EEET ECOLOGICAL ENGINEERING & ENVIRONMENTAL TECHNOLOGY

Ecological Engineering & Environmental Technology, 2025, 26(1), 316–332
https://doi.org/10.12912/27197050/196090
ISSN 2719–7050, License CC-BY 4.0

Received: 2024.11.01
Accepted: 2024.11.28
Published: 2024.12.01

Advanced uncertainty quantification in rainfall-intensity duration frequency curve modeling: A case study of Hilla City and surrounding regions, Iraq

Ameer Hashim Hussein[1], Thair Jabbar Mizhir Alfatlawi[2]

[1] Al-Mussaib Technical Institute, Al-Furat Al-Awsat Technical University,51009 Babylon, Iraq
[2] Department of Civil Engineering, College of Engineering, University of Babylon, Iraq
* Corresponding author's e-mail: inm.ame@atu.edu.iq

ABSTRACT

The rainfall-intensity duration frequency (IDF) relationship is mostly used in water resource engineering to plan, design, and operate water source projects and projects to manage flood dangers. Engineers must accurately calculate rainfall to design structures that effectively manage runoff collection, conveyance, and storage, as the hydrologic cycle relies on precipitation. An analysis is conducted on the annual rainfall measurements (mm) from five atmospheric observatories in Iraq (Najaf, Hilla, Kerbala, Diwaniya, and Baghdad) spanning 1989 to 2023. The objective is to determine the characteristics of the observed frequency distributions. The Gamma, Log Normal, and Normal distributions compare the data. Kolmogorov-Smirnov, Anderson-Darling, and Chi-Square are study tests. The IDF depict extreme rainfall values over 15, 30, and 60 minutes, with 5, 10, 15, and 50-year return periods. The results indicate that the Chi-Square test has the most optimal distribution among all the stations. The normal distribution was found to be the best for the years (5, 10, 15, and 50) through the IDF curves drawing of stations where the distributions were compared. Equations of Hilla station were found through an IDF equations curve to Hilla from its surrounding stations; and get the error. The result was good agreement, with a ratio Coefficient of Determination ranging from 83.2 to 94.7.

Keywords: intensity-duration-frequency (IDF) curves, Hilla City, frequency distributions, Kolmogorov-Smirnov, Anderson-Darling, hydrologic modeling.

INTRODUCTION

The IDF relationship serves as a reference for constructing drainage systems in engineering projects, enabling engineers to create flood control structures that are both secure and cost-effective (Alam et al., 2021). It indicates the frequency at which rainfall intensity will likely occur during a specified timeframe (Campos et al., 2020). IDF curves depict the highest rate at which rain falls during a specific time frame. Precise evaluation of rainfall is crucial for effective water resource administration. Hydrological models rely on accurate forecasts of mean precipitation. Intense precipitation events can significantly affect society, individuals, and the economy. Creating IDF curves for precipitation is a commonly employed technique in water resources projects and management, urban sewer design, geomorphological research, and the hydraulic design of facilities that regulate storm runoff, such as flood detention reservoirs and sewage networks. The Gumbel distribution is well-known for its application in developing IDF curves, which researchers and engineering services widely use (Ghahraman and Hosseini, 2005). Nevertheless, the creation of IDF curves is influenced by inherent errors stemming from multiple factors, including climatic conditions, the scarcity of long-term rainfall data, and the selection of statistical models employed in frequency analysis (Cooley et al., 2007; Huang et al., 2010). These uncertainties might cause differences in the expected levels of rainfall, which could lead to insufficient infrastructure planning and heightened susceptibility to severe weather events. Hence, this work aims to thoroughly investigate the uncertainties associated with the IDF curves for Hilla City and its environs. This project aims to improve the accuracy of rainfall predictions and support better water resource management in the region by analyzing the influence of various statistical models and data variability on the IDF curves.

Underestimating or overestimating rainfall intensities can have serious repercussions, resulting in insufficient infrastructure that collapses during intense occurrences or too expensive systems that are unnecessarily developed. Hence, conducting a comprehensive uncertainty analysis can effectively narrow the disparity between theoretical models and practical implementations, thereby guaranteeing the precision and relevance of the resulting IDF curves within the specific local setting. In their study, Silva et al. (2021) investigated the impact of non-stationarity and climate change on the future distribution of intense rainfall at six specific monitoring locations throughout Canada. By comparing the projected future rainfall patterns with historical data, the severity of extreme weather events is expected to rise in all the locations studied, elevating the risk level. Campos et al. (2020) examined the correlation between the partial-area effect and the equations used to calculate rainfall IDF. The partial-area effect occurs when the rainfall duration is shorter than the time it takes for water to reach the outflow, also known as the time of concentration. This phenomenon is relevant to the rational method used in hydrology. AL-Dulaimi et al., (2020) forecast monthly rainfall for specific regions by analyzing how long-term trends, seasonal changes, periodic patterns, and random fluctuations affect time series data. The study also studied northern Iraqi rainfall daily, monthly, and annually. It also linked rainfall strength, length, and frequency at 2, 5, 10, 15, 25, 50, and 100 years to understand rainstorms better and suggest water management and treatment solutions for general and ungagged basins.

El Hannoun et al. (2023) investigated the variability of rainfall intensity over various durations. This study presents a novel model that considers the interdependence of rainfall intensities during various periods, such as brief bursts and prolonged storms. The model employs a D-vine copula statistical technique to accurately represent and analyze these interactions. Miller et al. (2022) examined a West African city's heavy rainfall event. This study examined the relationship between large-scale weather patterns, climate change, and urban flooding. The purpose is to inform regional decision-makers about climate change and extreme weather occurrences that cause urban flooding. Noor et al. (2021) examined Nigeria's peak daily rainfall distribution using Pearson, Log-Pearson, Gumbel, Log-Gumbel, Normal, and Log-Normal distributions. They selected 20 locations with 54 years of annual rainfall data for frequency analysis. Schlef et al. (2023) examined Accra, Ghana's maximum daily and two- to five-day rainfall over a return period of 2 to 100 years. To find the best-fitting distribution for describing these rainfall patterns, they investigated three popular probability distributions: normal, log-normal, and gamma. Yüksek et al. (2022) developed IDF curves for Turkish regions. They targeted the rainy Eastern Black Sea Basin. The researchers tested nine IDF algorithms, including their own, using rainfall intensity data from 5 minutes to 24 hours. Their criteria were accuracy (determination coefficient) and error (mean relative error). Their findings indicate that these reliable algorithms may motivate regional IDF investigations. Akpen et al. (2019) developed rainfall intensity-duration-frequency models for the Lokoja Metropolis in Kogi State, Nigeria. Analyzed precipitation data from the Nigeria Meteorological Agency was utilized to ascertain the frequency. Kareem et al. (2022) investigated innovatively to create IDF curves and empirical IDF estimates designed for Erbil. The researchers utilized the sole accessible recorded data, consisting of 39 years' annual maximum rainfall data from 1980 to 2018. The empirical equations and intensity-duration-frequency curves for standard durations and return periods were derived from daily rainfall data using statistical methods such as Gumbel and Log-Pearson Type III.

The primary goals of this study are to fit probability distributions to rainfall data from five different locations in Iraq (Diwaniya, Baghdad, Kerbala, Hilla and Najaf) and determine the most suitable distribution of statistics based on goodness of fit tests. They are acquiring rainfall intensity-duration-frequency curves for the five stations, contrasting IDF curves for all stations, and predicting the IDF curve for Hilla based on the neighbouring regions.

MATERIAL AND METHODS

Data selection for the gauge stations

Information regarding the amount of rainfall, specifically the total annual rainfall, is collected from five meteorological stations throughout Iraq. Rainfall depth data is available annually from 1989 to 2023, except for specific years where missing data was disregarded. Figure 1 depicts the research study conducted at the five stations. Figure 2 displays data on the intensity of rainfall at the five meteorological stations for durations of 15, 30, and 60 minutes.

Fitting probability distributions to the data

Statisticians frequently employ a range of probability distributions for statistical research, including the Normal (N.D), Log Normal (LN.D), and Gamma distributions (G.D). This study concentrates explicitly on these three probability distributions.

Normal distribution

The Gaussian distribution, referred to as the normal distribution, is a crucial illustration of an ongoing probability distribution. The function for density is defined as follows:

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

where: σ denotes the standard deviation, and μ indicates the mean.

The distribution function that corresponds to it is provided by Equation 2.

$$\mathrm{F(x)} = \mathrm{P(X \leq x)} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(x-\mu)^2/2\sigma^2} \mathrm{dx} \quad (2)$$

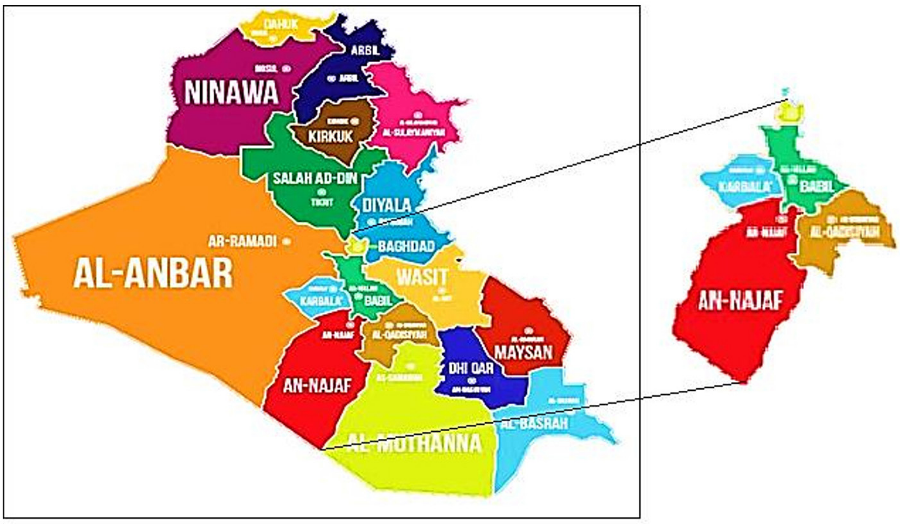

Figure 1. Location of station in the study

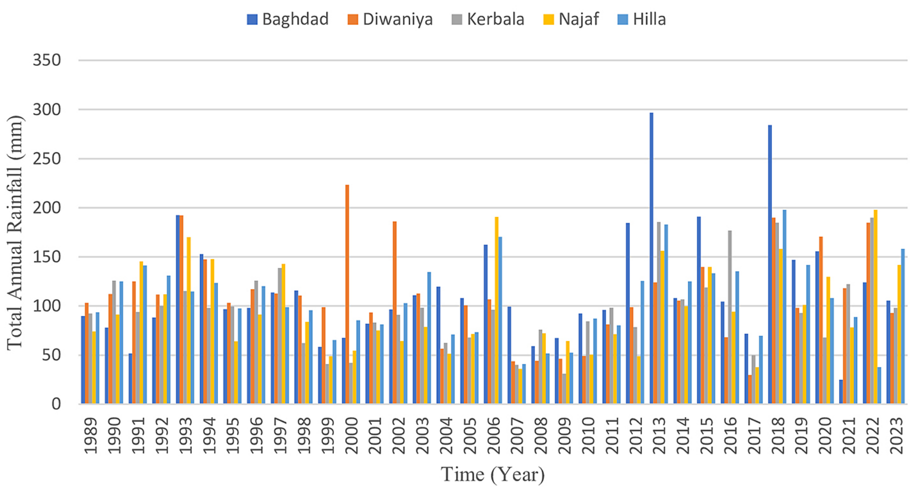

Figure 2. Total annual rainfall for the stations

The standardized variable z = (x − μ)/σ follows the conventional Normal distribution if x follows the N(μ, σ) distribution. N(0, 1) with a standard deviation of 1 and a mean of 0. The standard variable that corresponds to X is called Z. In these circumstances, the definition of a normal distribution can be used to obtain the density function for Z by setting μ = 0 and σ² = 1.

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} - \infty \leq z \leq \infty \quad (3)$$

The area under the curve is approximately 68% of the mean, with one standard deviation on either side. Within two standard deviations of the mean, the area under the curve is approximately 95%. The area of approximately 100% of the total area is within three standard deviations of the mean, as illustrated in Figure 3, (Takara and Takasao, 1988; Pecho et al., 2009).

Log-normal distribution

The log-normal distribution is similar to the Normal distribution, except for substituting the independent variable (x) with its logarithm. The log-normal distribution exhibits a pronounced positive skewness and is constrained on the left side by zero. If the logarithm of a variable x, usually the natural logarithm, follows a normal distribution, then the variable x has a log-normal distribution. The probability density function for a given variable is *y = ln(x)* (Shamkhi and Obeid, 2022).

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2{}_x}} exp\left(-\frac{(log\, x-\mu_y)^2}{2\sigma^2{}_y}\right) \; 0 \leq x \leq \infty \quad (4)$$

Gamma distribution

The gamma distribution is defined by a random variable X, which follows a gamma distribution with shape parameter (b > 0) and scale parameter (F > 0). The gamma distribution has two parameters, one for shape and one for scale. The formula for the probability density function (PDF) of this distribution, as stated in the research conducted by (Bolstad, 1998), is as follows:

$$f(x) = \frac{\lambda^{\beta} x^{(\beta-1)} e^{-\lambda x}}{\Gamma(\beta)} \; x > 0 \quad (5)$$

where: Γ denotes the gamma function.

Unlike the "Normal", the gamma distribution is not an exception to the rule that most distributions do not have mean and standard deviation as parameters. A gamma random variable has a mean of a variance of λ^{β}, (Bolstad, 1998).

Statistical estimators are evaluated on consistency, efficiency, and bias. These standards determine how well an estimated parameter matches its genuine value. As data increases, a method is consistent if the estimate approaches the actual value. An efficient technique estimates the parameter's true value with slight variance, and an unbiased estimator does not over- or under-estimate it. Two prevalent methodologies for deriving point estimators are the maximum likelihood technique and the method of moments. Maximum likelihood estimates are generally used due to their superior efficiency compared to moment estimators. Nevertheless, moment estimators can sometimes be calculated with greater ease. Both strategies can generate unbiased point estimators.

The moment approach compares relevant sample moments to population moments, which are anticipated values. Unidentified parameters determine population moments. After solving these equations, unknown parameter estimators are produced (Fadhel et al., 2017).

$$M_k = \frac{\sum_{i=1}^{n} xi^k}{n} \quad (6)$$

$$M_k = \mathrm{E}(x^k) = \int_{-\infty}^{\infty} x^k f(x)dx \quad (7)$$

where: k_{th} – sample moment about the origin;
k_{th} – population moment about the origin.

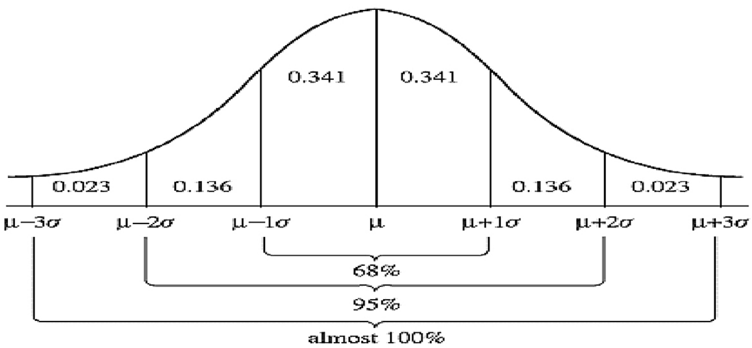

Figure 3. The normal distribution

The maximum likelihood method is more complicated than moments. Maximum Likelihood Estimation (MLE) calculates the likelihood of getting population parameters based on the measured record. MLE maximizes probability function. A likelihood function L is the joint density function of a set of random variables X-1 to X-n measured at f (x-1 to x-n), (Gu et al., 2022).

$$L = \prod_{i=1}^{n} f(x_i) \tag{8}$$

Since many probability density functions are exponential, utilizing the log-likelihood function is often more practical.

$$\ln L = \sum_{i=1}^{n} ln[f(x_i)] \tag{9}$$

The probability values for the three selected distributions (Normal, Log Normal, and Gamma) are computed using the moments method for all stations involved in this study, as illustrated in Figures 4 to 8. For the Normal distribution, the maximum likelihood method is the same as the moments method concerning the parameter estimates. The moments and the maximum likelihood estimates for all the stations are given in Table 1.

The mean and the standard deviation for the Log-Normal distribution are (Heidari et al., 2022)

$$\hat{\mu}_y = \ln \bar{x} - \frac{1}{2}(\hat{\sigma}_y)^2 \tag{10}$$

$$\hat{\sigma}_y{}^2 = \ln (C_v{}^2 + 1) \tag{11}$$

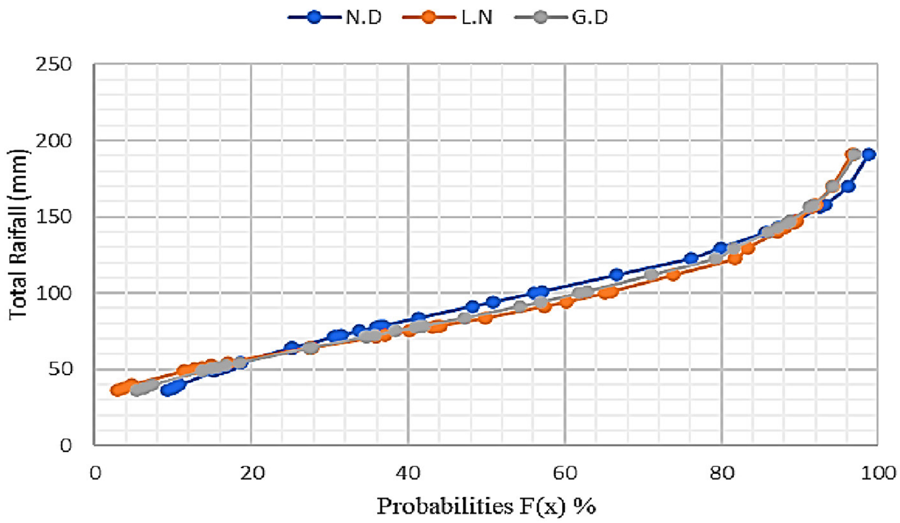

Figure 4. The probabilieties estimation using the moments method for (N.D.) at Najaf Station

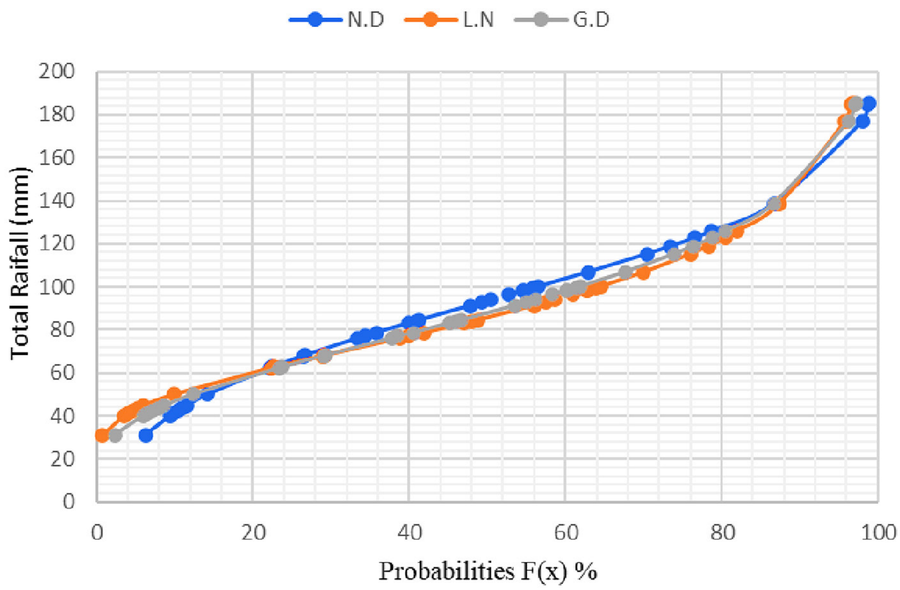

Figure 5. The probabilieties estimation using the moments method for (N.D.) at Karbala Station

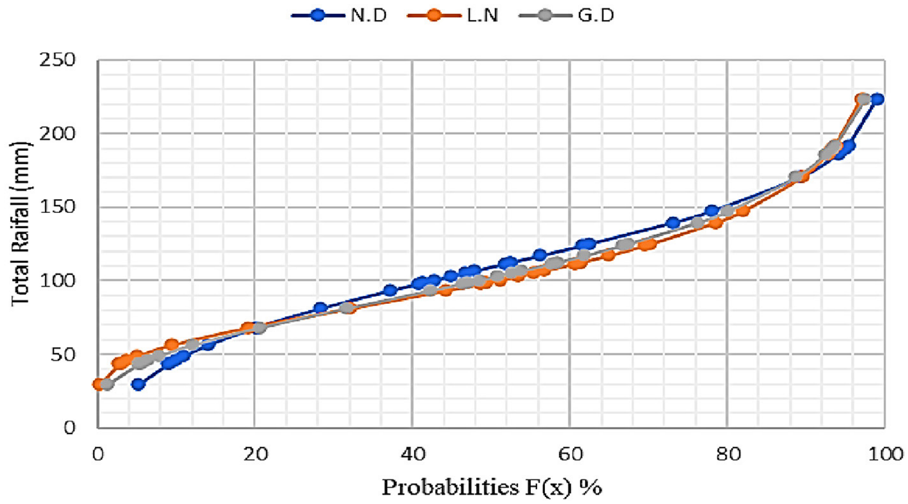

Figure 6. The probabilieties estimation using the moments method for (N.D.) at Diwana Station

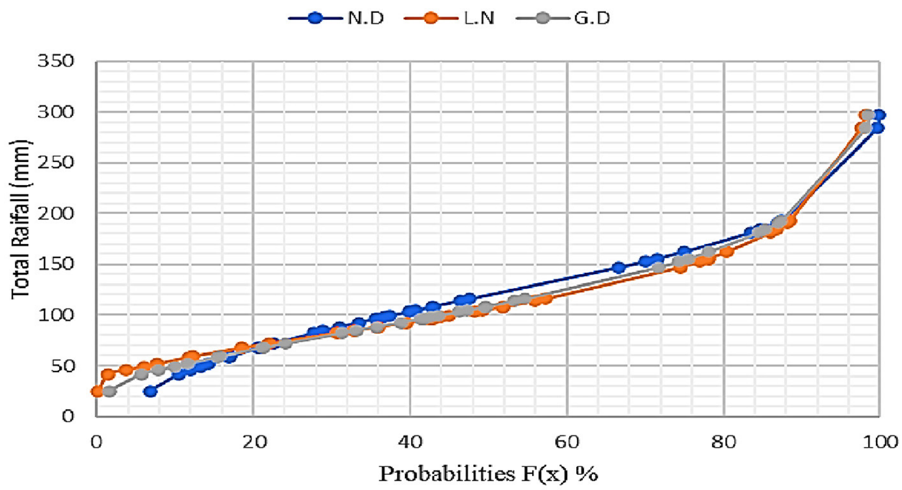

Figure 7. The probabilieties estimation using the moments method for (N.D.) at Baghdad Station

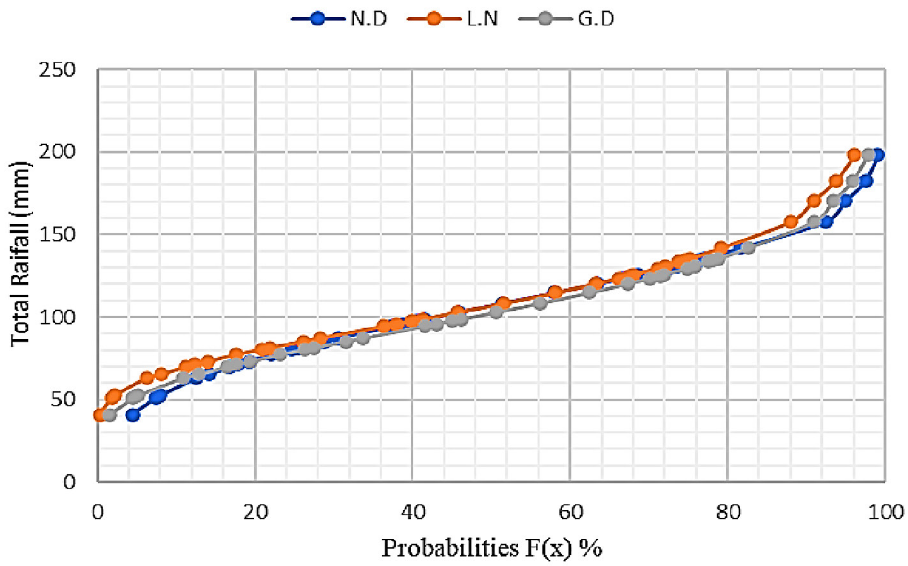

Figure 8. The probabilieties estimation using the moments method for (N.D.) at Hilla Sation

Table 1. Estimation of parameters for the Normal distribution

Stations	n	$\hat{\mu}$	$\hat{\sigma}$
Najaf	35	94.29	44.46
Kerbala	35	95.43	43.63
Diwaniya	35	110.54	52.08
Baghdad	35	121.56	65.48
Hilla	35	109.84	40.61

The two parameters $\hat{\mu}_y$ and $\hat{\sigma}_y$ of the Log-Normal distribution are estimated based on the coefficient of variation. Initially, the mean and standard deviation values are estimated using Equations 10 and 11, followed by the determination of the coefficient of variation C_v from Equation 12 (Chow, 1964).

$$C_v = \frac{s}{\bar{x}} \tag{12}$$

Figures 9 to 13 illustrate the probabilities associated with the log-normal distribution at the point of approach. The maximum likelihood technique is equivalent to the method of moments for parameter estimates. The moments and maximum probability estimates for all stations are presented in Table 2. The mean and variance for Gamma distribution can be estimated as follows [Bhakar et al., 2006]:

$$\mu = b \times F \tag{13}$$

$$\sigma^2 = b \times F \tag{14}$$

where: μ and σ are the mean and standard deviation, respectively, then substituted in Equations 13 and 14 and solved to find the value of $\hat{F}$ and $\hat{b}$.

The results for all the stations are given in Table 3. The MLE to the Gamma Distribution is listed in Table 4.

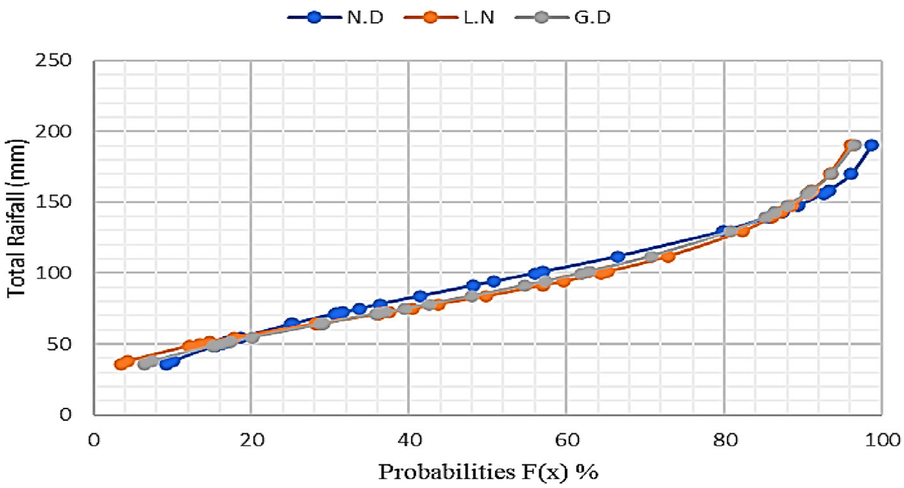

Figure 9. The probabilieties estimation using the MLE method fof (LN.D) at Najaf Station

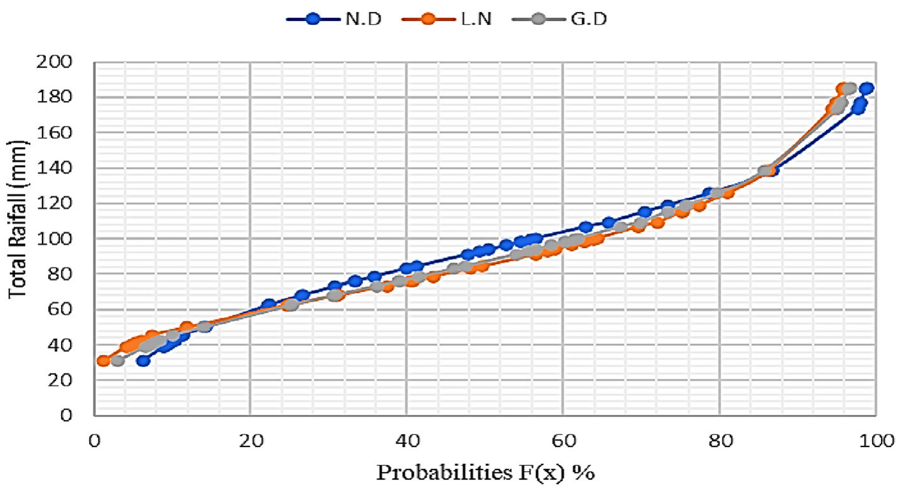

Figure 10. The probabilieties estimation using the MLE method fof (LN.D) at Karbala Station

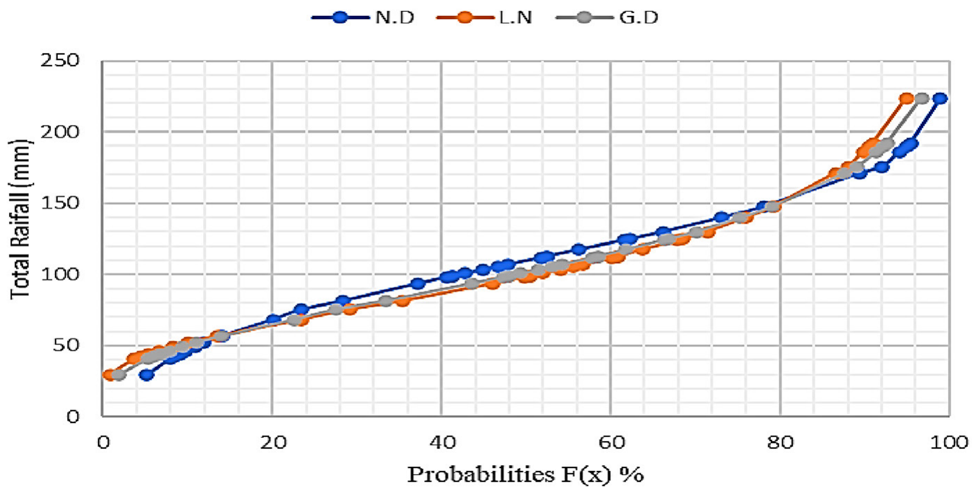

Figure 11. The probabilieties estimation using the MLE method fof (LN.D) at Diwana Station

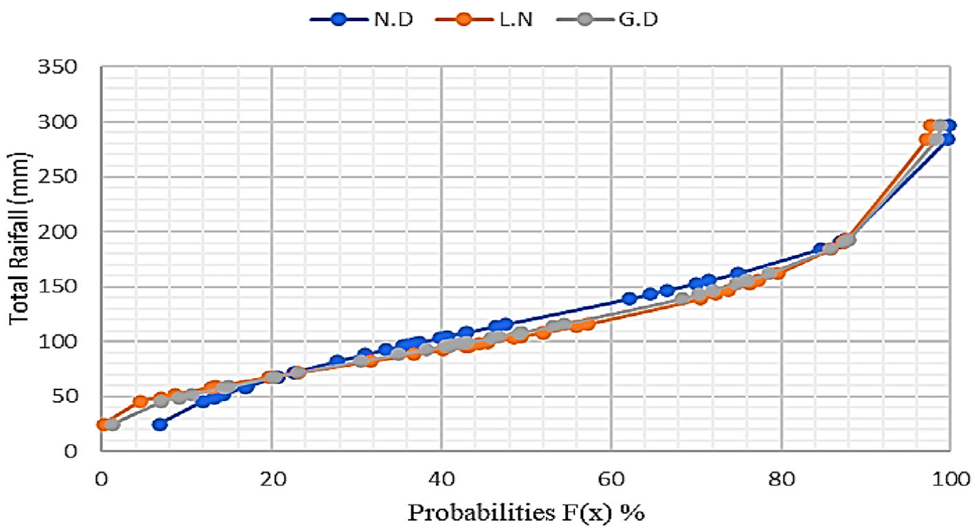

Figure 12. The probabilieties estimation using the MLE method fof (LN.D) at Baghdad Station

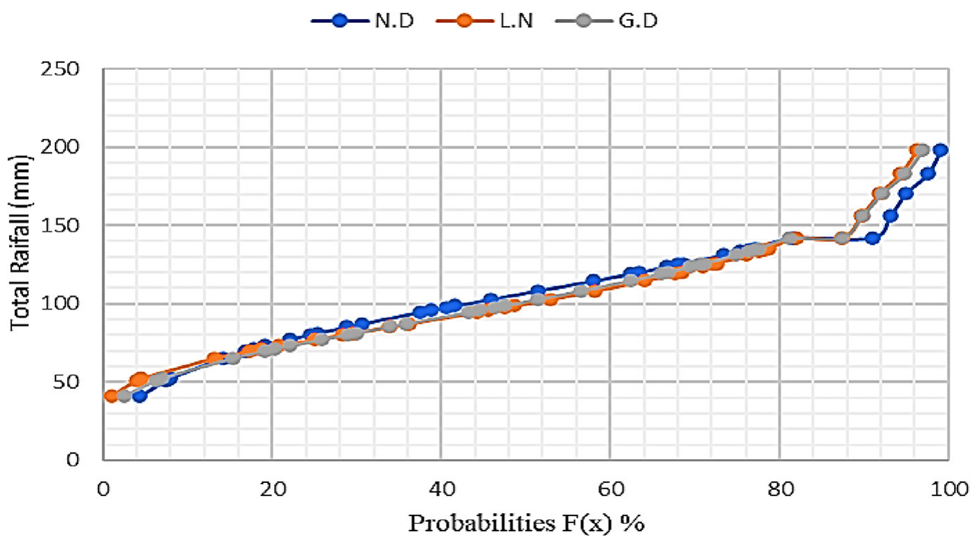

Figure 13. The probabilieties estimation using the MLE method fof (LN.D) at Hilla Sation

Table 2. Estimation of parameters for the log-Normal distribution (MLE)

Stations	n	$\hat{\mu}$	$\hat{\sigma}_y$
Najaf	35	4.43	0.47
Kerbala	35	4.44	0.45
Diwaniya	35	4.59	0.50
Baghdad	35	4.66	0.52
Hilla	35	4.60	0.38

THE GOODNESS OF FIT

The goodness of fit tests objectively determine if a theoretical distribution can effectively represent observable data. These tests can only reject reject models, not prove their correctness. An assumed model that deviates even little from an empirical distribution may be insufficient if the sample size is too small. The Anderson-Darling Index (ADI), Kolmogorov-Smirnov Index (KSI), and Chi-square Index (CHI) are examined in this study for diverse distributions.

Chi-square index

The amount of histogram classes into which the data is categorized influences the Chi-Squared statistic, and there is no definitive guideline about the appropriate quantity to employ (Huang et al., 2010). The relationship is utilized to generate the Chi-square test statistic.

$$\chi^2 = \frac{\sum_{i=1}^{k}(O_i - E_i)^2}{E_i} \qquad (15)$$

In the examined probability distribution, O_i denotes the observed frequency in the ith class interval, while E_i signifies the expected frequency. The product of the anticipated relative frequency and the total observations results in the predicted values. The results for all stations are summarised in Table 5, reflecting a confidence interval of 95%. The Najaf, Kerbala, Hilla, Baghdad, and Diwaniya stations had success in the chi-square index, except for the Baghdad station, which utilized the Normal distribution instead of the Gamma distribution through the maximum likelihood method, as well as

Table 3. Estimation of parameters for gamma distribution (moments method)

Stations	n	$\hat{\mu}$	$\hat{\sigma}$	$\hat{b}$	$\hat{F}$
Najaf	35	93.29	43.46	4.61	20.25
Kerbala	35	93.43	40.63	5.29	17.67
Diwaniya	35	109.54	49.08	4.98	22.00
Baghdad	35	119.56	63.48	3.55	33.71
Hilla	35	106.84	38.61	7.66	13.95

Table 4. Estimation of parameters for gamma distribution (max. Likelihood method)

Stations	n	$\hat{b}$	$\hat{F}$
Najaf	35	4.22	22.10
Kerbala	35	4.77	19.59
Diwaniya	35	4.37	25.10
Baghdad	35	3.77	31.71
Hilla	35	6.47	16.50

Table 5. Stations results for Chi-Sq. Test

Stations	v	Theo. Chi-Sq.	N. Dis.		L. N. Dis.		G. Dis.	
			Obs. Chi-Sq.		Obs. Chi-Sq.		Obs. Chi-Sq.	
			Mo.	Ma.L.	Mo.	Ma.L.	Mo.	Ma.L.
Najaf	5	12.33	7.343	7.7106	9.2803	8.5127	8.6571	7.8349
Kerbala	5	12.33	5.6198	5.6698	7.8556	6.6576	6.5388	5.7811
Diwaniya	5	12.33	9.2373	9.5473	18.7832	13.9346	13.3788	12.8582
Baghdad	5	12.33	16.4555	16.3255	10.165	9.4525	11.7517	11.5976
Hilla	5	12.33	8.8159	8.7459	9.8920	9.7757	8.8457	8.5128

moments and maximum likelihood procedures. Diwaniya station employs the Log-Normal and Gamma distributions utilizing moment and maximum likelihood approaches.

Kolmogorov-Smirnov index

A statistic that evaluates the difference between the proposed cumulative distribution function and the observed cumulative histogram is utilized to conduct the Kolmogorov-Smirnov (K-S) test's goodness of fit test (Reder et al., 2022). The distribution function that was observed of X at $x_{(1)}$, $x_{(2)}$, ..., denoted by $S(x_1)$, $S(x_2)$,…..,$S(x_i)$ are determined from the relation $S(x_i) = i/n$. The following equation is used to calculate the deviation D_2.

$$D_2 = S(x_i) - F(x_i) \quad (16)$$

The theoretical value displayed in statistical tables is compared with the highest absolute value of D_2, which may be calculated from the above equation. If D_2 is more than the theoretical value in the tables, reject hypothesis H; if not, accept hypothesis H. The values of the K-S index are obtained for all the stations. Figure 14 shows the method of calculating this index using the Normal distribution for the Baghdad station. Results are summarized for all the stations in Table 6 and represent a confidence interval equal to 95%. It is shown from this table that all the stations succeeded in the Kolmogorov-Smirnov index.

Anderson–Darling index

A statistical test, the Anderson-Darling normalcy test, is conducted to ascertain whether the examined data range corresponds to the theoretical range and supports or contradicts prior hypotheses. Compared to the K-S test, it gives the tails greater weight. The data is quantified by how closely it follows a specific distribution. Its benefit is that a more sensitive test may be conducted; nevertheless, each distribution's critical values must be determined. The smaller the statistic, the more suitable the data distribution is for the specified data collection and distribution, according to (Jäntschi and Bolboaca, 2018). The statistic for the AD test is computed using the relationship.

$$\mathrm{AD} = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)[\ln(F_i) + \ln(1 - F_{n-i+1})] \quad (17)$$

where: F is the specified distribution's cumulative distribution function, and n is the number of items in the sample.

The Anderson- Darling index values are obtained for all the stations used in this study. Results are summarized for all the stations in Table 7 and represent a confidence interval equal to 95%.

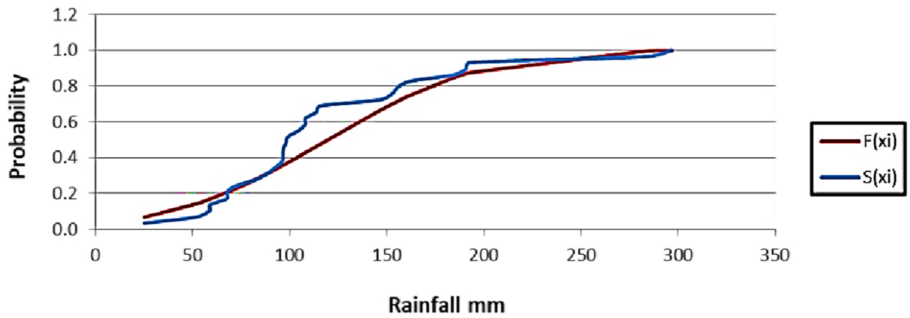

Figure 14. The deviation D_2 in the K-S test is calculated using the Normal distribiution (Station: Baghdad)

Table 6. Every station Kolmogorov-Smirnov index value, with a 95% confidence level

Stations	n	Theo. K-S	Obs. Dis.					
			N. Dis.		L. N. Dis.		Ga. Dis.	
			Mo.	Ma.L.	Mo.	Ma.L.	Mo.	Ma.L.
Najaf	35	0.246	0.15386	0.15386	0.08146	0.07968	0.10190	0.09207
Kerbala	35	0.246	0.15888	0.15888	0.09189	0.08009	0.10467	0.10451
Diwaniya	35	0.246	0.13501	0.13501	0.12213	0.08909	0.09492	0.08887
Baghdad	35	0.246	0.21327	0.21327	0.11676	0.11738	0.14316	0.14523
Hilla	35	0.246	0.09177	0.09177	0.11010	0.07683	0.07378	0.08618

It is shown in Table 7 that Najaf and Kerbala and Hilla and Baghdad and Diwaniya stations succeeded in the chi-square index except (for Baghdad and Najaf stations by utilizing the highest likelihood and Normal distribution by moments techniques. and Diwaniya station by using the Log Normal distributions by moments method).

INTENSITY DURATION FREQUENCY CURVES

This project aims to develop IDF curves for extreme rainfall data from five sources over 15, 30, and 60 minutes. When planning, developing, and building water resource projects, IDF curves are critical. Intensity-duration-frequency analysis requires data with different durations. After data collection, annual extremes are removed. Fitting the yearly extreme data to a probability distribution estimates rainfall. This study's annual extreme rainfall data are fitted with Normal and Log-Normal distributions to calculate the Gumbel extreme value.

Gumbel distribution

The Gumbel probability distribution can be expressed in the following way:

$$x_t = \mu + K_t \sigma \tag{18}$$

where: T = 1/(1-P_T) is the estimated chance that P_T will not be eliminated, which can be found using uniform random numbers created as follows (Sane et al., 2018).

$$K_t = \frac{-\sqrt{6}}{\pi} [0.5772 + ln\,(ln\,[\frac{T}{T-1}]\,)] \tag{19}$$

Figure 15 shows the relationship between intensity-duration-frequency for the stations used in this study.

Normal distribution

For the Normal distribution, the frequency factor whose value is dependent on the return period is given in the equations below (Wambua, 2019; Kourtis and Tsihrintzis, 2022):

$$K_t = w - \frac{2.515517 + 0.802853\,w + 0.010328 w^2}{1 + 1.432788\,w + 0.189269\,w^2 + 0.001308\,w^3} \tag{20}$$

where:

$$w = [ln(\frac{1}{p^2})]^{1/2} \quad [0 < p \leq 0.5] \tag{21}$$

where: P = probability of exceeding.

Equation 21 uses (1-p) instead of p when p exceeds (0.5). As an alternative, tables are used to calculate the frequency factor. This indicates that K_t value depends on the skew coefficient (Cs = 0) and the return period. The value of severe rainfall intensity is obtained by substituting the computed value of K_t into Equation 18 (Kourtis and Tsihrintzis, 2022). Figure 16 shows the relationship between intensity-duration-frequency for the stations used in this study.

Log-normal distribution

The same procedures as for the Normal distribution are used to calculate the frequency factor for the Log N. dis.; however, the logarithm of the data determines the value of the extreme intensity of rainfall. The value of extreme rainfall intensity is then obtained by using this value in Equation 22.

$$y_t = \overline{y} + K_t S_y \tag{22}$$

where: y_t shows the magnitude of the T-year event, while the logarithms of the data mean and standard deviation are represented by the terms $\overline{y}$ and Sy, respectively.

Table 7. Stations results for A-D test

Stations	N. Dis.		L. N. Dis.		Ga. Dis.	
	Obs. A-D.		Obs. A-D.		Obs. A-D.	
	A-D critical		A-D critical		A-D critical	
	0.704		0.795		0.752	
	Mo.	Ma.L.	Mo.	Ma.L.	Mo.	Ma.L.
Najaf	0.88017	0.88017	0.39483	0.32960	0.40410	0.37072
Kerbala	0.65471	0.65471	0.46292	0.42716	0.34311	0.36428
Diwaniya	0.57439	0.57439	0.99880	0.74637	0.56743	0.53171
Baghdad	1.21970	1.21970	0.35272	0.36978	0.50132	0.47645
Hilla	0.27105	0.27105	0.62566	0.27404	0.19938	0.21900

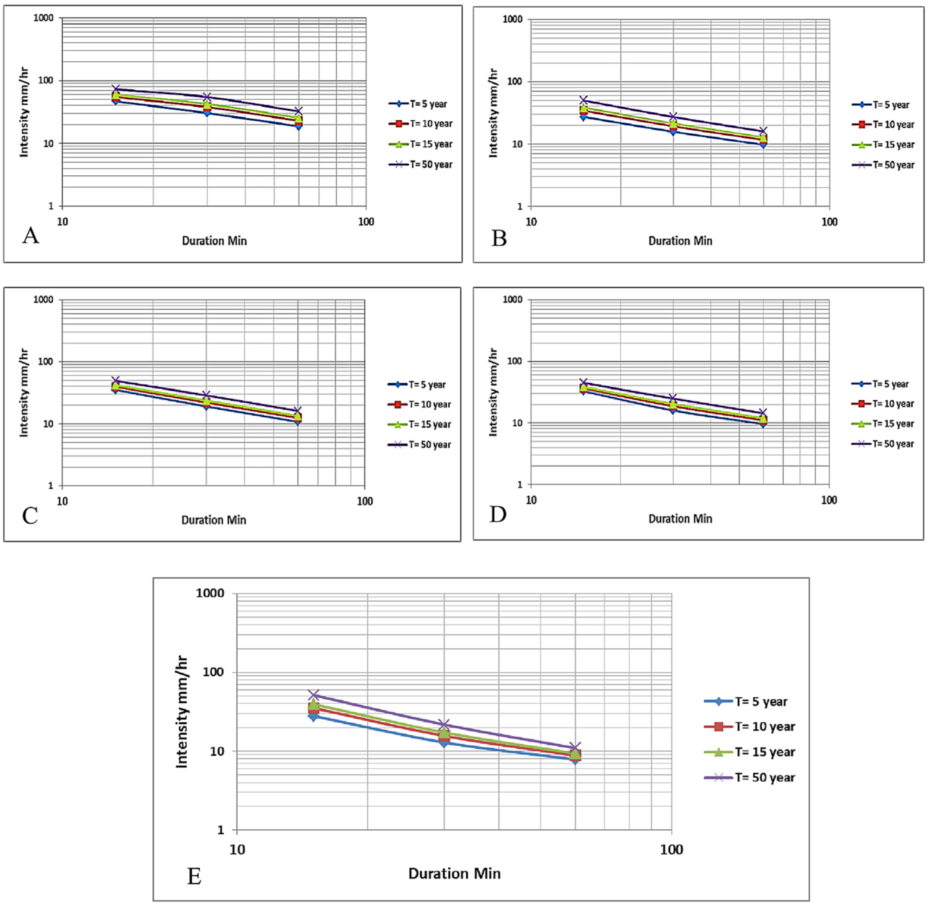

Figure 15. IDF curve by the Gumbel distribiution (A) Najaf, (B) Diwanyia, (C) Kerbala, (D) Hilla, (D) Baghdad

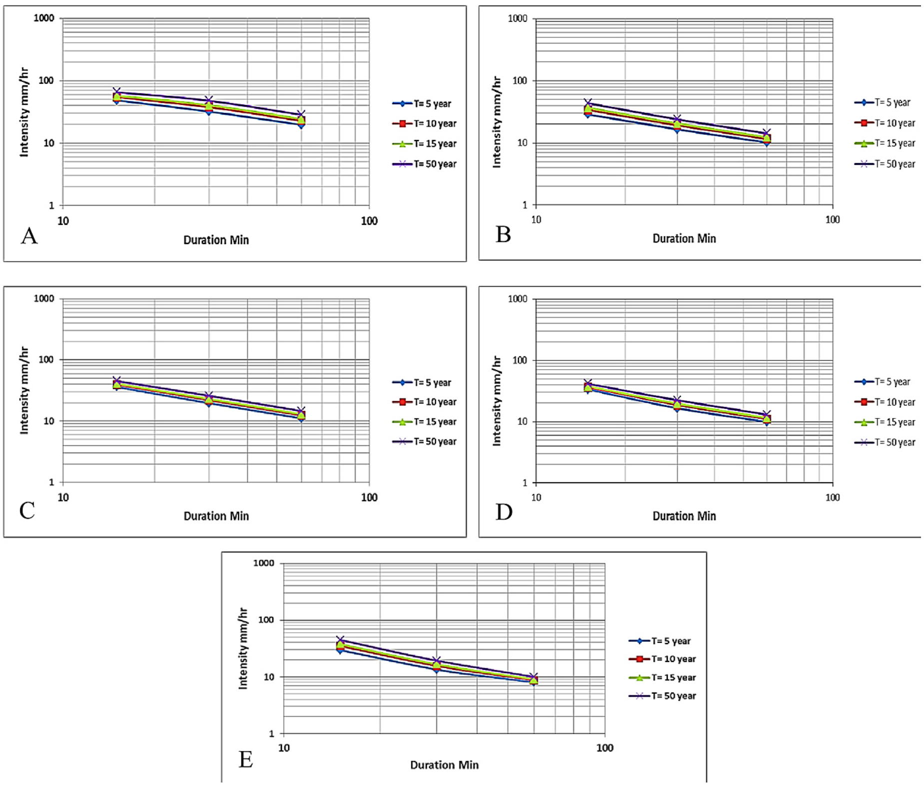

Figure 16. IDF curve by the Noraml distribiution (A) Najaf, (B) Diwanyia, (C) Kerbala, (D) Hilla, (D) Baghdad

Figure 17 shows the relationship between intensity-duration-frequency for the stations used in this study.

COMPARISON OF THE IDF CURVE

The study compares the tests conducted at different stations over the years. Specifically, it examines the comparison between the same year and inspection at five stations, represented by the Standard error in Figures 18 to 20. The comparison also includes tests conducted at 5, 10, 15, and 50 years.

IDF CURVE FORECASTING

Linear equations were discovered at the Hilla Station by examining the surrounding locations (Kerbala, Baghdad, Najaf, Diwaniya). Anova has derived the following linear equation:

$$y = a + bx \tag{23}$$

Table 8 shows the coefficients of the linear equation of neighbouring areas. The equations of the Hilla region were found through the total adjacent areas and their impact on the Hilla region and the linear equation below:

$$y = A + Bx_1 + Cx_2 + Dx_3 + Ex_4 \tag{24}$$

where: (A) is constant, and (B, C, D, E) is a Coefficient factor for (Kerbala, Baghdad, Najaf, and Diwaniya) respectively, and (x_1,x_2,x_3, and x_4) intensity for (Kerbala, Baghdad, Najaf, Diwaniya) respectively, with the coefficient of determination (R^2) ranging from 83.2 to 94.7.

$$y = -0.13 + 0.93x_2 + 0.09x_4 \text{ for 5 years} \tag{25}$$

$$y = 0.56 + 0.77x_2 + 0.16x_3 \text{ for 10 years} \tag{26}$$

$$y = 0.79 + 0.71x_2 + 0.18x_3 \text{ for 15 years} \tag{27}$$

$$y = 1.26 + 0.6x_2 + 0.2x_3 \text{ for 50 years} \tag{28}$$

The error percentage used for the 20 Equations between the theoretical IDF_{the} and the actual IDF_{act} (calculated from the equations) as mentioned in the formula below (Reder et al., 2022):

$$Error\% = \left(\frac{(Q_r)_{act} - (Q_r)_{the}}{(Q_r)_{act}}\right) \times 100 \tag{29}$$

Table 9 shows the error percentage for the all stations.

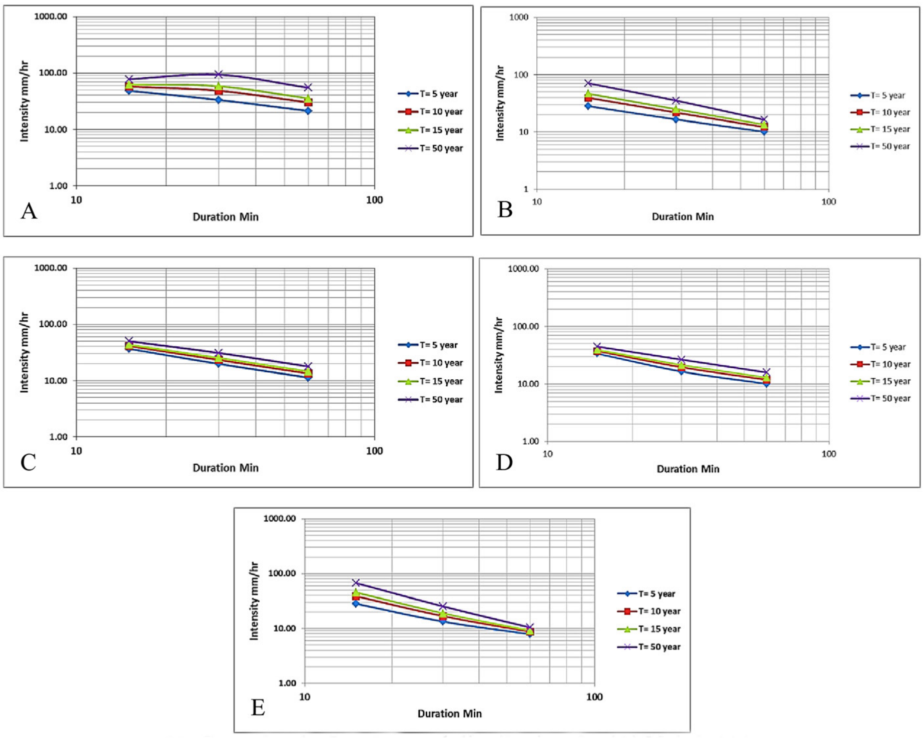

Figure 17. IDF curve by the Log Noraml distribiution (A) Najaf, (B) Diwanyia, (C) Kerbala, (D) Hilla, (D) Baghdad

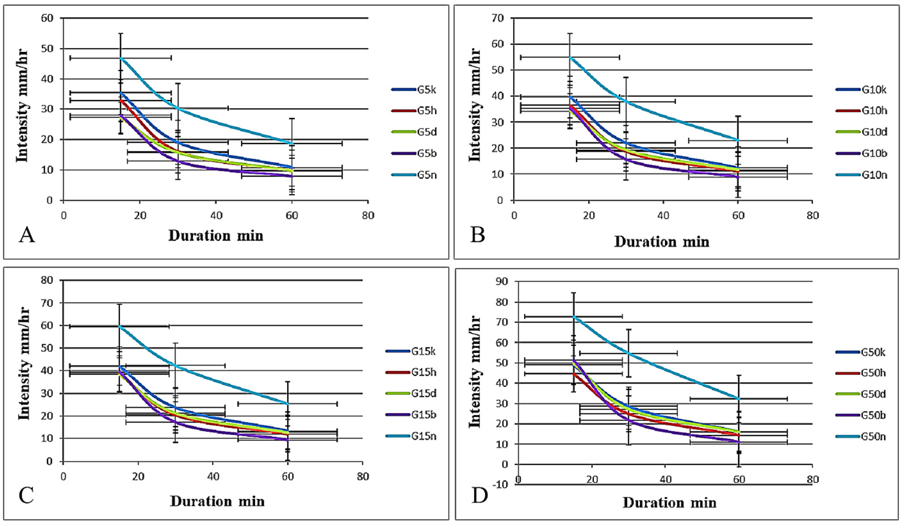

Figure 18. IDF curves by the Gumbel distribiution (A) 5, (B) 10, (C) 15, (D) 50 years

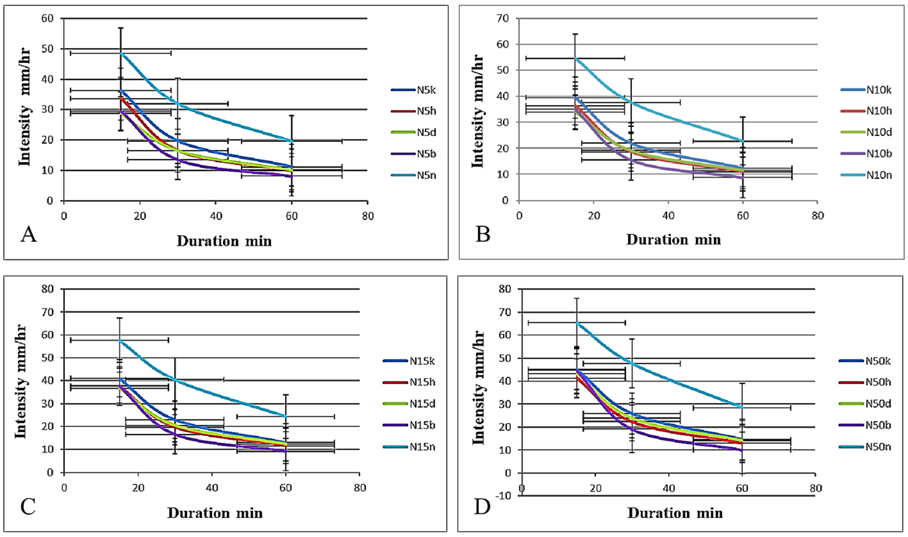

Figure 19. IDF curves by the Noraml distribiution (A) 5, (B) 10, (C) 15, (D) 50 years

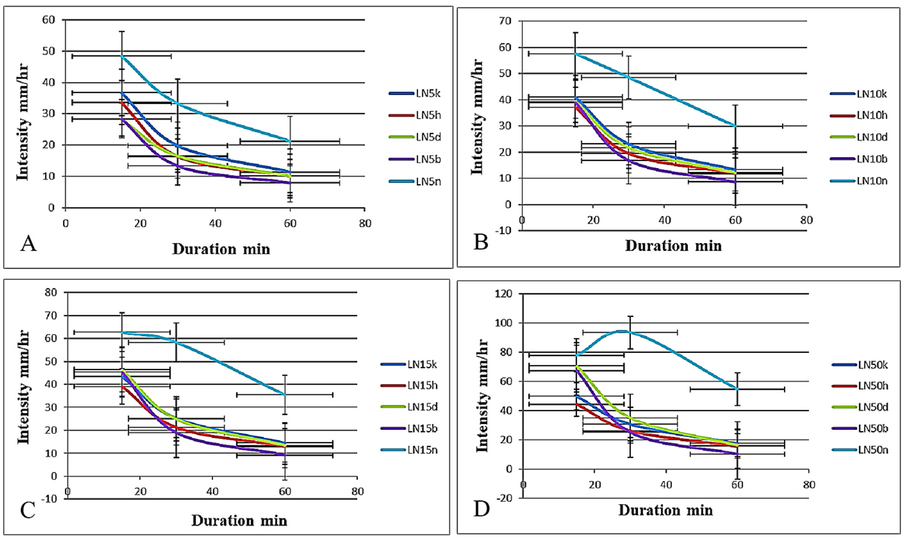

Figure 20. IDF curves by the Log Noraml distribiution (A) 5, (B) 10, (C) 15, (D) 50 years

Table 8. Coefficient of IDF curve for Hilla station

Station	5 years		10 years		15 years		50 years	
	a	b	a	b	a	b	a	b
Baghdad	1.37	1.10	3.14	0.96	3.94	0.90	5.83	0.80
Diwaniya	-3.85	1.30	-2.60	1.14	-2.05	1.08	-0.73	0.97
Kerbala	-1.29	0.95	-1.21	0.95	-1.16	0.94	-1.06	0.93
Najaf	-7.7	0.83	-8.72	0.80	-9.19	0.79	-10.27	0.76

Table 9. % Error for all stations

Station	5 years	10 years	15 years	50 years
	% Error	% Error	% Error	% Error
Baghdad	-13.1	-1.9	-1.0	-1.9
Diwaniya	6.9	1.0	1.0	4.2
Kerbala	8.1	7.5	8.2	8.7
Najaf	30.1	32.1	33.4	42.9
4 Stas	15.1	13.3	17.9	16.2

RESULTS AND DISCUSSION

This research employs three probability distributions: Gamma, Log-Normal, and Normal. The distributions mentioned above pertain to the rainfall depth statistics for five designated regions in Iraq: Hilla, Baghdad, Najaf, Kerbala, and Diwaniya. The maximum likelihood and method of moments techniques are employed to estimate the parameters of the chosen distributions. The Anderson-Darling, Chi-square, and Kolmogorov-Smirnov tests evaluate the adequacy of theoretical distributions to the data. The subsequent results pertain to the Chi-square index. The normal distribution is suitable for the max likelihood and moments methods applied to the Diwaniya, Hilla, Kerbala, and Najaf stations, except the Baghdad station. All stations, except for Diwaniya station, exhibit Log Normal distributions deemed acceptable when calculated via either the maximum likelihood or moments methods. Using the moments approach, the Gamma distribution is appropriate for the stations in Najaf, Kerbala, Hilla, and Baghdad. Nevertheless, the maximum likelihood technique is suboptimal for the Baghdad station. Furthermore, the Diwaniya station is unsuitable for the Gamma distribution. The Normal distribution is suitable for all stations when using the maximum likelihood and moments methods to get the Kolmogrov-Smirnov index.

The Log Normal distribution, utilizing maximum likelihood and moments methods, is suitable for the stations of Najaf, Kerbala, Hilla, Diwaniya, and Baghdad according to the Kolmogorov-Smirnov statistic. The Gamma distribution is suitable for the maximum likelihood and moments methods applied to the Diwaniya, Hilla, Baghdad, Kerbala, and Najaf stations concerning the Kolmogorov-Smirnov index. Except for the Diwaniya station, the Log-Normal (AD) regards the Moments method as disadvantageous. All stations have an appropriate distribution when employing the two estimate methods, specifically maximum likelihood and moments.

The five stations were compared by drawing and using Gumbel, Normal, and Log-Normal distributions. The Normal distribution was the best from the minor variance for the years (5, 10, 15, and 50). The equations of the Hilla region have been found. Through these equations, the extracted IDF curve Hilla from the surrounding regions and the comparison of the results have been in good agreement and by proportion (R^2) ranging from 83.2 to 94.7. The error percentage was calculated for the equations, representing that the best forecasting formula for the Hilla region could start from Diwaniya, Baghdad, Kerbala and Najaf, depending on the percentage. Utilizing the Gumbel, Normal, and Log-Normal distributions to generate IDF curves for five stations in Iraq (Hilla, Baghdad, Najaf, Kerbala, and Diwaniya) reveals that the Normal distribution consistently surpasses the Gumbel and Log-Normal distributions across all time intervals (15, 30, and 60 minutes). This superiority is attributed to the reduced deviation between the Normal distribution IDF curve and those of the Gumbel and Log-Normal distributions.

REFERENCES

1. Akpen, G. D., Aho, M. I., & Musa, A. A. (2019). Rainfall intensity-duration-frequency models for Lokoja Metropolis, Nigeria. *Global Journal of Pure and Applied Sciences, 25*(1), 81–90. https://doi.org/10.4314/gjpas.v25i1.11
2. Alam, F., Salam, M., Khalil, N. A., khan, O., & Khan, M. (2021). Rainfall trend analysis and weather forecast accuracy in selected parts of Khyber Pakhtunkhwa, Pakistan. *SN Applied Sciences, 3*, 1-14. https://doi.org/10.1007/s42452-021-04457-z
3. AL-Dulaimi, M. H. A., Alfatlawi, T. J. M., & Mohammed, W. A. (2020). Developing (IDF) curves models for Babylon City and alluvial fertile zone, Iraq. *Solid State Technology*, 2585–2598. https://www.solidstatetechnology.us/index.php/JSST/issue/view/50
4. Bhakar, S. R., Bansal, A. K., Chhajed, N., & Purohit, R. C. (2006). Frequency analysis of consecutive days maximum rainfall at Banswara, Rajasthan, India. *ARPN Journal of Engineering and Applied Sciences, 1*(3), 64–67.
5. Bolstad, B. M. (1998). *Comparing some iterative methods of parameter estimation for censored gamma data*. The University of Waikato, 55.
6. Campos, J. N. B., Studart, T. M. D. C., Souza Filho, F. D. A. D., & Porto, V. C. (2020). On the rainfall intensity–duration–frequency curves, partial-area effect and the rational method: Theory and the engineering practice. *Water, 12*(10), 2730. https://doi.org/10.3390/w12102730
7. Chow, V. T. (1964). Statistical and probability analysis of hydrologic data. *Handbook of applied hydrology*, 8–1.
8. Cooley, D., Nychka, D., & Naveau, P. (2007). Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association, 102*(479), 824–840. https://doi.org/10.1198/016214506000000780
9. El Hannoun, W., Boukili Makhoukhi, A., Zoglat, A., & El Adlouni, S. E. (2023). Intensity–Duration–Frequency Curves for Dependent Datasets. *Water, 15*(14), 2641. https://doi.org/10.3390/w15142641
10. Fadhel, S., Rico-Ramirez, M. A., & Han, D. (2017). Uncertainty of intensity–duration–frequency (IDF) curves due to varied climate baseline periods. *Journal of hydrology, 547*, 600–612. https://doi.org/10.1016/j.jhydrol.2017.02.013
11. Ghahraman, B., & Hosseini, S. M. (2005). *A new investigationon on the performance of rainfall IDF models*.
12. Gu, X., Ye, L., Xin, Q., Zhang, C., Zeng, F., Nerantzaki, S. D., & Papalexiou, S. M. (2022). Extreme precipitation in China: A review on statistical methods and applications. *Advances in Water Resources, 163*, 104144. https://doi.org/10.1016/j.advwatres.2022.104144
13. Heidari, H., Arabi, M., Ghanbari, M., & Warziniack, T. (2020). A probabilistic approach for characterization of sub-annual socioeconomic drought intensity-duration-frequency (IDF) relationships in a changing environment. *Water, 12*(6), 1522. https://doi.org/10.3390/w12061522
14. Huang, Q., Chen, Y., Xu, S., Liu, Y., & Li, X. (2010, August). Scaling models of a rainfall intensity-duration-frequency relationship. In *2010 Sixth International Conference on Natural Computation 7*, 3415–3419). IEEE.
15. Kourtis, I. M., & Tsihrintzis, V. A. (2022). Update of intensity-duration-frequency (IDF) curves under climate change: a review. *Water Supply, 22*(5), 4951–4974. https://doi.org/10.2166/ws.2022.152
16. Jäntschi, L., & Bolboacă, S. D. (2018). Computation of probability associated with Anderson–Darling statistic. *Mathematics, 6*(6), 88. https://doi.org/10.3390/math6060088
17. Kareem, D. A., M Amen, A. R., Mustafa, A., Yüce, M. I., & Szydłowski, M. (2022). Comparative analysis of developed rainfall intensity–duration–frequency curves for Erbil with other Iraqi Urban areas. *Water, 14*(3), 419. https://doi.org/10.3390/w14030419
18. Miller, J., Taylor, C., Guichard, F., Peyrillé, P., Vischel, T., Fowe, T.,... & Parker, D. J. (2022). High-impact weather and urban flooding in the West African Sahel–A multidisciplinary case study of the 2009 event in Ouagadougou. *Weather and Climate Extremes, 36*, 100462. https://doi.org/10.1016/j.wace.2022.100462
19. Noor, M., Ismail, T., Shahid, S., Asaduzzaman, M., & Dewan, A. (2021). Evaluating intensity-duration-frequency (IDF) curves of satellite-based precipitation datasets in Peninsular Malaysia. *Atmospheric Research, 248*, 105203. https://doi.org/10.1016/j.atmosres.2020.105203
20. Pecho, J., Faško, P., Lapin, M., & Gaál, L. (2009, April). Analysis of rainfall intensity-duration-frequency relationships in Slovakia (estimation of extreme rainfall return periods). In *EGU General Assembly Conference Abstracts*, 11240.
21. Reder, A., Raffa, M., Padulano, R., Rianna, G., & Mercogliano, P. (2022). Characterizing extreme values of precipitation at very high resolution: An experiment over twenty European cities. *Weather and Climate Extremes, 35*, 100407. https://doi.org/10.1016/j.wace.2022.100407
22. Sane, Y., Panthou, G., Bodian, A., Vischel, T., Lebel, T., Dacosta, H.,... & Diop Kane, M. (2018). Intensity–duration–frequency (IDF) rainfall curves in Senegal. *Natural Hazards and Earth System Sciences,*

18(7), 1849–1866.

23. Schlef, K. E., Kunkel, K. E., Brown, C., Demissie, Y., Lettenmaier, D. P., Wagner, A.,... & Yan, E. (2023). Incorporating non-stationarity from climate change into rainfall frequency and intensity-duration-frequency (IDF) curves. *Journal of Hydrology, 616*, 128757. https://doi.org/10.1016/j.jhydrol.2022.128757
24. Shamkhi, M. S., Azeez, M. K., & Obeid, Z. H. (2022). Deriving rainfall intensity–duration–frequency (IDF) curves and testing the best distribution using EasyFit software 5.5 for Kut city, Iraq. *Open Engineering, 12*(1), 834–843. https://doi.org/10.1515/eng-2022-0330
25. Silva, D. F., Simonovic, S. P., Schardong, A., & Goldenfum, J. A. (2021). Assessment of non-stationary IDF curves under a changing climate: Case study of different climatic zones in Canada. *Journal of Hydrology: Regional Studies, 36*, 100870.
26. Takara, K., & Takasao, T. (1988). *Criteria for evaluating probability distribution models in hydrologic frequency analysis*. Doboku Gakkai Ronbunshu, 393, 151–160.
27. Wambua, R. M. (2019). *Estimating rainfall intensity-duration-frequency (IDF) curves for a tropical river basin*.
28. Yüksek, Ö., Anılan, T., Saka, F., & Orgun, E. (2022). Rainfall intensity-duration-frequency analysis in Turkey, with the emphasis of Eastern Black Sea basin. *Teknik Dergi, 33*(4), 12087–12103. https://doi.org/10.18400/tekderg.727085