Academic Journal of Nawroz University (AJNU), Vol.13, No.1, 2024

https://doi.org/10.25007/ajnu.v13n1a1589

Developing Rainfall Intensity- Frequency - Duration Curves for Three Selected Sites in the Kurdistan Region, Iraq

Kareen Shajee and Evan Hajani

[1, 2] Water Resources Department, University of Duhok, Kurdistan Region, Iraq.

ABSTRACT: The relationship of rainfall intensity frequency duration (IFD) is one of the key tools used considerably in water resources engineering, whether for planning, designing, managing, and operating projects for water resources or flood control and management projects. The purpose of this study is to develop IFD curves at a site location for three selected governorates in Kurdistan Region, Iraq. The current study covers rainfall data recorded over 30 years between 1991-2020 for three rainfall stations located within three governorates (i.e., Duhok, Erbil, and Sulaymaniyah) in the Kurdistan Region, Iraq. The stationary IFD curves have been derived at the site location using the Generalized Extreme Value (GEV), Gumbel, and Log Pearson Type 3 (LPT3) distributions for the three selected stations. The fitness of the three selected distributions to the observed rainfall data has been tested by applying three goodness-of-fit tests (i.e., Kolmogorov-Smirnov (KS), Anderson-Darling (AD), and Chi-Square tests). The results indicated that the three selected distributions fit the rainfall data at the three significance levels (10%, 5%, and 1%) for all three stations. Overall, it has been discovered that at-site IFD curve data derived from the GEV distribution are generally higher than curves derived from the LPT3 and Gumbel distributions.
KEYWORDS: Rainfall; Gumbel; GEV; LPT3; IFD; Kurdistan.

1. Introduction

The IFD relationship is a mathematical relationship between rainfall intensity, duration, and return period. The establishment of such a relationship was done as early as 1931 and 1932 (Sherman, 1931; Bernard, 1932). The rainfall-IFD relationship is commonly required for planning and designing various water resource projects (El-Sayed, 2011). The engineering application of rainfall intensity is essential in the estimation of design discharge for flood control structures. This relationship is determined through a statistical analysis of rainfall data from meteorological stations. Quantification of rainfall was generally done using isopluvial maps and IFD curves (Chow *et al*., 1988). There has been considerable attention and research on the IFD relationship, such as Zainudini *et al.* (2011) collected the AMR values for a short duration from 12 different stations in the Sistan and Balochistan provinces of Iran and fitted them to the statistical Gumbel distribution to establish the IFD curves. The results for shorter durations appeared to be more acceptable than those for longer durations, where the IFD curves bent and flattened out.

Al-Anazil and El-Sebaie (2013) derived IFD relationships for Abha city in the Kingdom of Saudi Arabia (KSA) of 6 frequency periods of 2 to 100 years and 8 different durations of 10 min to 12 hr. Three probability distributions were used, the Gumbel, Lognormal, and LPT3 distributions. It has been revealed that there were tiny differences between the results gained from the three methods. Sulaiman (2015) estimated the IFD curves with various durations of 5 min to 1 day for Duhok city in Iraq. For different durations from 1976 to 2013, the AMR data was fitted using Gumbel and Weibull distributions. The obtained intensity values by the two methods were very close for the high duration at all periods, while for the low and medium durations there was a small difference between the intensity values gained by the two methods. Dakheel (2017) developed the IFD curves with different rainfall durations of 10 min to 1 day and 6 specified return periods of 2 to 100 years for Nasiriyah city, Iraq. The daily rainfall data for 36 years from 1980 to 2015 was converted into shorter duration by applying the empirical formula and then fitted into the Gumbel and LPT3 distributions to derive the curves. The obtained results showed that the results of the LPT3 distribution were more acceptable than those of the Gumbel distribution. Hajani and Rahman (2018) generated IFD curves from the two probability distributions, GEV and LPT3. Ten pluviography stations from eastern New South Wales (NSW) in Australia were selected for the study. The derived IFD curves were compared with the latest regional IFD curves in Australia. It has been found that regional IFD curves derived in 2003 were generally higher than the at-site IFD curves derived from the current study.

Suchithra and Agarwa (2020) generated IFD curves and developed formulas that could approximate the design rainfall intensity for the Krishna district in India. The Log-Normal, Normal, and Gumbel methods were applied to the daily rainfall data for the period from 1981 to 2018. The intensity values obtained from the Log-Normal and Normal distribution methods were close to each other, while the highest values were found in the Gumbel distribution method. There have been numerous studies on estimating extreme rainfall, as was already mentioned. However, there is a lack of studies comparing the IFD curves generated using various distributions and methodologies. The IFD curves must be developed to offer a tool for designing rainfall events that enable the calculation of peak flow required to design various hydraulic structures (such as storm sewers, culverts, and drainage systems), as well as for assessing and forecasting flood hazards and designing structures for flood protection. Hence, this study is devoted to comparing IFD curves obtained from different distributions. Consequently, the objective of the present study is to develop the stationary IFD curves for three stations based on daily rainfall data at a site located in the Kurdistan Region of Iraq. For the three selected stations, the stationary IFD curves were derived by using the three most commonly used probability distributions (i.e., Generalized Extreme Value, Gumbel, and Log Pearson Type 3). To the best of the author's knowledge, this is the first attempt to derive the IFD curves at the site location for the Kurdistan Region.

2. Study Area and Data

This study covers three rainfall stations within three main governorates (i.e., Duhok, Erbil, and Sulaymaniya) located in the Kurdistan Region of Iraq (see Figure 1). The current study covers the annual maximum rainfall (AMR) data series for 30 years (period of 1991-2020) at three rainfall stations in the Kurdistan Region (see Figure 1). The historical records of the AMR data, including a 24-hr rainfall dataset, have been provided by the directorate of Meteorology and Seismology in Duhok and the general directorate of Meteorology and Seismology in Erbil in the Kurdistan Region-Iraq are shown in Figure 2. The rainfall data were analyzed to determine the maximum amount of rainfall received in a year (i.e., 365 days period). The gaps in the rainfall data of each station were filled by the regression analysis method, where the gaps in the rainfall data of a particular station were filled by regression analysis with a nearby station that had no data gap (Hajani, 2020; Hajani and Klari, 2022).

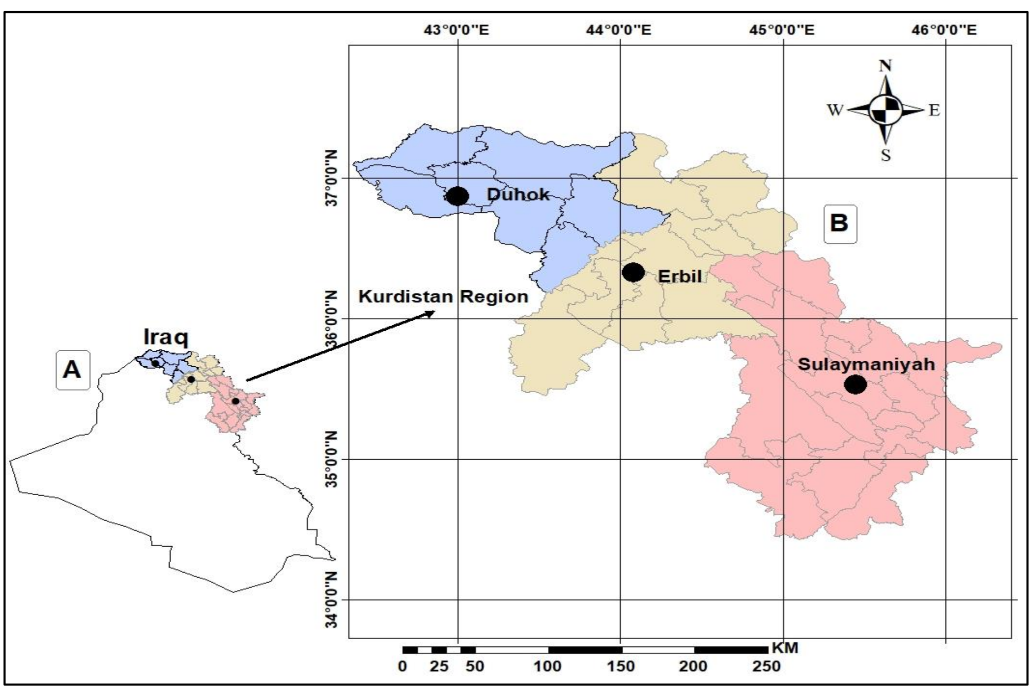


Figure 1: The location map of the study area. (A) Map of Iraq, (B) Location of the three selected rainfall stations.

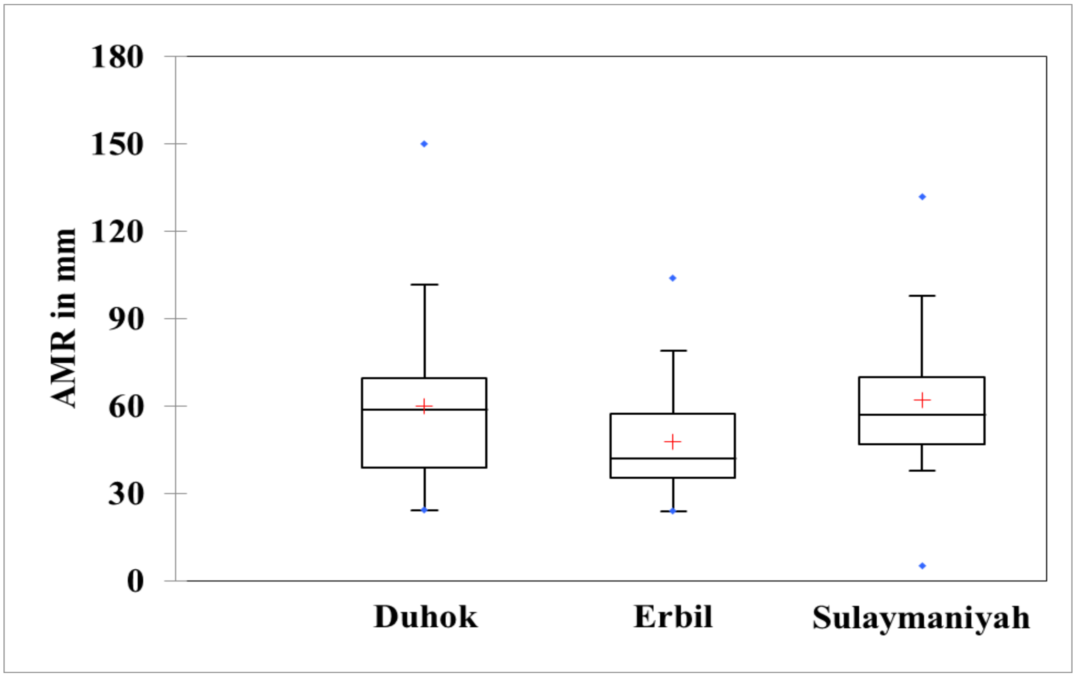


Figure 2: Boxplot of AMR for three stations in the Kurdistan region during1991 to 2020.

3. Methods

In this study, different statistical techniques have been applied to the rainfall data series of the three rainfall stations in the Kurdistan Region of Iraq. The AMR data at each station separately have been used in an empirical equation to estimate short-duration rainfall events (5 min to 3 day). In addition, three goodness-of-fit tests will be used to assess how well the selected statistical distributions fit the rainfall data, and the distributions will be used to drive the intensity quantiles of stationary IFD curves. A brief description of each of these statistical techniques used is presented in the sections below; see Appendix A and Appendix B for the list of Abbreviations and symbols.

3.1. Estimation of Short Duration Rainfall

The rainfall data consists of the maximum daily rainfall values from 1991 to 2020. Sub-daily rainfall values can be obtained from the daily AMR data of 24 hr using the Indian Meteorological Department (IMD) empirical reduction formula (Ramaseshan,1996) which is:

$$R_t = R_{24}\left(\frac{t}{24}\right)^{1/3} \qquad (1)$$

Where R_t is the required rainfall depth in *mm* at *t-hr* duration, R_{24} is the daily rainfall in *mm* and *t* is the duration of rainfall for which the rainfall depth is required in *hr*.

3.2. Distribution Methods for Developing Stationary IFD Curves

The first step in the construction of IFD curves is fitting some theoretical frequency distribution to the extreme rainfall amounts for several fixed durations. A logical step to proceed is to describe the change of the parameters of the distribution with duration by a functional relation. From the fitted relationships, the rainfall intensity for any duration and return period can be derived. In this research, AMR values for all the available durations which were estimated by Equation 1 (i.e., 5 min, 10 min, 15 min, 30 min,1 hr, 3 hr, 5 hr, 8 hr, 12 hr, 1 day and 3 day) have been statistically analyzed using three different distributions, namely: Generalized Extreme Value (GEV) distribution, Gumbel distribution, and Log Pearson Type 3 (LPT3) distribution.

3.2.1. Generalized Extreme Value (GEV) Distribution

GEV distribution is a continuous probability distribution developed within the extreme value theory to combine the Fréchet (1927), Weibull (1951), and Gumbel (1958) families of distributions. The GEV distribution has three parameters location (ξ), scale (α), and shape (κ), and it is used as an approximation to model the maxima of long sequences of random variables. The GEV distribution is equivalent to a Gumbel, Fréchet, or Weibull distribution depending on whether $\kappa = 0$, $\kappa > 0$, or $\kappa < 0$, respectively. The GEV distribution has the cumulative density function (CDF) and probability density function (PDF) (Hosking and Wallis, 1997) as shown below in Equation 2 and Equation 3 respectively :

$$F(x, \xi , \alpha, \kappa) = exp\left\{-\left(1 - \frac{\kappa(x-\xi)}{\alpha}\right)^{1/\kappa}\right\} \quad (2)$$

$$f(x, \xi , \alpha, \kappa) = \alpha^{-1} \left(1 + \frac{\kappa(x-\xi)}{\alpha}\right)^{-1/\kappa -1} \times exp\left\{-\left(1 + \frac{\kappa(x-\xi)}{\alpha}\right)^{-1/\kappa}\right\} \quad (3)$$

3.2.2. Gumbel Distribution

The Gumbel distribution also referred to as the Extreme Value type I (EV1), Gumbel distribution only uses two parameters, location (ξ), and scale (α). The cumulative density function (CDF) and probability density function (PDF) as defined in (Hosking and Wallis, 1997) are:

$$F(x, \xi, \alpha) = exp\left[-exp\left(-\frac{x-\xi}{\alpha}\right)\right] \quad (4)$$

$$f(x, \xi, \alpha) = \alpha^{-1}\, exp\left[\left(-\frac{x-\xi}{\alpha}\right) - exp\left(-\frac{x-\xi}{\alpha}\right)\right] \quad (5)$$

3.2.3. Log Pearson Type 3 (LPT3) Distribution

The LPT3 distribution is a three-parameter distribution (similar to GEV), it uses location (ξ), scale (a), and shape (κ) parameters (Millington et al., 2011). The cumulative density function (CDF) and probability density function (PDF) as defined in (Hosking and Wallis, 1997) are shown below:

If $\gamma \neq 0$, let $\kappa = 4 / \gamma^2$, $\xi = \mu - 2\,\sigma/\gamma$ and $\alpha = 0.5\,\sigma|\gamma|$

If $\gamma > 0$, then

$$F(x, \xi , \alpha, \kappa) = \frac{\bar{\Gamma}\left(\kappa,\ \frac{x-\xi}{\alpha}\right)}{\Gamma(\kappa)} \quad (6)$$

$$f(x, \xi , \alpha, \kappa) = \frac{(x-\xi)^{\kappa-1}\ e^{-(x-\xi)/\alpha}}{\alpha^{\kappa}\ \Gamma(\kappa)} \quad (7)$$

If $\gamma < 0$, then

$$F(x, \xi , \alpha, \kappa) = 1 - \frac{\bar{\Gamma}\left(\kappa,\ \frac{\xi-x}{\alpha}\right)}{\Gamma(\kappa)} \quad (8)$$

$$f(x, \xi , \alpha, \kappa) = \frac{(\xi-x)^{\kappa-1}\ e^{-(\xi-x)/\alpha}}{\alpha^{\kappa}\ \Gamma(\kappa)} \quad (9)$$

If $\gamma = 0$, then the distribution is normal distribution, where the CDF and PDF equal to:

$$F(x, \mu , \sigma) = f(x, \mu , \sigma) = \Phi\left(\frac{x-\mu}{\sigma}\right) \quad (10)$$

where x refers to the primary data series which is to be fitted by the GEV, Gumbel, and LPT3 distribution., Γ is the gamma function (i.e., is one of the extensions of the factorial function) ,$\bar{\Gamma}$ is the incomplete gamma function (i.e., is a type of special function, which is similar to the gamma function but with incomplete integral limits), and Φ is the Euler phi function (written $\Phi(n)$, is the number of non-negative integers less than n, that are relatively prime to n).

3.3. L-Moments Method

In this research, L-moments method is used to fit the GEV, Gumbel, and LPT3 distributions. L-moments are defined as liner combinations of Probability Weighted Moments equations (PWMs), which were introduced by Greenwood *et al*. (1979) and others (i.e., Landwehr *et al*., 1979; Wallis, 1980; Greis and Wood, 1981; Hosking *et al*., 1985; Hosking and Wallis, 1987) to develop statistical inference procedures for use as a tool for estimating the parameters of probability distributions. Sample PWMs are computed by the flowing equations:

$$b_o = n^{-1} \sum_{i=1}^{n} X_i \quad (11)$$

$$b_r = n^{-1} \sum_{i=r+1}^{n} \frac{(i-1)(i-2)\ldots\ldots(i-r)}{(n-1)(n-2)\ldots\ldots(n-r)} X_i \quad (12)$$

where data values X_1, X_2, ... X_n, are arranged in increasing order, n is the sample size, i is the rank of the value in ascending order, and b_o, and b_r two forms of the probability-weighted moments. L-moments are derived as the linear combinations of probability weighted moments, which can be used as measures of the location, scale and shape of the data to be analyzed. The first four L-moments, expressed as liner combination of PWMs, are defined by:

$$\hat{\lambda}_1 = b_o \quad (13)$$

$$\hat{\lambda}_2 = 2b_1 - b_o \quad (14)$$

$$\hat{\lambda}_3 = 6b_2 - 6b_1 + b_o \quad (15)$$

$$\hat{\lambda}_4 = 20b_3 - 30b_2 + 12b_1 - b_o \quad (16)$$

The first L-moment is simply the mean of the data i.e, it is related to the location, the second L-moment is standard deviation, which is a measure of the scale of the data values about the mean. L-moment ratio is defined as:

$$\tau_r = \frac{\hat{\lambda}_r}{\hat{\lambda}_2} \quad (17)$$

where $\hat{\lambda}_r$ is higher-order L-moments and $\hat{\lambda}_2$ is the dispersion measure (in case r =2, $\hat{\lambda}_1$ is the dispersion measure). These are dimensionless parameters and do not depend on the unit of measurement. τ_2 is a measure of L-variation, τ_3 is a measure of L-skewness and τ_4 is a measure of L-kurtosis. The L-coefficient of variation (L-CV) is defined by:

$$\tau_2 = \frac{\hat{\lambda}_2}{\hat{\lambda}_1} \quad (18)$$

L-Skewness gives indication about the symmetry, with a negative value indicating a longer left tail and vice versa. L-Skewness is defined by:

$$\tau_3 = \frac{\hat{\lambda}_3}{\hat{\lambda}_2} \quad (19)$$

L-kurtosis is much less biased than ordinary kurtosis and it is defined by:

$$\tau_4 = \frac{\hat{\lambda}_4}{\hat{\lambda}_2} \quad (20)$$

The parameters of the three distributions can be expressed by L-Moments method in the following sections (Hosking and Wallis, 1997).

3.3.1. Estimating the Parameters of GEV Distribution

$$\kappa = 7.8590\, c + 2.9554\, c^2 \tag{21}$$

where

$$c = \frac{2}{3+\tau_3} - \frac{ln2}{ln3} \tag{22}$$

$$\alpha = \frac{\hat{\lambda}_2 \times \kappa}{(1-2^{-\kappa})\ \Gamma(1+\kappa)} \tag{23}$$

$$\xi = \hat{\lambda}_1 - \alpha\,[1-\Gamma(1+\kappa)]/\kappa \tag{24}$$

Once all parameters have been estimated, calculating the *T*-year return period quantile (X_t) can be done using the following equation:

$$X_t = \xi + \left(\frac{\alpha}{\kappa}\right)\left[1-\left(-log\left(\frac{T-1}{T}\right)\right)^{\kappa}\right] \tag{25}$$

3.3.2. Estimating the Parameters of Gumbel Distribution

$$\alpha = \frac{\hat{\lambda}_2}{log\,2} \tag{26}$$

$$\xi = \hat{\lambda}_1 - (\alpha\,\gamma) \tag{27}$$

where $\gamma = 0.5772$ (Euler's Constant)

Calculating the *T*-year return period quantile (X_t) can be done using the following equation:

$$X_t = \xi + \alpha\,\left(-ln\left[-ln\left(1-\frac{1}{T}\right)\right]\right) \tag{28}$$

3.3.3. Estimating the Parameters of LPT3 Distribution

$$\gamma = 2\,\beta^{-0.5}\,sign(\tau_3) \tag{29}$$

$$\sigma = \frac{\hat{\lambda}_2\,\pi^{0.5}\ \beta^{0.5}\,\Gamma(\beta)}{\Gamma(\beta+0.5)} \tag{30}$$

$$\mu = \hat{\lambda}_1 \tag{31}$$

For estimating the parameter of the LPT3 distribution (β):

if $0 < |\tau_3| < \frac{1}{3}$, let $z = 3\,\pi\,{\tau_3}^2$ then

$$\beta = \frac{1+0.2906\,z}{z+0.1882\,z^2+0.044\,z^3} \tag{32}$$

if $\frac{1}{3} < |\tau_3| < 1$, let $z = 1 - |\tau_3|$ then

$$\beta = \frac{0.36067\,z-0.59567\,z^2+0.25361\,z^3}{1-2.7886\,z+2.56096\,z^2-0.7704\,z^3} \tag{33}$$

To calculate the *T*-year return period quantile (X_t), the data is converted to the logarithmic series, in which $y=log(x)$. The LPT3 distribution based on the mean (μ_y), standard deviation (σ_y), and skewness (γ) of the converted logarithmic series becomes:

$$X_t = \mu_y + K_T\,\sigma_y \tag{34}$$

where K_T is frequency factor for return period T, which depends on skewness and can be approximated by (Kite, 1977):

$$K_T = Z + (Z^2 - 1)\,\gamma + \frac{1}{3}(Z^3 - 6Z)\,\gamma^2 - (Z^2 - 1)\,\gamma^3 + Z\gamma^4 + \frac{1}{3}\gamma^5 \tag{35}$$

$$\gamma = \frac{C_{s(y)}}{6} \tag{36}$$

$$Z = w - \frac{2.515517 + 0.802853w + 0.010328w^2}{1 + 1.432788w + 0.189269w^2 + 0.001308w^3} \tag{37}$$

$$w = \left[ln\left(\frac{1}{P^2}\right)\right]^{1/2} \tag{38}$$

$$P = \frac{1}{T} \tag{39}$$

where $Cs_{(y)}$ is coefficient of skewness, P is probability, T is return period and γ, Z and w are the values used to calculate K_T.

3.4. Goodness-of-Fit Tests

Goodness-of-fit analysis based on probability plots were employed in this research. Three goodness-of-fit tests (i.e., Kolmogorov-Smirnov (KS), Anderson-Darling (AD) and Chi-Square tests) are adopted to assess the goodness-of-fit of the GEV, Gumbel, and LPT3 distributions. In this study, three goodness-of-fit tests are used to assess how well a given distribution fits the rainfall data series of a given duration. The tests are conducted at three significance levels (10%, 5%, and 1%). The critical values based on sample size for n are shown in Table 1. These methods assess the fitted distribution at a site by summarizing the deviations between observed and computed data series. The detail of each test is provided in the following sections.

Table 1: Critical values for the goodness-of-fit tests as a function of sample size (n=30).

Significance level (α)	Tests		
	KS test	AD test	x^2 test
0.01	0.28987	3.9074	11.345
0.05	0.2417	2.5018	7.8147
0.1	0.21756	1.9286	6.2514

3.4.1. Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov (KS) test (Kirkman, 1996) utilises the greatest vertical difference between the theoretical and the empirical cumulative distribution functions:

$$KS = \max_{1 \le i \le n}\left\{F(X_i) - \frac{i-1}{n}, \frac{i}{n} - F(X_i)\right\} \tag{40}$$

Where, F is the cumulative distribution function of the probability distribution being tested.

3.4.2. Anderson-Darling Test

Anderson-Darling (AD) test (Scholz and Stephens, 1987) gives more weight to the tails of the distribution and is defined by the following equation:

$$AD^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)[\ln(F(X_i) + \ln(1 - F(X_{n-i+1})))] \tag{41}$$

3.4.3. Chi-Squared Test

The Chi-Squared (χ^2) test (Preacher, 2001) is applied to binned data. The χ^2 test statistic is defined as:

$$\chi^2 = \sum_{i=1}^{N}\frac{(O_i - E_i)^2}{E_i} \tag{42}$$

where O_i is the observed frequency of the data sample and E_i is the expected frequency of data sample calculated by $E_i = F(X_2) - F(X_1)$.

3.5. Developing Stationary IFD Curves Formula

Once, the six rainfall quantiles (T = 2, 5, 10, 25, 50, and 100 year) are estimated from the fitted statistical distributions, the IFD curve is developed by making an empirical relationship among rainfall intensity, return period, and duration, it is expressed mathematically as follows:

$$I = F(T, D) \tag{43}$$

The empirical IFD relationship is given by the below equation:

$$I = (f \times T^v)/(l \times D^e) \tag{44}$$

where I refers to a rainfall intensity (mm/hr), T is return period (year), D is duration (min), and f, l, e, and v are coefficients.

3.6. Root Mean Squared Error

In this study, the relative accuracy of the new IFD curves is assessed by calculating the root mean squared error (RMSE) (expressed by Equation 45) which serves to aggregate the magnitudes of the errors in predictions for data points into a single measure of predictive power.

$$RMSE(\%) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{I}_i - I_i)^2} \times 100 \tag{45}$$

Where $\hat{\boldsymbol{I}}_{\boldsymbol{i}}$ is the forecast or expected value and $\boldsymbol{I}_{\boldsymbol{i}}$ is the observed value.

4. Results and Discussion

4.1. Estimation of the AMR Data for Short Rainfall Durations

The evaluation of IFD curves for the three adopted stations (Duhok, Erbil, and Sulaymaniyah) has started by reducing the values of the daily rainfall data (AMR of 24hr duration) into short durations (smaller time scale) of ten rainfall durations (i.e., 5 min, 10 min, 15 min,30 min, 1 hr, 3 hr, 5 hr, 8 hr, 12 hr, and 3 day) using Equation 1. The AMR data for three stations (i.e., Duhok, Erbil, and Sulaymaniyah) in mm/hr for all the evaluated rainfall durations mentioned above are shown in Figure 3.

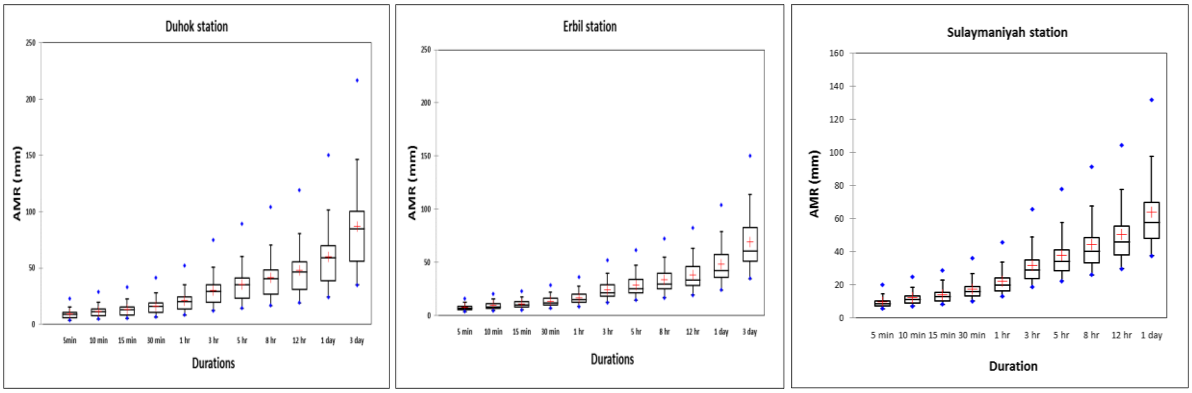


Figure 3: Boxplots of AMR data for different durations of Duhok, Erbil, and Sulaymaniyah stations.

4.2. The Goodness of Fit Tests Results

Based on the AMR series of the three adopted rainfall stations and with the 11 rainfall durations shown in Figure 3, a total of 33 data series (i.e., 3×11) were valid for the three goodness of fit testing. Table 2 shows the results of minimum and maximum values of KS, AD, and $\chi 2$ tests, which were discussed in section 3. Depending on the outcomes of the three goodness of fit tests, the three distributions (i.e., GEV, Gumbel, and LPT3) fit the AMR data series with only a slight variation. Out of the 33 data sets dependent on the critical values in Table 1 and at the three significance levels (10%, 5%, and 1%), none of the three goodness of fit tests rejected the GEV, Gumbel, and LPT3 distributions.

Table 2. Minimum (Min) and Maximum (Max) values of the goodness-of-fit tests.

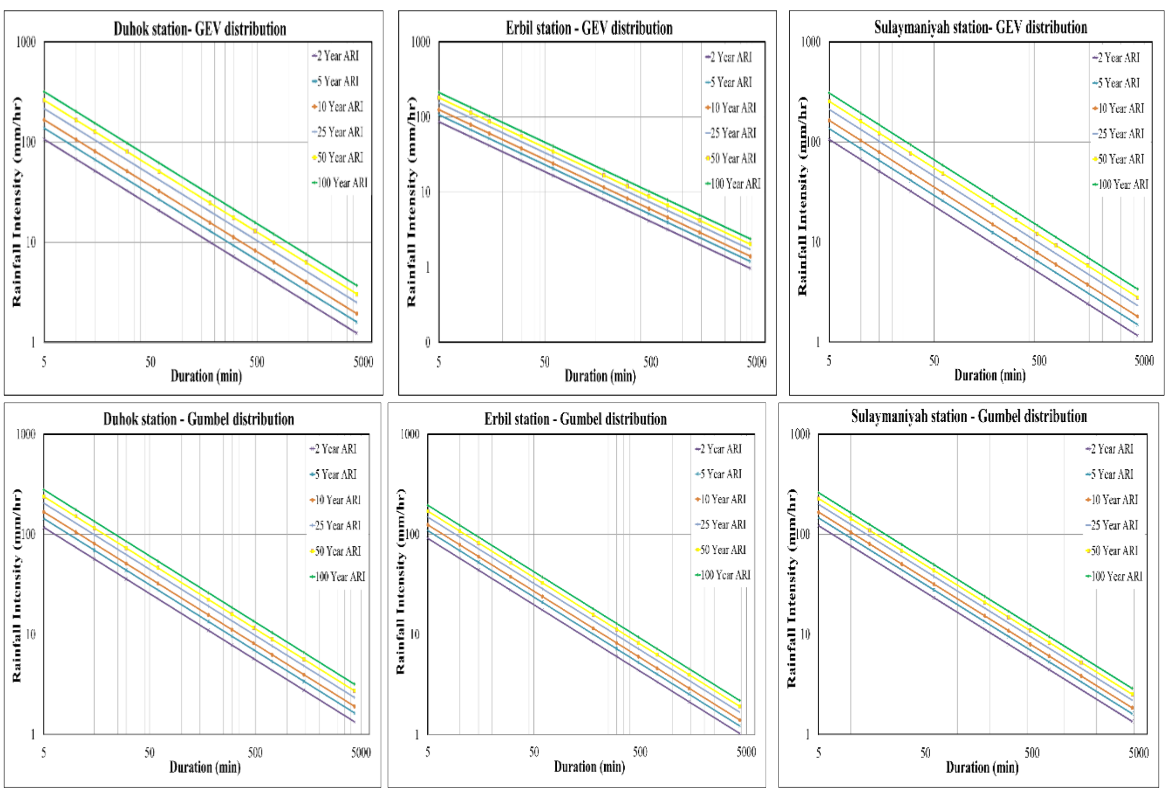
Test	**Kolmogorov-Smirnov**			**Anderson-Darling**			**Chi-Squared**		
Distribution	GEV	Gumbel	LPT3	GEV	Gumbel	LPT3	GEV	Gumbel	LPT3
Duhok station									
Min	0.120	0.120	0.114	0.457	0.586	0.441	0.725	0.609	0.829
Max	0.122	0.122	0.115	0.459	0.588	0.443	0.727	0.614	0.831
Erbil station									
Min	0.060	0.133	0.066	0.163	0.862	0.203	0.582	2.281	0.785
Max	0.068	0.137	0.069	0.169	0.867	0.208	0.589	2.284	0.788
Sulaymaniyah station									
Min	0.069	0.093	0.065	0.111	0.200	0.113	0.241	1.271	0.238
Max	0.075	0.099	0.079	0.116	0.204	0.136	0.248	1.275	0.535

4.3 The IFD curves generated by GEV, Gumbel, and LPT3 distributions

The empirical relationships of the IFD curves for all three adopted stations based on Equation 1 are presented in Table 3. In Table 3, the R^2 values of the fitted IFD curves range from 0.996 to 0.999 (average = 0.998). This demonstrates that the R^2 values for the empirical IFD curves are extremely high (i.e., the curves precisely fit the data). The range of the RMSE values (estimated by Equation 45) is 1.595 to 3.958 (average =1.922). The quantiles of the IFD curves were calculated based on empirical relationships shown in Table 3 by fitting the selected distribution to AMR data from the three selected rainfall stations, as shown in Figure 4. The three adopted distributions give a maximum intensity for the Duhok station at a 100 year return period with a 5 min rainfall duration (i.e., the GEV distribution gives a maximum intensity of 319.268 mm/hr; the Gumbel distribution gives a maximum intensity of 278.513 mm/hr; the LPT3 distribution gives a maximum intensity of 310.555 mm/hr). Figure 4 shows that there is a high degree of consistency in the developed IFD curves across different return periods. Moreover, it is noted that, in general, the IFD curves have a negative gradient, which is consistent with the experience that higher rainfall intensity occurs over shorter durations.

Table 3. Empirical relationship between I (mm/hr), T (years), and D (minutes).

Stations	Distribution	IFD Equation	R^2	RMSE
Duhok	GEV	$I = 308.800T^{0.280}/1.218D^{0.659}$	0.998	2.819
	Gumbel	$I = 179.678T^{0.221}/0.617D^{0.661}$	0.996	3.958
	LPT3	$I = 173.137T^{0.266}/0.660D^{0.658}$	0.998	3.242
Erbil	GEV	$I = 135.528T^{0.229}/0.627D^{0.664}$	0.998	1.789
	Gumbel	$I = 135.940T^{0.194}/0.587D^{0.663}$	0.997	2.258
	LPT3	$I = 135.404T^{0.219}/0.615D^{0.663}$	0.998	1.909
Sulaymaniyah	GEV	$I = 308.243T^{0.273}/1.198D^{0.667}$	0.999	0.818
	Gumbel	$I = 181.968T^{0.194}/0.587D^{0.665}$	0.997	2.969
	LPT3	$I = 177.330T^{0.248}/0.647D^{0.666}$	0.999	1.595



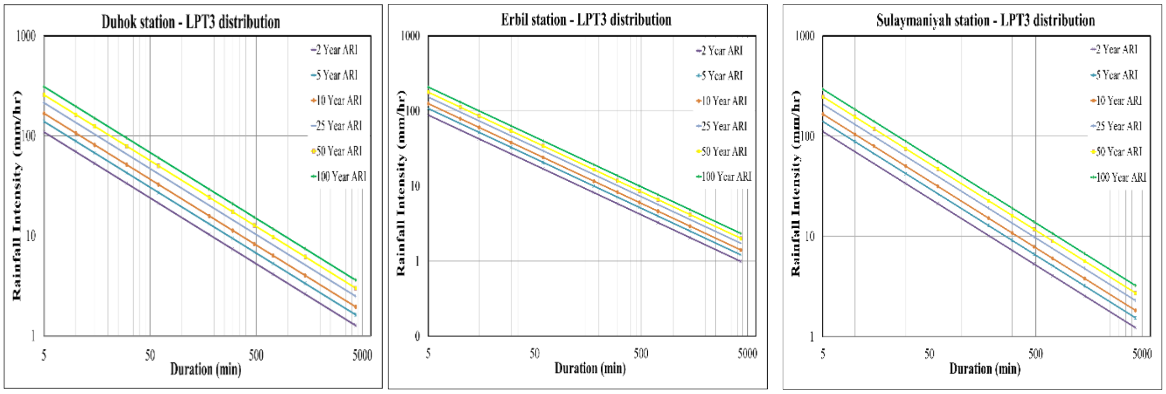


Figure 4. Developed stationary IFD curves based on empirical relationships shown in Table 3.

4.4 Comparison between IFD curves generated by GEV, Gumbel, and LPT3 distributions

The IFD curves based on the three adopted distributions (i.e., GEV, Gumbel, and LPT3) were compared. The comparison is illustrated for the 100 year return period for the thee selected stations (i.e., Duhok, Erbil, and Sulaymaniyah stations) in Figure 5. This figure shows that for all selected stations, there is quite a good match between the IFD curves based on three adopted distributions. Overall, it has been found that at-site IFD curves data derived from the GEV distribution are generally higher than the curves derived from the LPT3 and Gumbel distributions.

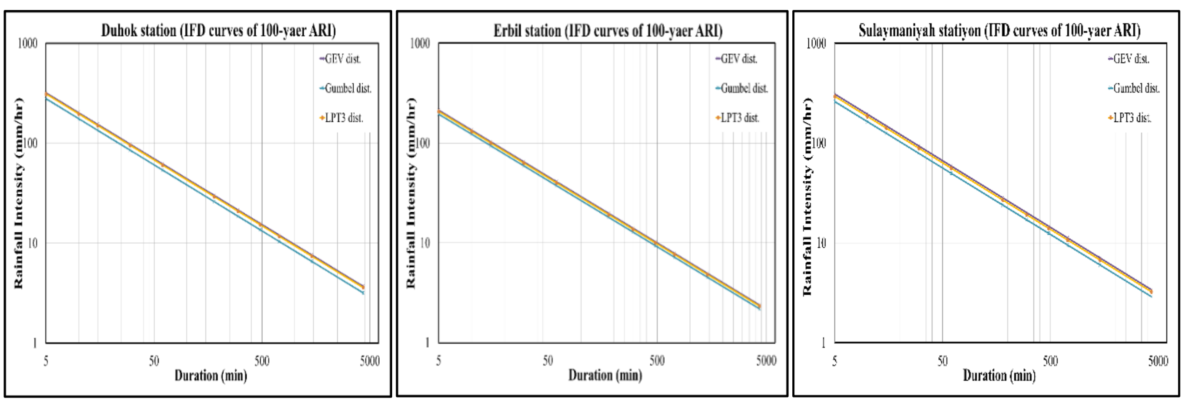


Figure 5. Stationary IFD curves of 100 year ARI derived from GEV, Gumbel, and LPT3 distributions for Duhok, Erbil, and Sulaymaniyah stations.

5. Conclusions

The annual maximum rainfall (AMR) data for three rainfall stations (Duhok, Erbil, and Sulaymaniyah) located in three main governorates, Kurdistan Region of Iraq, with a recorded length of 30 years covering the period 1991-2020, were used in this study. The stationary intensity frequency duration (IFD) curves have been derived at the site location using the Generalized Extreme Value (GEV), Gumbel, and Log Pearson Type 3 (LPT3) distributions for the three selected stations. Eleven rainfall durations (i.e., 5 min, 10 min, 15 min, 30 min, 1 hr, 3 hr, 5 hr, 8 hr, 12 hr, 1 day, and 3 day) and six return periods (i.e., 2, 5, 10, 25, 50, and 100 year) have been considered. The three adopted distributions are found to fit the AMR data (at 10%, 5%, and 1% significance levels) for all three of the selected stations based on the three goodness of fit tests (Kolmogorov-Smirnov test, Anderson-Darling test, and Chi-Square test). Depending on the results of the goodness-of-fit tests, the three adopted distributions (i.e., GEV, Gumbel, and LPT3) showed a minor difference in fitting the AMR data series. The results of the at-site stationary IFD curves show that for all the selected stations, there is quite a good match between the IFD curves based on three adopted distributions. The results of this study should improve guidance for choosing the appropriate IFD data for a particular application in the Kurdistan governorates. The methodology created here can be applied to other cities in Iraq and the Kurdistan Region

Acknowledgements

The authors would like to thank the Directorate of Meteorology and Seismology in Erbil and Duhok governorates for providing daily rainfall data that was used in this study.

References

1. Al-Anazil K. K. and El-Sebaie I. H. (2013). Development of Intensity-Duration-Frequency Relationships for Abha City in Saudi Arabia. International Journal of Computational Engineering Research, 3(10): 58-65.
2. Bernard M.M. (1932). Formulas for Rainfall Intensities of Long Duration. Transactions, ASCE, 96: 592-624.
3. Chow V.T., Maidment D. R. and Mays L.W. (1988). Applied Hydrology. New York: McGraw- Hill.
4. Dakheel A. A. (2017). Drawing Curves of the Rainfall Intensity Duration Frequency (IDF) and Assessment Equation Intensity Rainfall for Nasiriyah City, Iraq. Journal of University of Thi-Qar, 12(2): 63-83.
5. El-Sayed E. A. (2011). Generation of Rainfall Intensity Duration Frequency Curves for Ungauged Sites. Nile Basin Water Science and Engineering Journal, 4(1): 112-124.
6. Fréchet M. (1927). Sur La Loi De Probabilité De L'écart Maximum. Journal of Annales De La Société Polonaise De Mathématique, 6: 93.
7. Greenwood J.A., Landwehr J.M. Matalas N.C. and Wallis J.R. (1979). Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressible in Inverse Form. Water Resources Research, 15: 1049-1054.
8. Greis N.P. and Wood E.F. (1981). Regional Flood Frequency Estimation and Network Design, Water Resources, 17 (4): 1167-1177.
9. Gumbel E.J. (1958). Statistics of Extremes. Columbia University Press, New York.
10. Hajani E. (2020). Climate Change and Its Influence on Design Rainfall At-site in New South Wales State, Australia. Journal of Water and Climate Change, 11 (S1): 251–269.
11. Hajani E. and Klari Z. (2022). Trends Analysis in Rainfall Data Series in Duhok City, Kurdistan Region, Iraq. Modeling Earth Systems and Environment, s40808-022-01354-1
12. Hajani E., and Rahman A. (2018). Design Rainfall Estimation: Comparison Between GEV and LPT3 Distributions and at-site and Regional Estimates. Natural Hazards, 93: 67–88.
13. Hosking J.R. and Wallis J. (1987). Parameter and Quantile Estimation for the Generalized Pareto Distribution. Technometrics, 29: 339-349.
14. Hosking J.R. and Wallis J. (1997). Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press, New York, 224.
15. Hosking J.R., Wallis J. and Wood E. (1985). Estimation of the Generalized Extreme Value Distribution by the Method of Probability Weighted Moments. Technometrics, 27: 251-261.
16. Kirkman T.W. (1996). Statistics to Use: Kolmogorov Smirnov Test.
17. Kite, G.W. 1977. Frequency and Risk Analysis in Hydrology. Water Resources Publications, Fort Collins, Colorado, USA.
18. Landwehr J., Matalas N. and Wallis J. (1979). Probability Weighted Moments Compared with Some Traditional Techniques in Estimating Gumbel Parameters and Quantiles. Water Resources Research, 15: 1055-1064.
19. Millington N., Das S. and Simonovic S. P. (2011). The Comparison of GEV, Log-Pearson Type 3 and Gumbel Distributions in the Upper Thames River Watershed under Global Climate Models. Raziskovalno Poročilo O Vodnih Sredstvih. London, Ontario, Canada, University of Western Ontario, Department of Civil and Environmental Engineering, 52.

20. Preacher K. J., 2001. Calculation for the Chi-Square Test: An Interactive Calculation Tool for Chi-Square Tests of Goodness of Fit and Independence [Computer Software].
21. Ramaseshan, S. (1996). Urban Hydrology in Different Climatic Conditions, Lecture notes of the International Course on Urban Drainage in Developing Countries, Regional Engineering College, Warangal, India.
22. Scholz F.W. and Stephens M.A. (1987). K Sample Anderson Darling Tests. Journal of the American Statistical Association, 82: 918-924.
23. Sherman C.W. (1931). Frequency and Intensity of Excessive Rainfall at Boston. American Society of Civil Engineers, 95: 951-960.
24. Suchithra A. S. and Agarwal S. (2020). IDF Curve Generation for Historical Rainfall Events. Proceeding of National Conference on Emerging Trends in Civil Engineering. Andhra Pradesh, KL University, 898-907.
25. Sulaiman A. H. (2015). Development of Intensity Duration Frequency Curves for Duhok Province/Iraqi Kurdistan Region. Journal of University of Duhok, 19:1-15.
26. Wallis J.R. (1980). Risk and Uncertainties in the Evaluation of Flood Events for the Design of Hydraulic Structures, in Piene E Siccità (Guggino E., Rossi G., and Todini E.). Fondazione Politecnica Del Mediterraneo, Catania, Italy, 3-36.
27. Weibull W. (1951). A Statistical Distribution Function of Wide Applicability. Journal of Applied Mechanics, Transactions ASME, 18 (3): 293-297.
28. Zainudini M. A., Marriott M. J., Mirjat M. S. and Chandio A. S. (2011). Establishing Intensity Duration Frequency Curves for Sistan and Balochistan Provinces of Iran. Pakistan Journal of Agriculture, Agrilculture Engineering and Veterinery Sciences, 27 (2): 115-124

Appendix A: List of Abbreviations

IFD	- Intensity-Frequency-Duration
AMR	- Annual Maximum Rainfall
GEV	- Generalized Extreme Value
LPT3	- Log Pearson Type 3
CDF	- Cumulative Density Function
PDF	- Probability Density Function
PWMs	- Probability Weighted Moments
KS	- Kolmogorov-Smirnov
AD	- Anderson-Darling
RMSE	- Root Mean Squared Error

Appendix B: List of Symbols

T	- Time
R_t	- Rainfall depth of t (in minute) duration
R_{24}	- Daily rainfall
α	- Significance level
ξ	- Location parameter
a	- Scale parameter

κ	- Shape parameter
x	- Primary data series which is to be fitted by the distribution
b	- Form of the probability-weighted moments
$\hat{\lambda}$	- L-moment
τ_r	- L-moment ratio
τ_2	- L- coefficient of variation
τ_3	- L- coefficient of skewness
τ_4	- L- coefficient of kurtosis
μ	- Mean of the data time-series
σ	- Standard deviation of the data time-series
γ	- Skewness value
$Cs_{(y)}$	- Coefficient of skewness
K_T	- Frequency factor
Z, w	- Values for K_T calculation
χ^2	- Chi-Squared test statistic
O_i	- Observed frequency of the data sample
E_i	- Expected frequency of data sample
F	- Cumulative distribution function of the probability distribution
f	- Probability distribution function of the probability distribution
f, l, e, v	- Coefficients of empirical IFD relationship
$\hat{I}_i$	- Expected values of the data point
I_i	- Observed values of the data point