International Journal of Modern Engineering Research (IJMER)
www.ijmer.com Vol. 3, Issue. 5, Sep - Oct. 2013 pp-2848-2852 ISSN: 2249-6645

Fitting Probability Distribution Functions To Discharge Variability Of Kaduna River

Haruna Garba[1], Abubakar Ismail[2] and Usman Tsoho[1]
[1] *Department of Civil Engineering, Nigerian Defence Academy,*
Department of Water Resources and Environmental Engineering, Ahmadu Bello University Zaria Email:

ABSTRACT: *In reliability based design of flood protection structures, frequency analysis aimed at determining the occurrence frequency or return periods is an indispensable tool. A frequency analysis was performed by fitting probability distribution functions of Normal, Log Normal Log pearson type III and Gumbel to the discharge variability of Kaduna River at Kaduna South Water Works. The Kolmongonov- Smirnov (K-S) goodness-of-fit test was used to check whether the mean annual discharge variability of the river basin is consistent with a regional GEV distribution for the site. From the measure of discrepancy, it was observed that at selected level of significance of α=1%, α=5% and α=10%, all the four theoretical distributions functions were acceptable.*

Keywords: *Discharge, Flood, Frequency distribution, Probability distribution function, Statistical analysis*

I. INTRODUCTION

Hydrologist must plan for extreme events. Dams must be built high enough to constrain or limit extreme floods, while bridges must be built high enough to remain above high water mark. The problem of how to plan for extreme events considering that they are so rare and that there is no comprehensive data is therefore a challenge to hydrologist. In [1] the information and data recorded in the past is normally applied to obtain statistical parameters that can be used to forecast events that may occur in the future. Although in [2] it has been recognized that many annual flood series are too short to allow for a reliable estimation of extreme events because of difficulties in the identification of the appropriate statistical distribution for describing the data and to the estimation of parameters of the selected distribution. Probability distribution therefore is a statistical tool most widely used in flood prediction and estimation. Hydrological variables such as rainfall, temperature, discharges etc are statistical quantities that can be estimated from a given probability of occurrence. [3] observed that the study of river hydrology is important in the overall understanding of river systems, which is a key component for river engineering and restoration of water resources planning and for river ecosystem studies. [3] Further opined that understanding the hydrology and geomorphology of river systems is important in conserving their natural beauty, habitat and resources. River hydrology influences the temporal and spatial distribution of discharge including water availability within a region. The monitoring of river discharge according to [2] is ideally suited to detect and monitor changes resulting from climate change. [4] Observed that floods from rainfall may appear several times within one year in Kaduna metropolis. [4] Further adduced that urbanization and structural development into traditional flood prone areas of Kaduna River modifies the channel shape, reduces the width of flood prone areas relative to bank full width at every point along Kaduna River channel and modifies the flood prone containment characteristics of the channel. In determination of probability distribution of maximum discharges on river basins of observed floods, [5] adopted the hypothesis of mutual independence of floods events in the computation of the annual probability of maximum seasonal discharges to take into account all suitable define floods. Flood frequency analysis is used to predict design floods for sites along a river. The technique involved using observed annual peak flow discharge data to calculate statistical information such as mean values, standard deviations, skewness and recurrence intervals. The statistical data are then used to construct frequency distributions which are graphs and tables that tell the likelihood of various discharges as a function of recurrence interval or exceedence probability. If a flood of a particular size occurs on average ones every Tr years, then Tr is called the return period and the probability P of such an event in any year is given by equation (1).

$$P = \frac{1}{Tr} \qquad (1)$$

Then the probability that there will be no such flood on a particular year is $(1-P)$ and the probability that there will be no such flood over the next n years is given by equation (2).

$$P = (1 - P)^{n} \qquad (2)$$

By adopting the hypothesis of mutual interdependence of floods events, it becomes possible in computation of annual probability of maximum seasonal discharge to take into account all suitable defined flows given by equation (3).

$$P = P_2\,[1 - (1 - P^1)\lambda^1] \qquad (3)$$

Where P is the probability referred to the year, P^1 is the probability referred to all floods while P_2 and λ^1 are the transformation parameters.

Flood frequency analysis according to [6] is a statistical analysis of floods, their magnitude and frequency. According to [6], to derive the risk of occurrence of any flood event, the frequency distribution which can best describe the past characteristics of the magnitude and possibility of such flood must be known and this requires determination of the most appropriate flood frequency model which can be fitted to the available historical data or record. Flood frequency distribution can take on many forms according to the equations used to carry out the statistical analysis. [7] observed that in many areas of civil engineering the questions often arises which probability should be used to model the load and resistance, instead of choosing one probability distribution.[7] further observed that it is possible to consider various probability distributions and to attached weight to these distributions according to how good the fits are. A probability distribution for the which the standard deviation of its prediction is large should be given less weight relative to those distributions that exhibit less scatter. A probability distribution is a continues mathematical expression that determines the probability of a particular event. Four of the common forms of probability distribution selected for this study are Normal, Log Normal, Log Pearson type 111 and Gumbel max distributions. In this study, observed maximum discharges of Kaduna River were used as the basis for fitting probability distribution to the design discharges.

II. MATERIALS AND METHODS

2.1 The Study Area

Kaduna River (fig 1) is the main tributary of Niger River in central Nigeria. It rises on the Jos plateau south west of Jos town in a North West direction to the north east of Kaduna town. It then adopts a south westerly and southerly course before completing its flow to the Niger River at Mureji. Most of its course passes through open savanna woodlands but its lower section cut several gorges including the granite ravine at Shiroro above its entrance into the extensive Niger flood plains.

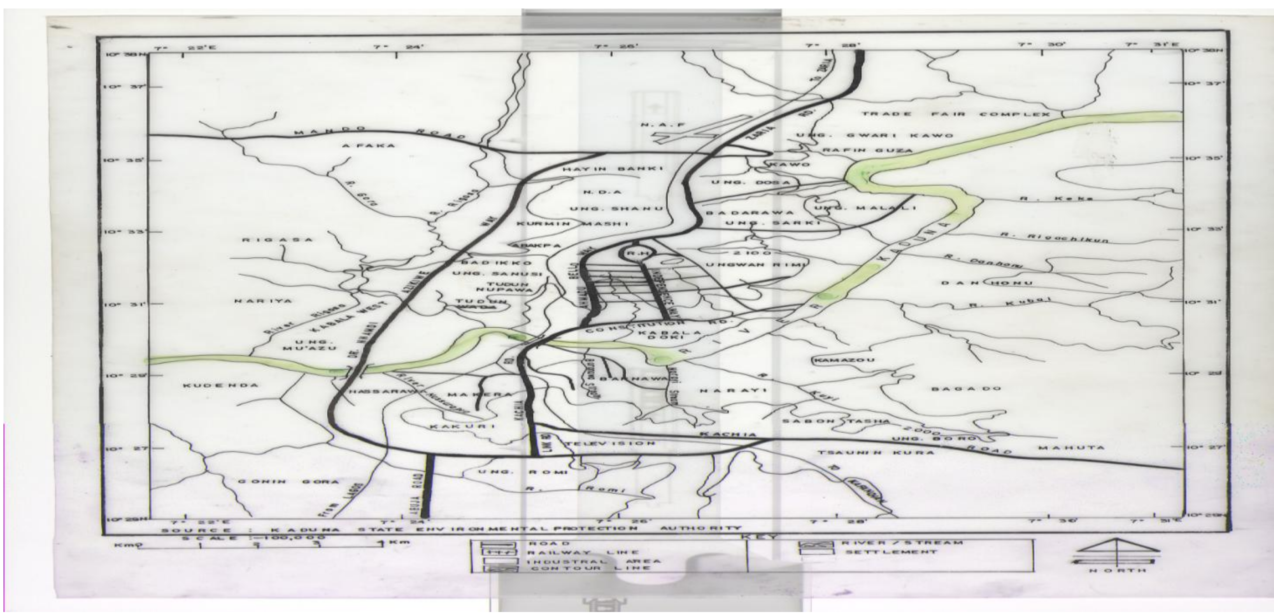


Fig 1 drainage map of Kaduna River

2.2. Data Used

In the study, the peak annual mean discharge data of Kaduna River for 11 years (2000-2010) were obtained from Kaduna State Water Board. To check the consistency of the data, a mass curve analysis was carried, to estimate missing record arising from human error, instrumental defect and improper citing.

2.3 The Flood Frequency Model

Hydrognomon is an open sources software tool used for the processing of hydrological data. Data are usually imported through standard text files, spread sheets or by typing. The available processing techniques for the tool includes time step aggregation and regularization, interpolation, regression analysis and infilling of missing values, consistency test, data filtering, graphical and tabular visualization of time series. Hydrognomon support several time step from the finest minutes scales up to decades. The programme also include common hydrological application such as evapotranspiration modeling, stage discharge analysis, homogeneity test, areal integration of point data series, processing of hydrometric data as well as lumped hydrological modeling with automatic calibration facilities (fig 2)

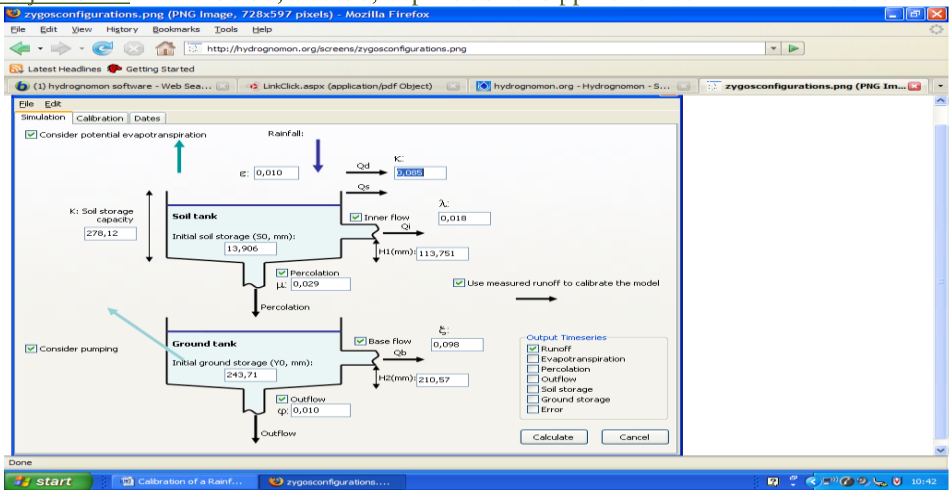


Figure 2 structure of the simulation module

2.4 The Procedure

Prior to the model application, the hydrognomon model was calibrated by [8] with historical data to obtained parameters to be used in studying the behaviors of Kaduna River. In order to assess the annual mean discharges with different return periods, or the occurrence frequencies of the return periods of extreme river floods reference was made flood frequency to predict the design floods at Kaduna South Water Works along Kaduna River as shown in fig 3. The hydrognomon probability distribution module was used to predict the likely values of discharges to expect along Kaduna River at various recurrence interval. For determining the most suitable statistical distribution that has the best fitting with the predicted values, the phytia statistical module was applied to the four theoretical distribution functions. The method of selecting the distribution and fitting values is the Kolmongonov- Smirnov (K-S) test.

2.5 Extrapolation of Point Measures to Watershed

The point measure of rainfall depth and intensity from a gauge is of value for estimating volume and runoff for large areas. For this, (a) the depth and intensity measured at a point must be considered as constant over an area, or (b) two or more point measures must be averaged .Intensity and Watershed Discharge. The calculated runoff considering the volume of water shed was obtained using the formula in equation (4).

$$Q = ICA. \qquad (4)$$

Where;Q is the Calculated runoff. I is the gauged water levels .C is a factor; the ratio maximum guage level at a point to the mean guage levels of Kaduna river. A is the drainage area of Kaduna river(18277.28km²).

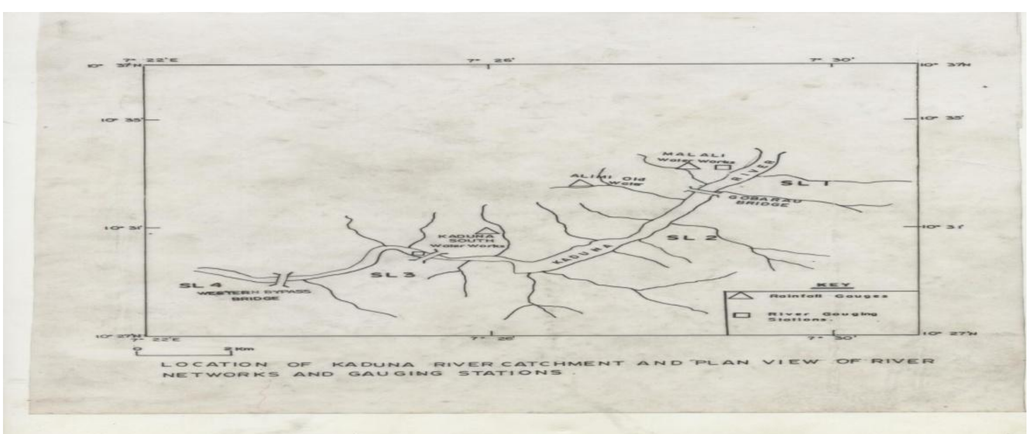


Figure 3 Kaduna South gauging station

III. GOODNESS- OF- FIT- TEST

In[9],it was recognized that many annual floods series ate too short to allow for reliable estimation of extreme events. However, [9] identified that the difficulty are related to both the identification of appropriate statistical distribution for describing the data and to the estimation of parameters of a selected distribution. [9] Therefore hypothesized that regionalization provides a means to cope with this problem. A flood frequency relationship according to [10] for a site with little or no stream flow record can be constructed by using a regionalized flood frequency model. In the method according to [9], the use of a generalized extreme value (GEV) distribution as a regional flood frequency model with an index flood approach has received considerable attention. In order to test whether the available flood data for a site are consistent with a proposed regional GEV distribution for that site, the Goodness – of- Fit Test as observed by [10] can be used to test whether the data at a particular site are consistent with a hypothesized regional distribution. [10] investigated the application of the Kolmongonov- Smirnov test, the probability plote correlation test and sample L-moment tests in testing whether a regional GEV Distribution is consistent with the available data for a site.

3.1 *The Kolmongonov-Smirnov Test*

The Kolmongonov- Smirnov goodness-of – fit – test can be used to check whether the observed flood sample at a site is consistent with a regional GEV flood distribution for the site. The advantage of the K-S test is that it gives simulation confidence interval for all the observations and therefore provides a visual goodness-of – fit- test. The K-S test procedure involves the comparison between the experimental cumulative frequency and the assumed theoretical distribution function. If the discrepancy is large compared to what is normally expected from a given sample size, the theoretical model is rejected. The test statistics for the K-S test involved determining the largest difference (Dmax) between the corresponding cumulative frequencies of the two samples. The maximum difference between the observed and theoretical predicted values (Dmax) is the measure of discrepancy between the theoretical model and observed data. Level of significance of α = 1%, α =5% and α =10% were selected and the critical value $D_{max}^{\propto}$ were computed based on the selected value of α . The K-S test determines whether for a specified level of significance α , the proposed distribution is an acceptable representation of the field data. If Dmax<$D_{max}^{\propto}$ the theoretical distribution is acceptable on the other hand, if Dmax>$D_{max}^{\propto}$ the theoretical distribution is rejected.

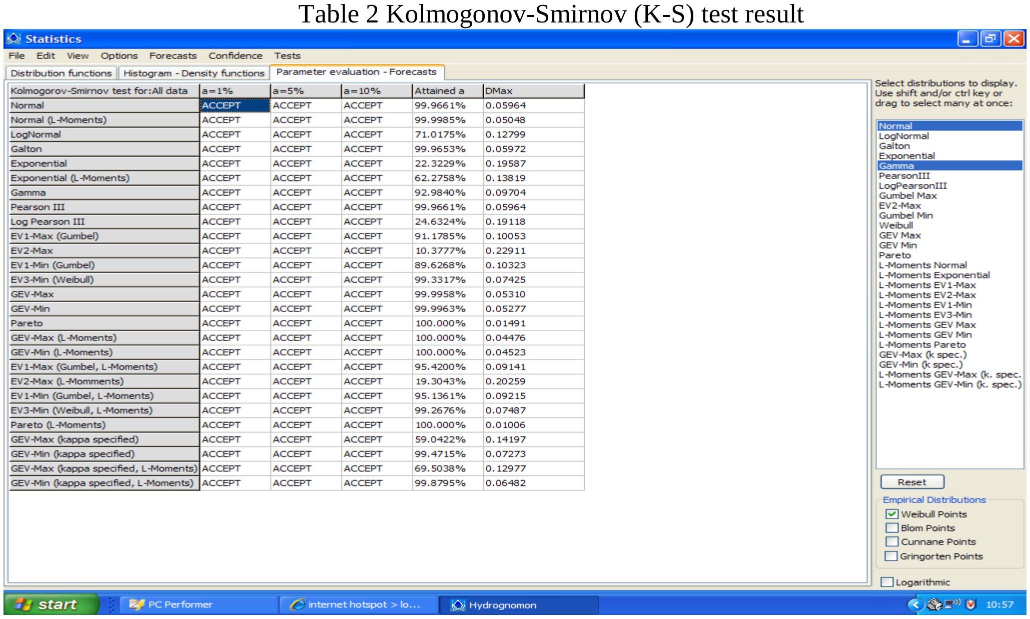
Results and Discussions

Table 1 Simulated mean discharge of Kaduna River

Recurrence Interval (years)	**Exceedence Probability (%)**	**95% Confidence Interval**			
		Normal Q (m^3/s)	**Log Normal Q (m^3/s)**	**Log Pearson Type 111 Q (m^3/s)**	**Gumbel Q(m^3/s)**
2	**50**	**45,000**	**40,000**	**45,000**	**40,000**
5	**20**	**75,000**	**65,000**	**70,000**	**70,000**
10	**10**	**90,000**	**85,000**	**90,000**	**90,000**
20	**5**	**100,00**	**105,000**	**115,000**	**105,000**
50	**2**	**115,000**	**135,000**	**145,000**	**130.000**
100	**1**	**120,000**	**140,000**	**150,000**	**145,000**

TABLE 1 show the annual mean discharge values in different return periods simulated with the four probability distributions functions. Evaluation of the statistical distribution of mean discharges showed that the Normal and Log Normal indicates the lowest fitting between observations and predicted values while the Log Pearson typeIII and the Gumbel showed the highest fitting.

Table 2 Kolmogonov-Smirnov (K-S) test result



Kolmogorov-Smirnov test for:All data	a=1%	a=5%	a=10%	Attained a	DMax
Normal	ACCEPT	ACCEPT	ACCEPT	99.9661%	0.05964
Normal (L-Moments)	ACCEPT	ACCEPT	ACCEPT	99.9985%	0.05048
LogNormal	ACCEPT	ACCEPT	ACCEPT	71.0175%	0.12799
Galton	ACCEPT	ACCEPT	ACCEPT	99.9653%	0.05972
Exponential	ACCEPT	ACCEPT	ACCEPT	22.3229%	0.19587
Exponential (L-Moments)	ACCEPT	ACCEPT	ACCEPT	62.2758%	0.13819
Gamma	ACCEPT	ACCEPT	ACCEPT	92.9840%	0.09704
Pearson III	ACCEPT	ACCEPT	ACCEPT	99.9661%	0.05964
Log Pearson III	ACCEPT	ACCEPT	ACCEPT	24.6324%	0.19118
EV1-Max (Gumbel)	ACCEPT	ACCEPT	ACCEPT	91.1785%	0.10053
EV2-Max	ACCEPT	ACCEPT	ACCEPT	10.3777%	0.22911
EV1-Min (Gumbel)	ACCEPT	ACCEPT	ACCEPT	89.6268%	0.10323
EV3-Min (Weibull)	ACCEPT	ACCEPT	ACCEPT	99.3317%	0.07425
GEV-Max	ACCEPT	ACCEPT	ACCEPT	99.9958%	0.05310
GEV-Min	ACCEPT	ACCEPT	ACCEPT	99.9963%	0.05277
Pareto	ACCEPT	ACCEPT	ACCEPT	100.000%	0.01491
GEV-Max (L-Moments)	ACCEPT	ACCEPT	ACCEPT	100.000%	0.04476
GEV-Min (L-Moments)	ACCEPT	ACCEPT	ACCEPT	100.000%	0.04523
EV1-Max (Gumbel, L-Moments)	ACCEPT	ACCEPT	ACCEPT	95.4200%	0.09141
EV2-Max (L-Momments)	ACCEPT	ACCEPT	ACCEPT	19.3043%	0.20259
EV1-Min (Gumbel, L-Moments)	ACCEPT	ACCEPT	ACCEPT	95.1361%	0.09215
EV3-Min (Weibull, L-Moments)	ACCEPT	ACCEPT	ACCEPT	99.2676%	0.07487
Pareto (L-Moments)	ACCEPT	ACCEPT	ACCEPT	100.000%	0.01006
GEV-Max (kappa specified)	ACCEPT	ACCEPT	ACCEPT	59.0422%	0.14197
GEV-Min (kappa specified)	ACCEPT	ACCEPT	ACCEPT	99.4715%	0.07273
GEV-Max (kappa specified, L-Moments)	ACCEPT	ACCEPT	ACCEPT	69.5038%	0.12977
GEV-Min (kappa specified, L-Moments)	ACCEPT	ACCEPT	ACCEPT	99.8795%	0.06482

TABLE 2 showed result of Kolmongonov- Smirnov test procedure which involves the comparison between observed cumulative frequency and the four theoretical distributions functions. The K-S goodness-of-fit test was used to check whether the annual flood discharge variability at Kaduna South Water Works is consistent with the regional GEV flood distribution for the site. From the measure of discrepancy, it was observed that the difference between observed cumulative frequency values and the theoretical predicted cumulative values at level of significance of α= 1%,α =5% and α=10% with all the theoretical distribution been acceptable.

IV. CONCLUSION

A frequency analysis was performed by fitting probability distribution functions to the discharge variability of Kaduna River at Kaduna South Water Works aimed at determining the occurance frequency or return periods of extreme river floods. Estimate of the return periods of river floods are necessary in a reliability based design of flood protection structures. However, although uncertainties are important in flood analysis reliability based design.Uncertaities could be statistical due to limited amount of flood data or model uncertainty due to limited descriptive capabilities of the physical flooding process.

REFERENCES

[1] Gh Khosravi, A. Magidi and A. Nohegar. Determination of suitable probability distribution for annual and peak discharge estimations (Case Study: Minab River-Barantin Gage, Iran). *International journal of probability and statistics 2012, 1(5): 160-163.*

[2] P.H.A.J.M van Gelder, N.M Neykov, P.I Neytcher, J.K Vrijling and H Chbab. Probability distribution of annual maximum river discharges in North-Western and Central Europe, TU Deft Report 1999.

[3] D Caissie. River discharges and channel width relationship for New Brunswick Rivers. Can. Tech. Rep. Fish. Sci. 2637:26p.

[4] U Looser. River discharged standard report www.fao.org/gtos/ecv-TO/html.

[5] A.W Alayande and J.C Agunwamba. Modeling the impact of urbanization on river flooding using ST. Venant equation. *ARPN Journal of Engineering and Applied Sciences, vol 5, no 7, 2010.*

[6] W Strupczewski. Determination of probability of maximum discharges on the basin of all observed floods. Technical University of Warsaw, Warsaw Poland.

[7] O.C Izinyon and H.N Ajumuka. Probability distribution models for flood prediction in Upper Benue River Basin. *Civil and Environmental Research vol,3,No,2, 2013.*

[8] P.H.A.J.M van Gelder, J.M Noortwirk and M.T Duits. Selection of probability distribution with case study of extreme Oder River discharges. *In G.I Schueller and P Kafta; Safety and reliability proceedings of ESREL. The Tenth European Conference on Safety and Reliability, Munich Garching Germany, 1999, 1475-1480 Rotterdam Balkema.*

[9] H Garba, A Ismail and F.O.P Oriola. Calibration of hydrognomon model for simulating the hydrology of urban catchment. *Open Journal of Modern Hydrology, 3, 75-78, 2013.*

[10] J.U Chowdhung, J.R Stedinger and L.I Hsiung LU. Goodness-of- Fit Test for regional generalized extreme value flood distribution. *Water Resources Research, Vol 27, No 7, Pp 1765-1776.*