Journal of Applied and Natural Science 9 (4): 2366 - 2371 (2017)

JANS

Regional analysis of maximum rainfall using L-moment and TL-moment: a comparative case study for the north East India

Dhruba Jyoti Bora* and Munindra Borah

Department of Mathematical Sciences, TezpurUniversity, Napaam, Tezpur, (Assam), INDIA
*Corresponding author. E-mail: dhrubabora@tezu.ernet.in

Received: February 28, 2017; Revised received: June 24, 2017; Accepted: October 28, 2017

Abstract: In this study it has been tried to develop a suitable model for maximum rainfall frequency analysis of the North East India using best fit probability distribution. The methods of L-moment have been employed for estimation of 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 TL-moment have 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 fit distribution using L-moment and GPA distribution using TL-moment method. Relative root mean square error (RRMSE) and Relative Bias (RBIAS) are employed to compare between the results found from L-moment and TL-moment analysis. It is found that PE3 distribution designated by L-moment method is the most suitable and the best fit distribution for rainfall frequency analysis of the North East India. Also the L-moment method is significantly more efficient than TL-moment.

Key words: L-moment, TL-moments, Probability distribution

INTRODUCTION

Every year most part of the North East India has been effected by flood caused by heavy rainfall (2000-4000mm) which causes destruction of crops and properties of people. Rainfall has a direct impact in the economy of this region. So proper analysis of maximum rainfall is necessary for this region. It is also important for construction of dam, bridge, road etc.

There are several methods such as L-moment, LQ-moment, LH-moment, TL-moment 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 different selected methods. For this study the methods of L-moment and TL-moment have been used to select the best fit distribution. Also RRMSE and RBIAS is used to make a comparison between the two best fitting distribution getting from L-moment and TL-moment analysis.

Application of extreme value distribution to rainfall data have been investigated by several authors from different parts of the world. Bora, D.J. *et al.* (2016) analysed annual maximum rainfall data of 12 gauged stations of the North East India using L-moment and LQ-moment. It is found that Pearson type 3 distribution designated by L-moment is the most suitable distribution for maximum rainfall analysis of the North East India. Shabri, A. B. *et al.* (2011) used L-moment and TL-moment to analyse 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, S. *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. Regional frequency analysis based on the index variable method and L-moments are utilized to analyse 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 analyse 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. Ogunlela (2001) used five probability distribution functions namely normal, log normal, log Pearson type 3, exponential and extreme value type I to analyse daily rainfall data for a period of 41 years (1955-1995) of Ilorin. He found that the log Pearson type III distribution is the best for describing peak daily rainfall data of Ilorin while the normal distribution best described the maximum monthly rainfall for Ilorin.

ISSN : 0974-9411 (Print), 2231-5209 (Online)

MATERIALS AND METHODS

Study region and data collection: For this study annual daily maximum rainfall data of 12 distantly situated gauged stations of the North East India for a period of 30 years has been considered. Data were collected from Regional Meteorological centre, Guwahati.

Method of L-Moment: L-moments are an alternative system of describing the shapes of probability distributions.

Let $X_1, X_2, \dots, X_n$ be a sample from a continuous distribution function $F(.)$ with quantile function $Q(F)$ and let $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ be the order statistics. Then the rth L-moment λ_r defined by Hosking and Wallis (1997) is given by

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} E(X_{r-k:r}), \quad r = 1,2,\dots \tag{1}$$

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

Coefficient of L-variation, $\tau = \lambda_2/\lambda_1$

Coefficient of L-skewness $\tau_3 = \lambda_3/\lambda_2$ (2)

Coefficient of L-kurtosis $\tau_4 = \lambda_4/\lambda_2$

Method of TL-Moment: In TL-moment defined by Elamir *et al.* (2003), the term $E(X_{r-k:r})$ in the above equation (1) is replaced by $(X_{r+t_1-k:r+t_1+t_2})$. That is for each r, the conceptual sample size will be increased from r to r+t_1+t_2 and work only with the expectation of r ordered statistics $Y_{t_1+1:r+t_1+t_2}, \dots, Y_{t_1+r:r+t_1+t_2}$ by trimming the t_1 smallest and t_2 largest from the conceptual sample. Thus the rth TL-moment is defined as

$$\lambda_r^{(t_1,t_2)} = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} E(X_{r+t_1-k:r+t_1+t_2}),\ r = 1,2\dots \tag{3}$$

For t_1=t_2=0, the TL-moment yields the original L-moment. When t_1=t_2=1then the rth TL-moment is defined as

$$\lambda_r^{(1)} = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} E(X_{r+t-k:r+2t}),\ r = 1,2\dots \tag{4}$$

The TL-co-efficient of variation, TL-co-efficient of skewness and TL-co-efficient of kurtosis are defined as

$$\tau_2^{(1)} = \frac{\lambda_2^{(1)}}{\lambda_1^{(1)}},\quad \tau_3^{(1)} = \frac{\lambda_3^{(1)}}{\lambda_2^{(1)}} \quad \text{and} \quad \tau_4^{(1)} = \frac{\lambda_4^{(1)}}{\lambda_2^{(1)}} \tag{5}$$

The rth sample TL-moment is given by

$$l_r^{(t)} = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} \hat{E}(X_{r+t-k:r+2t}),\ r = 1,2\dots \tag{6}$$

where unbiased estimator is given by

$$\hat{E}(X_{r+t-k:r+2t}) = \frac{1}{\binom{n}{r+2t}}\sum_{i=1}^{n} \binom{i-1}{r+t-k-1}\binom{n-i}{t+k} X_{i:n} \tag{7}$$

Regional rainfall frequency analysis

Screening of data: The Discordancy test D_i, proposed by Hosking and Wallis (1993) is given by

$$D_i = \frac{1}{3}N(u_i - \bar{u})^T S^{-1}(u_i - \bar{u} \tag{8}$$

Where, $S = \sum_{i=1}^{N}(u_i - \bar{u})(u_i - \bar{u})^T$ and $u_i = [t_2^i, t_3^i, t_4^i]^T$ 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.

Same procedure discussed above is employed for TL-moment also. Here L-moment ratios are replaced by respective TL-moment ratios.

Heterogeneity measure: 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. These tests are defined respectively as

$$V_1 = \sqrt{\sum_{i=1}^{N} n_i (t_2^{(i)} - t_2^R)^2 / \sum_1^N n_i} \tag{9}$$

$$V_2 = \sum_{i=1}^{N}\{n_i[(t_2^{(i)} - t_2^R)^2 + (t_3^{(i)} - t_3^R)^2]^{\frac{1}{2}}\} / \sum_{i=1}^{N} n_i \tag{10}$$

$$V_3 = \sum_{i=1}^{N}\{n_i[(t_3^{(i)} - t_3^R)^2 + (t_4^{(i)} - t_4^R)^2]^{\frac{1}{2}}\} / \sum_{i=1}^{N} n_i \tag{11}$$

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

$$t_2^R = \sum_{i=1}^{N} n_i\, t_2^i / \sum_{i=1}^{N} n_i,\quad t_3^R = \sum_{i=1}^{N} n_i\, t_3^i / \sum_1^N n_i,\quad t_4^R = \sum_{i=1}^{N} n_i\, t_4^i / \sum_1^N n_i \tag{12}$$

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

$$H_j = \frac{V_j - \mu_{V_j}}{\sigma_{V_j}},\quad j = 1,2,3 \tag{13}$$

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 the methods of TL-moment.

Goodness of fit measures

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 defined by Hosking and Wallis (1993) as follows

$$Z^{DIST} = (\tau_4^{Dist} - t_4^R)/\sigma_4 \tag{14}$$

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

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

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 TL-moment is same as above.

Moment ratio diagram: It is a graph of the skewness and kurtosis which compares the fit of 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

$$\tau_4 = \sum_{k=0}^{8} A_k \tau_3^k \tag{15}$$

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

For TL-moment ratio diagram in equation (15) L-skewness and L-kurtosis are replaced by TL-skewness and TL-kurtosis. The coefficients A_k are found in Shabri et al. (2011).

RESULTS AND DISCUSSION

For both L-moment and TL-moment methods it is observed from table-1 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 consid-

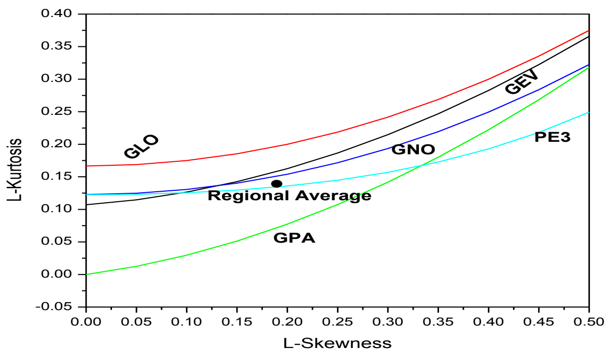

Fig. 1. *L-moment ratio diagram for NE region.*

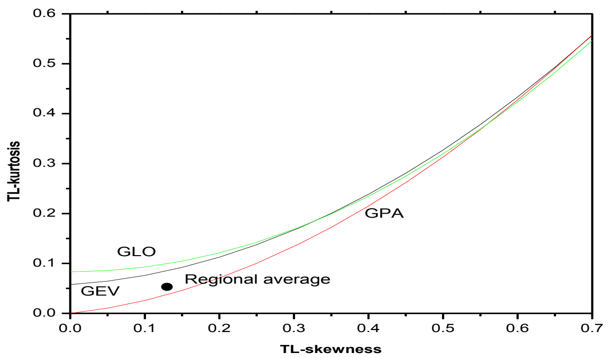

Fig. 2. *TL-moment ratio diagram for NE region.*

Table 1. Discordancy measures of each sites of the NE region using L-moments and TL-moments.

S. N.	Name of sites	No. of observation	L-moment D_i	TL-moment D_i
1	Guwahati	30	0.27	0.26
2	Mohanbari	30	0.09	0.95
3	Silchar	28	0.61	0.52
4	Lakhimpur	30	0.93	0.71
5	Passighat	30	1.82	0.52
6	Agartala	30	1.30	1.75
7	Imphal	30	0.19	0.04
8	Shillong	30	1.32	1.13
9	Itanagar	26	1.45	1.25
10	Dhubri	22	0.75	2.23
11	Jorhat	25	1.72	1.50
12	Lengpui	13	1.56	1.14

Table 2. Heterogeneity measures for L-moment and TL-moments.

Methods	H1	H2	H3
L-moment	1.54	-0.35	0.40
TL-moment	0.69	0.49	1.04

Table 3. Z-statistics values of the distributions.

Methods	Z-statistics values of probability distributions				
	GLO	GEV	GPA	PE3	LN3
L-moment	2.58	0.87	2.97	**0.19**	0.55
TL-moment	2.94	2.16	-0.51	--	--

ered for the development of regional frequency analysis.

It has been observed from heterogeneity measures (Table-2) that for both L-moment and TL-moment methods, our study region can be considered as a possibly homogeneous one.

From table-3 it is observed that for L-moment the absolute value of Z statistics less than the critical value 1.64 is occurred by three distributions GEV, PE3, and LN3. Out of these three distributions PE3 have the lowest Z statistics value. Also for TL-moment the absolute value of Z statistics less than the critical value 1.64 is occurred by GPA distribution only. Therefore, PE3 distribution is selected as the best fitting distribution for L-moment and GPA distribution for TL-moment method.

Also L-moment ratio diagram (fig-1) and TL-moment ratio diagrams (fig-2) show the same result.

The quantile function of the best fitting distribution PE3 designated by L-moment is given by

$$Q(F) = \mu + \sigma Q_0(F) \quad (16)$$

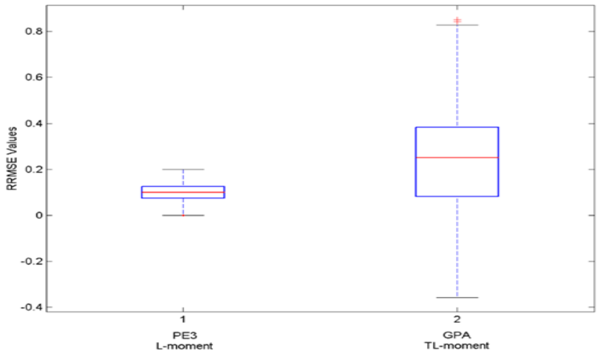

Fig. 3. *Box plot of RRMSE values.*

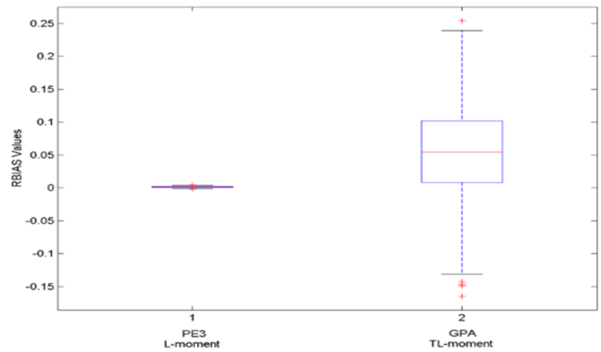

Fig. 4. *Boxplot of RBIAS values.*

where $Q_0(F) = \frac{2}{\gamma}\left[1 + \frac{\gamma \emptyset^{-1}(F)}{6} - \frac{\gamma^2}{36}\right]^3 - \frac{2}{\gamma}$ and $\emptyset^{-1}(.)$ has a standard normal distribution with zero mean and unit variance.Q(F) is the quantile estimate at return period T and $F = 1 - {}^{1}/_{T}$. Parameters γ, μ and σ are the standard parameterizations which can

Table 4. Regional parameters of best fitting distributions.

Methods	Best fitting distribution	Parameters		
		Location	Scale	Shape
L-moment	PE3	1.000	0.302	1.155
TL-moment	GPA	0.656	0.510	0.365

Table 5. Quantile estimates by using best fitting distributions.

Methods	Distribution	Return period (in years)				
		2	10	20	100	1000
L-moment	PE3	0.943	1.450	1.574	1.942	2.434
TL-moment	GPA	0.968	1.451	1.586	1.794	1.942

Table 6. RRMSE values of different quantiles of best fitting distributions.

Methods	Distribution	Return period (in year)				
		2	10	20	100	1000
L-moment	PE3	0.064	0.068	0.084	0.124	0.172
TL-moment	GPA	0.067	0.077	0.115	0.260	0.665

Table 7. RBIAS values of different quantiles of best fitting distributions.

Methods	Distribution	Return period (in year)				
		2	10	20	100	1000
L-moment	PE3	0.000	-0.002	-0.001	0.001	0.003
TL-moment	GPA	0.001	-0.001	0.009	0.056	0.184

be obtained by setting

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

The quantile function of the best fitting distribution GPA designated by TL-moment is given by

$$Q(F) = \xi + \frac{\alpha}{k}\{1 - (1 - F)^k\} \quad (17)$$

where Q(F) is the quantile estimate at return period T. x, α, k are the parameters and $F = 1 - 1/T$.

The Parameters of the best fitting distributions are given in table-4. Substituting the regional parameters of the distributions in respective quantile functions (16) and (17) the quantiles are estimated. Estimated quantiles are given in table 5.

The robustness of the two best fitting distributions designated by L-moment and TL-moment are also investigated. For this purpose, Monte Curlo simulation proposed by Meshgi and Khalili (2009) are used to evaluate error between simulated and calculated flood quantiles. Two error functions, relative root mean square error (RRMSE) and relative bias (RBIAS) are given by $RRMSE = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(\frac{Q_T^m - Q_T^c}{Q_T^c}\right)^2}$ and $RBIAS = \frac{1}{M}\sum_{m=1}^{M}\left(\frac{Q_T^m - Q_T^c}{Q_T^c}\right)$

Where, M is the total number of samples, Q_T^m and Q_T^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 1st quantile, the median, 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.

From table-6 and table-7it is observed that the RRMSE and RBIAS values of PE3 distribution designated by L-moment method are less than the respective RRMSE and RBIAS values of GPA distribution designated by TL-moment method. From the box plot of RRMSE and RBIAS values (fig-6 and fig-7) it is observed that PE3 distribution designated by L-moment has the minimum median RRMSE and RBIAS values as well as minimum dispersion. Hence PE3 distribution is selected as suitable and the best fit distribution for rainfall frequency analysis of the North East India. Also the L-moment method is significantly more efficient than TL-moment for rainfall frequency analysis of the North east India.

Development of Model: The regional rainfall frequency relationship is developed by using suitable and the best fitting distribution PE3. The form of regional frequency relationship or growth factor for PE3 distribution is

$$Q_T = \left[\mu + \sigma\frac{2}{\gamma}\left\{\left\{1 + \frac{\gamma\emptyset^{-1}(F)}{6} - \frac{\gamma^2}{36}\right\}\right\}^3 - \frac{2}{\gamma}\right] * \bar{Q} \quad (18)$$

where Q_T is the maximum rainfall at return period T, $\bar{Q}$ is the mean annual maximum rainfall of the site, $\emptyset^{-1}(.)$ has a standard normal distribution with zero mean and unit variance. Parameters γ, μ and σ are the standard parameterizations which are given in the table-4. Substituting these values in expression (18) rainfall frequency relationship for gauged sites of study area is expressed as:

$$Q_T = \left[1.000 + \frac{0.604}{1.155}\left\{\left\{1 + \frac{1.155\emptyset^{-1}(F)}{6} - \frac{1.334}{36}\right\}\right\}^3 - \frac{2}{1.155}\right] * \bar{Q}$$

For estimation of maximum rainfall for a desired return period above regional flood frequency relationship may be used.

Conclusion

For both the methods, L-moment and TL-moment, Discordancy measure shows that data of all the 12 gauging sites of the study region can be considered for analysis. Also from homogeneity test, it is found that the region is possibly homogeneous. From Regional rainfall frequency analysis using L-moment method it is found that PE3 distribution is the best fit distribution for rainfall frequency analysis of the North East India. Also using TL-moment it is found that GPA distribution is the best fit distribution for the region. Using RRMSE and RBIAS values it can be concluded that PE3 distribution designated by L-moment is more suitable distribution for rainfall frequency analysis of the North East India. Also the L-moment method is significantly more efficient than TL-moment 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

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.

Elamir, E.A.H. and Seheult, A.H., (2003). Trimmed L-moments. *Comput. Statist. Data Anal*. 43: 299–314.

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. and Wallis, J.R. (1997). Regional frequency analysis- An approach based on L-moments. *Cambridge University Press*, New York.

Ogunlela, A.O.(2001). Stochastic Analysis of Rainfall events in Ilorin, Nigeria. *Journal of Agricultural Research and Development.* 1:39-50.

Meshgi, A. and Khalili, D. (2009). 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.

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). Exploratory data analysis. *Addision-Wesley, Reading.*