ISSN NO. 2320-5407

Journal Homepage: -www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI:10.21474/IJAR01/ 9332
DOI URL: http://dx.doi.org/10.21474/IJAR01/9332

RESEARCH ARTICLE

THE INFLUENCE OF LAG-1 AUTOCORRELATION ON THE MK TEST AND APPLICATION IN THE CHANGE-POINT ANALYSIS OF FLOOD DATA.

Jian Wang and Qiuli Dong.
School of Mathematics and Statistics, Shandong University of Technology.

Manuscript Info

Manuscript History
Received: 03 May 2019
Final Accepted: 05 June 2019
Published: July 2019

Key words:-
Mann-Kendall test; Monte Carlo simulation; AR(1) process; Change-point analysis.

Abstract

In this paper, the effects of the AR(1) process with the different lag-1 serial correlation coefficients and different sample sizes on the Mann-Kendall test were analyzed by Monte Carlo simulation. Then, using pre-whitening method, we modify the primary data and eliminate the influence of lag-1 serial correlation on the MK Mutation Test. Finally, based on the flood data from 1949 to 2008 in Zibo, China, the change point and the trend of the flood data were analyzed and the results between the revised and the original method were compared and discussed.

Introduction:-

The Change-point Analysis of time series has important applications in many fields of life, such as studies on Change Point Analysis of hydrometeorological time series. The global climate change threaten the living environment of human beings, so the research on the hydrometeorological series is becoming more and more important, and the Change Point Analysis has become a focus of many scholars. The nonparametric statistical method requires less assumptions, and its operation process is relatively simple, which makes it easier for researchers to apply and understand. Therefore, it is of great significance to study the nonparametric statistical test method. At present, there are many nonparametric statistical methods for Change Point Analysis, such as Mann-Kendall method, pettitt method, Reverse Spearman's rho test, WMW method and so on. The majority of the previous studies have assumed that sample data are serially independent.[1–3] It is known that some hydrometeorological time series may show serial correlation. The MK test is one of the most popularly applied tests for change-points of hydrometeorological time series, and lag-1 serial correlation has a great influence on the MK test. Therefore, this paper will investigates this issue by Monte Carlo simulation and find a way to eliminate the influence of the lag-one serial correlation on the Mann-Kendall test in mutation test of hydrometeorological time series.

Mann-Kendall Mutation Test

Theorem1.1 Let $\{X_t\}$ be a time series which observations are denoted by $\{x_t\}$, where $t = 1, 2, \cdots, k, \cdots, n$

$$S_k = \sum_{i=1}^{k} r_i \qquad (k = 2, 3, \cdots, n)$$

Where

Corresponding Author:-Jian Wang.
Address:-School of Mathematics and Statistics, Shandong University of Technology.

$$r_i = \begin{cases} 1 & if \ x_i > x_j, \\ 0 & otherwise. \end{cases} \quad (j = 1, 2, \cdots, i)$$

Let us assume that the time series above is serially independent. Then, the statistic, denoted UF_k, is defined as

$$UF_k = \frac{[S_k - E(S_k)]}{\sqrt{Var(S_k)}} \quad (k=1,2,\cdots,n) \qquad (1)$$

Where $UF_1 = 0$.

If $X_1, X_2, \cdots, X_n$ are independent random variables of the continuous type, each having the same p.d.f.. Then we have

$$E(S_k) = \frac{k(k-1)}{4}$$

$$Var(S_k) = \frac{n(n-1)(2n+5)}{72}$$

Let $\{X_t^*\}$ be a time series which observations is $\{x_t^*, t = 1, 2, \cdots, n\}$, where $X_t^* = X_{n+1-t}$. Then we have

$$S_k^* = \sum_{i=1}^{k} r_i^* \quad (k = 2, 3, \cdots, n)$$

Where

$$r_i^* = \begin{cases} 1 & if \ x_i^* > x_j^*, \\ 0 & otherwise. \end{cases} \quad (j = 1, 2, \cdots, i)$$

Assume that all the conditions of the statistic UF_k hold. We still have

$$UB_k = -\frac{[S_k^* - E(S_k^*)]}{\sqrt{Var(S_k^*)}} \quad (k = 1, 2, \cdots, n) \qquad (2)$$

Let α be the significance level. If the observed value of UB_k is equal to that of UF_k and the value is within the $(1-\alpha)\%$ confidence interval, the change point of the time series will be k.

The Influence of the Lag-1 Serial Correlation on the MK Test

Simulation Study[4–5]

The study investigate this issue by Monte Carlo simulation.

To indicate the influence of the lag-1 serial correlation on the MK test, given lag-1 serial correlation coefficient $\rho_1 = (-0.9(0.1)0.9)$, AR(1) processes with mean $\mu_x = 1$ and variance $\sigma_x^2 = 0.5$ for each sample size $(n = 20, 50, 100, 150)$ are generated by

$$X_t = \mu_x + \rho_1 (X_{t-1} - \mu_x) + \sigma_x \sqrt{1-\rho_1^2 \varepsilon_t} \qquad (3)$$

Where ε_t is the white-noise process with mean $\mu_\varepsilon = 0$ and variance $\sigma_\varepsilon^2 = \sigma_x^2 (1-\rho_1^2)$. Given that ε_t is a normally distributed random variable with mean $\mu_\varepsilon = 0$ and variance $\sigma_\varepsilon^2 = 1$. The Monte Carlo simulation generated 3000 time series of AR(1) processes using Equation(3) for each sample size above with the given ρ_1. Then, using Monte Carlo simulation, this study investigates the influence of lag-1 serial correlation and sample size on the MK Mutation Test, respectively.

Simulation Results:-

Taking the probability of having more than six change-points as the research object, we analyzed the effects of the AR(1) process with the different lag-1 serial correlation coefficients and different sample sizes on it by using R statistic software for programming. And the change-points are obtained by the MK Test method in Theorem1.1. Here the significance level $\alpha = 0.05$ and the critical value $\mu_{0.05} = \pm 1.96$. Then, the simulation results were shown in Fig.1 and Fig.2, respectively.

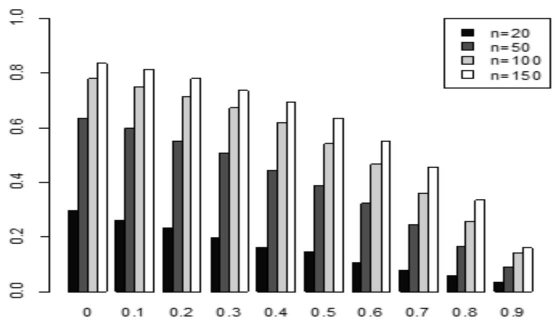


Fig 1:-Effect of positive serial correlation

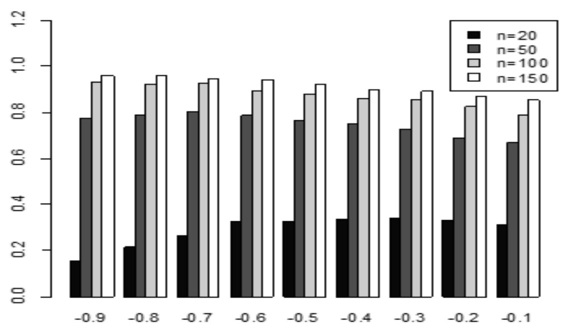


Fig 2:-Effect of negative serial correlation

The probability of having more than six change-points corresponding to given ρ_1 and the sample size n are shown in Figure 1 and Figure 2, which indicates that the higher the absolute value of positive lag-1 serial correlation, the lower the probability of having more than six change-points. And the influence of negative serial correlation on having more than six change-points is opposite to the positive cases. It also can be seen that the larger the sample size, the higher the probability of having more than six change-points. Hence, the lag-1 serial correlation has great effect on the Mann-Kendall Mutation Test. Eliminating the lag-1 serial correlation in a time series can produce more accurate results.

Pre-whitening

Pre-whitening [5–6] has been suggested by von Storch (1995) and has been used by Kulkarni and von Storch (1995), Douglas et al. (2000), Zhang et al. (2000,2001), and Burn and Hag Elnur (2001) to reduce the influence of an AR(1) component on the application of the MK Test. The series can be pre-whitened using the following formula:

$$Y_t = X_t - \rho_1 X_{t-1} \qquad (4)$$

Obviously, $\{Y_t\}$ is a time series which observation is the $\{y_t\}$. Then, using the Monte Carlo simulation in column 2, we show the probability of having more than six change-points corresponding to given ρ_1 and the sample size n in Figure 3.

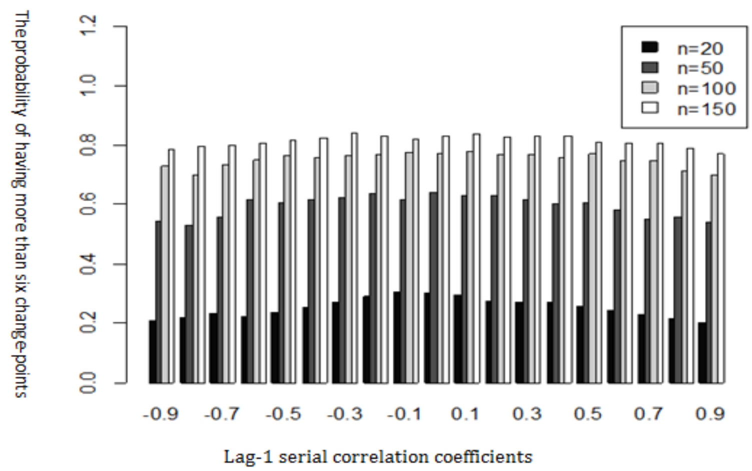


Fig 3:-Modified by Pre-whitening

Figure 3 indicates that the pre-whitened series no longer displays the existence of lag-1 serial correlation in a time series. Therefore, Pre-whitening method can eliminate the influence of lag-1 serial correlation on the MK Mutation Test.

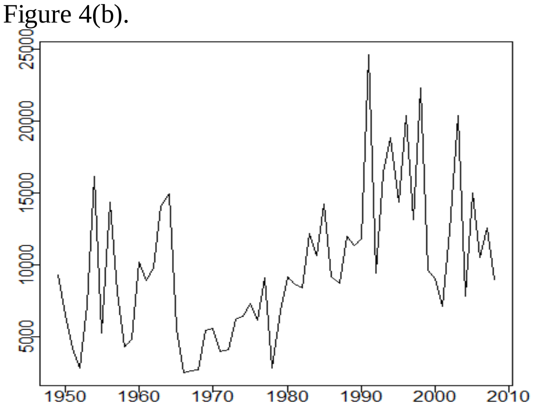
Case Study

To demonstrate pre-whitening method can eliminate the influence of lag-1 serial correlation on the MK Mutation Test, the flood data from 1949 to 2008 in Zibo, China were made use of. First, we need to judge whether the lag-1 serial correlation exists in the flood time series.

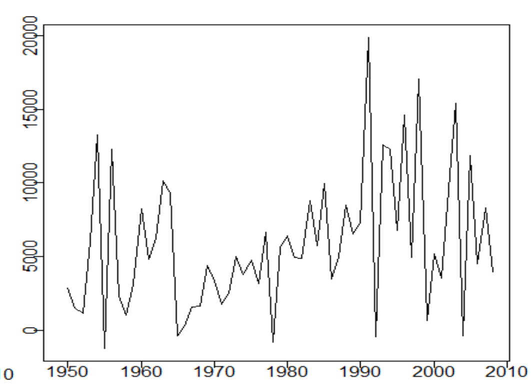
Tab 1:-The test of significance of lag-1 serial correlation coefficient

parameter	estimated value	Std. Error	t-Statistic	Pr > |t|	lag
MU	9763.1	994.85509	9.81	<.0001	0
AR1,1	0.40125	0.12030	3.34	0.0015	1

The existence of lag-1 serial correlation in the flood time series, obtained by conditional least squares estimator, is displayed in Table 1. Given that the lag-1 serial correlation is denoted by ρ and p denote the p-value in this issue. From Table 1, it can be seen that $\rho = 0.40125$ and $p < 0.01$. Therefore, lag-1 serial correlation exists in the flood time series. Then, to eliminate the influence of lag-1 serial correlation on the MK Mutation Test, we modified the flood data by pre-whitening. The line charts of flood data before and after pre-whitening are displayed in Figure 4(a) and Figure 4(b).



(a)

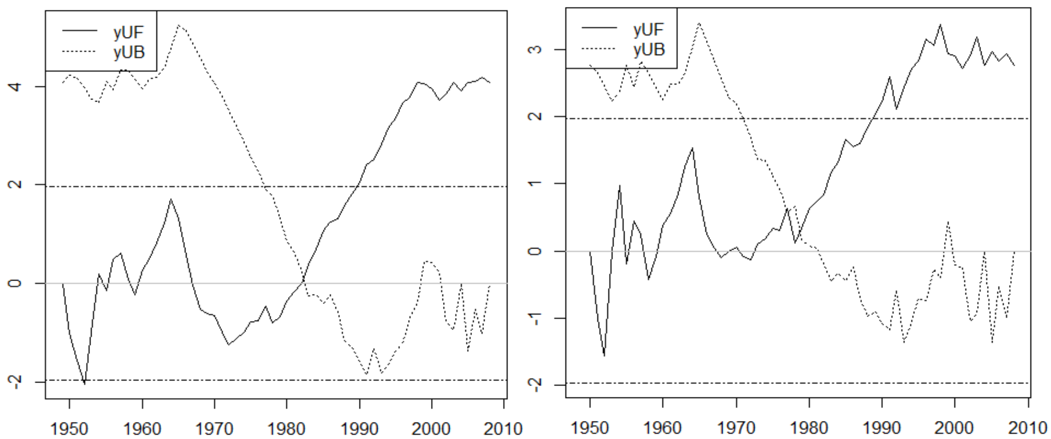


(b)

Fig 4:-Line chart of flood data before and after pre-whitening from 1949 to 2008 in Zibo, China.

By viewing Figure 4, it is evident that the fluctuation of the flood data modified by pre-whitening is the same as primary data. It indicates that the main information in the flood data is not lost.

Next the MK Mutation Test ascertains the change-points in the pre-whitened flood series and primary flood data, respectively. The results are shown in Figure 5.



(a) (b)

Fig 5:-Line chart of the MK test based on the flood data before and after pre-whitening from 1949 to 2008 in Zibo, China.

By comparing the change-points in Figure 5(a) with those in Figure 5(b), it can be seen that there is a change-point in the primary data, but there are two change-points in the pre-whitened series. It indicates that lag-1 serial correlation has a great influence on the MK test and Pre-whitening can eliminate the influence. We have known that the absolute value of lag-1 serial correlation coefficients and sample size have different degrees of influence on the probability of having more than six change-points in column 2.2. So if the sample size n (resp. the lag-1 serial correlation ρ) is not equal to the sample size 59(resp. the lag-1 serial correlation 0.4), the influence on the MK test will be different.

From Figure 5, it also can be seen that there is a change-point before 1980. We guess: The change-point before 1980 may result in the change-point 1980. This phenomenon deserves in-depth study. We conclude that pre-whitening method can eliminate the effect of lag-1 serial correlation on the MK Mutation Test. Therefore, we can more accurately find the change-points and what the trend is in the pre-whitened series.

Concluding Remarks:-

This study investigated the effect of lag-1 serial correlation on the MK Mutation Test. And pre-whiting can eliminate the influence. The Monte Carlo simulation results show that the absolute value of lag-1 serial correlation coefficients have different degrees of influence on the MK Mutation Test. Therefore, we must judge whether the lag-1 serial correlation exists in time series before ascertaining the change-points by MK Test. If lag-1 serial correlation exists in time series, we will modify the time series using the pre-whiting method. Then, we can ascertaining the change-points by MK Mutation Test.

From the simulation study in column 2, it can be seen that if the absolute value of lag-1 serial correlation coefficient $|\rho_1| \geq 0.2$, the lag-1 serial correlation will have a great influence on the MK Mutation Test. However, we can ignore the effect when the sample size $n \geq 100$ and the absolute value of lag-1 serial correlation coefficient $|\rho_1| < 0.1$.

The MK Mutation Test method studied in this paper can be applied to other fields such as hydrometeorology and finance, and has important research and application value.

References:-

1. Zhenlong Wang, Xi Chen, Zhenchun Hao: 2010, 'Long-team Trend and Jump Change for Major Climate Processes over Area of Huabei', Journal of Irrigation and Drainage. 29(5):52-56.
2. Fengying Wei: 1999, 'Statistical Diagnosis and Prediction Technology of Modern Climatic', China Meteorological press, Beijing.
3. Yinghua Zhang, Xianfang Song: 2015, 'Techniques of abrupt change detection and trends analysis in hydroclimatic time-series: Advances and evaluation', ARID LAND GEOGRAPHY. 38(4): 652-665.
4. YUE S, WANG C Y: 2004, 'The Man-Kendall test modified by effective sample size to detect trend in serially Correlated Hydrological Series [J] ', Water Resour Manage. 18:201-218.
5. YUE S, PILON P, PHINNEY B. et al: 2002, 'The influence of autocorrelation on the ability to detect trend in hydrological series [J]', Water Resour Manage. 16: 1807-1829.
6. YUE S, WANG C Y: 2002/2001, 'The applicability of pre-whiting to eliminate the influence of serial correlation on the Mann-Kendall test [J]', Water Res. 38(6): 10.1029/ WR000861, 4-1-7.
7. Xinming Yao, Yuanfang Chen, Shenghua Gu: 2014, 'Application of EEMD-NNBR Model in Rainfall Forecasting', Water Resources and Power. 32(12):11-13, 16.