개선된 GEV 분포를 이용한 강우량 빈도분석

Frequency Analysis of Rainfall Data Using Advanced GEV Distribution

이길성*, 강원구, 박경신***, 성진영******

Kil Seong Lee, Won Gu Kang, Kyungshin Park, Jinyoung Sung

요 지

강우는 수자원 확보 측면에서 근원이 되는 요소이다. 그러므로 정확한 확률강우량 산정은 미래의 가용 수자원량을 예측하는데 있어 중요한 사항중 하나이며 무엇보다 신중한 결정이 요구된다. 또한 하천의 범람에 의한 침수를 예방하는 수공구조물 등의 설계에 있어서는 신뢰할 수 있는 확률강우량 산정이 선행되어야 한다. 본 연구에서는 최근 우리나라 극치강우확률분포로서 많은 연구가 이루어지고 있는 GEV 분포(GEV-O)를 기반으로 위치 매개변수에 시간의 함수를 고려한 개선된 GEV 분포(GEV-A)를 이용하여 서울지점에 적용함으로서 GEV-O 분포에 의한 확률강우량과 GEV-A 분포로 산정된 확률강우량을 비교·검토하였다.

먼저 임의의 난수 발생을 통해 최우도추정법과 확률가중모멘트법으로 매개변수를 추정한 GEV-O 분포와 최우도추정법으로 매개변수를 추정한 GEV-A 분포의 상대평균제곱근오차(R-RMSE)를 계산하여 비교함으로서 GEV-A 분포의 효율성을 판단하였다.

사례연구는 1961년부터 2008년까지 서울강우관측소에서 측정된 연최대 1일 강우량으로 하였으며 χ^2-검정, PPCC-검정으로 적합도 검정을 실시하였다. 강우빈도분석 결과 GEV-A 분포가 GEV-O 분포로 산정된 결과 보다 대체로 재현기간 200년 이상일 경우, 과다 산정되는 경향을 보였다.

추후 개선된 GEV 분포를 서울 인근 지점에 적용함으로서 지역빈도해석(Regional Frequency Analysis)을 실행하기 위한 연구가 진행되어야 할 것이다. 또한 확률홍수량 산정 등에도 개선된 GEV 분포를 이용함으로서 보다 정확하고 신뢰성 있는 확률수문량을 예측하여야 할 것이다.

핵심용어 : GEV 분포, 개선된 GEV 분포, 강우확률분포, PPCC, 상대평균제곱근오차

1. 서 론

강우는 수자원 확보에 있어 근간이 되는 요소이다. 그러나 최근에 대두되고 있는 기후변화로 인하여 한반도 기후 양상이 아열대성으로 변화 될 것이라는 비관론이 점차 부상하고 있다. 이는 가용 수자원 확보 측면에서의 불리함 뿐만 아니라 기존의 수공 구조물 설계에 사용되어져 왔던 방법의 수정도 불가피 함을 의미한다. 보다 정확하고 신뢰성 있는 강우량빈도분석은 이와 같은 문

* 정회원・서울대학교 건설환경공학부 교수・공학박사・E-mail : kilselee@snu.ac.kr
** 서울대학교 건설환경공학부 석사과정・E-mail : safewater@snu.ac.kr
*** 서울대학교 건설환경공학부 석사과정・E-mail : diploma@snu.ac.kr
**** 서울대학교 건설환경공학부 석사과정・E-mail : blackleo83@snu.ac.kr

제점을 해결하는 실마리가 될 것이다. 본 연구는 극치확률분포 중 최근 활발한 연구가 진행되고 있는 GEV 분포(GEV-O)와 Coles (2001)가 제안한 개선된 GEV 분포(GEV-A)를 비교·분석함으로서 신뢰성 있는 확률강우량 산정하는데 목적이 있다.

2. 이론적 배경

2.1 GEV 분포(GEV-O)

일반적인 GEV 분포(GEV-O)의 누적분포함수(CDF)는

$$F_{\mathrm{GEV-O}}(x) = \exp\left[-\left\{1-\frac{\kappa}{\alpha}(x-\xi)\right\}^{1/\kappa}\right] \qquad \kappa \neq 0$$
$$= \exp\left[-\exp\left(-\frac{x-\xi}{\alpha}\right)\right] \qquad \kappa = 0$$

이며, 여기서 $\kappa < 0$ 일 때 $\xi + \alpha/\kappa \leq x \leq +\infty$, $\kappa = 0$ 이면 $-\infty \leq x \leq +\infty$ (Gumbel), $\kappa > 0$ 일 때 $-\infty \leq x \leq \xi + \alpha/\kappa$ (Weibull). ξ ($\in$ □), $\alpha(>0)$ 그리고 κ ($\in$ □)는 각각 위치(Location), 규모(Scale), 형상(Shape) 매개변수이며 x 는 분석하고 싶은 자료이다.

2.2 개선된 GEV 분포(GEV-A)

Coles (2001)는 기존 GEV 분포(GEV-O)의 위치 매개변수(ξ)에 시간(t)을 고려한 새로운 GEV 분포를 제안하였다. 본 연구에서는 이를 개선된 GEV 분포(GEV-A)라 지칭하였다. 개선된 GEV 분포(GEV-A)의 누적분포함수(CDF)는

$$F_{\mathrm{GEV-A}}(x) = \exp\left[-\left[1-\frac{\kappa}{\alpha}\{x-(\beta_0+\beta_1 t)\}\right]^{1/\kappa}\right] \qquad \kappa \neq 0$$
$$= \exp\left[-\exp\left\{-\frac{x-(\beta_0+\beta_1 t)}{\alpha}\right\}\right] \qquad \kappa = 0$$

이며, 여기서 β_0와 β_1은 GEV-O 분포의 위치매개변수를 구성하고 있으며 시간(t)에 대한 1차 함수의 형태를 가진다.

3. GEV-A 효율성 검증을 위한 모의실험

GEV-A 분포의 효율성 검증을 위해 상대평균제곱근오차(R-RMSE)를 이용한 모의실험을 실시하였다. 참값은 실제 강우현상과 유사한 경우를 모의하기 위하여 $\xi = 115, \alpha = 50$ 그리고 κ는 - 0.3부터 + 0.1 까지 0.1 단위로 변화시킨 GEV-O 분포의 변위치(Quantile)로 하였다. 모의값은 일양분포(Uniform distribution)를 따르는 0부터 1까지의 n = 30, 1000개 군의 난수를 발생시킨 후 이를 참값의 CDF로 가정하여 산출된 자료를 최우도추정법(MLE)과 확률가중모멘트법(PWMs)으로 매개변수를 추정한 GEV-O 분포와 최우도추정법으로 매개변수를 추정한 GEV-A 분포의 비초과확률 0.990(재현기간 100년)의 경우와 0.995(재현기간 200년)의 경우를 변위치로 하였다.

모의실험 결과는 표 1, 그림 1 그리고 그림 2와 같다.

표 1. GEV-O, GEV-A 분포의 R-RMSE 산정

분 포		GEV-O		GEV-A
p	재현기간(년)	MLE	PWMs	MLE
$\kappa = -0.3$				
0.990	100	0.167	0.265	0.145
0.995	200	0.192	0.328	0.165
$\kappa = -0.2$				
0.990	100	0.149	0.184	0.121
0.995	200	0.170	0.223	0.143
$\kappa = -0.1$				
0.990	100	0.130	0.124	0.098
0.995	200	0.148	0.146	0.111
$\kappa = 0$				
0.990	100	0.113	0.081	0.065
0.995	200	0.126	0.092	0.072
$\kappa = 0.1$				
0.990	100	0.095	0.050	0.044
0.995	200	0.106	0.055	0.046

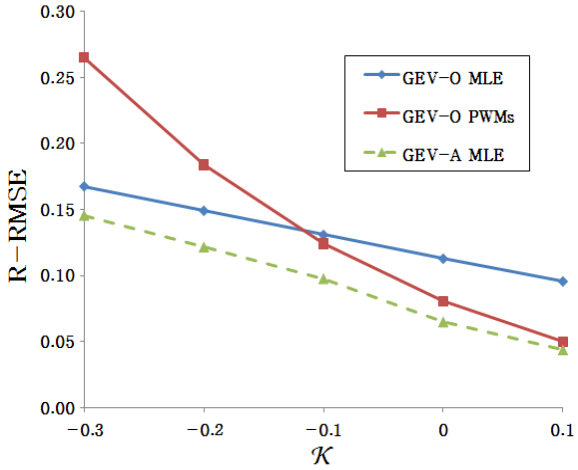

그림 1. 비초과확률 0.990일 때의 R-RMSE

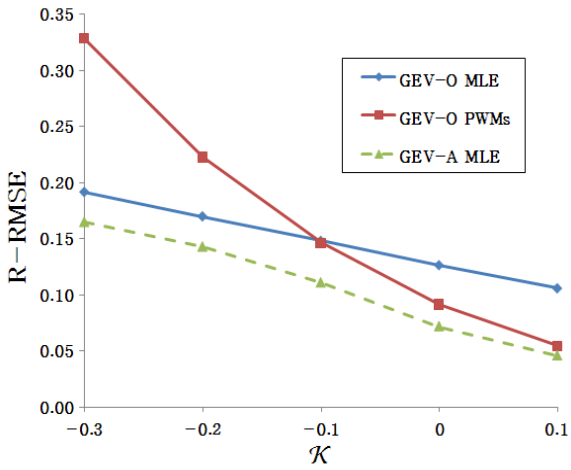

그림 2. 비초과확률 0.995일 때의 R-RMSE

n을 30으로 책정한 이유는 우리나라의 경우 대도시 혹은 주요 지점을 제외하고 아직까지 30년 이상의 자료가 기록된 경우가 많지 않기 때문이다. 결과를 살펴보면 모의를 실시한 κ의 전 구간($-0.3 \le \kappa \le 0.1$)에서 GEV-A 분포가 GEV-O 분포에 비하여 R-RMSE가 낮음을 알 수

있다. 한국건설기술연구원(2000)에 따르면 우리나라 극치강우사상은 $-0.2 \le \kappa \le 0.1$ 인 경우가 대부분이므로 GEV-A 분포 사용시 좀 더 정확하고 신뢰성 있는 강우량 빈도분석이 가능할 것이다.

4. 사례 연구

GEV-A 분포를 실례에 적용해 보기 위하여 서울강우관측소의 1961년부터 2008년까지의 연최대 1일 강우량을 채택하였다. 연구방법은 각 분포의 적합성 검증을 위하여 χ^2-검정과 PPCC-검정을 실시한 후 재현기간에 따른 확률강우량을 산정하여 비교하였다. χ^2-검정은 일반적으로

$$\chi^2 = \sum_{i=1}^{m} \frac{(\mathrm{O}_i - \mathrm{E}_i)^2}{\mathrm{E}_i}$$

이며, 여기서 O_i은 i번째 구간에서의 관측값이며 E_i는 i번째 구간에서의 기대값이므로 χ^2 이 작을수록 자료는 분포에 적합하다고 할 수 있다. 검정결과 GEV-O의 $\chi^2(\chi^2_{\mathrm{GEV-O}})$는 2.750, GEV-A의 $\chi^2(\chi^2_{\mathrm{GEV-A}})$는 2.083 으로 산정되었다. 좀 더 신뢰성있는 분석을 위해 χ^2 방법보다 검정능력이 우수하다고 판단되고 있는 PPCC-검정을 실시하였다. PPCC-검정은 i번째 자료값($x_{(i)}$)과 이에 상응하는 분포의 변위치(w_i)간의 상관관계를 계수(r)로 표현하는 방법을 말한다. 이를 식으로 나타내면

$$r = \frac{\sum\left(x_{(i)} - \overline{x}\right)\left(w_i - \overline{w}\right)}{\left[\sum\left(x_{(i)} - \overline{x}\right)^2 \sum\left(w_i - \overline{w}\right)^2\right]^{0.5}}$$

이며, 상관계수 r이 높을수록 자료는 분포에 적합하다고 할 수 있다. 검정결과 GEV-O의 r ($r_{\mathrm{GEV-O}}$)은 0.939이며 GEV-A의 r($r_{\mathrm{GEV-A}}$)은 0.951로 산정되었다. 이와 같이 두 검정결과 모두 GEV-A 분포가 GEV-O 분포보다 자료에 더욱 적합함을 말해주고 있다.

서울강우관측소의 연최대 1일 강우량을 통한 강우빈도분석을 위하여 MLE와 PWMs로 매개변수를 추정 한 GEV-O 분포와 MLE로 매개변수를 추정 한 GEV-A 분포를 이용하였다. 강우빈도분석 결과는 표 2와 같다. 양적인 측면에서 보았을 때 GEV-O 분포와 GEV-A 분포간의 강우량 차이는 크지 않지만 대체로 재현기간 200년 이상일 경우 GEV-A 분포로 산정된 확률강우량이 GEV-O 분포로 산정된 확률강우량 보다 다소 증가하는 경향을 보였다. 앞서 실시한 적합도 검정을 통하여 서울강우관측소의 경우 GEV-A 분포가 GEV-O 분포보다 적합함을 확인한 바, 재현기간 200년 이상의 강우는 일반적인 GEV 분포로 산정되는 확률강우량보다는 과다 산정 되는 것이 통계적 신뢰성이 높다고 할 수 있다.

표 2. GEV-O, GEV-A 분포로 산정한 서울강우관측소의 확률강우량(mm)

분 포		GEV-O		GEV-A
발생확률	재현기간(년)	MLE	PWMs	MLE
0.200	5	188.423	190.798	179.391
0.100	10	222.377	226.250	216.827
0.020	50	292.422	299.369	295.191
0.013	80	311.388	319.164	317.176
0.010	100	320.202	328.363	327.638
0.005	200	346.857	356.177	360.886
0.002	500	380.524	391.302	409.973

5. 결 론

앞으로의 기후양상은 점차 정상(Stationary)계열을 따르지 않을 것이라는 예상이 지배적인 가운데 Coles (2001)가 제안한 개선된 GEV 분포(GEV-A)는 기존의 GEV 분포(GEV-O)의 위치매개변수에 시간의 함수를 고려함으로서 비정상(Nonstationary)계열 자료의 경향성 분석에도 효과가 있음이 입증된 바 있다. 이에 GEV-A 효율성 검증을 위한 모의실험을 실시한 결과 $-0.3 \leq \kappa \leq 0.1$의 구간에서 GEV-O 분포보다 높은 효율성을 지닌 것을 확인하였다. 사례연구는 1961년부터 2008년까지 서울강우관측소에서 측정된 연최대 1일 강우량을 자료로 하였으며 그 결과 재현기간 200년 이상의 경우 GEV-O 분포에 비하여 GEV-A 분포로 측정된 확률강우량이 다소 과다 산정되는 경향을 보였다. 추후 GEV-A 분포를 서울 인근 지점에 적용함으로서 지역빈도해석(Regional Frequency Analysis)을 실시하기 위한 기본 연구와 더불어, 확률홍수량 산정 등에도 GEV-A 분포를 이용함으로서 확률수문량을 산정의 불확실성을 줄이고 신뢰성을 높이는 연구를 진행하여야 한다.

감 사 의 글

본 연구는 서울시와 계약 체결한 “대학과 연계한 하천관리에 대한 연구용역” 및 서울대학교 SIR BK21 (안전하고 지속가능한 사회기반건설)사업단의 연구비 지원으로 수행되었으며, 이에 감사의 뜻을 표합니다.

참 고 문 헌

1. 김영오 (1999). “기후변화를 고려한 수자원관리 기법.” **대한토목학회논문집**, 대한토목학회, 제 47권, 제 5호, pp. 29-33.
2. 한국건설기술연구원 (2000). 1999년도 수자원관리기법개발연구조사 보고서, 제 1권 한국확률강우량도 작성. 건설교통부, pp. 87-90.
3. Coles, G.S. (2001). *An Introduction to Statistical Modeling of Extreme Values.* Springer, NewYork, Washington, D.C., pp. 208-220.
4. El Adlouni, S., Ouarda, T., Zhang, X., Roy, R., and Bobee, B. (2007). “Generalized maximum likelihood estimators for the nonstationary generalized extreme value model.” *Water Resources Research,* Vol. 43, W03410.

5. Lee, K.S. (2000). *Fundamentals of Water Resource Systems.* Saeron, Seoul, Korea, pp. 175-177.
6. Maidment, D. R. (1993). *Handbook of Hydrology.* McGraw-Hill, press, New-York, Washinton, D.C., pp. 18.1-18.35.
7. Smith, R.L. (1985). "Maximum likelihood estimation in a class of nonregular cases." *Biometrika*, Vol. 72, No. 1, pp. 67-90.