MAUSAM, **68**, 3 (July 2017), 451-462

551.577.3 (540)

Regional analysis of maximum rainfall using L-moment and LH-moment: A comparative case study for the northeast India

DHRUBA JYOTI BORA, MUNINDRA BORAH and ABHIJIT BHUYAN*

Deptt. of Mathematical Sciences, Tezpur University, Napaam, Tezpur, Assam, India

**Department of Mathematics, C. K. B College, Teok, Assam, India*

(*Received 29 November 2016, Accepted 10 February 2017*)

e mail : dhrubabora@tezu.ernet.in

सार – वर्षा की बारम्बारता के आकलन के लिए सबसे सही मॉडल का चयन करने के लिए पूर्वोत्तर भारत के वर्षा के आँकड़ों पर कार्य किया गया है। प्राचलों के पाँच संभाव्य वितरणों जैसे कि सामान्यीकृत चरम मान (GEV), सामान्यीकृत लॉजिस्टिक (GLO), पीयरसन टाइप 3 (PE 3), 3 प्राचल लॉग सामान्य (LN_3) और सामान्यीकृत पेरेटो (GPA) वितरण के आकलन के लिए L- मोमेंट की पद्वति को अपनाया गया है। प्राचलों के तीन संभाव्य वितरणों जैसे कि सामान्यीकृत चरम मान (GEV), सामान्यीकृत लॉजिस्टिक (GLO) और सामान्यीकृत पेरेटो (GPA) वितरणों के आकलन के लिए चार क्रमों (L_1, L_2 L_3 & L_4 मोमेंट) के LH मोमेंट की पद्वति का उपयोग किया गया है। PE $_3$ वितरण सर्वोत्तम वितरण पाया गया है जिसमें L - मोमेंट का उपयोग किया गया है, L_1 मोमेंट का उपयोग करते हुए GPA वितरण और L_2 L_3 & L_4 का उपयोग करते हुए GLO वितरण। L मोमेंट और LH मोमेंट विश्लेषण से प्राप्त हुए परिणामों के बीच तुलना करने के लिए सापेक्षिक रूट मीन स्क्वेयर त्रुटि (RRMSE) और RBIAS को अपनाया गया है। यह पता चला है कि पूर्वोत्तर भारत में वर्षा की बारम्बारता के विश्लेषण के लिए L_1 मोमेंट पद्वति से प्राप्त GPA वितरण अत्यंत उपयुक्त और सही पद्वति है। पूर्वोत्तर भारत की वर्षा बारम्बारता विश्लेषण के लिए L मोमेंट और LH मोमेंट के अन्य क्रमों की अपेक्षा L_1 मोमेंट पद्वति विशेष रूप से अधिक सक्षम भी है।

ABSTRACT. Rainfall data of the northeast region of India has been considered for selecting best fit model for rainfall frequency analysis. The methods of L-moment has been employed for estimation of parameters five probability distributions, namely Generalized extreme value (GEV), Generalized Logistic(GLO), Pearson type 3 (PE3), 3 parameter Log normal (LN3) and Generalized Pareto (GPA) distributions. The methods of LH-moment of four orders (L_1 L_2, L_3 & L_4-moments) have also been used for estimating the parameters of three probability distributions namely Generalized extreme value (GEV), Generalized Logistic (GLO) and Generalized Pareto (GPA) distributions. PE3 distribution has been selected as the best fitting distribution using L-moment, GPA distribution using L_1-moment and GLO distribution using L_2, L_3 & L_4-moments. Relative root mean square error (RRMSE) and RBIAS are employed to compare between the results found from L-moment and LH-moment analysis. It is found that GPA distribution designated by L_1-moment method is the most suitable and the best fitting distribution for rainfall frequency analysis of the northeast India. Also the L_1-moment method is significantly more efficient than L-moment and other orders of LH-moment for rainfall frequency analysis of the northeast India.

Key words – L-moment, LH-moments, Probability distribution, Rainfall frequency analysis.

1. Introduction

There are several methods for maximum rainfall frequency analysis. To develop a suitable model for maximum rainfall for a certain return period for a particular region, it is necessary to make a comparative study among the methods.

For this study the L-moment has been employed to select the best fitting distribution among five probability distributions, namely generalized extreme value (GEV), generalized Logistic (GL), Pearson type 3 (PE3), 3 parameter Log normal (LN3) and generalized Pareto (GPA) distribution. Also LH-moment of four orders has been used to select the best fitting distribution among three probability distributions namely Generalized extreme value (GEV), Generalized Logistic (GLO) and Generalized Pareto (GPA) distributions. The homogeneity of the study region has been carried out by using heterogeneity measure proposed by Hosking and Wallis (1993). Two goodness of fitness measure namely Z-statistics and L-moment ratio diagram have been employed for identification of the best fitting distribution for our study region. Also RRMSE and RBIAS is used to make a comparison between the two best fitting distribution getting from L-moment and LH-moment analysis.

Application of extreme value distribution to rainfall data have been investigated by several authors from different parts of the world. Bora, *et al.* (2016) used L-moments and LQ-moments methods to analyze the maximum rainfall data of 12 stations of the North East India. L-moments method designated PE3 distribution as the best fit distribution whereas GPA distribution is selected as the best fit distribution by LQ-moments method. Comparative study between two methods showed that PE3 distribution designated by L-moments methods is more suitable distribution for maximum rainfall frequency analysis of the North East India. Shabri *et al.* (2011) used L-moment and TL-moment to analysis the maximum rainfall data of 40 stations of Selangor Malaysia. Comparison between the two approaches showed that the L-moments and TL-moments produced equivalent results. GLO and GEV distributions were identified as the most suitable distributions for representing the statistical properties of extreme rainfall in Selangor. Deka *et al.* (2011) fitted three extreme value distributions using LH moment of order zero to four and found that GPA distribution is the best fitting distribution for the majority of the stations in North East Region of India. Also Deka *et al.* (2009) tried to determine the best fitting distribution to describe the annual series of maximum daily rainfall data for a period of 42 years of nine stations of North East Region of India. Five extreme value distributions were fitted using L-moment and LQ-moment. Generalised Logistic distribution is empirically proved to be the most appropriate distribution for the majority of the stations in North East Region of India. Norbiato *et al*. (2007) tried to characterize the severity of a flash flood generating storm on 29th August, 2003 in the eastern Italian Alps which was characterized by extra ordinary rainfall. Regional frequency analysis based on the index variable method and L-moments are utilized to analyze annual maximum rainfall data for the region of north eastern Italy. It was found that the regional growth curves based on Kappa distribution may be useful for the region. Trefry *et al*. (2005) used L-moments method to analyze annual maximum rainfall and partial duration rainfall data of 152 stations of the state of Michigan. It was found that GEV distribution is the best fit distribution for annual maximum rainfall data and GPA distribution is the best fit distribution for partial duration rainfall data. Koutsoyiannes (2004) performed an extensive empirical investigation using a collection of 169 of the longest available rainfall records worldwide each having 100-154 year of data. This verified that the Gumbel distribution is not an appropriate distribution for rainfall distribution while Extreme value distribution of type II(EV2) is an appropriate distribution. Ogunlela (2001) studied the stochastic analysis of rainfall event in Ilorin using probability distribution functions. He found that the log Pearson type III distribution is the best for describing peak daily rainfall data of Ilorin. Adamowski *et al.* (1996) used L-moments method for regional rainfall frequency analysis of Canada and found that GEV distribution is the best fit distribution for rainfall frequency analysis of Canada.

2. Data

For this study 12 distantly situated gauged stations of the North East India *viz*., Imphal, Agartala, Shillong, Guwahati, Silchar, Jorhat, Dhubri, Lengpui, Lakhimpur, Pasighat, Mohanbari and Itanagar are considered. Annual daily maximum rainfall data of these stations for a period of 30 years from 1984 to 2013 are considered for this study. Data are collected from Regional Meteorological Centre, Guwahati.

3. Methodology

3.1. *Method of L-Moment*

The probability weighted moments (PWMs) of a random variable X with cumulative distribution function (CDF), $F(.)$ *were* defined by Greenwood *et al.* (1979) as:

$$\beta_r = M_{1,r,0} = E[X\{F(X)\}^r] \quad (1)$$

where, $M_{p,r,s} = E[X^p\{F(X)\}^r\{1-F(X)\}^s]$ (2)

and β_r can be rewritten as:

$$\beta_r = \int_0^1 x(F)F^r dF, \qquad r = 0,1,2\ldots \quad (3)$$

where, $x(F)$ is the inverse CDF of x evaluated at the probability F.

The general form of L-moments in terms of PWMs is given by Hosking and Wallis (1997) as

$$\lambda_{r+1} = \sum_{k=0}^{r} p^*_{r,k}\beta_k \quad (4)$$

where, $p^*_{r,k}$ defined by Hosking and Wallis (1997) as

$$p^*_{r,k} = \frac{(-1)^{r-k}(r+k)!}{(k!)^2(r-k)!} \quad (5)$$

The first four L-moments can be defined as:

$$\lambda_1 = \beta_0 \quad (6)$$

$$\lambda_2=2\beta_1-\beta_0 \quad (7)$$

$$\lambda_3=6\beta_2-6\beta_1+\beta_0 \quad (8)$$

$$\lambda_4=20\beta_3-30\beta_2+12\beta_1-\beta_0 \quad (9)$$

Hosking and Wallis (1997) defined L-moments ratios (LMRs) as:

$$\text{Coefficient of L-variation,} \quad \tau=\lambda_2/\lambda_1$$

$$\text{Coefficient of L-skewness} \quad \tau_3=\lambda_3/\lambda_2 \quad (10)$$

$$\text{Coefficient of L-kurtosis} \quad \tau_4=\lambda_4/\lambda_2$$

3.2. *Method of LH-Moment*

Wang (1997) introduced the concept of LH-moments as generalization of the L-moments and defined as:

$$\lambda_1^\eta=E\left[X_{(\eta+1):(\eta+1)}\right] \quad (11)$$

$$\lambda_2^\eta=\frac{1}{2}E\left[X_{(\eta+2):(\eta+2)}-X_{(\eta+1):(\eta+2)}\right] \quad (12)$$

$$\lambda_3^\eta=\frac{1}{3}E\left[X_{(\eta+3):(\eta+3)}-2X_{(\eta+2):(\eta+3)}+X_{(\eta+1):(\eta+3)}\right] \quad (13)$$

$$\lambda_4^\eta=\frac{1}{4}E[X_{(\eta+4):(\eta+4)}-3X_{(\eta+3):(\eta+4)}+3X_{(\eta+2):(\eta+4)} -X_{(\eta+1):(\eta+4)}] \quad (14)$$

When $\eta = 0$, LH-moments reduces to L-moments of Hosking (1990). As η increases, LH-moments reflect more and more the characteristics of the upper part of distribution and larger events in data (Wang, 1997). The LH-moments are denoted as L_1-moments, L_2-moments, … etc. for η=1,2,…, respectively. The LH-moments ratios (LHMRs) can be defined as:

$$\text{LH-coefficient of variation,} \quad \tau^\eta=\lambda_2^\eta/\lambda_1^\eta$$

$$\text{LH-coefficient of skewness,} \quad \tau_3^\eta=\lambda_3^\eta/\lambda_2^\eta \quad (15)$$

$$\text{LH-coefficient of skewness,} \quad \tau_4^\eta=\lambda_4^\eta/\lambda_2^\eta$$

For a given ranked sample, $x_{(1)}\leq x_{(2)}\leq\ldots\leq x_{(n)}$, the sample estimates of LH-moments defined by Wang (1997) as :

$$\hat{\lambda}_1^\eta=\frac{1}{\binom{n}{\eta+1}}\sum_{i=1}^{n}\binom{i-1}{\eta}x_{(i)} \quad (16)$$

$$\hat{\lambda}_2^\eta=\frac{1}{2}\frac{1}{\binom{n}{\eta+2}}\sum_{i=1}^{n}\left\{\binom{i-1}{\eta+1}-\binom{i-1}{\eta}\binom{n-i}{1}\right\}x_{(i)} \quad (17)$$

$$\hat{\lambda}_3^\eta=\frac{1}{3}\frac{1}{\binom{n}{\eta+3}}\sum_{i=1}^{n}\left\{\binom{i-1}{\eta+2}-2\binom{i-1}{\eta+1}\binom{n-i}{1}+\binom{i-1}{\eta}\binom{n-i}{2}\right\}x_{(i)} \quad (18)$$

$$\hat{\lambda}_4^\eta=\frac{1}{4}\frac{1}{\binom{n}{\eta+4}}\sum_{i=1}^{n}\{\binom{i-1}{\eta+3}-3\binom{i-1}{\eta+2}\binom{n-i}{1} \\ 3\binom{i-1}{\eta+1}\binom{n-i}{2}-\binom{i-1}{\eta}\binom{n-i}{3}\}x_{(i)} \quad (19)$$

Alternatively, Wang (1997) defined LH-moments as linear combination of normalized PWMs as:

$$\hat{\lambda}_1^\eta=B_\eta \quad (20)$$

$$\hat{\lambda}_2^\eta=\frac{1}{2}(\eta+2)\{B_{\eta+1}-B_\eta\} \quad (21)$$

$$\hat{\lambda}_3^\eta=\frac{1}{3!}(\eta+3)\{(\eta+4)B_{\eta+2}-2(\eta+3)B_{\eta+1}+(\eta+2)B_\eta\} \quad (22)$$

$$\hat{\lambda}_4^\eta=\frac{1}{4!}(\eta+4)\{(\eta+6)(\eta+5)B_{\eta+3} \\ -3(\eta+5)(\eta+4)B_{\eta+2}+3(\eta+4)(\eta+3)B_{\eta+1} \\ -(\eta+3)(\eta+2)B_\eta\} \quad (23)$$

where,

$$B_r=\frac{\int_0^1 x(F)F^r dF}{\int_0^1 F^r dF}=(r+1)\int_0^1 x(F)F^r dF=(r+1)\beta_r \quad (24)$$

The sample LH-moment ratios can be defined as follows:

$$\hat{\tau}^\eta=\hat{\lambda}_2^\eta/\hat{\lambda}_1^\eta,\ \hat{\tau}_3^\eta=\hat{\lambda}_3^\eta/\hat{\lambda}_2^\eta,\ \hat{\tau}_4^\eta=\hat{\lambda}_4^\eta/\hat{\lambda}_2^\eta \quad (25)$$

3.3. *Screening of data*

The Discordancy test D_i, proposed by Hosking and Wallis (1993) has been used to screen out data from stations whose point sample L-moments are markedly different from other stations. The objective is to check the

$$D_i=\frac{1}{3}N(u_i-\bar{u})^T S^{-1}(u_i-\bar{u}) \quad (26)$$

where, $S=\sum_{i=1}^{N}(u_i-\bar{u})(u_i-\bar{u})^T$ and $u_i=\left[t_2^i, t_3^i, t_4^i\right]^T$ for i-th station, N is the number of stations , S is covariance matrix of u_i and $\bar{u}$ is the mean of vector, u_i. Critical values of discordancy statistics are tabulated by Hosking and Wallis (1993), for $N = 12$, the critical value is 2.757. If the D-statistics of a station exceeds 2.757, its data is discordant from the rest of the regional data.

TABLE 1

Discordancy measures of each sites of the NE region using L-moments

No. of Site	No. of observation	Name of sites	L-CV	L-skewness	L-kurtosis	D_i
1	30	Guwahati	0.1509	0.2298	0.1551	0.27
2	30	Mohanbari	0.1458	0.1521	.1011	0.09
3	28	Silchar	0.1449	0.1420	0.1345	0.61
4	30	Lakhimpur	0.1394	0.2107	0.0976	0.93
5	30	Passighat	0.2224	0.3773	0.3455	1.82
6	30	Agartala	0.1763	0.1421	0.0529	1.30
7	30	Imphal	0.1772	0.2015	0.1711	0.19
8	30	Shillong	0.1863	0.1569	0.1778	1.32
9	26	Itanagar	0.1710	0.3629	0.2452	1.45
10	22	Dhubri	0.1798	0.1620	0.0904	0.75
11	25	Jorhat	0.1196	-0.0514	-0.0832	1.72
12	13	Lengpui	0.1166	0.1523	0.1265	1.56

TABLE 2

Discordancy measures of each sites of the NE region using L_1-moments

No. of Site	No. of observation	Name of sites	L_1-CV	L_1-skewness	L_1-kurtosis	D_i
1	30	Agartala	0.1284	0.1518	0.0386	1.78
2	22	Dhubri	0.1328	0.1931	0.1510	0.33
3	30	Guwahati	0.1209	0.2782	0.1108	0.27
4	30	Imphal	0.1356	0.2757	0.1458	0.10
5	26	Itanagar	0.1493	0.3966	0.1807	0.58
6	25	Jorhat	0.0760	0.1261	0.0298	2.56
7	30	Lakhimpur	0.1111	0.2264	0.0893	0.25
8	13	Lengpui	0.0903	0.2151	0.0180	1.58
9	30	Mohanbari	0.1100	0.1953	0.2101	1.69
10	30	Passighat	0.1880	0.4665	0.3527	2.25
11	30	Shillong	0.1363	0.2572	0.1355	0.19
12	28	Silchar	0.1084	0.2152	0.1346	0.43

Same procedure discussed above is employed for LH-moment also. For discordancy test L-cv, L-skewness and L-kurtosis are replaced by L_i-cv, L_i-skewness and L_i-kurtosis, i = 1, 2, 3, 4 respectively.

3.4. *Heterogeneity measure*

An essential task in regional frequency analysis is the determination of homogeneous regions. Hosking and Wallis (1993) suggested the heterogeneity test, H, where L- moments are used to assess whether a group of stations may reasonably be treated as belonging to a homogeneous region. The proposed heterogeneity tests are based on the L-co-efficient of variation (L-Cv), L-skewness (L-Sk) and L-kurtosis (L-Ck). These tests are defined respectively as:

$$V_1=\sqrt{\sum_{i=1}^{N} n_i\left(t_2^{(i)}-t_2^{R}\right)^2 / \sum_1^N n_i} \qquad (27)$$

TABLE 3

Discordancy measures of each sites of the NE region using L_2-moments

No. of Site	No. of observation	Name of sites	L_2-CV	L_2-skewness	L_2-kurtosis	D_i
1	30	Agartala	0.1054	0.1464	0.0732	2.52
2	22	Dhubri	0.1117	0.2448	0.2274	0.90
3	30	Guwahati	0.1082	0.2710	0.0933	0.28
4	30	Imphal	0.1201	0.2903	0.1525	0.07
5	26	Itanagar	0.1409	0.3724	0.1363	1.05
6	25	Jorhat	0.0456	0.1429	0.1035	1.68
7	30	Lakhimpur	0.0968	0.2273	0.1064	0.07
8	13	Lengpui	0.0791	0.1765	-0.0276	1.17
9	30	Mohanbari	0.0940	0.2830	0.3164	1.75
10	30	Passighat	0.1804	0.4952	0.3569	2.14
11	30	Shillong	0.1192	0.2736	0.1085	0.22
12	28	Silchar	0.0939	0.2483	0.1483	0.15

TABLE 4

Discordancy measures of each sites of the NE region using L_3-moments

No. of Site	No. of observation	Name of sites	L_3-CV	L_3-skewness	L_3-kurtosis	D_i
1	30	Agartala	0.0907	0.1615	0.1235	2.44
2	22	Dhubri	0.1011	0.3111	0.2817	0.49
3	30	Guwahati	0.0995	0.2579	0.0905	2.27
4	30	Imphal	0.1108	0.3004	0.1724	0.04
5	26	Itanagar	0.1342	0.3408	0.1108	0.81
6	25	Jorhat	0.0387	0.1120	0.1349	1.42
7	30	Lakhimpur	0.0875	0.2362	0.1291	0.06
8	13	Lengpui	0.0701	0.1287	-0.0665	1.39
9	30	Mohanbari	0.0875	0.3800	0.3892	2.23
10	30	Passighat	0.1788	0.5087	0.3512	2.17
11	30	Shillong	0.1092	0.2674	0.0881	0.28
12	28	Silchar	0.0860	0.2719	0.1620	0.41

$$V_2 = \sum_{i=1}^{N} \{n_i[(t_2^{(i)} - t_2^R)^2 + (t_3^{(i)} - t_3^R)^2]^{\frac{1}{2}}\} / \sum_{i=1}^{N} n_i \qquad (28)$$

$$V_3 = \sum_{i=1}^{N} \{n_i[(t_3^{(i)} - t_3^R)^2 + (t_4^{(i)} - t_4^R)^2]^{\frac{1}{2}}\} / \sum_{i=1}^{N} n_i \qquad (29)$$

The regional average L-moment ratios are calculated using the following formula:

$$t_2^R = \sum_{i=1}^{N} n_i\, t_2^i / \sum_{i=1}^{N} n_i$$
$$t_3^R = \sum_{i=1}^{N} n_i\, t_3^i / \sum_{1}^{N} n_i \qquad (30)$$
$$t_4^R = \sum_{i=1}^{N} n_i\, t_4^i / \sum_{1}^{N} n_i$$

where, N is the number of stations and n_i is the record length at i-th station. The heterogeneity test is then defined as :

TABLE 5

Discordancy measures of each sites of the NE region using L_4-moments

No. of Site	No. of observation	Name of sites	L_4-CV	L_4-skewness	L_4-kurtosis	D_i
1	30	Agartala	0.0811	0.1949	0.1694	2.73
2	22	Dhubri	0.0963	0.3707	0.3205	0.52
3	30	Guwahati	0.0926	0.2489	0.0875	0.05
4	30	Imphal	0.1046	0.3147	0.1975	0.07
5	26	Itanagar	0.1278	0.3133	0.0941	0.78
6	25	Jorhat	0.0402	0.0336	-0.2104	1.82
7	30	Lakhimpur	0.0810	0.2516	0.1521	0.16
8	13	Lengpui	0.0622	0.0773	-0.1209	0.81
9	30	Mohanbari	0.0862	0.4594	0.4253	2.34
10	30	Passighat	0.1795	0.5132	0.3399	2.27
11	30	Shillong	0.1016	0.2544	0.0676	0.19
12	28	Silchar	0.0809	0.2911	0.1736	0.27

$$H_j = \frac{V_j - \mu_{V_j}}{\sigma_{V_j}}, \quad j = 1, 2, 3 \tag{31}$$

where, μ_{V_j} and σ_{V_j} are the mean and standard deviation of simulated V_j values, respectively. The region is acceptably homogeneous, possibly homogeneous and definitely heterogeneous with a corresponding order of L-moments according as $H<1$, $1\leq H<2$ and $H\geq 2$.

The procedure discussed as above is similarly employed for LH-moment. For Heterogeneity test L-cv, L-skewness & L-kurtosis are replaced by L_i-cv, L_i-skewness and L_i-kurtosis, i = 1, 2, 3, 4 respectively.

3.5. *Goodness of fit measures*

3.5.1. *Z-statistics criteria*

The Z-test judges how well the simulated L-Skewness and L-kurtosis of a fitted distribution matches the regional average L-skewness and L-kurtosis values. For each selected distribution, the Z-test is calculated as follows:

$$Z^{DIST} = (\tau_4^{Dist} - t_4^R)/\sigma_4 \tag{32}$$

where, DIST refers to a particular distribution, τ_4^{DIST} is the L-kurtosis of the fitted distribution while the standard deviation of t_4^R is given by:

$$\sigma_4 = \left[(N_{sim})^{-1} \sum_{m=1}^{N_{sim}} \left(t_4^{(m)} - t_4^R\right)^2\right]^{\frac{1}{2}}$$

TABLE 6

Heterogeneity measures for L-moment and LH-moments (L_1, L_2, L_3& L_4)

Methods	H1	H2	H3
L-moment	1.54	-0.35	0.40
L1-moment	0.77	0.20	-0.13
L2-moment	1.72	-0.92	-0.43
L3-moment	1.57	-0.73	-0.23
L4-moment	1.57	-0.12	0.68

t_4^m is the average regional L-kurtosis and has to be calculated for the m^{th} simulated region. This is obtained by simulating a large number of kappa distribution using Monte Carlo simulations. The value of the Z-statistics is considered to be acceptable at the 90% confidence level if $|Z^{DIST}| \leq 1.64$. If more than one candidate distribution is acceptable, the one with the lowest $|Z^{DIST}|$ is regarded as the best fit distribution.

The Z-statistics criteria for LH-moments is similar as above. Here L-cv, L-skewness and L-kurtosis are replaced by L_i-cv, L_i-skewness and L_i-kurtosis, i = 1, 2, 3, 4 respectively.

L-moment ratio diagram is a graphical tool which can be used as another goodness of fit measure for selection of best fit distribution. It is a graph of the L - skewness and L - kurtosis which compares the fit of

TABLE 7

Z-statistics values of the distributions

Methods	S. No.	Name of the probability distribution	Z-Statistics values
L-moment	1	GLO	2.58
	2	GEV	0.87
	3	LN3	0.55
	4	PE3	0.19
	5	GPA	2.97
L_1-moment	1	GLO	1.81
	2	GEV	0.83
	3	GPA	-0.76
L_2-moment	1	GLO	-0.03
	2	GEV	-0.64
	3	GPA	-1.51
L_3-moment	1	GLO	-1.20
	2	GEV	-1.65
	3	GPA	-2.20
L_4-moment	1	GLO	-0.38
	2	GEV	-0.72
	3	GPA	-1.19

several distributions on the same graph. According to Hosking and Wallis (1997), the expression of τ_4 in terms of τ_3 for an assumed distribution is given by:

$$\tau_4=\sum_{k=0}^{8} A_k \tau_3^k \quad (33)$$

where, the coefficients A_k are tabulated by Hosking and Wallis (1997).

For LH-moment ratio diagram in equation (33) L-skewness and L-kurtosis are replaced by L_i-skewness and L_i-kurtosis, i = 1, 2, 3, 4 respectively. The coefficients A_k are calculated by Meshgi and Khalili (2009b).

3.6. *Quantile estimation*

The quantile function of the best fitting distribution PE3 is given by:

$$Q(F)=\mu + \sigma Q_0(F) \quad (34)$$

where, $Q_0(F)=\frac{2}{\gamma}\left[1+\frac{\gamma\phi^{-1}(F)}{6}-\frac{\gamma^2}{36}\right]^3-\frac{2}{\gamma}$ and $\phi^{-1}(.)$ has a standard normal distribution with zero mean and unit variance. Parameters γ, μ and σ are the standard parameterizations which can be obtained by setting:

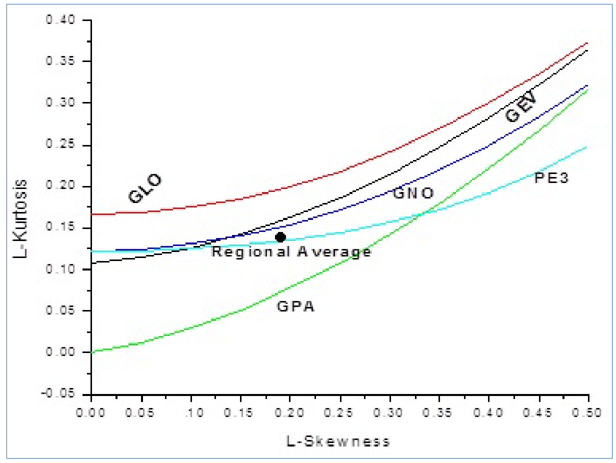

Fig. 1. L-moment ratio diagram for NE region

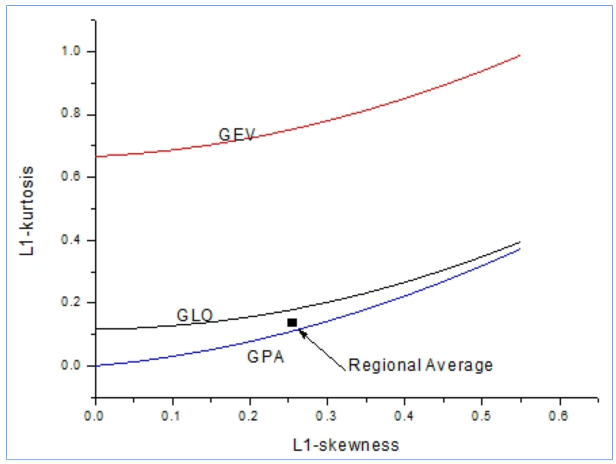

Fig. 2. L_1-moment ratio diagram for NE region

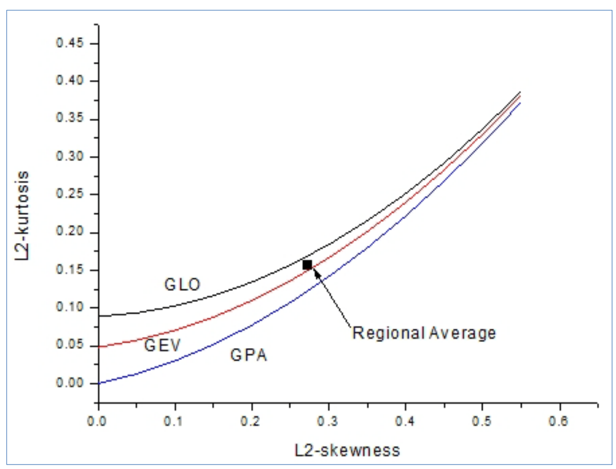

Fig. 3. L_2-moment ratio diagram for NE region

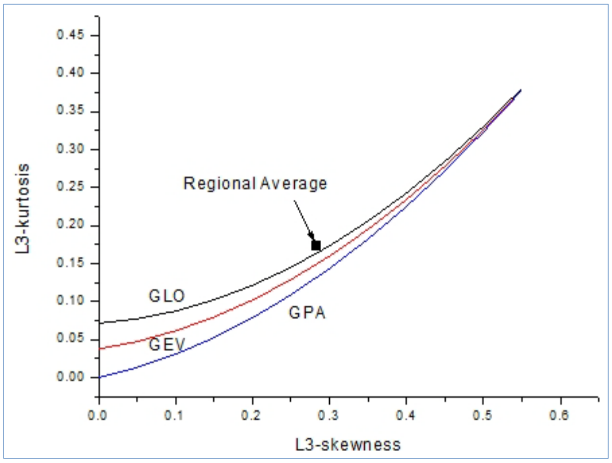

Fig. 4. L_3-moment ratio diagram for NE region

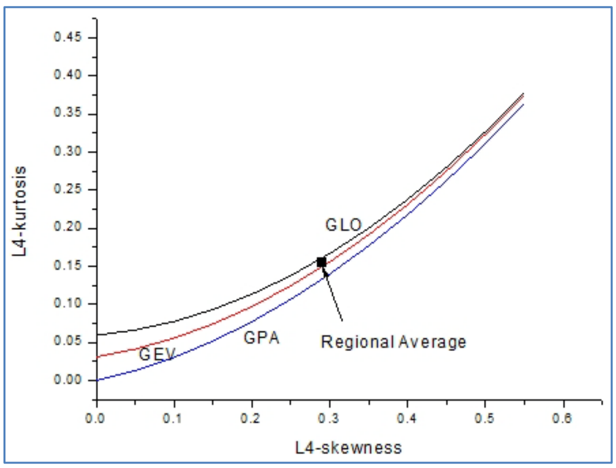

Fig. 5. L_4-moment ratio diagram for NE region

$$\alpha = \frac{4}{\gamma^2},\ \beta = \frac{1}{2}\sigma|\gamma| \text{ and } \xi = \mu - \frac{2\sigma}{\gamma}$$

Substituting the regional parameters of PE3 distribution in equation (33) quantiles are estimated.

3.7. *For L_1-moment*

The quantile function of the best fitting distribution GPA is given by:

$$Q(F) = \xi + \frac{\alpha}{k}\{1 - (1 - F)^k\} \tag{35}$$

where, *Q(F)* is the quantile estimate at return period F. ξ, α, k are the parameters.

The quantile function of the best fitting distribution GLO is given by:

$$Q(F)=\xi+\frac{\alpha}{k}\left[1-\{(1-F)/F\}^k\right] \tag{36}$$

TABLE 8

Regional parameters of best fitting distributions

Methods	Best fitting distribution	Parameters		
		Location	Scale	Shape
L-moment	PE3	1.000	0.302	1.155
L_1-moment	**GPA**	**0.555**	**0.386**	**0.254**
L_2-moment	GLO	0.754	0.141	-0.100
L_3-moment	GLO	0.711	0.141	-0.074
L_4-moment	GLO	0.676	0.143	-0.056

where, Q(F) is the quantile estimate at return period F. ξ, α, k are the parameters.

4. Results and discussion

For both L-moment and LH-moments (L_1, L_2, L_3 & L_4) methods it is observed from Tables 1-5, that the D_i values of all the twelve stations are less than critical value 2.757. Therefore, all the data of twelve stations are considered for the development of regional frequency analysis.

It has been observed from heterogeneity measures (Table 6) that for both L-moment and LH-moment methods, our study region can be considered as a possibly homogeneous one.

From Table 7 it is observed that the lowest absolute values of Z-statistics for L-moment, L_1-moment, L_2-moment, L_3-moment and L_4-moment are occurred by PE3, GPA, GLO, GLO and GLO distributions respectively. Therefore, PE3 distribution is selected as the best fitting distribution for L-moment, GPA distribution for L_1-moment and GLO distribution for L_2, L_3 & L_4-moments.

Also L-moment ratio diagram (Fig. 1) and LH-moment ratio diagrams (Figs. 2-5) show the same result.

The Parameters of the best fitting distribution are given in Table 8. Substituting the regional parameters of the distributions in respective quantile functions (34), (35) and (36), the quantiles are estimated. Estimated quantiles are given in Table 9.

The robustness of the five best fitting distributions designated by L-moment and LH-moment are also investigated for estimation of designed flood quantile. For this purpose, Monte Curlo simulation proposed by Meshgi

TABLE 9

Quantile estimates by using best fitting distributions

Methods	Best fit dist.	Quantiles									
		0.010	0.020	0.050	0.100	0.200	0.500	0.900	0.950	0.990	0.999
L-mom	PE3	0.553	0.576	0.620	0.669	0.745	0.943	1.405	1.574	1.942	2.434
L_1-mom	GPA	0.559	0.563	0.575	0.595	0.639	0.801	1.229	1.366	1.604	1.813
L_2-mom	GLO	0.235	0.300	0.395	0.476	0.572	0.754	1.101	1.237	1.577	2.158
L_3-mom	GLO	0.161	0.233	0.337	0.424	0.525	0.711	1.048	1.175	1.483	1.983
L_4-mom	GLO	0.096	0.176	0.288	0.380	0.485	0.676	1.010	1.133	1.425	1.881

TABLE 10

RRMSE values of different quantiles of best fitting distributions

Methods	Dist.	RRMSE error									
		0.010	0.020	0.050	0.100	0.200	0.500	0.900	0.950	0.990	0.999
L-mom	PE3	0.153	0.125	0.094	0.079	0.072	0.064	0.068	0.084	0.124	0.172
L_1-mom	GPA	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
L_2-mom	GLO	1.041	0.635	0.330	0.195	0.109	0.070	0.082	0.090	0.160	0.457
L_3-mom	GLO	4.503	2.126	0.854	0.425	0.203	0.089	0.077	0.087	0.161	0.417
L_4-mom	GLO	12.769	4.452	1.474	0.668	0.288	0.099	0.083	0.090	0.154	0.407

TABLE 11

RBIAS values of different quantiles of best fitting distributions

Methods	Dist.	RBIAS error									
		0.010	0.020	0.050	0.100	0.200	0.500	0.900	0.950	0.990	0.999
L-mom	PE3	-0.012	-0.005	-0.000	0.003	0.002	0.000	-0.002	-0.001	0.001	0.003
L_1-mom	GPA	0.000	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
L_2-mom	GLO	0.354	0.258	0.165	0.111	0.064	0.011	-0.019	-0.014	0.028	0.180
L_3-mom	GLO	-1.551	-0.709	-0.255	-0.104	-0.031	0.006	-0.005	-0.010	-0.008	0.049
L_4-mom	GLO	-4.272	-1.518	-0.481	-0.193	-0.061	0.004	-0.004	-0.010	-0.010	0.050

and Khalili (2009a) are used to evaluate error between simulated and calculated flood quantiles. Commonly used two error functions are relative root mean square error (RRMSE) and relative bias (RBIAS) are given by:

$$\mathrm{RRMSE} = \sqrt{\frac{1}{\mathrm{M}}\sum_{\mathrm{m}=1}^{\mathrm{M}}\left(\frac{\mathrm{Q}_{\mathrm{T}}^{\mathrm{m}}-\mathrm{Q}_{\mathrm{T}}^{\mathrm{c}}}{\mathrm{Q}_{\mathrm{T}}^{\mathrm{c}}}\right)^2}$$

$$\mathrm{RBIAS} = \frac{1}{\mathrm{M}}\sum_{\mathrm{m}=1}^{\mathrm{M}}\left(\frac{\mathrm{Q}_{\mathrm{T}}^{\mathrm{m}}-\mathrm{Q}_{\mathrm{T}}^{\mathrm{c}}}{\mathrm{Q}_{\mathrm{T}}^{\mathrm{c}}}\right)$$

where, M is the total number of samples, $\mathrm{Q}_{\mathrm{T}}^{\mathrm{m}}$ and $\mathrm{Q}_{\mathrm{T}}^{\mathrm{c}}$ are the simulated quantiles of mth sample and calculated quantiles from observed data respectively. The minimum RRMSE and RBIAS values and their associated variability are used to select the most suitable probability distribution function. For this purpose, box plots, a graphical tool introduced by Tukey (1977) are used.

Box plot is a simple plot of five quantities, namely, the minimum value, the 1^{st} quantile, the median,

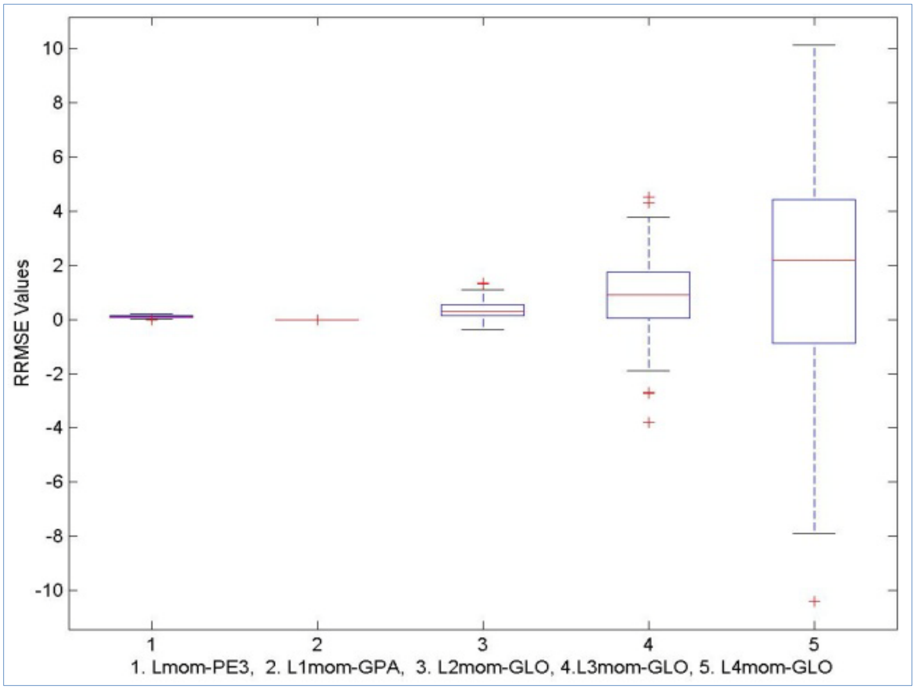

Fig. 6. Box plot of RRMSE values

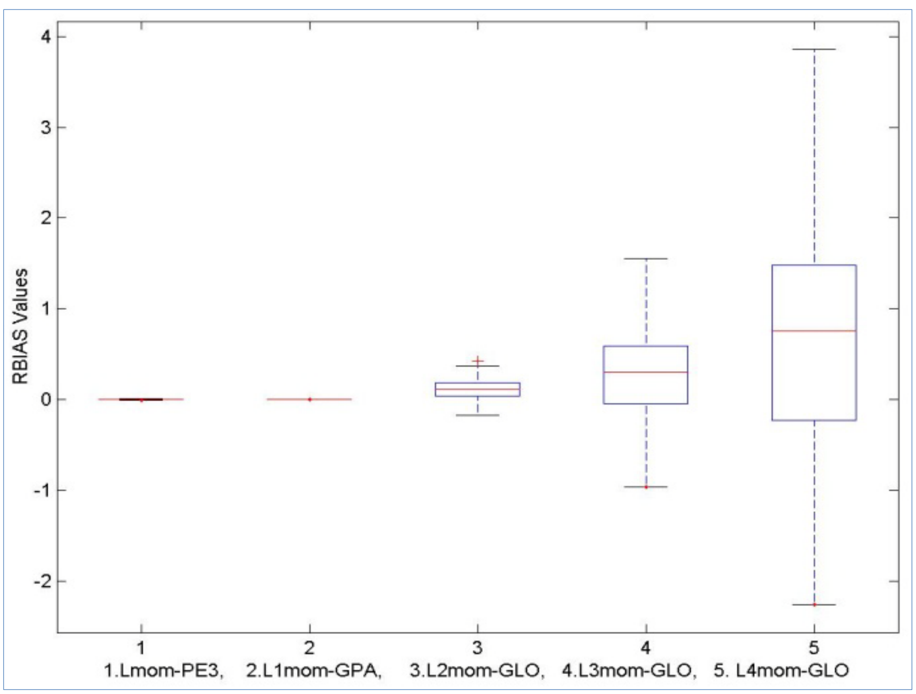

Fig. 7. Box plot of RBIAS values

the 3rd quantile and maximum value. This provides the location of the median and associated dispersion of the data at specific probability levels. The probability distribution with the minimum achieved median RRMSE or RBIAS values, as well as the minimum dispersion in the median RRMSE or RBIAS values, indicated by both ends of the box plot are selected as the suitable distribution.

RRMSE and RBIAS values are given in Tables 10 and 11 respectively. From Table 10 it is observed that the RRMSE values of GPA distribution are less than or equal to the respective RRMSE values of other distributions. Also from Table 11 it is observed that the RBIAS values (absolute) of GPA distribution are smaller than the respective RBIAS values of other distributions. Fig. 6 and Fig. 7 represent the box plot of RRMSE and RBIAS values respectively. From Fig. 6 and Fig. 7, it is observed that GPA distribution designated by L_1-moment has the minimum median RRMSE and RBIAS values as well as minimum dispersion. Hence GPA distribution is selected as suitable and the best fitting distribution for rainfall frequency analysis of the North East India. Also the L_1-moment method is significantly more efficient than L-moment and other orders of LH-moment for rainfall frequency analysis of the northeast India.

The regional rainfall frequency relationship is developed by using the best fitting GPA distribution. The form of regional frequency relationship or growth factor for GPA distribution is

$$Q(F) = [\varepsilon + \frac{\alpha}{k}\{1 - (1 - F)^k\}] * \overline{Q} \tag{36}$$

where, Q(F) is the quantile estimation at non-exceedance probability F, $\overline{Q}$ is the mean annual maximum rainfall of the site, ξ, α and k are the parameters of the GPA distribution. The regional parameters for the GPA distributions are presented in Table 8. Substituting the regional values of GPA distribution based on the data of 12 gauging sites the regional rainfall frequency relationship for gauged sites of study area is expressed as:

$$Q(F) = [0.555 + 1.449\{1 - (1 - F)^{0.254}\}] * \overline{Q}$$

For estimation of rainfall of desired non-exceedance probability for a small to moderate size gauged catchments of study area, above regional flood frequency relationship may be used. Alternatively, rainfalls of various non-exceedance probabilities may also be computed by multiplying the mean annual rainfall of a gauge station by corresponding values of growth factors based on the GPA distribution given in Table 9. The growth factor or site-specific scale factor (Q(F)/ $\overline{Q}$) is computed by dividing flood quantile (Q(F)) by the annual mean rainfall of a gauging site ($\overline{Q}$).

5. Conclusions

For both the methods, L-moment and LH-moment, Discordancy measure shows that data of all gauging sites of our study area are suitable for using regional frequency analysis. Also from homogeneity test, the region has been found to be possibly homogeneous. Regional rainfall frequency analysis was performed using five frequency distributions: *viz*., GLO, GEV, GPA, LN3 and PE3. Using L-moment ratio diagram and Z-statistic it is found that PE3 distribution is the best fitting distribution for rainfall frequency analysis of the North East India. Also using LH-moment ratio diagram and Z-statistic it is found that GPA distribution is designated as the best fitting distribution for L_1-moment and GLO distribution for L_2, L_3& L_4-moments.

Using RRMSE and RBIAS values it can be concluded that GPA distribution designated by L_1-moment is the most suitable distribution for rainfall frequency analysis of the North East India. Also the L_1-moment method is significantly more efficient than L-moment and LH-moment of other orders for rainfall frequency analysis of the North east India. The regional flood frequency relationship for gauged stations has been developed for the region and can be used for estimation of rainfalls of desired return periods.

References

Adamowski, K., Alia, Y. and Pilin, P. J., 1996, "Regional rainfall distribution for Canada", *Atmospheric Research*, **42**, 1-4, 75-88.

Bora, D. J., Borah, M. and Bhuyan, A., 2016, "Regional Analysis of Maximum Rainfall using L-moment and LQ-moment: A comparative case study for the North East India", *International Journal of Applied Mathematics & Statistical Sciences*, **5**, 6, 79-90.

Deka, S., Borah, M. and Kakaty, S. C., 2011, "Statistical analysis of annual maximum rainfall in North-East India: An application of LH-moments", *Theor Appl Climatol*., **104**, 1, 111-122.

Deka, S., Borah, M. and Kakaty, S. C., 2009, "Distributions of annual maximum rainfall series of North-East India", *European Water*, **27**, 28, 3-14.

Greenwood, J. A., Landwehr, J. M., Matlas, N. C. and Wallis, J. R., 1979, "Probability weighted moments: Defination and relation to parameters of several distributions expressible in inverse form", *Water Resour. Res*., **15**, 5, 1049-1054.

Hosking, J. R. M. and Wallis, J. R., 1997, "Regional frequency analysis: An approach based on L-moments", Cambridge University Press, New York.

Hosking, J. R. M. and Wallis, J. R., 1993, "Some statistics useful in regional frequency analysis", *Water Resour. Res*., **29**, 2, 271-281.

Hosking, J. R. M., 1990, "L-moments: Analysis and Estimation of Distribution Using Liner Combinations of Order Statistics", *Journal of the Royal Statistical Society*, Series B, **52**, 105-124.

Koutsoyiannis, D., 2004, "Statistics of extremes and estimation of extreme rainfall, empirical investigation of long rainfall records", *Hydrological science Journal*, **49**, 4, 591-610.

Meshgi, A. and Khalili, D., 2009a, "Comprehensive evaluation of regional flood frequency analysis by L- and LH-moments. I. A re-visit to regional homogeneity", *Stoch. Environ. Res. Risk Assess.*, **23**, 119-135.

Meshgi, A. and Khalili, D., 2009b, "Comprehensive evaluation of regional flood frequency analysis by L- and LH-moments. II. Development of LH-moments parameters for the generalized Pareto and generalized logistic distributions", *Stoch. Environ. Res. Risk. Assess.*, **23**, 137-152.

Norbiato, D., Borga, M.,Sangati, M. and Zanon, F., 2007, "Regional frequency analysis of extreme precipitation in the eastern Italian Alps and the August 29, 2003 flash flood", *Journal of Hydrology*, **345**, 3-4, 149-166.

Ogunlela, A. O., 2001, "Stochastic Analysis of Rainfall events in Ilorin, Nigeria", *Journal of Agricultural research and development*, **1**, 39-50.

Shabri A. B., Doud, Z. M. And Ariff, N. M., 2011 "Regional analysis of annual maximum rainfall using TL-moments method", *Theor. Appl. Climatol.*, **104**, 561-570.

Trefry, C., Watkins, D., Jr. and Johnson, D., 2005, "Regional rainfall frequency analysis for the state of Michigan", *Journal of Hydrolologic Engineering*, **10**, 6, 437-449.

Tukey, J. W., 1977, "Expolratory data analysis", *Addision-Wesley*, Reading.

Wang, Q. J., 1997, "LH-moments for statistical analysis of extreme events", *Water Resour. Res.*, **33**, 12, 2841-2848.