Flood Frequency Analysis at Hadejia River in Hadejia – Jama'are River Basin, Nigeria

Usman Abdullahi Ibrahim[1*] Sunday Gashi Yadima[1] Abdulhamid Nur Alkali[1]

1. Department of Civil and Water Resources Engineering, Faculty of Engineering, University of Maiduguri.
P. O. Box 1069, Maiduguri, Borno State, Nigeria

Abstract

The maximum annual streamflow data of River Hadejia gauging station obtained from the Hadejia Jama'are Ko madugu TrustFund, Damaturu for the period of 1963 to 2014 were subjected to flood frequency analysis. Three p robability distribution functions; Extreme Value Type 1 (EV-1), Lognormal (LN), and Log Pearson Type III (LP III) were used for the analysis. The models were used to predict and compare corresponding flood discharge esti mates at 2, 5, 10, 25, 50, 100 and 200 years return periods. The results for EV-1, LN and LP III at 200 years retur n period indicated predicted discharge values of 157.419, 169.43 and 135.21 respectively. From the results, logn ormal distribution model gives higher flood discharge estimates and therefore it is recommended to be utilized fo r safe design.

Keywords: Flood frequency, probability distribution, quantile, recurrence interval, discharge.

1. Introduction

A flood occurs when there is an unusual high stage of a stream or river. This is due to runoff from precipitations in quantities too large to be confined in the normal water surface elevations of the stream or river. This may result from unusual combination of meteorological factors (Mustapha and Yusuf, 2012).

How frequently a flood event of a given magnitude may be expected to occur is of great importance, because almost every activities on a particular flooded area might be controlled by it (Hosking and Wallis, 1997). Flood frequency analysis with various risks of exceedence is therefore needed for a wide range of engineering problems; planning for weather related emergencies, reservoir management, pollution control, and insurance risk calculations (Gottschalk and Krasovskaia, 2001; Kjeldsen *et al*, 2002; Saf, 2008).

In hydrological events, there are numerous and unpredictable sources of uncertainties about the physical processes (Hosking and Wallis, 1997). Thus, stochastic models (such as flood frequency analysis) are very important and desirable to estimate how often a specified event will occur on average in a particular area.

The primary objectives of flood frequency analysis are to determine the return periods and then to estimate the magnitudes of events for design return periods beyond the recorded range. The intermediary between these two objectives is the theoretical probability distribution. The fitted distribution is used to estimate event magnitudes corresponding to return periods less than or greater than those of the recorded events (Mustapha and Yusuf, 2012). However, it must be emphasized that the prediction is statistical and not guaranteed. Many factors such as a change in the precipitation pattern in the drainage basin, construction of artificial levees and dams, and deforestation and urbanization can introduce significant errors into the flood frequency analysis (Mustapha and Yusuf, 2012).

Return period, T, may be defined as the time interval for which a particular flood having magnitude Q_T (also known as quantiles) is expected to be exceeded (Mengistu, 2008).

Numerous probability distribution functions have been used to model phenomena such as stream flow, precipitation, etc., which are characterized by significant variability and not significantly explained by physical principles (Wurbs and James, 2009). However, the three most used probability distribution models for flood frequency analysis are: EV- I (Extreme value type -1), Log -normal, and Log Pearson Type III (Mustapha and Yusuf, 2012).

Extreme value type -1

EV-I (Extreme value type I) also referred to as Gumbel Maximum distribution is one of the most commonly used distribution in flood frequency analysis, the exponential probability function of largest values fits symmetrically the distribution of annual maximum flood events, and is given by:

$$P = exp[-exp(-y)] \quad \ldots \quad (1)$$

Where; P is the probability of occurrence of given flood being equaled or exceeded, and y is the reduced variate, which is a function of probability and is given by:

$$y = \frac{x-u}{\alpha} \quad \ldots \quad (2)$$

The parameters are estimated by the equation:

$$\alpha = \frac{\sqrt{6}}{\pi} S \quad \ldots \quad (3)$$

$$u = \bar{x} - 0.5772\alpha \quad \ldots \quad (4)$$

For a given return period *T*, the reduced variate y_T is given by:

$$y_T = -ln\left[ln\left(\frac{T}{T-1}\right)\right] \quad \ldots \quad (5)$$

x_T is related to y_T by:

$$x_T = u + \alpha y_T \quad \ldots \quad (6)$$

Lognormal

Natural phenomena have values greater than zero and may be unconstrained theoretically in the upper range. Lognormal distribution fits those conditions. The logarithms of the hydrological variables follow a normal distribution. Its probability distribution function (PDF) is given by:

$$p = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right) \quad \ldots \quad (7)$$

The two parameters; mean $\bar{x}$, and standard deviation, □ are given by:

$$\bar{x} = \frac{I}{N}\sum_{i=I}^{N} x_i \quad \ldots \quad (8)$$

Where x_i magnitude of the ith event and N the total number of events.

$$\sigma = \left[\frac{\sum_{i=I}^{N}(x-\bar{x})^2}{N-I}\right]^{\frac{I}{2}} \quad \ldots \quad (9)$$

The PDF of lognormal distribution is derived from substituting the equation in the normal:

$$y_i = \log_e x_i \quad \ldots \quad (10)$$

Since the logarithm of *x* follows a normal distribution, the mean and standard deviation becomes:

$$\overline{logx} = \frac{I}{N}\sum_{i=I}^{N} \log x_i \quad \ldots \quad (11)$$

$$\sigma \log x = \left[\frac{\sum_{i=I}^{N}(logx - \overline{logx})^2}{N-I}\right]^{\frac{I}{2}} \quad \ldots \quad (12)$$

Therefore, the probability of exceedence related to an occurrence period can be applied to the logarithms as:

$$\log x = \overline{\log x} + k\,□_{\log□} \quad \ldots \quad (13)$$

Where, *K* is the frequency factor.

Log Pearson Type III

This distribution involves the transformation of random variable x as a Pearson Type III from natural units to logarithmic units and computes the mean, standard deviation and skew coefficient parameters of the distribution. The mean is expressed as:

$$\overline{\log x} = \frac{\sum \log x}{n} \quad \ldots \quad (14)$$

Standard deviation:

$$\sigma \log x = \frac{\left[\sum(\log x - \overline{\log x})^2\right]^{\frac{1}{2}}}{n-1} \quad \ldots \quad (15)$$

Skew coefficient:

$$G = \frac{n\sum(\log x - \overline{\log x})^3}{(n-1)(n-2)(\sigma \log x)^3} \quad \ldots \quad (16)$$

For any probability level, the value of x is computed from:

$$\log x = \overline{\log x} + k\sigma_{\log x} \quad \ldots \quad (17)$$

The value of k can be obtained from tables in Standard Hydrology Textbooks.

2. Materials and Methods

2.1 Study Area

The study area is located in the North Western part of Nigeria on the Hadejia River. It falls within geohraphical coordinates of 12°26'N and 10°04'E and has a drainage area of 25,900km^2. The upstream section of the Hadejia River system lies on the largely impermeable Basement Complex rocks. The upstream Basement Complex region is hilly (with peaks of up to 1,200 m). In the upstream area, from 1980 onwards, there has been a tendency for the tree-dominated savannah to be replaced by land-use for rainfed agriculture and grazing (Afremedev, 1999). The middle and downstream parts are, except for some ancient sand dunes, relatively flat. Most of the flows in the Hadejia River system (~80%) is regulated by Tiga Dam. The Hadejia River splits into three channels in the Hadejia Nguru Wetland: The Marma channel which flows into the Nguru Lake, the old Hadejia River which joins up with the Jama'are River to become the Yobe River and the relatively small Burum Gana River (Goes, 2001). A map showing the Hadejia – Jama'are River Basin is presented in Figure 1.

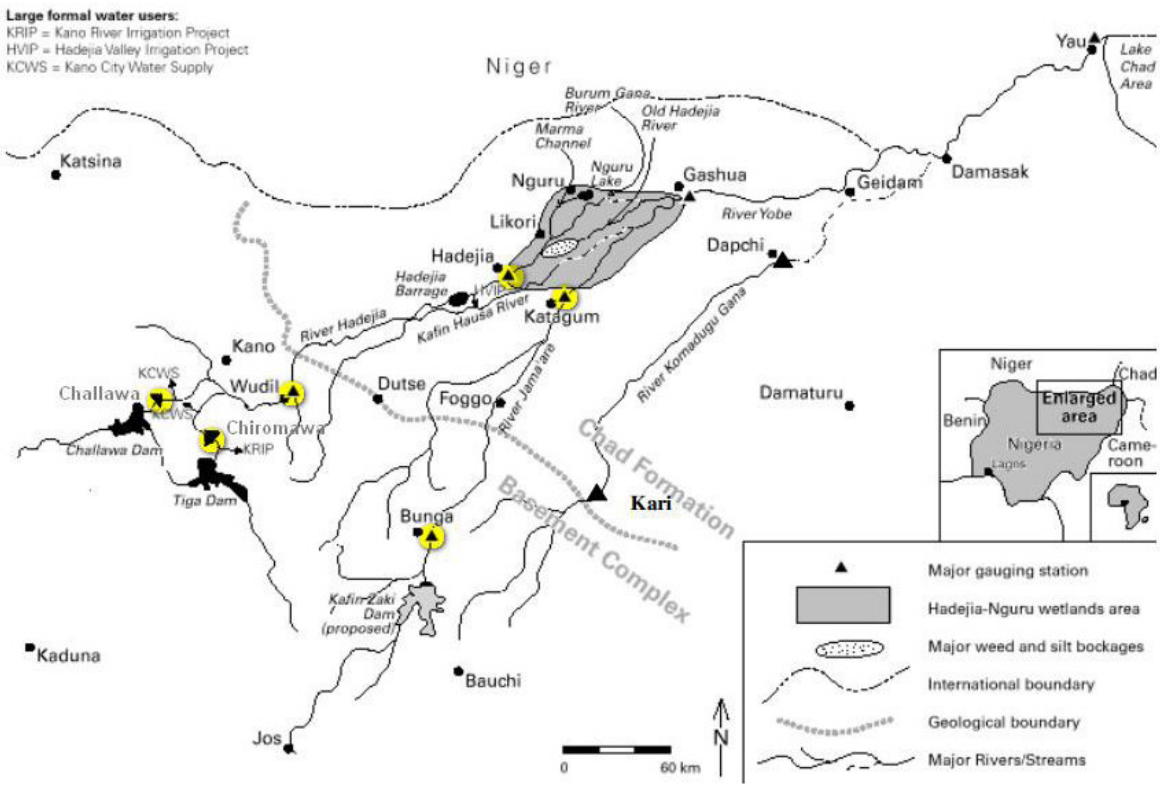

Figure 1: The Hadejia Jama'are Komadugu Yobe Basin

Source: Komadugu Yobe Basin Project (2006)

2.2 Methodology

2.2.1 Data Collection

Data were collected on discharge/streamflow for the period of 52 years (1963 – 2014). The data were subjected to Flood Frequency Analysis (FFA). Three probability distribution models namely: Extreme Value Type I (Gumbel), Lognormal and Log-Pearson Type III were utilized for the analysis.

2.2.2 Confidence Bands

Confidence band was constructed for the fitted distribution around the flood frequency curve to determine the reliability of points representing the annual maximum discharge X for a recurrence interval T (Chow *et al.*, 1988).

For estimating the event magnitude for return period T, the upper limit, $U_{T,\alpha}$ and lower limit, $L_{T,\alpha}$ may be specified by adjustment of equation (13).

$U_{T,}\alpha = \bar{y} + S_y K_T^U, \alpha$ … … … … … … … … … … … … …. (18)

And

$L_{T,}\alpha = \bar{y} + S_y K_T^L, \alpha$ … … … … … … … … … … … …. …. … (19)

Where K_T^U, α and K_T^L, α are the upper and lower confidence limit factors, which can be determined for normally distributed data using the non – central t – distribution (Kendall and Stuart, 1967). Approximate values for these factors are given by the following formulae (Natrella, 1963, U. S. Water Resources Council, 1981):

$$K_T^U, \alpha = \frac{K_T + \sqrt{K_T^2 - ab}}{a} \quad (20)$$

$$K_T^L, \alpha = \frac{K_T - \sqrt{K_T^2 - ab}}{a} \quad (21)$$

Where;

$$a = 1 - \frac{z_a^2}{2(n-1)} \quad (22)$$

$$b = K_T^2 - \frac{z_a^2}{n} \quad (23)$$

3. Results and Discussions

3.1 Extreme Value Type I Distribution

For the observed stream flow data, the mean ($\bar{x}$) = 72.62 and Standard deviation (S_x) = 23.0416. The parameters were estimated as; α= 17.97 and u= 62.25.

$$x_T = Q_T = 62.25 + 17.97 y_T$$

$$y_T = -ln\left[ln\left(\frac{T}{T-1}\right)\right]$$

For different values of y_T, the various Q_T were obtained.

Table 1. Predicted flood discharge values for different return periods for Extreme Value Type I Distribution

T (yrs)	y_T	*U*	α	αy_T	Q_T
2	0.367	62.25	17.97	6.595	68.845
5	1.500	62.25	17.97	26.955	89.205
10	2.250	62.25	17.97	40.433	102.683
25	3.199	62.25	17.97	57.486	119.739
50	3.902	62.25	17.97	70.119	132.369
100	4.600	62.25	17.97	82.662	144.912
200	5.296	62.25	17.97	95.169	157.419

Table 2. Calculation of 95% confidence limits for Extreme Value Type I Distribution

T (yrs)	2	5	10	25	50	100	200
Z_a	1.599	1.599	1.599	1.599	1.599	1.599	1.599
K_T	0	0.842	1.282	1.751	2.054	2.326	2.576
K_T^U, α	0.224	1.123	1.996	3.012	3.650	4.218	4.737
K_T^L, α	-0.224	0.604	0.633	0.580	0.563	0.554	0.547
$U_{T,}\alpha$	77.781	98.497	118.611	142.021	156.720	169.800	181.767
$L_{T,}\alpha$	67.459	86.537	87.213	85.970	85.602	85.380	85.227

3.2 Lognormal Probability Distribution

The observed annual maximum discharge values (*Q*) have been transformed to log values in Table 3 for subsequent analysis using lognormal distribution. The frequency factor (*K*) for normal and log normal for different return periods can be obtained from the tables in standard Hydrology textbooks for Pearson and Log Pearson distribution but with skew (*G*) equal to zero (Wurbs and James, 2009). The *K* values for Pearson Type III and Log- Pearson Type III distributions for zero skew coefficients are given in Table 4.

Table 3. Log transformed data for Lognormal/Log Pearson Type III Distributions

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Q	139.44	107.22	107.09	105.05	98.30	98.13	95.72	95.72	95.72	95.69	94.84	94.22
Log Q	2.14	2.03	2.03	2.02	1.99	1.99	1.98	1.98	1.98	1.98	1.98	1.97

	13	14	15	16	17	18	19	20	21	22	23	24	25	26
	93.57	92.05	84.87	84.73	83.62	82.94	79.48	79.48	78.61	78.11	77.15	77.15	76.67	74.60
	1.97	1.96	1.93	1.93	1.92	1.92	1.90	1.90	1.90	1.89	1.89	1.89	1.88	1.87

	27	28	29	30	31	32	33	34	35	36	37	38	39
	74.60	74.24	71.90	71.74	68.13	66.35	65.21	58.73	58.73	58.73	58.13	53.96	53.96
	1.87	1.87	1.86	1.86	1.83	1.82	1.81	1.77	1.77	1.77	1.76	1.73	1.73

	40	41	42	43	44	45	46	47	48	49	50	51	52
	52.70	52.64	52.64	49.68	48.31	48.19	47.96	47.15	41.30	38.98	34.56	33.21	24.44
	1.72	1.72	1.72	1.70	1.68	1.68	1.68	1.67	1.62	1.59	1.54	1.52	1.39

Sum	Mean	Stan. Dev	Skew
3776.34	72.62	23.0416	0.20703
95.54	1.84	0.15103	-0.7017

Table 4. K values for different return periods for lognormal distribution

Skew coefficient (G=0)	**Recurrence interval, T (yrs)**						
	2	5	10	25	50	100	200
K	0.000	0.842	1.282	1.751	2.054	2.326	2.576

From equation (17), the predicted flood discharge values for the lognormal distribution at different return periods and K values (Table 4) were obtained and presented in Table 5.

Table 5. Predicted flood discharge values for different return periods for lognormal distribution

T (yrs)	K_T	$\sigma_{\log x}$	$K_T\sigma_{\log x}$	$\overline{\log x}$	$y_T = \overline{\log x} + K_T\sigma_{\log x}$	$Q_T = 10(y_T)$(m³/s)
2	0	0.15103	0	1.84	1.840	69.18
5	0.842	0.15103	0.1272	1.84	1.967	92.68
10	1.282	0.15103	0.1936	1.84	2.034	108.14
25	1.751	0.15103	0.2645	1.84	2.105	127.35
50	2.054	0.15103	0.3102	1.84	2.150	141.25
100	2.326	0.15103	0.3513	1.84	2.191	155.24
200	2.576	0.15103	0.3891	1.84	2.229	169.43

3.3 Log Pearson Type III Distribution

In applying Log-Pearson type III probability distribution model to the annual maximum data of the river; the mean, standard deviation and skew coefficient of the log transformed data of Table 3, were obtained as; 1.84, 0.15103 and -0.7017 respectively. The K factor values for different return period values for skew coefficient (-0.7017) were obtained from tables available in standard Hydrology Text Books. The values of K_T for thesame return periods are presented in Table 6.

From equation (17), the predicted flood discharge values for the log Pearson Type III distribution at different return periods and K values (Table 6) were obtained and presented in Table 7.

Table 6. K values for different return periods for Log Pearson Type III distribution

Skew Coefficient (G= -0.7017)	**Recurrence interval, T (yrs)**						
	2	5	10	25	50	100	200
K	0.116	0.857	1.183	1.488	1.663	1.806	1.926

Table 7. Predicted flood discharge values for different return periods for log Pearson Type III distribution

T (yrs)	**K_T (G= -0.7017)**	$\sigma_{\log x}$	$K_T\sigma_{\log x}$	$\log x$	$y_T = \overline{\log x} + K_T\sigma_{\log x}$	$Q_T = 10^{(y_T)}$(m^3/s)
2	0.116	0.15103	0.018	1.84	1.858	72.11
5	0.857	0.15103	0.129	1.84	1.969	93.11
10	1.183	0.15103	0.179	1.84	2.019	104.47
25	1.488	0.15103	0.225	1.84	2.065	116.15
50	1.663	0.15103	0.251	1.84	2.091	123.31
100	1.806	0.15103	0.273	1.84	2.113	129.72
200	1.926	0.15103	0.291	1.84	2.131	135.21

The results of the 2, 5, 10, 25, 50, 100 and 200 years return periods frequency analysis based on maximum instantaneous flow recorded on the Hadejia River at Hadejia gauging station from 1963 to 2014 using Extreme value Type 1(EV-I), Lognormal (LN) and Log Pearson Type III (LPIII) distributions are summarized in Table 8.

Table 8. Summary of flood quantile estimates for different probability distributions

***T*(yrs)**	**Flood quantile estimates (m^3/s)**		
	Extreme Value Type 1	Lognormal	Log Pearson Type III
2	68.845	69.18	72.11
5	89.205	92.68	93.11
10	102.683	108.14	104.47
25	119.739	127.35	116.15
50	132.369	141.25	123.31
100	144.912	155.24	129.72
200	157.419	169.43	135.21

Table 9. Quantile estimates by distributions and percentage deviations from LN values

(yrs)	] V-1	N	PIII	% deviation of EV-1 values from values LN	% deviation of LPIII values from LN values
2	68.845	69.18	72.11	0.48	-4.24
5	89.205	92.68	93.11	3.75	-0.46
10	102.683	108.14	104.47	5.05	3.39
25	119.739	127.35	116.15	5.98	8.79
50	132.369	141.25	123.31	6.29	12.70
100	144.912	155.24	129.72	6.65	16.44
200	157.419	169.43	135.21	7.09	20.20

Table 9 presents the quantile estimates obtained for the specific return periods of 2, 5, 10, 25, 50, 100, 200 years obtained by fitting the three probability distribution models to the observed flood data and the computed percentage deviation of the EV-1 and LPIII quantile values from their corresponding LN values. As indicated in Table 9, the percentage deviation of the EV-1 predicted values from LN predicted values ranges from 0.48% at T= 2years to 7.09% at T =200years. However, the lognormal distribution predicted higher quantile values than the corresponding values predicted by EV-1 and LPIII distributions. Moreover, the percentage deviation of LPIII predicted values from LN predicted values for the corresponding return periods ranges from -4.24% at T=2 years to 20.20% at T= 200 years, and for lower return periods (up to T=10 years), LPIII predicted quantile estimates are higher than the corresponding values predicted by EV-1 distribution. Also, from the table, it can be seen that all the distributions gave similar magnitudes for corresponding return periods with the indicated degree of deviation, whereas, for corresponding values of return periods, the disparities in the percentage deviation of EV-1 and LPIII predicted quantile values increased with increasing return periods.

4. Conclusion

This study presents the flood frequency analysis of Hadejia River using the streamflow data at Hadejia gauging station recorded between 1963 and 2014, and subjecting same to three probability distribution models; Extreme Value Type 1 (EV-1), Lognormal (LN) and Log Pearson Type III (LPIII).

EV-1, LN and LPIII can be utilized for frequency analysis of the Hadejia River flood data. However, it is safer to use the lognormal distribution because it gives higher quantile magnitude.

5. Recommendations

At the end of this study, the following recommendations were made;

i. Establishment of Telemetric Data Collection Platform (DCP) station on the river is needed to carry out reliable design and operation of hydraulic structures and for flood plain and flood risk mapping.

ii. The use of GIS should be encouraged to introduce terrain analysis in flood prediction.

iii. Regional flood frequency analysis should be utilized to provide useful alternative to the single site analysis especially in cases where records are short and in situations where estimates are needed in ungauged sites.

References

Afremedev Consultancy Services Limited (1999): Hadejia – Jama'are – Komadugu Yobe Basin Catchment. Action Plan for Integrated Natural Resource Management, Nigeria.

Chow, V.T., D.R. Maidment and L.W. Mays. 1988. *Applied hydrology*. McGraw-Hill Book Company, New York, 572p.

Goes, B. J. M. (2001). Effects of damming the Hadejia River in semiarid Northern Nigeria lessons learnt for future Management. *Regional Management of Water Resources* (Proceedings of a symposium held during the Sixth IAHS Scientific Assembly at Maastricht, The Netherlands, July 2001. IAHS Publ. no. 268:73-153

Gottschalk, L., & Krasovskaia, I. (2001). *Regional flood frequency analysis, - a theoretical background*. Lecture notes in Stochastic Hydrology. Department of Geophysics. University of Oslo, Norway.

Hosking, J. R. M., & Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge: Cambridge University Press, England.

Kendall, M. G. and Stuart, A. (1967). The advanced theory of Statistics. Hafner, Volume 3. New York, USA.

Kjeldsen T.R., Rosbjerg D. (2002). Comparison of regional index flood estimation procedures based on the extreme value type I distribution. *Stochastic Environmental Research and Risk AssP a g e | 131essment.* 16(5):358-373.

KYBP (2006). *Komadugu Yobe Basin Water Audit Report by Afremedev Consultancy Services Limited for FMWR-IUCN-NCF Komadugu Yobe Basin Project, Kano, Nigeria.*

Mengistu, D. (2008). Regional Flood Frequency Analysis for Upper Awash Sub- Basin (Upstream of Koka). Dissertation. Civil Engineering, Engineering. Addis Ababa University. Ethiopia.

Mustapha, S. and Yusuf, M. I. (2012). *A Textbook of Hydrology and Water Resources*. Topsmerit page publishers, Abuja, Nigeria.

Natrella, M. G. (1963). Experimental Statistics NBS Handbook 91. U.S. Government printing office, Washington DC, USA.

Saf, B. (2008). Application of index procedures to flood frequency analysis in Turkey. *Journal of the American Water Resources Association, 44*(1):37-47.

United States Water Resources Council (USWRC) (1981) Guidelines for Determining Flood Flow Frequency. Bulletin No. 17B, 15-19.

Wurbs, R. A. and James, W. P. (2009). Water Resources Engineering, PHI Learning Private Limited, New Delhi, India.