

# THE BEST FITTED DISTRIBUTION FOR ESTIMATING ANNUAL PEAK FLOWS IN BENUE RIVER, UPPER BENUE RIVER BASIN TROUGH

BY

Tanko<sup>1</sup>, J.A. and Agunwamba<sup>2</sup>, J.C.

1. Department of Civil Engineering  
Federal University of Technology, Yola

2. Department of Civil Engineering,  
University of Nigeria, Nsukka

## ABSTRACT

*Frequency analysis is widely applied for estimating extreme flows and precipitation. However, its merits for applicability based on the various distributions available for different data and purposes have not been clearly established especially for River Benue in the Upper Benue River Basin. The annual peak flows of river Benue were taken at the Jimeta – Yola, bridge gauging station in Adamawa State for a 45 – years record from 1960 – 2004. The record was fitted with Normal, Log-Normal, Gumbel's Extreme values, Gamma, and Pearson type III distributions. Log- Normal distribution was found most suitable for annual peak flow estimation in the Benue River based on the goodness of fit test using Chi- Square test.*

**Key words:** Peak flows, Frequency analysis distributions and goodness of fit

## 1. INTRODUCTION

Large peak flows (floods) usually impact man and his property. Several hydrological models are in use among which are LISFLOOD.FP and GLUE models (1-3) used for flood forecasting and for assessing predictive uncertainty respectively. Frequency analysis is used for analyzing stations with more than 10 years peak flow data. However, the choice of the distribution to be fitted for analysis remains a problem for the Hydrologist. WMO (4) outlined the importance of adopting the best fitted distribution for each river Basin in various regions. Yadav and Lal (1998)

selected Log-normal distribution as the best fitted distribution for Rapti river system in Eastern Urta Padesh, India. Several distributions are available but the applicability of each to a particular situation had not been checked especially for the Upper Benue Basin Rivers. The use of appropriate distribution would ensure good estimation of design floods for hydraulic structures design.

Today more than 40% of the catastrophes are caused by flooding alone with an average of about 10,000 deaths per year as observed by Lafittle and Bartle (5) and Hunt (6). Notable flood associated

catastrophes include the Germany flood disaster in 2002 that claimed over 100 lives with property and infrastructure damaged as reported by the ravaging hurricane reported in Cartrina, USA (7), which left a number of deaths with the destruction of properties worth millions of dollars. Many similar cases were reported in Asia; Tsunami wave flood disaster of 2004 that claimed more than 177, 000 lives in addition to millions of dollars worth of property damaged as reported by Hunt (6).

The middle East and Africa are not left in Nigeria, flood incidences are on the increase as observed by Ojiako (8). The new Nigeria in her editorial column also stated that “floods are in the news again damaging millions of Naira worth of property, rendering hundreds homeless and occasionally with loss of lives (New Nigeria, 1986). The recent flood disasters include: Loko flood, 2004, Gombe flood, 2005; Jalingo flood, 2005 (10, 11). These floods were experienced in the Upper Benue River Basin trough alone. It is necessary, therefore, that flood estimation techniques be assessed with a view to establish more suitable techniques. This would ensure adequate preparations for the worst case. In addition, the hydraulic and structural design of dams, bridges, spurs, and river works are all based on the magnitude and probable frequency of occurrence of floods, as observed by WMO (4). The frequency of occurrence, maximum stage reached, areas inundated and the duration of the floods are important for planning and design of flood control works and other hydraulic structures as suggested by (12).

## 2. Theory of Frequency Analysis

The following are the theory of some selected distributions and their fitting techniques:

**2.1 Normal Distribution:** The general equation for the Normal distribution curve is given by (4, 13, 14) as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

For  $-\infty \leq x \leq +\infty$

Where,  $f(x)$  = the height of the curve of probability density distribution of variable  $x$

$x$  = random variable,  $\mu$  = variable mean,  $\sigma$  = standard deviation of variables and  $f$  = frequency of discharge  $Q_i$

The above equation was fitted to the peak flow variables as

$$y_i = \frac{N_i}{\sigma\sqrt{2\pi}} e^{-(Q_i - Q_m)^2/2\sigma^2}$$

Where  $N_i$  = frequency in class interval,  $i$ ;  $y_i$  = frequency for discharge;  $Q_i$  and  $Q_m$  = the discharge in class interval  $i$  and mean discharge respectively.

$$Q_m = \sum f Q_i / \sum f$$

$$\sigma = \left( \sum f (Q_i - Q_m)^2 / \sum f \right)^{1/2}$$

data ...(4)

**2.2. Long-Normal Distribution:** the peak flow discharge values ( $Q$ ) were replaced by their logarithmic values in equation (2) as suggested by Yevjevich (15)

$$f(Q) || ABS(dQ) = \Phi(\log Q) ABS(\log Q)$$

Where,  $ABS$  = absolute value of the parenthesis.

For  $Q \geq 0$  and  $f(Q) = \frac{1}{2.303} \phi \log Q$

And by analogy,  $(\log Q)$  could be expressed in terms of normal distribution and equation (5) was written as

$$f(Q) = \frac{0.4342}{Q\sigma_m\sqrt{2\pi}} I^{-(\log Q - \log Q_m)^2/2\sigma^2}$$

Where,  $\sigma_m$  = standard deviation of  $\log Q$ .  
 $\log Q_m$  = mean of  $\log Q$

A theoretical basis exists for the log – normal distribution by considering causative factors as having positive and multiplicative, rather than additive effects. Therefore, the logarithms of factors should satisfy the basic four conditions for normal distributions as suggested by Yadav and Lal (5). After simplification, equation (7) was fitted to the flood discharges as

$$y_i = \frac{N_i}{2.303 Q\sigma_m} e^{-(\log Q - \log Q_m)^2/2\sigma^2}$$

Where, all terms retain their earlier meaning.

**2.3. Gumbel’s Extreme Values:** The following assumptions were made for this distribution.

- P The distribution is of exponential type.
- P The number of observations should sufficiently large, and
- P Observations are assumed to be independent variables.

The Gumbel’s Extreme Value distribution is given as suggested by Yadav and Lal (5) and Subramanya (13) as

$$f(Q) = e^{-e^{-y}}$$

For  $- < Q <$  and  $Y = (Q - U)$  (10)  
 The parameters and  $U$  were estimated using the method of Maximum Likelihood

(MLH). The values of and  $U$  estimated from the peak flow discharge record were  $1.15 \times 10^{-3}$  and 3620 respectively.

The annual peak flow discharges were fitted with

$$y_i = N_i \alpha e^{-\alpha(Q_i - U)} e^{-e^{-\alpha(Q_i - U)}} \tag{7}$$

**2.4 Gamma Distribution with 2 Parameters**

The general form of this function is given by

$$f(x) = \frac{1}{\beta^\alpha \Gamma \alpha} Q^{\alpha-1} e^{-x/\beta}$$

Equation (12) was fitted to the peak flow record by

$$y_i = \frac{N_i}{\beta^\alpha \Gamma \alpha} Q_i^{\alpha-1} e^{-Q_i/\beta}$$

Where,  $N_i$  and  $y_i$  retain<sup>(8)</sup> their earlier meaning, = shape parameter and = scale parameter

The values of and were estimated in terms of  $Q_m$  and using the relations suggested by Yadav and Lal (5) as

$$\alpha = Q_m^2 / \sigma^2 \tag{14}$$

$$\beta = \sigma^2 / Q_m \tag{15}$$

Where,  $Q_m$  and retain their earlier meanings.

**2.5 Pearson Type III Distribution:** The

general function for this distribution is as suggested by Mustafa and Yusuf (14) Sabramanya<sup>(9)</sup> (13).

$$f_a(x) = \frac{1}{\beta^\alpha \Gamma \alpha} (x_i - r)^{\alpha-1} I^{(x-r)/\beta}$$

$$f_a(x) = \frac{1}{\beta^\alpha \Gamma \alpha} (x_i - r)^{\alpha-1} e^{-(x-r)/\beta}$$

Where  $f_a$  = three parameter probability density function,  $r$  = location parameter. The Pearson type III distribution is a special case of distribution often used in hydrological frequency analysis. The equation used for this distribution is given by.

$$y_i = \frac{N_i}{\beta^\alpha \Gamma \alpha} (Q_i - r)^{\alpha-1} e^{-yQ_i - r01B}$$

The skewness coefficient is given by Pearson (15) as

$$r_i = \sum (Q_i - Q_m)^3 / (N - 1)\sigma^3$$

Where,  $\alpha = \frac{4}{r_1^2}$

$$\beta = \sigma r_i / 2$$

$$r = Q_m - (2\sigma / r_1)$$

## RESULTS

### 4.0

The observed and estimated frequencies of the peak flow discharges are shown in Table 1.0. The plot is shown in figure 1.0. The values for  $r$  and  $Q_m$  were estimated from the peak flow data as  $1.8276 \times 10^3$  and  $4.1911 \times 10^3 \text{ m}^3/\text{s}$  respectively.

The estimated and observed frequencies of the peak flow discharges and the plot are shown in Table 1.0 Figure 1.0 respectively. The frequencies were re-estimated using equation (11). The observed and the estimated frequencies for the peak flow discharges are shown in Table 1.0. The plot is shown in Figure 1.0.

The parameters  $\alpha$  and  $\beta$  were estimated from the peak flow record as 5 and 0.7978 respectively. The estimated and

the observed frequencies for peak flows in the Benue River are shown in Table 1.0. The plot of estimated frequencies using various distributions is shown in Figure 1.0.

The parameters  $\alpha$ ,  $\beta$ ,  $r$  and  $r_1$  were estimated from the peak flow record as 2.0587, 1.2744, 1.5674 and 1.3939 respectively. The observed and estimated frequencies for the peak flow and the plot are shown in Table 1.0 and Figure 1.0 respectively.

$$(17)$$

### 3. Materials and methods

Flood frequency analysis using Annual Flood Series (AFS) was adopted because of its suitability for sites with available historical peak flow records. A peak flow record of 45 years from 1960 to 2004 taken at the Jimeta – Yola Bridge gauging station was used for the analysis.

Guidelines for using frequency analysis were also suggested by Yadav and Lal (5) as follows:

(20) After processing historical flow records, a theoretical frequency distribution is chosen.

Parameters for the distributions are estimated using the available techniques.

Goodness of fit criterion chosen for the best fitted distribution based on the criterion selected.

The floods for different recurrence intervals are then chosen using the estimated parameters of the best fitted distribution.

Peak flow estimation distributions selected in literature for this study include Normal, Log-Normal, Gumbel's Extreme Values, Gamma with 2 parameters, and Pearson Type III (5).

Table 1.0: The observed and estimated frequencies for the Benue River for the various distributions.

S/No	Class Inter (x1000 m <sup>3</sup> /s)	Mid values (x1000 m <sup>3</sup> /s)	Observed Freq. (1)	Estimated Frequencies				
				Normal (2)	Log-Normal (3)	Gumbel EVI (4)	Gamma with 2 Parameters (5)	Pearson type III (6)
1	0-1.59	0.80	3.00	2.80	4.20	0.70	2.90	3.00
2	1.60-2.59	2.10	7.00	4.00	6.40	7.40	6.30	9.00
3	2.60-3.59	3.10	6.00	5.50	8.40	10.00	7.30	10.90
4	3.60-4.59	4.10	11.0	12.00	13.00	20.00	11.80	15.60
5	4.60-5.59	5.10	9.00	8.50	8.00	8.00	6.60	8.30
6	5.60-5.59	6.10	4.00	2.60	3.00	1.30	2.00	2.20
7	6.60-7.59	7.10	3.00	0.90	2.00	0.30	0.70	0.90
8	7.60-8.59	8.10	2.00	0.20	1.00	0.03	0.20	0.30

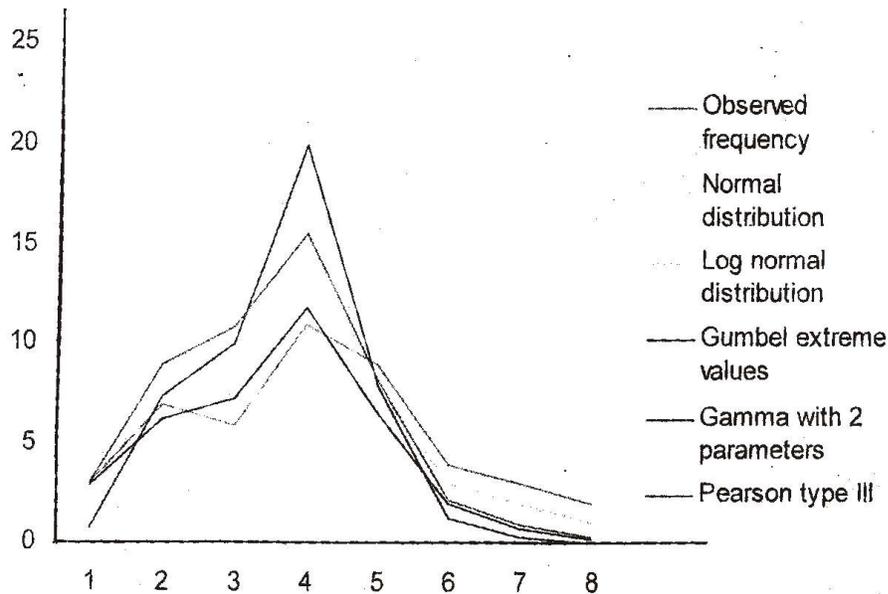


Figure 1.0: Comparison between estimated frequency and that for various distributions

**5. TEST OF GOODNESS OF FIT**

The goodness of fit of the distributions to the annual peak flows of the Benue River in Upper Benue River Basin trough was done using Chi-Square test as used by Yadav and Lal (5). It is given by

$$\chi^2 = \sum_{i=1}^k \frac{(b - c_{ij})^2}{c_i}$$

Where  $i = 1, \dots, k$  class interval covering the range.

$b_i$  = Number of observations actually in a

given class interval,  $c_i$  = Expected number of observations in a given class interval.

In a Chi-Square test, a critical value  $\chi^2_0$  of  $\chi^2$  for a significance level  $\alpha$  so that for  $\chi^2 < \chi^2_0$ , Null hypothesis of good fit is accepted. And for  $\chi^2 \geq \chi^2_0$ , Null hypothesis of good fit rejected. The value of  $\chi^2_0$  is usually obtained for a given number of degree of freedom (NDF) at a particular level of significance usually taken as 5% from standard statistics text book. The estimated values of  $\chi^2$  are shown in Table 2.

Table 2: Chi-Square test of fit for various distributions.

Distribution	Chi-Square $\chi^2$ for 5% level NDF = 2, $\chi^2_0 = 5.99$
1. Normal	24.27
2. Log-Normal	3.35
3. Gumbel's extreme Values	172.56
4. Gamma with 2 Parameters	26.99
5. Pearson Type III	20.30

## 6. CONCLUSION

This research determined the best distribution for estimation of the annual peak flows in Benue River. Records on annual peak flows of River Benue were collected at Jimeta – Yola, Adamawa State for a 45 – years record from 1960 – 2004. the data was fitted to five different distributions and their fitness compared using Chi-Square.

From the goodness of fit test, the Log-Normal distribution would be more suitable for annual peak flow estimation in the Benue River, Upper Benue Basin through.

**Acknowledgment:** This is to acknowledge the staff and management of the Upper Benue River Basin Development authority, Yola for providing all the necessary data

for the research. My gratitude also goes to Engr. (Prof.) N. Egbuniwe for guidance and encouragement.

## REFERENCES

- Gouweleeuw, B.T., Thielen, J., Franchello, G., De Roo, A.P.J., and Buizza, R., Food forecasting using probabilistic weather prediction. *Hydrol., Earth Syst. Sci.*, 9, 87-102, 2005.
- Hunter, N.M., Bates, P.D. and Horrieth, M.S. Utility of different data types for Calibrating flood inundation models within a GLUE frame work. *Hydrol., Earth Syst. Sci.* 9, 134-152, 2005.
- Pappenberger, F., Beven, K.J., Hunter, N., Bates, P., Gouweleeuw, B.T., and Thielen, J. Forecast (10 days) through a rainfall-runoff to flood inundation predictions within the European Flood Forecasting System (EFFS). *Hydrol., Earth Syst. Sci.*, 9, 103-115, 2005.
- WMO Statistical distribution for flood frequency analysis. *Operational Hydrology Report, No. 33, World Meteorological Organization, Geneva-Switzerland.* 1-29, 1989.
- Lafitte, R. and Bartle A (2001): Dams and Water Management. *Official Journal of the Water Works Council.* 3, 163, 2001.
- Yadav, R. and Lal, B.B. Pande Best fitted distribution for estimation of future flood for Rapti river system in eastern Uttar Pradesh. *Indian Journal of Engineering and Material Science.* 5, 22-27, 1998.
- Hunt, J.C. Inland and coastal flooding: Development in prediction and prevention. *Phil. Trans. Royal Society A* 363, 1475-1491, 2005.
- CNN “World News” Cable Network News, 8<sup>th</sup> August 2005.
- Ojiako G.U. “Floods in Nigeria and their optimal management strategy”. *Third World strategies for Technological Development, Federal University of*

- Technology, Yola. 360-367, 1990.
10. New Nigeria (1986): Editorial comment on floods of August 13<sup>th</sup>, 1, 1986.
  11. NTA (2004): Net work news. Nigerian Television Authority, Yola, 2004.
  12. NTA (2005): NET work news. Nigeria Television Authority, Yola, 2005.
  13. Lal, B.B. Pande and Tiwari, S.P. Generalized formula for flood estimation for Indian rivers. Fourth National Symposium on hydrology of Minor Water Resources scheme, Madras, India. 109-113, 1991.
  14. Sabramanya, S. Engineering Hydrology. Second Edition, Tata McGraw-Hill Publishing Company Ltd, New Delhi, India, 1994.
  15. Mustafa S. and Yusuf M.I. A text of hydrology and water resources. First Edition, Craft Technics Publishers, Nigeria, 1991
  16. Yevjevich, V. (1972): Probability and Statistics in Hydrology. Water Resources Publication, Port Collins, Colorado, USA
  17. Pearson K. Tables of the incomplete gamma function. Cambridge University Press, UK, 1921.