

# Comprehensive Diagnosis of Runoff Series Discontinued Point in Longchuan River Basin

JIANG Ru-cheng<sup>1</sup>, CHENG Jing<sup>1</sup>, GU Shi-xiang<sup>1</sup>

<sup>1</sup>Yunnan Water Conservancy and Hydroelectric Survey Design and Research Institute, Kunming 650021, China;

**Abstract:** With the influence of climate change and human activities, the runoff series will have discontinued point. The diagnosis of discontinued point is an important part of runoff analysis and calculation. This paper used the comprehensive diagnosis method to diagnosis the discontinued point of the tributary of the runoff series in Longchuan river basin, indicating that the flood mutated in 1979 and 1998, which are significant discontinued points. The result is consistent with the hydrological investigated results. The comprehensive diagnosis method can consider the advantages of various methods and avoid the limitation of single method.

## 1 Introduction

Under the influence of atmospheric circulation movement, underlying surface conditions and human activities, the hydrological situation will undergo temporal and spatial variability, breaking the original general change law, and once the hydrological sequence is abrupt, the consistency will change. Therefore, prior to the hydrological analysis calculation, prediction or simulation, the hydrological sequence must be identified for the discontinued point. Discontinued point identification, that is, using relevant methods and calculations to diagnose when or how the hydrological time series has undergone significant variation, and find corresponding discontinued points. At present, there are many diagnostic methods for hydrological sequence discontinued points, such as ordered cluster analysis method, T-test method, rank sum test method, Mann-Kendall method, Bayesian theory based Lee-Heghinan method, R/S analysis method, Slide the segmentation model, run-length testing method, etc. <sup>[1, 2]</sup>. These methods have their own advantages and disadvantages. Different methods have inconsistent diagnosis results. How to accurately determine the reliable discontinued points from various diagnostic results is the key to the diagnosis of discontinued points, and it is also the scientific problem faced by the current research. .

The Longchuan River originates from the Tianzi Temple slope in Nanhua County, Chuxiong Prefecture, Yunnan Province. The river source is about 2600m above the sea level. From the west to the east, from the west to the east, after flowing through the Shaqiao dam of Nanhua County, it is included in the north-south tributary of the left bank, Shuangdian River, and then eastward, passing through the Nanhua County to reach the

intersection of the Laidian Town of Chuxiong City and the Zidian River, the left bank tributary of the north-south direction. After about 14km, the process entered the Qingshanzui Reservoir area. After leaving the warehouse, it flows through the Chuxiong urban area in the southeast direction. After passing through the river, the river front and the Xiaohekou, it turns to the northeast to the sea to the Yuanmou. In the lower part of the small cucumber garden hydrological station, it is included in the river from the west to the east. Joined the Jinsha River near the village. The total length of the Longchuan River is 257km, the average river channel is 4.8‰, and the runoff area is 9185km<sup>2</sup>. The Longchuan River Basin involves 9 counties in Chuxiong Prefecture. The human activities in the basin are more significant, and there are many reservoirs and dams, and the channels are vertical and horizontal.

In order to investigate the influence of meteorological changes and underlying surface conditions on the runoff characteristics of the Longchuan River Basin, this paper uses the comprehensive diagnosis method of runoff sequence discontinued points to diagnose the discontinued point of the Longchuan River runoff series, specifically identifying the year in which the runoff is mutated, thereby quantifying describe the degree of variation in runoff and provide a basis for subsequent hydrological analysis calculations. This month, we collected the series of monthly measured runoffs from the seven hydrological stations in the Longchuan River Basin as the data base for the analysis of the discontinued points.

## 2 Information and methods

### 2.1 Data

The monthly hydrological stations of the seven hydrological stations built or built in the Longchuan River Basin have been measured monthly ( $m^3/s$ ); county water conservancy records, meteorological disasters survey statistics and historical materials.

### 2.2 Analytical method

The comprehensive diagnosis method of the runoff sequence discontinued point is used for analysis. The method includes three steps: preliminary diagnosis, detailed diagnosis and comprehensive diagnosis. The initial diagnosis is to test the pure randomness of the hydrological sequence [3], and to judge whether there is variation in the sequence from the level of qualitative analysis; the detailed diagnosis is based on various methods such as ordered cluster analysis, Lee-Heghinan method, Time-by-time Sliding Segmentation model, Mann-Kendall method, R/S analysis method, T-test method, rank sum test method, run-length test method,

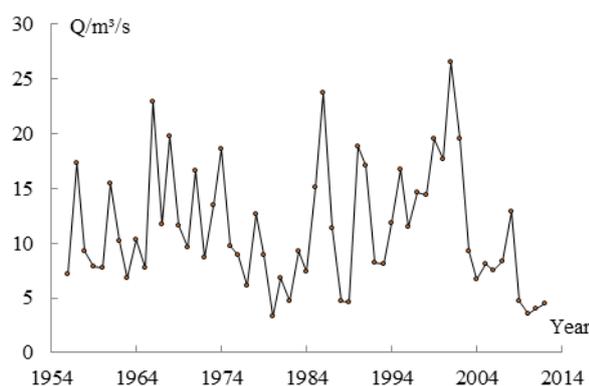


Fig. 1 Longchuan River runoff process line

According to the R/S analysis method [10], the Hurst coefficient of the 61-year runoff series of Xiaocuyuan Station is  $0.639 \in (0.5, 1)$ , which indicates that the runoff sequence is affected by external factors, not the pure random sequence, and it also has the feature of strong long-term memory.

Draw a cumulative runoff process line, as shown in Figure 2 (the ordinate on the ordinate is the cumulative runoff). From the figure, it can be initially diagnosed that the slope of the cumulative process line of the runoff series has turned around in 1974 and 2002, and the three-stage runoff accumulation process line before 1974, 1974-2002, and 2002 was linearly simulated. It can be seen that there are differences in the slopes of the three straight lines fitted, indicating that the external factors produced by the runoff changes have changed significantly, and the 1974 and 2002 changes in the slope are initially regarded as the discontinued points.

Through the above analysis, it is preliminarily judged that there is a variation characteristic of the runoff series

etc and analyzing and diagnosing runoff sequences to find possible discontinued points; comprehensive diagnosis is in the first two On the basis of the steps, the diagnostic results of various methods are statistically compared and compared, and the point with the most occurrences in the test results is taken as the most probable discontinued point. Finally, the discontinued points are comprehensively determined in combination with the actual hydrological survey situation in the basin.

## 3 Diagnosis of Runoff Discontinued point in Longchuan River Basin

### 3.1 Typical hydrological station diagnosis

#### 3.1.1 Preliminary diagnosis

According to the 61-year measured runoff sequence of the Kumquat Garden Station from 1954 to 2014, the line drawing of the runoff process over the years is shown in Figure 1. It can be seen from the figure that after the 1960s, 1970s, and 2000s, the series has obvious steepness, and it is speculated that there is variation.

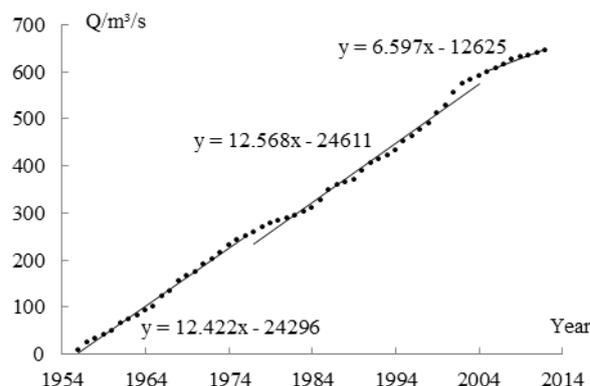


Fig. 2 Longchuan River runoff accumulation process line

of the cucumber garden of Longchuanjiang export control station.

#### 3.1.2 Detailed diagnosis

The ordered clustering method [1], Lee-Heghinan method [1], rank sum test method [1], Mann-Kendall method [7] were used to diagnose the breakpoints of the cucumber runoff sequence in detail.

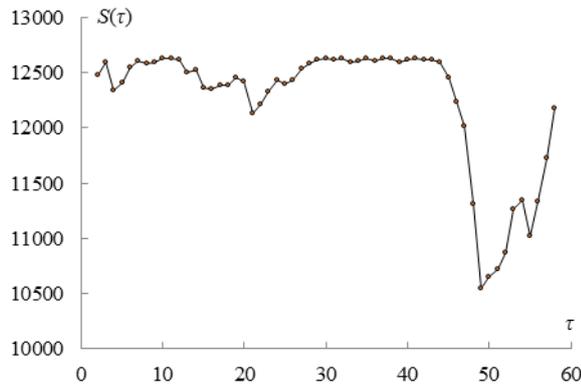
##### (1) Ordered clustering method

The orderly clustering analysis is based on the principle that the squared sum of the dispersions between the similarities is small, and the squared sum of the differences between the different classes is larger, and the optimal dividing point is derived as the most likely point of mutation. For the runoff series  $x_1, x_2, \dots, x_n$ , let the possible discontinued point be  $\tau$ , if it satisfies:

$$S = \min_{1 \leq \tau \leq n-1} \left\{ S_n(\tau) = \sum_{i=1}^{\tau} (x_i - \bar{x}_{\tau})^2 + \sum_{i=\tau+1}^n (x_i - \bar{x}_{n-\tau})^2 \right\} \quad (1)$$

Then  $\tau$  is the most likely point of mutation. Where  $\bar{x}_\tau$  and  $\bar{x}_{n-\tau}$  are the mean of the sequence around  $\tau$ .

Figure 3 shows the results of the ordered clustering method for the breakpoint test of the Longchuan River runoff sequence. It can be seen that when  $\tau=49$  is the smallest, the corresponding 2002 is the most likely discontinued point. Set 2002 as the boundary, the runoff



**Fig. 3** Inspection diagram of ordered clustering method

### (2) Lee-Heghinian method

For the runoff sequences  $x_1, x_2, \dots, x_n$ , Lee and Heg, assuming that the population is normally distributed and the a priori distribution of the abrupt points is evenly distributed, the posterior distribution of the possible discontinued point  $\tau$  is:

$$f(\tau/x_1, x_2, \dots, x_n) = k[n/\tau(n-\tau)]^{1/2} [R(\tau)]^{-(n-2)/2} (1 \leq \tau \leq n-1) \quad (2)$$

In the formula,  $k$  is a proportionality constant,  $n$  is the sample size,

$$R(\tau) = \left[ \sum_{i=1}^{\tau} (x_i - \bar{x}_\tau)^2 + \sum_{i=\tau+1}^n (x_i - \bar{x}_{n-\tau})^2 \right] / \sum_{i=1}^n (x_i - \bar{x}_n)^2$$

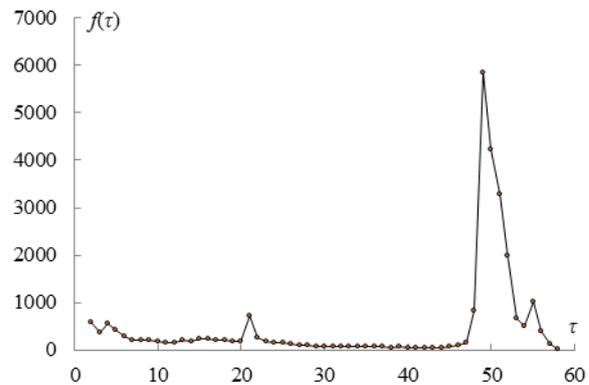
$\bar{x}_\tau$  and  $\bar{x}_{n-\tau}$  is the mean value around  $\tau$ ,  $\bar{x}_n$  is the mean of the entire sequence. When  $\max_{1 \leq \tau \leq n-1} f(\tau/x_1, x_2, \dots, x_n)$  is satisfied,  $\tau$  is the most likely point of mutation.

Figure 4 shows the results of the Lee-Heghinian test. It can be seen that when  $\tau=49$  is the largest, the corresponding 2002 is the most likely discontinued point; as can be seen from the figure, when  $\tau$  is 21 or 55, it is the local maximum point, and the corresponding 1974 and 2008 may also be the discontinued points.

### (3) Rank sum test method

The rank sum test method arranges the sequences from small to large or from large to small, and uniformly numbers them from 1 to the order of each data in the arrangement. The rank of the data is called the rank of the data, and the rank of the same data is averaged. value.

sequence of 1954-2002 and 2003-2014 is tested. The previous stage is  $\tau=21$ ,  $S_n(\tau)$  is a local minimum point. The corresponding 1974 is a possible discontinued point. In the stage test is  $\tau=55$ ,  $S_n(\tau)$  is a local minimum point, and the corresponding 2008 is a possible discontinued point.



**Fig. 4** Test diagram of Lee-Heghinian method

Let the variation point of the runoff sequence  $x_1, x_2, \dots, x_n$  be  $t$ , and the two parts before and after  $t$  have  $n_1$  and  $n_2$  values, and the sum of the data ranks of the part with smaller capacity is  $W$ , when  $n_1, n_2 > 10$ ,  $W$  approximates a normal distribution, as

$$W \sim N(n_1(n_1 + n_2 + 1)/2, n_1 n_2 (n_1 + n_2 + 1)/12) \quad (3)$$

Then the statistic is

$$U = (W - n_1(n_1 + n_2 + 1)/2) / \sqrt{n_1 n_2 (n_1 + n_2 + 1)/12} \sim N(0,1) \quad (4)$$

Where  $n_1$  is a small sample size. Suppose that two samples before and after  $\tau$  come from the same population, as  $F_1(x) = F_2(x)$ . Given the significance level  $\alpha$ , check  $u_{\alpha/2}$ , when  $|U| < u_{\alpha/2}$ , the null hypothesis is accepted, the two samples before and after the discontinued point  $\tau$  are from the same population, indicating that the variation is not significant; on the contrary, the variation is significant.

The point at which the  $U$  statistic reaches its maximum value, as the point of the most likely variability. Take the significance level  $\tau = 0.05$ . Figure 5 shows the results of the rank sum test. It can be seen that under the condition of satisfying the confidence level, when the  $\tau$  is 21, the fluctuation is most significant, and the corresponding year is 2002; When the  $\tau$  is 55, the fluctuation is also significant, and the corresponding year is 2008.

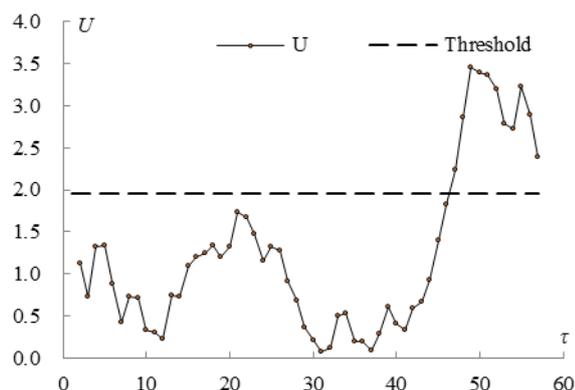


Fig. 5 Rank sum test test chart

#### (4) Mann-Kendall test method

Mann-Kendall test method has been widely used to analyze trends or jumps in time series of factors such as precipitation, runoff, and temperature. Define the statistic as

$$UF_k = [S_k - E(S_k)] / \sqrt{\text{var}(S_k)} \quad (5)$$

In the formula,  $S_k = \sum_{i=1}^k \sum_{j=1}^i a_{ij}$  ( $k=2,3, \dots, n$ ),

$E(S_k)=k(k-1)/4$ ,  $\text{var}(S_k)=k(k-1)(2k+5)/72$ , when  $x_i > x_j$ ,  $a_{ij}=1$ , or  $a_{ij}=0$ .

Calculate the statistic  $UF_k$  ( $k=2,3, \dots, n$ ) of the original sequence, and plot the curve UF, then invert the sequence, calculate the  $UF_{k^*}$  of the inverted sequence, and then according to  $UB_k = -UF_{k^*}(k^*=n+1-k)$  Calculate  $UB_k$  and plot curve UB. If there is an intersection point between the curves UF and UB, and the intersection point is within the critical line  $\pm 1.96$  (take the significance level  $\alpha$  as 0.05), then the moment corresponding to the intersection point is the time at which the mutation starts.

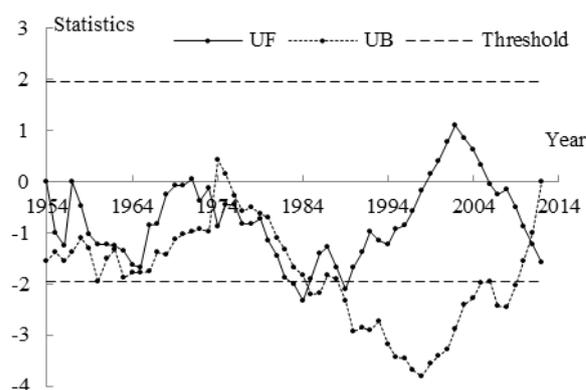


Fig. 6 Mann-Kendall test chart

Figure 6 shows the results of the Mann-Kendall test method. It can be seen from the Figure 6 that the intersection of the two curves within the critical line is 1974 and 2008, and these years are the discontinued points detected by the method.

#### 3.1.3 Comprehensive diagnosis

The comprehensive diagnosis of the discontinued point is to test the hydrological time series by various methods, and then to compare and compare the various test results, and consider that the year with the most number of discontinued points is the most likely discontinued point; and the conclusion of the hydrological survey is included in the analysis. Among them, the focus is on the point on the underlying surface where the significant change time interval may occur, and then the discontinued point is found more accurately.

Table 1 Comprehensive diagnosis analysis table

Method	Results	Method	Results
Ordered clustering	1974、2002、2008	Rank sum test	2002、2008
Lee-Heghinan	1974、2002、2008	Mann-Kendall	1974、2008

It can be seen from the summary of the test results of various test methods that all of the four methods tested the point of change in 2008, and each of the three methods tested the discontinued points in 1974 and 2002. Obviously, from the perspective of mathematical statistics, the above three years are all discontinued points. The comprehensive diagnosis method of the discontinued point tests the hydrological sequence from various aspects, solves the problem of inconsistency of various test methods, and the analyzed results are relatively reliable.

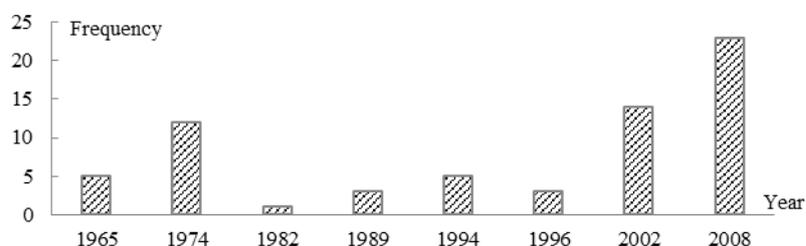
### 3.2 Full basin diagnosis

According to the above comprehensive diagnosis method, the hydrological sequence discontinued points were diagnosed in 7 hydrological stations in the Longchuan River Basin, and the diagnostic results of various diagnostic methods were counted. See Table 2 and Figure 7. It can be seen from the table that most of the sites have detected 1974, 2002, and 2008 as the discontinued points. According to the actual investigation and analysis, the

Longchuan River built a large number of reservoirs and dams in the 1970s, and the reservoir impeding function drastically changed the runoff. The natural condition shows the phenomenon of reduced runoff; before and after 2002, the rivers in Yunnan Province were generally in the period of abundant water, and the runoff increased sharply; after 2008, Yunnan Province began to suffer from drought and the river runoff was exhausted. The discontinued points determined by this comprehensive diagnosis are consistent with the actual investigation, which confirms the rationality of the analysis and calculation. Therefore, when the degree of influence of the change of the surface condition on the runoff of the Longchuan River Basin is quantitatively described, the runoff sequence can be divided into three segments. The sequence before 1974 is used as the "natural" sequence not affected by human activities; from 1974 to 2002 The impact of reservoir storage and storage is significant; after 2002, except for the impact of the reservoir, the runoff affected by climate change has changed accordingly in 2002 and 2008, respectively.

**Table 2** Summary of test results of each site

Hydrologic Station	Xiaohekou	Doubashi	Fengtun	Zhengxingba	Xiaohuangguayuan	Suhaichong	Duoke
Results	1982 2002 2008	1989 1994 1996	1974 2002 2008	1965 1974 1994 2008	1974 2002 2008	1965 1974 2002 2008	1974 2002 2008



**Fig. 7** Summary of test results of various methods

## 4 Summary

In this paper, the comprehensive diagnosis method of hydrological sequence variation is used to diagnose the runoff discontinued point in the Longchuan River Basin. The diagnosed discontinued points are 1974, 2002 and 2008, and the diagnostic results are consistent with the watershed investigation. Compared with the single test method, the method has many test indicators, can test the hydrological sequence from various aspects, and comprehensively reflects the variation characteristics of the time series. The analyzed results are more reliable, avoiding the unreasonable test results of the single test method. The problem of inconsistent inspection methods. At the same time, considering the conclusion of the actual hydrological survey, the analysis of the occurrence of the sudden change point from the physical cause makes the test result more reliable and scientific. In the future, when quantitatively describing the influence of the change of the surface condition on the runoff of the Longchuan River Basin, the sequence before 1974 can be used as the natural sequence, which can avoid the deficiency of artificially selected natural sequences, and is more scientific and objective.

## Acknowledgment

The study is financially supported by the Key Program of Yunnan Applied Basic Research Projects (No. 2017FA022 ), the Academician Workstation Project (No. 2015IC013) and Scientific Research and Technical Innovation Team Construction of Yunnan Province (No. 2017TD025).

## References

1. Wang W S, Jin J L, Li Y Q, et al. Stochastic Simulation Technology for Hydrology and Water Resources[M]. Chengdu: Sichuan University Press, 2007.
2. Xie P, Chen G C, Li D, et al. Comprehensive Diagnosis Method of Hydrologic Time Series Change-point

- Analysis [J]. Water Resources and Power, 2005, 23(2): 11-14.
3. Xie P, Chen G C, Lei H F, et al. Hydrological alteration diagnosis system [J]. Journal of Hydroelectric Engineering, 2010, 29(1): 86-91.
4. Xi Y J. Elementary Analysis on Features of Water and Sediment in Kuye River Basin [J]. Ground water, 2004, 26 (3): 226-228.
5. Zhu H F, Zhao W W, Kang M Y, et al. Effect of Human Activities on Flood Season Runoff in Water and Soil Conservation Region [J]. Advances in Water Science, 2008, 19(3): 400-405.
6. Ding J. Statistical Detection for Transition Point in Flood Time Sequences [J]. Journal of Wuhan College of Hydr. & Elec. Engineering, 1986 (3): 36-41.
7. Lei H F, Xie P, Chen G C, et al. Comparison and Analysis on the Performance of Hydrological Time Series Change-point Testing Methods [J]. Water Resources and Power, 2007, 25(4): 36-40.
8. Xiao Y, Xia J, Shen M L, et al. The Application of Difference Information Theory in Detecting Change-points of Hydrological Time Series [J]. China Rural Water and Hydropower, 2001, (11): 28-30.
9. Xia J. Research Progress of Hydrological Complexity and Uncertainty [M]. Beijing: China Water Power Press, 2004.
10. Wang X L, Hu B Q, Xia J. R /S Analysis Method of Trend and Aberrance Point on Hydrological Time Series [J]. Engineering Journal of Wuhan University, 2002, 35(2): 10-12.
11. Jiao E Z. Elementary Analysis on Trends of Water and Sediment in Kuye River Basin [M]. Zhengzhou: Huanghe Water Press, 2002.