International Journal of Scientific Research in Mathematical and Statistical Sciences
Vol.10, Issue.6, pp.09-17, December 2023
E-ISSN: 2348-4519
Available online at: www.isroset.org

Research Paper

Effect of Data Length on Assessment of Uncertainty of Error in Rainfall Estimation Using Log Normal Distribution

N. Vivekanandan[1]

[1]Scientist-C, Central Water and Power Research Station, Pune, Maharashtra, India

Author's Mail Id: anandaan@rediffmail.com

Received: 05/Nov/2023; **Accepted**: 09/Dec/2023; **Published**: 31/Dec/2023

Abstract— Estimation of rainfall for a given return period is considered as one of the important parameters in hydrological studies to estimate flood discharge, which is needed for the planning, design and management of civil engineering infrastructure projects. In this paper, a study on effect of data length on assessment of uncertainty of error in rainfall estimation by adopting 2-parameter Log Normal (LN2) distribution for Pune and Vadgaon Maval rain-gauge sites of Maharashtra was carried out. The parameters of LN2 were determined by Method of Moments (MoM), Maximum Likelihood Method (MLM) and Method of L-Moments (LMO), and used for estimation of rainfall. The selection of best fit method of LN2 for rainfall estimation was made through Goodness-of-Fit (viz., Chi-Square and Kolmogorov-Smirnov) and diagnostic (viz., correlation coefficient, Nash-Sutcliffe model efficiency and root mean squared error) tests. The outcomes of rainfall analysis of Pune and Vadgaon Maval indicated that the estimated rainfall increases when data length increases. The study showed that the standard error in the estimated rainfall using three (viz., MoM, MLM and LMO) methods of LN2 are in decreasing order when data length increases. The study also showed that the standard errors in rainfall estimates computed by MoM for Pune and MLM for Vadgaon Maval are less than those values of other methods. The results indicated that the estimated rainfall by LMO for Pune and MoM for Vadgaon Maval is higher than those values other methods. Based on GoF and diagnostic tests results, it was found that the LMO is superior to MoM and MLM, and hence adjudged as best method for rainfall estimation at Pune and Vadgaon Maval.

Keywords— Chi-Square, Kolmogorov-Smirnov, Log Normal, L-Moments, Mean Squared Error, Model Efficiency, Rainfall

1. Introduction

Extreme rainfall is one of the hydrology quantities that can be used for computing the flood discharge. This can be used for planning, design and management of civil engineering structures [1]. Moreover, the determination of extreme (i.e., annual 1-day maximum) rainfall would enhance the management of water resources projects as well as the effective utilization of water resources. Also, the extreme rainfall is one of the important parameters in hydrological studies and urban drainage systems, as it directly contributes to runoff. This can be estimated by recorded rainfall data over many stations in a given region, which may have some uncertainty that has traditionally been estimated by probability distribution model [2].

Number of attempts has been made by various researchers in estimating the uncertainty of error in rainfall estimation [3-6]. Generally, the uncertainty in hydrological modelling can be classified as (i) data and sampling errors and (ii) modelling or structural errors [7]. The use of limited quantity of rainfall data (i.e., data of short record length) in the frequency analysis introduces sampling uncertainty. Also, the estimates of higher order moments (skewness and kurtosis) become unstable, in particular due to the presence of extremes or outliers in data series. Hailegeorgis et al. [8] described that the regional frequency analysis of extreme rainfall events and the derivation of intensity-duration-frequency curves is subject to the uncertainties of different sources that includes (i) data quality (viz., stationary and independent), sampling of data related to the time period and length of data series and the sampling type (viz., annual maximum series or partial duration series); (ii) selection of frequency distribution to describe the data; (iii) parameter estimation; and (iv) regionalization and quantile estimation.

A number of probability distributions such as Extreme Value Type-1 (EV1) (commonly known as Gumbel), Extreme Value Type-2, 2-parameter Log Normal (LN2), Log Pearson Type-III (LP-III), Exponential (EXP), Generalized Extreme Value (GEV), Gamma (GAM), Generalized Pareto, etc., are generally available for extreme value analysis of rainfall. Cho et al. [9] made an attempt to investigate the properties of tropical precipitation by characterizing the gamma and LN2 models of rain rates using TRMM (Tropical Rainfall Measuring Mission) dataset. Chifurira and Chikobvu [10] applied LN2, GAM and EXP distributions to assess the effect of data length on the selection of a suitable distribution for rainfall estimation of Zimbabwe. Amin et al. [11] applied Normal, LN2, LP-III and Gumbel distributions for modelling annual maximum rainfall in the northern regions of Pakistan and identified the best fit distribution through Goodness-of-Fit tests. Rosmaini and Saphira [12] applied Normal, GAM and LN2 distributions for estimation of monthly rainfall of Tuntungan, Tanjung Selamat, and Medan Selayang Stations in Medan City, Indonesia. Hasan et al. [13] evaluated the GAM, LN2 and Weibull distributions for predicting rainfall in ungauged catchments of Australia. They found that the LN2 model overestimates extremely high rainfall events whereas the LN2 model underestimates low rainfall events. Study by Vivekanandan [14] revealed that LP-III is best fit distribution for estimation of peak flood at Haora site. Kaur et al. [15] applied Gumbel, LP-III, LN2 and Ven Te Chow distributions for the prediction of annual 1-day, 2-day and 3-day maximum rainfall values of Roorkee. In light of the above, LN2 distribution is adopted for rainfall estimation and the results are presented in this paper.

In statistical theory, the Method of Moments (MoM), Maximum Likelihood Method (MLM) and Method of L-Moments (LMO) are generally applied for determination of parameters of LN2 [16, 17]. Research reports indicated that MoM is a natural and relatively easy parameter estimation method [18, 19]. But, the studies carried by various researchers indicated that the estimated parameters of distributions fitted by MoM are often less accurate than those obtained by other parameter estimation procedures like MLM, PWM (Probability Weighted Moments) and LMO. MLM is considered the most efficient method, since it provides the smallest sampling variance of the estimated parameters and hence of the estimated quantiles compared to other methods [20]. But, the method has the disadvantage of frequently giving biased estimates and often failed to give the desired accuracy in estimating the extremes from hydrological data. To overcome from this issue, the LMO as suggested by Hosking [21] that has been widely applied to determine the parameters of various probability distributions. He also described that LMOs are linear combinations of the PWM tend to share similar characteristics with PWM and MLS (Method of Least Squares), and also the computations are simpler. However, there was no general agreement in applying particular method for a region because of the characteristics of the estimators of LN2. Hence, in the present study, MoM, MLM and LMO were considered for determination of parameters of LN2 that are used in estimating the rainfall. The characteristics of rainfall data series used in the analysis was examined through Wald-Wolfowitz runs test for independence, Mann-Whitney U-test for homogeneity and Grubbs' test for identifying the outliers in data series. The adequacy of fitting three methods of LN2 distribution to the data series was evaluated by Goodness-of-Fit (GoF) tests viz., Chi-Square (χ^2) and Kolmogorov-Smirnov (*KS*). The selection of best method of LN2 for rainfall estimation was made through diagnostic tests viz., correlation coefficient (*CC*), Nash-Sutcliffe Model Efficiency (*NSE*) and Root Mean Squared Error (*RMSE*). This paper presented a study on effect of data length on assessment of uncertainty of error in rainfall estimation by adopting three methods of LN2 distribution with an illustrative example and the results obtained thereon.

In this paper, Section 1 describes the significance of importance for rainfall estimation in hydrological studies and the assessment of uncertainty in model error. Section 2 details the methodology applied in determining the parameters of LN2 distribution by MoM, MLM and LMO, and evaluation of the results through GoF and diagnostic tests. Section 3 gives the description of the study area and data used in the study whereas the results and discussions obtained from rainfall data analysis are elaborated in Section 4. The conclusions and scope of future work are given in Section 5.

2. Methodology

The probability distribution function (*f(x)*) and cumulative distribution function (*F(x)*) of LN2 distribution is given by:

$$f(x)=\frac{1}{\beta x\sqrt{2\pi}}e^{-\left(\frac{(ln(x)-\alpha)^2}{2\beta^2}\right)},\ x>0,\ \alpha>0,\ \beta>0 \quad (1)$$

$$F(x)=\varphi\left(\frac{ln(x)-\alpha}{\beta}\right) \quad (2)$$

Where, x is the variable (i.e., annual 1-day maximum rainfall (AMR)), α is the scale parameter, β is the shape parameter, $f(x)$ is the probability distribution function of x, $F(x)$ is the cumulative distribution function (CDF) of x and $\phi(\ldots)$ is the CDF of the standard normal distribution [22]. The parameters of LN2 were determined by MoM, MLM and LMO that are further used to estimate the extreme (i.e., 1-day maximum) rainfall (ER) for different return period (T) and given by:

$$x(T)=e^{\alpha+K(T)\beta} \quad (3)$$

Where, $x(T)$ is the estimated extreme rainfall (*ER*), $K(T)$ is the frequency factor for a return period (T) that can be derived from the following equations.

$$K(T)=W-\frac{2.515517+0.802853\,W+0.00110328\,W^2}{1+1.43278\,W+0.189269\,W^2+0.001308\,W^3} \quad (4)$$

$$W=ln\left(\frac{1}{P^2}\right)^{0.5}\ with\ P=1-\frac{R}{N+1} \quad (5)$$

Where, P is the probability of exceedance R is the rank assigned to the sample arrange in ascending order in such a way that $x(1)<x(2)<x(3)<\ldots\ldots x(N)$ and N is the number of samples. Table 1 presents the equations used in determining the parameters of LN2 using MoM, MLM and LMO.

Table 1. MoM, MLM and LMO Estimators of LN2 distribution

Method	Scale parameter (α)	Shape parameter (β)
MoM	$\alpha=-\frac{1}{2}ln\left(\sum_{i=1}^{N}(x(i))^2\right)+2ln\left(\sum_{i=1}^{N}x(i)\right)-\frac{3}{2}ln(N)$	$\beta=\left[ln\left(\sum_{i=1}^{N}(x(i))^2\right)-2ln\left(\sum_{i=1}^{N}x(i)\right)+ln(N)\right]^{1/2}$
MLM	$\alpha=\frac{1}{N}\sum_{i=1}^{N}ln(x(i))$	$\beta=\left(\frac{1}{N}\sum_{i=1}^{N}(ln(x(i))-\alpha)^2\right)^{1/2}$
LMO	$\alpha=\lambda_1=b_0=\frac{1}{N}\sum_{i=1}^{N}ln(x(i))$	$\lambda_2=2b_1-b_0=\beta/\sqrt{\pi}$ $b_1=\frac{1}{N(N-1)}\sum_{i=2}^{N}(i-1)ln(x(i))$

In Table 1, $x(i)$ is the observed AMR for i^{th} sample, $ln(x(i))$ is the logarithmic transformed value of $x(i)$, λ_1 and λ_2 are the first and second LMO, b_0 and b_1 are the first and second moments of sample, and N is the number of samples.

2.1 Computation of Standard Error

The standard error (SE) in rainfall estimation using MoM, MLM and LMO of LN2 distribution was computed from the following equations that are expressed by:

$$SE(x(T)) = \left(\frac{1}{2}\right)\left[x(T)\left(e^{SE(y(T))} - e^{-SE(y(T))}\right)\right] \quad (6)$$

$$SE(y(T)) = \left(\frac{\beta}{\sqrt{N}}\right)\left(1+0.5K(T)^2\right)^{0.5} \quad (7)$$

The lower and upper confidence limits (LCL and UCL) of the estimated rainfall for a return period (T) at 99% level were computed from the following equations.

$LCL = x(T)-2.58*SE(x(T))$ and $UCL=x(T)+2.58*SE(x(T))$ (8)

2.2 Goodness-of-Fit Tests

The theoretical descriptions of GoF tests viz., Chi-square (χ^2) and Kolmogorov-Smirnov (KS) applied in checking the adequacy of fitting LN2 distribution to the AMR series are given as below:

$$\chi_C^2 = \sum_{j=1}^{NC} \frac{(O_j(x) - E_j(x))^2}{E_j(x)} \text{ and}$$

$$KS_C = Max \sum_{i=1}^{N} \left| F_e(x(i)) - F_D(x(i)) \right| \quad (9)$$

Where, χ_C^2 and KS_c are the computed *values of* χ^2 and KS tests statistic using three methods of LN2, $O_j(x)$ and $E_j(x)$ are the observed and expected frequency values of x in j^{th} class, $F_D(x(i))$ is the derived CDF of $x(i)$ using LN2 and $F_e(x(i))$ is the Empirical CDF of $x(i)$ using Weibull plotting position formula for $i=1,2,3,\ldots,N$ with $x(1)<x(2)<\ldots..x(N)$. The acceptance region of χ^2 statistic at the desired significance level (η) is given by $\chi_C^2 \le \chi^2_{1-\eta,NC-m-1}$ wherein m is the number of parameters of the distribution and NC is the number of frequency classes. The theoretical values of GoF tests statistic for different significance level can be read from the technical note on 'Goodness-of-Fit Tests for Statistical Distributions' [23]. If the computed values of GoF tests statistic given by the method is less than its theoretical value at the desired significance level (η) then the selected method is acceptable for rainfall estimation.

2.3 Diagnostic Test

The theoretical expressions of CC, NSE and $RMSE$ [24] applied in selecting the best fit method of LN2 distribution for rainfall estimation is given as below:

$$CC = \frac{\sum_{i=1}^{N}(x(i)-\mu(x))(z(i)-\mu(z))}{\sqrt{\sum_{i=1}^{N}(x(i)-\mu(x))^2 \sum_{i=1}^{N}(z(i)-\mu(z))^2}} \quad (10)$$

$$NSE(\%) = \left(1 - \frac{\sum_{i=1}^{N}(x(i)-z(i))^2}{\sum_{i=1}^{N}(x(i)-\mu(x))^2}\right) * 100 \quad (11)$$

$$RMSE = \left(\frac{1}{N}\sum_{i=1}^{N}(x(i)-z(i))^2\right)^{1/2} \quad (12)$$

Where, $z(i)$ is the predicted AMR of i^{th} sample, $\mu(x)$ is the average of observed AMR and $\mu(z)$ is the average of predicted AMR. The method with high CC (say, CC>0.9), good NSE and minimum $RMSE$ is adjudged as best method for estimation of rainfall.

3. Study Area and Data Used

In this paper, a study on effect of data length on assessment of uncertainty of error in rainfall estimation of Pune and Vadgaon Maval rain-gauge sites of Maharashtra was carried out. These sites are located on the western side of Deccan Plateau and are on leeward side of Sahyadri mountain range which forms a barrier from Arabian Sea. Figure 1 presents the index map of the study area with locations of Pune and Vadgaon Maval sites [25].

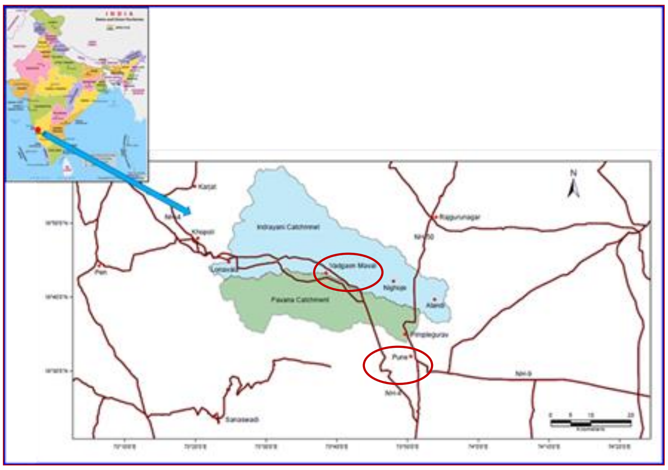

Figure 1. Index map of the study area with locations of rain-gauge stations

From Figure 1, it was witnessed that the Pune IMD (India Meteorological Department) rain gauge site is located approximately between the latitude 18° 31′ N and longitude 73° 51″ E and whereas the Vadgaon Maval site is located within latitude 18° 42′ N and longitude 73° 38″ E. In this paper, the daily rainfall observed at Pune (1901 to 2017) and Vadgaon Maval (1901 to 1965, 1968 to 1971 and 1973 to 2017) sites was used. By considering the importance of the hydrologic extremes, the missing data for the years 1966, 1967 and 1972 were not considered in rainfall data analysis for Vadgaon Maval. The AMR series is derived from the daily rainfall data and used to generate the data series (DS) with different data length viz., DS1 (series with 50 years data), DS2 (series with 75 years data for Pune and 72 years for Vadgaon Maval), DS3 (series with 100 years data for Pune and 97 years for Vadgaon Maval) and DS4 (series with entire data viz., 117 years for Pune and 114 years for Vadgaon Maval). The generated AMR series was used in rainfall estimation by applying three methods (viz., MoM, MLM and LMO) of LN2. Table 2 presents the descriptive statistics of AMR with different data length used in rainfall estimation for Pune and Vadgaon Maval. From Table 2, it was noted that the average and standard deviation of AMR of DS4 series is higher than those values of DS1, DS2 and DS3.

Table 2. Descriptive statistics of AMR of DS1, DS2, DS3 and DS4 series for Pune and Vadgaon Maval

Data series	Pune				Vadgaon Maval			
	AVG (mm)	SD (mm)	C_s	C_k	AVG (mm)	SD (mm)	C_s	C_k
DS1	69.8 (4.196)	22.2 (0.321)	0.695 (-0.184)	0.458 (-0.013)	89.3 (4.437)	31.8 (0.330)	1.241 (0.369)	1.683 (-0.126)
DS2	70.7 (4.201)	24.6 (0.344)	0.780 (-0.052)	0.148 (-0.155)	93.9 (4.480)	35.7 (0.347)	1.315 (0.460)	1.640 (-0.124)
DS3	72.2 (4.222)	25.0 (0.342)	0.764 (-0.030)	0.098 (-0.265)	97.8 (4.512)	40.7 (0.366)	1.598 (0.601)	2.813 (0.123)
DS4	74.2 (4.249)	26.1 (0.344)	1.010 (-0.042)	1.695 (-0.014)	102.5 (4.556)	42.9 (0.377)	1.420 (0.472)	2.143 (-0.187)

AVG: Average; SD: Standard Deviation; C_s: Coefficient of Skewness; C_k: Coefficient of Kurtosis. Numbers given within brackets indicates the descriptive statistics of log-transformed data.

4. Results and Discussion

By applying the procedures, as described above, the assessment of uncertainty of error in rainfall estimation using MoM, MLM and LMO of LN2 distribution with reference to data length for Pune and Vadgaon Maval sites was carried out and the results are presented in the following sections.

4.1 Analysis of AMR Series Based on Statistical Tests

In this paper, Wald-Wolfowitz runs test and Mann-Whitney U-test were used for checking the randomness and homogeneity of AMR series. Grubbs' test was used for detection of outliers in the data series. Tables 3 and 4 present the results of statistical tests applied to the AMR series of Pune and Vadgaon Maval respectively. From statistical tests results, it was found that the computed values of Wald-Wolfowitz runs test and Mann-Whitney U-test using DS1, DS2, DS3 and DS4 series of Pune and Vadgaon Maval are not greater than its theoretical values either at 1% or 5% level, and at this level, the data series used in rainfall data analysis is random as well as homogeneous. The Grubb's test results indicated that there were no outliers in the data series.

Table 3. Statistical tests results for randomness and homogeneity for Pune

Data series	Wald-Wolfowitz runs test				Mann-Whitney U-test			
	Computed	Theoretical	Significance level	Randomness	Computed	Theoretical	Significance level	Homogeneous
DS1	1.053	1.960	5 %	Yes	-1.174	1.960	5 %	Yes
DS2	2.564	2.580	1 %	Yes	0.207	1.960	5 %	Yes
DS3	1.709	1.960	5 %	Yes	-0.496	1.960	5 %	Yes
DS4	1.942	1.960	5 %	Yes	-1.325	1.960	5 %	Yes

Table 4. Statistical tests results for randomness and homogeneity for Vadgaon Maval

Data series	Wald-Wolfowitz runs test				Mann-Whitney U-test			
	Computed	Theoretical	Significance level	Randomness	Computed	Theoretical	Significance level	Homogeneous
DS1	-0.168	1.960	5 %	Yes	-1.261	1.960	5 %	Yes
DS2	-0.561	1.960	5 %	Yes	-2.123	2.580	1 %	Yes
DS3	-1.264	1.960	5 %	Yes	-1.872	1.960	5 %	Yes
DS4	-2.681	1.960	5 %	Yes	-2.845	1.960	5 %	Yes

4.2 Analysis of Results Based on Estimated Rainfall

The parameters of LN2 were determined by three methods and are used for estimation of rainfall for different return periods at Pune and Vadgaon Maval that are given in Tables 5(a & b) and 6(a & b) whereas the plots are shown in Figures 2 and 3. From the estimated rainfalls of Pune and Vadgaon Maval, it was found that:

i) The estimated rainfall increases when data length increases.
ii) The standard error in the estimated rainfall decreases when data length increases while using DS1, DS2, DS3 and DS4 series for estimation of rainfall at Pune and Vadgaon Maval.
iii) The standard errors in rainfall estimates computed by MoM for Pune and MLM for Vadgaon Maval are comparatively less than those values of other methods.
iv) The estimated rainfall by LMO for Pune and MoM for Vadgaon Maval is higher than those values of other methods.

4.3 Analysis of Results Based on GoF Tests

The GoF (viz., χ^2 and *KS*) tests were applied for checking the adequacy of fitting three methods (viz., MoM, MLM and LMO) of LN2 to the rainfall series with different data length of Pune and Vadgaon Maval sites, and are presented in Tables 7 and 8. These results indicated that the computed values are lesser than its theoretical value at 1% significance level, and at this level, all three methods are acceptable for rainfall estimation for Pune and Vadgaon Maval.

4.4 Analysis of Results Based on Diagnostic Tests

The results obtained from three methods of LN2 for Pune and Vadgaon Maval sites were evaluated through diagnostic tests using *CC*, *NSE* and *RMSE*. The diagnostic values were computed by MoM, MLM and LMO of LN2 and the results of Pune and Vadgaon Maval are presented in Tables 9 and 10. From the diagnostic tests results, it is noticed that (i) *CC* obtained through three methods of LN2 vary between 0.993 and 0.996 for Pune while 0.984 to 0.992 for Vadgaon Maval;

(ii) *NSE* in rainfall estimation vary from 98.2 % to 99.0 % for Pune while 94.8 % to 98.2 % for Vadgaon Maval; (iii) *RMSE* values computed by LMO for Pune and MoM for Vadgaon Maval are minimum when compared with those values of other methods considered in rainfall estimation.

4.5 Selection of Method for Estimation of Rainfall

On the basis of analysis of results using GoF and diagnostic tests, it was found that the *RMSE* of LMO for Pune and MoM for Vadgaon Maval are minimum than those values of other methods of LN2 distribution while estimating the rainfall using DS1, DS2, DS3 and DS4. But, as described earlier, the rainfall estimates computed by MoM were considered as less accurate than those values obtained by MLM and LMO though its *RMSE* is noted as minimum. By considering these aspects, after eliminating the MoM from the selection of best method, the MLM is found as the second best next to MoM on the basis of its *RMSE* values for Vadgaon Maval. However, on the basis of *CC* and *NSE* values, LMO was found as superior than MLM and hence LMO is identified as best method for estimation of rainfall at Vadgaon Maval. In light of the above, the LMO of LN2 distribution was adjudged as best suitable method for rainfall estimation of Pune and Vadgaon Maval.

Table 5(a). Estimated 1-day maximum rainfall (mm) with standard error by LN2 distribution using DS1 and DS2 series for Pune

Return period (year)	DS1						DS2					
	MoM		MLM		LMO		MoM		MLM		LMO	
	ER	SE	ER	SE	ER	SE	ER	SE	ER	SE	ER	SE
2	66.5	2.9	66.4	3.0	66.4	3.0	66.8	2.6	66.7	2.6	66.7	2.7
5	86.4	4.4	86.8	4.5	87.2	4.6	88.8	4.0	89.0	4.1	89.3	4.2
10	99.1	5.9	99.8	6.1	100.5	6.2	103.0	5.4	103.4	5.5	104.0	5.6
20	110.9	7.5	112.1	7.7	113.0	7.9	116.4	7.0	117.1	7.1	118.0	7.2
50	125.9	9.8	127.6	10.1	128.9	10.4	133.6	9.2	134.7	9.4	135.9	9.6
100	137.1	11.6	139.2	12.1	140.8	12.4	146.5	11.0	147.8	11.2	149.4	11.5
200	148.1	13.5	150.6	14.1	152.6	14.5	159.4	12.9	161.0	13.2	162.8	13.5
500	162.7	16.2	165.8	16.9	168.3	17.5	176.5	15.6	178.5	16.0	180.8	16.4
1000	173.8	18.4	177.4	19.2	180.2	19.8	189.6	17.8	192.0	18.2	194.6	18.7

Table 5(b). Estimated 1-day maximum rainfall (mm) with standard error by LN2 distribution using DS3 and DS4 series for Pune

Return period (year)	DS3						DS4					
	MoM		MLM		LMO		MoM		MLM		LMO	
	ER	SE	ER	SE	ER	SE	ER	SE	ER	SE	ER	SE
2	68.3	2.3	68.2	2.3	68.2	2.4	70.0	2.2	70.0	2.2	70.0	2.2
5	90.5	3.5	90.8	3.6	91.2	3.7	93.3	3.4	93.4	3.4	93.6	3.5
10	105.0	4.8	105.5	4.8	106.1	4.9	108.4	4.6	108.6	4.6	109.0	4.7
20	118.6	6.1	119.3	6.2	120.3	6.4	122.7	5.9	123.0	6.0	123.5	6.1
50	136.0	8.1	137.1	8.2	138.5	8.4	141.1	7.8	141.5	7.9	142.3	8.0
100	149.1	9.6	150.5	9.9	152.1	10.1	154.8	9.4	155.3	9.5	156.3	9.6
200	162.1	11.3	163.8	11.6	165.8	11.9	168.5	11.0	169.2	11.1	170.3	11.3
500	179.4	13.7	181.6	14.0	184.1	14.4	186.8	13.4	187.7	13.5	189.1	13.7
1000	192.7	15.6	195.2	16.0	198.1	16.4	200.8	15.2	201.8	15.4	203.4	15.6

Table 6(a). Estimated 1-day maximum rainfall (mm) with standard error by LN2 distribution using DS1 and DS2 series for Vadgaon Maval

Return period (year)	DS1						DS2					
	MoM		MLM		LMO		MoM		MLM		LMO	
	ER	SE	ER	SE	ER	SE	ER	SE	ER	SE	ER	SE
2	84.2	4.1	84.5	3.9	84.5	4.0	87.8	3.8	88.2	3.6	88.2	3.6
5	112.5	6.4	111.3	6.0	111.8	6.1	119.6	6.0	117.9	5.6	118.1	5.6
10	131.0	8.6	128.5	8.0	129.4	8.2	140.6	8.2	137.2	7.5	137.6	7.6
20	148.4	11.1	144.7	10.3	146.0	10.5	160.6	10.7	155.5	9.7	156.1	9.8
50	170.9	14.7	165.4	13.5	167.2	13.9	186.6	14.3	179.0	12.8	179.8	13.0
100	187.8	17.7	180.8	16.1	183.1	16.6	206.3	17.2	196.6	15.4	197.7	15.6
200	204.7	20.8	196.1	18.9	198.9	19.5	226.1	20.4	214.2	18.1	215.5	18.3
500	227.2	25.2	216.5	22.7	219.9	23.5	252.7	24.8	237.7	21.9	239.3	22.2
1000	244.4	28.7	232.1	25.8	236.0	26.7	273.1	28.5	255.8	25.0	257.6	25.3

Table 6(b). Estimated 1-day maximum rainfall (mm) with standard error by LN2 distribution using DS3 and DS4 series for Vadgaon Maval

Return period (year)	DS3						DS4					
	MoM		MLM		LMO		MoM		MLM		LMO	
	ER	SE	ER	SE	ER	SE	ER	SE	ER	SE	ER	SE
2	90.3	3.7	91.1	3.4	91.1	3.4	94.6	3.6	95.2	3.4	95.2	3.4
5	126.4	6.0	123.7	5.3	123.6	5.4	132.6	5.8	130.6	5.1	130.8	5.3
10	150.7	8.2	145.2	7.2	145.0	7.3	158.3	8.0	154.0	7.0	154.5	7.1
20	174.2	10.7	165.8	9.4	165.5	9.4	183.1	10.6	176.5	9.2	177.3	9.3
50	205.2	14.3	192.4	12.5	192.0	12.6	215.8	14.1	205.9	12.3	207.0	12.5
100	228.8	17.2	212.5	15.1	212.0	15.1	240.8	17.0	228.0	15.0	229.4	15.2
200	252.8	20.4	232.7	17.8	232.1	17.9	266.2	20.2	250.5	17.5	252.1	17.6
500	285.2	24.3	259.8	21.7	259.0	21.8	300.6	24.0	280.6	21.4	282.7	21.5
1000	310.5	28.4	280.7	24.8	279.8	25.0	327.3	28.3	303.9	24.5	306.3	24.8

Table 7. Theoretical and computed values of GoF tests statistic by LN2 distribution for Pune

Data series	Theoretical value at 1% level	Computed values of χ^2			Theoretical value at 1% level	Computed values of *KS*		
		MoM	MLM	LMO		MoM	MLM	LMO
DS1	18.475	7.200	7.203	7.205	0.231	0.041	0.040	0.044
DS2	26.217	8.000	8.012	8.030	0.188	0.045	0.044	0.047
DS3	18.475	6.000	6.014	6.020	0.163	0.043	0.042	0.045
DS4	23.209	7.556	7.558	7.559	0.151	0.038	0.037	0.039

Table 8. Theoretical and computed values of GoF tests statistic by LN2 distribution for Vadgaon Maval

Data series	Theoretical value at 1% level	Computed values of χ^2			Theoretical value at 1% level	Computed values of *KS*		
		MoM	MLM	LMO		MoM	MLM	LMO
DS1	13.277	3.760	3.766	3.763	0.231	0.069	0.071	0.070
DS2	16.812	9.250	9.258	9.254	0.192	0.083	0.085	0.084
DS3	20.090	16.516	16.519	16.517	0.166	0.093	0.095	0.094
DS4	23.209	19.421	19.425	19.423	0.153	0.070	0.072	0.071

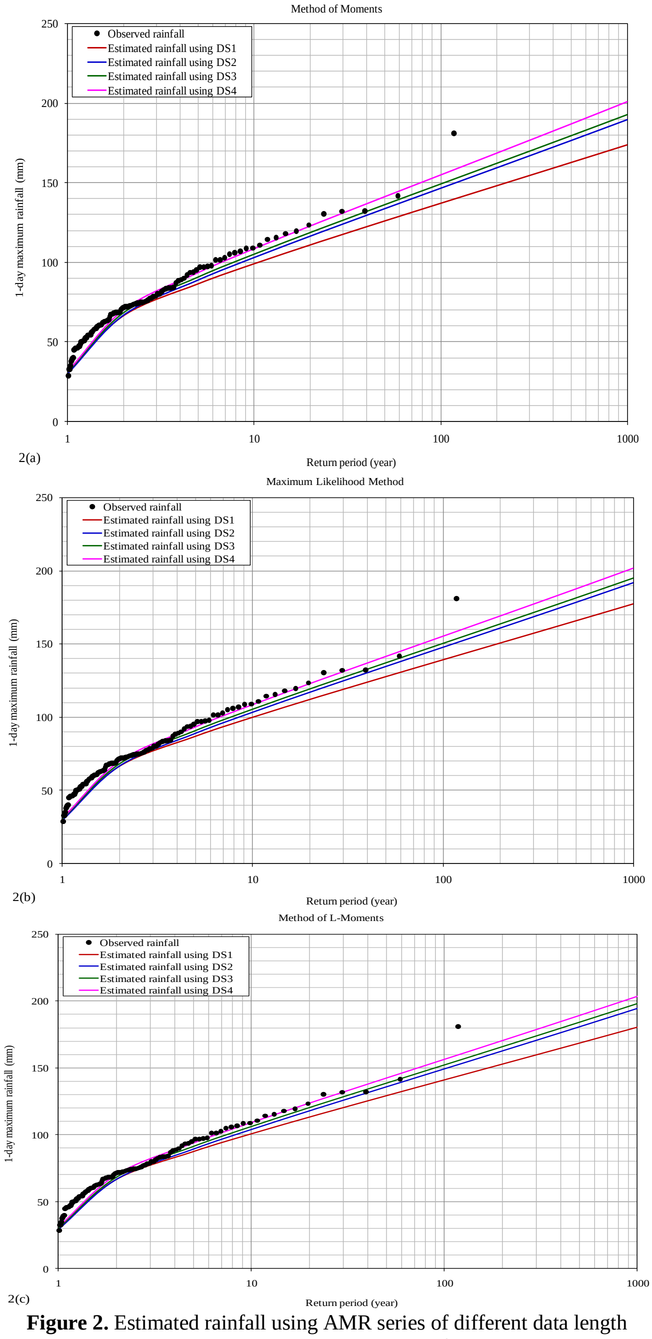

Figure 2. Estimated rainfall using AMR series of different data length by three methods LN2 distribution for Pune

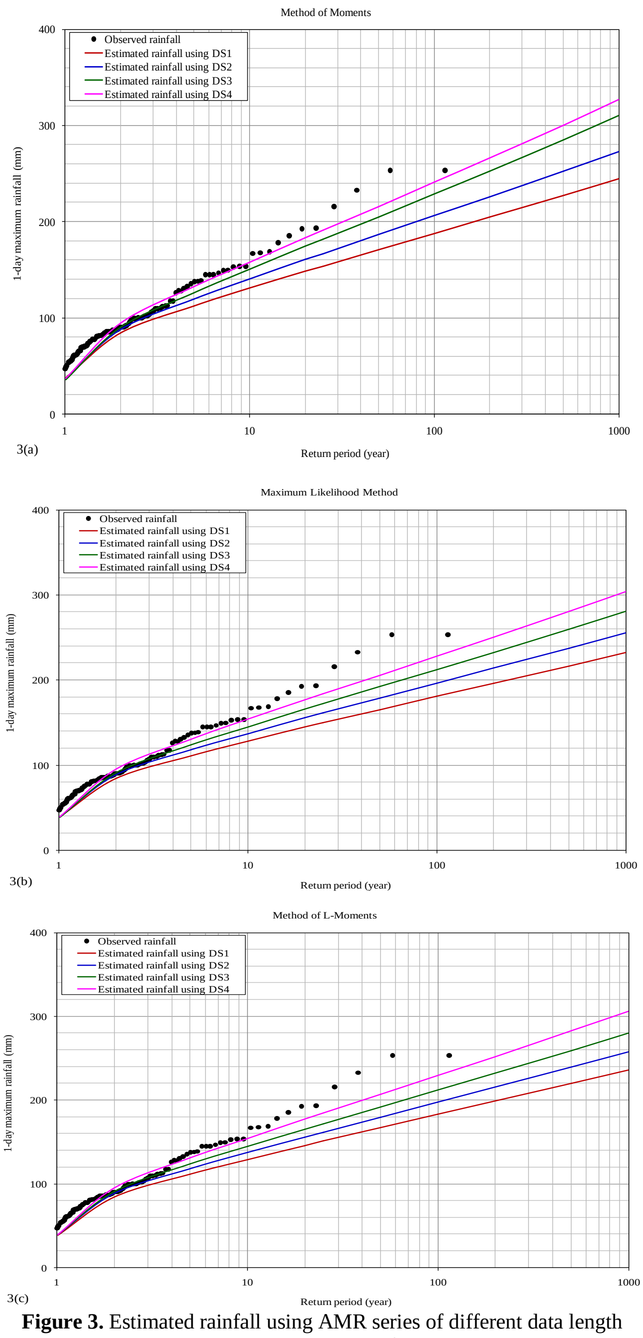

Figure 3. Estimated rainfall using AMR series of different data length by three methods LN2 distribution for Vadgaon Maval

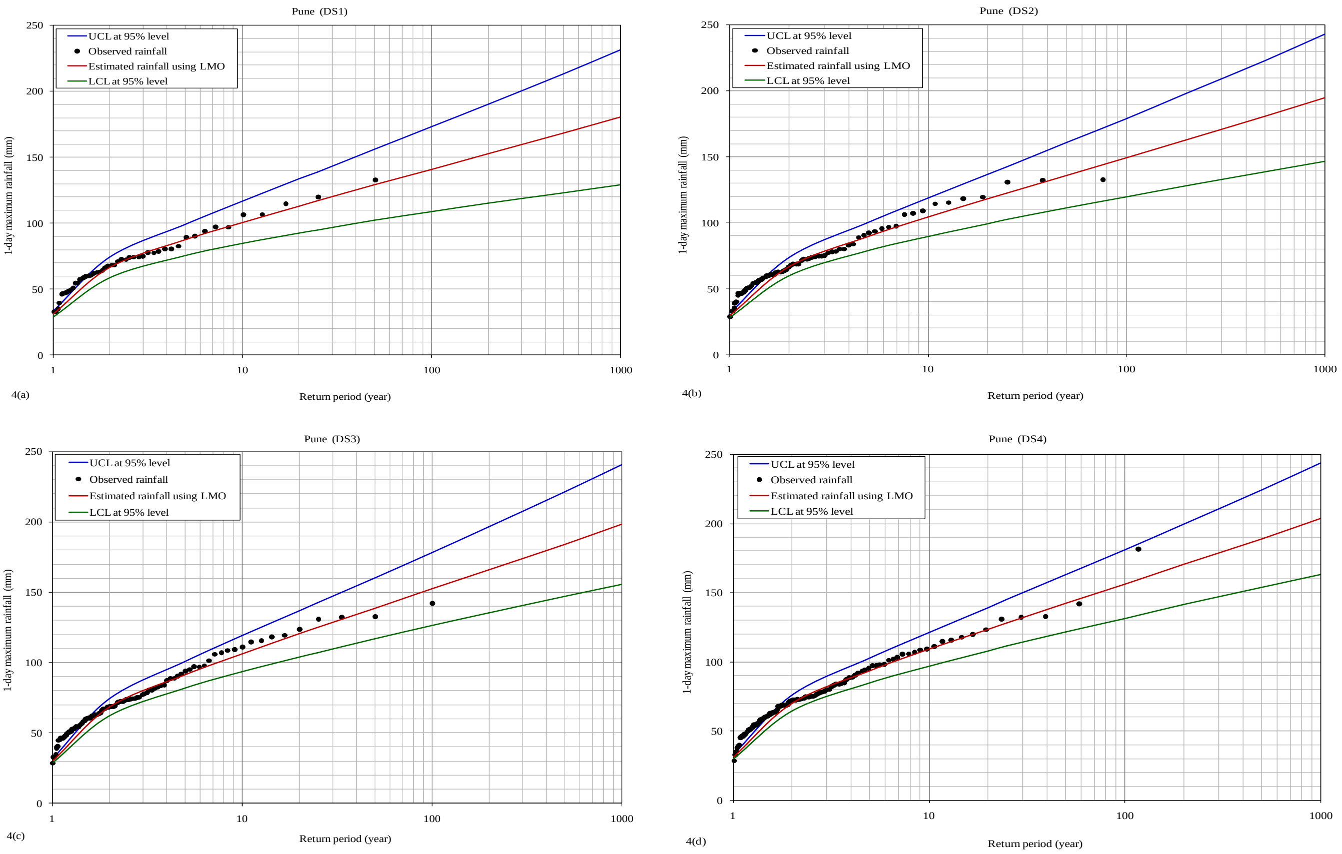

Figure 4. Estimated rainfall together with 99% confidence limits using AMR series of different data length by LMO of LN2 distribution for Pune

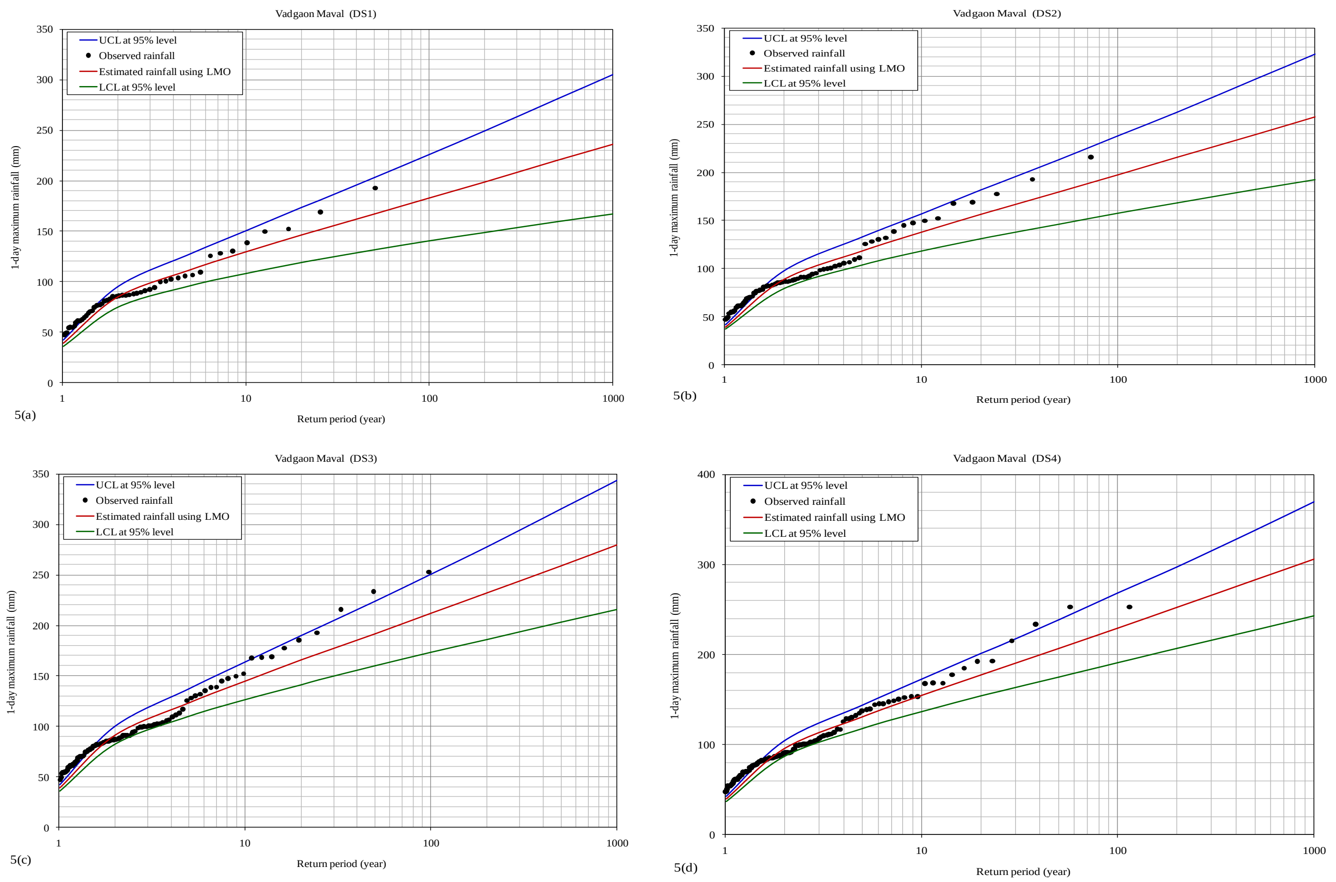

Figure 5. Estimated rainfall together with 99% confidence limits using AMR series of different data length by LMO of LN2 distribution for Vadgaon Maval

Table 9. *CC*, *NSE* and *RMSE* values given by three methods of LN2 distribution for Pune

Data series	MoM			MLM			LMO		
	CC	*NSE (%)*	*RMSE (mm)*	*CC*	*NSE (%)*	*RMSE (mm)*	*CC*	*NSE (%)*	*RMSE (mm)*
DS1	0.996	98.5	2.7	0.996	98.8	2.4	0.996	99.0	2.2
DS2	0.993	98.2	3.2	0.993	98.4	3.1	0.993	98.5	3.0
DS3	0.995	98.8	2.7	0.995	98.9	2.6	0.995	99.0	2.5
DS4	0.996	98.9	2.7	0.996	99.0	2.6	0.996	99.0	2.5

Table 10. *CC*, *NSE* and *RMSE* values given by three methods of LN2 distribution for Vadgaon Maval

Data series	MoM			MLM			LMO		
	CC	*NSE (%)*	*RMSE (mm)*	*CC*	*NSE (%)*	*RMSE (mm)*	*CC*	*NSE (%)*	*RMSE (mm)*
DS1	0.989	97.3	5.2	0.988	96.1	6.2	0.989	96.5	5.9
DS2	0.989	97.4	5.7	0.987	96.0	7.1	0.988	96.2	6.9
DS3	0.987	97.1	6.9	0.984	94.8	9.2	0.984	94.8	9.3
DS4	0.992	98.2	5.8	0.991	96.8	7.6	0.991	97.0	7.4

The plots of observed and estimated rainfall by LMO of LN2 distribution with 99% confidence limits for different return periods using DS1, DS2, DS3 and DS4 series for Pune and Vadgaon Maval are shown in Figures 4 and 5. From these figures, it was witnessed that the range of observed AMR viz., 70 to 180 mm of Pune and 80 to 240 mm of Vadgaon Maval are within the confidence limits of the estimated rainfall using LMO of LN2. From Figures 2 to 5, it was also witnessed that the fitted curves using MoM, MLM and LMO of LN2 are in the form of linear curves for Pune and Vadgaon Maval.

5. Conclusion and Future Scope

The paper presented a study on effect of data length on assessment of uncertainty of error in rainfall estimation using MoM, MLM and LMO methods of LN2 distribution for Pune and Vadgaon Maval rain-gauge sites of Maharashtra. The data series with different data length viz., DS1, DS2, DS3 and DS4 was used. The GoF tests (viz., χ^2 and *KS*) results indicated that all three methods of LN2 are acceptable for determination of parameters of LN2 that are used in rainfall estimation. The study showed that the estimated rainfall increases when data length increases whereas the standard error in the estimated rainfall decreases when data length increases while using DS1, DS2, DS3 and DS4 series for estimation of rainfall at Pune and Vadgaon Maval. Moreover, the standard errors in rainfall estimates computed by MoM for Pune and MLM for Vadgaon Maval are comparatively less than those values of other methods. The study also showed that the estimated rainfall obtained through LMO for Pune and MoM for Vadgaon Maval are higher than those values other two methods. From the diagnostic tests results, it was noticed that there is generally good correlation between the observed and estimated values using MoM, MLM and LMO of LN2. However, on the basis of *CC*, *NSE* and *RMSE* values, LMO was found superior than MoM and MLM, and hence LMO is adjudged as best method for estimation of rainfall at Pune and Vadgaon Maval. The results presented in this paper would be helpful to the stake holders for arriving at a design value for planning and design of civil engineering infrastructure projects in the study area.

The study focussed on three parameter estimation methods of LN2 for rainfall estimation. However, as a part of the research work in future, it is suggested that the study can possibly be carried out with rainfall series of different data length by adopting standard parameter estimation procedures of other distributions like EV1, GEV, LP-III, GAM, etc that are commonly used for estimation of rainfall.

Data Availability

The rainfall data of Pune and Vadgaon Maval was collected from IMD, Pune and used in this paper.

Conflict of Interest

The author does not have any conflict of interest.

Funding Source

None.

Author Contribution

The author has conducted the literature survey, formulated the methodology, processed and analyzed the rainfall data, carried out the study and prepared the research paper.

Acknowledgement

The author is thankful to the Director, Central Water and Power Research Station, Pune for providing the research facilities to carry out the study.

References

[1]. International Atomic Energy Agency (IAEA), "Meteorological events in site evaluation for nuclear power plants-IAEA Safety Guide No. Ns-G-3.4", International Atomic Energy Agency, Vienna, **2003**.

[2]. A. Montanari and A. Brath, "A stochastic approach for assessing the uncertainty of rainfall-runoff simulations", *Water Resources Research*, Paper ID. W0110640, Vol. **40**, Issue **1**, pp. **1-11**, **2004**. https://doi.org/10.1029/2003WR002540

[3]. B. Renard, D. Kavetski, G. Kuczera, M. Thyer and S.W. Franks, "Understanding predictive uncertainty in hydrologic modelling: The challenge of identifying input and structural errors", *Water Resources Research*, Paper ID. W05521, Vol. **46**, Issue **5**, pp. **1-22, 2010**. https://doi.org/10.1029/2009WR008328.

[4]. Y. Tung and C. Wong, "Assessment of design rainfall uncertainty for hydrologic engineering applications in Hong Kong", *Stochastic Environmental Research and Risk Assessment*, Vol. **28**, Issue **3**, pp. **583-592**, **2014**. https://doi.org/10.1007/s00477-013-0774-2

[5]. C.M. Tfwala, L.D. Van Rensburg, R. Schall, S.M. Mosia, and P. Dlamini, "Precipitation intensity-duration-frequency curves and their uncertainties for Ghaap plateau", *Climate Risk Management*,

Vol. **16**, Issue **1,** pp. **1-9, 2017**. https://doi.org/10.1016/j.crm. 2017.04.004.

[6]. M.N. Ibrahim, "Assessment of the uncertainty associated with statistical modelling of precipitation extremes for hydrologic engineering applications in Amman, Jordan", *Sustainability*, Paper ID. 17052, Vol. **14**, Issue **24**, pp. **1-20**, **2022**. https://doi.org/ 10.3390/ su142417052.

[7]. K. Haddad and A. Rahman, "Design rainfall estimation and changes, In Handbook of Engineering Hydrology: Modelling", *Climate Change and Variability*, Edited by Saeid Eslamian, CRC Press, pp. **173-190**, **2014**.

[8]. T.T. Hailegeorgis, S.T. Thorolfsson and K. Alfredsen, "Regional frequency analysis of extreme precipitation with consideration to uncertainties to update IDF curves for the city of Trondheim", *Journal of Hydrology*, Vol. **498**, Issue **1**, pp. **305-318**, **2013**. https://doi.org/10.1016/ j.jhydrol. 2013.06.019

[9]. H.K. Cho, K.P. Bowman and G.R. North, "A Comparison of Gamma and Lognormal distributions for characterizing satellite rain rates from the tropical rainfall measuring mission", *Journal of Applied Meteorology*, Vol. **43**, Issue **11**, pp. **1586-1597**, **2004**. https://www.jstor.org/stable/26186015

[10]. R. Chifurira and D. Chikobvu, "Extreme Rainfall: Candidature probability distribution for Mean annual rainfall data", *Research Journal of Pharmaceutical, Biological and Chemical Sciences*, Vol. **8**, Issue **3**, pp. **237-248**, **2017**.

[11]. M.T Amin, M. Rizwan and A.A. Alazba, "A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan", *Open Life Sciences*, Special Issue on CleanWAS 2015, Vol. **11**, Issue **1**, pp. **432-440**, **2016**. https://doi.org/10.1515/biol-2016-0057

[12]. E. Rosmaini and Y.Y. Saphira, "Probability distribution of rainfall in Medan", *Journal of Research in Mathematics Trends and Technology*, Vol. **1**, Issue **2**, pp. **39-45**, **2019**. https://doi.org/ 10.32734/jormtt.v1i2.2835

[13]. M.M. Hasan, B.F.W. Croke, S. Liu, K. Shimizu and F. Karim, "Using mixed probability distribution functions for modelling non-zero sub-daily rainfall in Australia", *Geosciences*, Vol. **10**, Issue **2**, pp. **1-11**, **2020**. https://doi.org/10.3390/geosciences10020043

[14]. N. Vivekanandan, "Evaluation of estimators of probability distributions for frequency analysis of rainfall and river flow data", *International Journal of Scientific Research in Civil Engineering*", Vol. **4**, Issue **4**, pp. **67-75**, **2020**. http://ijsrce.com/ IJSRCE204412

[15]. L. Kaur, Anvesha, M. Kumar, S.L. Verma and P. Kumar, "Annual maximum rainfall prediction using frequency analysis for Roorkee, Uttarakhand, India", *MAUSAM*, Vol. **72**, Issue **2**, pp. **359-372**, **2021**. https://doi.org/10.54302/mausam.v72i2.623

[16]. N. Bhagat, "Flood frequency analysis using Gumbel's distribution method: a case study of lower Mahi Basin, India", *Journal of Water Resources and Ocean Science*, Vol. **6**, Issue **4**, pp. **51-54**, **2017**. https://doi.org/10.11648/j.wros.20170604.1

[17]. R. Kumar, "Flood frequency analysis of the Tapti river basin using log Pearson type-III and Gumbel extreme value-1 methods", *Journal of the Geological Society of India*, Vol. **94**, No. **5**, pp. **480-484**, **2019**. https://doi.org/10.1007/s12594-019-1344-0

[18]. A.O. Opere and A.K. Njogu, "Using extreme value theory to estimate available water in the upper Awach-Kibuon catchment in Nyamira County, Kenya", *American Journal of Water Resources*, Vol. **8**, Issue **4**, pp. **200-210**, **2020**. https://doi.org/10.12691/ajwr-8-4-6

[19]. V. Manohar Reddy, "Flood frequency analysis using Gumbel's distribution method and Log-Pearson Type III distribution method in Krishna River, Andhra Pradesh", *Proceedings of Recent Developments in Sustainable Infrastructure and Part of the Lecture Notes in Civil Engineering Book Series*, Vol. **207**, **2022**. https://doi.org/10.1007/978-981-16-7509-6_36

[20]. J. Prabhu, T. George, B. Vijayakumar and P.M. Ravi, "Extreme value statistical analysis of meteorological parameters observed at Kudankulam site during 2004-2014", *Radiation Protection Environment*, Vol. **39**, Issue **2**, pp. **107-112**, **2016**. https://doi.org/10.4103/ 0972-0464.190388

[21]. J.R.M. Hosking, "L-moments: Analysis and estimation of distributions using linear combinations of order statistics", *Journal of Royal Statistics Society: Series B (Statistical Methodology)*, Vol. **52**, Issue **1**, pp. **105-124**, **1990**. https://doi.org/10.1111/j.2517-6161.1990.tb01775.x.

[22]. D. Bilkova, "Lognormal distribution and using L-moment method for estimating its parameters", International Journal of Mathematical Models and Methods in Applied Sciences, Vol. **6**, Issue **1**, pp. **30-44**, **2012.**

[23]. J. Zhang, "Powerful goodness-of-fit tests based on the likelihood ratio", *Journal of Royal Statistical Society: Series B (Statistical Methodology)*, Vol. **64**, Issue **2**, pp. **281-294**, **2002**. https://doi.org/10.1111/1467-9868.00337

[24]. N. Vivekanandan, F.T. Mathew and S.K. Roy, "Modelling of wind speed data using probabilistic approach", *Journal of Power and River Valley Development*, Vol. **62**, Issue **3-4**, pp. **42-45**, **2012**.

[25]. N. Vivekanandan, C. Srishailam and R.G. Patil, "Intercomparison of MoM, MLM and LMO estimators of probability distributions for assessment of extreme rainfall", In: Timbadiya PV, Singh VP and Sharma PJ (eds) Climate Change Impact on Water Resources. Hydro 2021. Lecture Notes in Civil Engineering, Vol. 313, Springer, Singapore, 2023. https://doi.org/10.007/978-981-19-8524-9_34.

AUTHOR PROFILE

N. Vivekanandan has post graduated in mathematics from Madurai Kamaraj University. In addition, he received M.Phil. (Mathematics), M.E. (Hydrology), M.B.A. (Human Resources). He is working as a Scientist of Central Water and Power Research Station, Pune and has more than 30 years experience in R&D. He has published many research papers in national/ international journals and conferences. His research interests include applied statistics, applied hydrology, irrigation planning, soft computing, climate change, etc. He has received a fellowship from Ministry of Water Resources for PG Diploma programme in Hydrology in 1999, best paper award from Institute of Engineers (India) in 2008, best paper award from Bal Krishna Institute of Technology in 2016, award from Vellore Institute of Technology in 2017 and Proudhadeveraya Institute of Technology in 2019.


Int. J. of Scientific Research in Biological Sciences
www.isroset.org
Int. J. of Scientific Research in Chemical Sciences
www.isroset.org
Int. J. of Scientific Research in Computer Science and Engineering
www.isroset.org
World Academics Journal of Engineering Sciences
ISSN:2348-635X
www.isroset.org
Journal of Physics and Chemistry of Materials
ISSN:2348-6341
www.isroset.org
Int. J. of Scientific Research in Mathematical and Statistical Sciences
www.isroset.org
International Scientific Research Organization for Science, Engineering and Technology
ISROSET
Submit your manuscripts at
www.isroset.org
email: support@isroset.org
Make a Submission
ISSN: 2349-3178 (Print), ISSN: 2349-3186 (Online)
International Journal of Medical Science Research and Practice
Published by ISROSET
Int. J. of Scientific Research in Multidisciplinary Studies
www.isroset.org
Int. J. of Scientific Research in Network Security and Communication
e-ISSN: 2321-3256
World Academics Journal of Management
ISSN: 2321-905X
www.isroset.org
Int. J. of Scientific Research in Physics and Applied Sciences
www.isroset.org
Int. J. of Computer Sciences and Engineering
www.ijcseonline.org