International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2643-640X
Vol. 8 Issue 5 May - 2024, Pages: 24-34

# Evaluation of Best –Fit Probability Distribution Models for the Prediction of Rainfall Intensity Duration in Umuahia, Nigeria

**Arimieari, L. W.[1*] and Omunakwe, C.[2]**

[1, 2,] Department of Civil Engineering, Faculty of Engineering, Rivers State University, Port Harcourt
* Correspondence author: love.arimieari@ust.edu.ng

***Abstract:*** *In this research, the Nigeria Hydrological Services Agency (NIHSA), in Abuja provided data on the amount, duration, and frequency of rainfall in Umuahia, Nigeria. The data was sorted for frequency analysis using two distinct techniques: the Gumbel and Normal distribution approaches. The Talbot Rainfall Intensity Duration Frequency (IDF) Equation, an empirical relationship between maximum rainfall intensity (as dependent variable) and other parameters of interest, such as rainfall duration and frequency (as independent variables), was used to develop the desired curves for the study area. According to the results, the Gumbel Distribution model best fits the data collection. Since these values exceed those of the Normal Distribution, it may be said that the Gumbel is a more reliable measure for the area of study. IDF curves for rainfall intensity duration frequencies were created using the aforementioned procedures. The duration of the curves varied from 10, 15, 30, 60, 120, and 240 minutes for each return period, which was 5, 10, 15, 20, 25, and 30 years. The essential storm duration (a concentration of times for determining peak discharge) for the catchment area, the chosen frequency (return period), and the intensity of rainfall are required variables for the hydrological design of hydraulic structures.*

**Keywords: Frequency Distribution, Gumbel, Normal, Goodness of fit test, Rainfall Intensity Duration**

## 1. INTRODUCTION

On occasion, a large amount of water builds up to the point where it overwhelms river banks and submerges a substantial area of land, creating a potentially dangerous situation. It's a natural calamity. It's also called a rainstorm at times. It is horrible and depressing when there is a large loss of life and property. In our country, excessive and persistently high rainfall is the main cause of flooding. After several days with heavy rain, the rivers are unable to hold and release as much water into the sea. When water levels rise above river banks, a significant portion of land is flooded. An infrequent large-scale water release from the mountains can result in a dangerous situation. Rainfall is one of the main elements of the hydrological cycle, which is closely related to human activities, especially in the agricultural sector [1].

Intensity Duration Frequency (series of events) refers to the empirical correlations between the intensity (amount), duration (time), and return period of rainfall. The abscissa (distance from a point to the vertical or y-axis, measured parallel to the horizontal or x-axis; the x-coordinate), ordinates (straight line drawn parallel to one coordinate), and return period or frequency are the three parameter curves that make up the IDF curve [2].

The absence of adequate or up-to-date national standards and procedures on urban drainage design across the nation is a major cause for concern. These procedures and standards determine the relationship between the intensity, duration, and frequency of rainfall at any given location through the analysis of rainfall records obtained there, which can then be utilized in a variety of civil and hydraulic engineering designs and planning. When considering how rainfall affects the hydrological cycle, a rough computation or evaluation needs to be done.

Design engineers have access to rainfall frequencies and return periods for any given time frame, from one minute to eighty-four hours, according to the UK flood studies study published by NERC in 1975 [3].

When building hydraulic structures and conducting economic planning, it can be helpful to look at the durations of periods during which rainfall is the highest [4]. Flood frequency analysis is a crucial and frequently researched topic in hydrology and water resources, especially for hydraulic design and the reduction and management of flood hazards [5]. The data obtained by (IDF) curves may be useful to the hydrologists and engineers working on the planning and design projects involving water resources. In this research, the IDF curve that best fit the set data was generated using the Gumbel distribution technique [6]. The development of the IDF curve that best fit the collection of data in this research required the assessment of flood risk in order to construct flood relief and protection works and to evaluate the safety of both proposed and existing infrastructure [7]. In Jordan's Mujib Basin, the relationship between rainfall, duration, and frequency was investigated. IDF equations were developed for each of the eight rainfall monitoring stations in the basin, and the results showed that the predicted values were typically consistent with the observed values when compared to the curves generated by the Gumbel method and the Water Authority of Jordan (WAJ) [8]. Investigating the best distribution model for each site is essential [9]. It will need additional investigation to ascertain which of the selection criteria works best, especially when dealing with heavy-tailed distributions and small sample sizes, which are common in flood frequency analysis. Additionally, the model selection criteria based on AD, BIC, and AIC were implemented [10].

## 2. MATERIALS AND METHODS

### Materials

### Data Collection

The rainfall data was provided by the Nigeria Hydrological Services Agency (NIHSA), which is situated in Abuja, Nigeria. The NIHSA is the agency responsible for monitoring and managing the rainfall data for the thirty years from 1981 to 2010. The database was used to extract and disaggregate the annual maximum series (AMS) from the collection of data.

For the length of the rainfall, the following intensities were selected: 10 minutes, 15 minutes, 30 minutes, 60 minutes, 120 minutes, and 240 minutes. In this case, the rainfall intensities were expressed in millimeters per minute, which were then divided by the hourly equivalent of the rainfall duration to convert to millimeters per hour. The standard deviation and average or mean of the intensities were calculated for each duration.

### Study Area

The capital of southeast Nigeria's Abia State is Umuahia. The rail line that connects Port Harcourt to the south and Enugu City to the north passes through Umuahia. 359,230 people call Umuahia home, according to the 2006 Nigerian census. The map of Umuahia is shown in Fig. 1.

**Fig. 1.** *Map of Umuahia, Abia State, Nigeria*

### Methods

### Flood Frequency Distribution Methods

We can examine frequency using a number of probability distribution functions, including the Pearson Type III, Gumbel Extreme Value type I, Normal, Log-Normal, and Log-Pearson Type II. In order to compute the necessary rainfall intensity using this equation, one must first determine the frequency factor $K_T$ using the techniques outlined below.

$$P_T = \bar{P} + k_T * \sigma \qquad (1)$$

Where $P_T$ is desired rainfall peak value for a specified frequency, $\bar{P}$ is mean of maximum rainfall corresponding to a specified duration, $K_T$ is frequency factor and σ is the standard deviation of observed rainfall. The rainfall intensity, *i* (mm/hr) of a specified return period, T is obtained from:

$$i = \frac{P_T}{d} \qquad (2)$$

where d = Duration in hours

### Gumbel (EV1)

The Extreme Value type I distribution (EV1), the frequency factor $K_T$ is given by Eqn. (3):

$$K_T = \frac{-\sqrt{6}}{\pi}\left\{0.5772 + \left[ln\left(\frac{T}{T-1}\right)\right]\right\} \qquad (3)$$

### Normal (N2)

A symmetric probability distribution about the mean, the two-parameter normal distribution (N2) approach—also called the Gaussian distribution—indicates that data near the mean occur more frequently than data far from the mean. When the normal distribution formula is plotted graphically, it looks like a bell curve.

$$z = w - \frac{(2.515517+0.80285*w+0.010328*w^2)}{(1+1.432788*w+0.189269*w^2+0.001308*w^3)} \qquad (4)$$

Where, z is the standard Variate as shown above and w is the intermediate variable with the formula

$$w = \sqrt{\left[ln\left(\frac{1}{p^2}\right)\right]} \qquad (5)$$

**Normal Distribution (N2) Procedure**

The method for calculating the rainfall intensity duration frequency curve for each returning period using the Normal Distribution methodology is described in detail below.

i. By using eqn. (5) to determine the value of an intermediate variable w.

ii. The normal variate Z can be found by using eqn. (4).

iii. Using the following mean values, durations, and standard deviations:

Obtain Rainfall intensity by applying equation

$$\mathrm{X_T} = \bar{X} + \mathrm{k_T} * S \qquad (6)$$

Where $\bar{X}$ is the mean, $\mathrm{k_T}$ is the return period, S is the standard deviation

iv. Plot a graph of all the Normal distribution values gotten alongside each durations as shown in Table 4

Frequency factor, $K_T$ as:

$$K_T = \frac{X_T - \mu}{\sigma} \qquad (7)$$

This is the same as the standard normal variate, z. The value of $z$ corresponding to an exceedance probability, $p = {}^{1}/_{T}$, can be calculated by finding the value of an intermediate variable w.

**Goodness-of-Fit Tests**

To ascertain how well the given data will fit the established models, a test or hypothesis known as the goodness of Fit (GoF) test is employed. The Gumbel and normal distribution models were used in this investigation. The goodness of fitness of a statistical model provides an objective assessment of how "close" the observed data (Robs) and rainfall simulated data (Rsim) are to each other. Several performance metrics are used to assess the goodness-of-fit of the model data, some of which are as follows:

i. Root Mean Square Error (RMSE)
ii. Coefficient of Efficiency (NSE)
iii. Coefficient of Determination ($R^2$)
iv. Chi-Square ($\chi^2$)
v. Probability Plot Correlation Coefficient (PPCC)

However, the Coefficient of Determination ($R^2$) and Nash-Sutcliffe Efficiency (NSE) were used in the Goodness of Fit test to choose the optimum fit distribution for this investigation.
Following the procedures as listed below.

Accomplishing for a duration of 10, 15, 30, 60, 120, and 240 minutes, for each return term ranging from 5, 10, 15, 20, 25, and 30 years.
i. Determining the mean observed rainfall intensity ($I_o$), the modelled rainfall intensity ($I_m$), and the observed rainfall intensity ($I_o$), all expressed in millimeters per hour, using the Talbot IDF equation as described in equation (10).

i. Finding their mean and sum values.

ii. Using the algorithm to obtain the NSE values, whose permissible range is between 0 and 1.0

$$NSE = 1 - \frac{\sum(I_o - I_m)^2}{\sum(I_o - \bar{I_o})^2} \qquad (8)$$

iv. Using the formula to get the correlation coefficient R in order to obtain the Coefficient of Determination, or $R^2$ value

$$R = \frac{N\sum XY - \sum X \sum Y}{\sqrt{N\sum_{X}2 - (\sum X)^2}\sqrt{N\sum Y^2 - (\sum Y)^2}} \qquad (9)$$

v. Plotting the modelled versus observed rainfall intensity graphs for each year, as indicated below, to see which best fits the test results, as indicated below in the derived NSE values.

## IDF Equations

The empirical equations that depict the relationship between the maximum rainfall intensity (as the dependent variable) and other relevant factors like the duration and frequency (as the independent variables) are known as the rainfall intensity-duration-frequency (IDF) formulae. The literature on hydrology applications has a number of frequently used functions. Four basic forms of equations used to describe the rainfall intensity duration relationship are summarized [3]:

Talbot Equation:

$$i = \frac{a}{(d+b)} \qquad (10)$$

Bernard Equation:

$$i = \frac{a}{d^e} \qquad (11)$$

Kimijima Equation:

$$i = \frac{a}{(d^e+b)} \qquad (12)$$

Sherman Equation:

$$i = \frac{a}{(d+b)^e} \qquad (13)$$

Where *i* is the rainfall intensity (mm/hr.): *d* is the duration (minutes): *a*, *b* and *e* are the constants parameters related to the metrological conditions. For this study, the Talbot was used to Estimate the Rainfall Intensity Duration in Umuahia, Nigeria [2].

## RESULTS AND DISCUSSION

### Results

The Annual Maximum Rainfall Intensity (mm/hr) for Umuahia, Nigeria, is shown in Table 1 for various years and times. The estimated mean and standard deviations for various years and durations are presented in Table 2.

**Table 1.** Annual Maximum Rainfall Intensity (mm/hr.) for Various Durations and Years for Umuahia Nigeria

| YEAR | 10 min | 15 min | 30 min | 60 min | 120 min | 240 min |
|---|---|---|---|---|---|---|
| 1981 | 509.4 | 339.6 | 169.8 | 84.9 | 42.4 | 21.2 |
| 1982 | 672.0 | 448.0 | 224 | 112.0 | 56.0 | 28.0 |
| 1983 | 453.0 | 302.0 | 151 | 75.5 | 37.8 | 18.9 |
| 1984 | 762.6 | 508.4 | 254.2 | 127.1 | 63.5 | 31.8 |
| 1985 | 552.0 | 368.0 | 184 | 92.0 | 46.0 | 23.0 |
| 1986 | 979.2 | 652.8 | 326.4 | 163.2 | 81.6 | 40.8 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 1987 | 979.2 | 652.8 | 326.4 | 163.2 | 81.6 | 40.8 |
| 1988 | 652.8 | 435.2 | 217.6 | 108.8 | 54.4 | 27.2 |
| 1989 | 774.0 | 516.0 | 258 | 129.0 | 64.5 | 32.3 |
| 1990 | 411.0 | 274.0 | 137 | 68.5 | 34.3 | 17.1 |
| 1991 | 794.4 | 529.6 | 264.8 | 132.4 | 66.2 | 33.1 |
| 1992 | 588.6 | 392.4 | 196.2 | 98.1 | 49.0 | 24.5 |
| 1993 | 690.6 | 460.4 | 230.2 | 115.1 | 57.5 | 28.8 |
| 1994 | 730.2 | 486.8 | 243.4 | 121.7 | 60.8 | 30.4 |
| 1995 | 971.4 | 647.6 | 323.8 | 161.9 | 80.9 | 40.5 |
| 1996 | 477.0 | 318.0 | 159 | 79.5 | 39.8 | 19.9 |
| 1997 | 971.4 | 647.6 | 323.8 | 161.9 | 80.9 | 40.5 |
| 1998 | 344.4 | 229.6 | 114.8 | 57.4 | 28.7 | 14.3 |
| 1999 | 906.0 | 604.0 | 302 | 151.0 | 75.5 | 37.8 |
| 2000 | 477.0 | 318.0 | 159 | 79.5 | 39.8 | 19.9 |
| 2001 | 748.2 | 498.8 | 249.4 | 124.7 | 62.3 | 31.2 |
| 2002 | 553.2 | 368.8 | 184.4 | 92.2 | 46.1 | 23.0 |
| 2003 | 484.2 | 322.8 | 161.4 | 80.7 | 40.3 | 20.2 |
| 2004 | 398.4 | 265.6 | 132.8 | 66.4 | 33.2 | 16.6 |
| 2005 | 1239.6 | 826.4 | 413.2 | 206.6 | 103.3 | 51.6 |
| 2006 | 822.0 | 548.0 | 274 | 137.0 | 68.5 | 34.3 |
| 2007 | 751.2 | 500.8 | 250.4 | 125.2 | 62.6 | 31.3 |
| 2008 | 584.4 | 389.6 | 194.8 | 97.4 | 48.7 | 24.3 |
| 2009 | 501.0 | 334.0 | 167 | 83.5 | 41.8 | 20.9 |
| 2010 | 1102.8 | 735.2 | 367.6 | 183.8 | 91.9 | 45.9 |

**Table 2.** Estimation of the Mean and Standard Deviations for Various Durations and Years for Umuahia, Nigeria

| **DURATION (MIN)** | **MEAN $\bar{x}$** | **STANDARD DEVIATION** |
|---|---|---|
| 10 | 696.0 | 226.9649 |

| | | |
|---|---|---|
| 15 | 464.0267 | 151.3099 |
| 30 | 232.0133 | 75.65495 |
| 60 | 116.0067 | 37.82748 |
| 120 | 58.00333 | 18.91374 |
| 240 | 29.00167 | 9.456869 |

**Frequency Distribution Methods**

Using Gumbel and Normal distribution methods, frequency distribution approaches are shown in Tables 3 and 4 to fit a probability distribution to the rainfall intensity data for Umuahia Metropolis, Nigeria, for various durations and periods. Figures 2 and 3 show rainfall intensity duration frequency curves that were produced by applying the normal distribution and Gumbel techniques.

**Table 3.** Gumbel Distribution Values for Various Durations and Periods for Umuahia, Nigeria

| **D(min)** | **10** | **15** | **30** | **60** | **120** | **240** |
|---|---|---|---|---|---|---|
| T=5 | 859.3103755 | 572.8735837 | 286.4367918 | 143.2183959 | 71.60919796 | 35.80459898 |
| T=10 | 992.0916092 | 661.3944061 | 330.6972031 | 165.3486015 | 82.67430076 | 41.33715038 |
| T=15 | 1067.005659 | 711.3371061 | 355.668553 | 177.8342765 | 88.91713826 | 44.45856913 |
| T=20 | 1119.458556 | 746.3057042 | 373.1528521 | 186.5764261 | 93.28821303 | 46.64410651 |
| T=25 | 1159.86105 | 773.2406999 | 386.62035 | 193.310175 | 96.65508749 | 48.32754374 |
| T=30 | 1192.729619 | 795.1530792 | 397.5765396 | 198.7882698 | 99.3941349 | 49.69706745 |

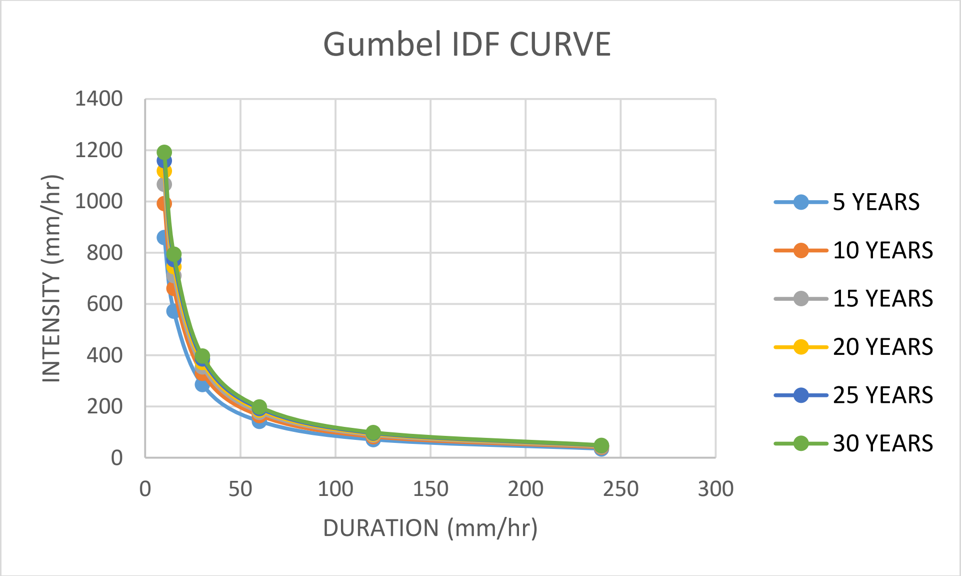

**Fig. 2.** *Rainfall Intensity Duration Frequency Curves by Gumbel Distribution Method*

**Table 4.** Normal Distribution Values for Various Durations and Periods for Umuahia, Nigeria

| **D(min)** | **10** | **15** | **30** | **60** | **120** | **240** |
|---|---|---|---|---|---|---|
| T=5 | 887.0213906 | 591.3475937 | 295.6737969 | 147.8368984 | 73.91844922 | 36.95922461 |
| T=10 | 986.9476715 | 657.9651144 | 328.9825572 | 164.4912786 | 82.24563929 | 41.12281965 |
| T=15 | 1036.801961 | 691.2013075 | 345.6006538 | 172.8003269 | 86.40016344 | 43.20008172 |
| T=20 | 1036.801961 | 691.2013075 | 345.6006538 | 172.8003269 | 86.40016344 | 43.20008172 |
| T=25 | 1093.473119 | 728.9820793 | 364.4910396 | 182.2455198 | 91.12275991 | 45.56137996 |
| T=30 | 1093.473119 | 728.9820793 | 364.4910396 | 182.2455198 | 91.12275991 | 45.56137996 |

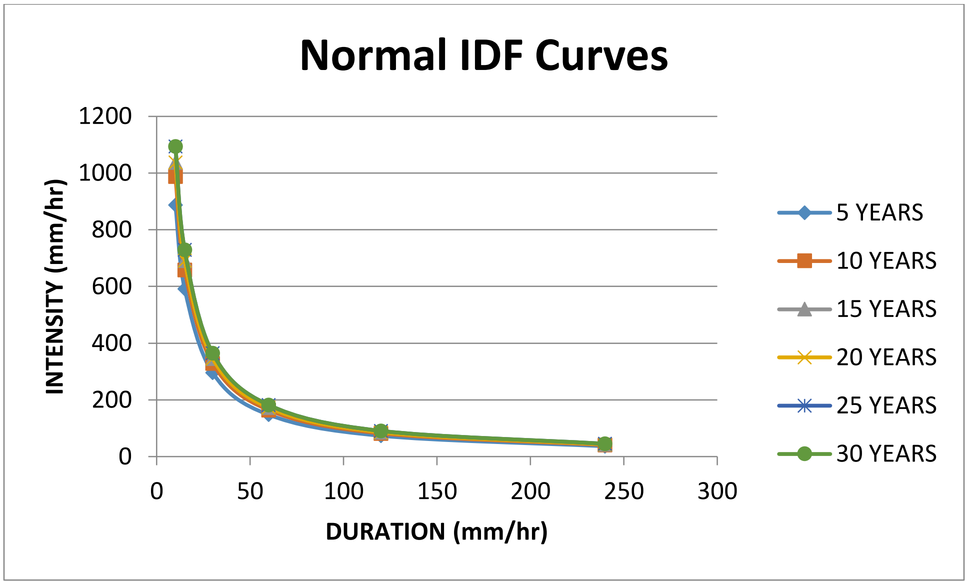

**Fig. 3.** *Rainfall Intensity Duration Frequency Curves by Normal Distribution Method*

### Goodness of Fit Test Results

The goodness of fit test findings, which were utilized to select the best fit distribution for this study, are presented in Table 5. The Gumbel and Normal Goodness of Fit test plots for each return time are shown in Figures 4 and 5.

**Table 5.** Goodness of Fit Test Results

| **Return Periods** | **5 years** | **10 years** | **15 years** | **20 years** | **25 years** | **30 years** |
|---|---|---|---|---|---|---|
| **Gumbel** | 0.877780369 | 0.598152718 | 0.369051954 | 0.178011092 | 0.013658986 | -0.1310878 |
| **Normal** | 0.832772273 | 0.611995706 | 0.467611824 | 0.467611824 | 0.27580705 | 0.27580705 |

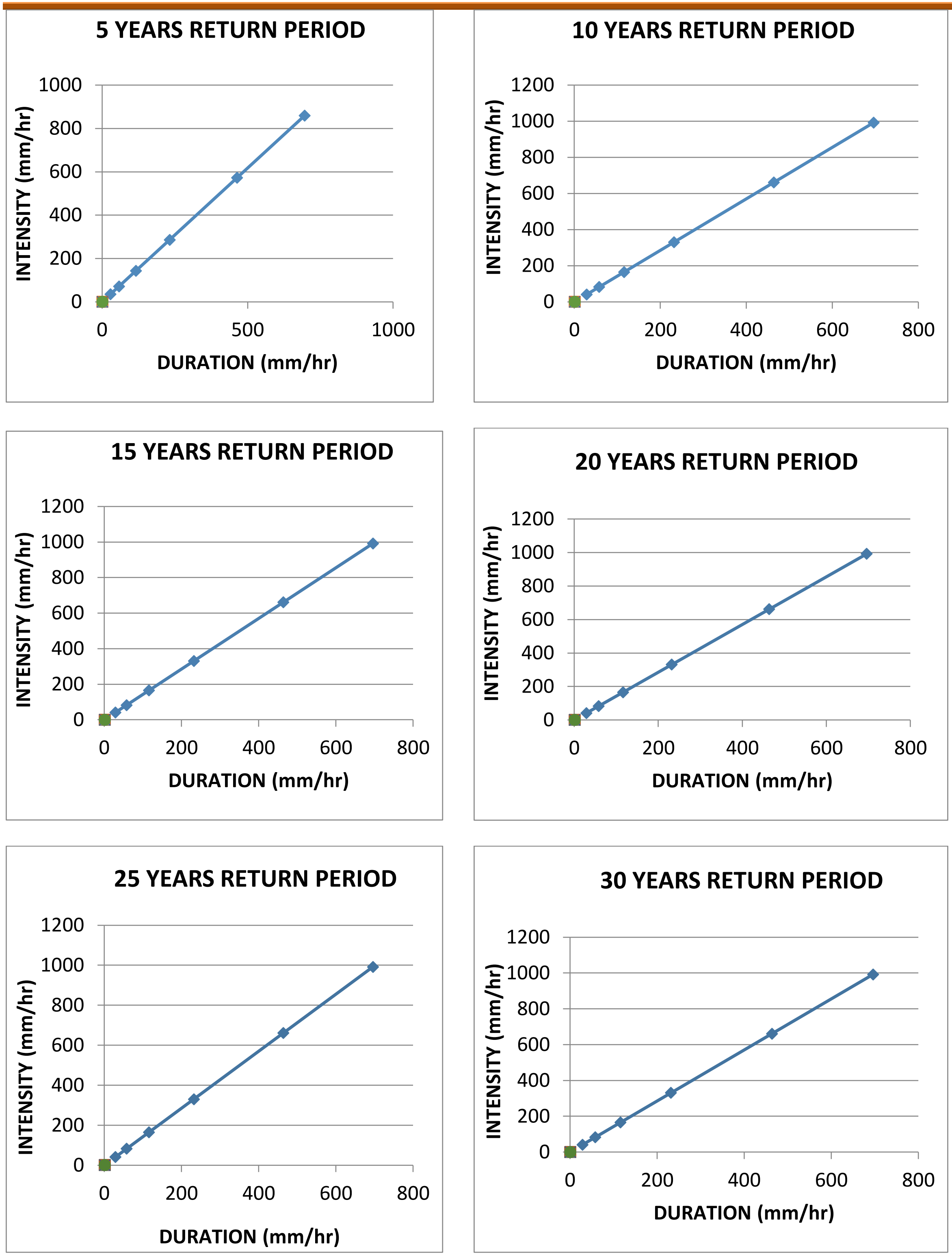

**Fig. 4.** *Gumbel Goodness of fit Test Graph for Each Return Period*

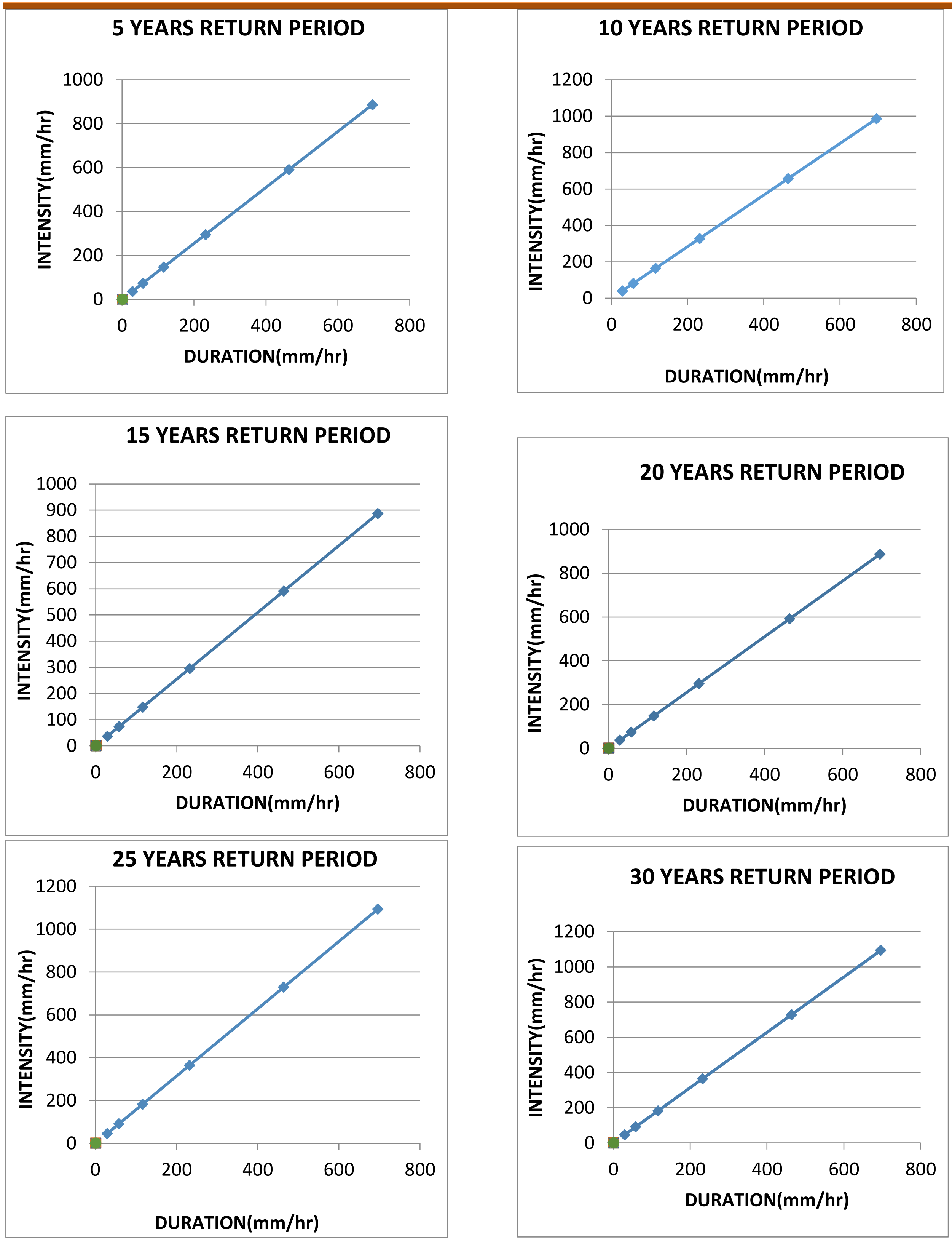

**Fig. 5.** *Normal Goodness of fit Test Graph for Each Return Period*

## Discussion of Results

The Gumbel and Normal distribution models, which are two IDF distribution models, were employed to evaluate the rainfall data quality spanning from 1981 to 2010.
The evaluation of the annual maximum, mean, and standard deviation for each of the time intervals of 10, 15, 30, 60, 120, and 240 minutes is presented in Tables 1 and 2. The Gumbel and Normal Distribution Values for Various Durations and Return Periods for Umuahia, Nigeria, are shown in Tables 3 and 4.
The IDF curves, which were also produced right after the study of the various durations, are shown in Figures 2 and 3.
Gumbel and the IDF distribution models, whose performances were evaluated using a model performance index, provided a coefficient of determination ($R^2$) above 90% and extremely good model efficiency, were used to examine the accuracy of the rainfall data for the years 1981–2010.
The goodness-of-fit of the distribution was computed in Excel. It shows the parameters of the model and the goodness-of-fit of the Normal and Gumbel distributions. The results of goodness-of-fit tests for the set data for Umuahia, Nigeria, are shown in Table 5 and Figures 4 and 5.
To evaluate the quality of rainfall data for the years 1981–2010, the IDF distribution was employed. According to the results, the Gumbel Distribution model best matches the data collection. Since these values exceed those of the Normal Distribution, it may be said that the Gumbel is a more reliable measure for the area of study. After the aforementioned steps, IDF curves were generated for each return period, which varied from 5, 10, 15, 20, 25, and 30 years, with a duration of 10, 15, 30, 60, 120, and 240 minutes.

## CONCLUSION

The Nigeria Hydrological Services Agency (NIHSA), located in Abuja, provided data for this study regarding the amount, duration, and frequency of rainfall in Umuahia, Nigeria. The Gumbel and Normal distribution procedures were the two distinct techniques utilized for arranging the data for frequency analysis. The required curves for the study area were created using the Talbot Rainfall Intensity Duration Frequency (IDF) Equation, an empirical relationship between maximum rainfall intensity (as a dependent variable) and other parameters of interest, such as rainfall duration and frequency (as independent variables).
The Gumbel Distribution model best fits the data collection, as indicated by the findings. It might be argued that the Gumbel is a more trustworthy metric for the field of study because these values are higher than those found in the Normal Distribution.
Using the above outlined methods, IDF curves for rainfall intensity duration frequencies were produced. For every return period of 5, 10, 15, 20, 25, and 30 years, the curves had varying durations: 10, 15, 30, 60, 120, and 240 minutes.
Required variables for the hydrological design of hydraulic structures are the essential storm duration (a concentration of times for determining peak discharge) for the catchment area, the selected frequency (return period), and the intensity of rainfall.

## REFERENCES

[1] Al Saji, M. O. (2015). Design impact and significance of non-stationarity of variance in extreme rainfall. *Proceedings of the International Association of Hydrological sciences*, 371(1), 117-123.
[2] Dar, A. Q., Maqbool, H., Raazia S. (2016). An empirical formula to estimate rainfall intensity in Kupwara region of Kashmir valley, J and K, India. MATEC Web of Conferences, 57. ICAET 2016, Available at:https://www.researchgate.net/publication/302977555_An_empirical_formula_to_estimate_rainfall_intensity_in_Kupwara_region_of_Kashmir_valley_J_and_K_India#:~:text=10.1051/matecconf/20165703010.
[3] Arimieari, L.W., Egop, S.E., & Arigbe, O.F. (2023). Estimation of Rainfall Duration Frequency Curve for Owerri, Nigeria. *Journal of Water Resource Research and Development*, 7(1), 1-12.

[4] Hailegeorgis, T. T., Alfredsen, K. (2017). Regional flood frequency analysis and prediction in ungauged basins including estimation of major uncertainties for mid Norway, *Journal of Hydrology: Regional Studies*, 9, 104-126, Available at: https://doi.org/10.1016/j.ejrh.2016.11.004.

[5] Hussain, Z. & Pasha, G. (2019). Regional flood frequency analysis of the seven sites of Punjab, Pakistan, using L-Moments. *Water Resour. Management*.

[6] Arimieari, L.W., Egop, S. E., & Sanderson, O. J. (2022). An Empirical Model to Estimate Rainfall Intensity Duration in Asaba Metropolis, Nigeria. *Journal of Water Resources and Pollution Studies*, 7(2), 38-50.

[7] Okoyeh E. I, Okeke H. C, et al (2015). Climate Change and Groundwater Resources of Part of Lower Niger Sub-Basin around Onitsha, Nigeria, *International Journal of Scientific &Engineering Research*, 6(9), 1463- 1467.

[8] Hadadin, N.A. (2005). Rainfall intensity–duration–frequency relationship in the Mujib basin in Jordan. *Journal of Applied Science*, 8(10), 1777–1784.

[9] Gabriele, V. & James, A. S. (2010). Flood Peak Distributions for the Eastern United States. *Water Resources Research*, 46(6), 1-17.

[10] Calenda, G., Mancini, C., & Volpi, E. (2009). Selection of the probabilistic model of extreme floods: The case of the River Tiber in Rome. *J. Hydrol*.