

상관성을 가진 시계열 자료의 경향성 분석에 관한 연구

Trend Detection of Serially Correlated Hydrologic Series

오 제 승* / 김 병 식** / 김 형 수*** / 서 병 하****

Oh, Je Seung / Kim, Byung Sik / Kim, Hung Soo / Seoh, Byung Ha

:: Abstract ::

The non-parametric Mann-Kendall(MK) statistical test has been widely used to assess the significance of trend in hydrologic time series. The test requires sample data should be serially independent. If sample data is serially correlated, the presence of serial correlation in a time series will affect the test ability for trend analysis. So, we would like to use the modified MK test which uses the effective sample size(ESS) to eliminate the effect of serial correlation in a series. This study investigates the ability of ESS to eliminate the influence of serial correlation of MK test by Monte Carlo simulation and by real series. As the results, MK test shows the increase of trend rate as the serial correlation is increased but the modified MK test shows ESS can eliminate the serial correlation for trend analysis. Therefore we confirmed the modified MK test is a very useful tool for trend analysis.

Keywords: Trend analysis, serial correlation, Mann-Kendall test, ESS

:: 요 지 ::

경향성을 분석하기 위한 여러 기법 가운데 비모수적 방법인 Mann-Kendall 검정(MK 검정)은 수문 시계열의 분석에서 널리 사용되어지고 있다. 이 검정 방법은 분석 대상 자료가 독립이라는 가정 하에 수행되며, 자료가 계열상관되어 있는 경우에는 그 상관성의 영향으로 경향성 존재의 유무를 정확하게 판단할 수 없게 된다. 따라서 MK 검정을 이용할 때 자료의 상관성으로 인해 받게 되는 영향을 소거시키기 위해 효과적인 자료의 크기(Effective Sample Size, ESS)를 이용하는 수정된 MK 검정 방법을 이용하고자 한다. 본 연구에서는 ESS가 얼마만큼 계열상관성을 제거할 수 있는지 Monte Carlo모의를 통해 검토하였다. MK 검정 결과 계열상관계수의 증가에 따라 자료가 경향성을 나타내는 비율이 높아 졌으나, ESS에 의해 수정된 MK 검정을 수행한 결과 계열상관계수의 영향이 제거되었음을 알 수 있었다. 또한 실측 자료의 적용을 통해 수정 MK검정이 경향성 분석에 매우 유용함을 확인 하였다.

핵심용어: 경향성 분석, 계열 상관성, Mann-Kendall 검정, 효과적인 자료 크기

* 인하대학교 토목공학과 석사과정 e-mail: truelove1801@hanmail.net

** 인하대학교 토목공학과 박사과정 e-mail: hydrokbs@kict.re.kr

*** 인하대학교 토목공학과 교수 e-mail: sookim@inha.ac.kr

**** 인하대학교 토목공학과 교수 e-mail: seohydro@inha.ac.kr

1. 서 론

수자원 계획 수립을 위하여 목적에 따라 추계학적 모형 구축이나 빈도분석등 수문시계열 해석을 시행한다. 이 때 수문시계열 자료는 정상성과 독립성을 지녀야 하고 우선적으로 이들에 대한 검정(test)을 수행하여야 한다. 그러나 국내의 경우 사전에 수문 시계열 자료가 정상성이나 독립성을 지니고 있다고 가정하고 수문 분석을 실시하는 경우가 많았기 때문에, 왜곡된 결과를 초래할 여지가 있었다고 사료된다. 그러므로, 수문자료를 확률적인 분석이나 추계학적 분석에 사용하기 위해서는 수문자료의 시계열 특성을 사전에 반드시 분석할 필요가 있다. 시계열 자료의 특성 중 하나인 경향성은 자료를 도시함으로써 쉽게 파악할 수 있지만, 경향성이 뚜렷하게 나타나지 않는 경우에는 통계적 방법을 이용하여 경향성 분석을 실시한다. 통계적 분석 방법으로는 T-검정, Hotelling-pabst 검정, 비선형 경향성 검정(nonlinear trend test), Mann-Kendall 검정(MK 검정), Sen 검정 등의 여러 가지 방법이 있다.

Yue and Hashino(2003)는 일본에서의 월, 년강우량에 대한 장기 경향성 분석에 대하여 연구하였으며, Yue등(2002)은 수문 시계열에서 자기 상관성이 경향성을 판단하는데 미치는 영향에 대하여, Yue and Wang(2002)은 자료의 계열 상관성이 Mann-Kendall 검정에 미치는 영향을 제거하기위한 방법에 대하여 연구한 바 있다. Hamed and Rao(1998)는 상관성이 존재하는 자료에 대한 수정된 Mann-Kendall 검정의 연구에서 효과적인 자료크기(Effective sample size, ESS)를 처음으로 제안 하였으며 국내에선 이상복(2004)등이 강수 자료에 대한 변동성 및 경향성 해석에 대하여

연구하였다.

본 연구에서는 수문 시계열 자료의 경향성분석 방법 중 하나인 Mann-Kendall 검정에서 자료의 자기 상관성이 검정에 미치는 영향을 파악하고, 자기 상관성의 영향을 소거시키기 위해 lag-1 계열상관계수에 의한 효과적 자료크기(Effective sample size, ESS)를 이용하여 수정된 Mann-Kendall 검정을 실시하였다. 즉, 수정 MK 검정이 계열 상관성의 영향 제거에 어떠한 효과를 나타내는지를 분석하였다.

2. Mann-Kendall 검정과 계열상관성

2.1 Mann-Kendall 검정

MK 검정 통계량 S를 정의하면 다음 식(1)과 같다.(Mann, 1945; Kendall, 1975)

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(X_j - X_i) \quad (1)$$

여기서, X_j : j번째 자료값

n : 자료의 개수

$$\text{sgn}(\theta) = \begin{cases} 1 & \text{if } \theta > 0 \\ 0 & \text{if } \theta = 0 \\ -1 & \text{if } \theta < 0 \end{cases} \quad (2)$$

그리고, 통계량 S의 분산은 다음 식(3)과 같다.

$$V(S) = \frac{n(n-1)(2n+5)}{18} \quad (3)$$

MK 검정의 표준 검정 통계량과 단측 검정을 위한 P-value등은 식(4) 및 식(5)와 같이 나타내어진다.

$$Z = \begin{cases} \frac{S-1}{\sqrt{Var(S)}} & S > 0 \\ 0 & S = 0 \\ \frac{S+1}{\sqrt{Var(S)}} & S < 0 \end{cases} \quad (4)$$

$$p = 0.5 - \Phi(|Z|)$$

$$\left(\Phi(|Z|) = \frac{1}{\sqrt{2\pi}} \int_0^{|Z|} e^{-\frac{t^2}{2}} dt \right) \quad (5)$$

위의 식에서 Z 값의 부호는 경향성의 증가 또는 감소 특성을 나타낸다. 유의수준이 0.05일 경우 $p \leq 0.05$ 이면 경향성의 존재는 통계적으로 의미 있는 값으로 고려되어진다.

2.2 MK 검정에 대한 계열 상관의 영향

Storch(1995) 와 Yue et al.(2002)는 경향성이 존재하지 않는 시계열 자료가 양의 계열상관성을 가지고 있을 경우, 자료가 비경향성 이라는 MK 검정의 귀무가설을 거부할 확률이 증가하며, 음의 계열 상관성을 가지고 있는 경우에는 귀무가설을 거부할 확률이 감소한다는 것을 연구하였다.

이러한 내용을 분석하기 위해 식(6)과 같은 AR(1)모형을 이용하는데 lag-1 계열상관계수 ρ_1 을 -0.9부터 0.1씩 증가시켜 +0.9까지의 값을 가지도록 하고, 평균 $\mu_A = 1.0$, 분산 $\sigma_A^2 = 1.0$, 그리고 표본은 10~80(중분 10)의 크기를 갖도록 하였다. 그리고, 자료의 크기별로 각각 10,000개씩의 자료군을 발생시켰다(Yue and Wang, 2004).

$$A_t = \mu_A + \rho_1(A_{t-1} - \mu_A) + \epsilon_t \quad (6)$$

여기서, ϵ_t 는 $\mu_\epsilon = 0$, $\sigma_\epsilon^2 = \sigma_A^2(1 - \rho_1^2)$ 의 값을

가지는 백색잡음이다.

따라서, 모의 발생된 자료에 대한 MK 통계량 S_i 는 각 자료크기와 상관계수 별로 10,000개씩 식 (1)에 의해 모의되어지고 MK 검정 결과 귀무가설을 기각할 확률은 다음의 식(7)로 계산된다.

$$R_{rej} = N_{rej} / N \quad (7)$$

여기서, N : 모의 발생된 자료군의 수(=10,000)

N_{rej} : 10,000개의 자료군중에 S_i 가 기각역에 위치하는 것의 개수

또한, MK 통계량에 대한 귀무가설의 기각역은 다음 식과 같으며, S값이 아래의 식(8)과 같은 경우 경향성이 있다고 판단하게 된다.

$$|S| > z_{1-\alpha/2} \sqrt{V(S)} \quad (8)$$

여기서, α : 유의수준

$z_{1-\alpha/2}$: 표준정규분포의 $1 - \alpha/2$ 분위수

위와 같이 모의발생된 자료를 이용한 MK 검정은 계열상관계수의 증가에 따라 귀무가설의 기각률이 지수함수적으로 증가하는 결과를 나타내었다. 이는 실제 자료가 비경향성임에도 불구하고 상관계수가 클수록 경향성으로 판단되어지는 횟수가 증가하는 것이며, 자료의 계열상관성이 MK 검정에 어떠한 영향을 미친다는 것을 알 수 있다. 또한 모의 발생된 자료를 통한 검정결과 자료의 크기에 따라서는 별다른 영향이 없는 것을 알 수 있었다.

상관성을 지닌 자료가 MK 검정에 미치는 영향은 MK 통계량에 대한 정규분포 적합 곡선을 통해 설명될 수 있는데, 자료의 크기가 일정한

경우, MK 통계량의 확률분포 곡선을 도시하면 상관계수의 값이 클수록 분산이 커지고 편평해지는 형태를 나타낸다. 즉, 상관계수가 큰 자료일수록 MK 통계량의 더 많은 값들이 기각역에 위치하게 되어 비경향성이라는 귀무가설을 기각하므로 경향성으로 판단하게 되는 것이다.

3. MK 검정에 대한 계열상관성의 영향 제거 방법

앞 절에서 시계열에 존재하는 계열상관성에 따라 MK 통계량의 분산이 변하게 됨을 설명하였다. 본 절에서는 분산을 수정하여 계열상관성의 영향을 제거할 수 있는 ESS방법과 그 효과를 기술하고자 하며, ESS에 의해 수정된 분산은 다음 식(9)와 같이 나타낼 수 있다.

$$V^*(S) = V(S) \frac{n}{n^*} \quad (9)$$

여기서, n : ASS(actual sample size)

n^* : ESS

n/n^* : 보정계수

3.1 ESS의 계산

BHMLL-ESS는 다음의 식으로 계산된다 (Bayley and Hammersley, 1946).

$$n^* = \frac{n}{1 + 2 \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \rho_k} \quad (10)$$

위의 식에서 ρ_k 는 lag-k인 계열상관계수이며 다음과 같이 계산된다(Salas et al., 1980).

$$r_k = \frac{\frac{1}{n-k} \sum_{t=1}^{n-k} (X_t - \bar{X}_t)(X_{t+k} - \bar{X}_t)}{\frac{1}{n} \sum_{t=1}^n (X_t - \bar{X}_t)^2} \quad (11)$$

$$\bar{X}_t = \frac{1}{n} \sum_{t=1}^n X_t \quad (12)$$

또한, 시계열 자료가 lag-1 AR 과정을 나타내는 경우 ESS는 다음의 식으로 계산된다 (Matalas and Langbein, 1962).

$$n^* = \frac{n}{1 + 2 \frac{\rho_1^{n+1} - n\rho_1^2 + (n-1)\rho_1}{n(\rho_1 - 1)^2}} \quad (13)$$

3.2 모의된 자료에 대한 ESS의 효과

Yue and Wang(2004)은 추계학적 모형을 이용해 모의한 자료를 가지고 ESS 방법의 효과를 검토하였다. 즉, 모의된 각각의 lag-1 계열상관계수와 자료 크기에 대하여 식 (13)을 사용하여 ESS를 계산하였으며 수정된 분산 $V^*(S)$ 는 식 (9)에 의해 계산되어진다. 이 값을 식 (8)에서의 $V(S)$ 값 대신 적용하여 새로운 기각역을 계산하고 MK 검정을 다시 수행한다. 이에 따라 앞절에서 모의된 자료에 대한 귀무가설이 맞을 때 귀무가설을 기각하는 오류(Type1 error-제1종 오류)를 재계산 하여 계열상관계수에 대한 기각률을 다시 계산한다.

그 결과 자료의 크기나 계열상관계수의 증가에 관계없이 매우 작은 기각률을 나타내었으며 이는 ESS의 적용을 통한 MK 검정의 결과가 자료의 계열상관성에 영향을 받지 않는다는 것을 의미한다. 이로부터 ESS 방법은 상관성을 나타내는 시계열 자료의 경향성 분석에 있어서 계열상관성의 영향을 제거하는데에 효과적인 방법임을 알 수 있다.

4. 수정 MK 검정을 이용한 지점 강우량 시계열의 경향성 분석

이상의 내용을 대관령지점과 금강유역에 속해 있는 29개 지점의 30년 동안의 연 강우량에 적용하여 경향성 분석을 실시하였으며 분석 과정은 우선, 원자료에 대한 MK 검정을 실시한 후 상관성의 영향을 알아보기 위하여 ESS에 의해 수정된 MK 검정을 실시하였다. 시계열의 경향성 정도를 나타내기 위하여 Theil (1950)과 Sen(1968)이 제안한 다음 식(14)를 사용하였다.

$$\beta = \text{Median}\left(\frac{X_j - X_l}{j - l}\right) \quad \forall 1 < j \quad (14)$$

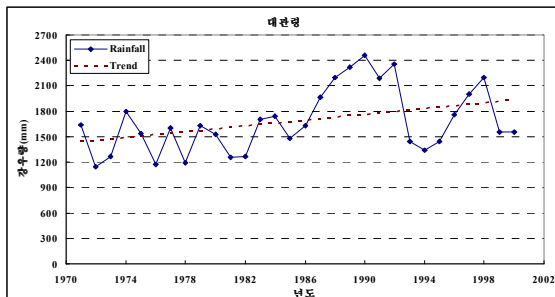
여기서, β : 경향성을 나타내는 추정값

X_l : l 번째 관측치

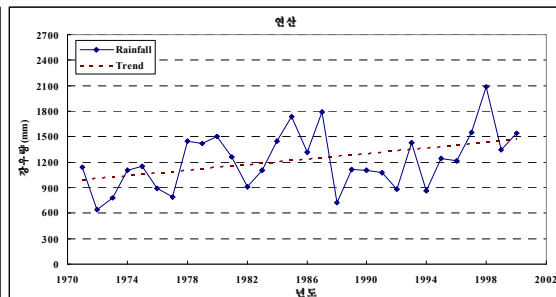
이를 이용하여 각 자료에 대한 경향성의 정도를 도시하였으며, 사용된 30개 지점의 연 강우자료중 MK 검정 결과 경향성이 존재하는 것으로 판단된 6개 지점의 원자료와 경향성, 경향성 제거 후의 자료값을 그림 1에 나타내었고, 그림 2에는 이 지점들의 lag별 계열상관계수의 값을 도시하였다. 최종 분석결과 경향성을 보이지 않는 대관령과 연산 지점의 경우는 원자료만을 도

시하였다.

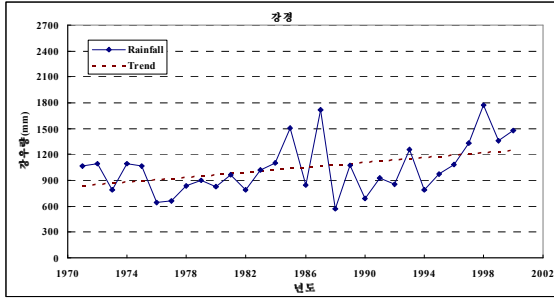
분석 결과는 표 1에 나타내었으며 첫 번째 열인 기울기는 식 (15)로 계산된 경향성의 정도를 나타낸다. MK 검정에서 주어진 지점의 자료에 대한 MK 통계량인 S를 먼저 계산하고 그에 해당하는 V(S)를 식 (1)과 식 (3)으로 계산한 후 ①열의 z는 식 (4)로부터, ②열의 p는 식 (5)로부터 계산한다. 이렇게 계산된 결과로부터 식 (8)을 만족하게 되면 경향성(T)을 나타내는 것으로 만족하지 않으면 비경향성(N)인 것으로 판단된다. ③열에는 이렇게 판단된 결과를 나타내었으며 30개 중 6개 지점의 연 강우자료에서 경향성을 보이는 것으로 판단되었다. 그러나, 자료에 상관성이 존재하는 경우에는 실제 비경향성인 자료임에도 귀무가설을 기각함으로써 경향성이라고 판단하는 제1종 오류가 발생할 수 있게 된다. 따라서, ESS에 의해 수정된 MK 검정을 다시 수행하였다. 그 결과는 ④~⑧열에 나타내었으며 ④열은 원자료가 가지는 lag-1 계열상관계수, ⑤열은 ④열의 값으로부터 계산되는 보정계수, ⑥~⑧열은 앞에서 설명한바와 같다. 이렇게 다시 분석한 결과 6개 지점 중 예산과 정안 지점만이 경향성을 보인다고 판단되었으며, 나머지는 경향성이 없는 것으로 나타났다. 상관성의 영향을 제거시킨 분석에서도 경향성을 보인 두 지점의 자료는 실제 경향성이 있는 것으로 판단된다.



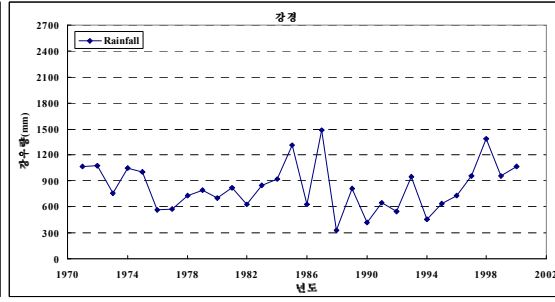
(a) 연강우 자료와 경향성(대관령)



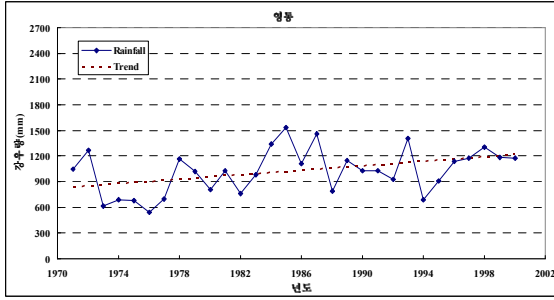
(b) 연강우 자료와 경향성(연산)



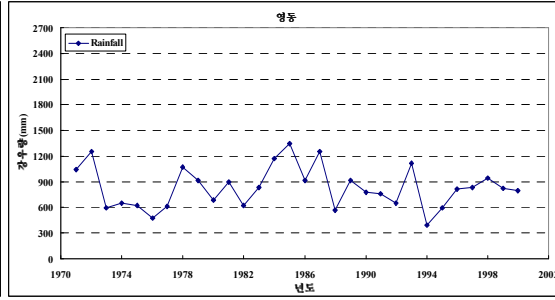
(c) 연강우 자료와 경향성(강경)



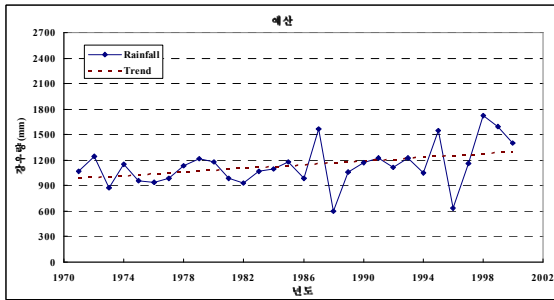
(d) 경향성 제거 후 자료(강경)



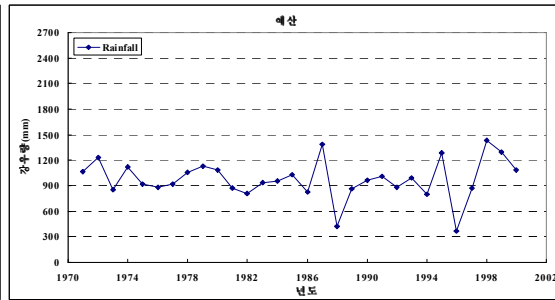
(e) 연강우 자료와 경향성(영동)



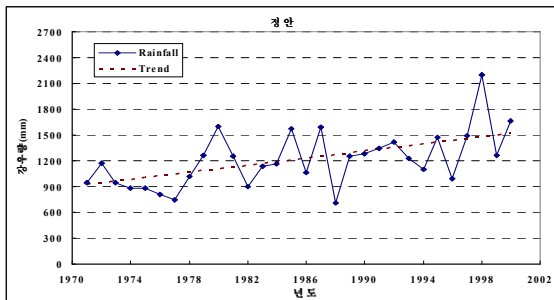
(f) 경향성 제거 후 자료(영동)



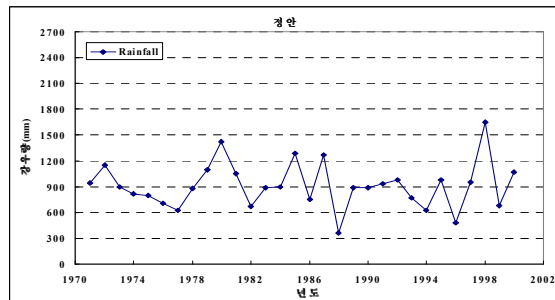
(g) 연강우 자료와 경향성(예산)



(h) 경향성 제거 후 자료(예산)



(i) 연강우 자료와 경향성(정안)



(j) 경향성 제거 후 자료(정안)

그림 1. 지점별 연 강우량과 경향성, 경향성 제거 후 자료

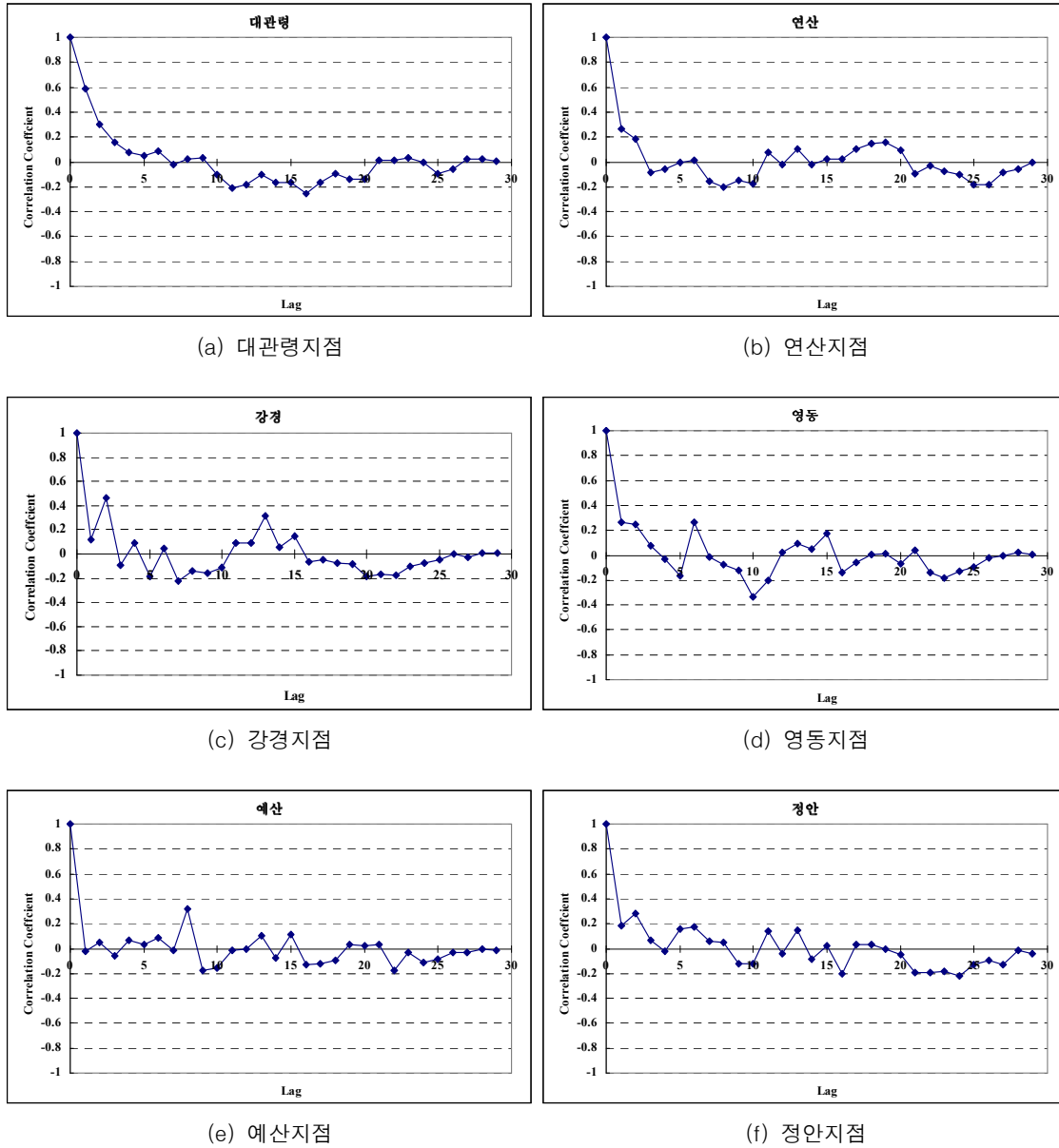


그림 2. 지점별 자기 상관계수

표 1. 30개 지점의 Mann-Kendall 검정 적용결과

관측소	기울기 (m ³ /s/yr)	MK 검정			수정 MK 검정				
		z	p	경향	r ₁	n/n*	z	p	경향
		①	②	③	④	⑤	⑥	⑦	⑧
강경	14.203	2.034	0.021	T	0.123	1.269	1.806	0.035	N
공주	9.578	1.249	0.106	N	0.109	1.235	1.124	0.131	N
군산	2.143	0.25	0.401	N	-0.022	0.958	0.255	0.399	N
금산	1.692	0.143	0.443	N	0.067	1.138	0.134	0.447	N

표 22. 30개 지점의 Mann-Kendall 검정 적용결과(계속)

관측소	기울기 (m ³ /s/yr)	MK 검정			수정 MK 검정				
		z	p	경향	r ₁	n/n*	z	p	경향
		①	②	③	④	⑤	⑥	⑦	⑧
대관령	17.025	2.355	0.009	T	0.585	3.592	1.243	0.107	N
대전	6.522	0.714	0.238	N	0.074	1.155	0.664	0.253	N
모서	9.641	1.178	0.12	N	-0.031	0.941	1.214	0.112	N
무주	2.517	0.464	0.321	N	-0.095	0.832	0.509	0.305	N
반포	14.212	1.802	0.036	N	0.034	1.068	1.744	0.041	N
병천	3.854	0.642	0.26	N	0.042	1.085	0.617	0.269	N
부강	5.994	1.213	0.113	N	-0.079	0.857	1.31	0.095	N
부여	7.144	1.213	0.113	N	-0.015	0.972	1.23	0.109	N
부향	9.326	1.57	0.058	N	0.192	1.456	1.301	0.097	N
생곡	3.222	0.571	0.284	N	0.024	1.048	0.558	0.289	N
안내	17.658	1.927	0.027	N	0.063	1.13	1.813	0.035	N
안성	9.228	1.677	0.047	N	0.152	1.346	1.446	0.074	N
연산	16.363	1.963	0.025	T	0.266	1.693	1.508	0.066	N
영동	13.113	2.284	0.011	T	0.268	1.698	1.752	0.04	N
예산	10.794	2.141	0.016	T	-0.026	0.951	2.196	0.014	T
옥천	6.935	1.035	0.15	N	-0.031	0.941	1.067	0.143	N
임피	6.017	0.607	0.272	N	-0.082	0.853	0.657	0.256	N
장수	10.62	1.178	0.12	N	0.205	1.494	0.963	0.168	N
정산	5.931	1.106	0.134	N	0.059	1.121	1.045	0.148	N
정안	20.576	3.033	0.001	T	0.185	1.436	2.531	0.006	T
진안	-3.465	-0.535	0.296	N	0.359	2.064	-0.373	0.355	N
진천	8.366	1.57	0.058	N	0.049	1.1	1.497	0.067	N
청양	7.696	1.32	0.093	N	0.03	1.06	1.282	0.1	N
청주	5.782	1.356	0.088	N	-0.074	0.866	1.457	0.073	N
추풍령	2.845	0.66	0.255	N	-0.038	0.93	0.685	0.247	N
홍산	8.015	0.714	0.238	N	0.257	1.662	0.554	0.29	N

5. 결론

본 연구에서는 기존의 MK 검정과 ESS에 의해 수정된 MK 검정을 통하여 시계열의 계열상관성이 경향성 분석에 미치는 영향을 분석하고 이를 30개 지점의 30년 연강우량 자료에 적용하였다. 본 연구에 대한 내용과 결론을 요약하면 다음과 같다.

- (1) 기존의 MK 검정은 시계열 자료에 상관성이 존재하는 경우 이것의 영향으로 비경향성 자료를 경향성 있는 자료로 판단하는 오류를 가지고 있다.

- (2) ESS에 의해 수정된 MK 검정은 자료가 가지고 있는 상관성의 영향을 제거하여 경향성 분석시 상관성을 나타내는 자료에 대해 기존의 MK 검정뿐만이 아니라 수정된 MK 검정을 수행하여 MK 검정의 오류를 제거해야 한다.

이상의 연구를 통하여 MK 검정이 경향성 분석에 미치는 상관계수의 영향을 파악하고, 수정 MK 검정이 경향성 분석의 다른 대안이 될 수 있음을 보였다.

참고문헌

1. Yue, S. and Wang, C., The Mann-Kendall Test Modified by Effective Sample Size to Detect Trend in Serially Correlated Hydrological Series. *Water Resources Management*, Vol.18, pp. 201-218, 2004.
2. Hamed, K. H. and Rao, A. R., A Modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology*, Vol.204, pp. 182-196, 1998.
3. Storch, V. H., Misuses of statistical analysis in climate research. *Analysis of Climate Variability: Applications of Statistical Techniques*, pp. 11-26, 1995.
4. Yue, S. and Wang, C., The applicability of pre-whitening to eliminate the influence of serial correlation on the Mann-Kendall test. *Water Resources Research*, Vol.38, 4-1-7, 2002.
5. Yue, S., Pilon, P., Phinney, B. and Cavadias, G., The influence of auto-correlation on the ability to detect trend in hydrological series. *Hydrological Processes*, Vol.16, pp. 1807-1829, 2002.
6. Yue, S. and Hashino, Long term trends of annual and monthly precipitation in Japan. *Water Resources Association*, Vol.39 No.3, pp. 587-596, 2003.
7. 이상복, 김경덕, 허준행. 강수자료에 대한 변동성 및 경향성 해석. 한국수자원학회 학술 발표회 논문집, pp. 137-137, 2004.