ON THE VARIABILITY OF TREND TEST RESULTS

[1]A. Ramachandra Rao and M. Azli

[1]Department of Civil Engineering
Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia.
[1]Tel: (603) 7967-5266, (6012) 314-5757 ; Fax: (603) 7967-5318
[1]E-mail: rao@um.edu.my, rao@ecn.purdue.edu

[1]Corresponding author: A. Ramachandra Rao

Abstract

Trend tests are used to investigate statistical significance of trends. The popular Mann–Kendall (MK) trend test was originally proposed for random data. It was later modified to handle correlated data. After the scaling hypothesis was introduced, the MK test was further modified to accommodate it. The results from these three versions of the MK test can be very different. The objective of the present paper is to illustrate these variations in the MK trend test results. Not considering these variations would lead to spurious conclusions about statistical significance of trends in data with associated erroneous deductions. Monthly temperature data from Malaysia are used for illustration.

Keywords: *Mann–Kendall trend test; Correlations; Scaling hypothesis; Monthly temperatures; Malaysia*

1.0 Introduction

Trend tests have been used to investigate the impacts of climate change and variability in hydrologic time series in different parts of the world. Trends in various series have been investigated: in Japanese precipitation series (Xu *et al.*, 2003); in Yangtze basin in China (Zhang *et al.*, 2006); in precipitation in Seoul, Korea (Wang *et al.*, 2006). Earlier studies include those by World Meteorological Organization (1988), Mitosek (1992), Chiew and McMahon (1993) and Burn (1994). In many of these studies tests based on assumption of randomness in data are used. With the exception of papers by Hamed (2008) and Kumar *et al*. (2009) the effect of scaling on trend detection is not considered.

A widely used non-parametric test for detecting trends in time series is the Mann–Kendall (MK) test (Mann, 1945; Kendall, 1975).The null hypothesis in the MK test is that data are random and independent, i.e. there is no trend or serial correlation among observations. However, observed hydrologic and climatic time series, especially monthly data, are generally autocorrelated. The autocorrelations in observed data will lead to misinterpretation of results of trend tests. This situation was recognized early by Cox and Stuart (1955) who stated that "positive serial correlation among the observations would increase the chance of significant answer even in the absence of a trend". Problems in interpreting confusing trend test results explain in part the variety and even contradictory results reported from them.

Modifying the tests for trends to account for the effect of serial correlation in data and using the modified tests has been the approach used by several investigators. Lettenmaier (1976) and Hirsch and Slack (1984) were early investigators who considered the effect of serial correlation on the results from trend tests. Hamed and Rao (1998) introduced a modified MK trend test for autocorrelated data with arbitrary correlation structure.

The effect of scaling on trend detection was investigated by Hamed (2008). By using simulated fractional Gaussian series, Hamed (2008) demonstrated that the null hypothesis of no trend was rejected by the MK trend test by as small a percentage as ten percent for random data to as high as sixty percent for data with the Hurst parameter H of 0.9. The number of rejections increases with increasing H and decreases with lower significance levels. Because of the symmetry of the test statistic, which is not affected by scaling, both the false positive and negative trends occur in equal proportions. These results point out the importance of testing for scaling effects in trend tests. The objective of the research reported herein is to present the variation in results from trend tests depending on the assumptions on which the tests are based.

Monthly Malaysian temperature data from two stations are used in the study. Temperature data from the past three decades have been selected for study because global warming and its effects became prominent during this period (Fig. 1) Climate change and its effects started attracting attention and investigation during this period. Monthly temperature data from Alor Setar in Kedah and Senai in Johor are used in the study (Fig. 2). Alor Setar is located in the north of Peninsular Malaysia while Senai is located in the south. The duration of data is from 1979 to 2007.

Fig. 1
Fig. 2

Three tests, the MK test (Mann, 1945; Kendall, 1975), the modified MK test (Hamed and Rao, 1998), and the MK test under the scaling hypothesis (Hamed, 2008), are used in the study. Because the details of the test are available in these references, they are briefly discussed next.

2.0 Tests used in the study

2.1 MK test

Consider a time series $X = [x_1, x_2, \ldots x_n]$. The test statistic *S* is computed by Eq. (1).

$$S = \sum_{i<j} a_{ij} \tag{1}$$

$$a_{ij} = sgn(x_i - x_j) = sgn(R_i - R_j) = \begin{cases} 1 & x_i < x_j \\ 0 & x_i = x_j \\ -1 & x_i > x_j \end{cases} \tag{2}$$

where R_i and R_j in Eq. (1) are the ranks of observations x_i and x_j respectively of the time series. Assuming that the data are independent and identically distributed, Kendall (1975) showed that

$$\left.\begin{array}{l} E(S) = 0 \\ V_0(S) - n(n-1)(2n+5)/18 \end{array}\right\} \tag{3}$$

Kendall (1975) also showed that the significance of trends can be tested by comparing the standardised variable u_1 in Eq. (4) with the standard normal variate at a significance level α.

$$u_1 = \begin{cases} \frac{S-1}{\sqrt{V_0(S)}}, & S > 0 \\ 0 & S = 0 \\ \frac{S+1}{\sqrt{V_0(S)}} & S < 0 \end{cases} \tag{4}$$

The basic assumption in this test is that the data are random. If the data are correlated then the correlation may be removed by pre-whitening the data. Alternatively, the variance $V_0(S)$ may be modified to account for the correlation. Such a modification to the MK test proposed by Hamed and Rao (1998) is discussed below.

2.2 Modified MK test

$V_0(S)$ in Eq. (4) is recalculated in this test as $V^*(S)$ by using Eq. (5).

$$V^*(S) = V_0(S) \times \frac{n}{n_s^*} = \frac{n(n-1)(2n+5)}{18} \times \frac{n}{n_s^*} \tag{5}$$

In Eq. (5), (n/n_s^*) represents a correction to $V_0(S)$ because of the autocorrelations in the data. The approximation used for (n/n_s^*) is the empirical expression in Eq. (6).

$$\frac{n}{n_s^*} = 1 + \frac{2}{n(n-1)(n-2)} \times \sum_{i=1}^{n-1} (n-i)(n-i-1)(n-i-2) \times \rho_s(i) \tag{6}$$

In Eq. (6), $\rho_s(i)$ are the autocorrelation coefficients of the ranks of the data.

As the ranks of the observations of observations are used in Eq. (6), $V^*(S)$ is computed without using either the data or their autocorrelation function. In the present study, significant correlation coefficients up to $N/10$ of N ranks are used. The modified statistic u_2 is computed and tested for significance.

$$u_2 = \begin{cases} \frac{S-1}{\sqrt{V^*(S)}}, & S > 0 \\ 0 & S = 0 \\ \frac{S+1}{\sqrt{V^*(S)}} & S < 0 \end{cases} \tag{7}$$

2.3 MK test under the scaling hypothesis

In this test, the data are detrended by using Sen's (1968) non-parametric trend estimator. The scaling coefficient H is obtained by maximising log likelihood function in McLeod and Hipel (1978). This estimate of H is approximately normally distributed for the uncorrelated case when true H is 0.5 with the mean and variance given by Eqs. (8).

$$\left.\begin{aligned} \mu_H &= 0.5 - 2.874n^{-0.9067} \\ \sigma_H &= 0.77654n^{-0.5} - 0.0062 \end{aligned}\right\} \tag{8}$$

The significance of H is tested by using μ_H and σ_H in Eqs. (8). If H is significant, the trend test under the scaling hypothesis is conducted. The modified variance of the test statistic is computed by using Eq. (9).

$$V(S) = \sum_{i<j} \sum_{k<l} 2/\pi \times \sin^{-1}(r_{ijkl}) \tag{9}$$

where:

$$r_{ijkl} = \frac{\rho_{jl} - \rho_{il} - \rho_{jk} + \rho_{ik}}{\sqrt{(2-2\rho_{ij})(2-2\rho_{kl})}} \tag{10}$$

The variance $V(S)$ in Eq. (9) is corrected for bias by multiplying it with the factor B in Eq. (11).

$$B = a_0 + a_1 H + a_2 H^2 + a_3 H^3 + a_4 H^4 \tag{11}$$

The coefficients a_0, a_1,... a_4 in Eq. (11) are functions of the sample size n and are found in Hamed (2008). The modified test statistic u_3 is computed by using the modified variance and Eq. (4). If u_3 is significant, then the trend is significant; otherwise, it is not. The test under the scaling hypothesis is conducted only if the decisions from MK or modified MK tests are significant.

3.0 Data analysis and results

3.1 Results of the MK test

The values of the statistic S and the variance $V_0(S)$ for the data from Alor Setar are 10,457 and 4,702,775, respectively. The statistic u_1 is 4.822, and is significant at 10%, 5% and 2.5% levels. Therefore the conclusion is that the Alor Setar temperatures have a strong positive trend. The values of S and $V_0(S)$ for data from Senai in Johor are, respectively, 11,650 and 4,702,775. The value of u_1 for Senai is 5.372 which is larger than u_1 for Alor Setar. Therefore the conclusion from this test may be that the positive trend in Senai data is stronger than that in Alor Setar. Depending only on these results one may conclude that there is a north-south gradient in the Malaysian temperature trend. But we will have to consider the strong correlation in monthly temperature data and perform the modified MK test.

3.2 Results of the modified MK test

For the data from Alor Setar, the values of the modified variance $V^*(S)$, the variance inflation factor $V^*(S)/V_0(S)$, and the statistic u_2 are 10,452,906, 2.223, and 3.234, respectively. u_2 is smaller than u_1 which is 4.822 due to the effect of correlation in the data. u_2 is also significant at 10%, 5%, and 2.5% levels, and is positive which indicates an increasing trend in temperature. The values of $V^*(S)$, $V^*(S)/V_0(S)$, and u_2 for Senai are 28,410,496, 6.041, and 2.186, respectively. u_2 for Senai has decreased from 5.372 to 2.186, a reduction of 59.3%. u_2 is positive for both data sets indicating increasing trends in temperature. For both sets of data, u_2 is significant at 10%, 5%, and 2.5% significance levels. u_2 for Senai is smaller than that for Alor Setar which is opposite to the behaviour of u_1. As u_2 is statistically significant, the MK test under the scaling hypothesis is conducted to test the significance of the test statistic.

3.3 Results from the MK test under the scaling hypothesis

Before performing the MK test under the scaling hypothesis, Hurst's parameter *H*, and mean and standard deviation of *H* are estimated. The statistical significance of *H* is tested and if *H* is found significant, the MK test is performed under the scaling hypothesis. Otherwise, inferences from the previous tests are accepted. Accordingly, the *H* value for Alor Setar, its mean and standard deviation are estimated to be 0.92, 0.486, and 0.035, respectively. The *H* value for Senai is 0.90, and its mean and standard deviation are the same as for Alor Setar data. The *H* estimates for both data sets are statistically significant. They are also close to unity which indicates that MK test should be run under the scaling hypothesis.

The bias-corrected variance $V(S)$, the variance inflation factor $V(S)/V_0(S)$, the bias correction factor *B*, and the statistic u_3 for the test under the scaling hypothesis are computed. For the Alor Setar data, the values of $V(S)$, $V(S)/V_0(S)$, *B*, and u_3 are 36,290,000, 7.717, 3.196, and 0.971, respectively. u_3 is statistically insignificant, and has decreased to 0.971 from u_2 of 3.234. Because of the high *H* value, the variance inflation factor $V(S)/V_0(S)$ is quite large and so is the bias correction factor *B*. Consequently the trend is statistically insignificant. The values of $V(S)$, $V(S)/V_0(S)$, *B*, and u_3 for the Senai data are 54,190,000, 11.523, 2.255, and 1.054 respectively. In this case, u_3 is also insignificant indicating the statistical insignificance of the trend in the temperature data.

4.0 Summary and conclusions

As the example discussed above clearly illustrates, the MK test statistic is strongly affected by correlation in the data and by the scaling factor *H*. Conclusions drawn without considering these factors can be misleading or even wrong. Although the trend statistic u_3 is insignificant, plots of the temperature data in Malaysia during these years show an overall, general, gradual warming trend. But this trend is statistically insignificant in the data from all the stations. The situation is "mixed" in the sense that there are increasing but statistically insignificant trends in Malaysian monthly temperature data.

Acknowledgements

We would like to extend our gratitude Dr. Khaled H. Hamed of Cairo University for clarifying several points in his paper, to the Vice-Chancellor Professor Datuk Dr. Ghauth Jasmon and the Dean of the Faculty of Engineering Professor Dr. Mohd Hamdi Abdul Shukor of the University of Malaya for their support. The Department of Irrigation and Drainage, Ministry of Natural Resources and

Environment, Malaysia, provided the data used in this study. Their help is acknowledged. We would also like to thank the University of Malaya for partially supporting this work.

References

Burn, D.H. (1994). Hydrologic effects of climatic change in West Central Canada. *J. Hydrol.*, 160, pp. 53–70.

Chiew, F.H.S., McMahon, T.A. (1993). Detection of trend or change in annual flow of Australian rivers. *Int. J. Climatol.*, 13, pp. 643–653.

Cox, D.R., Stuart, A. (1955). Some quick sign tests for trend in location and dispersion. *Biometrika*, 42, pp. 80–95.

Hamed, K.H. (2008). Trend detection in hydrologic data: the Mann-Kendall trend test under the scaling hypothesis. *J. Hydrol.*, 349, pp 350–363.

Hamed, K.H., Rao, A.R. (1998). A modified Mann-Kendall trend test for autocorrelated data. *J. Hydrol.*, 204, pp. 182–196.

Hansen, J., Ruedy, R., Sato, M., Lo, K. (2010). *Global surface temperature change*. Draft only (dated 1 June 2010). NASA Goddard Institute for Space Studies, New York, New York, USA. Available at: http://data.giss.nasa.gov/gistemp/ [Accessed 1 July 2010].

Hirsch, R.M., Slack, J.R. (1984). Non-parametric trend test for seasonal data with serial dependence. *Water Resour. Res.*, 20(6), pp. 727–732.

Kendall, M.G. (1975). *Rank correlation methods*. Griffin, London.

Kumar, S., Merwade, V., Kam, J., Thurner, K. (2009). Streamflow trends in Indiana: effects of long term persistence, precipitation and subsurface drains. *J. Hydrol.*, 374, pp. 171–183.

Lettenmaier, D.P. (1976). Detecting trends in water quality data from records with dependent observations. *Water Resour. Res.*, 12(5), pp. 1037–1046.

Mann, H.B. (1945). Nonparametric tests against trend. *Econometrica*, 13, pp. 245–259.

McLeod, A.I., Hipel, K.W. (1978). Preservation of the rescaled range: 1. a reassessment of the Hurst phenomenon. *Water Resour. Res.*, 14(3), pp. 491–508.

Mitosek, H.T. (1992). Occurrence of climate variability and change within the hydrologic time series: a statistical approach. CP-92-05, *International Institute for Applied Systems Analysis*, Laxenburg, Austria.

Sen, P.K. (1968). Estimates of the regression coefficient based on Kendall's tau, *J. Amer. Statistical Assoc.*, 63, pp. 1379–1389.

Wang, B., Ding, Q., Jhun, J.G. (2006). Trends in Seoul (1778–2004) summer precipitation. *Geophys. Res. Lett.*, 33, L15803

World Meteorological Organization. (1988). *Analyzing long time series of hydrological data with respect to climate variability*. WCAP-3 (WMO/TD-No.224).

Xu, Z.X., Takeuchi, K., Ishidaira, H. (2003). Monotonic trend and step changes in Japanese precipitation. *J. Hydrol.*, 279, pp. 144–150.

Zhang, Q., Liu, C., Xu, C.Y., Xu, Y., Jiang, T. (2006). Observed trends of annual maximum water level and stream flow during past 130 years in the Yangtze River basin, China. *J. Hydrol.*, 324, pp. 255–265.

Figures

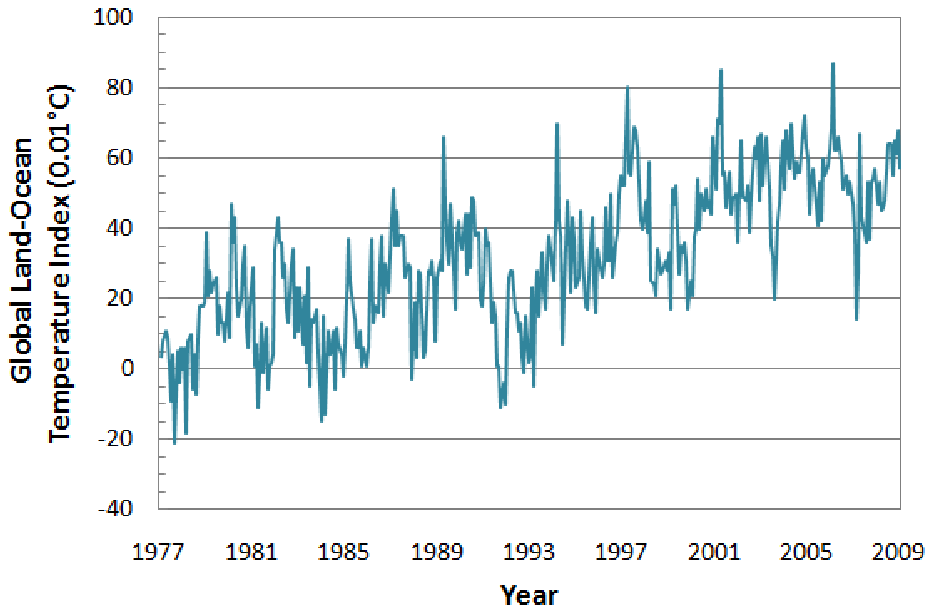

Fig. 1. *Global land-ocean temperature index for 1977–2009 (Hansen et al., 2010)*

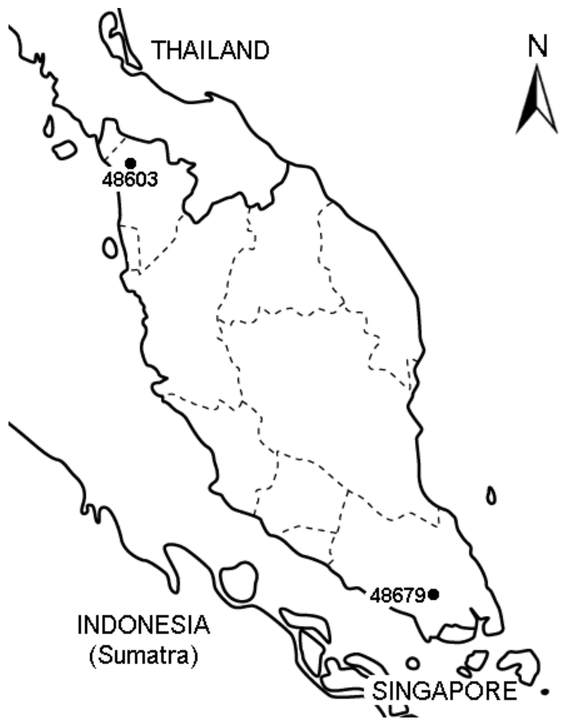

Fig. 2. *Locations of two meteorological stations in Peninsular Malaysia: Alor Setar in Kedah (48603) and Senai in Johor (48679)*