Non-stationary hydrological frequency analysis based on the reconstruction of extreme hydrological series

Y. M. Hu[1,2], Z. M. Liang[1,3], X. L. Jiang[1], and H. Bu[1]

[1]College of Hydrology and Water Resources, Hohai University, Nanjing, China
[2]Hydroinformatics Chair Group, UNESCO-IHE Institute for Water Education, Delft, the Netherlands
[3]State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering,
Hohai University, Nanjing, China

Correspondence to: Z. M. Liang (zmliang@hhu.edu.cn)

Abstract. In this paper, a novel approach for non-stationary hydrological frequency analysis is proposed. The approach is due to the following consideration that, at present the data series used to detect mutation characteristic is very short, which may only reflect the partial characteristic of the population. That is to say, the mutation characteristic of short series may not fully represent the mutation characteristic of population, such as the difference of mutation degree between short sample and population. In this proposed method, an assumption is done that the variation hydrological series in a big time window owns an expected vibration center (EVC), which is a linear combination of the two mean values of the two subsample series obtained through separating the original hydrological series by a novel optimal segmentation technique (change rate of slope method). Then using the EVC to reconstruct non-stationary series to meet the requirement of stationary, and further ensure the conventional frequency analysis methods is valid.

1 Introduction

Hydrological frequency analysis (HFA) is widely used to obtain designed flood with a given horizon level, which is extremely important for hydraulic engineering construction and water resources management. A fundamental assumption of conventional HFA relays on is that the hydrological extreme series should be stationary. However, due to the impact of climate change and human activities, the hydrology process shows non-stationary characteristic, which destroy the basic assumption and reduce the reliability of the conventional HFA (Milly et al., 2009). In recent years, in term of non-stationary HFA, large bodies of research works have been done. Xie et al. (2005) introduced decomposition and composition of time series method (DCCM) for hydrological frequency analysis. In this method, the time series is separated into stationary random component and non-stationary deterministic component, then using function to fit the deterministic component, finally, according to the theory of decomposition and composition in time series, the deterministic

forecasting value and stationary random component is synthesized. Strupczewski et al. (2001) established the stochastic relationship between time and parameters of distribution function, and then estimate the design value for a given design level. The two methods are very difficult to be used for engineering design because the designed value with a given probability are different at different time. Singh et al. (2005) used conditional distribution function for hydrological frequency analysis considering that floods are generated by different mechanisms. Singh and Sinclair (1972) employed the mixed distribution function for frequency analysis. However, it is difficult to ensure the reliability of parameters estimation when the parameters increase with the sub-distributions increase. Other studies on non-stationary frequency analysis includes Cunderlik and Burn (2003), Leclerc and Ouarda Taha (2007) and Gilroy and McCuen (2012). It is noted that at present, the data series used to detect the mutation is very short, which may only present the partial characteristic of the population. That is to say, the mutation characteristic of short series may not fully represent the mutation characteris-

tic of population, such as the difference of mutation degree between short sample and population. The thing is very possible to happen that in 50-year series, the trend or jump mutation can be detected, however, in 100a series including the 50-year series, there is no mutation (Koutsoyiannis, 2006).

Based on above consideration, in this paper, we propose a novel statistical method for reconstructing the non-stationary series to ensure the conventional frequency analysis methods is valid. In this method, an assumption is done that the mutation hydrological series in a big time window owns an EVC, which is a linear combination of the two mean values of the two subsample series obtained through separating the original hydrological series by a optimal segmentation technique (change rate of gradient method). Then using the EVC to reconstruct non-stationary series to meet the requirement of stationary. Finally, a study case will be shown.

2　Methodology

2.1　Basic assumption

The mutation analysis is significantly influenced by sample length. The degree and type of mutation always continue to change with time going. Based on these consideration, an assumption is done that the trend-mutation hydrological series in a big time window owns an EVC, which is a linear combination of the two mean values of the two subsample obtained through separating the original hydrological series by an optimal segmentation technique named change rate of gradient method.

Assuming a hydrological extreme series $x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_n$, its gradient is k. For the first sub-sample $x_1, x_2, \ldots, x_m$, its gradient is k_m and for the second sub-sample $x_{m+1}, \ldots x_n$, its gradient is k_{m+1}. The change rate of gradient kc_m can be calculated by the following expression.

$$kc_m = |\frac{k_{m+1} - k_m}{k}|, \ m \in \left[\frac{1}{4}n, \frac{3}{4}n\right] \quad (1)$$

In order to ensure the reliability of estimation on gradient. The m value is set as $m \in [\frac{1}{4}n, \frac{3}{4}n]$. And the optimal point τ of division can be obtained by,

$$\tau = \max(kc_m), \ m \in \left[\frac{1}{4}n, \frac{3}{4}n\right]. \quad (2)$$

Using the optimal point τ, the original series $x_1, x_2, \ldots, x_m, x_{m+1} \ldots, x_n$ can be separated into two sub-samples $x_1, x_2, \ldots, x_\tau$ and $x_{\tau+1} \ldots, x_n$. The EVC can be computed by,

$$\text{EVC} = \alpha \cdot \frac{1}{\tau} \sum_{i=1}^{\tau} x_i + (1 - \alpha) \frac{1}{n - \tau} \sum_{j=\tau+1}^{n} x_j. \quad (3)$$

Considering the recent environment better represent the future environment than that of past environment. The weight α can be estimated by Eq. (4).

$$\alpha = \frac{\frac{1}{n-\tau} \sum_{j=\tau+1}^{n} x_j}{\frac{1}{n} \sum_{i=1}^{n} x_i}, \text{ when } \frac{1}{\tau} \sum_{i=1}^{\tau} x_i \geq \frac{1}{n - \tau} \sum_{j=\tau+1}^{n} x_j$$

$$\text{or } \alpha = \frac{\frac{1}{\tau} \sum_{i=1}^{\tau} x_i}{\frac{1}{n} \sum_{i=1}^{n} x_i}, \text{ when } \frac{1}{\tau} \sum_{i=1}^{\tau} x_i < \frac{1}{n - \tau} \sum_{j=\tau+1}^{n} x_j \quad (4)$$

2.2　Reconstruction of hydrological extreme series

Firstly, the Eqs. (3)–(4) are used to calculate the EVC value, then based on EVC, the hydrological extreme series is reconstructed using Eqs. (4)–(5). For the sub-sample $x_1, x_2, \ldots, x_\tau$, the reconstructed series x_i', $i = 1, 2, \ldots, \tau$ can be calculated by,

$$x_i' = x_i - (f_1(x_i) - \text{EVC}), i = 1, 2, \ldots \tau. \quad (5)$$

And for the sub-sample $x_{\tau+1} \ldots, x_n$, the reconstructed series x_i', $i = \tau + 1, \tau + 2, \ldots, n$ is calculated by,

$$x_i' = x_i - (f_2(x_i) - \text{EVC}), i = \tau + 1, \tau + 2, \ldots, n \quad (6)$$

where, $f_1(x)$ and $f_2(x)$ are linear regression function of the two sub-samples respectively. In oder to make it better understood, the schematic diagram is shown (Fig. 1).

3　Study case

The annual runoff of Sanhuangmiao station covering the period of 1947–2008 is used to demonstrate the proposed method. Figure 1 shows the characteristic of runoff series. It can be seen that the series has a significant decreasing trend. The detect results done by change rate of gradient method shows that at 1965, a significant change of the series occurs. The gradient are -1.93 and -0.37 before and after 1965 respectively. This means that the two sub-samples before and after 1965 own different trend.

Using Eqs. (4)–(5), the series can be reconstructed. It seems that the reconstructed series is stationary (Fig. 2). The Man–Kendell method (M–K) is used to test the series. The test results with 0.05 significant level shows that the M–K statistical series is between the lines $+1.96$ and -1.96, which indicates that the series is stationary (Fig. 3).

4　Conclusions

In this paper, a simple statistical method is introduced to reconstruct the non-stationary hydrological extreme series, which consider that mutation characteristic of finite sample may only represent spatial characteristic of population and cannot fully reflect population's change. In this method, a
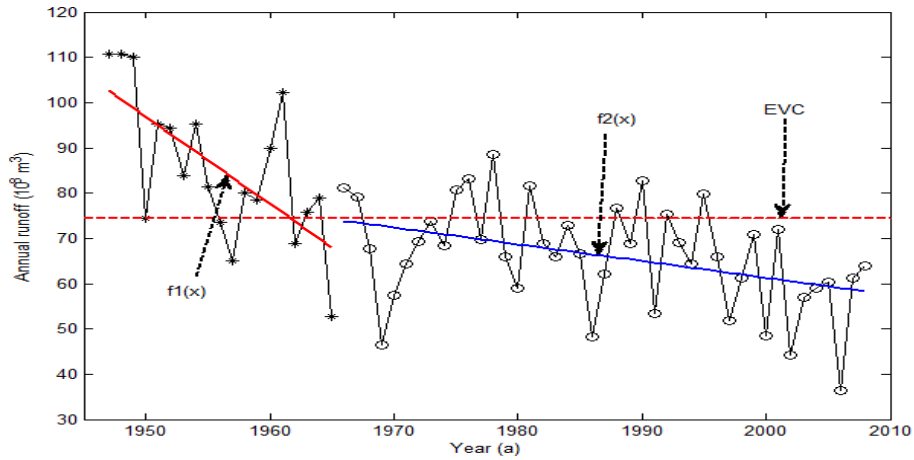
Figure 1. schematic diagram to introduce how to reconstruct the hydrological series. The red solid line is the trend line of sub-sample; the blue solid line is the trend line of sub-sample; the red imaginary line means EVC.
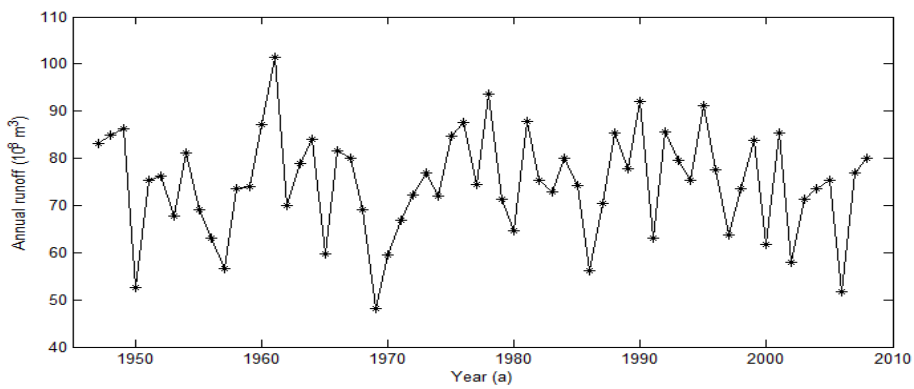
Figure 2. Reconstructed annual sreamflow series.
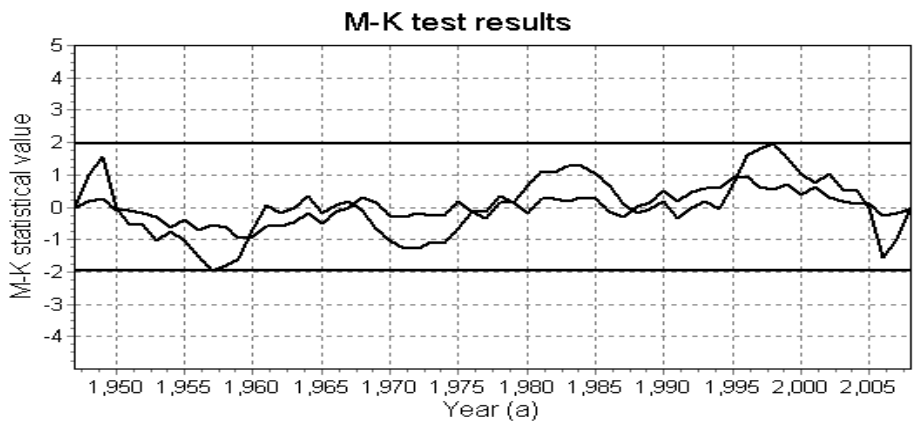
Figure 3. Man–Kendell test result for reconstructed annual sreamflow series.

basic assumption is done that non-stationary hydrological series in a big time window or population owns an EVC, which is a linear combination of the two mean values of the two sub-sample series obtained from the original hydrological series. In order to search for the optimal point to separate the orig-

inal series, a novel optimal segmentation technique named change rate of gradient method is proposed. A case is employed to demonstrate this method. And the results indicated that using this method to reconstruct the non-stationary hydrological extreme series can ensure the series used for hy-

drological frequency analysis meets the requirement of stationary.

References

Cunderlik, J. M. and Burn, D. H.: Non-stationary pooled flood frequency analysis, J. Hydrol., 276, 210–223, 2003.

Gilroy, K. L. and McCuen, R. H.: A nonstationary flood frequency analysis method to adjust for future climate change and urbanization, J. Hydrol., 414–415, 40–48, 2012.

Koutsoyiannis, D.: Nonstationarity versus scaling in hydrology, J. Hydrol., 324, 239–254, 2006.

Leclerc, M. and Ouarda Taha, B. M. J.: Non-stationary regional flood frequency analysis at ungauged sites, J. Hydrol., 343, 254–265, 2007.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J.: Stationarity is dead: whither water management?, Science, 319, 573–574, 2009.

Sing, K. P. and Sinclair, R. A.: Two-distribution method for flood frequency analysis, J. Hydraul. Div., 98, 29–44, 1972.

Singh, V. P., Wang, S. X., and Zhang, L.: Frequency analysis of nonidentically distributed hydrologic flood data, J. Hydrol., 307, 175–195, 2005.

Strupczewski W. G., Singh V. P., and Feluch, W.: Non-stationary approach to at-site flood frequency modeling I. aximum likelihood estimation, J. Hydrol., 248, 123–142, 2001.

Xie, P., Chen, G., and Xia, J.: Hydrological frequency calculation principle of inconsistent annual runoff series under changing environments, Eng. J. Wuhan Univers., 38, 6–9, 2005.