New York Science Journal, 2009, 2(3), ISSN 1554-0200
http://www.sciencepub.net/newyork , sciencepub@gmail.com

Best–fit Probability distribution model for peak daily rainfall of selected Cities in Nigeria

Olofintoye, O.O, Sule, B.F and Salami, A.W
Department of Civil Engineering, University of Ilorin, P.M.B 1515, Ilorin, Nigeria
geotoseen@yahoo.co.uk

Abstract

This study aims at studying the peak daily rainfall distribution characteristics in Nigeria, by using different statistical analyses such as Gumbel, Log-Gumbel, Normal, Log-Normal, Pearson and Log-Pearson distributions. 20 stations having annual rainfall data of fifty-four (54) years were selected to perform frequency analysis. Mathematical equation for the probability distribution functions were established for each station and used to predict peak rainfall, the predicted values were subjected to goodness of fit tests such as chi-square, Fisher's test, correlation coefficient and coefficient of determination to determine how best the fits are. The model that satisfies the tests adequately was selected as the best fit model. Results showed that the log-Pearson type III distribution performed the best by occupying 50% of the total station number, while Pearson type III performed second best by occupying 40% of the total stations and lastly by log-Gumbel occupying 10% of the total stations. [New York Science Journal. 2009;2(3):1-12]. (ISSN: 1554-0200).

Key words: Rainfall, peak rainfall, probability distribution models and goodness of fit tests

Introduction

The design and construction of certain projects, such as dams and urban drainage systems, the management of water resources, and the prevention of flood damage require an adequate knowledge of extreme events of high return periods (Tao et al, 2002). In most cases, the return periods of interest exceed usually the periods of available records and could not be extracted directly from the recorded data. Therefore, in current engineering practice, the estimation of extreme rainfalls is accomplished based on statistical frequency analysis of maximum precipitation records where available sample data could be used to calculate the parameters of a selected frequency distribution. The fitted distribution is then used to estimate event magnitudes corresponding to return periods greater than or less than those of the recorded events, hence accurate estimation of extreme rainfall could help to alleviate the damage caused by storms and can help to achieve more efficient design of hydraulic structures.

Tao et al (2002) reported that several probability models have been developed to describe the distribution of annual extreme rainfalls at a single site. However, the choice of a suitable model is still one of the major problems in engineering practice since there is no general agreement as to which distribution, or distributions, that should be used for the frequency analysis of extreme rainfalls. The selection of an appropriate model depends mainly on the characteristics of available rainfall data at the particular site. Hence, it is necessary to evaluate many available distributions in order to find a suitable model that could provide accurate extreme rainfall estimates. This study aimed to determine the best fit probability distribution for the prediction of peak daily rainfall data in selected cities in Nigeria.

Literature review

Phien and Ajirajah (1984) evaluated the applicability of the LP3 distribution to flood and maximum rainfall data, and its general use in fitting annual rainfall and stream flow sequences. The evaluation was carried out by selecting four types of data, which consist of annual flood, annual maximum rainfall of different durations, annual stream flow, and annual rainfall. Three methods of parameter estimation were used, these include; method of maximum likelihood (MML), method using the first two moments of x and first moment of y (MM1) and method using the first two moments of x and second moment of y (MM2). It was pointed out that the estimates of a parameter obtained from the methods MM1 and MM2 are relatively close to each other, while that obtained by maximum likelihood (MML) shows a larger difference. The applicability of the LP3 distribution to the four data types evaluated by the Chi-square (χ^2) and Kolmogorov – Smirnov (Δ) tests was said to be highly suitable. Chapman (1994) tested 5 daily rainfall generating models with several methods of evaluating the model parameters, and reported that the Srikanthan-McMahon (TPM) model performed particularly well when calibrated with long rainfall records.

Boughton (1999) reported that daily rainfall records form a major hydrological data base in Australia, but the common 50 – 100 years of available record at a station do not give adequate information about long term risks of droughts or floods. Transaction probability matrix (TPM) models have been used in prior studies to generate long sequences of daily rainfalls, but the model most commonly used in Australia seriously under estimates the variance of annual totals of rainfall. There has been a steadily increasing interest in the stochastic generation of long sequences of daily rainfalls for periods ranging from 100 to 1,000,000 years in order to give a better definition of extreme droughts and floods. Ogunlela (2001) studied the stochastic analysis of rainfall event in Ilorin using probability distribution functions. He concluded that the log-Pearson type III distribution best described the peak daily rainfall data for Ilorin.

Tao et al (2002) proposes a systematic assessment procedure to compare the performance of different probability distributions in order to identify an appropriate model that could provide the most accurate extreme rainfall estimates at a particular site. Nine probability models such as Beta-K (BEK), Beta-P (BEP), Generalized Extreme Value (GEV), Generalized Normal (GNO), Generalized Pareto (GPA), Gumbel (GUM), Log-Pearson Type III (LP3), Pearson Type III (PE3), and Wakeby (WAK) distributions were compared for their descriptive and predictive abilities to represent the distribution of annual maximum rainfalls. The suggested methodology was applied to 5-minute and 1-hour annual maximum rainfall series from a network of 20 raingages in Southern Quebec region. On the basis of graphical and numerical comparisons, it was found that the WAK, GNO, and GEV models could provide the most accurate extreme rainfall estimates. However, the GEV was recommended as the most suitable distribution due to its theoretical basis for representing extreme – value process and its relatively simple parameter estimation.

Topaloglu (2002) reported that the frequency analysis of the largest, or the smallest, of a sequence of hydrologic events has long been an essential part of the design of hydraulic structures. Therefore, the question of better fit among countless probability models used in frequency analysis is always a fresh one. In his study, he made a statistical comparison of currently popular probability models such as Gumbel, log-logistic, Pearson-3, log-Pearson-3 and log-Normal-3 distributions were applied to the series of annual instantaneous flood peaks and annual peak daily precipitation for 13 flow gauging and 55 precipitation gauging stations in the Seyhan basin, respectively. The parameters of the distributions were estimated by the methods of moments (MOM) and probability weighted moments (PWM). A detailed Chi-square and Kolmogorov-Smirnov (k-s) goodness -of-fit tests were also applied. According to the evaluations of Chi-squared tests, Gumbel (MOM) for both flow and precipitation stations in the Seyhan river basin were found to be the best models. As a result of the k-s test, log-Normal-3 (MOM) and log-Pearson-3 (MOM) models were determined to be the best for flow and precipitation stations, respectively.

Guevara (2003) carried out hydrologic analysis to estimate engineering design parameters of storms in Venezuela, which help hydrologists to improve their environmental designs. The analysis focused on storm advancement coefficient r (SAC) to establish storm Pluviographs; intensity Duration Frequency (IDF). ; and Area – Depth – Duration relationships (ADD). Based on the analysis of 275 storm events, values of r were calculated obtaining a mean value of 0.41 and a standard deviation of 0.075, and being 61, 30, and 9 % storms from the advanced, retarded and symmetric type respectively.

Salami (2004) indicated that Gary and Robert in1971 studied the normal, log-normal, square-root-normal and cube-root-normal frequency distributions of meteorological data for Texas. The results of this research shows that precipitation data conform to the square-root-normal distribution, while evaporation and temperature data conform to all of the frequency distributions tested. The evaporation, temperature and precipitation data were further fitted to the Gumbel extreme-value and log-Pearson type III distributions. The precipitation data fit the log-Pearson type III (LP3) distribution more adequately than the Gumbel distribution, while both the evaporation and temperature data conform very well to Gumbel distribution.

Lee (2005) studies the rainfall distribution characteristics of Chia-Nan plain area, by using different statistical analyses such as normal distribution, log-normal distribution, extreme value type I distribution, Pearson type III distribution, and log-Pearson type III distribution. Results showed that the log-Pearson type III distribution performed the best in probability distribution, occupying 50% of the total station number, followed by the log-Normal distribution and Pearson type III distribution, which accounts for 19% and 18% of the total station numbers respectively.

Bhakar et al (2006) studied the frequency analysis of consecutive day's maximum rainfall at Banswara, Rajasthan, India. Various probability distributions and transformations were applied to estimate one day and two to five consecutive days annual maximum rainfall of various return periods. Three commonly used probability distributions, Normal, Log Normal and Gamma distribution were tested by comparing the Chi-square value. Gamma distribution was found to be best fit for the region. The magnitudes of 1 day as well as 2 to 5 consecutive days annual maximum rainfall corresponding to 2 to 100 years return period were estimated using Gamma function. A maximum of 154.31 mm in 1 day, 250.88 mm in 2 days, 270.15 mm in 3 days, 284.18 mm in 4 days and 295.54 mm in 5 days is expected to occur at Udaipur, Rajasthan every 2 years. For a recurrence interval of 100 years, the maximum rainfall expected in 1 day, 2, 3, 4 and 5 days is 773.6 mm, 849.34 mm, 874.19 mm, 931.78 mm and 957.89 mm, respectively. Annual one day maximum rainfall and two to five days consecutive days maximum rainfall corresponding to return period varying from 2 to 100 years are used by design engineers and hydrologists for the economic planning, design of small and medium hydrologic structures and determination of drainage coefficient for agricultural fields.

Data and Analysis

The daily rainfall data for the selected cities were obtained from the Nigerian Meteorological Agency (NIMET), Oshodi Lagos, Nigeria. NIMET is the agency responsible for the measurement, control, and storage of rainfall data of the areas in Nigeria. The nature of data is rainfall depth (mm) recorded for everyday of the year. The rainfall data spanned between 1952 and 2005 and the peak daily values were extracted for the purpose of analysis. Twenty cities were selected for this study and include Bauchi, Gusau, Ilorin, Jos, Kaduna, Kano, Maiduguri, Makurdi, Minna, Potiskum, Sokoto, Yola, Benin City, Calabar, Enugu, Ibadan, Ikeja, Ondo, Owerri, and P. Harcourt. The summary of statistics for peak daily rainfall is presented in table 1.

The data are ranked according to Welbull's plotting position and the corresponding return period was estimated. The ranked data were evaluated with six methods of probability distribution functions to determine the best – fit functions. The methods include; Gumbel (EVI type1), Log-Gumbel (LG), Normal (N), Log-Normal (LN), Log-Pearson type III (LP_3) and Pearson type III (P) probability distribution models. Four statistical goodness of fit test were used for the selection of the best fit models.

Table 1 Summary of statistics for peak daily rainfall (1952 – 2005)

Selected Cities	Parameters					
	Mean value, $\bar{x}$ (mm)	Standard deviation,σ (mm)	Skewness coefficient (G)	Coefficient of variation (Cv)	Maximum (mm)	Minimum (mm)
Bauchi	73.42	11.71	1.43	0.16	108.60	61.70
Gusau	74.56	15.79	1.74	0.21	136.30	57.60
Ilorin	91.10	19.08	1.44	0.21	160.50	70.90
Jos	72.12	11.28	1.50	0.16	107.19	60.71
Kaduna	76.81	15.65	1.47	0.20	132.10	59.60
Kano	91.49	30.94	3.46	0.34	253.10	70.70
Maiduguri	74.46	17.91	1.30	0.24	128.00	56.40
Makurdi	101.97	24.68	2.32	0.24	206.25	80.70
Minna	88.92	18.12	1.15	0.20	144.00	69.30
Potiskum	73.35	16.35	1.85	0.22	129.30	57.80
Sokoto	72.87	16.00	1.96	0.22	139.40	56.10
Yola	78.17	16.16	1.45	0.21	126.24	61.72
Benin City	118.15	20.91	1.21	0.18	181.40	96.00
Calabar	139.65	24.74	1.14	0.18	216.00	112.20
Enugu	106.93	21.18	1.41	0.20	174.40	83.80
Ibadan	117.16	28.95	1.32	0.25	225.90	86.90
Ikeja	121.73	27.93	1.91	0.23	237.30	95.50
Ondo	100.24	29.93	2.73	0.30	246.30	73.90
Owerri	113.12	22.35	1.61	0.20	181.90	90.50
P. Harcourt	111.01	23.64	1.71	0.21	185.30	88.40

Probability distribution analysis was carried out in accordance with standard procedure [Warren et al, (1972); Viessman et al, (1989); Mustapha and Yusuf, (1999), and Topaloglu (2002)]. The mathematical expressions obtained for various probability distributions functions are presented in Table 2. The mathematical expressions obtained for each function were used to predict the peak rainfall data based on the estimated returned periods and were also used in performing the statistical tests (goodness of fit tests) for the selection of the best fit models.

Table 2 Mathematical expression for probability distributions models

Station	Distributions					
	Normal	Log-Normal	Pearson III	Log-Pearson III	Gumbel	Log-Gumbel
Bauchi	$R_p = 73.42 + 11.71K$	$\text{Log } R_p = 1.86 + 0.06K$	$R_p = 73.42 + 11.71K'$	$\text{Log } R_p = 1.86 + 0.06K'$	$R_p = 68.16 + 9.13Y_T$	$\text{Log } R_p = 1.83 + 0.05Y_T$
Gusau	$R_p = 74.56 + 15.79K$	$\text{Log } R_p = 1.86 + 0.08K$	$R_p = 74.56 + 15.79K'$	$\text{Log } R_p = 1.86 + 0.08K'$	$R_p = 67.46 + 12.31Y_T$	$\text{Log } R_p = 1.83 + 0.06Y_T$
Ilorin	$R_p = 91.10 + 19.08K$	$\text{Log } R_p = 1.95 + 0.08K$	$R_p = 91.10 + 19.08K'$	$\text{Log } R_p = 1.95 + 0.08K'$	$R_p = 82.51 + 14.88Y_T$	$\text{Log } R_p = 1.91 + 0.06Y_T$
Jos	$R_p = 72.12 + 11.28K$	$\text{Log } R_p = 1.85 + 0.06K$	$R_p = 72.12 + 11.28K'$	$\text{Log } R_p = 1.85 + 0.06K'$	$R_p = 67.04 + 8.80Y_T$	$\text{Log } R_p = 1.83 + 0.05Y_T$
Kaduna	$R_p = 76.81 + 15.65K$	$\text{Log } R_p = 1.88 + 0.08K$	$R_p = 76.81 + 15.65K'$	$\text{Log } R_p = 1.88 + 0.08K'$	$R_p = 69.76 + 12.21Y_T$	$\text{Log } R_p = 1.84 + 0.06Y_T$
Kano	$R_p = 91.49 + 30.94K$	$\text{Log } R_p = 1.95 + 0.11K$	$R_p = 91.49 + 30.94K'$	$\text{Log } R_p = 1.95 + 0.11K'$	$R_p = 77.56 + 24.13Y_T$	$\text{Log } R_p = 1.90 + 0.08Y_T$
Maiduguri	$R_p = 74.46 + 17.91K$	$\text{Log } R_p = 1.86 + 0.10K$	$R_p = 74.46 + 17.91K'$	$\text{Log } R_p = 1.86 + 0.10K'$	$R_p = 66.40 + 13.97Y_T$	$\text{Log } R_p = 1.82 + 0.07Y_T$
Makurdi	$R_p = 101.97 + 24.68K$	$\text{Log } R_p = 2.00 + 0.09K$	$R_p = 101.97 + 24.68K'$	$\text{Log } R_p = 2.00 + 0.09K'$	$R_p = 90.86 + 19.25Y_T$	$\text{Log } R_p = 1.96 + 0.07Y_T$
Minna	$R_p = 88.92 + 18.12K$	$\text{Log } R_p = 1.94 + 0.08K$	$R_p = 88.92 + 18.12K'$	$\text{Log } R_p = 1.94 + 0.08K'$	$R_p = 80.77 + 14.13Y_T$	$\text{Log } R_p = 1.90 + 0.06Y_T$
Potiskum	$R_p = 78.17 + 16.16K$	$\text{Log } R_p = 1.89 + 0.08K$	$R_p = 78.17 + 16.16K'$	$\text{Log } R_p = 1.89 + 0.08K'$	$R_p = 70.90 + 12.61Y_T$	$\text{Log } R_p = 1.85 + 0.06Y_T$
Sokoto	$R_p = 72.87 + 16.00K$	$\text{Log } R_p = 1.85 + 0.08K$	$R_p = 72.87 + 16.00K'$	$\text{Log } R_p = 1.85 + 0.08K'$	$R_p = 65.67 + 12.48Y_T$	$\text{Log } R_p = 1.82 + 0.07Y_T$
Yola	$R_p = 78.17 + 16.16K$	$\text{Log } R_p = 1.89 + 0.08K$	$R_p = 78.17 + 16.16K'$	$\text{Log } R_p = 1.89 + 0.08K'$	$R_p = 70.90 + 12.61Y_T$	$\text{Log } R_p = 1.85 + 0.06Y_T$
Benin-City	$R_p = 118.15 + 20.91K$	$\text{Log } R_p = 2.07 + 0.07K$	$R_p = 118.15 + 20.91K'$	$\text{Log } R_p = 2.07 + 0.07K'$	$R_p = 108.74 + 16.31Y_T$	$\text{Log } R_p = 2.03 + 0.06Y$
Calabar	$R_p = 139.65 + 24.72K$	$\text{Log } R_p = 2.14 + 0.07K$	$R_p = 139.65 + 24.72K'$	$\text{Log } R_p = 2.14 + 0.07K'$	$R_p = 128.52 + 19.29Y_T$	$\text{Log } R_p = 2.11 + 0.06Y$
Enugu	$R_p = 106.93 + 21.18K$	$\text{Log}R_p = 2.02 + 0.08K$	$R_p = 106.93 + 21.18K'$	$\text{Log}R_p = 2.02 + 0.08K'$	$R_p = 97.40 + 16.52Y_T$	$\text{Log}R_p = 1.99 + 0.06Y_T$
Ibadan	$R_p = 117.16 + 28.95K$	$\text{Log } R_p = 2.06 + 0.10K$	$R_p = 117.16 + 28.95K'$	$\text{Log } R_p = 2.06 + 0.10K'$	$R_p = 104.13 + 22.58Y_T$	$\text{Log } R_p = 2.01 + 0.08Y$
Ikeja	$R_p = 121.73 + 27.93K$	$\text{Log } R_p = 2.08 + 0.09K$	$R_p = 121.73 + 27.93K'$	$\text{Log } R_p = 2.08 + 0.09K'$	$R_p = 109.17 + 21.78Y_T$	$\text{Log } R_p = 2.04 + 0.07Y$
Ondo	$Rp = 100.24 + 29.93K$	$\text{Log } Rp = 1.99 + 0.10K$	$Rp = 100.24 + 29.93K'$	$\text{Log } Rp = 1.99 + 0.10K'$	$Rp = 86.78 + 23.35Y_T$	$\text{Log } Rp = 1.94 + 0.08Y$
Owerri	$R_p = 113.12 + 22.35K$	$\text{Log } R_p = 2.05 + 0.08K$	$R_p = 113.12 + 22.35K'$	$\text{Log } R_p = 2.05 + 0.08K'$	$R_p = 103.06 + 17.43Y_T$	$\text{Log } R_p = 2.01 + 0.06Y$
PortHarcourt	$R_p = 111.01 + 23.64K$	$\text{Log } R_p = 2.04 + 0.08K$	$R_p = 111.01 + 23.64K'$	$\text{Log } R_p = 2.04 + 0.08K'$	$R_p = 100.37 + 18.44Y_T$	$\text{Log } R_p = 2.00 + 0.06Y$

In order to determine the best-fit model(s) at each station, probability distribution models were subjected to four (4) statistical tests (goodness of fit tests). The statistical tests include chi-square (χ^2), Fisher's (F) test, probability plot coefficient of correlation (r) and coefficient of determination (R^2). The statistical tests were carried out in accordance with standard procedure (Chowdhury and Stedinger (1991); Adegboye and Ipinyomi (1995); Dibike and Solomatine (1999); Murray and Larry (2000)). The assessment of the probability distribution models was based on the total test score obtained from all the tests. Test scores ranging from zero to six (0-6) is awarded to each distribution model based on the criteria that the distribution (s) with the highest total score is or are chosen as the best distribution model(s) for the data of a particular city. In general, the distribution best supported by a test is awarded a score of six (6), the next best is awarded five (5), and so on in descending order. A distribution is awarded a zero (0) score for a test if the test indicates that there is a significant difference between the rainfall values estimated by the distribution model and the observed rainfall data. For every test category, overall ranks of each distribution were obtained by summing the individual point rank at each of the 20 stations. The overall ranking results are presented in Table 3 while the best fit probability distribution models are presented in Table 4.

Table 3 Summary of the statistical test score results at each station

Stations	Best fit distribution model	Chi-square test	Fisher's test	PPCC (r)	R^2	Total
Bauchi	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	5	6	5	5	21
	Log-Pearson	6	5	6	6	23
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	4	4	16
Gusau	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	6	6	4	6	22
	Log-Pearson	5	5	5	5	20
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	6	4	18
Ilorin	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	6	6	4	5	21
	Log-Pearson	5	5	5	6	21
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	6	4	18
Jos	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	5	6	5	5	21
	Log-Pearson	6	5	6	6	23
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	4	4	16
Kaduna	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	5	5	4	4	18
	Log-Pearson	4	4	5	5	18
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	6	6	6	6	24
Kano	Normal	0	2	1	1	4
	Log-Normal	0	1	2	2	5
	Pearson	6	6	5	6	23
	Log-Pearson	5	5	6	5	21
	Gumbel (EVI)	0	4	3	3	10
	Log-Gumbel	0	3	4	4	11
Maiduguri	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	4	6	4	5	19

Log-Pearson	6	5	6	6	23
Gumbel (EVI)	3	3	3	3	12
Log-Gumbel	5	4	5	4	18

Table 3 continue

Stations	Best fit distribution model	Chi-square test	Fisher's test	PPCC (r)	R^2	Total
Makurdi	Normal	0	1	1	1	3
	Log-Normal	2	2	2	2	8
	Pearson	6	6	5	6	23
	Log-Pearson	5	5	6	5	21
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	4	4	16
Minna	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	4	5	4	4	17
	Log-Pearson	6	6	6	6	24
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	5	4	5	5	19
Potiskum	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	6	6	5	6	23
	Log-Pearson	5	5	6	5	21
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	4	4	16
Sokoto	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	5	6	4	5	20
	Log-Pearson	6	5	6	6	23
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	5	4	17
Yola	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	5	6	4	5	20
	Log-Pearson	6	5	6	6	23
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	5	4	17
Benin City	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	4	5	4	5	18
	Log-Pearson	6	6	6	6	24
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	5	4	5	4	18
Calabar	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	4	5	4	5	18
	Log-Pearson	6	6	6	6	24
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	5	4	5	4	18

Table 3 continue

Stations	Best fit distribution model	Chi-square test	Fisher's test	PPCC (r)	R^2	Total
Enugu	Normal	1	1	1	1	4

	Log-Normal	2	2	2	2	8
	Pearson	6	6	4	5	21
	Log-Pearson	5	5	5	6	21
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	6	4	18
Ibadan	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	6	5	4	5	20
	Log-Pearson	5	4	5	4	18
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	6	6	6	22
Ikeja	Normal	0	1	1	1	3
	Log-Normal	2	2	2	2	8
	Pearson	6	6	5	6	23
	Log-Pearson	5	5	6	5	21
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	4	4	16
Ondo	Normal	0	2	1	1	4
	Log-Normal	0	1	2	2	5
	Pearson	5	6	5	6	22
	Log-Pearson	6	5	6	5	22
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	4	4	16
Owerri	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	6	6	4	6	22
	Log-Pearson	5	5	6	5	21
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	5	4	17
P. Harcourt	Normal	1	1	1	1	4
	Log-Normal	2	2	2	2	8
	Pearson	6	6	5	6	23
	Log-Pearson	5	5	6	5	21
	Gumbel (EVI)	3	3	3	3	12
	Log-Gumbel	4	4	4	4	16

Table 4 Goodness of fit tests and the selected model for the peak rainfall

	Best-Fit Model	Second Best-Fit Model	Best-Fit Model Equation
Bauchi	Log-Pearson III	Pearson III	$\text{Log } R_p = 1.86 + 0.06K'$
Gusau	Pearson III	Log-Pearson III	$R_p = 74.56 + 15.79K'$
Ilorin	Log-Pearson III	Pearson III	$\text{Log } R_p = 1.95 + 0.08K'$
Jos	Log-Pearson III	Pearson III	$\text{Log } R_p = 1.85 + 0.06K'$
Kaduna	Log-Gumbel	Pearson III	$\text{Log } R_p = 1.84 + 0.06Y_T$
Kano	Pearson III	Log-Pearson III	$R_p = 91.49 + 30.94K'$
Maiduguri	Log-Pearson III	Pearson III	$\text{Log } R_p = 1.86 + 0.10K'$
Makurdi	Pearson III	Log-Pearson III	$R_p = 101.97 + 24.68K'$
Minna	Log-Pearson III	Log-Gumbel	$\text{Log } R_p = 1.94 + 0.08K'$

Potiskum	Pearson III	Log-Pearson III	$R_p = 73.35 + 16.35K'$
Sokoto	Log-Pearson III	Pearson III	$\text{Log } R_p = 1.85 + 0.08K'$
Yola	Log-Pearson III	Pearson III	$\text{Log } R_p = 1.89 + 0.08K'$
Benin-City	Log-Pearson III	Log-Gumbel	$\text{Log } R_p = 2.07 + 0.07K'$
Calabar	Log-Pearson III	Log-Gumbel	$\text{Log } R_p = 2.14 + 0.07K'$
Enugu	Log-Pearson III	Pearson III	$\text{Log } R_p = 2.02 + 0.08K'$
Ibadan	Log-Gumbel	Pearson III	$\text{Log } R_p = 2.01 + 0.08Y_T$
Ikeja	Pearson III	Log-Pearson III	$R_p = 121.73 + 27.93K'$
Ondo	Pearson III	Log-Pearson III	$R_p = 100.24 + 29.93K'$
Owerri	Pearson III	Log-Pearson III	$R_p = 113.12 + 22.35K'$
Port-Harcourt	Pearson III	Log-Pearson III	$R_p = 111.01 + 23.64K'$

Application of results

The results obtained from the analysis may be useful for engineering planning and designs in that future year maximum daily rainfall events can be predicted. However the selected best distribution model were used to predict maximum daily rainfall depths for the twenty stations for return periods of 5, 10, 20, 50, 100, 200, and 500 years. The quantile estimates are presented in table 5.

Table 5 Quantile estimates for various return periods

		Reccurence interval in years						
Station	**Best-Fit Model**	**5**	**10**	**20**	**50**	**100**	**200**	**500**
Bauchi	Log-Pearson III	80.90	88.32	95.90	106.41	114.84	123.78	136.46
Gusau	Pearson III	84.74	95.15	105.45	119.07	129.42	139.84	153.74
Ilorin	Log-Pearson III	103.24	115.30	127.68	144.99	159.02	174.00	195.49
Jos	Log-Pearson III	79.19	86.39	93.78	104.08	112.39	121.22	133.81
Kaduna	Log-Gumbel	86.25	96.16	106.74	122.18	135.19	149.54	170.82
Kano	Pearson III	101.03	123.39	149.10	187.17	218.60	252.01	298.92
Maiduguri	Log-Pearson III	85.92	97.27	108.94	125.30	138.61	152.87	173.38
Makurdi	Pearson III	115.27	132.57	150.58	175.37	194.78	214.73	241.86
Minna	Log-Pearson III	101.06	112.34	123.68	139.21	151.56	164.55	182.86
Potiskum	Pearson III	83.59	94.52	105.45	120.01	131.14	142.39	157.45
Sokoto	Log-Pearson III	82.29	92.48	103.22	118.63	131.42	145.35	165.77
Yola	Log-Pearson III	88.26	98.57	109.24	124.29	136.57	149.76	168.81
Benin	Log-Pearson III	132.13	145.13	158.17	175.96	190.05	204.82	225.56
Calabar	Log-Pearson III	156.53	171.74	186.85	207.25	223.27	239.95	263.19
Enugu	Log-Pearson III	120.55	133.88	147.50	166.41	181.66	197.86	220.97
Ibadan	Log-Gumbel	134.31	153.36	174.17	205.36	232.33	262.73	309.01
Ikeja	Pearson III	138.92	157.72	176.62	201.90	221.30	240.94	267.30
Ondo	Pearson III	113.99	135.47	158.62	191.34	217.46	244.64	282.05

Owerri	Pearson III	128.02	142.48	156.63	175.15	189.13	203.12	221.70
PortHarcourt	Pearson III	126.36	141.89	157.24	177.48	192.84	208.30	228.89

Development of Isohyet maps of 100 and 200 year return period

Analysis of rainfall data requires handling of large volumes of data and repeated computation of a number of statistical parameters for distribution fitting and estimation of expected rainfall at different return periods. The use of rainfall frequency atlases may considerably reduce the computational tedium involved in the frequency analysis of rainfall to a greater extent. Maps have been prepared for some regions, particularly East Africa, showing the annual rainfall likely to be equaled or exceeded in 80% of years. These are extremely useful for planning purposes (Edwards et al, 1983). In this study isohyet maps were constructed using software (Surfer 8). Surfer is a contouring and 3D surface mapping program that runs under Microsoft Windows. Surfer version 8 is a product of Golden Software, Inc (www.goldensoftware.com). The quantile estimates for 100 and 200 years return periods presented in table 5 were used to construct isohyetal maps for the maximum daily rainfall for selected cities in Nigeria. The isohyetal map for 100 and 200 years return period are presented in Figure 1 and 2 respectively. With the established maps, the rainfall depth for 100 and 200 years return period for any location (longitude and latitude) in Nigeria may be estimated more easily and faster without having to go through the rigor of fitting probability distribution models all over again. These are very useful for design and planning purposes.

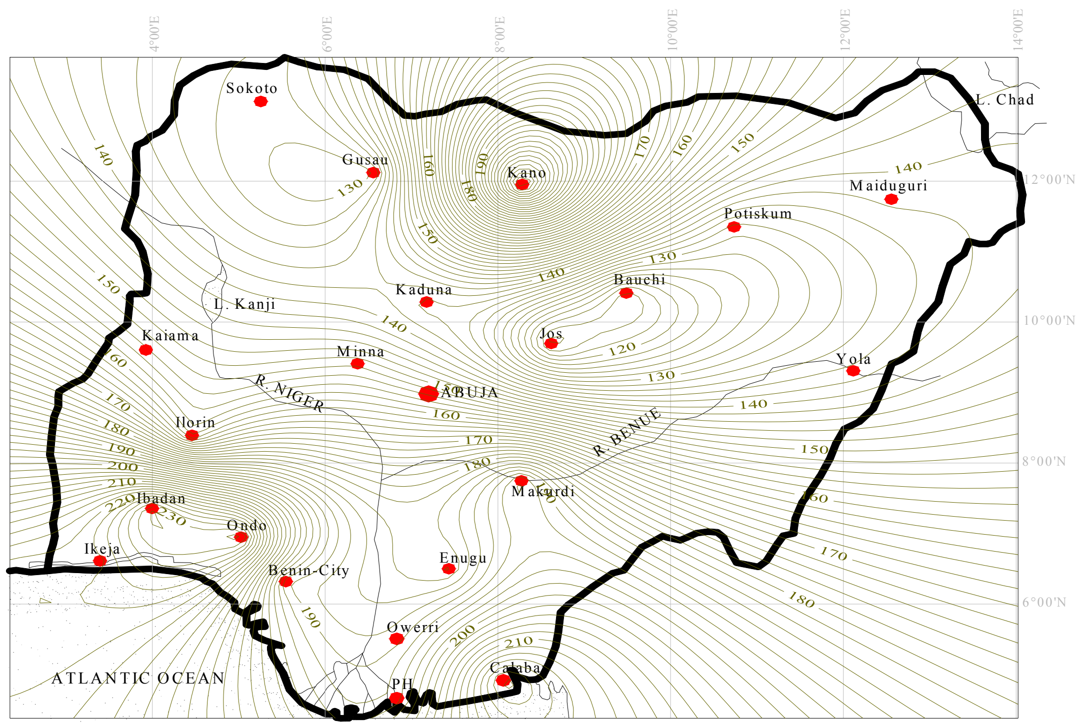

Figure 1 Peak daily rainfall (mm) Isohyet Map of 100 year return period

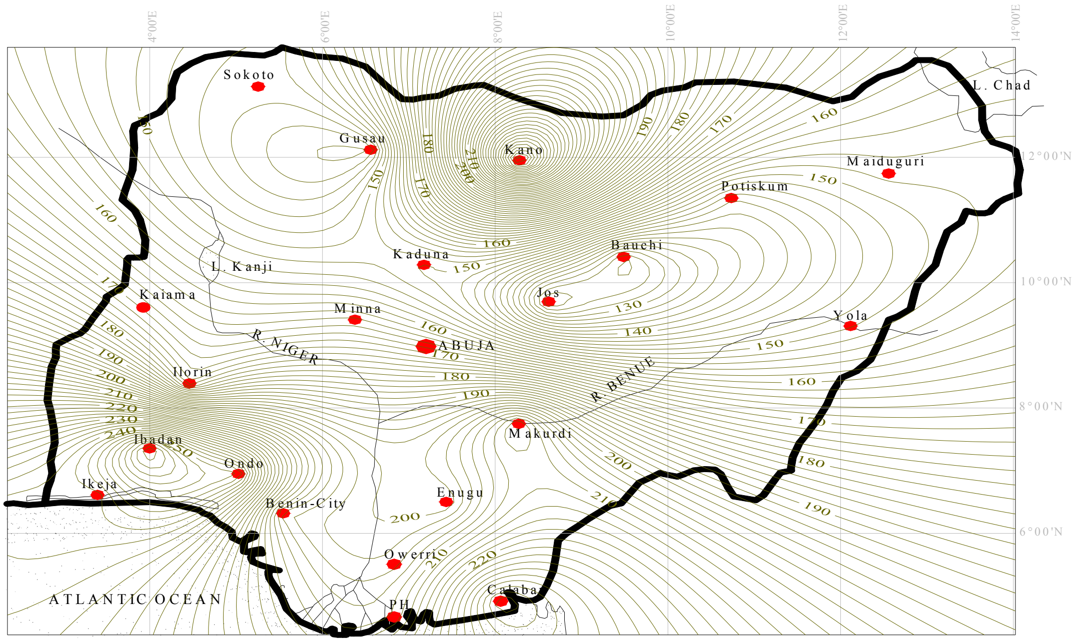

Figure 2 Peak daily rainfall (mm) Isohyet map of 200 year return period

Results and Discussion

The rainfall data was obtained for a period of fifty four years (1952 - 2005) for the selected Cities in Nigeria and the peak daily values were selected. The data were evaluated with various probability distribution functions to determine the best fitting model, the summary statistic is presented in Table 1 and the mathematical representations of the evaluated probability functions are presented in Table 2. For the purpose of theoretical determination of best fit probability function, statistical tools (goodness of fit tests) were adopted. The results of the statistical test score (the goodness of fit tests) and the best fit models are presented in Table 3 and 4 respectively. However, the quantile estimates for the selected Cities based on various return periods are presented in Table 5. The isohyet maps of 100 and 200 year return period were established and presented in Figure 1 and 2 respectively for the estimation of rainfall depth for any location with known longitude and latitude in Nigeria.

The statistical test score results at each station presented in Table 3 was used to decide which of the probability model (s) best fit the peak daily rainfall at each station. Examination of the goodness-of-fit test results reveals that in many cases there was very little difference between the various distributions for each station. Furthermore, the good fit assessment for all 20 stations also indicated that no one distribution ranked consistently best at all locations. However, the overall ranks for the 10 stations combined show that Log-Pearson type III was best to describe the peak rainfall. While, the overall ranks for the 8 stations combined show that Pearson type III was best to describe the peak rainfall. The overall ranks for 2 stations indicate that Log-Gumbel distribution is the best. The computed skew coefficients from the observed data in this study (Table 1) revealed that the peak daily rainfall distributions at all the stations were positively skewed. Also, the 3-parameter Pearsonian distributions, which take into cognizance the use of the skew coefficient in the estimation of future rainfalls, were found to be the best-fit models at most stations. In view of this, it can be said that the peak daily rainfall distribution in Nigeria are positively skewed and the Pearson III and Log-Pearson III distributions may be conveniently used for the prediction of future peak daily rainfall events anywhere in Nigeria.

Conclusions

From the results of six frequency distributions applied in this study, it suggests that the best frequency distribution obtained for the peak daily rainfall in Nigeria is the log-Pearson type III distribution, which occupies 50% of the total station number, followed by the Pearson type III distribution and log-Gumbel distribution, which accounts for 40% and 10% of the total station number, respectively. The outcome was relied on the results of four goodness-of-fit tests, Chi-square, Fisher's test, correlation coefficient and coefficient of determination. The proposed assessment procedure has been successfully used to identify the best probability distributions that could provide accurate peak daily rainfall estimates for Nigeria. The results of the frequency analysis suggest log-Pearson type III and Pearson type III distributions has the primary distribution pattern for this study site and should be used as a universal distribution model for the prediction of peak daily rainfalls in Nigeria.

References

1. Adegboye, O.S and Ipinyomi, R.A (1995) "Statistical tables for class work and Examination. Tertiary publications Nigeria Limited, Ilorin, Nigeria, pp. 5 - 11

2. Bhakar, S.R., Bansal, A.K, Chhajed, N, and Purohit, R.C (2006). Frequency analysis of consecutive days maximum rainfall at Banswara, Rajasthan, India. ARPN Journal of Engineering and Applied Sciences.

3. Boughton, W.C (1999). A daily rainfall generating model for water yield and flood studies. Cooperative Research Centre for Catchment Hydrology, Australia.

4. Chapman, T (1994). Stochastic models for daily rainfall. The institution of Engineers, Australia, National Conference Publication 94 (15), 7-12.

5. Chowdhury, J.U. and Stedinger, J.R. (1991) "Goodness of fit tests for regional generalized extreme value flood distributions." Water Resource. Res., 27(7) : 1765 – 1776.

6. Dibike, B.Y and Solomatine, D.P (1999) "River flow forecasting using Artificial Neural Networks". Paper Presented at European Geophysical Society (EGS) XXIV General Assembly. The Hague, The Netherlands, 1 – 11.

7. Edwards, K.A, Classen, G.A, and Schroten, E.H.J (1983). The water resource in tropical Africa and its exploitation. International Livestock Centre for Africa (ILCA) Research Report No. 6. Chapter 2: Climate and rainfall in the tropics.

8. Guevara, E (2003). Engineering Design Parameters of storms in Venezuela. Hydrology Days, 80-91.

9. Lee, C (2005). Application of rainfall frequency analysis on studying rainfall distribution characteristics of Chia-Nan plain area in Southern Taiwan. Journal of Crop, Environment & Bioinformatics (2). 31-38. 189 Chung-Cheng Rd, Wufeng.

10. Mustapha, S. and Yusuf, M.I. (1999) "A Textbook of Hydrology and Water Resources", Jenas Prints & Publishing Co., Abuja, Nigeria, pp. 164 – 184, First edition.

11. Murray, R.S and Larry, J.S (2000) "Theory and problems of statistics" Tata Mc Graw – Hill Publishing Company Limited, New Delhi, pp. 314 – 316, Third edition.

12. Ogunlela, A.O (2001) " Stochastic Analysis of Rainfall Events in Ilorin, Nigeria" Journal of Agricultural Research and Development. Volume 1: 39 -50.

13. Phien, H.N and Ajirajah, T, J (1984). Applications of the log-Pearson Type-3 distribution in hydrology. Journal of hydrology, 73; 359-372.

14. Salami, A.W (2004). Prediction of the annual flow regime along Asa River using probability distribution models. AMSE periodical, Lyon, France. Modelling C- 2004, 65 (2), 41-56. (http://www.amse-modeling.org/content_amse2004.htm)

15. Tao, D.Q, Nguyen, V.T and Bourque, A (2002). On selection of probability distributions for representing extreme precipitations in Southern Quebec. Annual Conference of the Canadian Society for Civil Engineering. 5th -8th June 2002. pp 1-8.

16. Topaloglu, F. (2002) "Determining Suitable Probability Distribution Models for Flow and Precipitation Series of the Seyhan River Basin". Turk Journal of Agric.26: 189 – 194.

17. Viessman, .W, Krapp, J.W and Harbough, T. E (1989) "Introduction to Hydrology". Harper and Row Publishers Inc., New York. pp. 675 – 695, Third edition.

18. Warren, .V, Terence, E.H and. John, W.K (1972) "Introduction to hydrology" Educational Publishers, New York, pp. 106 – 141, Second edition.

1/24/2009