Matematika, 2002, Jilid 18, bil. 1, hlm. 33–43

Comparisons of the LH Moments and the L Moments

Ani Shabri
Department of Mathematics
Universiti Teknologi Malaysia,
81310 UTM Skudai, Johor, Malaysia

Abstract This paper discusses comparisons of the LH moments method with L moments method. LH moments, a generalization of L moments, based on linear combinations of higher-order statistics was introduced for charactering the upper part of distributions and larger events in data by Wang (1997). Analysis of observed data shows that using LH moments estimates of the upper part of distribution events are expected to be more reasonable than the L moments estimates. A comparison of the LH moment diagram and the L moment diagram of the data also shows that the GEV distribution describes the LH moment ratios better than the L moments.

Keywords LH moments, GEV, PWM, Linear Combination, Higher-Order Statistics

Abstrak *Kertas ini membincangkan perbandingan antara kaedah LH momen dengan kaedah L momen. LH momen adalah L-momen umum yang berasaskan kepada gabungan linear bagi statistik tertib tinggi digunakan untuk dipadankan dengan ciri-ciri hujung taburan dan data ekstrim. Analisis terhadap data menggunakan penganggar LH momen bagi hujung atas taburan didapati LH lebih baik berbanding penganggar L momen. Perbandingan bagi gambarajah LH momen dan gambarajah L momen ke atas data juga menunjukkan bahawa taburan GEV dapat dipadankan dengan baik dengan nisbah LH momen berbanding dengan nisbah L momen.*

Katakunci *LH Moments, GEV, PWM, Gabungan Linear, Statistik Peringkat-Tinggi*

1 Introduction

Floods are extreme events, which may kill people and destroy property. The probability that such floods occur is an essential input for the design of hydraulic structures and also for estimating the risks involved. Statistical analysis of extreme events are often conducted for predicting large return period events. The relationship between flood magnitude and its return period is of great importance to avoid high damages. However, estimating floods of large return periods is difficult because extreme events are by definition rare and the relevant data record is often short. Regional frequency using conventional moments [5], [8] and L-Moment analysis can solve this problem. However conventional moments have been criticized for being sensitive to the upper part of the distributions and thus sample outliers [12]. L moments were found as an alternative to product moments for characterizing distributions and data [3]. Since 1990, many researchers have investigated application of L moments to analyzing extremes. Examples of works L moments can be found in the works by Ani [1], Al-Khudhairy [2], Hosking [3], Hosking and Wallis [4] and Vogel et al. [10,11]. LH moments are generalization of L moments have been introduced to characterize the upper part of distributions and larger events in data [12, 13]. Wang [12, 13] found the method of LH moments provides an analytical means of fitting a distribution to the larger sample events without explicit sample censoring. In this paper, L moments and LH moments are used for characterizing the upper part of distributions and larger events in data. LH moments and L moments are compared by using real data. LH moments and L moments diagrams are constructed using annual maximum flow data from 31 streams in Malaysia to examine the appropriateness of the General Extreme Value (GEV) distribution in modeling the extreme events.

2 General Extreme Value Distribution

The most widely accepted distribution for describing flood frequency data in U.K.[7], Australia and United Stated [9, 10] is the GEV. It was introduced by Jenkinson in 1955, and recommended by the Natural Environmental Research Council [7]. The cumulative distribution function of the GEV can be written

$$F(x) = \exp\left\{-\left[1-\frac{k}{\alpha}(x-\xi)\right]^{1/k}\right\} \qquad \text{if} \qquad k \neq 0 \tag{1}$$

$$= \exp\left\{-\exp\left[-\frac{1}{\alpha}(x-\xi)\right]\right\} \qquad \text{if} \qquad k = 0 \tag{2}$$

or in inverse form

$$x(F) = \xi + \frac{\alpha}{k}\left[1-(-\ln F)^{-k}\right] \qquad \text{for} \qquad k \neq 0 \tag{3}$$

$$x(F) = \xi - \alpha \ln - \ln F \qquad \text{for} \qquad k = 0 \tag{4}$$

where ξ is a location parameter, α is a scale parameter and k is shape parameter. For $k = 0$, the distribution is unbounded both above and below, and it referred to as the extreme value Type I (EV1) or Gumbel distribution. For $k < 0$, the distribution has a finite lower bound at $\xi+\alpha/k$ and a thick upper tail, and it referred to as the Type II extreme value distribution. For $k > 0$, the distribution has a finite upper bound at $\xi + \alpha/k$, and is referred to as the Type III extreme value distribution.

Normalized Probability Weighted Moments (PWM) of the GEV distribution are given for $k \neq 0$

$$B_r = \xi + \frac{\alpha}{k}\left[1 - \Gamma(1+k)(r+1)^{-k}\right] \quad (5)$$

and for $k = 0$

$$B_r = \xi + \alpha\left[\epsilon + \ln(r+1)\right] \quad (6)$$

where $\Gamma()$ is Gamma function and $\epsilon = 0.5772$ is the Euler constant [12].

3 Higher Probability Weighted Moments and LH moments

The LH moments are based on linear combination of higher probability weighted moments was introduced by Wang 1997 [12]. For the first four LH moments are defined as

$$\lambda_1^\eta = \mathrm{E}\left[X_{(\eta+1):(\eta+1)}\right] \quad (7)$$

$$\lambda_2^\eta = \frac{1}{2}\mathrm{E}\left[X_{(\eta+2):(\eta+2)} - X_{(\eta+1):(\eta+2)}\right] \quad (8)$$

$$\lambda_3^\eta = \frac{1}{3}\mathrm{E}\left[X_{(\eta+3):(\eta+3)} - 2X_{(\eta+2):(\eta+3)} + X_{(\eta+1):(\eta+3)}\right] \quad (9)$$

$$\lambda_4^\eta = \frac{1}{4}\mathrm{E}[X_{(\eta+4):(\eta+4)} - 3X_{(\eta+3):(\eta+4)} + 3X_{(\eta+2):(\eta+4)} - X_{(\eta+2):(\eta+4)}] \quad (10)$$

where

$$\mathrm{E}\left[X_{j:m}\right] = \frac{m!}{(j-1)!(m-j)!}\int_0^1 x(F)F^{j-1}dF$$

When $\eta = 0$, LH moments become identical to L moments. As η increases LH moments reflect more and more the characteristics of the upper part of distributions and larger event in data. LH moments are called L1 moments, L2 moments, ... for $\eta = 1, 2$, respectively. LH coefficient of variation, skewness, and kurtosis are

$$\tau_2^\eta = \frac{\lambda_2^\eta}{\lambda_1^\eta} \quad (11)$$

$$\tau_3^\eta = \frac{\lambda_3^\eta}{\lambda_2^\eta} \quad (12)$$

$$\tau_4^\eta = \frac{\lambda_4^\eta}{\lambda_2^\eta} \quad (13)$$

respectively. Given a ranked sample $x_1 \leq x_2 \leq \ldots \leq x_n$, the unbiased estimator of LH moments are given by

$$\hat{\lambda}_1^\eta = \frac{1}{\binom{n}{\eta+1}} \sum_{i=1}^{n} \binom{i-1}{\eta} x_i \quad (14)$$

$$\hat{\lambda}_2^\eta = \frac{1}{2\binom{n}{\eta+2}} \sum_{i=1}^{n} \left\{ \binom{i-1}{\eta+1} - \binom{i-1}{\eta}\binom{n-1}{1} \right\} x_i \quad (15)$$

$$\hat{\lambda}_3^\eta = \frac{1}{3\binom{n}{\eta+3}} \sum_{i=1}^{n} \left\{ \binom{i-1}{\eta+2} - 2\binom{i-1}{\eta+1}\binom{n-1}{1} + \binom{i-1}{\eta}\binom{n-i}{2} \right\} x_i \quad (16)$$

$$\hat{\lambda}_4^\eta = \frac{1}{4\binom{n}{\eta+4}} \sum_{i=1}^{n} \left\{ \binom{i-1}{\eta+3} - 3\binom{i-1}{\eta+2}\binom{n-i}{1} + 3\binom{i-1}{\eta+1}\binom{n-i}{2} - \binom{i-1}{\eta}\binom{n-i}{3} \right\} x_i \quad (17)$$

where

$$\binom{m}{j} = \frac{m!}{j!(m-j)!} \quad (18)$$

is equal to zero when $j > m$. The first four LH moments for GEV distribution are given for $k \neq 0$

$$\lambda_1^\eta = \xi + \frac{\alpha}{k}\left[1 - \Gamma(1+k)(\eta+1)^{-k}\right] \quad (19)$$

$$\lambda_2^\eta = \frac{(\eta+2)\alpha\Gamma(1+k)}{2!k}\left[-(\eta+2)^{-k} + (n+1)^{-k}\right] \quad (20)$$

$$\lambda_3^\eta = \frac{(\eta+3)\alpha\Gamma(1+k)}{3!k}\left[-(\eta+4)(\eta+3)^{-k} + 2(\eta+3)(\eta+2)^{-k} - (\eta+2)(\eta+1)^{-k}\right] \quad (21)$$

$$\lambda_4^\eta = \frac{(\eta+4)\alpha\Gamma(1+k)}{4!k}\left[3(\eta+5)(\eta+4)(\eta+3)^{-k} - 3(\eta+4)(\eta+3)(\eta+2)^{-k} + (\eta+3)(\eta+2)(\eta+1)^{-k} - (\eta+6)(\eta+5)(\eta+4)^{-k}\right] \quad (22)$$

and for $k = 0$

$$\lambda_1^\eta = \xi + \alpha\left[\epsilon + \ln(\eta+1)\right] \quad (23)$$

$$\lambda_2^\eta = \frac{(\eta+2)\alpha}{2!}\left[\ln(\eta+2) - \ln(\eta+1)\right] \quad (24)$$

$$\lambda_3^\eta = \frac{(\eta+3)\alpha}{3!}\left[(\eta+4)\ln(\eta+3) - 2(\eta+3)\ln(\eta+2) + (\eta+2)\ln(\eta+1)\right] \quad (25)$$

$$\lambda_4^\eta = \frac{(\eta+4)\alpha}{4!}\left[-3(\eta+5)(\eta+4)\ln(\eta+3) + 3(\eta+4)(\eta+3)\ln(\eta+2) - (\eta+3)(\eta+2)\ln(\eta+1) + (\eta+6)(\eta+5)\ln(\eta+4)\right] \quad (26)$$

The three parameters ξ, α and k in the GEV distribution can be estimated by matching to the first three LH moments to their sample estimates for a selected η. By using the approximation

$$k = \alpha_0 + \alpha_1[\tau_3^\eta] + \alpha_2[\tau_3^\eta]^2 + \alpha_3[\tau_3^\eta]^3 \quad (27)$$

where η as given in the table 1.

Table 1: Values of Coefficients Vary with $\eta = 0$ and 2

η	α_0	α_1	α_2	α_3
0	0.2849	−1.8213	0.8140	−0.2835
2	0.5914	−2.3351	0.7442	−0.1616

k is obtained using (21) and (22) or (24) and (25). This gives a solution for α and ξ.

4 Application To Annual Flood Data

Characteristics of stream flow records from 31 stations in Malaysia used this study are summarized in Table 2. The data was obtained from the Department of Irrigation and Drainage Malaysia. The record length ranges are from 14 to 37 years, with average of 22 years and median of 20 years. The data are used here to demonstrate the effect of L2 moments (L2-mom), and L moments (L-Mom) in estimating the GEV distribution for flood frequency analysis.

Each of the data series is fitted to the GEV distribution by using L-Moment and L2-Moment and the GEV for $k = 0$ (EV1) distribution by using L-moment. As a first example, the annual maximum flows at site 9 was analysed. The data are plotted as filled circles against the Gumbel reduced variable in the probability plot in Figure 1 (a).

The nonexceedance probability is calculated using the Gringorten formula [13]

$$F_i = \frac{i - 0.44}{n + 0.12}$$

where i is the rank and n is the sample size. The Gumbel reduced variable is calculated as

$$g_i = -\ln\{-\ln(F_i)\}$$

The fitted distribution is plotted in Figure 1(a) as the continuous line. There are three lines will give very different estimates of floods. The EV1 distribution has a straight line, while GEV distribution having a concave shape. The graph show that the fitting is in serious error, especially for larger flows. The GEV using L2-mom was found is the best followed by GEV using L-mom and EV1 using L-mom. Second and third examples are analyses using data from site number 16 and 29 respectively. The data are plotted against the Gumbel reduced variable in Figure 1 (b) and Figure 1 (c). The advantages of using L2 moments are also evident here. Root mean square error (RMSE) was used to evaluate the precision of the methods of parameter estimate of GEV distribution using L-Mom (GEV-Lmom) and

Table 2: Characteristics of Streamflow Records of 31 Stations in Malaysia

Station Number	Site	Years of Record	L Moment Ratio		LH Moment Ratio $\eta = 2$	
			τ_3	τ_4	τ_3''	τ_4''
1732401	1	16	0.141	0.001	0.075	−0.019
1737451	2	32	0.276	0.111	0.272	0.106
1836402	3	18	0.286	0.156	0.292	0.096
2130422	4	21	0.09	0.207	0.288	0.152
2235401	5	18	0.358	0.140	0.281	0.051
2237471	6	34	0.865	0.768	0.726	0.564
2527411	7	22	0.189	0.156	0.318	0.290
2928401	8	15	0.323	0.384	0.545	0.450
3030401	9	16	−0.062	0.188	0.229	0.135
3224433	10	21	0.243	0.189	0.334	0.184
3329401	11	14	0.077	0.007	0.140	0.246
3519426	12	29	0.160	0.157	0.286	0.148
3629403	13	23	0.158	−0.086	−0.043	−0.142
4019462	14	34	0.332	0.268	0.413	0.267
4023412	15	26	0.096	0.012	0.119	0.139
4121413	16	21	−0.04	0.016	0.024	0.063
4131453	17	14	0.636	0.464	0.577	0.423
4218416	18	15	−0.201	0.113	−0.023	0.073
4219415	19	16	0.171	0.217	0.333	0.163
4223450	20	19	0.427	0.147	0.316	0.117
4232452	21	19	−0.018	0.095	0.181	0.233
4732461	22	16	0.079	0.068	0.163	0.155
4832441	23	25	0.769	0.589	0.623	0.386
5129437	24	18	0.223	0.168	0.302	0.126
5130432	25	32	0.440	0.237	0.404	0.219
5229436	26	15	0.152	0.032	0.112	0.010
5320443	27	23	0.365	0.285	0.440	0.261
5428401	28	18	0.197	0.128	0.243	0.077
5721442	29	37	0.368	0.168	0.332	0.150
5724411	30	20	0.788	0.812	0.777	0.692
6019411	31	31	0.062	0.183	0.240	0.086

L2-Mom (GEV-L2Mom) and EV1 distribution using L-Mom (EV1-Lmom). RMSE can be expressed as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=j}^{n}\left(\frac{x_i - x_{F_i}}{x_i}\right)^2}$$

where x_i is ordered set observation values and x_{F_i} is computed observation values for a given value of F_i. For a complete data series $(0 \leq F \leq 1)$, $j = 1$ while for j values was determined based on sample size data and Gringorten formula.

Table 3 gives the values of RMSE for a complete data series $(0 \leq F \leq 1)$ and the upper part of distributions $(0.9 \leq F \leq 1)$).

The 2 methods were ranked for all the stations according to the values of RMSE on a scale 1 to 3, with one being the best method and the result is given in Table 4 and 5. Clearly, the GEV using L-mom methods was the best of all, followed by the EV1 using L-mom and GEV using L2-mom for a complete data series, $(0 \leq F_i \leq 1)$. While for the upper part of distributions $(0.9 \leq F_i \leq 1)$ the GEV using L2-mom is the best followed by GEV using L-mom and EV1 using L-mom.

5 Comparison of L-Moment and LH moments Ratio Diagram

Figure 2 compares the kurtosis versus skewness curve of L moment diagram based on equations (12) and (13) for $-0.65 \leq k \leq 0.65$ and $\eta = 0$. Also plotted in the diagram is L skewness and L kurtosis for 31 stations. Most of the sample estimates of L kurtosis and L skewness are seen to fall far from the theoretical GEV curve. Figure 3 is the L2 moment diagram of skewness and kurtosis for the same 31 stations. L2 moment diagram also found based on equations (12) and (13) for $-0.65 \leq k \leq 0.65$ but with $\eta = 2$. The L2 kurtosis estimates scatter much more evenly above and below and also close to the theoretical GEV curve as compared to the L moments diagram of Figure 2. This shown that the GEV distribution using L2 moment ratio better than L moment ratios.

6 Conclusion

LH moments are based on linear combination of higher order statistics can be used for characterizing the upper part of distributions and larger events in a sample. L moments are a special case of LH moments. A comparison of the L2 moment diagram and the L moment diagram of annual maximum flow of 31 stations shows that the GEV distribution using the L2 moment ratios better than the L moments ratios with respect to this data set. It is suggested that LH moments be used as a matter of routine for estimating floods of large return periods.

7 Acknowledgments

The author would like to thank the Department of Irrigation and Drainage Malaysia for providing the flood flow data of streams in Peninsular Malaysia. The computing facili-

Table 3: Values of RMSE for A Complete Data Series $0 \leq F \leq 1$
And The Upper Part of Distribution ($0.9 \leq F \leq 1$)

Station No	$0 \leq F \leq 1$			$0.9 \leq F \leq 1$		
	GEV		EV1	GEV		EV1
	L-MOM	L2-MOM	L-MOM	L-MOM	L2-MOM	L-MON
1	0.023	0.544	0.019	0.009	0.003	0.012
2	0.024	0.146	0.061	0.014	0.006	0.007
3	0.005	0.050	0.020	0.007	0.003	0.004
4	0.016	0.077	0.034	0.001	0.002	0.002
5	0.022	0.580	0.148	0.027	0.008	0.012
6	0.076	0.341	26.600	0.391	0.323	0.806
7	0.004	0.006	0.004	0.016	0.013	0.017
8	0.005	0.020	0.008	0.016	0.016	0.027
9	0.009	0.147	0.047	0.001	0.000	0.005
10	0.005	0.004	0.021	0.009	0.009	0.009
11	0.017	0.037	0.011	0.002	0.003	0.002
12	0.013	0.015	0.012	0.000	0.001	0.000
13	0.108	8.396	0.097	0.039	0.003	0.042
14	0.005	0.006	0.007	0.010	0.009	0.025
15	0.051	0.330	0.026	0.003	0.003	0.008
16	0.000	0.000	0.000	0.000	0.000	0.000
17	0.037	0.150	3.764	0.086	0.067	0.151
18	0.003	0.066	0.013	0.008	0.002	0.005
19	0.000	0.000	0.000	0.000	0.000	0.000
20	0.280	7.647	2.854	0.028	0.001	0.012
21	0.030	0.043	0.023	0.006	0.004	0.010
22	0.034	0.077	0.022	0.006	0.006	0.007
23	0.409	11.839	104.741	0.374	0.226	0.137
24	0.078	0.116	0.150	0.005	0.004	0.004
25	0.047	0.692	1.293	0.047	0.049	0.081
26	0.022	1.404	0.016	0.007	0.002	0.008
27	0.008	0.016	0.073	0.015	0.016	0.032
28	0.050	5.090	0.304	0.009	0.006	0.007
29	0.028	0.422	0.275	0.007	0.001	0.016
30	0.378	2.283	12.843	0.628	0.256	4.546
31	0.015	0.046	0.037	0.000	0.001	0.004

Table 4: Ranking of the Methods of parameter Estimator for the 31 Stations by RMSE for a full data series, (on a scale of 1 to 3 with 1 being the best method)

	Number of Stations Receiving Ranking		
Method	1	2	3
GEV - L MOM	21	8	2
GEV - L2 MOM	2	12	17
EV1 - L MOM	9	10	12

Table 5: Ranking of the Methods of parameter Estimator for the 31 Stations by RMSE for the upper part of distributions, $(0.9 \leq F_i \leq 1)$ (on a scale of 1 to 3 with 1 being the best method)

	Number of Stations Receiving Ranking		
Method	1	2	3
GEV - L MOM	10	13	8
GEV - L2 MOM	21	7	3
EV1 - L MOM	0	12	19

ties especially MATHCAD software provided by College University Tun Hussein Onn are gratefully acknowledged.

References

[1] S. Ani, *Penggunaan Analisis Frekuensi Banjir*, Matematika, 16 (2000), 47-59.

[2] D. H. A. Al-Khudhairy, *Regional Flood Frequency Analysis, European Commission*, Joint Research Centre, Institute for System, Informatics and Safety, 1998.

[3] J.U. Chowhury, J.R. Stedinger & L. Lu, *Goodness-of Fit for Regional Generalized Extreme Value Flood Distributions*, Water Resour. Res., 27(7)(1991), 1765-1776.

[4] J.R.M. Hosking, *Analysis and Estimation of Distributions Using Linear Combination of Order Statistics*, IBM Research Division, T.J. Watson Research Center Yorktown Heights, New York 10598, 1989.

[5] J.R.M. Hosking & J.R. Wallis, *Some Statistics Useful In Regional Frequency Analysis*, Water Resour. Res., 29(2)(1993), 271-281.

[6] A. Sankarasubramaniam & K. Srinivasan, *Investigation and Comparison of Sampling Properties of L moments and ConventionaL moments*, Journal of Hydrology, 218(1999), 13-34.

[7] C.D. Sinclair and M.I. Ahmad, *Location-Invariant Plotting Positions For PWM Estimation of the Parameters of the GEV Distribution*, Journal of Hydrology, 99(1988), 271-279.

[8] J.R. Stedinger, R.M. Vogel & G.E. Foufoula, *Frequency Analysis of Extreme Events*, Handbook of Applied Hydrology, Mc-Graw Hill Book Co., New York, chapter 18, 1993.

[9] R.M. Vogel & N.M. Fennessey, *L-Moment Diagrams Should Replace Product Diagrams*, Water Resour. Res. 29(6)(1993), 1745-1752.

[10] R.M. Vogel, W.O. Thomas & T. A. McMahon, *Floodflow frequency model Selection in Southwestern U.S.A.*, J. Water Resour., Planning Manage., ASCE, 119(3)(1993).

[11] R.M. Vogel, W.O. Thomas & T. A. McMahon, *Floodflow frequency model Selection in Australia.*, Journal of Hydrology, 146(1993), 421-449.

[12] Q.J. Wang, *Using Higher Probability Weighted Moments For Flood Frequency Analysis, Journal of Hydrology*, 194(1997), 95-106.

[13] Q.J. Wang, *LH moments For Statistical Analysis of Extreme Events*, Water Resour. Res. Vol.33, (1997), 2841-2848.