Al-Azhar University Civil Engineering Research Magazine (CERM)
Vol. (41) No. (1) January, 2019

Assessing Adequacy Frequency Distribution Functions Based on L-Moments Approach for Annual Peak Flow on Nile River

Sabry M. Foudah*, Saad H. Moharram, Tarek A. Saafan and Mahmoud M. El-Gamal
Irrigation and Hydraulics Engineering Department, Faculty of Engineering, Mansoura University, Egypt.

ملخص البحث
أساس التصميم الإقتصادي للمنشآت الهيدروليكية يعتمد على التقدير الصحيح للتصرفات القصوى التصميمية لفترة زمنية مستقبلية. و حيث أن تقنية تحليل التوزيعات الإحتمالية تعتبر نمذجة رياضية مهمة و آداة فعالة في التصميم, لذا تم التركيز في هذا البحث على نمذجة تحليل التوزيعات الإحتمالية في الموقع بطريقة العزوم-إل لأربعة دوال توزيعات إحتمالية و هي: توزيع (EV1)، توزيع (LOG)، توزيع (GPA)، وتوزيع (GEV) كما تم تطبيقها على تصرفات مسجلة لأربعة محطات قياس على نهر كاجيرا, نهر بحر الجبل, النيل الأبيض و النيل الأزرق على حوض نهر النيل. بتحليل و مقارنة نتائج التصرفات لكل محطة قياس تبين أن دالة توزيع GEV تعتبرالأفضل و الأكثر دقة لتطبيقها فى التنبؤ بتقدير التصرف التصميمي.

Abstract
Economical design of pivotal hydraulic structures are almost possible with correct and realistic estimation of future design quantile magnitude for a certain return period. If appropriate and enough data records are available, at-site frequency distribution analysis is applied to estimate a design flow on a river or stream. In this paper, ’at-site method’ of frequency analysis is localised and discussed including the L-moments method on four distributions functions; namely: Gumbel's Extreme value type 1 distribution (EV1), Logistic distribution (LOG), Generalized Pareto distribution (GPA), and Generalized Extreme Value distribution (GEV) were selected. L-moments method is applied for fittings these distributions to annual peak flow of four selected stations in Nile River tributaries, and outcome estimates thus obtained are compared. Results from this comparative study represented that the quantile estimates by GEV provides a better fit to peak/maxima flow than other distributions, whereas the Gumbel distribution gives less fitting accuracy.

Keywords: Frequency Analysis, L-moments, Probability Distribution Function, Quantile Estimates.

1. Introduction
The frequency distribution analysis is the generally used execution for estimating the design peak discharge of a certain return period on a river or stream in many water resource valley projects. In other words, frequency analysis is an efficient tool in design which reduces the cost of the projects due to employment of the forecasting techniques. Many Frequency analysis approaches exist in the literature. Benson (1968) and Vogel et al. (1993) used several distributions for describing flood flows in the USA. Felício Cassalho et al. (2018) studied at-Site flood frequency analysis coupled with multiparameter probability distributions.

The parameters of frequency distribution functions can be evaluated by several methods. The method of moments and the maximum likelihood method have been widely used. Recently, the L-moments technique, Hosking (1990), has been vastly utilized for estimating the parameters, spatially for the inverse form probability distributions because it have some important advantages such as capability of yielding good parameter estimates that are almost unbiased and more accurate, Stedinger et al. (1993).

For the design of control structures, a design return period T-years is often required to calculate the estimation discharges from the best-fit distribution function. The selection of best distribution function in the studied four stations on Nile River tributaries with the available records will definitely help the planners and administrators to estimate the design flow discharge. This paper offered a comparative consideration made to decide which distribution represented the statistical feature of observed flow discharge data in order to be utilized for predicting design discharge of different return periods for 4 sets of peak/maxima discharges on the Nile River tributaries.

2. Methodology

Frequency distribution functions:

There are numerous forms of theoretical probability frequency distribution functions that had been successfully applied to hydrologic data. The frequency distribution functions commonly used for this study are: Gumbel's Extreme value type 1 distribution (EV1), Logistic distribution (LOG), Generalized Pareto distribution (GPA), and Generalized Extreme Value distribution (GEV).

Extreme Value type 1 distribution (EV1) or Gumbel distribution

EV1 distribution is a double exponential distribution. It reads as in Eq. (1).

$$X = \xi - \alpha \ln(-\ln F) \tag{1}$$

$$\alpha = \frac{l_2}{\ln 2}; and$$

$$\xi = l_1 - 0.5772\alpha \tag{2}$$

where, α and ξ are the EV1 distribution parameters, l_1 and l_2 are L-moments.

Logistic distribution (LOG)

The log-logistic distribution is a continuous probability distribution which are used to model stream flow in hydrology. The inverse cumulative distribution function (quantile function) of the LOG distribution is defined as follows:

$$X = \xi + \alpha \ln\left[\frac{F}{1-F}\right] \tag{3}$$

$$\xi = l_1; and$$

$$\alpha = l_2 \tag{4}$$

where, α and ξ are the LOG distribution parameters, l_1 and l_2 are L-moments.

Generalized Pareto distribution (GPA)

The generalized Pareto distribution is a three-parameter distribution that contains uniform, exponential, and Pareto distributions as special cases. The GPA is also a special case of the Wakeby distribution. The GPA reads as in Eq. (7).

$$X = \xi + \frac{\alpha}{k}\left[1 - (1-F)^k\right] \tag{5}$$

$$k = \frac{(1-3\tau_3)}{(1+\tau_3)};$$
$$\alpha = l_2(1+k)(2+k); and$$
$$\xi = l_1 - l_2(2+k) \quad (6)$$
where, α, ξ and k are the GPA distribution parameters, τ_3 is L-SKEWNESS (L-CS), l_1 and l_2 are L-moments.

Generalized Extreme Value distribution (GEV)

The Generalized Extreme Value (GEV) distribution is the basic distributional tool, in contrast to classical statistical inference which focuses on central measures of a distribution and where the normal distribution is the reference. The GEV distribution is often used as an approximation to model the maxima of long (finite) sequences of random variables. The distribution function of the GEV is of the form:

$$X = \xi + \frac{\alpha}{k}\left[1-(-\ln F)^k\right] \quad (7)$$
$$k = 7.8590C + 2.9554C^2;$$
$$C = \frac{2}{3+\tau_3} - \frac{\ln 2}{\ln 3};$$
$$\alpha = \frac{l_2.k}{\Gamma(1+k)(1-2^{-k})}; and$$
$$\xi = l_1 + \frac{\alpha}{k}\left[\Gamma(1+k)-1\right] \quad (8)$$
where, α, ξ and k are the GEV distribution parameters, τ_3 is L-SKEWNESS (L-CS), l_1 and l_2 are L-moments.

Plotting position

Different Plotting positions have been widely used in older to hydrologic applications; the method of plotting positions is based on the first creating a visualization of the sample distribution and then performing a curve-fit between the chosen distribution and the sample. In this study, Gringorten Formula (1963) is applied on to the samples.

$$P(x) = \frac{m-0.44}{N+0.12} \quad (9)$$

Where: $P(x)$ is the exceedance probability, N is the sample size, and m is the rank of the observations in ascending order.

3. The Study Area

A total of 4 series flow quantile data have been adopted for this study for the calculation of theoretical frequency distribution function, on gauging stations of Nile River tributaries namely Kagera River at Kyaka Ferry, Bahr el Jebel at Mongalla, White Nile at Malakal and Blue Nile at Sennar, as shown in Figure 1.

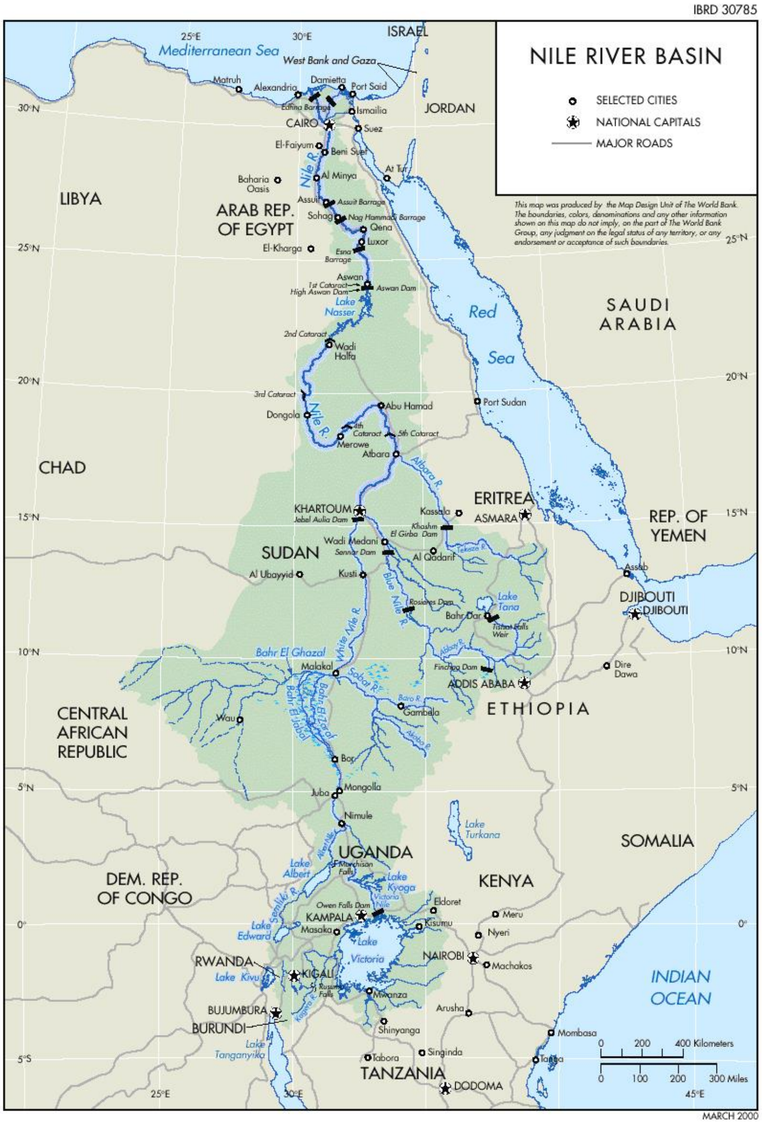

Figure 1.Political Map of the Nile River Basin.

The result of frequency distribution depends on the length of the hydrological series. If the observed data sample is too small the predictions about the future discharges cannot be expected to be reliable, Therefore, Studies have revealed that records shorter than 20 years should not be used in the frequency analysis. In the present study, the available discharge data records length differs from 32 till 62 years. The data used are summarized in table 1.

Table 1: Discharge data of 4 - stations of the Nile River Basin and its tributaries.

Serial number	Station and River	Period of Record	Number of years
1	Kagera River at Kyaka Ferry	1940-1971	32
2	Bahr el Jebel at Mongalla	1912-1973	62
3	White Nile at Malakal	1912-1973	62
4	Blue Nile at Sennar	1912-1973	62

4. Results and Analysis

This section offered the at-site frequency analysis results of four selected probability distributions (i.e. EV1, LOG, GPA and GEV) for four sets of peak flow of Nile River, Gringorten Formula and L-Moments Method. A computer programs in Microsoft FORTRAN-90 PowerStation 4.0 language and Microsoft Visual Basic 6.0 program of the selected distributions have been created. The calculation of T-years 2, 5, 10, 20, 50 and 100 years quantiles for 4-stations using the above mentioned distributions is as tabulated in Table 2.

The results are used to create the probability plot and frequency curves for different rivers. Figures 2-5 illustrated the outcome curves of four distributions used with Gringorten Formula. From the results and analysis, it is seen that the estimation quantiles of GEV distribution is always nearest to all observed data except at Kagera River at Kyaka Ferry station followed by GPA distribution. To get accurate assess, the comparison criterion of Root Mean Square Error (RMSE) is used to provide the best fit of a distribution function. The RMSE was computed as:

$$RMSE = \sqrt{\frac{1}{n}\sum\left(\frac{Q_{obs.} - Q_{comp.}}{Q_{obs.}}\right)^2} \quad (10)$$

The comparison of four distributions has been based on the results at upper tail $P(x) \geq 0.5$ when $T(x) \geq 2$. Table 2 gives the summarization of goodness of fit-results for the frequency distributions used at each station. This analysis also showed that GEV distribution provided the best fit results of all while EV1 distribution presented inability of fitting.

Table 2: Estimation of Peak Discharge by Different Frequency Distributions for various return periods and the comparison criteria.

Return period (Year)	**2**	**5**	**10**	**20**	**50**	**100**	**RMSE**
(a) Kagera River at Kyaka Ferry							
$Q_{obs.}$	559.17	918.53	1127.21	1295.97	1331.20	2334.50	
LOG: X_T	653.59	813.12	906.43	992.42	1101.43	1182.36	***0.176423***
EV1: X_T	688.57	828.42	887.88	931.56	975.87	1002.65	***0.141984***
GEV: X_T	569.23	860.19	1066.10	1264.58	1518.73	1706.16	***0.089342***
GPA: X_T	572.59	868.54	1002.31	1068.58	1134.39	1996.95	***0.067662***
(b) Bahr el Jebel at Mongalla							
$Q_{obs.}$	3060.00	4844.76	5253.11	6255.81	7341.43	7348.46	
LOG: X_T	3486.32	3936.49	4199.82	4442.46	4750.10	4978.48	***0.192322***
EV1: X_T	3585.03	3979.67	4147.46	4270.74	4395.77	4472.19	***0.175111***
GEV: X_T	3133.07	4428.09	5322.21	6172.48	7248.59	8034.83	***0.062556***
GPA: X_T	2997.29	4773.43	5653.28	6152.54	6471.70	6581.42	***0.055475***
(c) White Nile at Malakal							
$Q_{obs.}$	3313.44	3871.27	4572.64	4869.87	6111.54	6364.67	
LOG: X_T	3520.66	3696.87	3799.95	3894.92	4015.34	4104.74	***0.137254***
EV1: X_T	3559.30	3713.77	3779.45	3827.71	3876.65	3906.56	***0.128691***

GEV: X_T	3284.13	3948.28	4467.35	4995.90	5705.65	6249.32	***0.054138***
GPA: X_T	3191.14	4136.80	4709.04	5056.56	5286.49	5366.86	***0.023744***
(d) Blue Nile at Sennar							
$Q_{obs.}$	15100.00	17239.86	18355.74	19498.71	24071.58	24423.16	
LOG: X_T	15378.63	15475.48	15531.94	15583.42	15644.78	15695.94	***0.133942***
EV1: X_T	15357.39	15471.62	15547.01	15618.54	15705.57	15778.79	***0.134658***
GEV: X_T	15240.74	17663.56	18953.47	20019.81	21215.96	22009.24	***0.057945***
GPA: X_T	15189.56	18147.61	19195.34	19729.12	20052.40	20160.66	***0.036385***

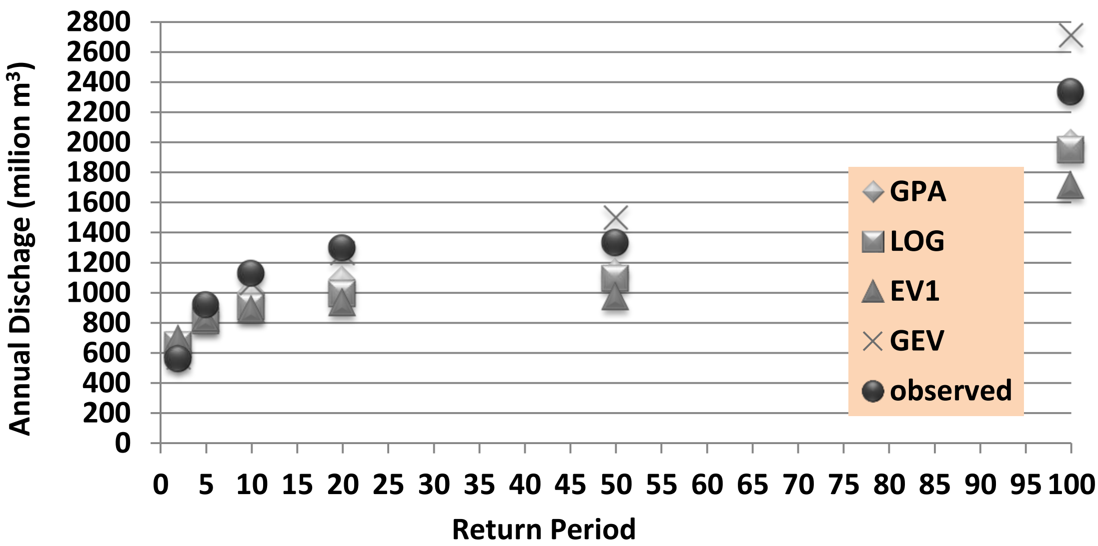

Figure 2: Frequency Curve at the Kagera River at Kyaka Ferry

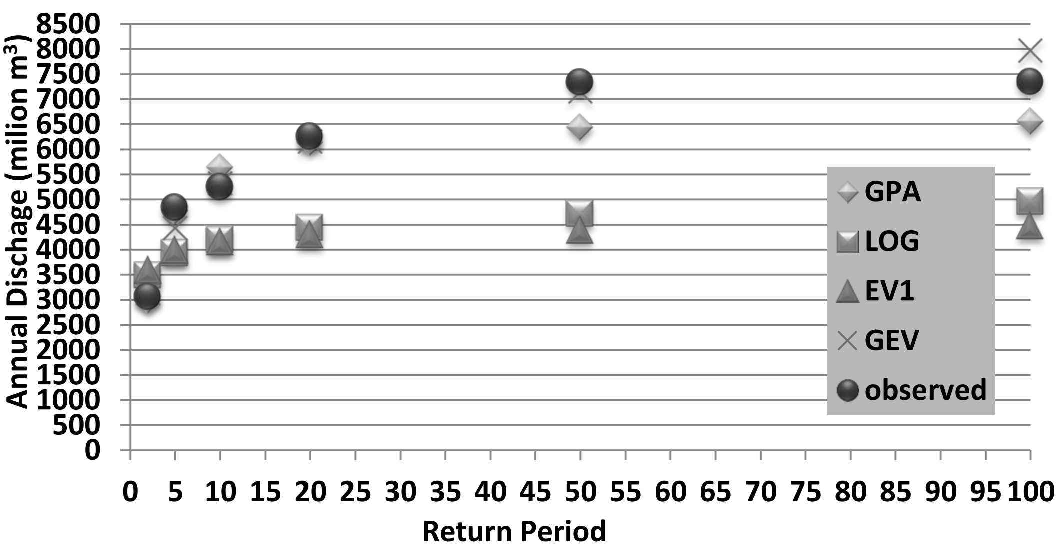

Figure 3: Frequency Curve at the Bahr el Jebel at Mongalla

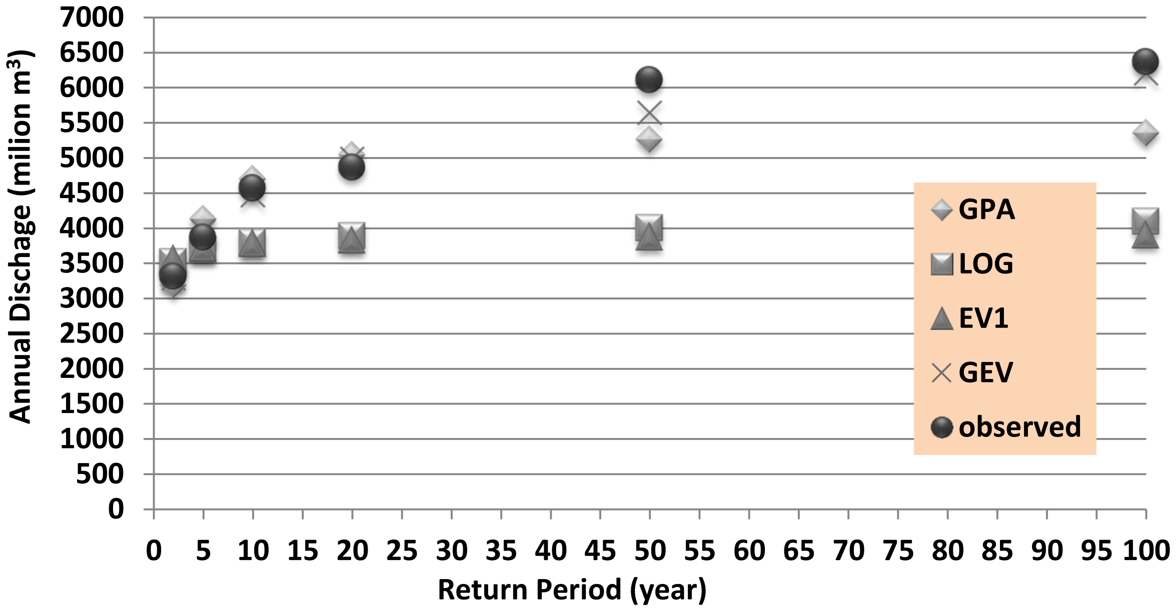

Figure 4: Frequency Curve at the White Nile at Malakal

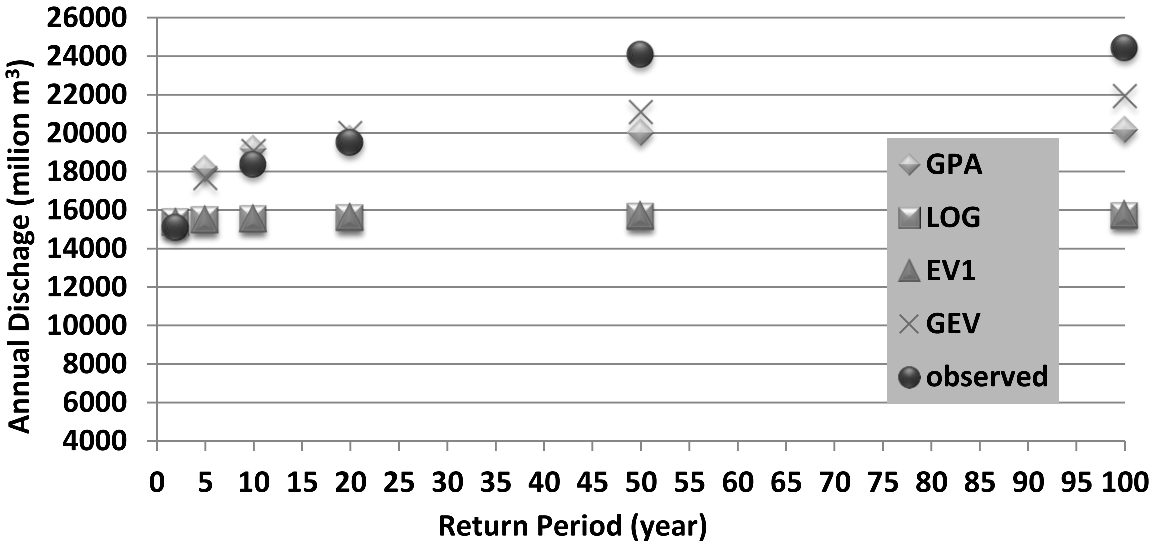

Figure 5: Frequency Curve at the Blue Nile at Sennar

5. Conclusions

The results submitted a precept for selecting the GEV distribution using L-moments for modeling at-site quantile series as the best distribution than other candidates for estimating the design discharges in Nile River tributaries. However the GPA distribution could also be used as followed by GEV distribution. Thus, both GEV and GPA distributions are good estimating annual peak discharge in the Nile River Basin.

6. References

Abdul Karim M., and Chowdhury J. U., (1995), "A comparison of four distributions used in flood frequency analysis in Bangladesh", Hydrological Sciences, 40(1), 55-66.

Abida H., & Ellouze M., (2008), "Probability distribution of flood flows in Tunisia", Hydrology and Earth System Sciences, 12(3), 703-714.

Analysis F., (2000), "Frequency Analysis." 408-413.
Arnell N.W., (2002), "Hydrology and Global Environmental Change", Prentice-Hall, Harlow.
Benson M.A., (1968), "Uniform Flood Frequency Estimating Methods for Federal Agencies", Water Resour. Res, 4, 891-908.
Bhattarai K., (2005), "Flood frequency analysis of Irish river flow data using variants of L-moments", National Hydrology Seminar, (Civil), 46-55.
Cannarozzo M., D'Asaro F., and Ferro V., (1995), "Regional rainfall and flood frequency analysis for Sicily using the two component extreme value distribution", Journal of Hydrological Sciences, 40(1), 19-42.
Cassalho F., Beskow S., Rogério C., Martim M., Kerstner L., and Ávila L. F., (2018), "At-Site Flood Frequency Analysis Coupled with Multiparameter Probability Distributions", Water Resour Manage, 32:285-300.
Chow V.T., Maidment D.R., and Mays L.R., (1988), "Applied Hydrology", McGraw Hill Inc., New York.
Cunnane C., (1989), "Statistical Distributions for Flood Frequency Analysis", Operational Hydrology Report no. 33, World Meteorological Organization.
Das L. M., & Qureshi Z. H., (2014), "Flood Frequency Analysis for Jiya Dhol River of Brahmaputra Valley", International Journal of Sciences: Basic and Applied Research (IJSBAR), 14(2), 14-24.
Ferdows M., and HossainM., (2005), "Flood Frequency Analysis at Different Rivers in Bangladesh: A comparison Study on Probability Distribution Functions", Thammasat Int. J. Sc. Tech., Vol. 10, No. 3, July-September.
Ferro V., & Porto P., (2006), "Flood Frequency Analysis for Sicily, Italy", Journal of Hydrologic Engineering, 11(2), 110-122.
Greenwood J. A., Landwehr J. M., Matalas N. C., Wallis J. R., (1979), "Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form", Water Resources Research, 15, 1049-1054.
Hosking, J. R. M., (1990), "L-moments: Analysis and estimation of distributions using linear combinations of order statistics", J. R. Stat. Soc., Ser. B, 52, 105-124.
Hosking, J. R. M., Wallis J. R., (1997), "Regional frequency analysis; An approach based on L-Moments", Cambridge University Press, United Kingdom.
Kurun A., (2005), "Frequency Analysis of Low Flow Series from Çekerek Stream Basin" 11(1), 72-77.
Liongson L. Q., (2000), "Regional Flood Frequency Analysis for Philippine Rivers", World, 1-8.
Moharram S. H., (1995), "Frequency Analysis of Discharge Data in River Nile Basin", Inter. Symposium on Environmental Hydrology, ASCE EIG, Egypt.
Moharram S. H., Gosain A. K., and Kapoor P. N., (1993), "A Comparative Study for Estimators of the Generalized Pareto Distribution", Journal of Hydrology, Vol. 150.
Onni S., Selaman, Said S., and Putuhena F.J., (2007), "Flood Frequency Analysis for Sarawak using Weibull, Gringorten and L-moments Formula", Jornal-The Institution of Engineers, Malaysia. Vol. 68, No. 1.
Orman U., (2000), "Application of at Site and Regional Frequency Analyses by Using the L-Moments Technique", 11(3).
Pearson C.P., McKerchar A.I., and Woods R.A., (1991), "Regional flood frequency analysis of Western Australian data using L-moments", International Hydrology and Water Resources Symposium, Part 2, v 2, n 91, pt 22, p 631-632.

Rakesh K., and Chandranath C., (2006), "Regional Flood Frequency Analysis Using L-Moments for North Brahmaputra Region of India" J. Hydrologic Engrg., 11(4), 380-382.
Saf B., (2009), "Regional Flood Frequency Analysis Using L-Moments for the West Mediterranean Region of Turkey" Water Resources Management, 23(3), 531-551.
Singh V.P. and Guo H., (1995), "Parameter estimation for 3-parameter generalized Pareto distribution by the principle of maximum entropy (POME)" Hydrol. Sci., 40: 165-181.
Stedinger J. R., Vogel R. M., & Georgiou E. F., (1993), "Frequency analysis of extreme events", Chapter18 In: Handbook of Hydrology, ed. D. R. Maidment. McGraw-Hill, New York, USA.
Vogel R.M., and Wilson I., (1996), "The Probability Distribution of Annual Maximum, Minimum and Average Streamflow in the United States", J. Hydrol.Engg, ASCE 1, 69-76.
Vogel R.M., Thomas W.O., and McMahon T.A., (1993), "Flood Flow Frequency Model Selection in the Southwestern United States" J. Wat.Resour. Plan. Manage, ASCE, 119, 353-366.
Vogel R. M., McMahon T. A., and Chiew F. H. S., (1993), "Floodflow frequency model selection in Australia", J. Hydro., Amsterdam, The Netherlands, 146, 421-449.
Wang Q.J., (1990), "Estimation of the GEV Distribution from Censored Samples by Method of Partial Probability Weighted Moments", Journal of Hydrology, 120 103-114.