

강우빈도해석에서의 측우기자료의 유용성 평가

Evaluation for usefulness of Chukwookee Data in Rainfall Frequency Analysis

김기욱*, 유철상**, 박민규***, 김대하****, 박상형*****, 김현준*****
Keewook Kim, Chulsang Yoo, Minkyu Park, Daeha Kim, Sanghyoung Park, Hyeon Jun Kim

Abstract

In this study, the chukwookee data were evaluated by applying that for the historical rainfall frequency analysis. To derive a two parameter log-normal distribution by using historical data and modern data, censored data MLE and binomial censored data MLE were applied. As a result, we found that both average and standard deviation were all estimated smaller with chukwookee data than those with only modern data. This indicates that rather big events rarely happens during the period of chukwookee data than during the modern period. The frequency analysis results using the parameters estimated were also similar to those expected. The point to be noticed is that the rainfall quantiles estimated by both methods were similar, especially for the 99% threshold. This result indicates that the historical document records like the annals of Chosun dynasty could be valuable and effective for the frequency analysis. This also means the extension of data available for frequency analysis.

Key words: historical rainfall frequency analysis, censored data MLE, binomial censored data MLE

1. 서 론

한정된 자료는 빈도해석을 수행함에 있어서 가장 큰 문제가 된다(Condie and Lee, 1982; Hosking and Wallis, 1986; Thorndycraft et al., 2003; Reis Jr. and Stedinger, 2005). 이 때, 역사적 증거나 문서, 물리적 증거 등의 역사적 자료는 자료의 확장, 정확도의 증가 등에 기여하여 빈도해석 결과의 개선에 도움을 준다(Tasker and Thomas Jr., 1978; Stedinger and Cohn, 1986).

국내에서도 강우나 홍수의 빈도해석 시에 항상 자료의 부족 문제에 부딪히게 된다. 따라서 우리나라의 역사적 자료인 측우기판측자료와 조선왕조실록에 기록된 강우 및 홍수자료를 빈도해석에 이용한다면 자료의 확장에 도움이 될 것으로 판단된다. 이에 본 연구에서는 대표적인 역사적 빈도해석방법들을 이용하여 측우기판측자료의 적용성을 검토해 보았다.

2. Historical Freqeuncy Analysis

본 연구에서는 근대관측자료와 역사적관측자료를 이용하여 2변수 대수정규분포를 유도하기 위한 두 가지 방법을 고려하였다.

* 정회원·고려대학교 사회환경시스템공학과 박사수료·E-mail: kkw5287@korea.ac.kr
** 정회원·고려대학교 사회환경시스템공학과 교수·E-mail: envchul@korea.ac.kr
*** 정회원·고려대학교 사회환경시스템공학과 박사과정·E-mail: minkyu@korea.ac.kr
**** 정회원·한국건설기술연구원 유령조사사업단 연구원·E-mail: violt@kict.re.kr
***** 정회원·고려대학교 사회환경시스템공학과 석사과정·E-mail: sh79863@korea.ac.kr
***** 정회원·한국건설기술연구원 수자원연구부 수석연구원·E-mail: hijkim@kict.re.kr

2.1 Censored data MLE

역사적 자료를 이용하여 빈도해석을 수행하는 방법 중 가장 간단한 방법인 모멘트조정법은 정량적인 관측자료와 정량, 정성적인 역사적 자료가 함께 존재하는 경우에 적용에 문제점을 가진다(Stedinger and Cohn, 1986). 따라서 일반적으로 범주형자료의 분석을 위해서는 최우도법을 많이 사용하게 된다(Bishop et al., 1975). 최우도법은 연속적인 관측자료와 역사적관측자료, 고자료 등이 섞여 있는 자료를 분석하는데 적절한 방법이다(Stedinger and Cohn, 1986).

s 개의 근대관측자료의 대수값 $\{X_1, X_2, \dots, X_s\}$ 과, 특정한 임계치 $X_0 = \ln(P_0)$ 를 초과하는 k 개의 역사적관측자료의 대수값 $\{Y_1, Y_2, \dots, Y_k\}$ 이 있다고 가정하면 2변수 대수정규분포의 우도함수는 식 (1)과 같이 나타낼 수 있다.

$$L(\mu, \sigma) = \prod_{i=1}^s f_x(x_i) \left\{ \binom{h}{k} F_x(X_0)^{(h-k)} [1 - F_x(X_0)]^k \right\} \prod_{j=1}^k f_Y(y_j) \quad (1)$$

여기서, $f_x(\cdot)$, $F_x(\cdot)$ 는 각각 X 의 확률밀도함수, 누가확률밀도함수이며, $f_Y(\cdot)$ 는 Y 의 확률밀도함수이다. Y 가 X_0 보다 항상 크므로 이를 적용하고 식 (1)의 양변에 대수를 취하면 대수우도함수를 얻을 수 있다.

$$L(\mu, \sigma) = \sum_{i=1}^s \ln[f(X_i|\mu, \sigma)] + \ln \binom{h}{k} + (h-k) \ln[F(X_0|\mu, \sigma)] + \sum_{j=1}^k \ln[f(Y_j|\mu, \sigma)] \quad (2)$$

수치해법을 통해 식 (2)를 최대로 하는 μ , σ 를 얻을 수 있다.

2.2 Binomial censored data MLE

역사적관측자료는 그 특성상 정성적인 자료로 존재하는 경우가 많이 있다. 이러한 경우에는 그 정확한 값을 알 수 없으므로 앞에서 설명한 방법을 그대로 적용하는 것은 어렵다. 따라서 역사적관측자료의 기간(h) 중 특정한 임계치 $X_0 = \ln(P_0)$ 를 초과한 횟수(k)만을 이용하여 분석을 수행해야 한다. 이러한 자료를 binomial censored data라 한다(Stedinger and Cohn, 1986).

Binomial censored data MLE에 대한 우도함수는 식 (1)의 censored data MLE의 우도함수에서 역사적관측자료의 값을 생략한 형태로 나타난다.

$$L(\mu, \sigma) = \prod_{i=1}^s f_x(x_i) \left\{ \binom{h}{k} F_x(X_0)^{(h-k)} [1 - F_x(X_0)]^k \right\} \quad (3)$$

식 (3)의 양변에 대수를 취하면 대수우도함수를 얻을 수 있다.

$$L(\mu, \sigma) = \sum_{i=1}^s \ln[f(X_i|\mu, \sigma)] + \ln \binom{h}{k} + (h-k) \ln[F(X_0|\mu, \sigma)] + k \ln[1 - F(X_0|\mu, \sigma)] \quad (4)$$

수치해법을 통해 식 (4)를 최대로 하는 μ , σ 를 얻을 수 있다.

3. 측우기 관측자료를 이용한 Historical Frequency Analysis

앞에서 설명한 censored data MLE와 binomial censored data MLE를 이용하여 historical frequency analysis를 수행하였다. 분석에 필요한 임계치로 전체 자료 분포의 상위 90%, 99%의 두 값은 고려하였다. 그 결과 90%의 임계치는 201.5mm로 이보다 큰 10개의 측우기 관측자료가 선정되었으며, 99%의 임계치는 354.7mm로 1개의 측우기 관측자료가 선정되었다(그림 1).

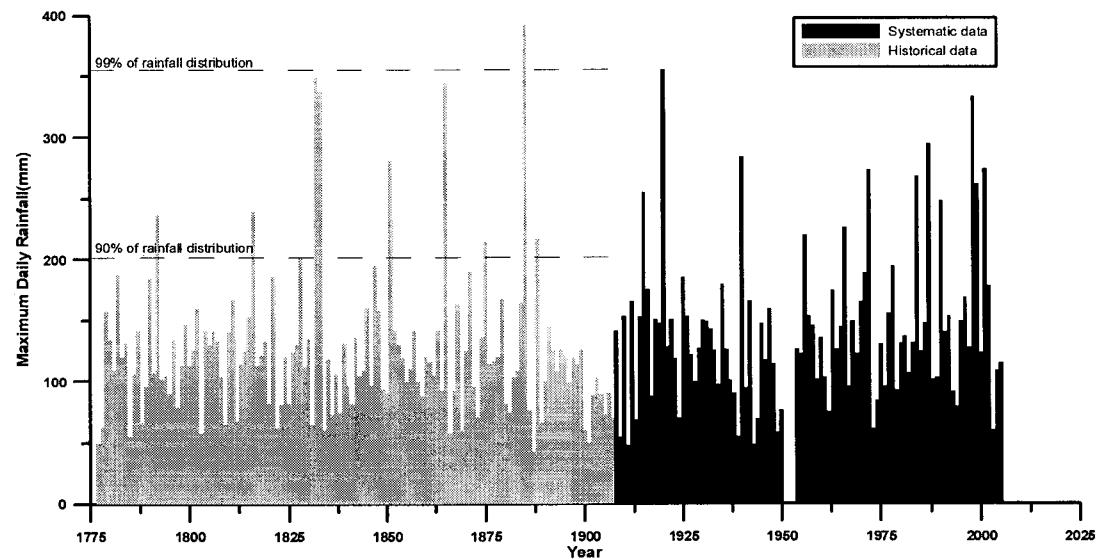


그림 1. 전체자료에 대한 90%, 99%의 임계치

그림 1의 근대 관측자료와 임계치에 따라 선정된 측우기 관측자료를 이용하여 2변수 대수 정규분포의 매개변수를 추정해 보았다. 이를 위하여 도해법을 이용하여 식 (2)와 식 (4)를 최대로 만드는 μ , σ 를 산정하였다. 그 결과 censored data MLE를 이용한 방법에서는 90%의 임계치를 이용하는 경우 $\mu = 4.8240$, $\sigma = 0.4179$, 99%의 임계치를 이용하는 경우 $\mu = 4.8639$, $\sigma = 0.4217$ 를 얻었으며 binomial censored data MLE를 이용한 방법에서는 90%의 임계치를 이용하는 경우 $\mu = 4.8241$, $\sigma = 0.4024$, 99%의 임계치를 이용하는 경우 $\mu = 4.8641$, $\sigma = 0.4227$ 를 얻을 수 있었다. 그림 2는 각 방법에서 90%의 임계치를 이용하는 경우의 예를 보여준다.

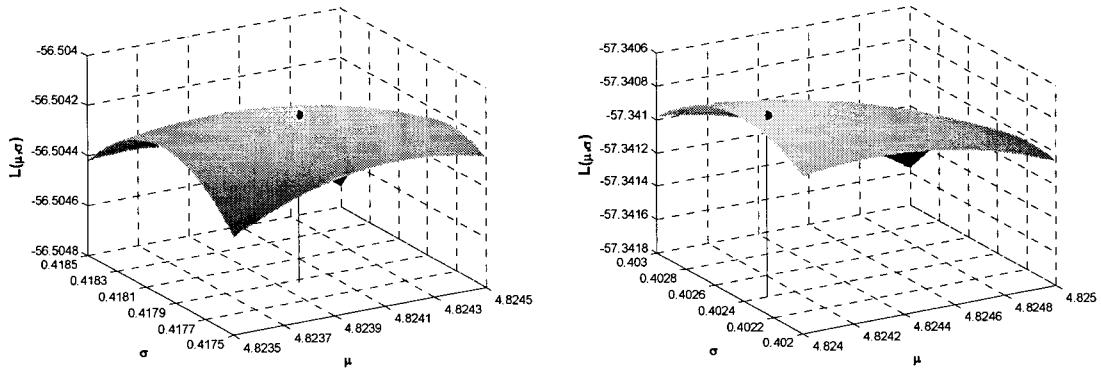


그림 2. 도해법을 이용한 확률밀도함수의 매개변수 추정(좌: censored data MLE, 우: binomial censored data MLE)

4. 확률강우량의 산정

앞에서 얻어진 2변수 대수정규분포의 매개변수와 근대관측자료만을 이용하여 얻어진 매개변수를 이용하여 확률강우량을 산정해보았다. 재현기간 2, 5, 10, 50, 100, 500년에 대한 확률강우량을 산정하였다. 그 결과를 표 1에 나타내었다.

표 1. 각 방법별 재현기간에 따른 확률강우량

재현 기간	Systematic Data	MLE(90%)		MLE(99%)	
		Censored	Binomial	Censored	Binomial
2	129.7	124.5	124.5	129.5	129.6
5	185.8	176.9	174.6	184.7	184.9
10	224.2	212.6	208.5	222.4	222.7
50	311.7	293.6	284.4	308.0	308.7
100	350.2	329.0	317.4	345.5	346.3
500	443.2	414.4	396.3	436.0	437.3

표 1을 보면 근대관측자료를 이용한 경우가 가장 큰 확률강우량을 보이는 것을 알 수 있다. MLE를 이용한 방법의 경우 근대관측자료를 이용한 경우에 비하여 확률강우량이 작게 산정되지만 이는 앞에서 설명한 바와 같이 증가된 자료의 기간에 비하여 관측된 자료에서 큰 값이 드물게 발생하기 때문인 것으로 판단된다. 주목할 점은 측우기관측자료의 값을 이용한 censored data MLE와 자료의 개수만을 이용한 binomial censored data MLE로부터 얻어진 확률강우량이 큰 차이를 보이지 않는데 있다. 특히, 99%의 임계치를 이용한 경우에는 두 방법에서 거의 유사한 확률강우량을 얻은 것을 알 수 있다. 이는 빈도해석에 있어서 정량적인 값을 가지지 않는 자료의 적용성을 나타내는 것으로서 우리나라의 조선왕조실록과 같은 흥수기록을 빈도해석에 이용할 수 있다는 것을 의미한다.

5. 결 론

본 연구에서는 우리나라의 측우기자료에 대한 historical frequency analysis의 적용성을 검토하기 위해 censored data MLE와 binomial censored data MLE를 이용하여 2변수 대수정규분포의 매개변수를 추정하였다. 또한 추정된 매개변수를 이용하여 확률강우량을 산정한 뒤 이를 비교해 보았다. 그 결과는 다음과 같다.

- (1) 매개변수의 추정결과, 근대관측자료와 측우기관측자료를 이용하여 MLE를 적용한 경우가 근대자료만을 이용한 경우와 비교하여 평균과 표준편차가 모두 작게 산정되는 것을 알 수 있었다. 이는 측우기관측자료를 이용함으로서 증가된 자료의 기간에 비하여 관측된 자료에서 큰 값이 드물게 발생한다는 것을 의미한다.
- (2) 얻어진 매개변수를 이용하여 확률강우량을 추정해 본 결과, 매개변수 추정결과와 유사한 결론을 얻을 수 있었다.
- (3) 주목할 점은 측우기관측자료의 값을 이용한 censored data MLE와 자료의 개수만을 이용한 binomial censored data MLE로부터 얻어진 확률강우량이 큰 차이를 보이지 않는데 있으며, 특히 99%의 임계치를 이용한 경우에는 두 방법에서 거의 유사한 확률강우량을 얻을 수 있었다.

본 연구의 결과는 정량적인 값으로 나타나지 않는 조선왕조실록과 같은 국내의 역사적 자료를 빙도해석에 효과적으로 이용할 수 있다는 것을 의미하며 이는 기존에 비하여 방대한 양의 자료를 추가로 확보할 수 있도록 해 줄 것이다.

감 사 의 글

본 연구는 건설교통부 한국건설교통기술평가원의 이상기후대비시설기준강화 연구단에 의해 수행되는 2005 건설기술기반구축사업(05-기반구축-D03-01)에 의해 지원되었습니다. 이에 감사드립니다.

참 고 문 헌

1. Bishop, Y. M., Fienberg, S. E., and Holland, P. W. (1975). *Discrete Multi-variate Analysis: Theory and Practice*. MIT Press, Cambridge, Mass.
2. Tasker, G. D. and Thomas Jr., W. O.(1978). Flood-Frequency Analyses with Prerecord Information. *Journal of the Hydraulics Division, ASCE*, Vol. 104, No. 2, pp. 249-259.
3. Condie, R. and Lee, K. A.(1982). Flood Frequency Analysis with Historic Information. *Journal of Hydrology*, Vol. 58, pp. 47-61.
4. Hosking, J. R. M. and Wallis, J. R.(1986). Paleoflood Hydrology and Flood Frequency Analysis. *Water Resources Research*, Vol. 22, No. 4, pp. 543-550.
5. Stedinger, J. R. and Cohn, T. A. (1986). Flood Frequency Analysis with Historical and Paleoflood Information. *Water Resources Research*, Vol. 22, No. 5, pp. 785-793.
6. Thorndycraft, V. R., Benito, G., Llasat, C., and Barriendos, M.(2003). Palaeofloods, Historical Data and Climatic Variability: Applications in Flood Risk Assessment. *Proceedings of the PHEFRA Workshop*, Barcelona, 16-19th, October, 2002.
7. Reis Jr., D. S. and Stedinger, J. R.(2005). Bayesian MCMC Flood Frequency Analysis with Historical Information. *Journal of Hydrology*, Vol. 313, pp. 97-116.