

TL-Moment 법을 이용한 Gumbel분포형에 관한 연구

A Study of Gumbel Distribution by TL-Moment Method

이정식*, 신창동**, 임성재***

Jung Sik Lee, Chang Dong Shin, Seong Jae Lim

요 지

본 연구에서는 실무에서 적용하고 있는 매개변수 산정방법인 L-moment 방법과 자료계열의 이상치를 고려하여 수정·제시된 TL-moment 방법을 비교·분석하여 빈도해석을 수행하는 경우, 보다 효율적인 매개변수 산정방법으로 수공구조물 설계 및 재난방계획에 합리적인 기준을 제시하고자 하였다. 산정방법의 비교를 위하여 서울특별시를 대상으로 강우자료를 수집하였으며, 국내 강우에 적합한 확률분포형으로 선정되어 있는 Gumbel 분포형에 적용하여 모멘트와 확률강우량을 산정하였다. 또한, 산정된 모멘트와 지속기간 25개, 재현기간 8개년의 확률강우량을 비교하여 타당성을 검토하였다. 본 연구의 수행으로 TL-moment 방법에 의한 모멘트의 변동폭이 L-moment 방법 보다 작게 산정되었으며, 산정된 확률강우량의 비교에서 이상치의 보정이 확인되었다.

핵심용어 : L-moment 방법, TL-moment 방법, Gumbel 분포형

1. 서론

현재 실무에서 확률강우량을 산정하는 방법으로 대상지점의 자료를 수집하고 일련의 통계적인 절차를 거쳐 빈도해석을 수행하는 방법인 지점빈도해석 기법이 이용되고 있다. 통계적인 절차는 강우자료의 독립성 및 경향성, 이상치 등을 검정하는 예비적 해석과 매개변수 산정을 통한 확률분포형의 적용과 적합도검정에 따른 최적분포형 선정과정을 의미한다. 매개변수 산정을 위하여 적용되는 방법으로는 모멘트법, 최우도법, 확률가중모멘트법이 적용되고 있다. 특히 자료를 오름차순으로 재배열하고 크기가 큰 순서로 가중치를 주어 매개변수를 산정하며, 자료가 충분하지 못한 경우의 매개변수 산정에 합리적인 방법으로 알려져 있는 확률가중모멘트의 선형조합인 L-moment 방법이 주로 이용되고 있다.

L-moment 방법은 차수가 다른 확률가중모멘트들을 선형으로 가중하여 편의가 제거된 모멘트를 구하는 방법이다. 산정된 표본추정량은 순차로 정리된 관측치의 선형조합이므로 일반모멘트 방법에서와 같이 2차나 3차 모멘트를 산정할 경우, 관측치의 제곱 및 세제곱이 불필요하므로 무차원화 된 변동계수나 왜곡도 계수의 추정치는 거의 편이되지 않게 산정된다. 그러나 수문계열자료의 빈도해석을 하는 경우, 이상치에 따른 자료편의의 발생으로 확률량이 왜곡되어 산정될 수 있으며, L-moment 방법의 경우에도 재현기간이 큰 경우에서 다른 방법에 비하여 확률량이 지나치게 높게 산정되는 경우가 있다. 이와 같이 이상치 발생에 따른 L-moment 방법의 왜곡을 수정하기 위하여 Elamir와 Seheult(2003)는 수정된 TL-moment 방법을 제시하였다. 기존 L-moment 방법은 4가지 식의 4개 변수로 구성되어 있는 반면, 수정된 TL-moment 방법

* 정희원 · 금오공과대학교 토목환경공학부 교수 · E-mail : jungsik@kumoh.ac.kr

** 정희원 · 금오공과대학교 건설기술연구소 연구원 · E-mail : sichado88@kumoh.ac.kr

*** 정희원 · 금오공과대학교 대학원 토목공학과 석사과정 · E-mail : tjdwojkrtd@hanmail.net

은 이상치 발생을 고려하여 4가지 식의 5개 변수로 구성되어 있다.

본 연구에서는 현재 실무에서 주로 적용하고 있는 매개변수 산정방법인 L-moment 방법과 자료계열의 이상치를 고려하여 수정·제시된 TL-moment 방법을 비교하였다. 산정방법의 비교를 위하여 서울특별시를 대상으로 강우자료를 수집하였으며, 국내 강우에 적합한 확률분포형으로 선정되어 있는 Gumbel분포형에 적용하여 모멘트와 확률강우량을 산정하였다. 또한, 방법별 산정된 모멘트와 확률강우량을 비교하여 본 연구의 타당성을 검토하였으며, 본 연구의 수행으로 빈도해석의 경우, 보다 효율적인 매개변수 산정방법을 통하여 수공구조물 설계 및 재난방재계획에 합리적인 기준을 제시할 것으로 판단된다.

2. 기본이론

2.1 Gumbel 분포형

Type-I 극치분포로 GEV분포형에서 $\alpha = 0$ 이며, GEV-1분포로도 알려져 있는 GUM 분포는 자료의 극치자료중에서도 최대치에 해당하는 연 최대홍수량 및 강우량자료의 분석에 많이 이용되고 있다. 국내에서도 건설교통부(2000)의 발표 이후로, 국내 강우의 빈도해석에서 널리 사용되고 있는 분포형이다. GUM 분포의 누가분포함수와 확률밀도함수는 식 (1)과 식 (2)와 같이 정의된다.

$$F(x) = \exp \left\{ - \exp \left[- \frac{x - x_0}{\alpha} \right] \right\} \quad (1)$$

$$f(x) = \frac{1}{\alpha} \exp \left\{ - \frac{(x - x_0)}{\alpha} - \exp \left[- \frac{(x - x_0)}{\alpha} \right] \right\}, \quad \infty < x < \infty \quad (2)$$

여기서 α 는 축척 모수, x_0 는 위치 모수이다.

2.2 L-moment 방법

L-moment 방법은 Hosking(1990)이 확률가중모멘트의 해석을 보다 용이하게 하기 위하여 개발하였으며, 선형조합의 형태로 이루어져 있다.

무작위 변량 X 가 누가분포함수 $F(x)$ 를 가질 때 확률가중모멘트는 식 (3)과 같이 나타낼 수 있다 (Maidment, 1992).

$$M_{ijk} = E[X^i F^j (1 - F)^k] = \int_0^1 X^i F^j (1 - F)^k dF \quad (3)$$

여기서 $E[\cdot]$ 는 기대값, $F(x)$ 는 X 의 누가분포함수(비초과확률), $1 - F(x)$ 는 초과확률, 그리고 i, j, k 는 양의 지수이다.

$M_{ijk} = M_{1r0}$ 로 하면, 확률가중모멘트는 식 (4)와 같다.

$$\beta_r = E\{X[F(X)]^r\} \quad (4)$$

또한, 분포형에 대한 β_r 의 불편의 표본추정치는 식 (5)로 계산할 수 있다.

$$b_r = \frac{1}{n} \sum_{j=1}^{n-r} \binom{n-j}{r} X_{(j)} = \frac{1}{(r+1)} \sum_{j=1}^{n-r} \frac{\binom{n-j}{r} X_{(j)}}{\binom{n}{r+1}} \quad (5)$$

여기서 β_r 은 모멘트 값, $X_{(j)}$ 는 크기순으로 큰 것부터 작은 것으로 재배열한 관측치이다.

따라서, L-모멘트는 식 (6)에서 식 (9)와 같이 나타낼 수 있다.

$$\lambda_1 = b_0 \quad (6)$$

$$\lambda_2 = 2b_1 - b_0 \quad (7)$$

$$\lambda_3 = 6b_2 - 6b_1 + b_0 \quad (8)$$

$$= 20b_3 - 30b_2 + 12b_1 - b_0 \quad (9)$$

여기서 λ_1 은 평균, λ_2 는 L-측척, $\tau_2 = \lambda_2/\lambda_1$ 는 L-변동계수, $\tau_3 = \lambda_3/\lambda_2$ 는 L-왜곡도계수, $\tau_4 = \lambda_4/\lambda_2$ 는 L-첨예도계수이다.

2.3 TL-moment 방법

TL-moment 방법은 빈도해석에 따른 최종 확률량이 이상치의 발생에 따라 많은 영향을 받으므로 표본 자료를 대상으로 L-moment 산정시 이상치에 대한 고려를 함으로서 자료의 왜곡 및 편의를 감소시키는 방법으로, Elamir와 Seheult(2003)가 제시한 방법이다. Asquith(2007)와 Hosking(2007)이 관련연구를 수행하였으며, TL-moment 식은 L-moment 식에 바탕을 두고 이상치 발생을 고려하여 다음과 같이 제시되었다.

$$\lambda_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} (Y_{r+1-k:r+2}) \quad (10)$$

여기서 λ_r 은 모멘트 값, $E(Y_{i:r})$ 은 다음과 같다.

$$E(Y_{i:r}) = \frac{r!}{(i-1)!(r-i)!} \int_0^1 x(F)^{i-1} (1-F)^{r-1} dF \quad (11)$$

$$\lambda_1 = E(X_{2:3}) = 6\beta_1 - 6\beta_2 \quad (12)$$

$$\lambda_2 = \frac{1}{2} E(X_{3:4} - X_{2:4}) = 6(-2\beta_3 + 3\beta_2 - \beta_1) \quad (13)$$

$$\lambda_3 = \frac{1}{3} E(X_{4:5} - 2X_{3:5} + X_{2:5}) = \frac{20}{3} (-5\beta_4 + 10\beta_3 - 6\beta_2 + \beta_1) \quad (14)$$

$$\lambda_4 = \frac{1}{4} E(X_{5:6} - 3X_{4:6} + X_{3:6} - X_{2:6}) = \frac{15}{2} (-14\beta_5 + 35\beta_4 - 30\beta_3 + 10\beta_2 - \beta_1) \quad (15)$$

여기서 $\tau_2 = \lambda_2/\lambda_1$ 는 TL-변동계수, $\tau_3 = \lambda_3/\lambda_2$ 는 TL-왜곡도계수, $\tau_4 = \lambda_4/\lambda_2$ 는 TL-첨예도계수이다.

3. 적용 및 분석

3.1 대상지역

강우량 기록년수는 일반적으로 30년 이상으로 하고 있으며, 본 연구에서는 국내의 대표적 도시인 서울지역을 대상으로 지속기간별 매년 최대치계열 강우자료를 이용하였다. 지속기간 10분 강우자료는 기상청의 기상연보를 이용하여 수집하였으며, 지속기간 1~24시간의 24개 지속기간 자료는 관측소에서 매시간 측정된 강우자료를 이용하였다. 수집된 강우자료의 기록연수는 1961~2011년까지 51년이며, 지속기간은 10분, 1~24시간까지의 25개 지속기간을 적용하였다.

3.2 모멘트 분석

확률밀도함수의 모멘트는 매개변수와 밀접한 관련을 가지고 있으며, 매개변수는 확률량 산정에 있어 반드시 필요한 자료이다. 원점에 대한 1차 모멘트는 평균, 2차 모멘트는 분산, 고차모멘트로 왜곡도와 첨예도가 있으며, 본 연구에서는 전술한 식들에서 이상치 발생에 따른 자료의 산포와 편의 정도를 나타내는 변동계수와 왜곡도계수를 표 1과 같이 비교하였다. 표 1에서는 전체지속기간 중, 지속기간 10분, 1, 6, 12, 18, 24시간의 변동계수 및 왜곡도계수를 나타내었으며, 각각 표본자료, L-moment, TL-moment 방법을 적용한 결과이다. 표 1을 살펴보면, TL-moment 방법의 변동계수 범위는 0.09734~0.14307, 왜곡도계수 범위는 0.01986~0.22714로서 다른 방법에 비하여 변동폭이 작은 것으로 나타났다. 따라서 TL-moment 방법을 적용하였을 경우, L-moment 방법에 비하여 상대적으로 작은 산포와 편의가 발생하는 것으로 판단된다.

표 1. 변동계수 및 왜곡도계수 비교

지속기간	변동계수(CV)			왜곡도계수(SKEW)		
	CV	L-CV	TL-CV	skew	L-skew	TL-skew
10분	0.37115	0.20665	0.11437	0.51819	0.20165	0.17632
1hr	0.36005	0.19584	0.09734	0.94047	0.10737	0.02879
6hr	0.39976	0.21918	0.11752	1.04536	0.21154	0.13972
12hr	0.38110	0.21291	0.11960	0.78380	0.20338	0.17710
18hr	0.44767	0.24209	0.13222	1.16561	0.26754	0.20315
24hr	0.48817	0.25973	0.14231	1.30996	0.29907	0.22714
최 소	0.36005(1hr)	0.19584(1hr)	0.09734(1hr)	0.51819(10분)	0.08233(2hr)	0.01986(2hr)
최 대	0.48817(24hr)	0.25973(24hr)	0.14307(23hr)	1.32770(23hr)	0.29907(24hr)	0.22714(24hr)

3.3 확률강우량 분석

본 절에서는 전술한 식들을 이용하여 L-moment 방법과 TL-moment 방법에 의하여 확률강우량을 산정하였으며, 5, 10, 20, 30, 50, 70, 100, 200년의 8개 재현기간과 10분 1, 6, 12, 18, 24시간의 6개 지속기간 확률강우량을 나타내면 표 2와 같다.

표 2. 확률강우량 비교

지속기간	5년		10년		20년		30년		50년		70년		100년		200년	
	L-mo	TL-mo	L-mo	TL-mo	L-mo	TL-mo	L-mo	TL-mo	L-mo	TL-mo	L-mo	TL-mo	L-mo	TL-mo	L-mo	TL-mo
10분	21.55	21.72	24.68	24.77	27.68	27.69	29.41	29.37	31.57	31.47	32.98	32.85	34.48	34.30	37.38	37.13
1hr	65.13	65.08	76.09	75.53	86.59	85.56	92.64	91.33	100.19	98.54	105.15	103.27	110.38	108.27	120.54	117.96
6hr	155.63	154.66	184.22	183.34	211.64	210.84	227.42	226.67	247.14	246.45	260.07	259.42	273.74	273.13	300.24	299.71
12hr	192.44	193.98	227.00	230.41	260.16	265.36	279.23	285.47	303.08	310.60	318.71	327.07	335.23	344.49	367.28	378.27
18hr	222.51	218.67	266.61	262.86	308.92	305.25	333.26	329.63	363.68	360.12	383.63	380.10	404.72	401.24	445.60	442.20
24hr	244.76	238.06	295.90	288.76	344.96	337.39	373.18	365.36	408.46	400.34	431.59	423.26	456.04	447.51	503.45	494.51

주) L-mo : L-moment 방법, TL-mo : TL-moment 방법

표 2를 살펴보면, 지속기간 10분의 경우 재현기간 5~20년, 지속기간 12시간의 경우 전체 재현기간에서 TL-moment 방법에 의한 확률강우량이, 기타 지속기간의 경우에는 L-moment 방법에 의한 확률강우량이 크게 산정되었다. 전체 강우지속기간인 10분, 1~24시간과 8개 재현기간의 경우에는 10분, 2, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17시간 지속기간에서 TL-moment 방법에 의한 확률강우량이 크게 산정되었다. 즉 산정된 전체 200개 확률강우량 중에서 87개 확률강우량에서 TL-moment 방법에 의한 값이 크게 산정되었다.

따라서 서울지점의 강우자료에서 TL-moment 방법에 의하여 산정된 확률강우량의 변동이 발생하여 이상치 자료의 보정을 확인할 수 있었으며, 추후 대상지점 및 확률분포형의 추가 검토에 따라 TL-moment 방법의 활용성을 증대 시킬수 있으리라 판단된다.

4. 결론

본 연구에서는 실무에서 적용하고 있는 매개변수 산정방법인 L-moment 방법과 자료계열의 이상치를 고

려하여 수정·제시된 TL-moment 방법을 비교·분석하여 빈도해석을 수행하는 경우, 보다 효율적인 매개 변수 산정방법으로 수공구조물 설계 및 재난방재계획에 합리적인 기준을 제시하고자 하였다. 산정방법의 비교를 위하여 서울특별시를 대상으로 강우자료를 수집하였으며, 국내 강우에 적합한 확률분포형으로 선정되어 있는 Gumbel 분포형에 적용하여 모멘트와 확률강우량을 산정하였다. 또한, 산정된 모멘트와 지속기간 25개, 재현기간 8개년의 확률강우량을 비교하여 타당성을 검토하였다. 본 연구의 수행으로 얻어진 결과를 요약하면 다음과 같다.

- (1) 자료의 산포정도를 나타내는 변동계수를 L-moment 방법 및 TL-moment 방법과 비교하였을 경우, TL-moment 방법에 의하여 산정된 변동계수의 변동폭이 0.09734~0.14307로서 제일 작게 발생하였다.
- (2) 자료의 왜곡정도를 나타내는 왜곡도계수를 L-moment 방법 및 TL-moment 방법과 비교하였을 경우, TL-moment 방법에 의하여 산정된 왜곡도계수의 변동폭이 0.01986~0.22714로서 제일 작게 발생하였다.
- (3) 확률강우량 비교에서는 산정된 200개의 강우량 중에서 87개의 확률강우량이 크게 산정되어, TL-moment 방법에 의한 이상치 보정이 확인되었으며, 추후 지점 및 자료의 보완이 수행되면 활용성을 증대시킬 수 있을 것으로 판단된다.

감 사 의 글

본 연구는 국토해양부가 출연하고 한국건설교통기술평가원에서 위탁 시행한 건설기술혁신사업(08기술혁신 F01)에 의한 차세대홍수방어기술개발 연구단의 연구비 지원에 의해 수행되었습니다.

참 고 문 헌

1. 건설교통부(2000), “한국 확률강우량도 작성”, 1999년도 수자원관리기법개발연구조사 보고서, 제 1권.
2. Asquith W.H.(2007). L-Moments and TL-Moments of The Generalized Lambda Distribution. Computational statistics & Data Analysis, 51, pp.4484-4496.
3. Elamir E.A.H. & Seheult A.H.(2003). Trimmed l-Moments. Computational statistics & Data Analysis, 43, pp.299-314.
4. Hosking J.R.M.(1990). L-moments-Analysis and estimation of distribution using linear combinations of order statistics, Journal Royal Statistical Society B, v.52, no.1, pp.105~124.
5. Hosking J.R.M.(2007). Some Theory and Practical Uses of Trimmed L-Moments. Journal of Statistical planning and Inference, 137, pp.3024-3039.
6. Maidment, D.R.(1992), Handbook of Hydrology, McGraw-Hill.