Stochastic Environmental Research and Risk Assessment (2019) 33:1751–1764
https://doi.org/10.1007/s00477-019-01727-6

ORIGINAL PAPER

Moving correlation coefficient-based method for jump points detection in hydroclimate time series

Ziyi Wu[1] · Ping Xie[1,2] · Yan-Fang Sang[3] · Jie Chen[1] · Wei Ke[1] · Jiangyan Zhao[1] · Yuxi Zhao[4]

Published online: 6 September 2019

Abstract

The jump points detection is critical to the understanding of hydrologic variability, especially in investigating the anthropogenic effects. Conventional methods are mainly statistical and cannot directly reflect the jump change degrees. This article proposes a moving correlation coefficient-based detection (MCCD) method for the detection of jump points (JPs) in hydroclimate data. The correlation coefficient (CC) between the potential jump component and the original data is calculated, and the CC series is realized by moving from the starting to the ending points of the original time series. Bigger CC value reflects higher jump degree; the position with the biggest absolute CC value is the JP that is the most expected. Its significance is evaluated by comparing its value with the CC threshold value at the relevant significance level. Monte-Carlo experimental results verify the MCCD method's higher efficiency compared with four commonly used conventional methods. It is especially noteworthy that the results indicate its stable efficiency, even when encountering the influences of some unfavorable factors. By applying the MCCD method to the Lancang River Basin, the JP of runoff in 2004 is detected at the Yunjinghong station in the lower reach. It is mainly attributed to the construction and operation of some major water hydropower projects, while the stable variations of areal precipitation and actual evapotranspiration, as well as the stable land-cover conditions, contribute little to the abrupt decrease in runoff. The MCCD method can be an effective alternative for the detection of JPs in hydroclimate data.

Keywords Jump point · Correlation analysis · Significance evaluation · Hydroclimatic process · Upper-Mekong River

1 Introduction

As an important type of abrupt change, a jump is a phenomenon frequently encountered in hydroclimate studies. It usually indicates some physical process that causes an abrupt switch or change from one behavior to another. For example, the eruptions of El Chichón and Mt. Pinatubo caused rapid stratospheric warmings over the last decades (Gagné et al. 2017; Karpechko et al. 2010). Anthropogenic effects, including land use and cover change, dam constructions and hydropower operations, often cause rapid changes of water levels and streamflow in many rivers (Cockburn and Garver 2015; Furey et al. 2010; Şen 2017). The detection of trends and periodicities in hydroclimate data could misrepresent the true underlying behavior, as the results would be destroyed by jump points and be apparently contradictory. How to detect the jump points (JPs) is therefore a critical step as part of analyzing hydroclimate

✉ Ping Xie
pxie@whu.edu.cn

✉ Yan-Fang Sang
sangyf@igsnrr.ac.cn

1 State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China

2 Collaborative Innovation Center for Territorial Sovereignty and Maritime Rights, Wuhan 430072, China

3 Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

4 Southwest Branch of State Grid, Chengdu 610000, China

variability (Li et al. 2016; Perreault et al. 1999; Yang et al. 2009).

Dozens of methods have been proposed for detecting JPs in observed hydroclimate data. Based on different mathematical hypotheses and scopes of application, these methods can be divided into two types: parametric and non-parametric tests (Ahmadi et al. 2018; Sang et al. 2012; Zhang et al. 2013). Some priori distribution information must first be estimated from the observed data when using the parametric tests, which limits their application (Kundzewicz and Robson 2004). In contrast, non-parametric tests do not require any priori probability distribution information and thus they perform better (Domonkos 2013; Reeves et al. 2007). A jump point is a part of the overall hydrologic variability, thus its significance is determined not only by its own change magnitude, but also by its proportion in the whole hydroclimate data. Besides, the sample size, dispersion degree and other factors would influence the detection of JPs and the evaluation of their significance (Montanari 2012; Villarini et al. 2009). Conventional methods, however, are mainly statistical but can not clarify these problems, and thus do not contribute to the understanding and detection of JPs. They perform differently and cannot adequately meet practical hydroclimatic needs (Zhou et al. 2011; Rougé et al. 2013). How to accurately detect JPs in a hydroclimate data is still a challenge.

Considering that the significance of a jump is mainly determined by its proportion to a complete data set, the correlation coefficient (CC) between the jump component and the original data may be an effective index to quantify their significance. The correlation coefficient can be mathematically determined through hypothesis testing and is easily calculated (Wikle 2003). The objective of this study is therefore to propose a CC-based method for the detection of JPs in hydroclimate data. We deduce the mathematical relationship between the CC and the JPs of a time series and utilize that relationship in our CC-based method. We evaluate our method's performance and compare it with that of traditional methods for verification. We then apply this method to investigate the abrupt changes in the precipitation–runoff processes in the Upper-Mekong River basin. Our conclusions and recommendations for further work are presented in the last section.

2 Moving correlation coefficient detection method proposed

In order to develop the moving correlation coefficient-based method for the detection of JPs, the correlation between the jump component and original data is first deduced. Based on the principle of linear superposition in stochastic hydrology (Ding and Deng 1988), if the hydroclimate time series X_t has a JP, the original series can be divided into two subseries, with length of n_1 (before the JP) and n_2 (after the JP). The jump component can be described as:

$$Y_t = \begin{cases} 0, & t \leq n_1 \\ q, & n_1 < t \leq n \end{cases} \tag{1}$$

where q is the difference of the mean value between the two subseries.

The residual of the series X_t is noted as S_t, and $X_t = Y_t + S_t$. The original series X_t can be described as:

$$X_t = \begin{cases} S_t, & t \leq n_1 \\ q + S_t, & n_1 < t \leq n \end{cases} \tag{2}$$

The mean value of S_t is noted as $\overline{S}$, and that of the jump component Y_t can be expressed as:

$$\bar{Y} = \frac{n_2 q}{n} \tag{3}$$

The mean value of the original series X_t can then be described as:

$$\bar{X} = \frac{n_1 \bar{S} + n_2(q + \bar{S})}{n} \tag{4}$$

The correlation coefficient r between the X_t and Y_t series is calculated as (Murphy and Myors 2004):

$$r = \frac{\sum X_t Y_t - n\bar{X}\bar{Y}}{\sqrt{\sum X_t^2 - n\bar{X}^2}\sqrt{\sum Y_t^2 - n\bar{Y}^2}} \tag{5}$$

By substituting Eqs. (3) and (4) into Eq. (5), a new equation of r is described as:

$$r^2 = \frac{n_1 n_2 q^2}{n^2 \sigma_X^2} \tag{6}$$

where $\sigma_X^2 = \frac{1}{n}\sum X_i^2 - \bar{X}^2$.

The variance of Y_t can then be described as:

$$\begin{aligned} \sigma_Y^2 &= \frac{n_1\left(0 - \frac{n_2 q}{n}\right)^2 + n_2\left(q - \frac{n_2 q}{n}\right)^2}{n} \\ &= \frac{n_1 \frac{n_2^2 q^2}{n^2} + n_2 \frac{n_1^2 q^2}{n^2}}{n} \\ &= \frac{n_1 n_2 q^2}{n^2} \end{aligned} \tag{7}$$

By substituting Eq. (7) into Eq. (6), r^2 can be described as:

$$r^2 = \frac{\sigma_Y^2}{\sigma_X^2} \tag{8}$$

The σ_Y^2 is determined by the jump component q, and since Eq. (8) intuitively indicates the positive relationship between the index r and σ_Y^2, a bigger r value indicates a more significant jump in the original series X_t.

Furthermore, From Eq. (2) it is assumed that Y_t and S_t are independent of each other, so σ_X^2 can be expressed as:

$$\sigma_X^2 = \sigma_Y^2 + \sigma_S^2 \tag{9}$$

The correlation coefficient can be described as:

$$\begin{aligned} r^2 &= \frac{\sigma_Y^2}{\sigma_Y^2 + \sigma_S^2} \\ &= \frac{1}{1 + \frac{\sigma_S^2}{\sigma_Y^2}} \end{aligned} \tag{10}$$

Equation (10) shows that when the jump component occupies a bigger proportion than residual components, the correlation coefficient r has a larger value. Thus, it can be an indicator to quantify the JPs significance in hydroclimate time series.

Based on the information conveyed in Eq. (10), we propose a new method for detecting JPs in hydroclimate data, the moving correlation coefficient-based detection (MCCD) method. The seven steps of this this method are detailed below:

Step 1. Set a starting point $k_0 (k_0 > 1)$ and assume it functions as a JP and then divide the hydroclimate time series $X_t (t = 1, 2, \ldots, n)$ into two subseries at that JP so they can be analyzed;

Step 2. Calculate the mean values of the two subseries before and after the JP, and obtain the jump component $Y_{t\,k_0}$ by Eq. (1);

Step 3. Calculate the correlation coefficient r_{k_0} between the jump component $Y_{t\,k_0}$ and the original series X_t using Eq. (5);

Step 4. Set $k_{n+1} = k_n + 1$, and repeat steps 2–3 to obtain the series $Y_{t\,k_1}$; then, find the maximum value $|r_{\max}|$ in the series r_k;

Step 5. Consider the degree of freedom (DOF) of the series X_t and the relevant significance level α, and then determine the threshold r_α value with which to assess the statistical significance of the JP. The present study considers a 5% significance level;

Step 6. Compare $|r_{\max}|$ with the threshold r_α value. If $|r_{\max}| > r_\alpha$, the JP is viewed as significant at significance level α; otherwise, the JP is not viewed as significant;

Step 7. The first jump component that is the most significant can be removed from the original series, and then new series $X_t' = X_t - Y_{t\,k_{\max}}$ obtained. Steps 1–6 may be repeated to identify more significant JPs if they are existing until all JPs are identified.

The procedure of the MCCD method is presented in Fig. 1. It should be noticed that that if the starting or ending point was close to either end of the series X_t, the sample size of the shorter subseries would be too small to achieve reliable results. Therefore, in this study the data points from 11 to n-10 are considered to detect the JPs in hydroclimate data.

3 Verification of the developed MCCD method through statistical experiments

A set of statistical experiments were designed and used to verify the efficiency of the MCCD method for JP detection. These were also used to investigate the influence of some of the main factors that affect JP detection: data length n, the mean value EX, the variation coefficient Cv and the skewness coefficient Cs of a series. Because a JP is represented by both its jump degree q and position a (calculated by n_1/n), these two parameters are also considered ("Appendix 1"). The statistical experiments are designed as follows:

1. Set EX = 500, Cv = 0.5, and Cs = 2.0. For the jump component, q = 300 and the JP is set at the middle of the series (a = 0.5). Set n = 50,100,150,200,250,300,350,400,450 and 500 to investigate its influence on the detection of JPs.
2. Set n = 200, Cv = 0.5, and Cs = 2.0. For the jump component, set q = 300 and a = 0.5. Set EX = 100,300,500,700 and 900 to investigate its influence on the JP detection.
3. Set n = 200, EX = 500, and Cs = 2.0. For the jump component, set q = 300 and a = 0.5. Vary Cv from 0.1 to 0.9 in increments of 0.1 to investigate its influence on JP detection.
4. Set n = 200, EX = 500, and Cv = 0.5. For the jump component, set q = 300 and a = 0.5. Change Cs from 1.0 to 3.0 in increments of 0.5 to investigate its influence on the detection of JPs.
5. Set n = 200, EX = 500, Cv = 0.5 and Cs = 2.0. For the jump component, a = 0.5. Vary q by 50 from 100 to 500 to investigate its influence on the detection of JPs.
6. Set n = 200, EX = 500, Cv = 0.5, and Cs = 2.0. For the jump component, set q = 300. Vary a by 0.1 from 0.1 to 0.9 investigate its influence on the detection of JPs.

For each group of statistical experiments, we generate 100 synthetic series with known JPs using the Monte-Carlo method (Veihe and Quinton 2015), and use the MCCD method to detect their JPs. Two typical non-parametric methods, the Pettitt test (Pettitt 1979) and the Mann–Kendall (M–K) test (Ahmadi et al. 2018; Mann 1945), and two common parametric methods, the Brown–Forsythe (B–F) method (Brown and Forsythe 1974) and the Bayesian method (Nasseribrahim et al. 2005), are also used f here ("Appendix 2"). As these four conventional methods are

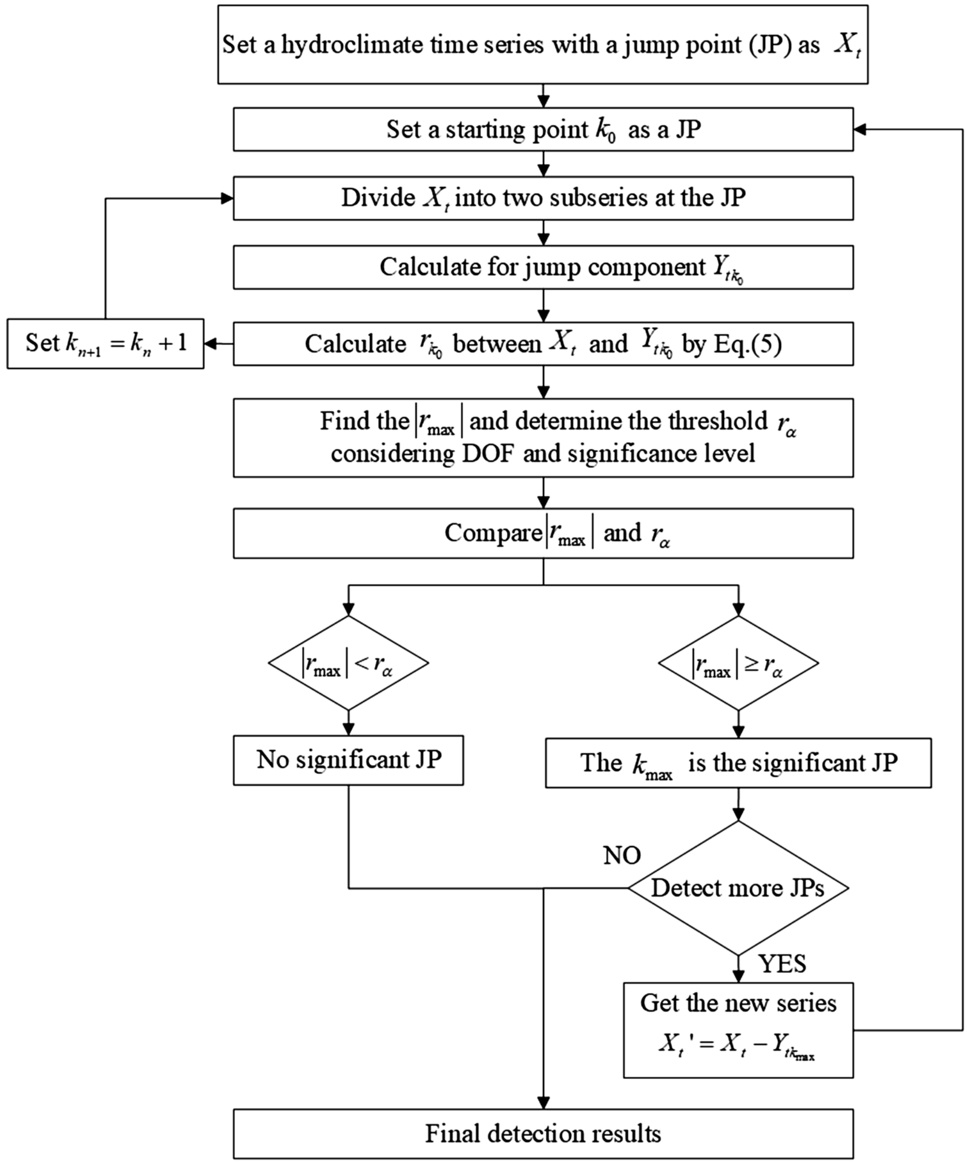

Fig. 1 Steps for jump point (JP) detection in hydroclimate data using the moving correlation coefficient-based detection (MCCD) method proposed in Sect. 2. "DOF" indicates degree of freedom

used widely and perform well, their results can be compared with the MCCD method for verification.

Comparing the JPs detected by different methods and the preset JPs, we can use the indicator of efficiency coefficient η $(0 \leq \eta \leq 100\%)$ to evaluate the performance of each method:

$$\eta = \frac{d}{D} \times 100\% \tag{11}$$

where d indicates the acceptable detection times and D represents the total detection times, here D = 1000. An acceptable detection time means that the difference between the identified JPs and the preset JPs should not be larger than 1.

We repeated the above steps 100 times and obtained the η_{final} from each group of experiments:

$$\eta_{final} = \frac{\sum \eta}{100} \tag{12}$$

The evaluation results are shown in Fig. 2. They indicate that the MCCD method performs better overall than the M–K test, but worse than the Pettitt test, and that it has a similar performance to that of the Bayesian method and the B–F method.

When the data length n is increased the efficiency of the M–K test goes down, but that of the B–F test slightly increases. The Pettitt test, Bayesian method and MCCD method perform stably and are little influenced by the data length n (Fig. 2a). The performance of all five methods becomes worse when EX is increased (Fig. 2b). This phenomenon is likely attributed to the fact that when the value of EX increases, the relative jump degree is less and so the corresponding JP becomes more difficult to identify. The

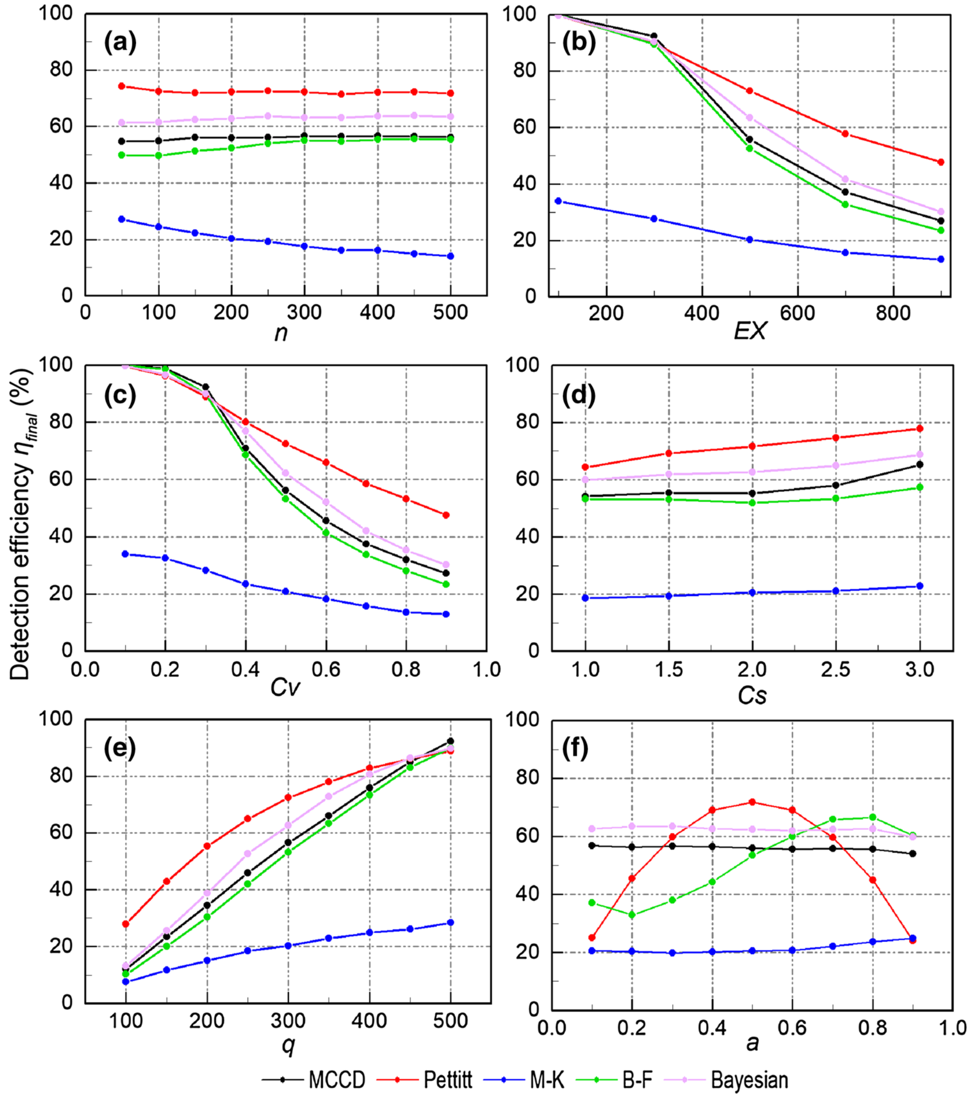

Fig. 2 Detection efficiencies of the moving correlation coefficient-based detection method, the Pettitt test, Mann–Kendall test, Brown–Forsythe method and the Bayesian method, and their variations with the changes of data length (n), mean value (EX), variation coefficient (Cv) and skewness coefficient (Cs) of a series, and with the changes of jump degree (q) and change point position (a)

performance of all five methods obviously decline with the increase of Cv (Fig. 2c). This implies that for those hydroclimate time series with higher variability degrees, the jump degrees would become weaker and so their JPs will be more difficult to identify. Meanwhile, the results indicate that the factor of Cs has little influence on the efficiencies ($\eta_{final} > 60\%$) of these methods (Fig. 2d).

Besides, the Pettitt test, the Bayesian method, the MCCD method and the B-F method all perform better with the increase of q, especially the MCCD method (Fig. 2e). That is, the more significant jumps are easier to identify. However, the position of the JP has a strong effect on the accuracy of the detection results. The Pettitt test performs much better at detecting the JP in the middle position of a series than in other positions (Fig. 2f), causing its higher efficiency in Fig. 2a–d where we set a = 0.5 (i.e., the JP is set at the middle of a series). The B–F method's efficiency becomes higher when a increases from 0.2 to 0.7, and then decreases to the end of the series. Meanwhile, the Bayesian method and the MCCD method maintain their high efficiency at any JP positions.

The above results show that the MCCD method performs well at detecting JPs in hydroclimate time series, even when encountering the influences of some adverse factors. It shows stable efficiency compared to four other commonly used methods. In addition, the correlation coefficient, as the core of the MCCD method, is easier to calculate than the calculations required for the other methods. The correlation coefficient directly reflects the proportion of the jump component in the original hydrologic data, and so it can also be an effective indicator for comparing the jump variation significance among different hydroclimate time series (Xie et al. 2018).

4 Case study

4.1 Study area and data

The hydroclimate data measured in the Lancang River Basin (LRB) are used to further verify the performance of the proposed MCCD method. The LRB is the up-reach of the Mekong River Basin, reaching from its source in the Tibet Plateau to the China–Myanmar border. Its long and narrow basin has a drainage area of 167,000 km^2 (Shi et al. 2013). Due to the plunge of 4700 km from the upper to the lower reaches, this basin is abundant in hydropower resources (Fan et al. 2015). It provides rich natural resources (Dugan et al. 2010) not only to China but also to South Asia. Under the influences of human activities over recent years (Lauri et al. 2012), especially the implementation of some large water conservancy projects, the natural hydrologic variability in the LRB has been changing. The jump points in observed hydrologic data have therefore become a hot issue and are being investigated in Mekong River studies.

There are four primary hydrologic stations in the main stream of the LRB: Liutongjiang (LIU), Gongguoqiao (GON), Jiajiu (JIA) and Yunjinghong (YUN) (shown in Fig. 3), with drainage areas of 78,391 km^2, 91,302 km^2 112,091 km^2 and 145,295 km^2, respectively, and runoff data measured at the four stations are used as our case study. The four sub-basins corresponding to these

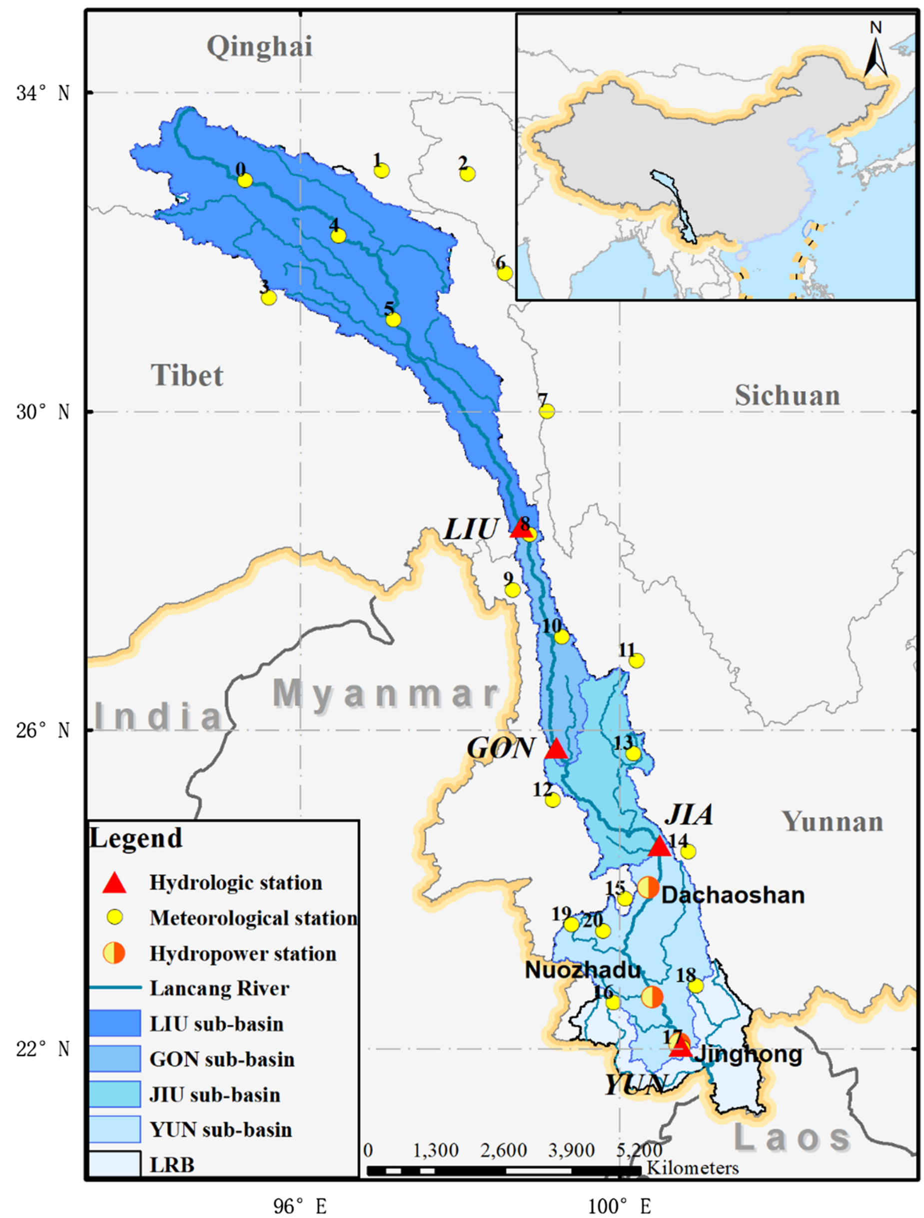

Fig. 3 Location of the Lancang River Basin (LRB) and its four sub-basins in China, and the distributions of hydrologic stations, meteorological stations and hydropower stations in and around the basin

Table 1 Areas and weights used for computing the annual area precipitation in four sub-basins (LIU, GON, JIA and YUN) of the Lancang River Basin

No.	Station	LIU		GON		JIA		YUN	
		Area (km^2)	Weight (%)	Area (km^2)	Weight (%)	Area (km^2)	Weight (%)	Area (km^2)	Weight (%)
0	Zaduo	24,067	30.70	24,067	26.36	24,067	21.47	24,067	16.56
1	Yushu	3139	4.00	3139	3.44	3139	2.80	3139	2.16
2	Shiqu	588	0.75	588	0.64	588	0.52	588	0.40
3	Dingqing	5588	7.13	5588	6.12	5588	4.99	5588	3.85
4	Nangqian	15,164	19.34	15,164	16.61	15,164	13.53	15,164	10.44
5	Changdu	27,882	35.57	18,914	20.72	18,914	16.87	18,914	13.02
6	Dege	1963	2.50	1114	1.22	1114	0.99	1114	0.77
7	Batang			7091	7.77	7091	6.33	7091	4.88
8	Deqin			4154	4.55	4154	3.71	4154	2.86
9	Gongshan			1611	1.76	1611	1.44	1611	1.11
10	Weixi			9872	10.81	7166	6.39	7166	4.93
11	Lijiang					2219	1.98	2219	1.53
12	Baoshan					9090	8.11	6812	4.69
13	Dali					12,186	10.87	10,122	6.97
14	Jingdong							5109	3.52
15	Lincang							7950	5.47
16	Lancang							5018	3.45
17	Jinghong							4535	3.12
18	Simao							7269	5.00
19	Gengma							2236	1.54
20	Shuangjiang							5429	3.74
	Total	78,391	100	91,302	100	112,091	100	145,295	100

21 Meteorological stations are considered, and the Thiessen Polygon Method is used. The Liutongjiang sub-basin is abbreviated as "LIU"; the Gongguoqiao sub-basin abbreviated as "GON"; the Jiajiu sub-basin as "JIA"; and the Yunjinghong sub-basin is abbreviated as "YUN"

hydrologic stations are also denoted as LIU, GON, JIA and YUN, respectively. Twenty-one meteorological stations are distributed in and around the whole LRB (as shown in Fig. 3). The annual area precipitation time series were obtained by using the Thiessen Polygon method (Bayraktar et al. 2005; Dessie et al. 2015), with the areas and weights shown in Table 1. The annual actual evapotranspiration E at each meteorological station is estimated using Fu's equation (Fu 1996; Sun 2007), which is suitable for mountainous areas:

$$E = E_0\left\{1 + \frac{p}{E_0} - \left[1 + \left(\frac{p}{E_0}\right)^m\right]^{1/m}\right\} \tag{13}$$

where E_0 is the potential evaporation, p is the precipitation and m is a parameter calculated by:

$$m = 1 + cp_d^a\left(\frac{1-Y}{Y}\right)^b \tag{14}$$

where p_d is the daily precipitation (mm/d), and Y is the runoff coefficient (the ratio between runoff depth and precipitation). For the LRB, the parameters of a, b, and c are set as 1.210, 0.393 and 0.293, respectively (Fu 1996). Again, the areal actual evapotranspiration time series in the sub-basins are obtained by using the Thiessen Polygon method. All the runoff, precipitation and evaporation data were analyzed to detect their JPs.

4.2 Detection of JPs in precipitation and runoff data

The MCCD method is used to detect the JPs in the annual precipitation and runoff series corresponding to the four sub-basins. The Pettitt test, Mann–Kendall test (M–K), Brown–Forsythe method (B–F) and the Bayesian method were again used for comparison. The JP detection results are summarized in Table 2 and represented in Fig. 4. These results show that the JPs vary with the methods used. For the four precipitation series, every method but t the M–K test produces similar results of no jump. This indicates that the precipitation variability in the LRB has remained stable over the last five decades. As for the runoff data, all five methods give the same results: no JP is identified in the LIU runoff time series, but a JP in about 2004 is identified

Table 2 Jump points identified in the annual area precipitation and runoff series of four sub-basins in the Lancang River Basin

Variable	Sub-basin	Year	MCCD	Pettitt	M–K	B–F	Bayesian
Precipitation	LIU	1961–2014	–	1984	1997	1997	–
	GON	1961–2014	–	–	1988	–	–
	JIA	1961–2014	–	–	1988	–	–
	YUN	1961–2014	–	–	1995	2004	–
Runoff	LIU	1987–2014	–	–	–	–	–
	GON	1955–2010	–	1986	1997	–	–
	JIA	1965–2014	–	2005	1979	2004	–
	YUN	1956–2014	2004	2001	–	2004	2004

5% Significant level is considered here. "–" means no jump point in the series

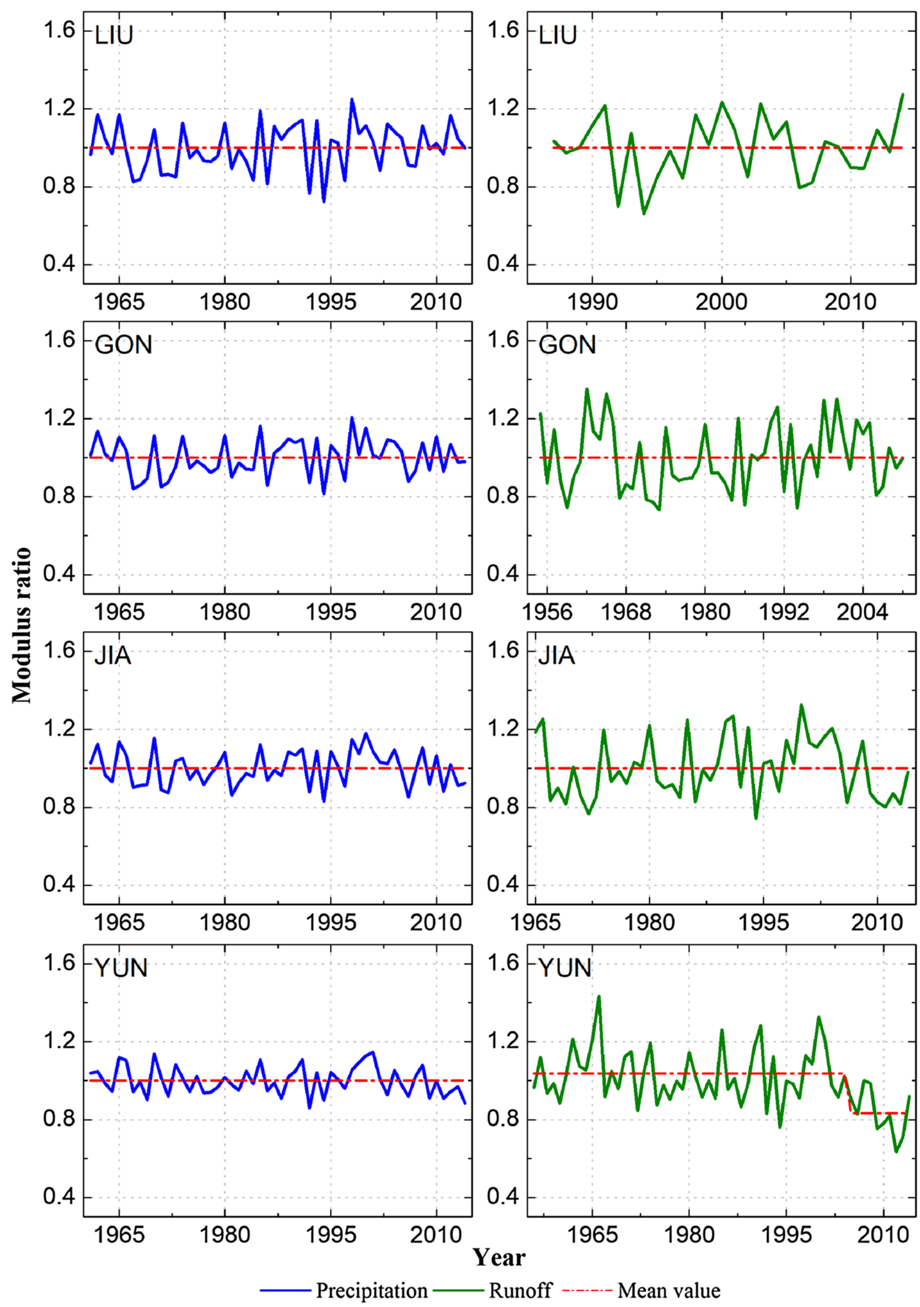

Fig. 4 Detection results of jump components in the precipitation and runoff time series (in modulus ratio) of four sub-basins (LIU, GON, JIA and YUN) in the Lancang River Basin. The modulus ratio is the ratio between the origin and mean value of a series with no unit

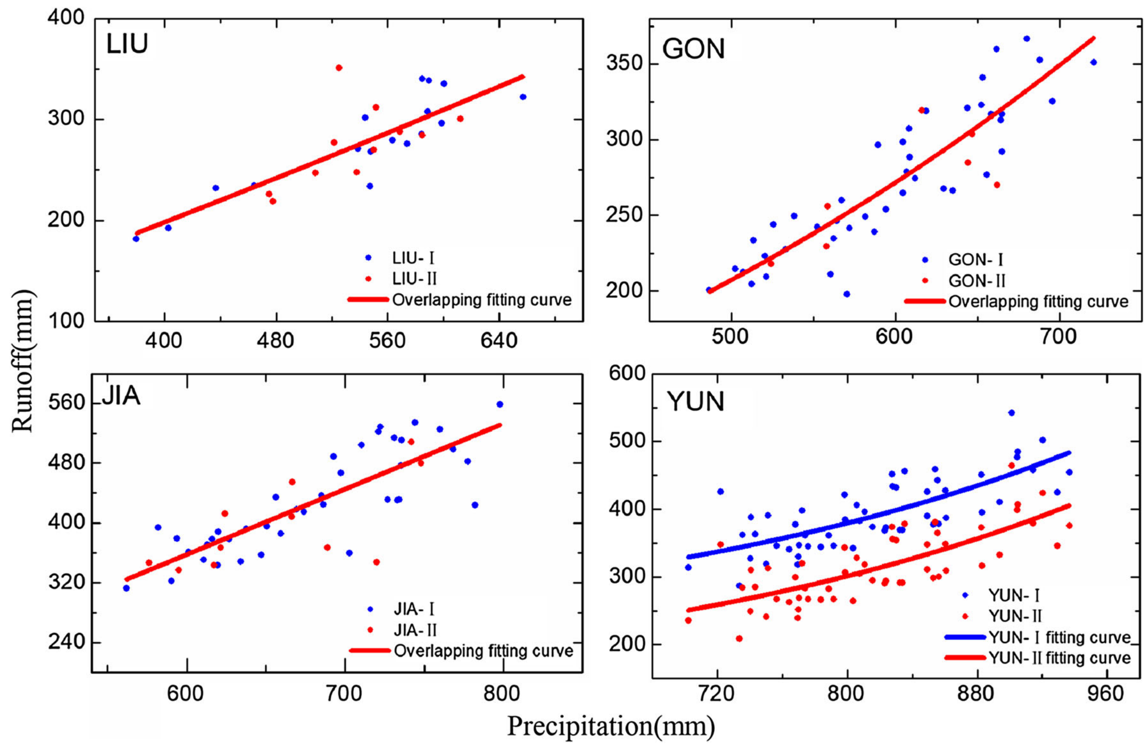

Fig. 5 Exponential fitting curves for describing the precipitation–runoff relationship in four sub-basins (LIU, GON, JIA and YUN) in the Lancang River Basin. Each series is divided into two parts: before the jump point in 2004 (I) and after 2004 (II)

in the YUN runoff time series. Figure 4 shows the slight decrease of annual runoff since 2004 at the YUN station, but indicates no changes at other stations. By considering Table 2 and Fig. 4 together, it can be perceived that the JPs detected by the MCCD and by the Bayesian methods are more reliable, but the results of the Pettitt, M–K and B–F methods have bias.

The jump in the precipitation–runoff relationship is further analyzed to investigate its changes from upstream to downstream in the LRB (see Fig. 5). Here the exponential curve is used to fit the precipitation–runoff relationship for each sub-basin thanks to its high fitting efficiency. Since the runoff series at the YUN station has a JP in 2004, it can be divided into two sections: a basic section (before the jump) and a changed section (after the jump). Considering that shorter sub-series would cause larger errors when analyzing the precipitation–runoff relationship, we restored the series according to the mean value difference before and after the JP in 2004. The restored YUN series before the JP (i.e., the original series) is denoted as the YUN-I series, while the series after the JP is denoted as the YUN-II series. We also divide the other three runoff series in 2004, but they give overlapping fitting curves because there are no JPs. Overall, it indicates that the precipitation–runoff relationship remains stable in the LIU, GON and JIA sub-basins, but there is a big difference before and after 2004 in the YUN sub-basin. A comparison of the two curves in the YUN sub-basin reveals that the runoff generation ability declines in the lower reach of the LRB after 2004.

The efficiency coefficient R^2 of each exponential curve is calculated (in Table 3).and the efficiency coefficient

Table 3 Efficiency coefficients (R^2) and exponential equations for describing the precipitation–runoff relationship in four sub-basins in the Lancang River Basin

Sub-basin	R^2 (%)	Equation
LIU	65.64	$Ru = -1487.40 + 67.49EXP((P + 9668.26)/3128.90)$
GON	79.71	$Ru = -109.03 + 11.11EXP((P + 1309.67)/540.33)$
JIA	68.39	$Ru = -24942.79 + 7371.42EXP((P + 35137.00)/28978.77)$
YUN	55.80	$Ru_{before} = 197.58 + 1.07EXP((P + 750.16)/301.64)$
	55.80	$Ru_{after} = 119.43 + 1.07EXP((P + 750.16)/301.64)$

P is precipitation and Ru is the corresponding runoff. For the YUN sub-basin, Ru_{before} (Ru_{after}) represents the precipitation–runoff relationship before (after) the runoff change in 2004

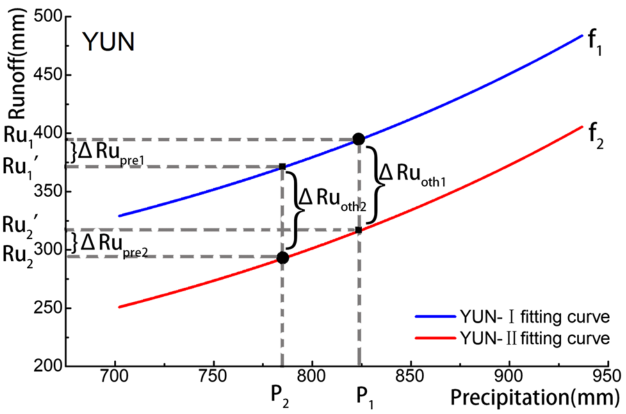

Fig. 6 Variables used to calculate the contribution rates of precipitation change and other factors to the runoff jump based on the precipitation–runoff relationship before and after the JP in the YUN sub-basin. P_1 (P_2) and Ru_1 (Ru_2) are the average precipitation and the related runoff before (after) the JP, respectively. Ru_1' (Ru_2') is the runoff calculated with P_2 (P_1) according to the precipitation–runoff equation before (after) the JP

decreases from upstream to downstream in the LRB, indicating that the precipitation–runoff relationship becomes weaker when moving downstream of the basin. This is especially notable in the YUN sub-basin, where the R^2 value is only 55.80%, implying that there should be some factors acting on the precipitation–runoff relationship and changing the natural runoff-generating conditions in this sub-basin.

4.3 Discussion

As shown in Fig. 3, there are three main hydropower stations (Dachaoshan, Nuozhadu and Jinghong) located between the JIA and YUN stations. Among them, the Dachaoshan hydropower station has been operating since 2003, and the Jinghong and Nuozhadu hydropower stations has been constructing since 2003 and 2004 respectively. Hence, they inevitably cause the changes (especially the abrupt decrease) in the runoff variability at around 2004, considering that there were no changes in precipitation and runoff upstream of the JIA station. However, can we be sure that the abrupt decrease in runoff is only caused by the effects of the three hydropower projects, or have other hydroclimate factors also contributed?

To further explore the physical causes, we first calculate the contribution rates of precipitation change and other factors to the runoff change based on the equations in Table 3 (Xie et al. 2012). For a brief overview, we rewrite the equations as:

$$\begin{cases} Ru_1 = f_1(P_1) \\ Ru_2 = f_2(P_2) \\ Ru_1' = f_1(P_2) \\ Ru_2' = f_2(P_1) \end{cases} \tag{15}$$

where f_1 (f_2) refers to the Ru_{before} (Ru_{after}) equation from Table 3; P_1 and P_2 are the average precipitation before and after the JP, respectively, versus the runoff Ru_1 and Ru_2; and Ru_1' (Ru_2') is the runoff from P_2 (P_1) according to the precipitation–runoff relationship before (after) the JP. The total runoff change ΔRu is caused by both the precipitation change (denoted as ΔRu_{pre}) and other factors (denoted as ΔRu_{oth}), that is, $\Delta Ru = |Ru_2 - Ru_1|$. From Fig. 6, we can obtain $\Delta Ru_{pre1} = |Ru_1 - Ru_1'| = |f_1(P_1) - f_1(P_2)|$ and $\Delta Ru_{pre2} = |Ru_2' - Ru_2| = |f_2(P_1) - f_2(P_2)|$. Considering the contribution(s) of other factors, it can be described as $\Delta Ru_{oth1} = |Ru_1 - Ru_2'| = |f_1(P_1) - f_2(P_1)|$ and $\Delta Ru_{oth2} = |Ru_1' - Ru_2| = |f_1(P_2) - f_2(P_2)|$. Thus, ΔRu can be obtained by $\Delta Ru_{pre1} + \Delta Ru_{oth2}$ or $\Delta Ru_{pre2} + \Delta Ru_{oth1}$. The contribution rate of C_{pre} and C_{oth} can be expressed as:

$$\begin{cases} C_{pre} = (\Delta Ru_{pre1} + \Delta Ru_{pre2})/2\Delta Ru \\ C_{oth} = (\Delta Ru_{oth1} + \Delta Ru_{oth2})/2\Delta Ru \end{cases} \tag{16}$$

The results in Table 4 indicate that the abrupt decrease of runoff in the YUN sub-basin is due much more to other factors (with a contribution rate of 76.75%) rather than the precipitation change (with a contribution rate of 23.25%).

Regarding the other factors, the underlying surface conditions are further considered to clarify the physical causes. Land-cover change is an important type of human activity. The land-cover maps in the periods 2000, 2005 and 2010 (Fig. 7) were collected, and the areas of six land-cover types in the whole LRB are shown in Table 5. Forest, grassland and farmland are the main land-cover types in the LRB, with an area of about 153,344 km^2 (in 2010), while urban areas only account for a small part. Forest accounts for 42.02% (in 2000), 42.03% (in 2005) and 42.12% (in 2010), grassland accounts for 42.53%, 42.54% and

Table 4 Contribution rate of precipitation change (C_{pre}) and other factors (C_{oth}) to the runoff abrupt in the YUN sub-basin

Jump point	P_1	P_2	Ru_1	Ru_2	Ru_1'	Ru_2'	C_{pre}	C_{oth}
2004	823.59	785.04	394.93	293.10	371.25	316.78	23.25%	76.75%

The unit for precipitation and runoff is millimeter (mm)

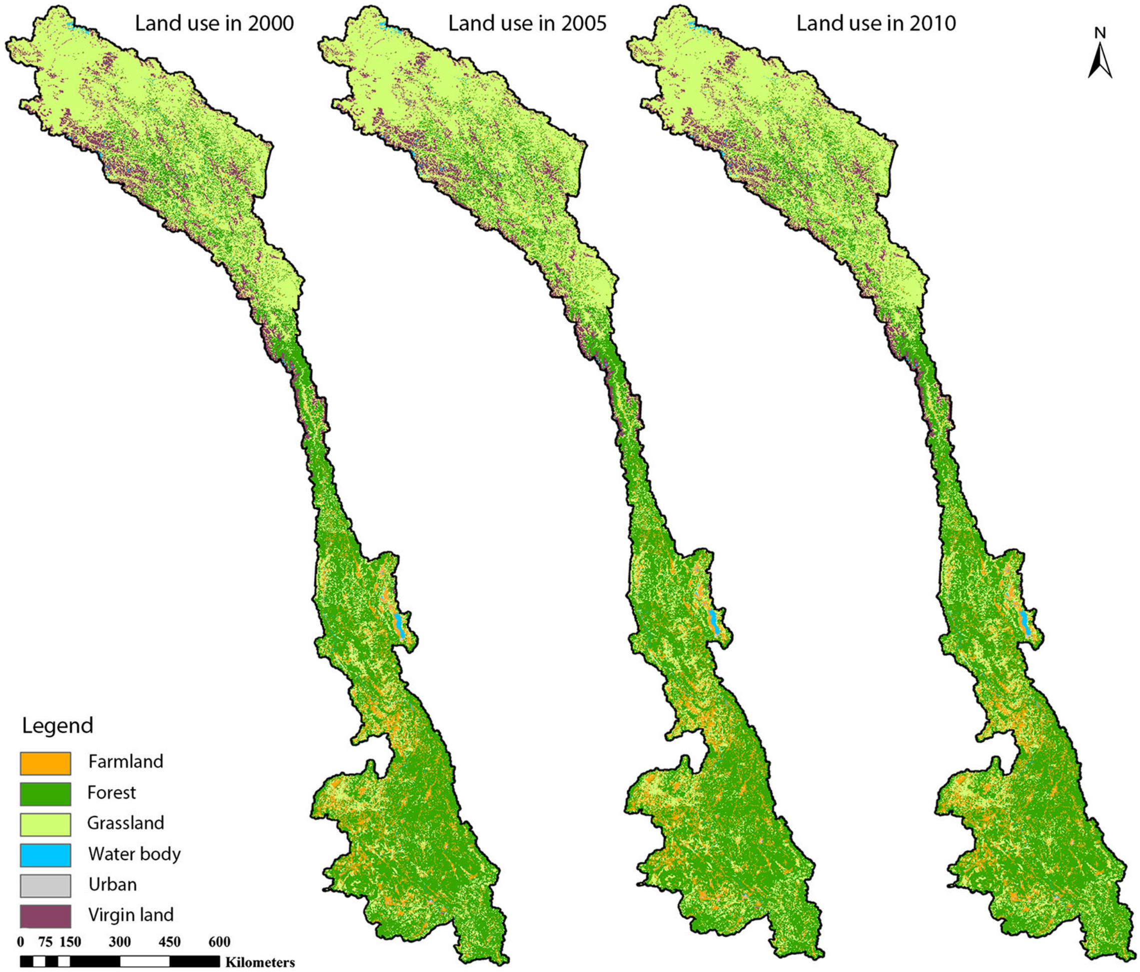

Fig. 7 Land-cover types for three periods (2000, 2005 and 2010) in the Lancang River Basin. Types of land-cover are farmland, forest, grassland, water body, urban and virgin land

Table 5 Areas of different land-cover types during three periods (2000, 2005, and 2010) in the Lancang River Basin

Year	Area and ratio	Farmland	Forest	Grassland	Water body	Urban	Virgin land
2000	Area (km^2)	14,501	69,030	69,856	1005	249	9629
	Proportion (%)	8.83	42.02	42.53	0.61	0.15	5.86
2005	Area (km^2)	14,442	69,040	69,883	1005	271	9629
	Proportion (%)	8.79	42.03	42.54	0.61	0.17	5.86
2010	Area (km^2)	14,329	69,192	69,823	1005	292	9629
	Proportion (%)	8.72	42.12	42.51	0.61	0.18	5.86

42.51%, and farmland accounts for 8.88%, 8.79% and 8.72% in the three periods, respectively. Furthermore, Further, no JP is identified in the areal actual evapotranspiration time series by MCCD method (as shown in Fig. 8). This indicates that both land surface conditions and precipitation (given in Fig. 4) show no significant change. They could not cause significant changes of actual evapotranspiration in the study area, and also could not cause the jump decrease of runoff in the YUN sub-basin. Therefore, we deduce that the abrupt decrease of runoff at the YUN station in 2004 was mainly caused by the operation and construction of the three hydropower projects, and that other hydroclimate factors contributed only very slightly.

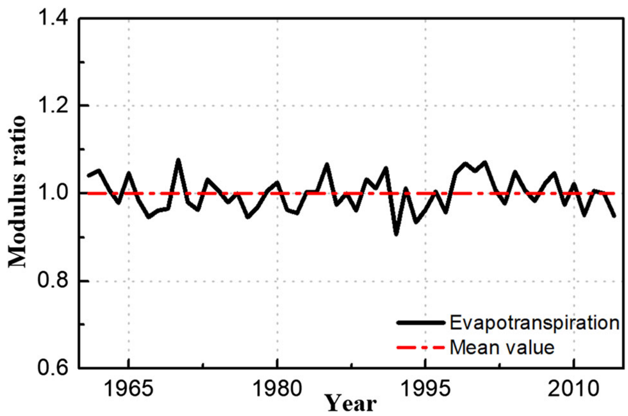

Fig. 8 Detection result of jump component in the area actual evapotranspiration series (in modulus ratio) in the Lancang River Basin. Here the modulus ratio is the ratio between the original evapotranspiration and the mean value of a series with no unit

5 Conclusions

In this study, we developed the MCCD method for the detection of jump points in hydroclimate data. In the MCCD method, the correlation coefficient (CC) between the potential jump component and the original hydrologic data is calculated. The position corresponding to the biggest absolute CC value is the expected JP, and its significance can be evaluated by comparing its biggest absolute CC value with the CC threshold at the significance level concerned. The results of statistical experiments have verified the higher efficiency of the MCCD method compared to four conventional methods. The results also showed that some parameters (including the mean value *EX*, the coefficient of variation *Cv*, the jump degree *q* and the change point position *a*) influence the JP detection result. The proposed MCCD method maintained a high efficiency level and was little influenced by the above-mentioned factors; it performed well for the detection of JPs in hydroclimate data.

By applying the MCCD method to the precipitation–runoff process in the LRB, the JP in 2004 was detected in runoff at the Yunjinghong station. The precipitation–runoff relationship was getting weaker from the upstream to downstream, and the runoff generation ability declined after 2004 in the YUN sub-basin. This was investigated, and the results showed that the land-cover conditions and the area actual evapotranspiration did not significantly change in the whole basin, and that there was little change of runoff and area precipitation in the areas above the JIA station. Therefore, the abrupt decrease of runoff in 2004 was mainly caused by the operation of some major hydropower projects, although precipitation change attributed a part of only 23.25%.

In summary, the detection of jump points is important to the understanding of the spatiotemporal variability of the hydrologic process and is especially useful in investigating anthropogenic effects. The MCCD method developed in this study performs better compared to four commonly used conventional methods, and thus it can be an effective alternative for the detection of JPs in hydroclimate data. However, the limited efficiency of the MCCD method should be noticed when detecting those JPs close to ends of time series, which should be further studied for clarification. Besides, the efficiencies of the MCCD method proposed are mainly investigated by detecting abrupt changes of mean values, while its performances for detecting the abrupt changes of *Cv* and *Cs* should be further verified. More hydrologic case studies are also needed to further verify the validity and applicability of the MCCD method. More efforts are needed to figure out the physical mechanisms of runoff abrupt changes in this basin, where additional factors (such as soil moisture) should be further considered.

Acknowledgements The authors gratefully acknowledge the valuable comments and suggestions given by the editors and the anonymous reviewers. This study was financially supported by the National Natural Science Foundation of China (No. 91547205, 91647110, 51579181, 51779176), the Strategic Priority Research Program of Chinese Academy of Sciences (No. XDA20060402), and the Youth Innovation Promotion Association CAS (No. 2017074).

Appendix 1: Groups of parameters used in statistical experiments

Group	n	EX	Cv	Cs	q	a
1	50, 100, 150, 200, 250, 300, 350, 400, 450, 500	500	0.5	2.0	300	0.5
2	200	100, 300, 500, 700, 900	0.5	2.0	300	0.5
3	200	500	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9	2.0	300	0.5
4	200	500	0.5	1.0, 1.5, 2.0, 2.5, 3.0	300	0.5
5	200	500	0.5	2.0	100, 150, 200, 250, 300, 350, 400, 450, 500	0.5
6	200	500	0.5	2.0	300	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9

Appendix 2: Mathematical expressions of the four jump point detection methods used in the statistical experiments

Method	Category	Mathematical expression	Core statistics	Indicator for jump point
Pettitt	Non-parametric	$U_{t,N} = U_{t-1,N} + \sum_{i=1}^{N} \operatorname{sgn}(x_t - x_i), \quad t = 2, 3, \ldots, N$ $K_{t,N} = \max\lvert U_{t,N}\rvert, \quad (1 \le t \le N)$	$K_{t,N}$	$p \le 0.05$
M–K	Non-parametric	$UF_k = \frac{s_k - \overline{s_k}}{\sqrt{\operatorname{var}(s_k)}}, \quad s_k = \sum_{i=1}^{k}\sum_{j}^{i-1} a_{i,j} (k = 2, 3, \ldots, n), \quad a_{i,j} = \begin{cases} 1, & x_i > x_j \\ 0, & x_i \le x_j \end{cases}$	s_k	$\begin{cases} \lvert UB_k\rvert < U_{\alpha/2} \\ \lvert UF_k\rvert < U_{\alpha/2} \\ UB_k = UF_k \end{cases}$
B–F	Parametric	$F = \sum_{i=1}^{m} n_i (\overline{x_1} - \bar{x})^2 / \sum_{i=1}^{m} (1 - n_i/N) S_i^2$	F	$F > F_\alpha$
Bayesian	Parametric	$P(H\mid E) = \frac{P(E\mid H)P(H)}{P(E)}$	–	Max posterior probability

References

Ahmadi F, Tahroudi MN, Mirabbasi R, Khalili K, Jhajharia D (2018) Spatiotemporal trend and abrupt change analysis of temperature in Iran. Meteorol Appl 25(2):314–321

Bayraktar H, Turalioglu FS, Şen Z (2005) The estimation of average areal rainfall by percentage weighting polygon method in Southeastern Anatolia Region, Turkey. Atmos Res 73(1–2):149–160

Brown MB, Forsythe AB (1974) Robust tests for equality of variances. J Am Stat Assoc 69(346):364–367

Cockburn JMH, Garver JI (2015) Abrupt change in runoff on the north slope of the Catskill Mountains, NY, USA: above average discharge in the last two decades. J Hydrol Reg Stud 3:199–210

Dessie M et al (2015) Water balance of a lake with floodplain buffering: Lake Tana, Blue Nile Basin, Ethiopia. J Hydrol 522:174–186

Ding J, Deng Y (1988) Statistical hydrology. Press of University of Science and Technology of Chengdu, Chengdu

Domonkos P (2013) Efficiencies of inhomogeneity-detection algorithms: comparison of different detection methods and efficiency measures. J Climatol. https://doi.org/10.1155/2013/390945

Dugan PJ et al (2010) Fish migration, dams, and loss of ecosystem services in the Mekong Basin. Ambio 39(4):344

Fan H, He D, Wang H (2015) Environmental consequences of damming the mainstream Lancang-Mekong River: a review. Earth Sci Rev 146:77–91

Fu B (1996) On the calculation of evaporation from land surface in mountainous areas. J Meteorol Sci 04:328–335

Furey P, Kampf SK, Lanini J, Dozier A (2010) Examining the vulnerability of hydropower production in meso-scale snowmelt-runoff basins under different climate change scenarios. AGU fall meeting

Gagné MÈ, Kirchmeier-Young MC, Gillett NP, Fyfe JC (2017) Arctic sea ice response to the eruptions of Agung, El Chichón, and Pinatubo. J Geophys Res Atmos 122(15):8071–8078

Karpechko AY, Gillett NP, Dall'Amico M, Gray LJ (2010) Southern hemisphere atmospheric circulation response to the El Chichón and Pinatubo eruptions in coupled climate models. Q J R Meteorol Soc 136(652):1813–1822

Kundzewicz ZW, Robson AJ (2004) Change detection in hydrological records—a review of the methodology. Int Assoc Sci Hydrol Bull 49(1):7–19

Lauri H et al (2012) Future changes in Mekong River hydrology: impact of climate change and reservoir operation on discharge. Hydrol Earth Syst Sci 16(12):4603–4619

Li S, Xiong L, Li H et al (2016) Attributing runoff changes to climate variability and human activities: uncertainty analysis using four monthly water balance models. Stoch Env Res Risk Assess 30(1):251–269. https://doi.org/10.1007/s00477-015-1083-8

Mann HB (1945) Nonparametric tests against trend. Econometrica 13(3):245–259

Montanari A (2012) Detecting and quantify non-stationarity: a look to a changing hydrology for a changing society and environment and a perspective on the next IAHS Scientific Decade 2013–2022. AGU fall meeting

Murphy KR, Myors B (2004) Statistical power analysis: a simple and general model for traditional and modern hypothesis tests. Lawrence Erlbaum Associates, New Jersey, pp 539–542

Nasseribrahim A, Yasushi M, Masahiro M (2005) Remote sensing algorithms for monitoring eutrophication in Ishizuchi storm water reservoir in Kochi Prefecture, Japan. Hydrol Sci J 50(3):542. https://doi.org/10.1623/hysj.50.3.525.65024

Perreault L et al (1999) Detection of changes in precipitation and runoff over eastern canada and U.S. using a Bayesian approach. Stoch Env Res Risk Assess 13(3):201–216. https://doi.org/10.1007/s004770050039

Pettitt AN (1979) A non-parametric approach to the change-point problem. J Roy Stat Soc 28(2):126–135

Reeves J, Chen J, Wang XL, Lund R, Lu Q (2007) A review and comparison of change point detection techniques for climate data. J Appl Meteorol Climatol 46(6):900

Rougé C, Ge Y, Cai X (2013) Detecting gradual and abrupt changes in hydrological records. Adv Water Resour 53:33–44. https://doi.org/10.1016/j.advwatres.2012.09.008

Sang YF, Wang ZG, Li ZL (2012) Entropy aided detection of abrupt climate change: a case study in the Haihe River Basin, China. Entropy 14(7):1274–1284

Şen Z (2017) Innovative trend methodologies in science and engineering. Springer, Heidelberg

Shi W, Yu X, Liao W, Wang Y, Jia B (2013) Spatial and temporal variability of daily precipitation concentration in the Lancang River Basin, China. J Hydrol 495:197–207

Sun F (2007) Study on watershed evapotranspiration based on the Budyko Hypothesis. Tsinghua University, Beijing

Veihe A, Quinton J (2015) Sensitivity analysis of EUROSEM using Monte Carlo simulation I: hydrological, soil and vegetation parameters. Hydrol Process 14(5):915–926

Villarini G, Serinaldi F, Smith JA, Krajewski WF (2009) On the stationarity of annual flood peaks in the continental United States during the 20th century. Water Resour Res 45(8):2263–2289

Wikle CK (2003) Modeling hydrologic change: statistical methods. Technometrics 45(4):366–367

Xie P, Liu Y, Yang GL et al (2012) Attribution analysis of water resources alteration in the Third Grade Wulijimurenhe. J China Hydrol 32(2):40–43

Xie P, Wu ZY, Sang YF et al (2018) Evaluation of the significance of abrupt changes in precipitation and runoff process in China. J Hydrol 560:451–460

Yang T et al (2009) Spatio-temporal changes of hydrological processes and underlying driving forces in guizhou region, southwest china. Stoch Env Res Risk Assess 23(8):1071–1087. https://doi.org/10.1007/s00477-008-0278-7

Zhang J, Huang Q, Zhao X (2013) Comparative research on abrupt change analysis methods for hydrological time series in Zhangze reservoir. J Basic Sci Eng 21(5):837–844

Zhou Y, Shi C, Fan X, Du J (2011) Advances in the research methods of abrupt changes of hydrologic sequences and their applications in drainage basins in China. Progr Geogr 30(11):1361–1369