trend

experimental results imply that WBS is superior to other methods: It reliably identifies the correct number and locations of change points, and it takes less computational time for identification. We analyzed whether the change point methods are situation dependent, and we provided different experiments to challenge change point detection. We have tried to see whether change point masks the trend detection or vice versa. Our experiment shows that if both are present in the data set, analyzing only trend or change point may mislead the results. The situation potentially exists when changes in one phenomenon may affect the other.

## References

- Adams RP, MacKay DJ (2007) Bayesian online changepoint detection. University of Cambridge, UK
- Baranowski R, Fryzlewicz P (2014) wbs: wild binary segmentation for multiple change-point detection. R package version 1.1
- Erdman C, Emerson JW (2007) bcp: an R package for performing a Bayesian analysis of change point problems. *J Stat Softw* 23(3):1–13
- Fryzlewicz P (2013) Wild binary segmentation for multiple change-point detection. Department of Statistics, London School of Economics, London
- Gerard-Merchant PG, Stooksbury DE, Seymour L (2008) Methods for starting the detection of undocumented multiple changepoints. *J Clim* 21:4887–4899
- Hamed KH (2008) Trend detection in hydrologic data: the Mann Kendall trend test under the scaling hypothesis. *J Hydrol* 349:350–363
- Hessa A, Iyera H, Malm W (2001) Linear trend analysis: a comparison of methods. *Atmos Environ* 35:5211–5222
- James NA, Matteson DS (2013) ecp: an R package for nonparametric multiple change point analysis of multivariate data. ArXiv e-prints
- Liu S, Yamada M, Collier N, Sugiyama M (2013) Change-point detection in time-series data by relative density-ratio estimation. *Neural Netw* 43:72–83
- Matteson DS, James NA (2013) A nonparametric approach for multiple change point analysis of multivariate data. *J Am Stat Assoc* 109(505):334–345. doi:10.1080/01621459.2013.849605
- Partal T, Kahya E (2006) Trend analysis in Turkish precipitation data. *Hydrol Process* 20:2011–2026
- Reeves J, Chen J, Wang XL, Lund R, Lu Q (2007) A review and comparison of changepoint detection techniques for climate data. *J Appl Meteorol Climatol* 46(6):900–915
- Xiong L, Guo S (2004) Trend test and change-point detection for the annual discharge series of the yangtze river at the Yichang hydrological station. *Hydrol Sci J* 49(1):99–112
- Yue S, Pilon P, Cavadias G (2002) Power of the Mann–Kendall and Spearman’s rho tests for detecting monotonic trends in hydrological series. *J Hydrol* 259:254–271