

Generalized extreme value, Generalized logistic 및 Generalized pareto 분포 모형에 의한 홍수빈도분석

이순혁*, ○맹승진**, 류경식***, 지호근***

I. 서론

과거 국내·외에서 확률밀도함수의 매개변수 유도방법으로 많이 사용되어 왔던 방법으로는 모멘트법(method of moments), 최우법(method of maximum likelihood) 등이 있으나 모멘트법은 왜곡된 분포의 경우 불확실한 매개변수 추정치를 갖게되며 고차모멘트로 갈수록 더욱 부정확하게 되며 최우법은 수치적으로 해를 구하는 과정이 복잡하고 매개변수 추정치를 얻을 수 없는 경우가 많은 단점이 있다. 그래서 최근에는 모멘트법과 최우법의 단점을 보완할 수 있는 대안으로 Greenwood 등은 매개변수 추정법으로 확률가중모멘트법을 제시하였다. 또한, Hosking은 확률가중 모멘트의 선형조합(linear combination)으로 이루어진 L-모멘트법을 제시하므로써 정확성이 높은 확률밀도함수의 매개변수 유도를 가능하게 하였다.

따라서, 본 연구에서는 홍수유출빈도분석을 위해 홍수자료의 동질성, 독립성 및 Outlier 검정과 K-S 검정 및 L-모멘트비도에 의한 Generalized extreme value(GEV), Generalized logistic(GLO) 및 Generalized pareto(GPA) 분포모형의 적합성을 판정한다. 적정분포모형의 매개변수를 L-모멘트법에 의해 추정하고 이에 의해 산정된 빈도별 설계홍수량과 Monte Carlo 기법에 의해 모의발생된 홍수자료를 이용한 빈도별 설계홍수량과의 상대평균제곱근오차(relative root-mean-square error, RRMSE)와 상대편의(relative bias, RBLAS)에 의한 오차분석을 수행함으로서 적정한 설계홍수량을 제시하고자 한다.

II. 기본수문자료

본 분석에 사용된 대상유역은 한강수계의 여주, 낙동강수계의 산양, 월포, 금강수계의 공주, 영산강수계의 남평, 나주 및 섬진강수계의 암록의 7개유역을 선정하였으며 설계홍수량 유도를 위해 각 유역의 연최고치 자료를 사용하였다. 각 유역에 대한 유역 특성의 결과에서 유역면적은 215.0 ~ 11,036.0 km²의 범위를 나타내고 있고 하천연장은 36.70 ~ 347.10 km에 이르며 유역평균 폭은 5.86 ~ 31.79 km이고 형상계수는 0.081 ~ 0.365의 범위를 나타내고 있다.

III. 결과 및 고찰

1. 사용수문자료의 기본통계치 분석

* 충북대학교 농과대학 농공학과 교수

** 한국수자원공사 수자원연구소 연구원

*** 충북대학교 농과대학 농공학과 박사과정

여주, 산양, 월포, 공주, 남평, 나주 및 압록유역에 대한 평균치, 표준편차, 왜곡도계수, 변동계수 및 첨예도의 기본통계치를 각 유역별로 구하였다. 전 유역에 걸쳐 평균치와 표준편차는 각각 170.656 ~ 4261.024 및 84.805 ~ 3235.191의 범위를, 왜곡도계수와 변동계수는 각각 0.355 ~ 1.741 및 0.350 ~ 0.759의 범위를 그리고 첨예도는 2.101 ~ 6.119의 범위를 나타내었다.

2. 연최대홍수량의 독립성, 동질성 검정 및 Outlier 검정

대상유역별 연최대홍수량의 독립성 및 동질성 검정을 각각 Wald-Wolfowitz Test 및 Mann-Whitney Test 검정법에 의해 실시하였으며, 본 분석에서 적용된 각 유역의 연최대홍수량자료에 대한 독립성 및 동질성 검정에 대한 결과에서 대상유역의 연최고치 자료에 관한 독립성과 동질성이 인정되었다.

일련의 수문자료계열에서 자료의 일반적인 균형분포의 상태에서 훨씬 위나 또는 아래로 격리되는 Data Point를 나타내는 Outlier의 유발로 인해 부적절한 통계학적 매개변수의 유도로서 설계수문량의 제시에 불확실성을 초래하게 된다. 이에 본 분석에서는 Grubbs-Beck에 의한 검정법을 사용하여 Outlier의 유무를 검정하였다.

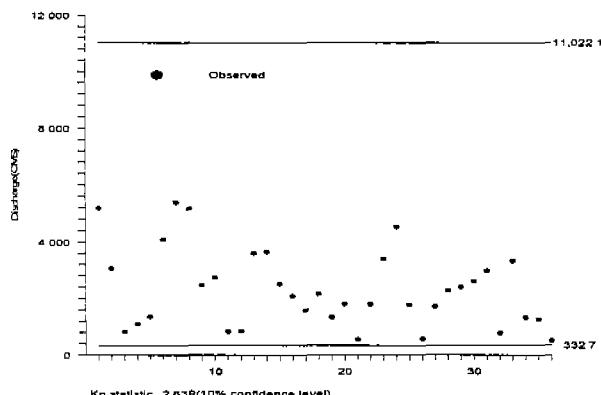


Fig. 1. G-B test for detection of high and low outliers at Abrog watershed of Seomjin river.

본 분석에 적용된 유역의 연최대홍수량자료에 대한 검정결과를 보면 표본 예로 Fig. 1에서 보는 바와 같이 전 유역에 걸쳐 High Outlier 및 Low Outlier가 없음이 확인되었다.

이상의 결과에서 보건데 독립성, 동질성 및 Outlier 검정 결과 대상유역의 연최고치 자료에 관한 독립성과 동질성이 인정됨과 동시에 연최고치 자료의 균형분포에서 상하로 격리되는 Outlier가 없음이 검정되므로서 수문자료로서 분석 타당성이 인정되었다.

3. L-모멘트비에 의한 L-변동계수, L-왜곡도 및 L-첨예도 계산

화률가중모멘트의 선형조합으로서 L-모멘트값을 각 유역별로 구하고 L-모멘트값의 비인 L-변동계수, L-왜곡도 및 L-첨예도를 구하였다. 전 유역에 걸쳐 L-모멘트비인 L-변동계수, L-왜곡도 및 L-첨예도는 각각 0.1971 ~ 0.3883, 0.0990 ~ 0.3149 및 0.0734 ~ 0.2154의 범위를 나타내고 있다.

4. 적합도 검정

가. Kolmogorov-Smirnov(K-S) 검정에 의한 적합도 검정

L-모멘트법에 따른 GEV, GLO 및 GPA 분포모형에 대한 적합도 판정은 K-S 검정을 이용하였으며 각 유역에 대한 결과를 보면 대상유역 전반에 걸쳐 GEV, GLO 및 GPA 분포모형에 대한 적합도가 인정되었다.

나. L-모멘트비도에 의한 적용 분포모형의 적합성 판별

7개 유역의 연최대홍수량에 대한 적정 확률분포모형의 선정은 L-모멘트비도를 도시함으로써 이루어질 수 있다. 따라서 본 분석에서 사용된 GEV, GLO 및 GPA 분포모형의 L-왜곡도와 L-첨예도의 관계인 L-모멘트비도를 도시하고, 각 유역별 연최대홍수량의 L-왜곡도와 L-첨예도를 구하여 L-모멘트비도를 작성하고 이를 근거로 적정한 확률분포모형을 선정한다.

각 유역별 연최대홍수량의 L-모멘트비인 L-왜곡도, L-첨예도를 L-모멘트비도에 플로팅한 결과는 Fig. 2와 같다.

Fig. 2에 나타난 바와 같이 유역별 연최대홍수량의 L-왜곡도와 L-첨예도의 무차원 L-모멘트 플로팅 포지션점이 GEV와 GPA 분포모형의 곡선의 사이에 고루 산재해 있고 L-왜곡도를 평균한 값이 0.21로서 GLO 확률분포모형의 L-모멘트비 곡선 보다 GEV와 GPA 분포모형의 L-모멘트비 곡선 사이에 근접되어 있음을 볼 수 있다.

결론적으로 유역별 연최대홍수량에 L-모멘트법을 적용하여 L-모멘트비를 플로팅한 결과 GEV 와 GPA 분포모형이 GLO 분포모형 보다 적합한 것으로 나타났다.

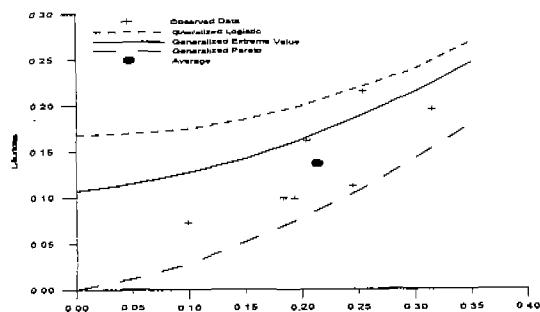


Fig. 2. L-skewness versus L-kurtosis for 7 sites in GEV, GPA and GLO distributions

이상의 결과 L-모멘트비도에 의한 검정에서는 적용된 3개 분포모형 중 GEV와 GPA 분포모형이 GLO 분포모형에 비해 적정한 것으로 검정되었고 K-S 검정에서는 적용된 3개 확률분포모형이 모두 적정한 것으로 검정된 사실에 비추어 비매개변수적 방법인 K-S 검정보다는 객관적인 검증방법에 의해 도출된 L-모멘트비도에 의한 GEV 및 GPA 분포모형을 본 분석에서는 최종적으로 선정하였다.

5. 적정 분포모형인 GEV 및 GPA 분포모형의 매개변수 추정

적정 확률분포모형으로 선정된 GEV 및 GPA 분포모형의 매개변수인 척도매개변수, α 위치매개변수, ξ 및 형상매개변수, k 를 유역별로 L-모멘트법에 의해 구하였으며 그 결과는 Table-1과 같다.

6. 모의 발생 및 상대평균제곱근오차(RRMSE)와 상대편의(RBIAS)에 의한 적정 설계홍수량 선정

각 유역에 대하여 Monte Carlo 알고리즘에 의해 표본크기 20, 40, 60, 100개 별로 1,000회 모의 발생된 분포모형별 연최대홍수량을 사용하여 추정된 빈도별 설계홍수량과 실측자료에 의한 빈도별 설계홍수량과의 오차분석을 상대평균제곱근오차와 상대편의에 의해 수행함으로서 각 유역에 따른 빈도별 적정 설계홍수량을 선정할 수 있게 된다.

따라서 본 분석에서는 각 유역의 연최대홍수량과 모의 발생된 유역들의 연최대홍수량이 GEV와 GPA 분포모형에 의해 각각의 빈도별 설계홍수량이 추정되었고, 실측치와 모의발생 자료에 의해 추정된 각각의 빈도별 설계홍수량을 대상으로 상대평균제곱근오차와 상대편의에 의한 오차분석을 실시하여 GEV와 GPA 분포모형에 의한 빈도별 설계홍수량간의 적정성 우열에 관한 비교분석을 수행하고자 하였다.

상대평균제곱근오차와 상대편의의 일반식은 각각 다음의 식 (1) 및 식 (2)와 같다.

$$RRMSE = \sqrt{\frac{1}{M} \sum_{m=1}^M \left\{ \frac{Q_j^m - Q_j}{Q_j} \right\}^2} \quad (1)$$

$$RBIAS = \frac{1}{M} \sum_{m=1}^M \left\{ \frac{Q_j^m - Q_j}{Q_j} \right\} \quad (2)$$

여기에서 Q_j 와 Q_j^m 은 각각 실측치에 의한 빈도별 설계홍수량과 모의발생된 자료에 의한 빈도별 설계홍수량이다.

유역별로 구한 실측치의 빈도별 설계홍수량과 표본의 크기별로 구한 모의발생된 빈도별 설계홍수량의 상대평균제곱근오차 및 상대편의를 구하였으며, 표본 예로 실측치에 의한 GEV와 GPA 분포모형의 빈도별 설계홍수량과 표본의 크기를 20과 100으로 하여 모의발생된 빈도별 설계홍수량과의 상대평균제곱근오차를 나타낸 결과는 각각 Table-2와 같고 이들을 표본 예로 암록유역에 대하여 도시한 결과는 Fig 3과 같다.

Table-2에서 보는 바와 같이 상대평균제곱근오차는 각 지역별로 재현기간이 커짐에 따라 증가한 반면 모의발생의 표본크기가 증가함에 따라 감소되는 양상을 나타내었다.

또한, Fig. 3에서의 결과를 보면 GEV 분포모형에 의한 상대평균제곱근오차가 GPA 분포모형에 의한 것보다 크게 나타났으며, 이러한 결과는 표본의 크기가 작을수록 심화되는 경향을 나타내었다.

최종적으로 이상의 결과를 종합해 볼 때 GPA 분포모형에 의해 유도된 빈도별 설계홍수량이 GEV 분포모형에 의한 빈도별 설계홍수량 보다 신빙도가 높음을 보여 주었다.

Table-1. Parameters calculated by GEV and GPA distributions and using method of L-moments

Distribution	River	Station	ξ	α	k
GEV	Han	Yeoju	2680.2468	1876.9049	-0.2137
	Nagdong	Sanyang	129.2254	67.5063	-0.0357
		Wolpo	397.8398	239.2314	-0.1264
	Geum	Kongju	1942.9361	1195.5068	-0.1126
	Yeongsan	Nampyeong	300.9603	98.0918	-0.0522
		Naju	1153.2218	714.2199	0.1135
GPA	Seomjin	Abrog	1661.7627	1104.6227	-0.0207
	Han	Yeoju	881.9523	3521.4979	0.0421
	Nagdong	Sanyang	56.7541	154.0733	0.3527
		Wolpo	155.7452	492.9456	0.1904
	Geum	Kongju	722.4136	2501.0952	0.2145
	Yeongsan	Nampyeong	196.7993	219.6678	0.3225
		Naju	303.8329	1949.3746	0.6397
	Seomjin	Abrog	463.9970	2565.2984	0.3803

GEV : Generalized extreme-value distribution

GPA : Generalized pareto distribution

7. 유역별 적정 설계홍수량의 유도

GEV 및 GPA 분포모형의 L-모멘트법에 따른 재현기간별 설계홍수량을 유역별로 구한 결과는 Table-3과 같고 GPA 분포모형의 적합도 검정을 위하여 L-모멘트법에 의해 유도된 GEV 및

GPA 분포모형에 의한 설계홍수량을 플로팅 포지션법인 Weibull 식에 의해 비초과확률을 구하고 이를 정규확률지상에 도시하였으며 표본 예로 섬진강 수계의 압록에 대해 정규확률지에 도시한 결과는 Fig. 4와 같다.

Table-2. RRMSE of the simulated quantile following the sample size of GEV and GPA distributions

Distribution	Sample size	River	Station	Q_5	Q_{10}	Q_{20}	Q_{50}	Q_{100}	Q_{200}
GEV	20	Han	Yeoju	0.1762	0.2073	0.2589	0.3559	0.4531	0.3754
		Nagdong	Sanyang	0.1212	0.1378	0.1664	0.2214	0.2759	0.3426
		Geum	Wolpo	0.1459	0.1698	0.2056	0.2741	0.3418	0.4252
		Yeongsan	Kongju	0.1461	0.1599	0.1867	0.2395	0.2915	0.3543
		Seomjin	Nampyeong	0.0969	0.1158	0.1441	0.1961	0.2464	0.3071
	100	Han	Naju	0.1168	0.1213	0.1388	0.1777	0.2164	0.2622
		Nagdong	Abrog	0.1347	0.1472	0.1752	0.2322	0.2896	0.3603
		Geum	Yeoju	0.0787	0.0926	0.1169	0.1606	0.2006	0.2463
		Yeongsan	Sanyang	0.0549	0.0609	0.0732	0.0966	0.1181	0.1421
		Seomjin	Wolpo	0.0660	0.0774	0.0970	0.1325	0.1653	0.2029
GPA	20	Han	Kongju	0.0664	0.0776	0.0963	0.1295	0.1596	0.1933
		Nagdong	Nampyeong	0.0440	0.0520	0.0655	0.0903	0.1135	0.1398
		Geum	Naju	0.0524	0.0547	0.0625	0.0783	0.0927	0.1084
		Yeongsan	Abrog	0.0615	0.0680	0.0810	0.1049	0.1266	0.1506
		Seomjin	Yeoju	0.1899	0.1942	0.2167	0.2720	0.3310	0.4049
	100	Han	Sanyang	0.1288	0.1225	0.1293	0.1576	0.1893	0.2272
		Nagdong	Wolpo	0.1610	0.1558	0.1651	0.1988	0.2374	0.2857
		Geum	Kongju	0.1598	0.1508	0.1557	0.1848	0.2209	0.2672
		Yeongsan	Nampyeong	0.1045	0.1045	0.1124	0.1370	0.1638	0.1959
		Seomjin	Naju	0.1215	0.1007	0.0992	0.1214	0.1471	0.1752

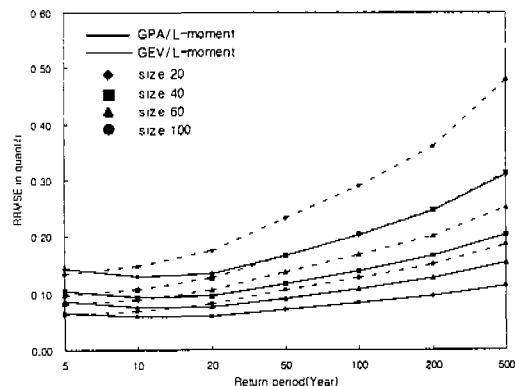


Fig. 3 Comparison of RRMSE estimated by GEV and GPA distribution using method of L-moments at Abrog watershed of Seomjin river.

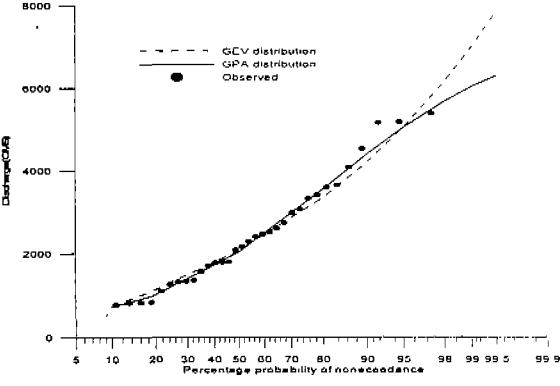


Fig. 4. Comparison of design floods estimated by GEV and GPA distribution using method of L-moments with Weibull plotting position at Abrog watershed of Seomjin river

Table-3. Comparison of design floods calculated by GEV and GPA distributions (Unit : m^3/s)

Distribution	River	Station	Return period(yrs)					
			5	10	20	50	100	200
GEV	Han	Yeoju	5,999.04	8,104.05	10,466.52	14,117.14	17,370.87	21,133.37
	Nagdong	Sanyang	233.24	287.41	340.75	411.86	466.72	522.75
		Wolpo	792.93	1,020.54	1,260.11	1,604.40	1,890.30	2,201.39
	Geum	Kongju	3,896.45	5,004.83	6,159.73	7,800.61	9,148.12	10,600.32
	Yeongsan	Nampyeong	454.01	535.19	616.12	725.49	811.00	899.36
		Naju	2,138.25	2,571.56	2,953.91	3,404.61	3,712.43	3,995.79
GPA	Seomjin	Abrog	3,344.65	4,206.45	5,045.79	6,151.00	6,993.30	7,844.73
	Han	Yeoju	6,361.62	8,609.45	10,792.56	13,582.29	15,622.22	17,603.40
	Nagdong	Sanyang	245.97	299.68	341.74	383.68	407.52	426.19
		Wolpo	839.05	1,074.65	1,281.11	1,515.39	1,667.35	1,800.52
	Geum	Kongju	4,126.47	5,267.10	6,250.15	7,344.41	8,040.48	8,640.39
	Yeongsan	Nampyeong	472.61	553.81	618.74	685.07	723.71	754.62
		Naju	2,262.78	2,652.61	2,902.82	3,101.70	3,191.07	3,248.43
	Seomjin	Abrog	3,551.79	4,399.27	5,050.35	5,685.47	6,038.48	6,309.68

표본 예인 Fig. 4의 결과와 타 대상유역을 종합해 보건데 GPA 분포모형에 의해 유도된 설계홍수량이 GEV 분포모형에 의한 것 보다 좋은 결과를 보여 주었다.

IV 결론

본 연구는 한강, 낙동강, 금강, 영산강 및 섬진강 수계의 7개 유역에 대한 연최대홍수량을 사용하여 이들 자료에 대한 신뢰도를 판정하고, L-모멘트비도에 의한 L-변동계수, L-웨곡도 및 L-첨예도를 구한 후 L-모멘트비도와 K-S검정에 의해 GEV, GLO 및 GPA 분포모형의 적합도를 검정하였으며 L-모멘트법에 의해 적정 분포모형으로 선정된 GEV 및 GPA 분포모형의 매개변수를 구하였고, L-모멘트법에 따른 설계홍수량을 유도하였다. 또한, 각 유역에 대하여 Monte Carlo 알고리즘에 의해 표본크기 20, 40, 60, 100개 별로 1,000회 모의발생된 분포모형별 연최대홍수량을 사용하여 추정된 빈도별 설계홍수량과 실측자료에 의한 빈도별 설계홍수량과의 오차분석을 상대평균제곱근오차와 상대편의에 의해 수행하였으며 이에 대한 결과를 요약하면 다음과 같다.

1. 본 연구에 사용된 홍수량 자료에 관한 Wald-Wolfowitz의 독립성검정과 Mann-Whitney의 동질성 검정 및 Grubbs-Beck의 Outlier검정을 수행한 결과 이의 분석타당성이 인정되었다.
2. GEV, GLO 및 GPA 분포모형의 적합성을 판별하기 위해 K-S 검정 및 L-모멘트비도의 분석을 도시한 결과 적정 분포모형으로 GEV 및 GPA 분포모형이 선정되었다.
3. GEV 및 GPA 분포모형의 L-모멘트법에 따른 매개변수를 각각 유도하였다.
4. 각 유역에 대하여 Monte Carlo 알고리즘에 의해 표본크기 20, 40, 60, 100개 별로 1,000회 모의발생된 분포모형별 연최대홍수량을 사용하여 추정된 빈도별 설계홍수량과 실측자료에 의한 빈도별 설계홍수량과의 오차분석을 상대평균제곱근오차와 상대편의에 의해 수행한 결과에서 GPA 분포모형에 의해 유도된 빈도별 설계홍수량이 GEV 분포모형에 의한 빈도별 설계홍수량 보다 신빙도가 높음을 보여 주었다.
5. 유역별 재현기간에 따른 적정 설계홍수량을 GPA 분포모형의 L-모멘트법에 의해 산정하였다.