

확률강우량의 정상성 판단: 1. 기존 방법의 적용 및 평가

On the Stationarity of Rainfall Quantiles:

1. Application and Evaluation of Conventional Methodologies

정 성 인* / 유 철 상** / 윤 용 남***

Jung, Sung In / Yoo, Chulsang / Yoon, Yong Nam

Abstract

This study evaluated the statistical stationarity of rainfall quantiles as well as the rainfall itself. The conventional methodologies like the Cox-Stuart test for trend and Dickey-Fuller test for a unit root used for testing the stationarity of a time series were applied and evaluated their application to the rainfall quantiles. As results, first, no obvious increasing or decreasing trend was found for the rainfall in Seoul, which was also found to be a stationary time series based on the Dickey-Fuller test. However, the Cox-Stuart test for the rainfall quantiles show some trends but not in consistent ways of increasing or decreasing. Also, the Dickey-Fuller test for a unit root shows that the rainfall quantiles are non-stationary. This result is mainly due to the difference between the rainfall data and rainfall quantiles. That is, the rainfall is a random variable without any trend or non-stationarity. On the other hand, the rainfall quantiles are estimated by considering all the data to result in high correlation between their consecutive estimates. That is, as the rainfall quantiles are estimated by adding a stationary rainfall data continuously, it becomes possible for their consecutive estimates to become highly correlated. Thus, it is natural for the rainfall quantiles to be decided non-stationary if considering the methodology used in this study.

key words : Stationarity, Rainfall, Rainfall Quantiles, Cox-Stuart, Dickey-Fuller

요 지

본 연구에서는 강우량 자체 및 확률강우량의 통계학적 정상성을 판단하였다. 시계열의 정상성을 판단하는 대표적인 방법인 Cox-Stuart의 추세검정과 Dickey-Fuller의 단위근 검정 방법이 적용되었으며, 특히 확률강우량의 정상성 평가상의 문제점을 평가하였다. 결과적으로, 먼저 서울지점 강우량 자료에 대한 분석에서는 강우량이 증가하거나 감소하는 추세가 없다는 판단을 할 수 있었고, 아울러 단위근 검정에서도 정상적인 시계열이라는 결론을 얻을 수 있었다. 그러나 Cox-Stuart 검정에 의하면 확률강우량이 전체적으로 어떤 상승 또는 하향의 추세가 있는지에 대해서 일관된 판단을 하기가 어려웠다. 그러나 Dickey-Fuller의 단위근 검정에서는 확률강우량이 비정상시계열임을 판단할 수 있었다. 이러한 결과는 근본적으로 강우량과 확률강우량의 차이에서 비롯된 것이다. 즉, 강우는 무작위 변량으로서 어떤 경향성이나 비정상성을 찾기 힘들다. 반대로 확률강우량은 계산시점까지 관측된 모든 자료를 고려하여 추정되므로 전-후의 값의 상관성이 매우 커지게 된다. 즉, 정상시계열인 강우자료가 연속적으로 추가되며 확률강우량이 추정되므로 전-후 값이 높은 상관성이 가능

* 동부엔지니어링 수자원부 (E-mail : sijung@dbeng.co.kr)

** 정희원 · 고려대학교 건축·사회환경공학과 교수

*** (주)삼안 고문(前 고려대학교 사회환경시스템공학과 교수)

하다, 따라서 확률강우량이 비정상 시계열로 판단되는 것은 본 연구에서 적용된 판단기법으로는 당연한 결과라 할 수 있다.

핵심용어 : 정상성, 강우량, 확률강우량, Cox-Stuart, Dickey-Fuller

1. 서론

최근 들어 우리나라뿐 아니라 전 세계 곳곳에서 가뭄, 홍수, 한파 등 이상 기상이 빈번히 발생하고 있다. 강우 패턴에도 변화를 보여, 특히 1990년대 이후 한 반도에서는 여름철의 대표적인 장마전선이 약화되고 장마가 끝난 후 대류성 폭풍우를 동반한 산발적이고 국지적인 소나기성 강우가 주를 이루고 있다(권태영 등, 1998 ; 서규우, 1998 ; 이상복 등, 2000 ; 차은정과 최영진, 2000 ; 차은정, 2001). 이로 인해 계획 홍수량을 초과하는 극치 강우량의 발생이 빈번하고, 따라서 기존 홍수 방어 시설물의 치수 안전도가 저하되고 있는 상황에 있다.

이런 현상의 발생 원인이 지구온난화로 인한 기후 변화라는 것에는 크게 이견이 없다. 이런 기후변화의 큰 원인은 17세기 산업 혁명 이후 나날이 가속되고 있는 화석 연료의 사용증가로 인해 대기 중 온실 기체의 농도가 급속히 증가되었기 때문이다. 온실 효과로 인한 지구평균기온의 상승은 지구 전체적인 수문순환에도 영향을 미쳐 강수 형태의 급격한 변화를 야기하고 있다고 알려져 있다(Jhones와 Wigley, 1990 ; Mansell, 1997). 이러한 변화는 국지적으로 더 클 수도 또한 작을 수도 있으나 명확한 원인의 규명은 아직 이루어지고 있지 않은 실정이다. 현재 수자원분야에서도 지구온난화와 관련하여 지구환경의 변화에 따른 강수량 및 증발산량의 변화와 같은 양적인 변화뿐만 아니라 홍수나 가뭄과 같은 극단기상의 빈도변화까지 다양하게 연구되고 있다(Nemec과 Schaake, 1982 ; Rind 등, 1989 ; McCabe와 Ayers, 1989 ; Wardlaw 등, 1996). 그러나 아직까지 물리적 이론에 근거한 명확한 설명은 충분하지 않은 상황이다(차은정, 2001).

그럼에도 불구하고 Trenberth(1998, 1999)의 다음과 같은 간단한 설명은 지구 온난화가 왜 집중호우를 증가시키는지에 대해 쉽게 이해할 수 있게 해준다. “지구온난화는 대기 중 온실기체 증가에 의한 지구 대기온도가 증가하는 현상을 말한다. 대기 중 증가하는 온실기체는 지구온난화를 유발한다. 기온의 증가는 대기가 물을 함유할 수 있는 능력이 증가하는 것보다 더불어 증발량도 증가함을 의미한다. 지구 온난화로

인해 증가된 증발량은 물 수지 균형을 이루기 위하여 강수량을 증가시킨다”. 최근의 연구들도 빈번한 최근의 홍수와 가뭄이 지구 온난화로 인한 것이며, 이런 추세가 앞으로도 계속 될 것이며, 또한 보다 극심한 홍수와 가뭄의 발생도 가능하다고 제시하고 있다(Lettenmaier와 Gan, 1990 ; Lettenmaier와 sheer, 1991 ; Nash와 Gleick, 1991 ; Ng와 Marsal다, 1992 ; McCabe와 Hay, 1995).

강우의 특성변화와 관련하여 다음 두 가지의 경우를 생각할 수 있다. 첫 번째는 강우가 계속적으로 증가하거나 감소하는 경우이고, 두 번째는 강우의 변동 정도가 변화하는 경우이다(최병선, 2001 ; 제갈돈과 류지성, 1998). 두 경우 모두 강우의 정상성에 반하는 경우이다. 또한, 강우의 특성이 변하면 수공구조물의 수문학적 설계기준인 확률강우량이 변하게 된다(안재현 등, 2000 ; 2002 ; Ahn 등, 2001 ; Thomas, 2001 ; Richard 등, 2002). 이상기후로 인하여 강우량이 증가하였다면 확률강우량도 변하게 될 것이고 그렇게 된다면 기존에 설계된 수공구조물의 치수 안전도는 저하되게 된다. 이런 이유로 강우 사상의 변화와 함께 확률강우량의 변화를 파악하는 것이 중요하다. 그러나 확률강우량의 분석과 강우량의 분석에는 근본적인 차이점이 있다. 무엇보다도 빈도해석에 의해 추정되는 확률강우량 시계열은 각 시점 및 이전의 모든 자료가 함께 고려되어 추정된다는 점을 들 수 있다. 즉, 나중 시점의 확률강우량은 이전 시점의 확률강우량과 대부분의 자료를 공유하고 있다. 따라서 확률강우량은 강우량에 비해 보다 평활화되어 있을 가능성이 크며, 또한 연속된 두 자료사이의 독립성도 만족시키기 어렵다.

본 연구에서는 강우량 및 확률강우량의 변화 특성을 알아보고 최근에 빈번하게 관측된 이상강우를 고려하는 경우에도 확률강우량이 정상성을 갖고 있다고 판단할 수 있는지에 대해 분석해 보고자 한다. 본 연구에서는 특히 시계열의 정상성을 판단하는 대표적인 방법인 Cox-Stuart의 추세검정과 Dickey-Fuller의 단위근 검정 방법(김명직과 장국현, 1998 ; 이상열, 2001)을 강우량 자료 및 확률강우량에 적용하여 그 결과 및 방법론 적용의 문제점 등을 평가해 보고자 한다.

2. 빈도해석

강우나 홍수 혹은 갈수의 발생빈도를 확률론적으로 예측하는 방법을 빈도해석(frequency analysis)라 한다(정중호와 윤용남, 2002 ; Rao와 Hamed, 2000). 빈도해석법은 어떤 수문사상이 발생하는 원인과 과정 등에 관해서는 전혀 상관하지 않고 오직 어떤 크기를 가진 사상이 발생할 확률을 결정한다는 점에서 확률론적 방법과 다르다. 빈도해석에서는 일반적으로 확률밀도함수(PDF: Probability Density Function) 또는 누가확률분포함수(CDF: Cumulative Probability Distribution Function)를 사용한다. 이는 모집단의 표본인 수문자료 계열을 적정 확률분포에 적합시켜서 모집단이 갖는 확률 정보를 함수의 형태로 간결하게 집약시켜 용이하게 나타내기 위함이다. 모집단의 성질을 간단한 확률분포형으로 완벽하게 나타내는 쉽지 않으나 이 보다 현저하게 우월한 대안을 찾기는 쉽지 않다(차영준, 1993).

확률분포형이 가정된 상황에서 매개변수의 추정에는 모멘트법(Method of Moments: MOM), 확률가중모멘트법(Method of Probability Weighted Moments: PWM), 최우도법(Method of Maximum Likelihood: MML)에 의해 수행될 수 있다(Hosking, 1989 ; Rao와 Hamed, 2000). 먼저, 모멘트법은 적용이 매우 간단한 매개변수 추정방법이며, 모집단의 모멘트인 확률밀도함수의 모멘트와 표본자료의 모멘트가 같다고 보고 확률분포의 매개변수를 추정하는 방법이다. 그러나 주어진 자료의 모멘트는 큰 값에 지나치게 민감하고, 반대로 대수치를 취하는 경우에는 작은 값에 지나치게 민감한 단점이 있다. 또한 표본의 크기가 작을 경우 변동계수와 왜곡도계수가 매우 크게 편(bias)되고 아울러 그 변동성이 커 적용이 곤란하다. 이에 반해 최우도법은 주어진 확률밀도함수의 우도함수를 관측 자료에 대해 최적화하여 매개변수를 결정하므로 안정적인 매개변수를 제공하는 장점이 있다. 그러나 우도함수의 유도과정이나 이의 최적화 과정이 복잡하다는 단점이 있다.

확률가중모멘트 법은 모멘트 법이나 최우도법의 대안으로 최근에 널리 쓰이는 매개변수 추정방법이며 모멘트법과 마찬가지로 모집단의 확률가중모멘트와 표본자료의 확률가중모멘트가 일치한다고 보고 매개변수를 추정하는 방법이다. 모멘트법과 확률가중모멘트법의 차이는 모멘트 법이 자료의 크기에 상관없이 일정한 가중치를 주는 반면 확률가중모멘트법은 자료를 크기순으로 나열하여 크기에 따라 가중치를 다르게 주

어 매개변수를 추정한다는 데 있다. 확률가중모멘트법은 자료수가 작아서 생기는 영향이 적으며 안정적이고 극대강우사상이나 이상치가 포함된 자료계열에도 왜곡특성을 크게 나타내지 않는 장점을 가지고 있다.

임의의 확률분포에 대한 적합도 검정은 그 확률분포의 상대도수함수(relative frequency function)와 누가도수함수(cumulative frequency function)의 이론값과 표본 값을 비교하여 수행하게 된다(Rao와 Hamed, 2000). 이에 대한 검정 방법으로는 전구간에 대한 적합도를 나타내는 χ^2 (Chi-Square) 검정과 각각의 소구간별 적합도 분석을 위한 Kolmogorov-Smirnov(K-S) 검정, Probability Plot Correlation Coefficient(PPCC) 검정, Robustness 검정 등이 있다. 특히, 최근에 와서 적정 확률분포형의 선정이나 매개변수 추정 방법의 선택 기준은 단순히 적합도 검정의 결과가 아니라, 분포형을 잘못 가정했거나 자료에 오류가 있음에도 불구하고 원래의 통계특성을 얼마나 잘 유지하는가(robustness)로 나타내는 것이 타당하다는 Robustness 검정에 관한 이론이 부각되고 있다(정중호와 윤용남, 2002).

일반적으로 많이 사용되는 χ^2 검정 및 K-S 검정을 이용한 적합도 검정 결과를 살펴보면 고려된 대부분의 확률분포형이 검정을 통과하게 되어 검정 후 채택 대상으로 남는 확률분포형이 너무 많은 문제점이 있다. 또한, χ^2 검정, K-S 검정 등은 기각 여부를만 판단할 수 있는 검정 방법이어서 계산된 통계량이 작다고 더 나은 확률분포형임을 의미하지도 않는다. 즉, 이 통계량 값을 이용하여 최적 분포형을 채택하는 것은 어렵다. 특히, 적합도 검정을 통과한 확률분포형 중 어떤 것을 채택하느냐에 따라 확률강우량 등의 결과에 많은 차이를 나타내고 있기 때문에 채택에 있어서 신중을 요한다. 그러나 별다른 대안이 없기 때문에 설계자 주관에 의한 임의 채택, 가장 큰 값을 주는 것을 채택, 또는 '한국 확률강우량도 작성(건설교통부, 1988 ; 2000)'에서 채택된 Gumbel 분포(Type-I 극치분포)를 채택하는 것이 일반적인 설계관행이다.

본 연구에서도 서울지점의 강우량을 대상으로 하며, 아울러 가장 적합하다고 판단되는 Gumbel 분포를 빈도해석에 사용하였다. Gumbel 분포의 확률밀도함수는 다음과 같이 표현된다.

$$f(x) = \frac{1}{\hat{\alpha}} \exp \left[-\frac{x - \hat{x}_c}{\hat{\alpha}} - \exp \left(-\frac{x - \hat{x}_c}{\hat{\alpha}} \right) \right], \quad (1)$$

$$-\infty < x < \infty$$

여기서 $\hat{\alpha} (> 0)$ 는 척척(scale)변수이며 \hat{x}_0 는 최대확률의 발생위치를 표시한다. 모멘트법을 이용하는 경우 Gumbel 분포함수의 매개변수는 다음과 같이 표현된다.

$$s = \hat{\sigma} = m_2^{1/2} = \frac{\pi}{\sqrt{6}} \hat{\alpha} \quad (2)$$

$$\hat{\alpha} = 0.7797 s \quad (3)$$

$$\hat{x}_0 = \bar{x} - 0.5772157 \hat{\alpha} = \bar{x} - 0.45005 s \quad (4)$$

여기서 \bar{x} , s 는 대상 자료계열(x)의 평균, 표준편차이며, 0.5772157은 Euler 상수(ε)이다. 아울러 확률가중모멘트법을 이용하는 경우의 매개변수는 다음과 같다.

$$\hat{l}_1 = \hat{b}_0 = \bar{x} \quad (5)$$

$$\hat{b}_1 = \widehat{M}_{1,1,0} = \frac{1}{N} \sum_{i=1}^N F_i x_i \quad (6)$$

$$\hat{l}_2 = 2\hat{b}_1 - \hat{b}_0 \quad (7)$$

$$\hat{\alpha} = \frac{\hat{l}_2}{\ln 2} \quad (8)$$

$$\hat{x}_0 = \hat{l}_1 - 0.5772157 \hat{\alpha} \quad (9)$$

여기서 \hat{l}_1 , \hat{l}_2 는 대상 자료계열(x)의 1차, 2차 L-모멘트이며, \hat{b}_1 , \hat{b}_0 는 표본 자료의 확률가중 모멘트이다.

3. 강수량 및 확률강우량의 정상성 판단

3.1 대상자료

본 연구에서는 서울시 1908~2002년(자료기간: 95년)까지의 서울지점 일과 월 강수량 자료를 이용하여 일 최대, 주 최대, 월 최대와 연 강수량자료를 구축하였다. 그 결과는 그림 1과 같다. 한편 1950~1953년은 일 강우 관측 자료의 결측으로 인해 일 최대와 주 최대 자료에서 제외되었다.

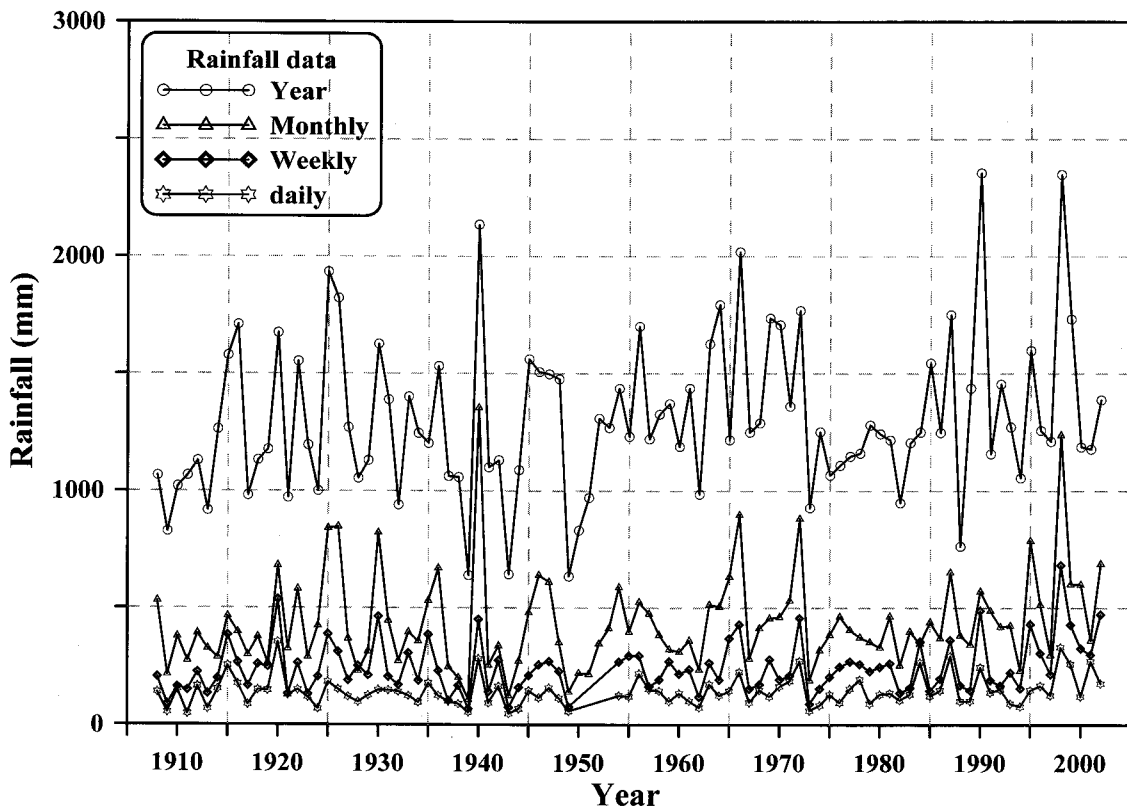


그림 1. 서울관측소의 일, 주, 월 최대 및 연 강수량 자료

3.2 확률강우량의 정상성 판단

시계열자료는 시간이 흐름에 따라 관찰된 값들로 정의된다. 이 자료를 분석하기 위해서는 이 자료를 발생시킨 확률구조(probability structure)에 대해 생각해 볼 필요가 있다. 시계열 분석에서 중요한 개념의 하나는 정상성이다. 어떤 확률변수계열 $\{\dots, y_{-1}, y_0, y_1, \dots\}$ 이 엄격한 의미의 정상성을 갖기 위한 필요충분 조건은 임의의 자연수 m, k 와 임의의 시점들 t_1, t_2, \dots, t_k 에 대해 $y(t_1), y(t_2), \dots, y(t_k)$ 와 $y(t_1+m), y(t_2+m), \dots, y(t_k+m)$ 이 동일한 결합확률분포를 갖는 것이다. 여기서 $y(t_i)$ 는 시점 t_i 에서의 확률변수이다. 또한 이 정상성의 정의는 확률변수계열, 즉 확률과정 $\{y_t\}$ 에 대한 것이나, 편의상 이 확률과정의 실현 값인 시계열자료가 정상성을 갖는다고 말한다. 그러나 이러한 엄격한 조건을 만족하기는 매우 어려우며 따라서 다음과 같은 덜 엄격한 기준을 가지고 정상성을 판단하게 된다. 즉, 시계열 자료 $\{y_t\}$ 가 정상성을 가지면 다음을 만족한다: (1) 평균값 $E(y_t)$ 는 일정하다. 즉, 시점 t 에 의존하지 않는다. (2) 분산 $Var(y_t)$ 도 시점 t 에 의존하지 않는다. (3) 모든 t, h 에 대하여 $Cov(y_t, y_{t+h}) = Cov(y_0, y_h)$.

각 조건이 의미하는 바를 고찰해 보면 다음과 같다. 먼저 (1)은 자료의 수준(level) 또는 평균이 일정함을 의미하고, (2)는 자료 사이에 변화하는 폭에 증감이 없이 일정하다는 분산의 안정성을 의미한다. 마지막으로 조건 (3)은 정상성의 조건 중 가장 중요한 것으로 시간 t 와 무관하게 일정한 간격을 두고 떨어져 있는 두 변수들의 공분산이 동일함을 의미하며, 특별히 분산이 동일함을 의미한다. 일반적으로 정상성을 갖는 시계열이라 함은 (2)번 조건이 (3)번 조건에 포함되므로 시간에 따라 평균의 증가나 분산의 변화가 없다는 것을 의미한다. 이러한 조건은 각 자료의 성질 뿐 아니라 자료들간의 구조 또는 패턴의 항상성을 의미하는 것이다.

정상성 판단 방법에는 일반적으로 변동성 분석방법인 Dickey-Fuller의 단위근 검정, Mann-Whitney 검정, Sign 검정, Abbe 검정, T 검정, F 검정 등이 있으며, 또한 경향성 또는 추세 분석방법으로 Hotelling-Pabst 검정, Mann-Kendall 검정, Sen 검정, Cox-Stuart 검정 등이 있다. 이들 방법 중 본 연구에서는 변동성 분석방법으로 Dickey-Fuller의 단위근 검정을, 추세검정 방법으로 Cox-Stuart 검정을 각각 선택하였다. 본 연구에서 선택한 방법은 타

방법과 약간의 차이를 보일 수도 있으나 본 연구가 검정방법을 비교하는 것이 아닌 확률강우량의 추세분석 자체에 목적을 두고 있으므로 각각의 경우에 하나씩을 선택하여 유의한 결과를 유도하는데 크게 무리가 없을 것으로 판단된다. 참고로 이들 방법의 특성은 참고문헌(Krishnaiah와 Sen 1984; Sheskin, 1997; Kishor, 2001)에서 자세히 살펴볼 수 있으며, 또한 서울시 수해백서(1998)에서는 다양한 적용 예를 살펴볼 수 있다.

먼저, Cox-Stuart 추세검정은 부호검정의 변형된 형태이다(Cox와 Stuart, 1955). 즉, 먼저 관측된 관측 값이 나중에 관측된 관측 값보다 클 때 그 쌍은 (-)부호를 부여하고, 그 반대이면 (+) 부호를 부여한다. 이때, (+)부호들이 많으면 많을수록 상향적 추세를 의미하고 (-)부호들이 많으면 많을수록 하향적 추세를 의미하며, (+)부호와 (-)부호들의 개수가 동일하면 추세가 없는 것을 의미한다는 것이 Cox-Stuart 추세검정의 기본 원리이다. Dickey-Fuller의 단위근 검정도 자료의 정상성을 판단하기 위하여 많이 쓰이는 방법의 하나이다(Dickey와 Fuller, 1979). 먼저, 다음과 같은 단순한 모형에 최소자승법을 적용하고 회귀계수에 대한 가설검정을 수행하는 경우를 생각해 보자

$$y_t = \alpha y_{t-1} + u_t \quad u_t \sim i.i.d. (0, \sigma_u^2) \quad (10)$$

만일 위 식에서 α 의 참값이 절대값으로 1보다 작은 경우 최소자승추정량 $\hat{\alpha}$ 는 다음과 같이 정규분포로 근사(approximate) 가능하다.

$$\begin{aligned} & \sqrt{T}(\hat{\alpha} - \alpha) \\ & \rightarrow N(0, \sigma_u^2 \cdot [\sigma_u^2 / (1 - \alpha^2)]) = N(0, 1 - \alpha^2) \end{aligned} \quad (11)$$

기술적으로 말하면 식 (11)은 최소자승추정량 $\hat{\alpha}$ 이 \sqrt{T} 의 비율로 모수 α 에 수렴하며 $\hat{\alpha}$ 의 분산은 $(1 - \alpha^2)/T$ 가 됨을 의미한다. 이때 통상의 통계적 검정절차를 따라 T 검정 등을 수행할 수 있다.

Dickey-Fuller 단위근 검정의 단위근 검정에서 귀무가설과 대립가설은 다음과 같다.

$$\begin{aligned} & \text{귀무가설} \rightarrow H_0 : \alpha = 1 \\ & \text{(즉, 단위근이 존재한다)} \\ & \text{대립가설} \rightarrow H_1 : \alpha \neq 1 \\ & \text{(즉, 단위근이 존재하지 않는다)} \end{aligned}$$

단위근 검정 결과 유의 확률이 0.05보다 작으면 귀무가설이 기각되어 단위근을 갖는다는 것이고 이는 자료가 정상 시계열이라는 의미가 된다. 그러나, 유의 확률이 0.05보다 크면 귀무가설을 기각할 수 없으므로 자료가 단위근을 갖지 않는다고 본다. 즉, 정상 시계열이 아니라는 의미가 된다.

3.3 서울지점 강수량 자료의 정상성 판단

3.3.1 Cox-Stuart의 추세검정

먼저 Cox-Stuart 추세검정을 이용하여 서울지점 강우관측 자료에 대하여 추세검정을 수행한 결과는 표 1과 같다.

위 식에서 Z값의 추정치는 Cox와 Stuart(1955)의 유도결과에 따라 계산된 값이다. 유의수준 α 가 0.05이므로 기각영역(Z_α : 단측검정 or $Z_{\alpha/2}$: 양측검정)은 1.96이다. 대표본 근사이론에 의해, 계산된 Z값이 -1.96보다 작으면 서울지점 강우관측자료는 상향추세가 있는 것이고 +1.96보다 크면 하향 추세가 있는 것이다. 또한 Z값이 -1.96과 +1.96사이에 있으면 상향적 추세나 하향적 추세가 없다고 할 수 있다.

표 1을 보면 모든 자료계열에서 Z값이 -1.96과 +1.96사이에 있는 것을 볼 수 있다. 따라서 서울지점 강우관측 자료에 대해서는 상향적 추세나 하향적 추세가 없다는 결론을 얻을 수 있다.

3.3.2 Dickey-Fuller의 단위근 검정

서울지점 강우관측 자료가 정상성을 가지고 있는가를 판단하기 위하여 통계 프로그램인 SAS (Statistical Analysis System)를 이용하여 Dickey-Fuller의 단위근 검정을 수행하였다.

검정 결과는 표 2에 수록하였다. 표 2를 보면 모든

자료계열에서 계산된 유의 확률 p가 유의 수준 0.05보다 작으므로 귀무가설은 기각된다(Dickey와 Fuller, 1979). 따라서 대립가설인 $\alpha \neq 1$ 즉, 단위근이 존재하지 않는다는 결론을 얻을 수 있고 이는 원자료계열이 모두 정상시계열임을 의미한다.

지금까지 Cox-Stuart의 추세검정과 Dickey-Fuller(DF)의 단위근 검정 결과 서울지점 강우 관측 자료에 대해서는 어떠한 추세도 발견되지 않았으며 정상시계열임을 확인 할 수 있었다. 참고로 1998년 서울시 수해백서에서도 본 연구와 유사한 결과가 도출된 바 있다.

3.4 서울지점 확률강우량의 정상성 판단

3.4.1 Cox-Stuart의 추세검정

서울지점의 확률강우량을 시계열자료로 만들기 위하여 자료의 수를 증가시켜 가면서 빈도해석을 수행하였다. 우선 자료의 최소 개수를 21개(1908~1928)로 하여 각각 일, 주, 월 최대강우량 및 년 강우량에 대해서 재현기간 10, 20, 50, 100년의 확률강우량을 산정하였다. 다음으로 자료를 한 개 더 추가하여 22개 자료를 가지고 동일한 방법으로 확률강우량을 산정하였다. 이런 방법으로 자료를 하나씩 계속 추가하면서 모든 자료에 대해서 빈도별로 확률강우량을 산정하였다. 이렇게 하여 총 75개의 확률강우량을 산정하였는데 이것은 1928년 이후 확률강우량이 어떻게 변동하고 있는지를 보여주는 시계열 자료가 된다.

전절에서와 동일한 방법으로 유의 수준 0.05에서 확률강우량에 대해 Cox-Stuart의 추세검정을 수행하였다. 그 결과는 표 3과 같다. 참고로 표 3에서의 부호의 총 개수는 자료기간과 같지 않으며, 아울러 결측기간이 고려된 결과이다. 예를 들어 주자료와 일자료의

표 1. 서울지점 관측자료의 Cox-Stuart 추세검정 결과

자료계열	(-)부호의 개수	(+)부호의 개수	동점인 경우	Z	비고
연 강수량	17	30	0	-1.896	추세 없음
월 최대	18	29	0	-1.605	추세 없음
주 최대	19	26	0	-1.043	추세 없음
일 최대	20	25	0	-0.745	추세 없음

표 2. Dickey-Fuller 단위근 검정을 이용한 서울지점 관측자료의 정상성 판단

강수량 자료 계열	유의 확률 (p-value)	유의 수준	비고
년 자료	0.0000955	0.05	정상 시계열
월 최대 자료	0.0001165	0.05	정상 시계열
주 최대 자료	0.0034609	0.05	정상 시계열
일 최대 자료	0.0000955	0.05	정상 시계열

표 3. 확률강우량의 Cox-Stuart 추세검정 결과

강수량 자료계열	분포형	빈도	(-)부호의 개수	(+)부호의 개수	동점인 경우	Z	추세
년 자료	EV1	10	7	30	0	-3.780	상향
		20	10	27	0	-2.800	상향
		50	12	25	0	-2.140	상향
		100	13	24	0	-1.810	없음
월 자료	EV1	10	19	18	0	0.164	없음
		20	19	18	0	0.164	없음
		50	19	18	0	0.164	없음
		100	19	18	0	0.164	없음
주 자료	EV1	10	29	6	0	3.888	하향
		20	27	8	0	3.212	하향
		50	28	7	0	3.550	하향
		100	27	8	0	3.212	하향
일 자료	EV1	10	30	5	0	4.226	하향
		20	27	8	0	3.212	하향
		50	19	15	1	0.507	없음
		100	25	10	0	2.535	하향

경우, 총 부호의 개수는 [총 자료기간(95년)-결측된 자료기간(4년)-자료의 최소 개수(21년)]/2, 즉 35개가 된다. 반면에 연자료와 월자료의 경우에는 결측된 자료기간이 없으므로 [총 자료기간(95년)-자료의 최소 개수(21년)]/2, 즉 37개가 된다.

표 3을 보면 월 최대 자료의 경우에는 추세가 없는 것으로 나타나나 주 최대 자료의 경우에는 하향 추세가 있는 것을 알 수 있다. 년 자료와 일 최대 자료의 경우에는 재현기간에 따라 상향 추세나 하향 추세가 발견된 것도 있으나 추세가 없는 경우도 살펴볼 수 있다. 따라서 확률강우량에 대해서는 전체적으로 일관된 '상향' 또는 '하향' 추세는 없는 것으로 파악된다. 아울러 '추세가 있다' 또는 '추세가 없다' 하는 것 자체도 쉽게 판단하기가 어렵다.

3.4.2 Dickey-Fuller의 단위근 검정

서울지점의 확률강우량에 대해서도 정상성을 판단하기 위하여 전절에서와 같이 SAS(Statistical Analysis System)를 이용하여 Dickey-Fuller의 단위근 검정을 수행하였다. 검정 결과는 표 4에 수록하였다. 표 4는 각 자료계열 별로 빈도에 따른 유의 확률을 표시한 것이다.

아래 표 4를 보면 모든 경우에서 유의 확률(p-value)이 유의수준인 0.05보다 큰 것을 볼 수 있다. 따라서 '단위근이 존재한다' 라는 귀무가설을 기각할 수가 없다. 즉, $\alpha = 1$ 이라는 귀무가설에 따라

확률강우량은 '단위근이 존재한다' 라는 결론을 얻을 수 있고, 이는 확률강우량이 비정상 시계열임을 의미한다.

이러한 결과는 서론에서도 언급한 것과 같이 확률강우량이 추정된 시점 및 그 이전의 모든 자료가 함께 고려되어 산정되기 때문이다. 즉, 나중 시점의 확률강우량은 이전 시점의 확률강우량과 대부분의 자료를 공유하고 있다. 따라서 확률강우량은 강우량에 비해 보다 평활화되어 있을 가능성이 크며, 또한 연속된 두 자료사이의 독립성도 만족시키기 어렵다. 이동평균된 자료나 백색잡음을 연속적으로 누가하여 유도하는 확률보행과정(random walk process)의 경우도 유사한 형태로 이해할 수 있다. 따라서, 현재 일반적으로 사용되는 Cox-Stuart의 추세검정과 Dickey-Fuller의 단위근 검정 방법을 이용하는 경우 확률강우량이 정상적이라고 판단되기는 어렵다.

그러나 이러한 결과에도 불구하고 확률강우량이 '통계학적 정상성'을 만족시키지 못한다고 판단하는 데는 무리가 있다. 실제 확률강우량의 변동정도도 매우 작을 뿐만 아니라 장기적인 증가 또는 감소추세를 확인할 수 없기 때문이다. 만일 그런 추세가 확인하다면 확률강우량은 더 이상 설계기준으로 사용되기 어렵게 된다. 따라서 현재 일반적으로 사용되는 Cox-Stuart의 추세검정과 Dickey-Fuller의 단위근 검정 방법으로 확률강우량의 정상성을 판단하는 데는 무리가 있음을 파악할 수 있다. 즉, 확률강우량의 특성

표 4. Dickey-Fuller의 단위근 검정을 이용한 확률강우량의 정상성 판단

강수량 자료계열	분포형	빈도	유의 확률	유의 수준	비 고
년 자료	EV1	10	0.380	0.05	비정상 시계열
		20	0.131	0.05	비정상 시계열
		50	0.081	0.05	비정상 시계열
		100	0.039	0.05	비정상 시계열
월 자료	EV1	10	0.368	0.05	비정상 시계열
		20	0.116	0.05	비정상 시계열
		50	0.296	0.05	비정상 시계열
		100	0.186	0.05	비정상 시계열
주 자료	EV1	10	0.709	0.05	비정상 시계열
		20	0.631	0.05	비정상 시계열
		50	0.634	0.05	비정상 시계열
		100	0.528	0.05	비정상 시계열
일 자료	EV1	10	0.714	0.05	비정상 시계열
		20	0.626	0.05	비정상 시계열
		50	0.121	0.05	비정상 시계열
		100	0.487	0.05	비정상 시계열

을 반영하는 새로운 정상성 판단기준이 필요하며, 새롭게 제시되는 기준을 통해 기후변화 또는 기후변동의 영향으로 가능한 확률강우량의 비정상성을 판단해 볼 수 있어야 할 것이다.

4. 결 론

본 연구에서는 강우량 및 확률강우량의 변화추이를 알아보고 최근에 빈번하게 관측된 이상강우를 고려하는 경우에도 확률강우량이 정상성을 갖고 있는지에 대해서 분석해 보았다. 본 연구에서는 특히 시계열의 정상성을 판단하는 대표적인 방법인 Cox-Stuart의 추세검정과 Dickey-Fuller의 단위근 검정 방법을 강우량 자료 및 확률강우량에 적용하여 그 결과 및 방법론 적용의 문제점 등을 평가해 보았다.

먼저 Cox-Stuart의 추세검정을 이용한 서울지점 강우량 자료에 대한 분석에서는 강우량이 증가하거나 감소하는 추세가 없다는 판단을 할 수 있었고, 아울러 Dickey-Fuller의 단위근 검정에서도 정상적인 시계열이라는 결론을 얻을 수 있었다. 그러나 서울지점의 확률강우량에 대해서는 전체적으로 어떤 상승 또는 하향의 일관된 추세를 판단하기가 어려웠다. 즉, 강우 자료에 따라 상승, 하향의 특성이 다르게 나타나는 결과를 확인할 수 있었다. 반면에 Dickey-Fuller의 단위근 검정에서는 확률강우량이 비정상시계열임을 판단할 수 있었다.

이러한 결과는 강우량과 확률강우량의 차이에서 비롯된 것이다. 먼저, 확률강우량이 비정상 시계열로 판정된 이유에 대해 다음과 같은 판단이 가능하다. 기본적으로 강우는 무작위 변량으로서 어떤 경향성이나 비정상성을 찾기 힘들다. 반대로 확률강우량은 관측된 모든 자료를 고려하여 추정되므로 전·후의 값의 상관성이 매우 커지게 된다. 즉, 정상시계열인 강우자료가 연속적으로 추가되며 확률강우량이 추정되므로 전·후 값이 높은 상관성이 가능하다, 따라서 확률강우량이 비정상 시계열로 판단되는 것은 본 연구의 판단기법으로는 당연한 결과라 할 수 있겠다.

그러나 문제는 확률강우량이 Cox-Stuart의 추세검정 결과 일부 경우에서 추세가 나타나고 Dickey-Fuller의 단위근 검정결과 비정상 시계열로 판단되었다고 해서 확률강우량이 비정상 시계열이라고 판단해야 하느냐 하는 것이다. 비정상의 의미는 상향 또는 하향의 추세가 유의하다는 것 또는 그 변동성이 크게 변화한다는 것을 의미한다. 최근 발생한 이상 강우사상이 그런 모습을 보여주고 있는 것이 사실이나, 기본적으로 강우량 자체는 정상시계열로 판단되므로 확률강우량 또한 정상성을 갖는 것으로 판단되는 것이 적절하다. 그러나 현재의 방법론에 의해 확률강우량의 정상성을 판단하는 것은 다소 무리가 있으므로, 새로운 정상성 판단 방법의 개발이 필요하다고 하겠다.

5. 감사의 글

본 연구는 건설교통부가 출연하고 한국건설교통기술평가원에서 위탁시행 한 2003년도 건설기술혁신사업 (03산학연C01-01)에 의한 도시홍수재해관리기술연구사업단의 연구성과입니다.

참 고 문 헌

건설교통부 (1988) 수자원관리기법개발 연구조사보고서. 제2권 한국 확률강우량의 작성.

건설교통부 (2000) 수자원관리기법개발 연구조사보고서. 제1권 한국 확률강우량의 작성.

권태영, 오성남, 박상욱 (1998) 한반도 여름 강우의 장기변동과 지역적 특성. 한국기상학회지, 한국기상학회, 제34권, 제1호, pp. 20-30.

김명직, 장국현 (1998) 금융시계열 분석. 경문사.

서규우 (1998) 지구온난화에 따른 수문환경의 변화. 한국수자원학회지, 한국수자원학회, 제31권, 제4호, pp. 103-106.

안재현, 김태용, 유철상, 윤용남 (2000) 자료기간 증가에 따른 확률강우량의 거동특성분석. 한국수자원학회논문집, 한국수자원학회, 제33권, 제5호, pp. 569-580.

안재현, 유철상, 윤용남 (2002) 확률강우량의 상승 및 감소와 관련하여 : 홍수량과 홍수위에 대한 영향 분석. 한국수자원학회 학술발표회 논문집(II), 한국수자원학회, pp. 1274-1279.

이상복, 이동률, 허준행 (2000) 남한지역 강수량의 변동성 및 경향성 해석. 대한토목학회 논문집, 대한토목학회, 제20권, 제7호, pp. 103-106.

이상열 (2001) 시계열 분석의 원리. 자유아카데미.

제갈돈, 류지성 (1998) 간여 시계열 분석에서 분산의 비정상성(non-stationarity)과 Log 변형. 정책분석평가학회보, 제8권, 제2호, pp. 5-18.

정중호, 윤용남 (2002) 수자원 설계 실무. 구미서관.

차영준 (1993) SPSS/PC+를 이용한 비모수 통계학. 자유아카데미.

차은정 (2001) 집중호우의 원인은 지구온난화 때문인가? 한국 수자원학회지, 한국수자원학회.

차은정, 최영진 (2000) 한반도 여름철 집중호우의 시간공간 변동 특성. 한국 수자원학회지, 한국수자원학회, 제33권, 제4호, pp. 47-56.

최병선 (2001) 단변량 시계열 분석. 세경사.

Ahn, J.H., Kim, T.W., Yoo, C., and Yoon, Y.N.

(2001) On the variation of frequency-based rainfall amounts: A case study for evaluating recent extreme rainfalls in Korea. *Stochastic Environmental Research and Risk Assessment*, Vol. 17, No. 3, pp. 217-227.

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

Dickey, D.A. and Fuller, W.A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, Vol. 74, pp. 427-431.

Hosking, J.R.M. (1989) *The theory of probability weighted moment*. Research Report, RC 12210. IBM Research Division, T.J. Watson Research Center. Now York.

Jhones, P.D. and Wigley, T.M.L. (1990) Global warming trends. *Scientific American*, Vol. 263, No. 2, pp. 84-91.

Kishor, S.T. (2001) *Probability and statistics with reliability, queueing, and computer science applications*. Wiley Interscience.

Krishnaiah, P.R. and Sen, P.K. (1984) *Handbook of statistics 4: Nonparametric methods*. North-Holland, Elsevier.

Lettenmaier, D.P. and Gan, T.Y. (1990) Hydrologic sensitivities of the Sacramento-San Joaquin river basin, California, to global warming. *Water Resources Research*, Vol. 26, No. 1, pp. 69-86.

Lettenmaier, D.P. and Sheer, D.P. (1991) Climate sensitivity of California water resources. *Journal of Water Resources Planning and Management*, Vol. 117, No. 1, pp. 108-125.

Mansell, M.G. (1997) The effect of climate change on rainfall trend and flooding risk in the west of Scotland. *Nordic Hydrology*, Vol. 28, pp. 37-50.

McCabe, G.J. and Ayers, M.A. (1989) Hydrologic effect of climate change in the Delaware river basin. *Water Resources Bulletin*, Vol. 25, No. 6, pp. 1231-1241.

McCabe, G.J. and Hay, L.E. (1995) Hydrologic

- effect of hypothetical climate change in the East river basin, Colorado, USA. *Hydrological Science Journal*, Vol. 40, pp. 303–318.
- Nash, L.L. and Gleick, P.H. (1991) Sensitivity of streamflow in the Colorado basin to climate changes. *Journal of Hydrology*, Vol. 125, pp. 221–241.
- Nemec, J. and Schaake, J. (1982) Sensitivity of water resources systems to climate variation. *Hydrological Science Journal*, Vol. 27, No. 3, pp. 327–343.
- Ng, H.Y.F. and Marsalek, J. (1992) Sensitivity of streamflow simulation to changes in climate inputs. *Nordic Hydrology*, Vol. 23, pp. 257–272.
- Rao, A.R. and Hamed, K.H. (2000) *Flood frequency analysis*. CRC Press LLC.
- Richard, W.K., Marc, B.P., and Philippe N. (2002) Statistics of extremes in hydrology. *Advances in Water Resources*, Vol. 25, pp. 1287–1304.
- Rind, D., Goldberg, R., and Ruedy, R. (1989) Change in climate variability in the 21st century. *Climatic Change*, Vol. 14, pp. 5–37.
- Sheskin, D.J. (1997) *Handbook of parametric and nonparametric statistical procedures*. CRC Press, Boca Raton, Fl.
- Thomas, E.C.II. (2001) Climate-biased storm-frequency estimation. *Journal of Hydrologic Engineering*, Vol. 6, pp. 275–283.
- Trenberth, K.E. (1999) Atmospheric moisture recycling: Role of advection and local evaporation. *Journal of Climate*, Vol. 12, No. 5, pp. 1368–1381.
- Trenberth, K.E. and Guillemot, C.J. (1998) Evaluation of the atmospheric moisture and hydrological cycle in the NCEP/NCAR reanalyses. *Climate Dynamics*, Vol. 14, No. 3, pp. 213–231.
- Wardlaw, R.B., Hulme, M., and Stuck, A.Y. (1996) Modeling the impacts of climate change on water resources. *Journal of Chartered Institution of Water and Environmental Management (J.CIWEM)*, Vol. 10, pp. 355–364.

© 논문접수일 : 2007년 07월 27일

© 심사의뢰일 : 2007년 07월 30일

© 심사완료일 : 2007년 10월 17일