Journal of Engineering Research

Volume 7 | Issue 1 Article 2

2023

Assessment Performance of 3-Parameter Probability Distributions for At-site Annual Streamflow in the Blue Nile Basin

Amira El metwaly

Journal of Engineering Research (ERJ)
Vol. 7 – No. 1, 2023
©Tanta University, Faculty of Engineering
ISSN: 2356-9441 https://erjeng.journals.ekb.eg/ e ISSN: 2735-4873

Assessment Performance of 3-Parameter Probability Distributions for At-site Annual Streamflow in the Blue Nile Basin

A. E. El metwaly[1], H. Abd. El gawad[2] and S. H. Moharram[2]

[1]Researcher, Delta Institute for Eng. & Tech., [2]Assoc. Prof. and Professor, Faculty of Eng. Mansoura University
Email: amirametwali914@yahoo.com

Abstract: **Many investigators have applied various probability distributions for flood discharges at-site or region, however, there is no scientific judge about the best distribution to accurate the flood discharge estimations. Different probability distributions are considered, and the best distribution is then applied to create the percentile quantiles. This paper introduces the assessment of three probability distributions that have three parameters; Generalized Extreme Value (GEV), Generalized Pareto (GPA) and Generalized Logistic (GLO) using L-moments (LM) method to estimate their parameters using annual peak discharge series of three hydrological stations on Blue River Basin and Atbara River in Sudan. Cunnane plotting position formula is considered to test the applicable probability distribution that gives good estimations in tails. The Q-Q relation with coefficient of determination (R^2) is adopted to present the consistency process of the estimates and their corresponding of observed annual peak data. L-moment ratio diagram (LMRD) as suggested by Hosking and Wallis [1] is also performed to measure the discordance of probability distributions. Further, the evaluation performance of probability distributions can be measured by using three comparison criteria; root mean square error (RMSE), mean absolute deviation index (MADI) and relative root mean square error (RRMSE). The results indicated that GLO distribution generally shows the best fit followed by GEV distribution; however, the GEV distribution gave more realistic in upper tail than others. It may be recommended as the appropriate probability distribution for annual peak discharge at-site in Blue Nile**

Keywords: **probability distributions; L-moment; Comparison criteria; L-moments diagram; return period.**

I. INTRODUCTION

Annual Maximum flow discharge and ifs frequency are required in hydrological and hydraulic engineering applications such as flood risk assessment projects, water resource managing and design of both hydraulic structure and water systems. When rainfall events or stream flow records are unavailable at or near the point of interest, hydrologists struggle to create trustworthy flood estimates. Overestimated flood magnitudes result in high hydraulic structure costs, whereas high flood damage costs and human lives are affected by underestimated flood magnitudes. Estimation of return duration of extreme hydrological phenomena such as floods at a site or regional is a common subject in several fields of water resources engineering [2].

In the hydrological literature, several probability distributions have been performed. For regional flood frequency assessments and at-site, the statistical distributions are usually monitored as the better-fitted distributions [2]. In Europe regions, Castellarin, Kohnová [3] offered the statistical methodologies used for flood frequency analysis. From their study, one of the recommended distributions was the GEV for, Italy, Austria, Spain and Germany. In Finland and Turkey, Gumbel (GUM) and GEV distributions were recommended as appropriate modelling for annual maximum flood flow. Rahman, Rahman [4] carried out his study of frequency analysis on 127 locations in Australia and concluded that the log-Pearson Type III, GEV, and GPA distributions were the preferable probability distributions. On the other hand, GEV distribution has generally been adopted as a robust distribution in Turkey's Upper-Lower Mediterranean sub-regions [5].

There have been several ways established for estimating the parameters of hydrologic frequency distributions. In hydrology, the method of moments (MOM), maximum likelihood (ML), probability weighted moment (PWM), and L-Moments are popular used approaches for estimating parameters. The most powerful approach for parameter estimation is the widely reasonable ML method. It's a good method since it results the estimation parameters with the least amount of sample variance [6].

Greenwood, Landwehr [7] appointed probability weighted moments (PWM) and parameter estimation using the PWM approach, which has subsequently been widely used in application and research. For the extreme values (GEV) distribution, Hosking [8] constructed an analytical strategy for developing those parameters of using the PWM method. Using Monte-Carlo simulation with the GEV distribution, they demonstrated that the PWM technique was better than the maximum-likelihood (ML) method, particularly for the prediction of the upper tail estimates, i.e. for recurrence intervals with non-exceedance probability greater than 0.5. The magnitudes of distribution parameters determined using the PWM method are the same as those obtained using the L-moments method. Hosking [9] determined L-moments as linear combinations with PWM technique. Moreover, Hosking [10] employed L-moment ratio diagrams to correspond underlying parent distributions and L-moment ratios to test hypotheses regarding the applicability of various statistical probability distributions in a homogeneous region.

Hosking and Wallis [1] proposed an L-moment ratio-based on homogeneity test. In their relevant study, Bobée and Rasmussen [11] indicate that L-moment ratio diagram has become famous methods for identification of regional distribution, particularly for examining the outlier stations.

DOI: 10.21608/ERJENG.2023.183841.1137

Journal of Engineering Research
Vol. 7 – No. 1, 2023
©Tanta University, Faculty of Engineering
ISSN: 2356-9441 https://erjeng.journals.ekb.eg/ e ISSN: 2735-4873

Abdo, Sonbol [12] developed the studies that underpin statistical distributions at numerous sites in the Eastern Nile basin which includes the Blue Nile and its tributaries (rehad and dinder) and also river Atbara. Various flood frequency distributions (EV1, LN2, LN3, and LP3) and their parameter estimate approaches were investigated for seven stations in the basin to create homogeneous hydrological regions and to perform the frequency analysis for instantaneous annual flood discharges. The 3-parameter log normal distribution (LN3) was determined to be the superior fit distribution for testing instantaneous annual maximum discharges, while regional analysis revealed that the seven stations may be classified as homogenous zones. Gubareva and Gartsman [13] provided a systematic explanation of the theoretical framework of L-moments method. Relationships were also outlined between various characteristics distribution shapes (L-Cv, L-Cs, and L-Ck) and parameters of distribution functions (GEV, GPD, GLN, and P3). Keshtkar, Salajegheh [14] used the existing 17 hydrometric stations in Iran sub-region for peak discharges to identify relevant probability distributions (LP3, P3, LN2 and LN3) using ordinary moment and L-moment methods. It was concluded that the L-moment approach was appropriate for determining the statistical distributions for peak flow sets of the Iranian central region, and P3 was the optimal statistical distribution for modelling peak sets in the region.

In Sweden, Kousar, Khan [15] conducted a frequency analysis for annual flood data five-gauge stations on the Une River. Peak annual flow data has been fitted using GEV, GLO, LN3, and Gumble (EV1) distributions. In their work, ML and LM approaches were considered to estimate the distribution parameters. The GEV distribution with LM gave the best fit to peak annual data. Badyalina, Mokhtar [16] studied how to determine the best frequency distribution on annual peak flow data of the Segamat River. They estimated distribution parameters using the L-moment method for five commonly distributions (GPA, GEV, GLO, LP3, and LN3). The results demonstrated that the LN3 distribution is a better probability distribution for the Segamat River's annual peak flow series.

Several probability distribution functions, such as Gumbel Extreme Value type 1 (EV1), Generalized Pareto (GPA), Generalized Extreme Value (GEV), Pearson Type 3 (P3), Log-Pearson Type 3 (LP3) and Lognormal (LN2 and LN3), have been examined in various regions for the probabilistic distribution of extreme hydrological events. Önöz and Bayazit [17], Karim and Chowdhury [18], Vogel and Wilson [19], [20], Amin, Rizwan [21], Hossain [22], and Ahmad, Fawad [20] were among those who have studied and analyzed the relevant literature.

Although many researches have been done in the area of flood frequency analysis, the studies and its applications would be continually investigated to stand the best probability distribution function. Hence, this study aims to discover the optimum distribution for annual peak discharge series of four hydrological stations on Blue River Basin and Atbara River in Sudan. The distributions are among the three parameters; generalized pareto distribution (GPA), generalised logistic distribution (GLO), and generalised extreme value (GEV), all of which have their parameters calculated using LM method. Another objective is to estimate the quantiles for different recurrence intervals (*T = 2, 5, ..., 100, and 200 years*) using the calculated parameters of the candidate probability distributions for each station.

II. MATERIALS AND METHODS

A. Area of Study

In this research, the selected area of study includes the Blue Nile basin that have mainly the Blue Nile River and its tributaries (dinder and rahad). Also, River Atbara has been adopted. They are located in the north-east of Sudan (Figure 1). In this study, annual peak discharge series (APDS, m^3/s)) from four-gauge stations on these rivers are considered. The source of consideration is based on Ahmed and Ismail (2008). Some information of the gage stations contains the catchment characteristics have been described in Table 1. The main statistical characteristics of the APDS are summarized in Table 2.

It is noted that coefficient of variation for Blue Nile at Eddeim station has a good variation than others. The variation coefficient is a useful metric for determining the degree of variability of hydrological events [23]. Moreover, the skew coefficient (C_s) of Atbarra at Khashm el-Girba station is more positive than that of the normal, but at the Eddiem station may be appropriate to the normal distribution.

Table 1. Information available on-site used on Blue Nile Basin.

River	Station	Recording Period	Longitude	Latitude	Area, km^2
Blue Nile	Eddeim	(1920-2006)	34° 59' E	11° 14' N	179,486
Rahad	Hawata	(1912-2005)	34° 55' E	13° 40' N	35,000
Dinder	Gwasi	(1912-2005)	34° 10' E	13° 20' N	37,000

Table 2. Summarizes of the main statistical characteristics of annual peak discharge series.

River/Station	N	$\overline{X}$	Max	Min	S^2	C_S	C_k	C_V
Blue Nile/Eddeim	87	1562.33	2243.89	948.34	74013.95	0.166	2.799	0.174
Rahad/ Hawata	95	36.49	80.758	2.112	100.46	0.551	7.528	0.275
Dinder/ Gwasi	95	87.68	180.73	11.74	1172.57	0.486	2.975	0.391
Atbara/Khashm el-Girba	20	407.96	826.72	157.71	27865.08	0.929	3.353	0.409

DOI: 10.21608/ERJENG.2023.183841.1137

Journal of Engineering Research (ERJ)
Vol. 7 – No. 1, 2023
©Tanta University, Faculty of Engineering
ISSN: 2356-9441 https://erjeng.journals.ekb.eg/ e ISSN: 2735-4873

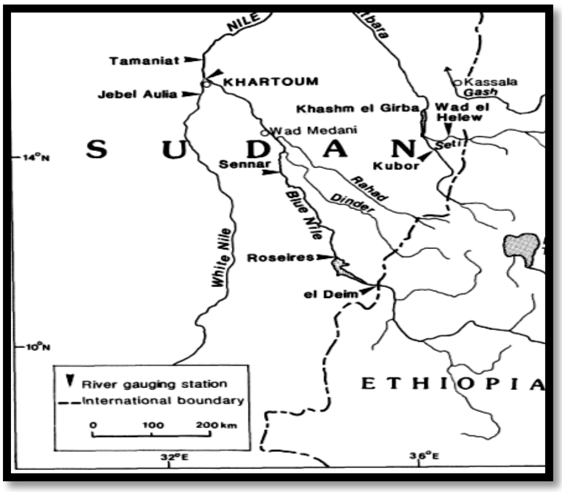

Figure 1. Area of study location [12]

B. Methods

B.1 Probability Distributions

When describing flood frequency at a certain location, an appropriate probability distribution function is critical. Since the most discharge data has a skew shape, the probability distributions with the shape parameter may be preferable to exam the flood frequency analysis. In this study, the candidate distributions; GPA, GLO and GEV are selected for examining of annual peak discharge series at four Blue River Basin gauging sites. These distributions have been advocated in prior research for flood frequency analysis [4, 19, 24, 25]. Table 3 describes the estimation of the probability density function (PDF) and its parameters using the LM approach. The sub-sections that follow provide a brief description of various methods.

B.2 L-Moments Method

To estimate the distribution parameters the L-moment approach is used here where it is take place on the annual peak discharge data to perform flood frequency analysis. L-moments theory was developed by Hosking [10] and instructions for their practical application were also introduced. L-moments are essentially linear functions of PWM's that can be interpreted directly as scale and form metrics. L-moments and PWM's are like conventional moments, summarize theoretical distributions and observed samples. Greenwood, Landwehr [7] outline PWM theory and characterize them as follows:

$$\beta_r = E\{x[\mathbb{F}_{\mathbb{S}}(x)]^r\} \qquad (1)$$

where β_r is the r^{th} order PWM's and $F_s(x)$ is the cumulative distribution function (CDF) of x.

For a sample rank, $x(1) \leq x(2) \leq \ldots . \leq x(n)$, PWM's as an unbiased estimate of sample can be written as

$$\beta_0 = \frac{1}{n}\sum_{j=1}^{n} x_{(j)} \qquad (2)$$

$$\beta_1 = \sum_{j=1}^{n-1}\left[\frac{(n-j)}{n(n-1)}\right] x_{(j)} \qquad (3)$$

$$\beta_2 = \sum_{j=1}^{n-2}\left[\frac{(n-j)(n-j-1)}{n(n-1)(n-2)}\right] x_{(j)} \qquad (4)$$

$$\beta_3 = \sum_{j=1}^{n-3}\left[\frac{(n-j)(n-j-1)(n-j-2)}{n(n-1)(n-2)(n-3)}\right] x_{(j)} \qquad (5)$$

where n is the sample size, $x_{j:n}$, and $b_r(r = 0,1,2,3)$ is the unbiased estimator of β_r

Table 3. L-Moment Parameter estimates for selected probability distributions.

Distribution	CDF	Quantile Estimator	Parameters
GEV	$F_X(X) = exp\left[-\left[1-\frac{K(x-\zeta)}{\alpha}\right]^{\frac{1}{k}}\right]$	$X(F) = \zeta + \frac{\alpha}{K}\{1-(-lnF)^{k}\}$	$\alpha = \frac{l_2 K}{\Gamma(1+K)(1-2^{-k})}$ $\zeta = L_1 + \frac{\alpha(\Gamma(1+K)-1)}{K}$ $K = 7.8590C + 2.9554C^2$ $C = \frac{2}{3+\tau_3} - \frac{Ln2}{Ln3}$
GPA	$F_x = (\frac{1}{\alpha})\left(\frac{\gamma^{(1-x)}}{1+\gamma}\right)^2$	$X(F) = \zeta + \frac{\alpha}{k}\{1-(1-F)^{k}\}$	$\alpha = l_2[(K+1)(K+2)]$ $\zeta = l_1 - l_2(K+2)$ $K = \frac{4}{\tau_3+1} - 3$
GLO	$F_x = 1-\left(1-\frac{k(x-\zeta)}{\alpha}\right)^{\frac{1}{k}}$	$X(F) = \zeta + \frac{\alpha}{K}\left\{1-\left(\frac{1-F}{F}\right)^{k}\right\}$	$\alpha = \frac{l_2}{\Gamma(1+k)\Gamma(1-k)}$ $\zeta = l_1 + \frac{(l_2-\alpha)}{k}$ $k = -\tau_3$

DOI: 10.21608/ERJENG.2023.183841.1137

Journal of Engineering Research (ERJ)
Vol. 7 – No. 1, 2023
©Tanta University, Faculty of Engineering
ISSN: 2356-9441 https://erjeng.journals.ekb.eg/ e ISSN: 2735-4873

For any distribution, the first four L- moments ($\lambda_1, \lambda_2, \lambda_3 and \lambda_4$) are given by:

$$\lambda_1 = b_0 \quad (6)$$

$$\lambda_2 = 2b_1 - b_0 \quad (7)$$

$$\lambda_3 = 6b_2 - 6b_1 + b_0 \quad (8)$$

$$\lambda_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \quad (9)$$

where β_r (r = 0, 1, 2, 3) is the PWMs. Estimates of the $\lambda_r$are established by applying the unknown β_r by sample estimates b_r from equation. A parallel theory of L-statistics exists to estimate the L- variation coefficient, L- skewness and L-kurtosis; to identify parent distribution; and to estimate distribution parameters [1].

Hence, the dimensionless samples of L- moment ratios are written as:

$$L - cv\ (\tau_2) = \frac{\lambda_2}{\lambda_1} \quad (10)$$

$$L - skewness(\tau_3) = \frac{\lambda_3}{\lambda_2} \quad (11)$$

$$L - kurtosis(\tau_4) = \frac{\lambda_4}{\lambda_2} \quad (12)$$

The *L*-moments $\lambda_1 and\ \lambda_2$, the *L-cv*, τ_2 and L-moment ratios $\tau_3\ and\ \tau_4$ are the most important qualities for outlining probability distribution. Table 3 shows the CDF and inverse form of three candidate distributions and its equations of sample moment parameters using L-moments method.

The parameters and quantiles estimation of the candidate probability distributions with available annual peak discharge data is based on Python3.7 software which is prepared by the authors. Python 3.7 software includes the parameters estimates with L-moments method for the of GPA, GLO, and GEV distributions and their quantiles estimation at certain return periods. The Python3.7 software includes comparison criteria tests for fitting the best distribution. Furthermore, Python3.7 program offers graphical presentation of the three probability distributions used, allowing for a clear visual evaluation of the fitted distributions to the given annual peak discharges.

III. PERFORMANCE OF FITTED DISRIBUTIONS

This section is classified into three items; comparison criteria, Q-Q relation with determination coefficient (R^2) and L-moment ratio Diagram (LMRD) to judge the accuracy performance and consequently to identify the optimum distribution at the four sites of both Blue Nile basin and Atbara River.

A. Comparison Criteria

To characterize how well a specific distribution matches the observation data, various criteria are taken. Three standard criteria are utilized in this study to assess how well a given distribution and a particular estimation technique match the observed discharge data: root mean squared error (RMSE), mean absolute deviation indicator (MADI), and relative root mean square error (RRMSE).

B. Comparison Criteria

To characterize how well a specific distribution matches the observation data, various criteria are taken. Three standard criteria are utilized in this study to assess how well a given distribution and a particular estimation technique match the observed discharge data: root mean squared error (RMSE), mean absolute deviation indicator (MADI), and relative root mean square error (RRMSE).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i-1}^{n}\left(\frac{x_i - y_i}{x_i}\right)^2} \quad (13)$$

$$MADI = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - y_i}{x_i}\right| \quad (14)$$

$$RRMSE = \left[\frac{\sum\left(\frac{x_i - y_i}{x_i}\right)^2}{(n-m)}\right]^{\frac{1}{2}} \quad (15)$$

$$R^2 = \frac{\sum(X_i - \bar{x})(y_i - \bar{y})}{\left[\sum(X_i - \bar{x})^2(y_i - \bar{y})^2\right]^{\frac{1}{2}}} \quad (16)$$

where x_i are sample records, i = 1, 2,...., n, and y_i, i = 1, 2,...., n are computed magnitudes from the selected probability distribution at different return periods and m is the parameter number of a distribution.

C. Quantile-Quantile Relation (Q-Q relation)

Beside the comparison criteria used, Q-Q relation is also adopted for fitting the best distribution. It can be assessed by calculating the coefficient of determination (R^2), which is necessary to measure the probability fitting's linearity. Python3.7 program allows for graphical (Q-Q relation) fitting of the probability distributions of interest. Here, a linear relationship is established between the observation data and its probability, as well as the corresponding fitted quantiles given by each distribution with regard to the chosen unbiased plotting location. Cunnane [2] provides the plotting position formula as follows:

$$P(x_i) = \frac{m - 0.40}{N + 0.2} \quad (17)$$

where $P(x_i)$ is the plotting probability, *n* is the size of sample, and *m or i* is the order rank of observation data with *m* = 1 for the smallest value of sample observations.

D. L-Moment Ratio Diagram (LMRD)

L-Moment Ratio Diagram (LMRD), Hosking and Wallis [1], is considered as another approach to assess the goodness of fit to obtain an ideal probability distribution. For the three distributions used, the LMRD illustrates the conceptual framework of relationship between τ_4 (*L-kurtosis*) and τ_3 (*L-skewness*). For this purpose, sample estimates of *both L-kurtosis* and *L-skewness* for the observation data are obtained to present the coordinates of point (τ_3, τ_4) in the domain of

DOI: 10.21608/ERJENG.2023.183841.1137

Journal of Engineering Research (ERJ)
Vol. 7 – No. 1, 2023
©Tanta University, Faculty of Engineering
ISSN: 2356-9441 https://erjeng.journals.ekb.eg/ e ISSN: 2735-4873

LMRD which represents by $\tau_4 = f(\tau_3)$ as a series of dependencies for various theoretical distributions [26].

IV. RESULTS AND DISCUSSION

This work involves three probability distributions which are suitable to annual peak discharge series in the field of frequency analysis, namely GEV, GPA and GLO. The distributions parameters have been obtained by using the L-moment method. The estimation parameters for GPA, GLO and GEV distributions are illustrated in Table 4.

A. Quantile Estimates

From the results of this work for annual peak discharge, the quantile estimates for each distribution are obtained using Cunnane unbiased plotting position formula for all stations. The determination coefficient (R^2) is calculated to establish the accuracy performance of each distribution as shown in Table 5. It is observed that the GEV has a relatively good correlation for three stations followed by GLO distribution, however GLO correlation may be approximately comparable with GEV distribution. Moreover, the correlation of GPA distribution has the less correlation for all stations.

In order to support the observation based on the Q-Q relation as a correlation method, other criteria named MADI, RRMSE and RMSE are considered to evaluate the accuracy of Candidate distributions. Table 6 addresses the evaluation results of performance the candidate distributions where the score order is obtained in order to get a better distribution. As notation, the smallest sum scores indicate the best distribution and vice versa, Table 6. From the Table 6, it is observed that GLO distribution has an agreement to annual peak discharge series of three stations in four used followed by GEV. GPA distribution is out fitting to the data used as it has poor results for all comparison criteria.

B. Performance of LMRD

The position of a point with coordinates (τ_3, τ_4) for the observation data under consideration is the LMRD may be suggested as appropriate test to check which distribution will be appropriate for the annual discharge series.

Table 4. Parameters estimates with L-moments for of distributions used at four stations

River/ Station	Distribution	Parameters		
		K	α	ζ
Blue Nile/ Eddiem	GEV	0.2127	8295.8202	45887.3179
	GPA	0.8433	25635.2894	35298.9145
	GLO	-0.0408	4878.0292	48878.5978
Rahad/ Hawata	GEV	0.2449	283.03	980.74
	GPA	-0.0224	163.60	1081.66
	GLO	0.9125	912.09	610.77
Dinder/ Gwasi	GEV	0.1046	961.19	2238.12
	GPA	-0.1047	599.26	2597.48
	GLO	0.6207	2591.91	1102.83
Atbara/ Khashm el-Girba	GEV	-0.0691	128.1044	324.6655
	GPA	0.2926	281.6323	190.0743
	GLO	-0.2149	87.9857	375.1274

Table 5. Determination coefficient (R2) for the probability distributions with Cunnane plotting position

River/station	(R^2)			best
	GEV	GPA	GLO	
Blue Nile/Eddeim	0.9971	0.9803	0.9942	GEV
Rahad/ Hawata	0.9494	0.8917	0.9733	GLO
Dinder/ Gwasi	0.9964	0.9854	0.9915	GEV
Atbara /Khashm el-Girba	0.9909	0.9863	0.9898	GEV

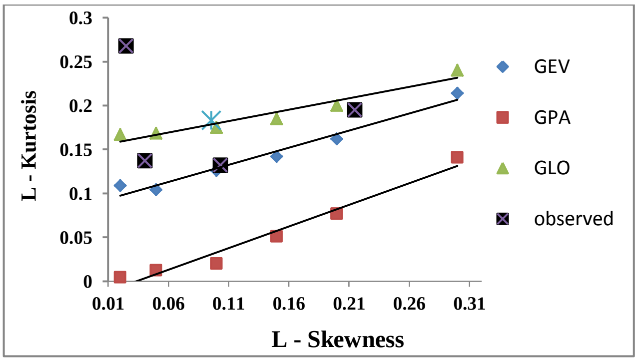

Figure 2. L-Moment Ratio Diagram

Table 6. Comparison criteria and their score for the probability distributions

River/ Station	Distribution	Comparison criteria						sum
		RMSE	rank	MADI	rank	RRMSE	rank	score
Blue Nile/ Eddiem	GEV	0.0148	1	0.011	1	0.015	1	3 (1st)
	GPA	0.0406	3	0.0292	3	0.0413	3	9 (3rd)
	GLO	0.0186	2	0.012	2	0.0189	2	6(2nd)
Rahad/ Hawata	GEV	0.6462	2	0.1098	2	0.6567	2	6(2nd)
	GPA	0.9544	3	0.1705	3	0.9699	3	9(3rd)
	GLO	0.4656	1	0.077	1	0.4731	1	3 (1st)
Dinder/ Gwasi	GEV	0.0801	2	0.0329	2	0.0814	2	6 (2nd)
	GPA	0.2529	3	0.0774	3	0.2569	3	9(3rd)
	GLO	0.0493	1	0.0324	1	0.0501	1	3 (1st)
Atbara/ Khashm el-Girba	GEV	0.0567	1	0.0465	2	0.0615	1	4 (2nd)
	GPA	0.0861	3	0.0651	3	0.0934	3	9(3rd)
	GLO	0.0567	1	0.042	1	0.0615	1	3 (1st)

DOI: 10.21608/ERJENG.2023.183841.1137

Journal of Engineering Research (ERJ)
Vol. 7 – No. 1, 2023
©Tanta University, Faculty of Engineering
ISSN: 2356-9441 https://erjeng.journals.ekb.eg/ e ISSN: 2735-4873

LMRD is developed for each of the four gaging stations, as well as discussed before. Figure 2 presents the LMRD for the data points of four stations used. The average of the four data is also shown as well as the candidate distributions; GPA, GLO and GEV for measuring those performance. The diagram showed that the GLO distribution is better than GEV and GPA distributions, and it can adequately characterize the annual peak stream flows data for four stations.

C. Upper Quantile Estimates

After getting the best probability distribution as a whole, it is necessary to check the upper tail of distribution. This is basically applicable to flood frequency analysis and flood design. Therefore, the quantile estimates with a certain return periods presents this goal. The quantile estimates for specific return periods (*T = 2, 5,.., 100, and 200 years*) are found using cumulative function and its parameters values of particular gauging site. Quantile estimates for various certain return periods, *T*, are listed in Table 7. As the results, the GEV distribution has perfect estimates of Blue Nile /Elddime and Dinder/ Gwasi station followed by GLO. Moreover, GPA has a good estimated for Atbara/ Kha-shm el-Girba station. For Rahad station, the candidate distributions gave lowest estimates; i.e. underestimated values. Figure 3 shows the frequency curves of upper tail for probability distributions. Further, the evaluation performance of the upper tail with specific return periods, *T* of probability distributions are approximately closed with the observation values of annual peak discharge for GEV followed by GLO and GPA. Furthermore, by using MADI criteria, the character of probability distributions has agreed with the visual inspection in upper tail trend, as shown in Table 8.

V. CONCLUSIONS

Since the climate conditions contribute in the change the rainfall events, expecting flows in future will be have fraught with uncertainty. In some catchments even, long data groups cannot treat the hydrological problems. This work, however, explores the possibility of three probability distributions and its usage for flood frequency analysis at single site in the Blue Nile Basin using the annual peak flows of four stations. The L-moment technique is used to estimate parameters for each selected distribution in this study. Three tests goodness-of-fit are employed for this purpose, namely root mean square error (RMSE), mean absolute deviation indicator (MADI), and relative root mean square error (RMSE). Moreover, Q-Q probability relation with coefficient of determination (R^2) and L-moment ratio diagram (LMRD) are suggested for visible test as goodness-of fit criteria.

Results from this comparative investigation establish that GLO distribution has relatively suitable performance than other distributions. More specific conclusions include the following: (i) the GPA distribution is not suitable for Blue Nile Basin data, since this distribution demonstrably leads to an underestimation of stream flow quantiles; (ii) the estimation of GEV distribution are most suitable in quantile estimates of the upper tail followed by GLO distribution and (iii) Further it can be recommended GEV as a good fit probability distribution for Blue Nile basin.

Table 8. MADI values for upper tail of distributions used.

River/Station	Distribution			Best distribution
	GEV	GPA	GLO	
Blue Nile/Eddeim	0.00773	0.03373	0.20468	GEV
Rahad/ Hawata	0.10126	0.14209	0.07641	GLO
Dinder/ Gwasi	0.03276	0.05630	0.06607	GEV
Atbara/Khashm el-Girba	0.10258	0.04746	0.12876	GPA

Table 7. Quantile estimates for specific return periods, T-year

River/ Station	Distribution	T-year						
		2	5	10	25	50	100	200
BlueNile/ Eddeim	Obs.	1531.405	1788.151	1952.356	2084.786	2132.242	2214.285	2261.740
	GEV	1552.009	1794.474	1924.052	2059.299	2142.495	2212.989	2272.997
	GPA	1545.897	1839.355	1952.725	2030.815	2060.983	2078.041	2087.686
	GLO	1568.016	1573.014	1574.681	1575.974	1576.588	1577.012	1577.309
Rahad/ Hawata	Obs.	36.4325	43.0478	46.4454	53.1652	63.1718	75.3488	82.6635
	GEV	36.2198	44.3493	48.6346	53.0536	55.7395	57.9925	59.8913
	GPA	36.1928	45.751	49.2072	51.4445	52.2535	52.6864	52.918
	GLO	36.2762	43.5899	47.9874	53.4269	57.4702	61.5255	65.6226
Dinder/ Gwasi	Obs.	85.1331	113.9874	138.0097	160.0172	166.2937	175.6901	182.5093
	GEV	83.9749	115.3302	134.0936	155.7456	170.4471	183.986	196.5062
	GPA	83.3206	120.6729	137.6548	151.3048	157.5107	161.5334	164.1409
	GLO	84.4074	112.7591	131.3266	155.9651	175.5312	196.2743	218.4125
Atbara/ Kha-shm el-Girba	Obs.	381.3823	509.3206	694.3689	806.6387	845.641	865.1422	874.8928
	GEV	372.2167	527.1208	636.5595	783.1996	898.3617	1018.3427	1143.7812
	GPA	366.7659	551.5782	661.9147	777.3222	846.2227	902.4763	948.4043
	GLO	375.1274	517.213	622.2001	776.2266	910.5782	1064.7229	1242.6081

DOI: 10.21608/ERJENG.2023.183841.1137

Journal of Engineering Research (ERJ)
Vol. 7 – No. 1, 2023
©Tanta University, Faculty of Engineering
ISSN: 2356-9441 https://erjeng.journals.ekb.eg/ e ISSN: 2735-4873

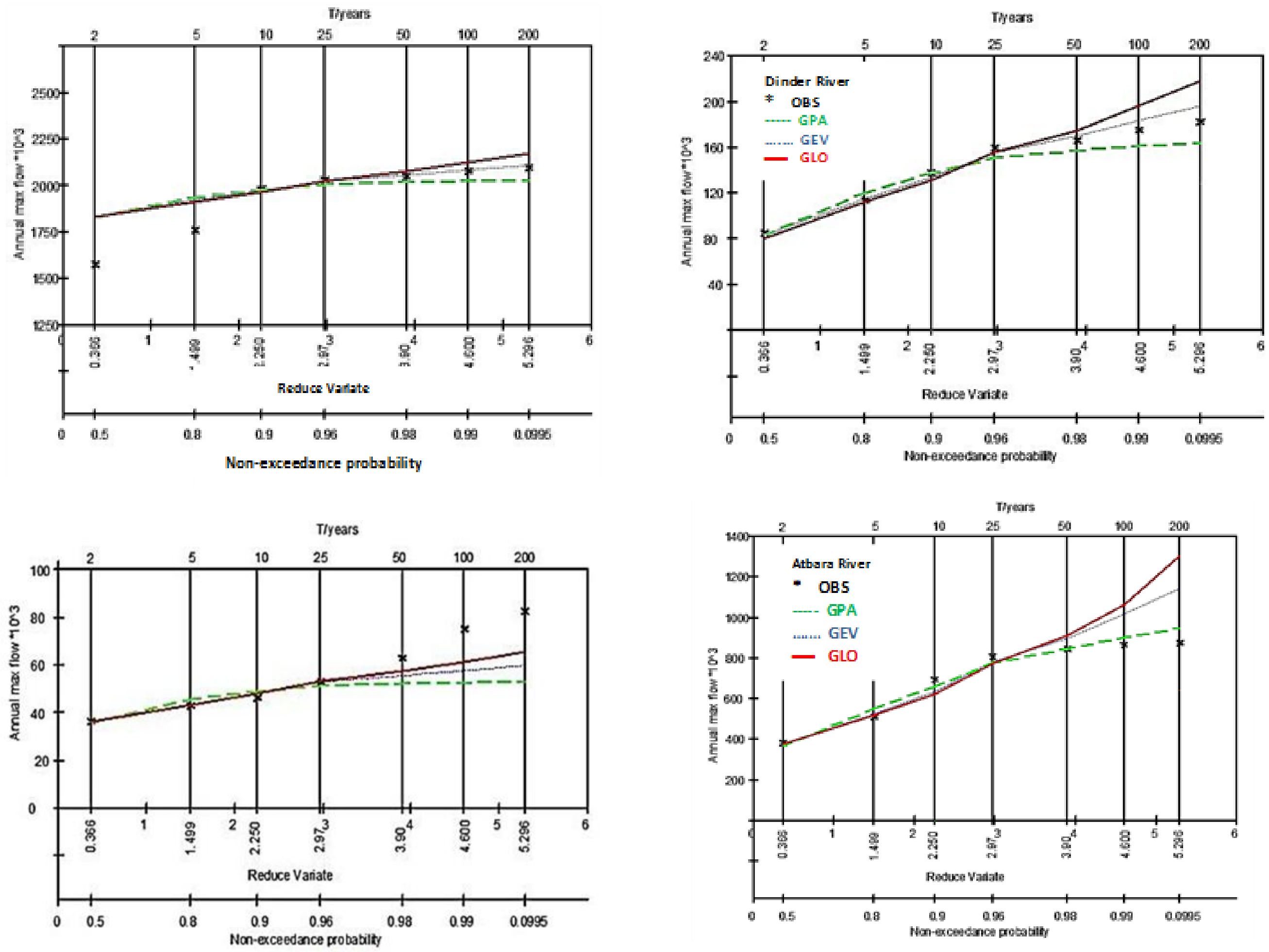

Figure 3. Frequency curves of upper tail for probability distributions.

Funding: This research has not been conducted under any fund.

Conflicts of Interest: The authors declare that there is no conflict of interest.

REFERENCES

1. Hosking, J. and J. Wallis, *Some statistics useful in regional frequency analysis.* J Water resources research, 1993. 29(2): p. 271-281.
2. Cunnane, C., *A particular comparison of annual maxima and partial duration series methods of flood frequency prediction.* Journal of hydrology, 1973. 18(3-4): p. 257-271.
3. Castellarin, A., et al., *Review of applied statistical methods for flood frequency analysis in Europe.* 2012.
4. Rahman, A.S., et al., *A study on selection of probability distributions for at-site flood frequency analysis in Australia.* J Natural hazards, 2013. 69(3): p. 1803-1813.
5. Saf, B., *Regional flood frequency analysis using L-moments for the West Mediterranean region of Turkey.* J Water Resources Management, 2009. 23(3): p. 531-551.
6. Rao, A.R. and K.H. Hamed, *Flood frequency analysis*. 2019: CRC press.
7. Greenwood, J.A., et al., *Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form.* J Water resources research, 1979. 15(5): p. 1049-1054.
8. Hosking, J.R., *Algorithm as 215: Maximum-likelihood estimation of the parameters of the generalized extreme-value distribution.* Journal of the Royal Statistical Society. Series C, 1985. 34(3): p. 301-310.
9. Hosking, J.R., *The theory of probability weighted moments*. 1986: IBM Research Division, TJ Watson Research Center New York, USA.
10. Hosking, J.R., *L- moments: Analysis and estimation of distributions using linear combinations of order statistics.* Journal of the Royal Statistical Society: Series B, 1990. 52(1): p. 105-124.
11. Bobée, B. and P.F. Rasmussen, *Recent advances in flood frequency analysis.* J Reviews of Geophysics, 1995. 33(S2): p. 1111-1116.
12. Abdo, G., M. Sonbol, and P. Willems. *Flood frequency analysis of the eastern Nile Rivers.* in *International Conference of UNESCO Flanders FIT FRIEND/Nile Project–"Towards a better cooperation", Sharm El-Sheikh, Egypt*. 2005.
13. Gubareva, T. and B. Gartsman, *Estimating distribution parameters of extreme hydrometeorological characteristics by L-moments method.* J Water Resources, 2010. 37(4): p. 437-445.
14. Keshtkar, A., A. Salajegheh, and M. Najafi Hajivar, *Flood flow frequency model selection using L-moment method in arid and semi arid regions of Iran.* J Desert, 2012. 17(1): p. 41-48.
15. Kousar, S., et al., *Some best- fit probability distributions for at- site flood frequency analysis of the Ume River.* Journal of Flood Risk Management, 2020. 13(3): p. e12640.
16. Badyalina, B., et al., *Flood Frequency Analysis using L-Moment For Segamat River.* J MATEMATIKA: Malaysian Journal of Industrial Applied Mathematics, 2021: p. 47-62.
17. Önöz, B. and M. Bayazit, *Best-fit distributions of largest available flood samples.* Journal of Hydrology, 1995. 167(1-4): p. 195-208.
18. Karim, M.A. and J.U. Chowdhury, *A comparison of four distributions used in flood frequency analysis in Bangladesh.* Hydrological Sciences Journal, 1995. 40(1): p. 55-66.

DOI: 10.21608/ERJENG.2023.183841.1137

Journal of Engineering Research (ERJ)
Vol. 7 – No. 1, 2023
©Tanta University, Faculty of Engineering
ISSN: 2356-9441 https://erjeng.journals.ekb.eg/ e ISSN: 2735-4873

19. Vogel, R.M. and I. Wilson, *Probability distribution of annual maximum, mean, and minimum streamflows in the United States.* Journal of hydrologic Engineering, 1996. 1(2): p. 69-76.
20. Ahmad, I., M. Fawad, and I. Mahmood, *At-Site Flood Frequency Analysis of Annual Maximum Stream Flows in Pakistan Using Robust Estimation Methods.* Polish Journal of Environmental Studies, 2015. 24(6).
21. Amin, M., M. Rizwan, and A. Alazba, *A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan.* J Open Life Sciences, 2016. 11(1): p. 432-440.
22. Hossain, A., *Flood frequency analysis: a case study for the Brisbane River catchment.* 2020.
23. Mkhandi, S., S. Kachroo, and D. es Salaam, *Regional flood frequency analysis for Southern Africa.* J Southern African FRIEND, Technical Documents in Hydrology, 1997(15).
24. Cunnane, C., *Statistical distributions for flood frequency analysis.* J Operational hydrology report, 1989.
25. Guru, N. and R. Jha, *Flood frequency analysis of Tel Basin of Mahanadi river system, India using annual maximum and POT flood data.* J Aquatic Procedia, 2015. 4: p. 427-434.
26. Hosking, J.R.M. and J.R. Wallis, *Regional frequency analysis*. 1997.

DOI: 10.21608/ERJENG.2023.183841.1137