J. Indian Water Resour. Soc.,
Vol. 42, No. 1, Jan., 2022

A COMPARATIVE STUDY ON THE PERFORMANCE OF PROBABILITY DISTRIBUTIONS WITH CONVENTIONAL AND L-MOMENTS FOR ANNUAL MAXIMUM DAILY STREAMFLOW AT THE GAUGING SITES OF GODAVARI RIVER SUB-BASINS

D Srinivasa Murthy[1], S Aruna Jyothy[2] and P Mallikarjuna[3]

ABSTRACT

Estimation of streamflow is essential for effective planning and management of water resources of a river basin. The present study aims at the probability distribution analysis of annual maximum daily streamflow at the gauging sites of Godavari sub-basins. The analysis indicated that lognormal or gamma distribution with conventional moments fitted the maximum daily streamflow data at the gauging sites of Godavari sub-basins. Among 2-parameter distributions with L-moments, Generalized Pareto followed by Gamma/Lognormal fitted annual maximum daily streamflow data at the upstream gauging sites of sub-basins. At the downstream-most gauging sites of Pranahitha, Indravathi and Godavari lower sub-basins, the data followed Weibull probability distribution. Among 3-parameter distributions with L-moments, Generalized Pareto at seven gauging sites, Weibull and Pearson Type III at five gauging sites each, Generalized Logistic at four gauging sites and Generalized Extreme Value at two gauging sites fitted maximum daily streamflow data.

Based on the performance evaluation of probability distributions with conventional and L-moments, 2 – parameter distributions with L-moments at the upstream, 3 – parameter distributions at the middle and, probability distributions with conventional moments at the downstream gauging sites performed better in the Godavari upper and middle sub-basins. Probability distributions with conventional moments and 3-parameter distributions with L-moments fitted the annual maximum daily streamflow data at the gauging sites of Pranahitha, Indravathi and Godavari lower sub-basins satisfactorily.

Keywords: *Annual maximum daily streamflow - Conventional moments - Statistical tests - L-moments - Performance evaluation*

INTRODUCTION

Streamflow in a basin represents an integrated response to catchment heterogeneity and spatial variability of key hydrological processes such as precipitation, infiltration and evapotranspiration, and provides an insight into long-term hydro-climatic changes. Maximum streamflow (flood) analysis plays an important role in hydrologic and economic evaluation of water resources projects and its prediction is useful for the design of hydraulic structures and for flood management studies. Probability analysis of streamflow provides means to capture its statistical structure and suggests appropriate distributions.

Durrans et al. (2003) presented frequency analysis of streamflow in U.S. Tennessee Valley using Log-Pearson Type III distribution (LP3). Kumar et al. (2003) carried out regional flood frequency analysis based on L-moments and concluded that Generalized Extreme Value (GEV) distribution is the robust distribution at the sites in the middle Ganges plains of India. Yue and Wang (2004) applied the method of L-moments to identify the probability distribution of annual streamflow in different climatic regions of Canada and recommended different probability distributions for different regions. Atiem and Harmancioglu (2006) derived hydrologically homogeneous regions and identified the regional statistical distributions for streamflows at the gauging sites on the Nile River tributaries and showed that hydrologically homogeneous region followed Generalized Logistic (GLO) distribution. Gamage (2006) made a study to evaluate the goodness-of-fit of alternative probability distributions to sequences of annual maximum streamflow in Sri Lanka through L-moment ratio diagrams. The study revealed that annual maximum streamflow is best approximated by GEV distribution. Abida and Ellouze (2008) identified regional flood frequency distributions for the sites in different flood zones of Tunisia. Flood data in Northern Tunisia was observed to follow Generalized Normal distribution while the Generalized Normal and GEV distributions were the best-fit distributions at the sites in central and southern Tunisia respectively. Haddad and Rahman (2008) compared a number of distributions for the catchments in Australia and found that GEV distribution is the best distribution. Bettil Saf (2009) derived regional flood frequency estimates for the gauged sites in West Mediterranean River Basins in Turkey and identified P3 distribution for the Antalya and Lower-West Mediterranean sub-regions and GLO for the Upper-West Mediterranean sub-region as the best-fit distributions. Haddad and Rahman (2010) found that two-parameter distributions were preferable to three-parameter distributions for Tasmania in Australia with lognormal

1. *Assistant Professor, Department of Civil Engineering, Sree Vidyanikethan Engineering College, A. Rangampet, Andhra Pradesh, India.*
Corresponding Author E-mail: srinivasskit@gmail.com.
2. *Assistant Professor, Department of Civil Engineering, S.V.U College of Engineering, Tirupati-517502,Andhra Pradesh, India.*
Email : aruna_jyothy@yahoo.co.in
3. *Professor(Retired), Department of Civil Engineering, S.V.U College of Engineering, Tirupati-517502, Andhra Pradesh, India.*
Email: mallikarjuna_perugu@yahoo.co.in
Manuscript No. 1566
Received: 5 February 2022 ; Accepted : 20 August 2022

appeared to be the best-selection by examining seven different probability distributions.

Hussain and Pasha (2011) carried out flood frequency analysis at seven stations located on the main stream of Indus river and found P3 and GLO as the robust distributions. Gubareva (2012) compared P3, lognormal, GEV and Generalized Pareto (GPA) distributions in the estimation of maximum flows in the rivers of Austria and Siberia and concluded that P3 distribution was observed to be the best-fit distribution. Ayesha Rahman et al. (2013) found that a single distribution couldn't be specified as the best-fit distribution for flood flows in Australian states and identified that LP3, GEV and GPA distributions as the top three best-fit distributions. Mamman et al. (2017) fitted various probability distribution models to river flows of the Kainji reservoir in New Busca, Niger state, Nigeria and recommended the Gumbel probability model. Drissia et al. (2019) carried out flood frequency analysis at regional level in Kerala, India and found that GPA and GLO were the best-fit distributions for the stations in the study area.

Most of the earlier studies recommended probability distributions using either conventional or L-moments for streamflows at the gauging sites in different regions. In the present investigation, probability distribution analysis of annual maximum daily streamflow at the selected gauging sites of Godavari sub-basins has been carried out using both conventional and L-moments and compared their performance in the selection of suitable distributions.

MATERIALS AND METHODS

The river, Godavari, originating in the Western Ghats in Nasik district of Maharashtra flows in the south easterly direction and joins Bay of Bengal. The river basin with tributaries of Purna, Manjira, Penganga, Wardha, Waingangā, Pranahitha, Indravathi and Sabari spreads over an area of 3,12,812 km². The daily streamflow data at Chass, Ashwi and Pachegaon of Godavari upper; Manjalegaon, Dhalegaon, Zari, GR Bridge, Purna and Yelli of Godavari middle; Gandlapet, Mancherial, Somanpally and Perur of Pranahitha; Pathagudem, Chindnar, Sonarpal, Jagdalpur and Nowrangpur of Indravathi; Sardaput, Injaram, Konta, Koida and Polavaram of Godavari lower sub-basins for the period varying between 1965-2011 were collected from Central Water Commission (CWC) and used in the analysis. The location map of gauging sites is shown as Fig.1 and a brief description of the sites is presented in Table 1.

Fig.1: Location map of gauging sites

Normal, Lognormal, Exponential, Extreme Value and Gamma distributions with conventional moments were employed for fitting the streamflow series. A brief description of these probability distributions is presented in Table 2.

Table 1: Brief description of gauging sites

Sub-basin	Gauging site	Latitude °N	Longitude °E	Catchment area (km²)	Mean annual maximum daily streamflow (cumec)
Godavari Upper	Chass	19° 57' 20"	74° 19' 15"	5230	1253.49
	Ashwi	19° 33' 00"	74° 36' 00"	1820	318.492
	Pachegaon	19° 32 '07"	74° 50' 01"	5800	563.08
Godavari Middle	Manjalegaon	19° 10' 00"	76° 15' 00"	3960	1126.70
	Dhalegaon	19° 13' 13"	76° 21' 52"	30840	2379.55
	Zari	19° 23' 43"	76° 46' 15"	5550	930.65
	GR Bridge	19° 01' 20"	76° 43' 45"	33934	2029.33
	Purna	19° 10' 33"	77° 00' 50"	15000	2801.51
	Yelli	19° 02' 38"	77° 27' 10"	53630	4210.67
Pranahitha	Gandlapet	18° 49' 16"	78° 26' 17"	1360	480.32
	Mancherial	18° 50' 00"	79° 27' 00"	102900	10038.33
	Somanpally	18° 38' 30"	79° 49' 35"	12691	1381.56
	Perur	18° 33' 00"	80° 22' 00"	268200	29943.64

Indravathi	Pathagudem	18° 49' 00"	80° 21' 00"	40000	14532.79
	Chindnar	19° 05' 00"	81° 18' 00"	17270	5332.89
	Sonarpal	19° 16' 00"	81° 52' 00"	1523	811.45
	Jagdalpur	19° 06' 30"	82° 01' 30"	7380	1768.98
	Nowrangpur	19° 12' 00"	82° 31' 00"	3545	1346.12
Godavari lower	Sardaput	18° 36' 00"	82° 08' 00"	4800	2657.09
	Injaram	17° 50' 00"	81° 23' 00"	12925	5375.50
	Konta	17° 14' 45"	81° 39' 35"	19550	5972.11
	Koida	17° 48' 00"	81° 23' 00"	305460	31343.01
	Polavaram	17° 14' 45"	81° 39' 35"	307800	31600.84

Table 2: Brief description of probability distributions with conventional moments (Ven Te Chow 1988)

Distribution	**Probability density function**	**Parameters of the distribution**	**Equations for parameters in terms of sample moments**
Normal	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$	μ, σ	$\mu = \bar{x}, \sigma = s_x$
Lognormal	$f(x) = \frac{1}{x\sigma_y\sqrt{2\pi}} \exp(-\frac{(y-\mu_y)^2}{2\sigma_y^2})$ where $y = log\ x$	μ_y, σ_y	$\mu_y = \bar{y}, \sigma_y = s_y$
Extreme value Type I	$f(x) = \frac{1}{\alpha} \exp[-\frac{x-u}{\alpha} - \exp(-\frac{x-u}{\alpha})]$ $u = \bar{x} - 0.5772\alpha$	α	$\alpha = \frac{\sqrt{6}s_x}{\pi}$
Exponential	$f(x) = \lambda e^{-\lambda x}$	λ	$\lambda = \frac{1}{\bar{x}}$
Gamma	$f(x) = \frac{\lambda^{\beta} x^{\beta-1} e^{-\lambda x}}{\Gamma(\beta)}$ where $\Gamma = gamma\ function$	λ, β	$\lambda = \frac{\bar{x}}{s_x^2}$ $\beta = \frac{\bar{x}^2}{s_x^2}$

ẍ = Sample mean, s_x = Sample standard deviation

Statistical tests such as Kolmogorov-Smirnov (K-S), Anderson-Darling (A-D) and Chi-square (χ^2) were used to assess the reasonableness of the selected distribution. A general brief description of these tests is presented below.

Kolmogorov – Smirnov (K-S) test

$x_1, x_2, \ldots, x_m, \ldots, x_n$, the ordered values of the random variable in a sample of size *n*, are arranged in the descending order of magnitude. The cumulative probability $P(x_m)$ for each observation, x_m is computed using Weibull's formula, $T = \frac{(n+1)}{m}$, where, *n* is sample size and *m* is the order or rank of the observation. Theoretical cumulative probability $F(x_m)$ for each ordered observation, x_m is calculated using the assumed distribution. The absolute difference between $P(x_m)$ and $F(x_m)$ is computed for each x_m. The Kolmogorov - Smirnov statistic, Δ is the largest value of these absolute differences.

$$\Delta = \text{maximum of } |P(x_m) - F(x_m)| \quad (1)$$

The statistic thus obtained is compared with the critical value, Δ_0 at 95% level of significance. If $\Delta < \Delta_0$, the hypothesis that the distribution is a good fit, is accepted. If more than one distribution passes the test, then the distribution which gives the least value of Δ is taken to be the most appropriate choice.

Anderson– Darling (A-D) test

The Anderson-Darling test evaluates whether a sample of data originates from a population with a specific distribution. The test statistic is expressed as

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\ln F(x_i) - \frac{1}{n}\sum_{i=1}^{n} ln(1 - F(x_{n+1-i})) \quad (2)$$

where, $F(x_i)$ is the CDF of the specific distribution at x_i, for i=1,2,….n. The null hypothesis is that the data sampled from a specific distribution is rejected if the test statistic, A, is greater than the critical value of 0.787 at 5% level of significance.

Chi – Square (χ^2) test

Chi–Square (χ^2) test statistic is based on the comparison of number of observed events and number of expected events of a specific probability distribution in class intervals covering the range of the data. This test is for continuous sample data and is used to determine if a sample is originated from the population with a specific distribution. The critical value of χ^2- test statistic, ${\chi^2}_c$ is given by

$$ {\chi^2}_c = \sum_{i=1}^{m} \frac{n[f(x_i)-p(x_i)]^2}{p(x_i)} \tag{3} $$

where, m is number of intervals, $nf(x_i) = n_i$, observed number of occurrences in an interval i, and $np(x_i)$ is the corresponding expected number of occurrences in an interval i.

The null hypothesis for the test is that the proposed probability distribution fits the data adequately. This hypothesis is rejected if the value of ${\chi^2}_c$, determined from χ^2 distribution with ν (= m-p-1) degrees of freedom, is larger than the limiting value at the chosen level of significance.

Commonly adopted 2-parameter distributions such as Generalized Pareto (GPA2), Log-normal (LN2), Gamma (GAM2) and Weibull (W2) and, 3-parameter distributions such as Generalized Extreme Value (GEV), Generalized Pareto (GPA3), Generalized Logistic (GLOG), Log-Normal (LN3), Pearson (P3) and Weibull (W3) using L-moments (Hosking, 1990) were selected to fit the streamflow series. The performance measure in terms of the deviation between observed and computed L-moment ratios (L-coefficient of variation in case of 2- parameter and L-kurtosis in case of 3-parameter distributions) is considered in the selection of appropriate probability distributions. A brief description of commonly adopted 2-parameter and 3-parameter probability distributions based on L-moments is presented in Table 3 and Table 4 respectively.

Table 3: Brief description of 2 - parameter distributions with L-moments (Hosking, 1990)

Distribution	Distribution function	Parameters of the distribution	Equations for parameters
GPA2	$X_T = \{\xi - \alpha \ln(1-(1-1/T))\}$	α, ξ	$\alpha = \lambda_2(1+k)(2+k)$ $k = \frac{1-3\tau_3}{1+\tau_3}$ $\xi = \lambda_1 + \lambda_2(k+2)$ L-skewness (τ_3) = λ_3 / λ_2 Where λ_1 = First L-Moment = β_0, λ_2 = Second L-Moment = $2\beta_1 - \beta_0$, λ_3 = Third L-Moment = $6\beta_2 - 6\beta_1 + \beta_0$, where, β_r (r = 0, 1, 2, 3) are the probability weighted moments given by $\beta_r = n^{-1}\sum_{i=r+1}^{n}\binom{j-1}{r}\binom{n-1}{r}^{-1}.x(j,n)$, r = 0, n -1
LN2	$X_T = exp(\mu_Y + \sigma_Y z_T)$ $w = \sqrt{-2\ln(1-F)}$, (1 > F ≥ 0.5) When F<0.5, F is substituted for (1-F) and the value of z is given a negative sign, F= 1-1/T $Z_T = \emptyset^{-1}(F)$ $= w - \frac{2.515517 + 0.802853w + 0.010328w^2}{1 + 1.432788w + 0.189269w^2 + 0.001308w^3}$	μ_y, σ_y	$\mu_y = \ln\lambda_1 - \frac{\sigma_y^2}{2}$ $\sigma_y = 2erf^{-1}\left(\frac{\lambda_2}{\lambda_1}\right) = 2erf^{-1}(\tau_2)$
GAM	$X_T = \lambda_1 + K_T\sqrt{\alpha^2\beta}$ $C_s = \frac{\alpha}{\lvert\alpha\rvert}\frac{2}{\sqrt{\beta}}$ $K_T = \frac{2}{C_s}\left\{\left[\frac{C_s}{6}\left(z - \frac{C_s}{6}\right)+1\right]^3 - 1\right\}$, if $C_s > 0$	α,β	$\alpha = \frac{\lambda_1}{\beta}$ $\beta = \frac{1-0.3080z}{z-0.5812z^2+0.01765z^3}$, if $0<\tau_2<0.5$ $= \frac{0.7213z-0.5947z^2}{1-2.1817z+1.2113z^2}$ if $0.5 \le \tau_2 < 1$ $z = \begin{cases}\pi{\tau_2}^2, & if\ 0<\tau_2<0.5\\ 1-\tau_2 & if\ 0.5 \le \tau_2 < 1\end{cases}$
W2	$X_T = \alpha\{-ln(1-1/T)\}^{1/k}$	k, α	$k = \frac{-ln2}{ln\left[1-\frac{\lambda 2}{\lambda 1}\right]}$, $\alpha = \frac{\lambda 1 - \xi}{\Gamma\left[1+\frac{1}{k}\right]}$ $\xi = \lambda_2 - (\alpha\,\Gamma(1+1/k))$

Table 4: Brief description of 3 - parameter distributions with L-moments (Hosking, 1990)

Probability distribution	Distribution function	Parameters of the distribution	Equations for parameters
GEV	$x_T = \xi + \frac{\alpha}{k}\{1-(-\ln(1-1/T))^k\}$, k≠0 $= \xi - \frac{\alpha}{k}\{ln(-\ln(1-1/T))^k)\}$, k=0	k, α, ξ	$k = 7.8590c + 2.9554c^2 \;\; where, c = \frac{2}{3+\tau_3} - \frac{\ln 2}{\ln 3}$ $\alpha = \frac{\lambda_2 k}{\Gamma(1+k)(1-2^{-k})}$ $\xi = \lambda_1 + \frac{\alpha}{k}[\Gamma(1+k)-1]$
GLOG	$x_T = \xi + \frac{\alpha}{k}\left\{1-\left[\frac{1}{T-1}\right]^k\right\}$, k≠0 $= \xi - \frac{\alpha}{k} ln\left\{\left[\frac{1}{T-1}\right]^k\right\}$, k=0	k, α, ξ	$k = -\tau_3$ $\alpha = \frac{\lambda_2}{\Gamma(1+k)\Gamma(1-k)}$ $\xi = l_1 + \frac{(\lambda_2-\alpha)}{k}$
GPA3	$x_T = \xi + \frac{\alpha}{k}\{[1/T]^k\}$, k≠0 $= \xi - \frac{\alpha}{k} ln\{[1-1/T] \;\; \}$, k=0	k, α, ξ	$k = \frac{1-3\tau_3}{1+\tau_3}$ $\alpha = \lambda_2(1+k)(2+k)$ $\xi = \lambda_1 - \lambda_2(2+k)$
LN3	$X_T = a + exp(\mu_Y + u\sigma_Y)$	α, μ_y, σ_y	$u = w - \frac{2.515517 + 0.802853w + 0.010328w^2}{1 + 1.432788w + 0.189269w^2 + 0.001308w^3}$ $F = \frac{1+\tau_3}{2}$, where τ_3 $is\ L-skewness$ $w = \sqrt{-2\ln(1-F)}$ $a = \lambda_1 - exp\left(\mu_y + \frac{\sigma_y^2}{2}\right)$ $\mu_y = \ln\left[\frac{\lambda_2}{erf\,{}^{\sigma_y}/_2}\right] - \frac{\sigma_y^2}{2}$ $\sigma_y = 0.999281z - 0.006118z^3 + 0.000127z^5, \quad z = \sqrt{{}^8/_3}(u)$
P3	$X_T = \alpha\beta + \xi + k_T\sqrt{\beta^2\alpha}$	α, β, ξ	$C_s = \frac{2}{\sqrt{\alpha}} K_T = \frac{2}{C_s}\left\{\left[\frac{C_s}{6}\left(z - \frac{C_s}{6}\right)+1\right]^3 - 1\right\}$, if $C_s > 0$ $w = \sqrt{-2\ln(1-F)}$ $(1 > F \geq 0.5)$, When F<0.5, F is substituted for (1-F) and the value of z is given a negative

	$C_s = \frac{\alpha}{\lvert\alpha\rvert}\frac{2}{\sqrt{\beta}}$ $K_T = \frac{2}{C_s}\left\{\left[\frac{C_s}{6}\left(z - \frac{C_s}{6}\right)+1\right]^3 - 1\right\}, \text{if } C_s > 0$		sign, F = 1-1/T $z = \emptyset^{-1}(F) = w - \frac{2.515517 + 0.802853w + 0.010328w^2}{1 + 1.432788w + 0.189269w^2 + 0.001308w^3}$ $t_m = 3\pi\tau_3^2$, if $0 \leq \lvert\tau_3\rvert < 1/3$ $\alpha = \frac{0.36067t_m - 0.59567t_m^2 + 0.25361t_m^3}{1 - 2.7886t_m + 2.56096t_m^2 - 0.77045t_m^3},$ $t_m = 1 - \tau_3$, if $\frac{1}{3} \leq \lvert\tau_3\rvert < 1$ $\alpha = \frac{1 + 0.2906t_m}{t_m + 0.1882t_m^2 + 0.0442t_m^3}$ $\beta = \frac{\lambda_2\sqrt{\pi}\Gamma(\alpha)}{\Gamma\left(\alpha + \frac{1}{2}\right)}$ $\xi = \lambda_1 - \alpha\beta$
W3	$x_T = B + A\left[-\ln\left(\frac{1}{T}\right)\right]^{1/k}$	k, A, B	$k = \frac{-ln2}{ln\left[1 - \frac{\lambda 2}{\lambda 1}\right]}$ $A = \frac{\lambda_2}{\left(1-2^{-\frac{1}{k}}\right)\Gamma(1+\frac{1}{k})}, B = \lambda_1 - A\Gamma(1 + \frac{1}{k})$

PERFORMANCE EVALUATION CRITERIA

The performance indicators such as Root Mean Square Error (RMSE), Relative Root Mean Square Error (RRMSE), Mean Absolute Deviation Index (MADI), Efficiency Coefficient (EC), Probability Plot Correlation Coefficient (PPCC) and Volumetric Error (VE) were used to evaluate the performance of the distributions with conventional and L-moments. A brief description of the performance indicators is presented below.

Root Mean Square Error (RMSE)

It measures the differences between observed and estimated values. It yields the residual error in terms of mean square error (Yu et al., 1994) and is expressed as

$$\text{RMSE} = \left(\frac{\sum(x_i - y_i)^2}{\text{N}}\right)^{\frac{1}{2}} \qquad (4)$$

where, x_i and y_i respectively denote the observed and estimated values. It indicates the relative performance of different models. It gives the quantitative model error in units of the variate. The smallest RMSE value indicates the best-fit model of the variate and gives the standard deviation of the model prediction error.

Relative Root Mean Square Error (RRMSE)

This indicator is calculated by dividing RMSE with average value of observed data. The model accuracy is considered excellent when RRMSE < 10%, good if 10% < RRMSE <20%, fair if 20% < RRMSE < 30% and poor if RRMSE > 30%.

$$\text{RRMSE} = \frac{RMSE}{\bar{x}} \qquad (5)$$

Mean Absolute Deviation Index (MADI)

It gives mean of absolute deviations of estimated values from observed values with respect to observed data.

$$\text{MADI}= \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i-y_i}{x_i}\right| \quad (6)$$

Probability Plot Correlation Coefficient (PPCC)

It evaluates the adequacy of a fitted distribution and is a measure of the linearity of a probability plot. It gives the correlation between the ordered observations and the corresponding fitted values determined by a plotting position. A value of the coefficient near one suggests that the observations are mostly drawn from the fitted distribution. It is given by

$$\text{PPCC} = \frac{\sum\left[(x_i-\bar{x})(y_i-\bar{y})\right]}{\left[\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2\right]^{1/2}} \quad (7)$$

where, $\bar{x}$ and $\bar{y}$ are the means of observed and estimated values respectively.

Volumetric Error (EV)

It is an absolute prediction error (Yu et al., 1994), expressed as

$$\text{VE} = \frac{\sum_{i=1}^{N}(y_i-x_i)}{\sum_{i=1}^{N}x_i} \text{ x } 100 \quad (8)$$

It measures the percent error in volume (bias) under the observed and estimated, summed over the data period. The negative EV values indicate under-estimation of the variable.

RESULTS AND DISCUSSION

Probability distributions with conventional moments

Goodness-of-fit statistics using K-S, A-D and χ^2 statistical tests for different distributions were calculated. A maximum weight of 5 for minimum value and a minimum weight of 1 for maximum value of each statistic for the distributions were given and the best-fit distribution was selected based on maximum total weight computed from the weights given to the statistics of the tests. Lognormal or gamma distribution fitted well the maximum daily streamflow data at most of the gauging sites of Godavari sub-basins selected for the present study.

Probability distributions with L-moments

Orthogonal deviations between sample L-moments (L-coefficient of variation and L-skewness) of the data at each gauging site of sub-basins with that of L-moments obtained from theoretical relationships for different 2-parameter distributions were calculated. The distribution which has the minimum value of the deviation at a gauging site indicates reasonably the best distribution. It is observed that no single distribution fitted the data at all the gauging sites. However, GPA2 followed by GAM2/LN2 fitted the data at most of the gauging sites barring a few exceptions. At the downstream-most gauging sites of Pranahitha, Indravathi and Godavari lower sub-basins, the data followed W2 probability distribution.

Orthogonal deviations between sample L-moments (L-skewness and L-kurtosis) at each gauging site of sub-basins with that of L-moments obtained from theoretical relationships for different 3-parameter distributions were calculated. The minimum value of the deviation at a gauging site indicates reasonably the appropriate distribution for describing the streamflow series. It is noticed that GPA3 at seven gauging sites, W3 and P3 at five gauging sites each, GLOG at four gauging sites and GEV at two gauging sites fitted the data in the sub-basins.

Performance evaluation of probability distributions

The performance of probability distributions recommended based on conventional and L-moments was evaluated using the performance indicators as presented in Table 5. 2 –parameter distributions with L-moments at the upstream-most, 3–parameter distributions at the middle and probability distributions with conventional moments at the downstream-most gauging sites performed better in the Godavari upper and middle sub-basins. Probability distributions with conventional moments and 3-parameter distributions with L-moments fitted the annual maximum daily streamflow data at the gauging sites in the Pranahitha, Indravathi and Godavari lower sub-basins satisfactorily.

The recommended distributions indicated relatively low values of RMSE and MADI. Further, values of RRMSE less than 10%, PPCC values above 90% and VE values less than 5% at most of the gauging sites substantiated the reasonableness of selected distributions.

It is also observed that at the gauging sites where the data showed moderate to large variability (0.075 < L-cv≤ 0.4) and moderate skewness (0.05 < L-skew ≤ 0.15), probability distributions with conventional moments seem to be a better choice compared to the distributions with L-moments. However, 2- and 3- parameter distributions with L-moments performed satisfactorily at the gauging sites where the data showed very large variability (L-cv >0.4) and skewness (L-skew>0.3).

CONCLUSIONS

The applicability of probability distributions with conventional and L-moments for annual maximum daily streamflow at the gauging sites of Godavari river basin selected for the presented study is examined. 2 - parameter distributions with L-moments (lognormal/Generalized Pareto) at the upstream, 3- parameter distributions (Pearson/Generalised Pareto) at the middle and probability distributions with conventional moments (lognormal/Gamma) at the downstream gauging sites mostly indicated satisfactory performance in the Godavari upper and middle sub-basins. Probability distributions with conventional moments (lognormal/Gamma) and 3-parameter distributions with L-moments (Generalized Pareto/Pearson/Weibull) performed better at most of the gauging sites of Pranahitha, Indravathi and Godavari lower sub-basins. The probability distributions suggested may be adopted to estimate the annual maximum daily streamflow reasonably at the gauging sites of Godavari river basin.

Table 5: Comparison of performance indicators of best-fit probability distributions with conventional and L-moments

Sub-basin	Gauging site	Performance Indicator																	
		Best-fit probability distribution			RMSE (cumec)			RRMSE			MADI			PPPC			VE		
		Conventional moments	L-moments		Conventional moments	L-moments		Conventional moments	L-moments		Conventional moments	L-moments		Conventional moments	L-moments		Conventional moments	L-moments	
			2-parameter distribution	3-parameter distribution		2-parameter distribution	3-parameter distribution		2-parameter distribution	3-parameter distribution		2-parameter distribution	3-parameter distribution		2-parameter distribution	3-parameter distribution		2-parameter distribution	3-parameter distribution
Godavari upper	Chass	Lognormal	**LN2**	GLOG	498.72	**473.74**	533.19	0.04	**0.03**	0.05	0.13	**0.12**	0.15	0.93	**0.95**	0.91	-4.97	**-3.68**	-8.56
	Ashwi	Gamma	GAM2	**P3**	63.32	68.92	**51.98**	0.04	0.05	**0.03**	0.12	0.20	**0.11**	0.97	0.96	**0.98**	-5.12	-15.15	**-4.28**
	Pachegaon	**Lognormal**	LN2	W3	**294.29**	329.29	393.52	**0.03**	0.10	0.12	**0.12**	0.42	0.45	**0.98**	0.97	0.95	**-1.42**	-2.03	-4.54
Godavari middle	Manjalegaon	Gamma	**GPA2**	GLOG	666.04	**652.29**	720.55	0.05	**0.04**	0.09	0.70	**0.56**	0.83	0.93	**0.96**	0.92	-8.56	**-3.99**	-9.29
	Dhalegaon	Gamma	**GPA2**	GPA3	276.83	**205.33**	217.93	0.05	**0.01**	0.04	0.15	**0.10**	0.12	0.96	**0.99**	0.97	-3.31	**-1.91**	-1.45
	Zari	Gamma	**GPA2**	GPA3	158.37	**109.87**	112.65	0.05	**0.03**	0.04	0.16	**0.12**	0.13	0.97	**0.99**	0.98	-4.38	**-3.34**	-3.44
	GR Bridge	Exponential	GPA2	**GPA3**	352.96	301.82	**257.45**	0.05	0.04	**0.03**	0.16	0.15	**0.13**	0.97	0.98/	**0.99**	-2.39	-2.25	**-2.14**
	Purna	Lognormal	GAM2	**P3**	767.03	846.22	**651.66**	0.04	0.05	**0.03**	0.21	0.20	**0.15**	0.96	0.97	**0.98**	-7.10	-12.98	**-6.84**
	Yelli	**Gamma**	GAM2	W3	**449.54**	565.05	598.76	**0.03**	0.04	0.05	**0.13**	0.14	0.15	**0.99**	0.96	0.95	**-2.82**	-9.32	-10.79
Pranahitha	Gandlapet	Lognormal	GAM2	**GPA3**	487.48	138.01	**123.58**	0.11	0.09	**0.02**	0.43	0.18	**0.17**	0.96	0.98	**0.99**	17.92	13.95	**-8.62**
	Mancherial	**Gamma**	GPA2	P3	**2311.42**	3526.30	3445.43	**0.02**	0.05	0.04	**0.13**	0.21	0.14	**0.99**	0.97	0.98	**-14.26**	-17.67	-17.91
	Somanpally	Lognormal	GPA2	**P3**	271.24	246.64	**240.99**	0.07	0.08	**0.03**	0.15	0.19	**0.18**	0.98	0.97	**0.99**	-1.31	-5.42	**-0.76**
	Perur	**Gamma**	W2	GPA3	**1931.83**	4155.56	9288.53	**0.01**	0.02	0.06	**0.06**	0.08	0.19	**0.99**	0.98	0.95	**-0.99**	-2.25	13.02
Indravathi	Pathagudem	**Lognormal**	LN2	GLOG	**1078.42**	1275.95	1240.85	**0.01**	0.03	0.02	**0.05**	0.07	0.06	**0.99**	0.98	0.97	**-1.26**	-4.27	-3.49
	Chindnar	**Gamma**	LN2	GEV	**528.49**	541.87	1844.76	**0.01**	0.02	0.08	**0.05**	0.07	0.19	**0.99**	0.98	0.97	**-1.38**	-3.35	-4.03
	Sonarpal	Gamma	**GAM2**	GPA3	98.84	**88.33**	235.92	0.03	**0.02**	0.09	0.10	**0.08**	0.11	0.98	**0.99**	0.97	-2.23	**-2.14**	-9.07
	Jagdalpur	Gamma	GAM2	**W3**	93.69	229.64	**68.50**	0.02	0.04	**0.01**	0.14	0.15	**0.12**	0.98	0.97	**0.99**	-0.86	-12.33	**-0.66**
	Nowrangpur	**Normal**	W2	GPA3	**94.70**	157.48	472.20	**0.01**	0.02	0.08	**0.02**	0.07	0.09	**0.99**	0.98	0.97	**-0.75**	-2.54	-5.07
Godavari lower	Sardaput	Gamma	LN2	**P3**	295.00	285.98	**259.07**	0.04	0.03	**0.01**	0.06	0.07	**0.09**	0.96	0.98	**0.99**	-12.34	-6.43	**-2.36**
	Injaram	**Gamma**	GPA2	W3	**571.90**	1162.97	626.23	**0.02**	0.04	0.03	**0.08**	0.14	0.09	**0.97**	0.95	0.96	**-1.01**	4.39	2.63
	Konta	**xtreme value**	GPA2	GLOG	**1239.62**	1559.05	1250.25	**0.02**	0.05	0.03	**0.08**	0.24	0.09	**0.96**	0.94	0.95	**-1.38**	-17.12	-3.18
	Koida	**Gamma**	GPA2	GEV	**2603.19**	6131.19	11376.46	**0.02**	0.13	0.12	**0.07**	0.16	0.08	**0.99**	0.96	0.98	**-4.92**	2.30	-34.93
	Polavaram	Normal	W2	**W3**	2009.51	4898.76	**1696.20**	0.02	0.03	**0.01**	0.06	0.14	**0.05**	0.98	0.97	**0.99**	-1.31	0.65	**-0.24**

REFERENCES

1. Abida and Elllouze, Probability distribution of flood flows in Tunisia, Journal of Hydrological Earth System Science, 2008, 12,703-714
2. Atiem I A and Harmancioglu N.B., Assessment of regional floods using L-moments approach: The case of the River Nile, Journal of Water Resources Management, 2006, 20, 723–747.
3. Ayesha S. Rahman, Ataur Rahman, Mohammad A. Zaman, Khaled addad, Amimul Ahsan & Monzur Imteaz, A study on selection of probability distributions for at-site flood frequency analysis in Australia, Natural Hazards, 2013, 69,1803-1813
4. Bettil Saf., Regional Flood Frequency Analysis Using L-Moments for the West Mediterranean Region of Turkey, Journal of Water Resources Management, 2009, 23,531–551.
5. Drissia T.K, Jothiprakash.V and Anitha A.B., Flood frequency analysis using L moments: a comparison between At Site and Regional Approach, Journal of Water Resources Management, 2019, 33, 1013-1037.
6. Durrans S. R, Eiffe M. A, Thomas Jr. W. O & Goranflo H. M., Joint Seasonal/Annual Flood Frequency Analysis, Journal of Hydrologic Engineering, 2003, 4(181),181-189.
7. Gamage N.P.D., Analysis of Probability Distribution of Annual Maximum, Mean, and Minimum Streamflows in Catchments, The Institution of Engineers, Sri Lanka, Annual Transactions, 2006, 11-20.
8. Gubareva T. S., Types of probability distributions in the evaluation of extreme floods, Journal of Water Resources, 2012, 38, 962-971.
9. Haddad K, Rahman A., Investigation on at-site flood frequency analysis in south-east Australia, The Journal of Institution of Engineers, 2008, Malaysia. 69(3), 59–64.
10. Haddad K, Rahman A., Selection of the best fit flood frequency distribution and parameter estimation procedure: A case study for Tasmania in Australia, Stochastic Environmental Research and Risk Assessment, 2010, 25(3),415–428
11. Hosking J. R. M., L-moments: Analysis and estimation of distributions using linear combinations of order statistics, Journal of Royal Statistical Society, Series B, London, England, 1990, 52, 105-124.
12. Hussain Z & Pasha GR., Regional flood frequency analysis of the seven sites of Punjab, Pakistan, using L moments, Journal of Water Resources Management, 2011, 23(10), 1917–1933.
13. Kumar R, Chatterjee C, Kumar S, Lohani A.K and Singh R. D., Development of regional flood frequency relationships using L-moments for Middle Ganga Plains Subzone 1(f) of India, Journal of Water Resources Management, 2003, 17,243–257.
14. Mamman M. J, Martins O .Y, Ibrahim J & Shaba M. J., Evaluation of best-fit probability distribution models for the prediction of inflows of Kainji reservoir, Niger State, Nigeria, Journal of Air, Soil and Water Research, 2017, 10, 1-7.
15. Ven Te Chow, David R. Maidment & Larry W. Mays., Applied Hydrology, Mc-Graw Hill Series in Water Resources and Environmental Engineering, 1988.
16. Yue S & Wang C. Y., Regional probability distributions of Canadian flood flows by L moments, Journal of Hydrology. (NZ), 2004, 43(1), 59–73.
17. Yu, P.S., Liui. C.L. & Lee. T.Y., Application of transfer function model to a storage runoff process. Stochastic and Statistical Methods in Hydrology and Environmental Engineering. 1994, 3, 87-97.