

# Identification for the Variant Pattern of Water Resources Quantity with Variance Variation

Binbin Li<sup>1,2</sup> and Ping Xie<sup>1,2,\*</sup>

<sup>1</sup>State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China; <sup>2</sup>Hubei Collaborative Innovation Center for Water Resource Security, Wuhan 430072, China

**Abstract:** In this paper, we propose a method for the issue of identification for the variant pattern of hydrological series with variance variation: Breusch-Pagan test is used to check the inconsistency of the variance of the series, and Hydrological Variation Diagnosis System is designed to identify the variant pattern of variance. We take the third grade region on the upper reaches of Cetian Reservoir of Yongding River in the Haihe River as the study area, and discuss the variance variation of annual runoff series in this area. The result shows that the variance of annual runoff series in this area has a middle variation with a form of jump down in 1967, which is consistent with the real situation.

**Keywords:** Breusch-pagan test, hydrological variation diagnosis system, identification, variance variation.

## 1. INTRODUCTION

From long time ago, the research of hydrological regulation is based on the hydrological series, which is observed long enough and under the same physical causes. Therefore, the hydrological series should have the characteristics of representation, reliability and consistency. With the increasing severely influence from human beings, the hydrological cycle and the physical causes have been changed. So the hydrological regulation is no longer consistent, that means the style and the parameters of distribution are different than before within the hydrological series, which is named as hydrological alteration in the scale of statistics [1]. Consistency implies that all the collected hydrologic time series data belong to the same statistical population having a time invariant mean [2-4].

There are definite component and stochastic component contained in the hydrological series, and the definite component could be decomposed as periodic, tendency and jump component. The hydrological series is stationary if periodic, tendency and jump component have nothing to do with it, which means the physical causes have no change with the time. The statistical regulations, such as the distribution style and parameters, are consistent with no alteration in the time scale of the hydrological series and the hydrological data fluctuates in random based on the mean value. Otherwise, the hydrological series is non-stationary and statistical laws are inconsistent, the physical causes and hydrological parameters are changed in the time scale of the hydrological series. Therefore, the definition of hydrological alteration in the statistical scale is the distribution style and parameters are changed obviously within the hydrological series [5].

There are many different methods to diagnose the hydrological alteration, the methods that used to diagnose the jump alteration such as Sequential Cluster, Rank-Sum Test, Sliding T Test etc. and the Spearman Rank Correlation Test, Kendall Rank Correlation Test etc. used to diagnose the trend alteration [6]. Use of one or two statistical tests for time series analysis is quite common for easy decision making [7-10]. However, the alteration situation is complex in the hydrological series because the alteration may happened in the elements of circle, trend and jump, the methods above could not diagnose the alteration comprehensively. Meanwhile, the classification of alteration degree is not considered in the hydrological series with alteration. Machiwal and Jha recommended that an adequate number of statistical tests must be applied for detecting a particular time series characteristic and the results should be analyzed critically to arrive at a reliable decision [11, 12]. In order to improve the methods above, Xie put forward Hydrological Alteration Diagnosis System (HADS) [1].

In the past, plenty of analysis about hydrological alteration mostly focused on the variations in the mean [13-15], but ignored that the variations in the variance could also lead a hydrological series lost their consistency. In fact, the length of the hydrological series is usually not long, and large section is less impacted by human activity, so the volatility features of the series is not obvious. As series length increases, with the influence of global climate changes and frequent human activities, the issue of inconsistency of hydrological series in the variance or high order moment stands out.

However, most of the research works on the variation of hydrological series in the variance or high order moment are based on time-varying parameters, missing diagnosing the variation in the variance of the series [16]. When the variation form is jump form or other complex forms, method based on time-varying parameters would get a counter-intuitive result. So the method for the issue of identification

for the variant pattern of hydrological series with variance variation was proposed: Breusch-Pagan test is used to check the inconsistency of the variance of the series, and Hydrological Variation Diagnosis System is designed to identify the variant pattern of variance. We take the third grade region on the upper reaches of Cetian Reservoir of Yongding River in the Haihe River as the study area, and detect the variation pattern of variance of the annual flow series in this area.

**2. DIAGNOSIS FOR THE VARIANCE VARIATION OF HYDROLOGICAL SERIES**

Breusch-Pagan test (BP test in short), also called Lagrange multiplier, is used to test for heteroscedasticity in a linear regression model [17-18]. It tests whether the estimated variance of the residuals from a regression are dependent on the values of the independent variables. Suppose that we estimate the regression model

$$Y = \beta_0 + \beta_1 X + \varepsilon \tag{1}$$

where  $\{x_i\} (i = 1, \dots, n)$  is a hydrological series, and obtain from this fitted model a set of values for  $\hat{\varepsilon}$ , the residuals. Then calculate the standardized square of the residuals  $s_i^2$ :

$$s_i^2 = \hat{\varepsilon}_i^2 / \left( \sum_{i=1}^n \hat{\varepsilon}_i^2 / n \right) \tag{2}$$

Perform the auxiliary regression:

$$s^2 = \gamma_0 + \gamma_1 X + \delta \tag{3}$$

Calculate the regression sum of squares  $RSS$  with least square method. The test statistic  $RSS/2$  is asymptotically distributed as  $\chi^2(1)$  under the null hypothesis of homoscedasticity. We extract its value  $\chi_c^2$  from the  $\chi^2$  distribution table, and compare it with the statistic value  $\chi_*^2 = RSS / 2$ . If  $\chi_*^2 > \chi_c^2$ , reject the null hypothesis, and heteroscedasticity (*i.e.* variation in the variance) in the original hydrological series exist. Otherwise, there is no variation in the variance of the original hydrological series.

**3. IDENTIFICATION FOR THE VARIANT PATTERN OF HYDROLOGICAL SERIES**

Hydrological Variation Diagnosis System could diagnose the trend and jump alteration, and it is composed of three parts, *i.e.* preliminary diagnose, detailed diagnose and comprehensive diagnose.

The method of Hydrograph Analysis, Sliding Average Analysis and Hurst Coefficient Analysis are used to diagnose whether the hydrological alteration exist or not in the process of preliminary diagnose, and based on the relationship between Hurst coefficient and the parameter of Fractional Brownian Motion to classify the degree of hydrological alteration, the principles of diagnosis and classification as follows.

The long-term relativity characteristics of temporal series could be described by the Hurst coefficient  $h$ . If  $h=0.5$ , the temporal series is random and the variation tendency is not impacted by now; if  $h$  is approaching to 1.0, it means the variation tendency of future is same to now and the durative of temporal series is positive; if  $h$  is approaching to 0.0, the situation is reverse. The stochastic feature of temporal series is disturbed by the long-term relativity and the alteration happens, and the farther the  $h$  deviated from 0.5, the more apparent the alteration is. Therefore, based on the Hurst coefficient  $h$ , the stochastic feature and alteration of temporal series could be diagnosed, and with the observation temporal series, the  $h$  could be calculated by the Rescaled Range Analysis (Hurst 1965) method as follows:

$$R / S = (c\tau)^h \tag{4}$$

Fractional Brownian motion (Mandelbrot 1969) was born with the combination of random process and fractal theory and it was put forward by Mandelbrot from the ordinary Brownian motion. The long-term relativity could be described by the motion parameter  $H$ , the relationship between  $H$  and correlation function  $C(t)$  as follows (LI 1993):

$$C(t) = - \frac{E[B_H(-t)B_H(t)]}{E[B_H(t)]^2} = 2^{2H-1} - 1 \tag{5}$$

XU *et al.* indicated that the motion parameter  $H$  is equivalent to Hurst coefficient  $h$  in logic and definition, so the correlation function  $C(t)$  of temporal series could be calculated with the  $h$  already known. The hypothesis testing of  $C(t)$  could be carried out under a certain confidence degree  $\alpha$ : if  $C(t)$  is smaller than the critical value  $r_\alpha$ , the long-term relativity is not significant and there is no alteration happens in the temporal series; else the long-term relativity is significant and the alteration exists. There is a situation existing in the application of  $C(t)$ , even if the hypothesis testing passed the testing with the confidence degree  $\alpha$ , it may not passed the testing with the confidence degree  $\beta (\alpha > \beta)$ , that means  $r_\alpha < C(t) < r_\beta$ . In this situation, the alteration is significant with the confidence degree  $\alpha$  and not significant with the confidence degree  $\beta$ , it is named weak alteration in this paper. If  $C(t)$  is more than and equal to  $r_\beta$ , the alteration is obvious and the classification could be considered based on the correlation degree, as Table 1 shows below.

If there is no alteration in the hydrological series by the result of preliminary diagnose, the analysis and investigation of causes should be carried out to confirm the result; else the process of detailed diagnose will be activated.

In the process of detailed diagnose, much more methods are considered to detect the trend alteration and jump alteration. In order to detect the trend alteration, the methods of Linear Correlation Coefficient, Spearman Rank Correlation Test and Kendall Rank Correlation Test are used. In order to detect the jump alteration, the methods of Sequential Cluster, Lee-Heghinian, Rank-Sum Test, Sliding F Test, Sliding T Test, Runs Test, R/S Analysis, Brown-Forsythe, Mann-Kendall and Bayesian are used. After the trend alteration and jump alteration diagnose, the process of comprehensive diagnose will be activated.

Table 1. Classification of alteration degree.

$C(t)$	$h$	Alteration Degree	$C(t)$	$h$	Alteration Degree
$0 \leq C(t) < r_\alpha$	$0.5 \leq h < h_\alpha$	none	$0.6 \leq C(t) < 0.8$	$0.839 \leq h < 0.924$	strong
$r_\alpha \leq C(t) < r_\beta$	$h_\alpha \leq h < h_\beta$	weak	$0.8 \leq C(t) \leq 1.0$	$0.924 \leq h \leq 1.0$	giant
$r_\beta \leq C(t) < 0.6$	$h_\beta \leq h < 0.839$	moderate			

Note:  $h_\alpha = \frac{1}{2}[1 + \ln(1 + r_\alpha) / \ln 2]$ ,  $\alpha > \beta$

In the process of comprehensive diagnose, the comprehensive results of trend diagnose and jump diagnose will be achieved based on the results of detailed diagnose. The fitting degree of hydrological series and the proportion of trend or jump will be estimated by the coefficient of efficiency as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2} \tag{6}$$

where  $Q_{obs,i}$  ( $i = 1, \dots, n$ ) is the observed hydrological series,  $\overline{Q_{obs}}$  is the mean of the observed hydrological series.

For jump variation diagnosis, we assume that  $k$  is the observed hydrological series, then

$$Q_{sim,i} = \frac{1}{k} \sum_{i=1}^k Q_{obs,i} \quad (i = 1, \dots, k) \tag{7}$$

$$Q_{sim,i} = \frac{1}{n-k} \sum_{i=k+1}^n Q_{obs,i} \quad (i = k+1, \dots, n) \tag{8}$$

For trend variation diagnosis,  $Q_{sim,i}$  is the value on the fitting trend line (simple regression line), and the bigger one will be accepted as the result of hydrological alteration. With the investigation and analysis of the actual conditions, the conclusion and the style of alteration will be confirmed at last.

Compared with the traditional simplex alteration diagnosis methods, Hydrological Alteration Diagnosis System could diagnose the alteration comprehensively and considered the alteration degree classification in the hydrological series, it is comprehensive in the indexes and the alteration analysis result is credible. Therefore, the HADS was used to diagnose the hydrological alteration in this paper.

We will make adjustments to the Hydrological Alteration Diagnosis System when identifying the Variant Pattern of Hydrological Series. The function of former preliminary diagnose have been done by Breusch-Pagan test, and other part remains in place and perform the same functions as before. It is important to note, however, that the series inserted into the recognition system is the square of the residual series, and the definite component is the mean component.

#### 4. CASE SUTDY

##### 4.1. Diagnosis the Alteration in the Mean of Annual Run-off Series

In this paper, we take the third grade region on the upper reaches of Cetian Reservoir of Yongding River in the Haihe River as the study area, and choose the data of annual flow series from 1956 to 2000 as research data, as shown in Fig. (1). Under the condition of the first degree of confidence  $\alpha=0.05$  and the second degree of confidence  $\beta=0.01$ , the Hydrological Variation Diagnosis System was used to analyze the alteration of the annual flow series of study area, and alteration diagnosis for the mean of annual runoff series was shown in Table 2.

##### 4.2. Diagnosis the Alteration in the Variance of Annual Runoff Series

The result of preliminary diagnose for the annual flow series  $Y_t$  showed that moderate alteration existed from the results of Table 2: the Hurst coefficient  $h=0.738$ , the first Hurst coefficient confidence limit  $h_\alpha=0.688$  and the second Hurst coefficient confidence limit  $h_\beta=0.735$ ,  $h > h_\beta$ , which did not meet the consistent demand. Detailed diagnose revealed that jump and tendency of the annual flow series were both significant, then entered the process of comprehensive diagnose. The larger of the jump efficiency coefficient or the trend efficiency coefficient (trend efficiency coefficient 24.64%) was selected as final variation form through jump comprehensive and trend comprehensive in the comprehensive diagnose. The final diagnose conclusion was the annual flow series from 1956-2000 had a downward moderate variation, and its trend form was shown in Fig. (2).

The trend line equation was able to get by least-squares regression:

$$\bar{Y} = 837.06 - 0.41t \quad (t = 1, \dots, 45) \tag{9}$$

The variance of the annual runoff series was  $(Y_t - \bar{Y})^2$  and the variance variation chart was shown in Fig. (3).

Using Breusch-Pagan test to check the variance variation of annual runoff series, we got a vlue for the test statistic  $\chi_*^2$  of 5.15, and we found the test value  $\chi_c^2$  of 3.84 from the distribution table,  $\chi_*^2 > \chi_c^2$ . In addition, we can see that the fluctuation of annual runoff series has been a significant

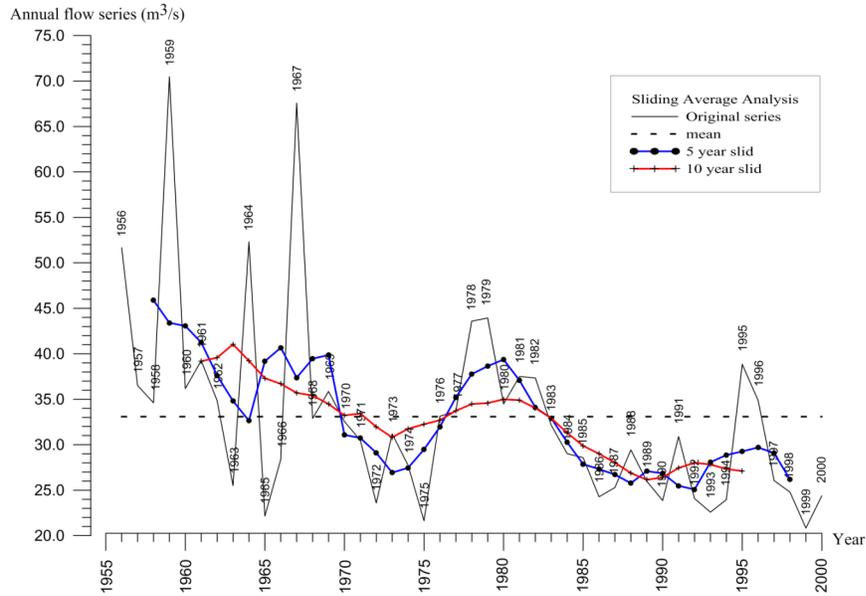


Fig. (1). Sliding average analysis of annual runoff series in study area.

Table 2. The alteration diagnosis for the mean of annual runoff series in the study area.

Hydrological Elements (1956-2000)			Annual Flow Series (m <sup>3</sup> /s)	
Preliminary diagnose	Hurst coefficient		0.738	
	Ateration degree		moderate	
Detailed diagnose	Jump Diagnose	Sliding F Test	1997(+)	
		Sliding T Test	1967(+)	
		Lee-Heghinian	1967(0)	
		Sequential Cluster	1967(0)	
		R/S Analysis	1972(0)	
		Brown-Forsvthe	1983(+)	
		Runs Test	1997(+)	
		Rank-Sum Test	1983(+)	
		Optimal information two degmentation	1996(0)	
		Mann-kendall	1980(-)	
		Bayesian analysis	1967(+)	
		Trend Comprehensive	Linear Correlation Coefficient	+
			Spearmen Rank Correlation	+
Kendall Rank Correlation	+			
Comprehensive diagnose	Jump Comprehensive	Jump Point	1967	
		Comprehensive weight	0.43	
		Comprehensive significant	2(+)	
	Trend Comprehensive	Comprehensive significant	3(+)	
	Compare	Efficiency coefficient (%)	Jump	23.43
Trend			24.64	
Diagnose Conclusion			Trend↓	

Note: “+” means significant, “-” means not significant, “0” means the significant test could not be tested

Table 3. The variance alteration diagnosis of annual runoff series in the study area.

Hydrological Elements (1956-2000)			Square Residual of Annual Flow Series (m <sup>3</sup> /s)	
Preliminary diagnose		Hurst coefficient	0.727	
		Ateration degree	moderate	
Detailed diagnose		Jump Diagnose	Sliding F Test	1996(+)
			Sliding T Test	1967(+)
			Lee-Heghinian	1967(0)
			Sequential Cluster	1967(0)
			R/S Analysis	1970(0)
			Brown-Forsvthe	1967(+)
			Runs Test	1968(+)
			Rank-Sum Test	1975(+)
			Optimal information two degmentation	1964(0)
			Mann-kendall	1972(+)
			Bayesian analysis	1967(+)
			Trend Comprehensive	Linear Correlation Coefficient
		Spearman Rank Correlation		+
Kendall Rank Correlation	+			
Comprehensive diagnose		Jump Comprehensive	Jump Point	1967
			Comprehensive weight	0.55
			Comprehensive significant	3(+)
		Trend Comprehensive	Comprehensive significant	
Compare	Efficiency coefficient (%)		Jump	20.98
		Trend	11.32	
Diagnose Conclusion			1967(+) $\downarrow$	

Note: "+" means significant, "-" means not significant, "0" means the significant test could not be tested

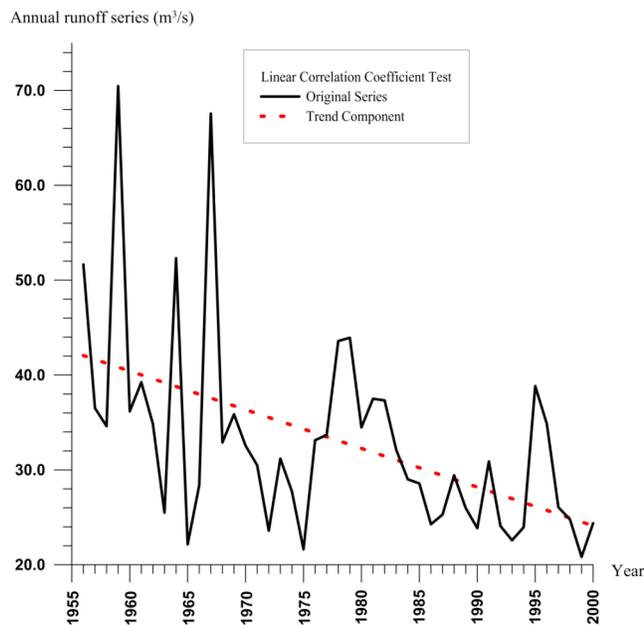


Fig. (2). Trend variation of annual runoff series.

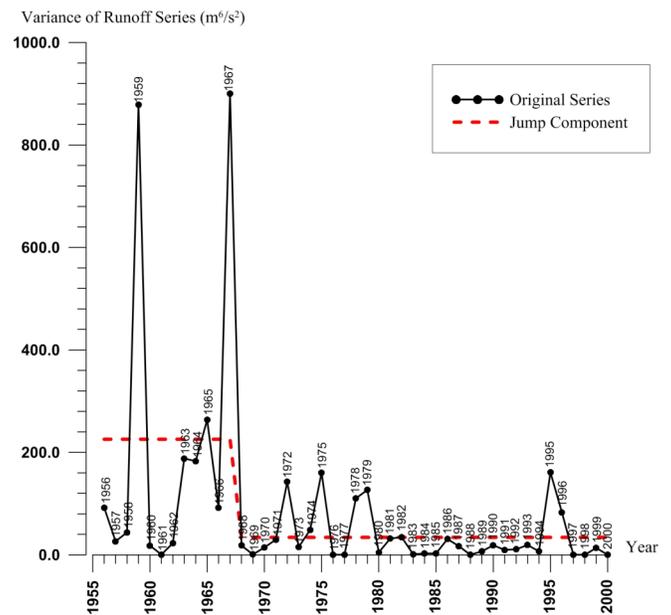


Fig. (3). Jump variation of variance series.

slowdown in Fig. (2), so the variance of the annual runoff series existed a significant variation. Last, Hydrological Variation Diagnosis System was used to analyze the alteration of the variance of the annual flow series, and alteration diagnosis for the variance of annual runoff series was shown in Table 3.

## 5. DISCUSSION

The final diagnose conclusion from Table 3 was that the variance of the annual flow series had a downward moderate variation in 1967, and its jump form was shown in Fig. (3). From Fig. (3), we can see that the variance located in the high value before 1967, then got a significant downward in 1967 and a big fluctuate from 1970 to 1980, and maintained a relatively lower level after 1980 finally.

Cetian reservoir was founded in March 1958, and started to retard flood in 1960. The data of this paper was adapted from 1956 to 2000. Runoff of this area was basically affected by reservoir of detention, in addition to the increased amount of water supply and consumption, so annual runoff series gradually decreased as a whole. Cetian reservoir was a dangerous reservoir before seepage prevention and reinforcement with low operational water level, and its regulation function was limited before 1970, leading a huge variance on the runoff series. Dam heightening and reservoir expanding made the variance a similar rate of decline from 1970 to 1976. The dam operated a low water level for the quality problem in 1976, so a small sustenance was shown from 1976 to 1980. The project of seepage prevention and reinforcement was started after 1980, and operational water level became increased in the after years, so the variance of the annual runoff series maintained a relatively lower level after 1980. Besides, the reservoir held biggest flood in 1967 and played the biggest adjustment function since its putting into operation. This change point in variance is consistent with previous calculation results.

## 6. CONCLUSION

We take Breusch-Pagan test to detect the homogeneity of the variance of the hydrological series, and combine Hydrological Alteration Diagnosis System to identify the variant form. The characteristic of this method is not only able to judge whether or not it occurs, but also able to get the specific variation form and degree, which lays a foundation for the research on the frequency calculation for the hydrological series with the variation in high order moment. However, Hydrological Alteration Diagnosis System is imperfection, which needs to be enhanced.

In this paper, we discuss the variation in the mean and the variance of the hydrological series, and analyze the variation in the third grade region on the upper reaches of Cetian Reservoir. Results show that the calculation is consistent with the real situation, and the method can be further improved and extended.

## CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

## ACKNOWLEDGEMENTS

This work is supported by the National Nature Science Foundation of China (Grant No.51190094, Grant No.50979075 and Grant No.51179131), the special research project of hubei province water conservancy (HBSLKJZX201301) and Guangdong province water conservancy science and technology innovations (2011-01).

## REFERENCES

- [1] P. Xie, G. C. Chen, H. F. Lei, *et al.* "Surface water resources evaluation methods on changing environment," Beijing: science press, 2009.
- [2] M. K. Goyal, and V. Gupta, "Identification of Homogeneous Rainfall Regimes in Northeast Region of India using Fuzzy Cluster Analysis," *Water Resources Management*, vol. 28, pp. 4491-4511, 2014.
- [3] J. Olsson, and K. Foster, "Short-term precipitation extremes in regional climate simulations for Sweden," *Hydrology Research* vol. 45, pp. 479-489, 2014.
- [4] D. Raje, "Changepoint Detection in Hydrologic Series of the Mahanadi River Basin Using a Fuzzy Bayesian Approach." *Journal of Hydrologic Engineering* vol. 19, pp. 687-698, 2014.
- [5] Q. Zhang, *et al.*, "Reservoir-induced hydrological alterations and environmental flow variation in the East River, the Pearl River basin, China," *Stochastic Environmental Research and Risk Assessment* vol. 28, pp. 2119-2131, Dec. 2014.
- [6] F. Chebana, *et al.*, "Testing for multivariate trends in hydrologic frequency analysis," *Journal of Hydrology*, vol. 488, pp. 519-530, 2013.
- [7] L. Z. Jiang, *et al.*, "Assessment of Hydrologic Alterations Caused by the Three Gorges Dam in the Middle and Lower Reaches of Yangtze River, China," *Water*, vol. 6, pp.1419-1434, 2014.
- [8] G. L. Xu, *et al.*, "Temporal and spatial variation of water level in urbanizing plain river network region." *Water Science and Technology*, vol. 69, pp. 2191-2199, 2014.
- [9] H. Du, *et al.*, "Variations and statistical probability characteristic analysis of extreme precipitation events under climate change in Haihe River Basin, China," *Hydrological Processes*, vol. 28, pp. 913-925, 2014.
- [10] T. P. Worrall, *et al.*, "The identification of hydrological indices for the characterization of macroinvertebrate community response to flow regime variability," *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* vol. 59, pp. 645-658, 2014.
- [11] D. Machiwal, M.K. Jha, "Hydrologic Time Series Analysis: Theory and Practice," Berlin: Springer-Verlag, 2012
- [12] A.W. Jayawardena, and W.H. Lau, "Homogeneity tests for rainfall data," *Journal of the Hong Kong Institution of Engineers*, pp 22-25, September 1990.
- [13] G. Lacombe, and M. M.Cartney, "Uncovering consistencies in Indian rainfall trends observed over the last half century," *Climatic Change*, vol. 123, pp. 287-299, 2014.
- [14] H. M. Kang, and F. Yusuf, "Homogeneity tests on daily rainfall series in Peninsular Malaysia," *International Journal of Contemporary Mathematical Sciences*, vol. 7, pp. 9-22, 2012.
- [15] M. Piccarreta, *et al.*, "Changes in daily precipitation extremes in the Mediterranean from 1951 to 2010: the Basilicata region, southern Italy," *International Journal of Climatology*, vol. 33, pp. 3229-3248, 2013.
- [16] V. Gabriele, A. S. James, S. Francesco, *et al.* "Flood frequency analysis for nonstationary annual peak records in an urban drainage basin," *Advances in Water Resources*, vol. 32, pp. 1255-1266, 2009.

- [17] H. Herwartz, "Testing for random effects in panel data under cross sectional error correlation - A bootstrap approach to the Breusch Pagan test," *Computational Statistics & Data Analysis*, vol. 50, pp. 3567-3591, 2006.
- [18] R. Modarres, and T. Ouarda, "Modelling heteroscedasticity of streamflow times series," *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, vol. 58, pp. 54-64, 2013.

---

Received: September 16, 2014

Revised: December 23, 2014

Accepted: December 31, 2014

© Li and Xie; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.