Journal Homepage: - www.journalijar.com
INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)
ISSN NO. 2320-5407
Article DOI:10.21474/IJAR01/2588
DOI URL: http://dx.doi.org/10.21474/IJAR01/2588

RESEARCH ARTICLE

THE SPEARMAN's RHO TEST FOR DETECTING TRENDS IN SERIALLY CORRELATED HYDROLOGICAL SERIES.

Jian Xin Wang and Jian Wang.
School of science, Shandong University of Technology, Zibo, China.

Manuscript Info

Manuscript History

Received: 27 October 2016
Final Accepted: 25 November 2016
Published: December 2016

Key words:-
Spearmans rho test, Serial correlation, Linear trend, Pre-whitening, ESS

Abstract

When sample data are serially correlated, the presence of serial correlation in time series will affect the ability of the test to correctly assess the significance of trend. Hence, this paper discussed two methods to eliminate the influence of serial correlation on the SR test about a time series only consist of an AR(1) process and combined series consist of a linear trend and an AR(1) process, respectively. This study investigated using Monte Carlo simulation generated some time series. Simulation demonstrates that when no trend exists within time series, Pre-whitening and ESS can effectively limit the influence of serial correlation on the SR test. When trend exists within time series, these two kinds of approach has a similar results, both the pre-whitening and the ESS approaches cannot properly limit the influence of serial correlation on the SR test.

Introduction:-

In recent years, with growing concerns about the implication of green-house gases on the environment, more and more researches has been studied the hydrological series by some nonparametric statistical test. For example, the Spearman's rho (SR) test and the Mann-Kendall(MK) test has been popularly used to assess the significance of trend in hydrological time series.The Spearman's rho (SR) test was initially proposed by C.Spearman in 1904.It has been widely used in trend detection studies. The previous papers has been large studied some properties of the MK test in hydrological time series. Examples include the studies of Yue(2002), Yue and Wang(2004),Hamedand Rao(1998),and others. Therefore, this paper studies the Spearman's rho test for detecting trends in serially correlated series.

As the Pre-whitening approach, it is in order tolimit the influence of serial correlation. The procedure has applied by Douglas et al.,2000;Hamilton et al.,2001; Burn and Hag Elnur,2002; Yue et al(2002b) and others in trend detection studies. Such as the study of Yue et al(2002b) documented that per-whitening eliminating the influence of serial correlation on the MK test. Besides, other method eliminating the effect of serial correlation is called ESS. Hamed and Rao(1998) developed an empirical formula for computing ESS to modify the variance of the MK statistic. In some literatures, such as Yue et al.(2002b), effective or equivalent sample size(ESS) has been proposed to modify the variance of the MK statistic to reduce the influence of the presence of serial correlation on the MK test.

The Spearman's rho (SR) test is one of the most popularly applied tests to detects trends in hydrologic time series. The remainder of the paper is organized as follows: in the second section we provide a brief description of the SR test, then we discuss the influence of the AR(1) process on the SR test when there is no trend. Next, we study the

Corresponding Author:- Jian Xin Wang.
Address:- School of science, Shandong University of Technology, Zibo, China.

pre-whitening and the ESS approaches of the time series only contain an AR(1) process and time series consist of a linear trend and an AR(1) process on the SR test, respectively.

Spearman's rho test:-

The most famous rank statistics methods is Wilcoxon Rank-Sum test by F.Wilcoxon in 1945. But, the Spearman's rho test method was initially proposed by C.Spearman in 1904. The Spearman's rho test, also called SR test, it has been widely used in hydrological series trend detectionstudies. Assuming, we have a paired-sample data$\binom{x_1}{y_1}, \binom{x_2}{y_2}, \ldots, \binom{x_n}{y_n}$.The R_i are rankof x_i in the data set $\{x_1, x_2, \ldots, x_n\}$, and Q_i are rank of y_i in the data set $\{y_1, y_2, \ldots, y_n\}$. Here's the basic idea: replace the data x_i and data y_i with rank R_i and Q_i,respectively. And, we can generatea new paired-sample data$\binom{R_1}{Q_1}, \binom{R_2}{Q_2}, \ldots, \binom{R_n}{Q_n}$.

Then, compute the rank correlation coefficient

$$r_s = \left[\sum_{i=1}^{n}(R_i - \bar{R})(Q_i - \bar{Q})\right] / \left[\sqrt{\sum_{i=1}^{n}(R_i - \bar{R})^2}\sqrt{\sum_{i=1}^{n}(Q_i - \bar{Q})^2}\right]$$

where

$$\bar{R} = \sum_{i=1}^{n} R_i \,/\, n = (n+1)/2, \bar{Q} = \sum_{i=1}^{n} Q_i \,/\, n = (n+1)/2$$

and

$$\sum_{i=1}^{n}(R_i - \bar{R})^2 = \sum_{i=1}^{n}(Q_i - \bar{Q})^2 = n(n^2 - 1)/12$$

Therefore

$$r_s = [12\sum_{i=1}^{n} R_i Q_i - 3n(n+1)^2] / [n(n^2 - 1)]$$

It is obvious that the paired-sample data$\binom{x_1}{y_1}, \binom{x_2}{y_2}, \ldots, \binom{x_n}{y_n}$ and$\binom{g(x_1)}{y_1}, \binom{g(x_2)}{y_2}, \ldots, \binom{g(x_n)}{y_n}$have a equal rank correlation coefficient, when the g (.) is a strictly monotone increasing function. Therefore, the rank correlation coefficient can describe whether it had upward or downward trend between two variable. Here, we assume that the population (X,Y) are continuous random variables. The null hypothesis H_0 is that X and Y is independent. The alternative hypothesis H_1 is that X and Y is positive correlation, or X and Y is negative correlation.

The Influence of the AR(1) process on the SR test:-

Yue(2002b) verify Von Storch(1995) demonstrated that the existence of positive serial correlation in a time series increases the probability that the MK test will detect a significant trend, i.e. serial correlation increases the type I error. Now, we test the existence of serial correlation in a time series the influence of the SR test detect a trend.

Similarly, we generate various series with different lag-1 serial correlation coefficients using Monte Carlo simulation as follows:

$$X_t = \mu_X + r_1(X_{t-1} - \mu_X) + \sigma_X\sqrt{1 - r_1^2}\,\varepsilon_t \rightarrow (1)$$

Where $\mu_X$is the mean of X_t,$\sigma_X^2$is the variance of X_t,andε_t is a normally distributed random variable with mean $\mu_\varepsilon = 0$ and variance$\sigma_\varepsilon^2 = 1$.The simulation generated 5000 time series of the AR(1) processes using Equation (1) for each sample size (n=10,50,100,150) with different given r_1. Besides, we assume that $r_1 = -0.9(0.1)0.9$,$X_0 = \mu_X$,$\mu_X = 1.0$ and $\sigma_X^2 = 0.5$.Given that the series is generated without trend, the rejection rate can be computed as$P_{rej} = N_{rej} / N$, in which N_{rej} is the number of the computed SR statistics falling in the critical regions and N is the number of simulations(N =5000).

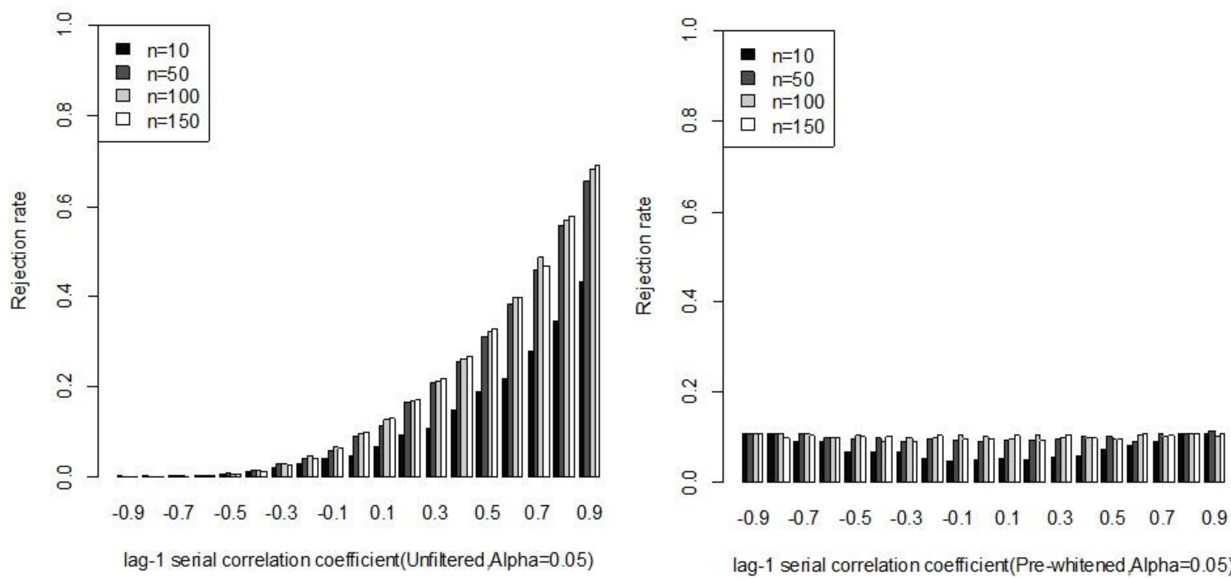

Figure 1:-Effect of serial correlation on the rejection rate(a:Unfiltered b:Pre-whitened)

Through the Figure 1a, it can be seen that along with the positive lag-1 serial correlation increases, the rejection rate also increases. But, we also find that with the negative lag-1 serial correlation increases, the rejection rate contrary decreases. Therefore, similar results with on the MK test, there is an increasing tendency to detect a significant trend, whereas in fact it is no trend,as the positive lag-1 serial correlation increases.Besides, the influence of the negative lag-1 serial correlation on the SR test may tend to underestimate the probability of detecting trends. In addition, we also see that as the positive lag-1 serial correlation coefficient increases, the P_{rej} of the minimum sample sizes have a significantly different for others. But, the P_{rej} of larger sample sizes is not sensitive to the sample sizes selected.

Pre-whitening has been used by von Storch(1995), Zhang et al.(2000,2001), Yue(2002b), and so on to reduce the influence of an AR(1) component on the application of the MK test. Now, we still make the method to reduce the influence of an AR(1) component on the application of the SR test by the following formula :

$$Y_t = X_t - r_1 X_{t-1}$$

The rejection rates can be estimated from the residual time series Y_t. Through the Figure 1b,we can see that these rejection rate are almost the same as the nominal significance level, i.e. the serial correlation has effectively been removed from the series on the SR test. Similarly, Figure 1b also demonstrates that the sample sizes can’t be too small.

The SR Test improved by effective sample size;-

How does the existence of serial correlation influence the rejection rate of the null hypothesis of no trend for the SR test? The probability density function of SR statistic of the generated series with sample size n=100 for difference serial correlation are plotted in Figures 2. These diagrams indicate that the probability density function of SR statistic were asymptotic normal distribution. Besides, we find that as the positive lag-1 serial correlation coefficient increases, the variance of SR statistic also increases. Hence, the numbers of SR statistic falling in the critical regions increases as positive serial correlation increases. On the contrary, less number of samples falls in the critical regions. This explains why the existence of serial correlation causes these results for SR test in Figures 1.

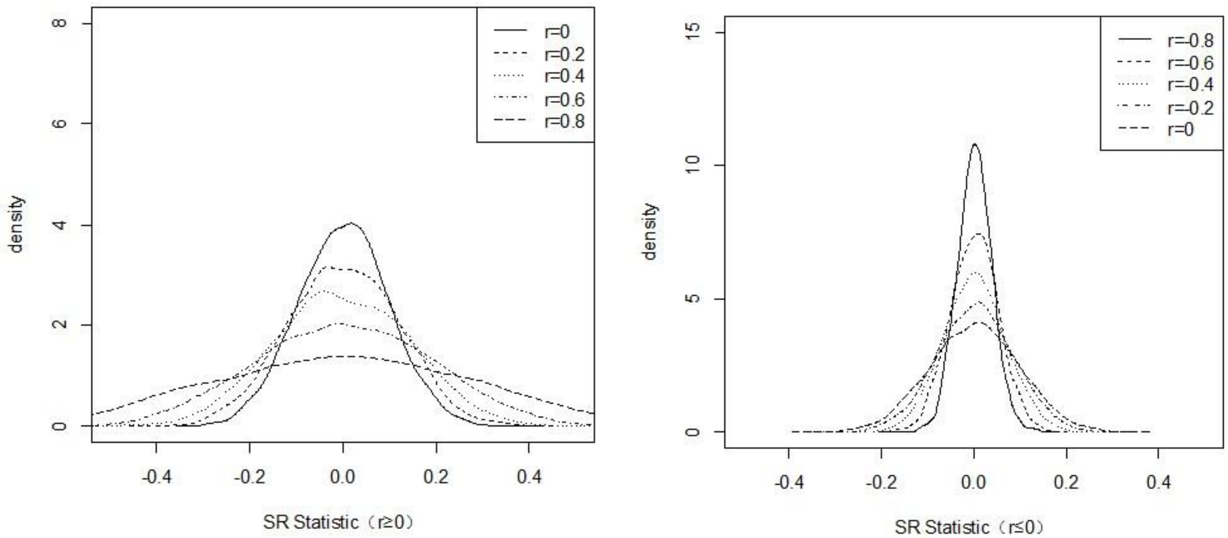

Figure 2:- Probability density functions of the SR statistics

Next we discuss the ESS method to eliminate the influence of serial correlation on the SR test. Bayley and Hammersley(1946) provided a formula for the ESS:

$$n^* = \frac{n}{1 + 2 \times \sum_{k=1}^{n-1}(1 - \frac{k}{n}) \cdot r_k}$$

Where$r_k$is the lag-k serial correlation coefficient, $n^*$is the ESS. In the paper, we discuss the influence of lag-1 serial correlation on the SR test, and it is given by (Matalas and Langbein,1962)

$$n^* = \frac{n}{1 + 2 \times \frac{r_1^{n+1} - n \cdot r_1^2 + (n-1) \cdot r_1}{n(r_1 - 1)^2}}$$

The preceding section demonstrated that serial correlation in time series alter the variance of the SR statistic. The modified variance $V^*(r_s)$ by ESS is given by

$$V^*(r_s) = V(r_s) \cdot \frac{n}{n^*}$$

where n is the actual sample data. $V(r_s)$is the variance of the SR statistic, and it is given by

$$V(r_s) = \frac{1}{n-1}$$

Therefore, we can obtain that the new critical regions of the SR statistic can be approximately given by

$$r_s > u_{1-\alpha}\sqrt{V^*(r_s)} \text{ or } r_s < u_{\alpha}\sqrt{V^*(r_s)}$$

The Figure 3 is the rejection rate of the null hypothesis after correction use the ESS approach on the SR test by Monte Carlo simulation. It΄s demonstrated that the ESS approach is an effective approach for eliminating the influence of serial correlation on the SR test with the large sample sizes when there is no trend in a time series. In addition, we can find that the Pre-whitening approach is better than ESS approach for eliminating the influence of serial correlation by contrasting the Figure 2 and Figure 3. We find thatthe rejection rate of the null hypothesis after correction use the ESS approachare still much higher than the pre-assigned significance level for n=10 and 50 in $-0.9 < r < -0.3$.

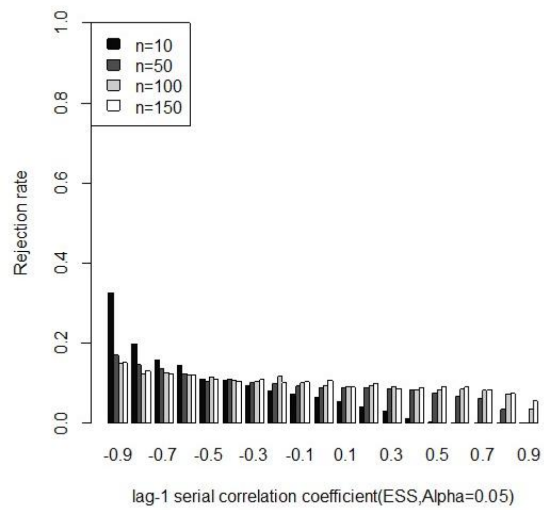

Figure 3:- Rejection rate after correction by the ESS approach

Combined series:-

In the preceding section, we discussed two methods to eliminating the influence of serial correlation on the SR test when there is no trend in the time series. However, in practice, we need to explore trend detection is to assess by SR test method whether or not trend is statistically significant when it exists in a tested series. Therefore, we will research a time series that consists of a linear trend and an AR(1)process.

$$X_t = T_t + A_t$$

where the $T_t = \beta(t-1), t = 1,2,3,...,n$ is a linear trend and A_t is an AR(1) process. Simulation generated 5000 AR(1) samples with n=100 and $\mu_A = 1.0$ for each given $r_1 = 0(0.1)0.9$. A linear trend with slope $\beta = 0(0.002)0.008$ is superimposed on each of the generated AR(1) processes.

The rejection rates corresponding to the given $r_1 = 0(0.1)0.9$ are displayed in the Figure 4. The Figure 4 indicates that when both upward linear trend an positive AR(1) exist, the case of β=0.002 and β=0.004, the effect of serial correlation on the SR test is similar to the case of β=0. Besides, we find that as the slope increases, the rejection rates of SR test also increases for same lag-1 serial correlation coefficient. As the case of β=0.006 and β=0.008, it can be seen that with the lag-1 serial correlation coefficient increases, the rejection rates decreases a little. However, the rejection rates are very big for difference lag-1 serial correlation coefficient. These results are similar to rejection rates of time series with both trend and AR(1) processes on MK test, which is the same as observed in Yue et al. (2004).

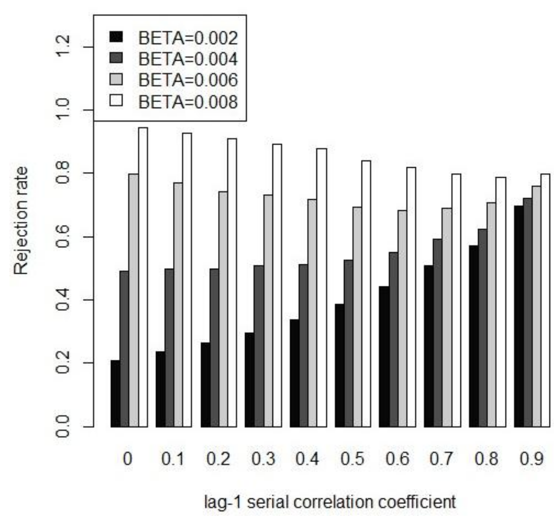

Figure 4:- Rejection rates of time series with both trend and AR(1) processes on the SR test at Alpha=0.05

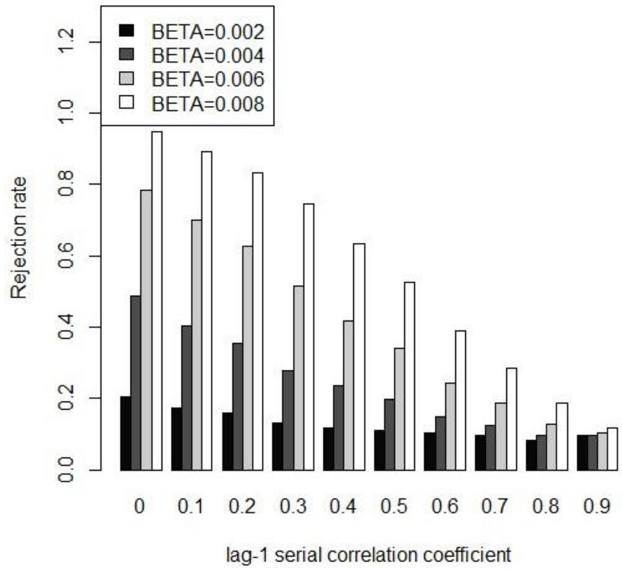

Figure 5:-Rejection rates of time series with both trend and AR(1) processes on the SR test after pre-whitening approach at Alpha=0.05

The rejection rates of time series with both trend and AR(1) processes on the SR test after pre-whitening approach and correction by the ESS approach at α=0.05 are displayed in the Figure 5 and Figure 6, respectively. We are very surprised to find that the rejection rates of after these two kinds of approach are similar. However, we have to realize that as the lag-1 serial correlation coefficient increases, the rejection rates decreases, i.e. the rejection rates in the case of $r_1 \neq 0$ are much less than those of $r_1 = 0$. This indicates that both the pre-whitening and the ESS overcorrected the effect of the true serial correlation on the SR test, and resulted in underestimation of the significance of the existing trend. Therefore, it can be seen that when some trend does exist within a time series, both the pre-whitening and the ESS approaches

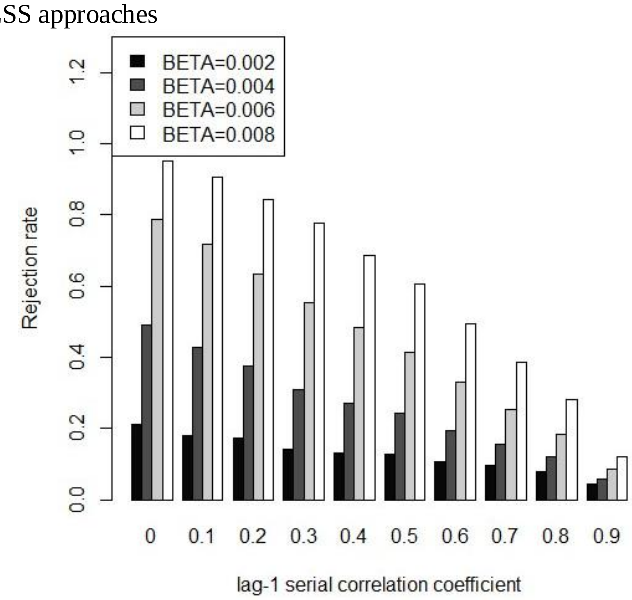

Figure 6:- Rejection rates of time series with both trend and AR(1) processes on the SR test after correction by the ESS approach at Alpha=0.05

cannot properly limit the influence of serial correlation on the SR test. This result is similar to the ESS approach limit the influence of serial correlation on the MK test. The previous paper of Yue et all.(2002b) demonstrated that removal of trend from time series does not affect the true AR(1)process. Hence, we may be to remove the existing trend first from the series if it exists.

Conclusions:-

This study investigated the influence of AR(1)process and both exist AR(1)process and a linear trend in a time series for the rejection rates by the Monte Carlo simulation experiments on the SR test. The simulation experiments demonstrated that the existence of serial correlation alters the variance of the SR statistic, whereas it does not alter the mean and the distribution type of the SR statistic. In such cases the existence of positive serial correlation in a time series will increase the probability that the SR test detects a significant trend. On the contrary, we find that as the negative serial correlation decreases, the rejection rates also decreases, i.e. the effect of the negative serial correlation on the SR test may tend to underestimate the probability of detecting trends. Besides, we find that the sample sizes can't be too small. Furthermore, this paper demonstrated that the pre-whitening process has effectively removed the autocorrelation from the AR(1) process on the SR test. However, should a trend exist, the pre-whitening process is inappropriate to eliminate the impact of the existence of serial correlation on the SR test. Similarly, it shown that when is not present in a time series, the ESS approach can effectively eliminate the influence of serial correlation on the SR test. And if exists a linear trend in time series, the ESS approach is also inappropriate. Therefore, if trend can be approximated by a linear trend, we can removed the trend as a first step.

This paper studies the Spearman's rho test for detecting trends in serially correlated series. Through comparative analysis(Yue 2002,Yue 2004 and so on), we found that the property of detecting trends in serially correlated series are similar between the Spearman's rho test and the Mann-Kendall test.

References:-

1. **Burn,D.H.and Hag Elnur, M.A. (2002):**Detection of hydrological trends and variability. J. Hydrol., 255(1-4), 107-122.
2. **Hamed, K.H. and Rao, A.R. (1998):**A modified Mann-Kendall trend test for autocorrelated data. J. Hydrol., 204, 182-196.
3. **Hirsch, R.M. and Slack, J.R. (1984):** Non-parametric trend test for seasonal data with serial dependence. Water Resources Research., 20(6),727-732.
4. **Jing Long,W.and Ling,X.J. (2006):**Nonparametric statistical analyses.Higher Education Press, Beijing.
5. **Kendall,M.G.(1975):** Rank Correlation Methods. Griffin,London.
6. **Mann,H.B. (1945):**Nonparametric tests against trend. Econometrica., 13,245-259.
7. **Yue, S., Pilon, P. and Cavadias, G. (2002a):**Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. J. Hydrol., 259(1-4),254-271.
8. **Yue,S.,Pilon,P.,Phinney,B.and Cavadias, G. (2002b):**The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrol. Processes., 16,1807-1829.
9. **Yue, S.and Wang, C.Y. (2004):**The Mann-Kendall Test Modified by Effective Sample Size to Detect Trend in Serially Correlated Hydrological Series.Water Resour. Manage., 18,201-218.
10. **Yue,S.and Wang,C.Y. (2002):**The applicability of pre-whitening to eliminate the influence of serial correlation on the Mann-Kendall test.Water Res.Res., 38(6),10.1029/2001WR000861,4-1-7.