Estimating the Intensity Equations for Rain Intensity Frequency Curves (Mosul /Iraq)

Danya N. Hussein a, Shatha H.D. AL-Zakar b *, Abdulwahab Mohammad Yonis c

a Engineer, Nineveh Sewers Directorate, Nineveh, Iraq.
b College of Engineering, University of Mosul, Mosul, Iraq.
c College of Engineering, AL-Kitab University, Kirkuk, Iraq.

Keywords:
Intensity Duration Frequency (IDF) - Curves; Probability Distributions; Return Period; Goodness of Fit; Mosul City.

Abstract: The relationship between the rainfall intensity, duration, and frequency (IDF) is widely used in water resources engineering in designing hydraulic structures, such as culverts and sewage systems, and for reducing and controlling floods. In this study, the curves of (IDF) were found for five rain stations registered in different regions in northern Iraq, i.e., Mosul, Tal-Afar, Sinjar, Rabia, and Tal-Abta, and for different periods extending from 1990 to 2019. The Indian Meteorological Department (IMD) empirical equation was utilized to obtain data during short periods, i.e., 10, 20, 30, 60, 120, 180, 360, 720, and 1440 minutes. Also, three probability distributions were used: the Gumble Distribution, the Log Pearson Type III Distribution, and the Log-Normal Distribution, for various return periods (2, 5, 10, 25, 50, and 100) years. Easy fit 5.6 software includes the tests (Kolmogorov-Smirnov, Anderson -Darling (AD), Chi-Squared ($\chi^2$)) was used to determine the most suitable distribution for the observed data among the three used distributions. The results demonstrated an insignificant difference between the three applied distributions and the statistics values that fell within the significance level, with priority to the Log Pearson-III distribution.
Tikrit Journal of Engineering Sciences Volume 30 No. 3 2023 Page 39


Tikrit Journal of Engineering Sciences Volume 30 No. 3 2023 Page 39


Tikrit Journal of Engineering Sciences Volume 30 No. 3 2023 Page 39

Iraq. Maximum daily rainfall was collected and disaggregated to short duration for different periods, i.e., 10, 20, 30, 60, 120, 180, 360, 720, and 1440 minutes. Three methods were used to examine the estimated rainfall in (mm) and its intensities in (mm/h) for various return intervals and durations (Gumbel, LPT III, and Log-Normal).

2. STUDY AREA AND METHODS
2.1. Study Area and Rainfall Collection Data
The study area is located in north Iraq and bordered by longitudes (41°-43°) East and latitudes (35°-36°) North. The region includes six selected stations, i.e., Mosul, Tal- Afar, Sinjar, Rabia, and Tal -Abta. The region is surrounded by a mountain range from the north and east and plains from the west and south inside Iraq. The average annual precipitation ranges between 350 and 1000 mm [8]. Fig. 1 is a map showing the climate monitoring stations’ locations included in the study. Table 1 is a map showing the climate monitoring stations’ locations included in the study. Table 1 shows a description of the statistical and hydrological characteristics of the data of those stations.

Fig.1 Location Map of the Selected Stations in the study.

Table 1 Statistical and Hydrological Features of Data at the Selected Stations in the Study Area.

<table>
<thead>
<tr>
<th>Station Name</th>
<th>Year (M)</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Duration (Year)</th>
<th>Maximum Daily Rain (Mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mosul</td>
<td>1990-2019</td>
<td>36°19'</td>
<td>43°09'</td>
<td>96.2</td>
<td></td>
</tr>
<tr>
<td>Tal-Afar</td>
<td>2000-2019</td>
<td>36°22'</td>
<td>42°99'</td>
<td>95.6</td>
<td></td>
</tr>
<tr>
<td>Sinjar</td>
<td>1990-2019</td>
<td>36°19'</td>
<td>41°59'</td>
<td>83.2</td>
<td></td>
</tr>
<tr>
<td>Rabia</td>
<td>1990-2019</td>
<td>36°48'</td>
<td>42°06'</td>
<td>95.6</td>
<td></td>
</tr>
<tr>
<td>Tal-Abta</td>
<td>2005-2015</td>
<td>33°33'</td>
<td>42°36'</td>
<td>83.8</td>
<td></td>
</tr>
</tbody>
</table>

To generate IDF curves, daily rainfall depth in mm was obtained from the Iraqi Meteorological Organization and Seismology (unpublished data) for different periods at five stations distributed in Mosul.

2.2. Generation Short-Duration Rainfall
The rainfall depth of short periods can be derived from the daily maximum rainfall data, given that it is unavailable at the meteorological station, depending on the Indian equation that was proposed by Indian Meteorological Department, IMD, which is known as the empirical reduction formula, as given in Eq. (1), from Ref’s. [3, 9]. IMD applies to calculating the precipitation depth for different periods of less than 24 hours.

\[ P_t = P_{24} \times \left( \frac{t}{24} \right)^{1/3} \]  

Where

- \( P_t \): rainfall depth in mm for time \( t \) (hours).
- \( P_{24} \): maximum daily rainfall depth in (mm)
- \( t \): the duration (hours) of rain for which rainfall depth is needed.

The rainfall was disaggregated into shorter durations of 0.1666, 0.333, 0.5, 1.0, 2.0, 3.0, 6.0, 12.0, and 24.0 hours.

3. ANALYSIS OF FREQUENCY DISTRIBUTION AND DEVELOPMENT IDF CURVES
Obtaining the best probability distribution among the different theoretical distributions is necessary to analyze the rainfall intensity curves of rainfall amounts at constant duration. For this study, three distributions, i.e., the Gumbel distribution, the Log Pearson III distribution, and the Log Normal distribution, were used to assess the annual maximum values for all accessible periods statistically.

3.1. Gumbel Distribution (GD)
This method is one of the most used probability distribution functions in hydrological studies to predict flood peaks and maximum rain intensity. GD method is used to analyze the (IDF) curves due to its suitability for extreme data modeling. The Gumbel distribution calculates for various return periods (2, 5, 10, 25, 50, and 100 years) for each period, as shown in the following equations [10, 11].

\[ P_t = P_{\text{ave}} + K \times S \]  

Where

- \( P_t \) is the frequency of rainfall in (mm) for each period and for a return period (T) in years.
- \( P_{\text{ave}} \) is the average of maximum rainfall values for specific periods, as in Eq. (3).

\[ P_{\text{ave}} = \frac{1}{n} \sum P_i \]  

\[ K \] is the Gumbel frequency coefficient found in special tables [12] or from Eq. (4).

\[ K = -\sqrt{\frac{\pi}{2}} \left( 0.5772 + \ln \left( \frac{\ln \frac{T}{T_c}}{\ln T} \right) \right) \]  

\( S \) is the standard deviation of \( P \) rainfall data, Eq. (5):

\[ S = \sqrt{\frac{1}{n-1} \sum (P_i - P_{\text{ave}})^2} \]  

Then rain intensity \( I_t \) in (mm/hour) for any return period \( T \) and duration \( t \) can be found by applying Eq. (6):

\[ I_t = \frac{P_t}{T} \]  

Where

- \( T \) is the duration in hours.

3.2. Log-Pearson Type III
The Log-Pearson Type III (LP III) distribution method is applied to calculate the rainfall intensity for different durations and return
periods to draw IDF curves using the same Gumbel method, as follows [12].

\[ \log x = \frac{1}{n} \sum_{i=1}^{n} \log x_i \]  

\[ \sigma_{\log x} = \sqrt{\frac{\sum (\log x_i - \log x)^2}{n-1}} \]  

\[ C_s = \frac{n \sum (\log x_i - \log x)^3}{(n-1)(n-2)(\sigma_{\log x})^3} \]  

\[ \log x = \log x + k \sigma_{\log x} \]  

where \( \log x \) and \( \sigma_{\log x} \) are the mean and standard deviation based on logarithmic transformation. \( C_s \) is the coefficient of skewness obtained from Ref. [12].

3.3. Log-Normal Distribution

To use the log-Normal distribution approach, the rainfall quantities must be transformed into logarithmic numbers, i.e., the logarithm of the statistical variables. The frequency coefficient \( K \) taken from Ref. [13] follows the same process as the (LPIII) distribution; however, it is equivalent to the standard natural variable \( Z \).

3.4. Derivation of IDF Empirical Equation

The IDF relationship is a widely used tool in water resources engineering, i.e., planning, designing, and operating water facilities. This equation expresses the linkage between the maximum rainfall intensity as a dependent variable and other parameters, such as duration and frequency of rainfall, as independent variables. So, (Bernard Equation) was applied to find the rainfall intensity [1], which is:

\[ I_T = \frac{c \times T_r \times m}{a \times e} \]  

where \( I_T \) is intensity (mm/hr), \( T_r \) is the return period (years), and \( d \) is the duration (hours). The coefficients \( c, m, \) and \( e \) are regional coefficients [10]. A log-log graph was created using non-linear regression analysis between rainfall intensity and duration for each return period to determine the constant (e). Values derived from interception points for each obtained equation were plotted on a log-log scale with recurrence intervals to determine the \( c \) and \( m \) values [11].

4. GOODNESS OF FIT

The relationship between observed and expected frequencies was examined using the goodness of fit test to select the best suitable type of probability distribution function (PDF) for the rainfall data in the research area. Three different goodness of fit tests were used in this investigation. An explanation of each is reported below.

4.1. Chi-Square Test

The chi-square distribution is one of the continuous probability distributions that was first described by Karl Pearson in 1900. It is expressed as follows.

\[ (x^2) = \frac{\sum (O_i - E_i)^2}{E_i} \]

where \( x^2 \) is a random variable for which the sampling distribution is very close to the chi-square distribution. The symbols \( O_i \) and \( E_i \) refer to the observed and expected frequencies for the interval of the i-th histogram class. The symbol \( k \) denotes the number of class intervals. If the observed frequencies are near the corresponding expected frequencies, the \( x^2 \) value will be small, showing a good fit; if not, it is a bad fit. Acceptance of the zero hypothesis results from a good fit, while rejection results from a bad fit. The critical region will therefore be in the right tail of the \( x^2 \). The critical value is found by \( x^2 \) tables [11, 14]. The chi-square test of fit was run using the software Simple Fit 5.6 [15].

4.2. Kolmogorov-Smirnov Test

The test determines whether there is a good agreement between the observed data’s frequency of occurrence and the expected frequency inferred from the distribution. Easy Fit 5.6 was utilized. The Kolmogorov-Smirnov test result that is lower value is the best fit for the selected distribution, and vice versa. Some steps can be taken to pass this test:

- The data is put in descending order.
- The observed data’s cumulative probability \( P(xi) \) is found using the Weibull equation [11].

\[ P(xi) = (m/n) + 1 \]

Where \( n \) is the number of historical data, and \( m \) is the descending order of values.

- For all observed data, the assumed distribution was used to determine the theoretical cumulative probability \( F(xi) \). The maximum absolute difference determined from the following equation is the Kolmogorov-Smirnov test statistic (\( \Delta \)), calculated from the following equation [16].

\[ \Delta = | P(xi) - F(xi) | \]

- The tables saved in the Easy Fit 5.6 program are used to get the tabular (\( \Delta O \)) value of the Kolmogorov-Smirnov statistic for a specific degree of probability.

- If the statistic’s value (\( \Delta \)) is less than the value of (\( \Delta O \)), the hypothesis that the distribution fits with the assumed probability level is accepted.

4.3. Anderson-Darling (AD) Test

This test found good agreement between the observed data’s frequency of occurrence and the predicted frequency derived from the distributions. If the \( A_2 \) statistic’s value exceeds the test’s critical value, the null hypothesis is rejected at level (\( \alpha \)) of probability. It is calculated from the following equation:

\[ A^2 = -n - S \]
The data are $Y_1, Y_2, ..., Y_n$, and $n$ is the number of data.

5. RESULTS FINDING AND DISCUSSION

Figures (2)-(16) display the results from the IDF curves obtained from the three used techniques, i.e., Gumbel, LPT III, and Log Normal, in all stations. The results obtained by the three approaches were consistent. After analyzing the equations derived from the intensity-duration-frequency (IDF) curves for the three probabilistic distributions that were applied for all stations, the coefficients ($c$), ($m$), and ($e$) were found. These parameters’ values with the rain intensity formula obtained are shown in Table 2 for all stations and the three probability distributions. The goodness-of-fit of the three probability distributions was used in the study for 24 hours to choose the best suitable type of probability distribution function (PDF) for the studied area’s rainfall data. Table 3 shows the test results. It can be shown that the Log Person III distribution was the most reliable in the study area because it was the first, i.e., the lowest value of the statistic among most distributions.
Fig. 4 Log-Normal Method IDF Curves at Mosul Station.

Fig. 5 Gumbel Method IDF Curves at Tal-Afar Station.

Fig. 6 Log Pearson III Method IDF Curves at Tal-Afar Station.
Fig. 7 Log Normal Method IDF Curves at Tal-Afar Station.

Fig. 8 Gumbel Method IDF Curves at Sinjar Station.

Fig. 9 Log Pearson III Method IDF Curves at Sinjar Station.
Fig. 10 Log Normal Method IDF Curves at Sinjar Station.

Fig. 11 Gumbel Method IDF Curves at Rabia Station.

Fig. 12 Log Pearson III Method IDF Curves at Rabia Station.
Fig. 13 Log Normal Method IDF Curves at Rabia Station.

Fig. 14 Gumbel Method IDF Curves at Tal-Abta Station.

Fig. 15 Log Pearson III Method IDF Curves at Tal-Abta Station.
Normal Distribution, and for different return
i.e., Gumble, Log Pearson Type III, and Log -
duration-frequency (IDF) curves for five
The study included drawing rain intensity -
7. CONCLUSION
The study included drawing rain intensity- 
duration-frequency (IDF) curves for five
selected stations in north Iraq. Also, curve
equations were derived for the periods of 10,
20, 30, 60, 120, 180, 360, 720, and 1440
minutes using three probability distributions,
i.e., Gumble, Log Pearson Type III, and Log- 
Normal Distribution, and for different return
periods, i.e., 2, 5, 10, 25, 50, and 100 years. The
results showed that the rainfall intensity increased with return periods and decreased with increasing duration for all return periods.
Using the Gumbel distribution method, it was
found that the highest rain intensity was obtained from Sinjar station for all periods and return periods, followed by Mosul station, while Tal- 
Abta station recorded the lowest rain
intensity values for the same duration above. Also, it was noted that there was a slight difference in the results values between the three probability distributions. The Log Person III distribution method produced better results than the Gumbel distribution, followed by the Log Normal distribution method at the stations (Mosul, Sinjar, Tal- Afar, Rabia, and Tal-Abta) for the return periods (2,5,10) years, while the Log Normal distribution method gave higher values than the other two distributions in return periods (25, 50, and 100). According to the three goodness-of-fit tests' results, none of the distributions were rejected at the 5% and 1% probability levels, as the statistics values were less than the critical values for each test. Thus, it is possible to rely on the three probability distributions, with preference given to the Log Person III distribution for having the first order, i.e., the lowest value, in most tests and stations. The calculated rainfall intensities can be used when designing hydraulic structures, such as culverts, sewage works, and other water resources projects in Mosul.

REFERENCES


