Malaysian Journal
Of Civil
Engineering

ASSESSMENT OF THE BEST PROBABILITY DISTRIBUTION METHOD IN RAINFALL FREQUENCY ANALYSIS FOR A TROPICAL REGION

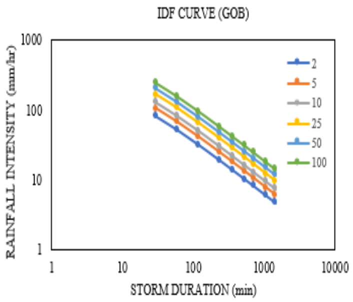
Stephen Luo Sheng Yong[a], Jing Lin Ng[a*], Yuk Feng Huang[b], Chun Kit Ang[a]

[a]Faculty of Engineering, Technology and Built Environment, UCSI University Kuala Lumpur, Malaysia
[b]Department of Civil Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Selangor, Malaysia

Article history
Received
7 October 2020
Received in revised form
7 January 2021
Accepted
11 January 2021
Published online
31 March 2021

*Corresponding author
ngjl@ucsiuniversity.edu.my

Graphical abstract



Abstract

The Intensity-Duration-Frequency (IDF) curve defines the relationship between rainfall intensities at certain durations and with the frequencies. The IDF Curve is extensively used in many applications such as flood modelling and peak discharge estimation. Over the years, the frequent occurrence of flood has become a great challenge in Kelantan river basin. Herein, IDF curves using frequency analyses based on different distributions were developed and compared. The historical rainfall data at eight rainfall stations for the period of 1985-2019 were selected for the assessment purpose. The Gumbel, Normal and Log Pearson Type III distributions were fitted into the annual maximum rainfall series for durations varying from 30 minutes to 24 hours. The goodness of fit tests were then used to evaluate the performances of each frequency distribution. It was found that the Gumbel distribution gave the highest passing rate followed by the Log Pearson Type III and then the Normal distributions. The Gumbel distribution resulted in respective 86% and 75% passing rate since most of the p-values generated by both the K-S and the Mann-Whitney test were greater than 5% of significance level leading to the acceptance of the null hypothesis. Thus, the Gumbel distribution is suggested for the frequency analyses in this study.

Keywords: IDF curve, Gumbel, Log Pearson Type III, Normal, Goodness of fit tests

1.0 INTRODUCTION

The IDF curve is expressed as the relationship between rainfall intensity and rainfall duration as well as the frequency of the downpour event (Lee, 2005; Huang *et al.*, 2014). It is one of the common approaches for rainfall intensity estimation at various duration and recurrence interval in hydrology analyses (Nhat *et al.*, 2006; Chang *et al.*, 2015). Each curve shows the rainfall intensity, duration and elements at various frequencies or average recurrence intervals (ARIs) (Mohd Ariff *et al.*, 2013). The IDF curve compromises three parameters which are duration (d), rainfall intensity (i), and the return period (T) of a downpour event. The parameters (storm duration and intensity) of IDF curve have great significance in the field of hydrology, as they are the basic elements when it comes to hydrologic risk analysis and urban infrastructural development such as stormwater drainage system, sewage and culverts (Liew *et al.*, 2014). Design engineers are able to deduce the falling time of maximum rainfall intensity within a specific time interval by using the IDF curve (Wagesho *et al.*, 2016). It determines the likelihood of return period of downpour occurrence within an aggregated time. The most severe case situation of rainfall intensity for different durations is explored and highly depended for the sake of urban drainage or sewage infrastructure development. Design engineers require to study the pattern of the IDF curve to analyse the how frequent downpour occurs at various time span for design frequency chosen. Therefore, occurrence of flooding events in Malaysia can be minimized.

In developing the IDF curve, frequency analysis is the application of probability distribution to link the magnitude of rainfall events to the probability of occurrence, graphically. Not only rainfall intensity generated by IDF curve is essential for urban infrastructure project design, but it is also necessary for weather estimation, flood risk assessment, reservoir management and stormwater management systems within cities to prevent economic loss and the list goes on (Ng, *et al.*, 2016; Kaboosi *et al.*, 2017; Leščešen *et al.*, 2019). Normal distribution, Extreme-value distribution, Binomial distribution, Log-Pearson Type III distribution, Log-normal distribution are several examples of probability distribution that are engaged for the purpose. Amongst all the probability distributions, the Normal and the Generalized Extreme-value distributions are mostly utilized for frequency analysis. The Normal distribution is extensively utilized in average annual temperature, annual rate of river flow, and so on which categorized under the central tendency observations (Mirzaei *et al.*, 2014). The Generalized Extreme distribution on the other hand, is more effective in modelling input data with less parameters. It displays a clear connection among rainfall maximum and return periods when fitted to the annual maximum rainfall series (Ng *et al.*, 2015; Ng *et al.*, 2019). However, the Gumbel distribution is always the priority for frequency analysis purpose when annual maximum series data is used (National Hydraulic Research Institute of Malaysia, 2010). It is noted that since the Gumbel distribution has not been compared with other probability distribution to identify the reliability of its distribution, therefore a few probability distributions will be used in this study in order to deduce the reliability of each distribution.

Accordingly, the main objective of this study is to determine and compare the reliability of the Gumbel, Normal and Log Pearson Type III distributions. The Intensity-duration frequency (IDF) curve for the Kelantan River Basin is to be developed using frequency analysis. The historical rainfall data of eight rainfall stations for the period of 1985-2019 were acquired. The Gumbel, Normal and Log Pearson Type III distributions were fitted into the annual maximum rainfall series for durations of 30 minutes, 1 hour, 2 hours, 4 hours, 6 hours, 9 hours, 12 hours and 24 hours.

2.0 METHODOLOGY

2.1 Study Area

Kelantan, one of the states in Peninsular Malaysia, is situated at the north-eastern part facing towards the South China Sea. Kelantan occupies an approximate area of 15101 km^2 (4.4% of Malaysia's land surface area) with a population of 1.86 million people. The biggest river in the Kelantan state is the Kelantan River. There are four dominant tributaries in Kelantan River, specifically the Galas, Nenggiri, Lebir and Pergau Rivers. They are frequently flooded by heavy rainfall from November until March (Northeast Monsoon period) (Ng *et al.*, 2020). All these four dominant tributaries has further tributaries of their own (Pradhan and Youssef, 2011). The mean width of river in Kelantan state ranges from 180 – 300 m and approximately 68.5% of the citizens dwell around Kelantan River Basin area (Adnan *et al.*, 2014).

In this study, the rainfall data at the various rainfall stations scattered around Kelantan state were obtained from the Department of Irrigation and Drainage (DID) Malaysia. There are a total number of eight stations chosen in order to develop the IDF curve as well as analysing the relationship between rainfall intensity, duration and return periods at each station. Detail information such as latitude, longitude, duration and station code for each station are indicated in Table 1. In general, rainfall data cover the years ranging from 1985 to 2019

2.2 Estimation of Missing Rainfall Data

When it comes to rainfall data collection at the rainfall stations, there will be missing rainfall event data, inevitably. This situation is credited to common mistakes such as human error, malfunction of instruments and communication lines and a host of other possibilities. Therefore, an estimation of missing data needs to be conducted. The approach used in this paper is nearest neighbour approach through the application of XLSTAT statistical software. According to XLSTAT, there are three types of missing values, and they are; data Missing Completely At Random (MCAR), data Missing At Random (MAR) and data Not Missing At Random (NMAR). Missing rainfall data which is categorized under data missing at random can be solved through nearest neighbour approach.

Table 1 Details of each rainfall station

	Station name	Record Period	Duration	Latitude	Longitude
4614001	Brook	1985-2019	34.3	04° 40' 35'' N	101° 29' 05'' E
4717001	Blau	1985-2019	34.3	04° 46' N	101° 45' 25'' E
5216001	Gob	1985-2019	34.3	05° 15' 05'' N	101° 39' 45'' E
5322044	Kpg. Lalok	1985-2019	34.3	05° 18' 30" N	102° 16' 30" E
5522047	JPS Kuala Krai	1985-2019	34.3	05° 31' 55" N	102° 12' 10" E
5718033	Kpg. Jeli	1985-2019	34.3	05° 42' 05" N	101° 50' 20" E
6019004	Rumah Kastam	1985-2019	34.3	05° 00' N	101° 34' E
6122064	Setor JPS Kota Bahru	1985-2019	34.3	06° 06' 30'' N	102° 15' 25'' E

2.3 Empirical IDF curves

One of the crucial roles for the empirical equation is to reduce mistakes in IDF curves rainfall intensity values evaluation. The empirical equation is expressed as (Government of Irrigation and Drainage Malaysia, 2012):

$$i = \frac{\lambda T^k}{(d+\theta)^n} \qquad (1)$$

where i represents the average rainfall intensity (mm/hr), T represents the average recurrence interval – ARI (0.5 < T < 12 month and 2 < T < 100 year), and λ, θ, κ, η represent the fitting constants dependent on the rainfall station location.

2.4 Fitting of Frequency distribution - Gumbel Distribution

The Extreme-value type I distribution is also known as the Gumbel distribution was created by Gumbel in 1940's (Alghazali *et al.*, 2014). Apart from its utilization in flood frequency estimation, it also plays a vital part in meteorological frequencies analysis. The estimation of recurrence period of rainfall event can be deduced through the characteristics posed by the Gumbel distribution whereas the likelihood of rainfall event to occur can be predicted through Gumbel distribution's parameter, mean, standard deviation and the list goes on. The equations for frequency precipitation with return period (T) at different duration are indicated as below (Elsebaie, 2012):

$$P_T = P_{ave} + KS \qquad (2)$$

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

$$S = \left[\frac{1}{n-1}\sum_{i=1}^{n}(P_i - P_{ave})^2\right]^{0.5} \qquad (4)$$

where P_{ave} represents average maximum rainfall for each duration, K represents Gumbel frequency factor and S represents standard deviation.

2.5 Log Pearson type III Distribution

The Log Pearson type III distribution is also known as the Gamma distribution. In the United States, this method is mandatory for projects that involve precipitation frequency analysis. The U.S. Weather Resources amended and improvised the Pearson type III distribution by log transforming volume of water into Log Pearson type III distribution (Lee, 2014). Log Pearson type III has a flaw which provides unsatisfied low upper bounds of rainfall magnitude (Millington *et al.*, 2011). It has two shape parameters which interact with one another and three parameters such as:

i. Shape parameter (γ)
ii. Location parameter (μ)
iii. Scale parameter (σ)

The equations for Log Pearson type III distribution are shown below:

$$P^* = \log(P_t) \qquad (5)$$

$$P{*}T = P{*}ave + KTS{*} \qquad (6)$$

$$P^*_{ave} = \frac{1}{n}\sum_{i=1}^{n} P^* \qquad (7)$$

$$S^* = \left[\frac{1}{n-1}\sum_{i=1}^{n}(P^* - P^*_{ave})^2\right]^{0.5} \qquad (8)$$

$$C_X = \frac{n\sum_i^{ni}(P^*_i - P^*_{ave})^3}{(n-1)(n-2)(S^*)^3} \qquad (9)$$

where P^*_{ave} represents average maximum rainfall for each duration, K_T represents Pearson frequency factor, S^* represents standard deviation and C_X represents skewness coefficient.

2.6 Normal Distribution

The Normal distribution, also namely the Gaussian distribution was developed by Gauss whilst conducting research on measurement errors. Many probability distributions can be approached with the Normal distribution since it possesses many fine features. The variation in the characteristic under consideration are determined by independent factors with limited magnitudes (Gubareva, 2011). Not only it is the basis of probability distributions like the Log-normal distribution and the three-parameter Log-normal distribution, it is also useful in hydrology field when it comes to describing events such as mean annual stream flow and annual pollutant loadings.

$$P_T = P_{ave} + KS \qquad (10)$$

$$w = \left[ln\left(\frac{1}{p^2}\right)\right]^{1/2} \quad (0 < p < 0.5) \qquad (11)$$

$$K_T = z = w - \frac{2.515517 + 0.802853w + 0.010328w^2}{1 + 1.432788w + 0.18929w^2 + 0.001308w^3} \qquad (12)$$

$$S = \left[\frac{1}{n-1}\sum_{i=1}^{n}(P_i - P_{ave})^2\right]^{0.5} \qquad (13)$$

where P_{ave} represents average maximum rainfall for each duration, S represents standard deviation, w represents intermediate variable and z represents normal variable.

2.7 Goodness of fit tests – Kolmogorov-Smirnov Test

The formula of the Kolmogorov-Smirnov test is derived through the largest vertical distance from empirical and theoretical cumulative density function source. The Kolmogorov-Smirnov value is considered a rejection if it exceeds the critical value which is 0.12555 provided significance level is $\alpha = 0.05$ (Mohd Ariff *et al.*, 2013). It is expressed as follows:

$$D = max\left[F(x) - \frac{i-1}{n}, \frac{i}{n} - F(X_i)\right] \tag{14}$$

where D represents test statistics, F(X) represents observed cumulative frequency distribution of a random sample of n observations, *i* represents number of observations and n represents total number of observations.

2.8 Mann-Whitney Test

The Mann- Whitney test or the Wilcoxon-Mann-Whitney test is a non-parametric test that is used to compare two independent variables. This test was improved by three researchers, Mann, Whitney and Wilcoxon, to identify whether the samples are identical or otherwise, based on their ranks. The equation for Mann-Whitney test is expressed as follow:

$$U_1 = R_1 - n_1 n_2 + \frac{n_1(n_1+1)}{2} \tag{15}$$

$$U_2 = R_2 - n_1 n_2 + \frac{n_1(n_1+1)}{2} \tag{16}$$

2.9 Graphical Test

The quantile-quantile (Q-Q) plot is chosen for this study. The Quantile-Quantile test is a graphical method that plots observed data against theoretical data. An accurate model theoretical distribution will provide linear pattern on the graph (Galoie *et al.*, 2013). The equations are shown as follow:

$$Q_F(P_i) = F^{-1}(P_t) \tag{17}$$

$$P_t = \frac{i - 0.5}{n} \tag{18}$$

$$P_i = \frac{i - c}{n - 2c + 1} \tag{19}$$

where Pi represents plotting position, Pt denotes plotting position, $F^{-1}(P_t)$ represents inverse function of P_t, N represents sample of size, i represents nth of sample size and c represents constant value.

The X-axis of the graph represents observed data values, X_i (i = 1, 2, 3, …n) whereas y-axis is expressed as the equation below:

$$F^{-1}\left(F_n(x_i) - \frac{0.5}{n}\right) \tag{20}$$

where F^{-1} (x) represents inverse cumulative distribution function and F(x) denotes cumulative distribution function.

3.0 RESULTS AND DISCUSSION

Figure 1 shows the IDF curves for the Gob station using the empirical, Gumbel, Normal and Log Pearson Type III distributions respectively. Rainfall estimates were increasing with the return period whereas rainfall intensities (mm/hr) were decreasing with rainfall durations in all return periods varying from 2 to 100 years. From Figure 1(a), a good consistency was shown with the results obtained, as the IDF curve was illustrating a similar decline trend of rainfall intensity for both the Gumbel and Normal distributions. The IDF curve for the Gob station decreased significantly as storm duration increases corresponding to ARI 2, 5, 10, 25, 50 and 100 years. For the Log Pearson Type III distribution, the rainfall intensity at ARI 2, 5, 10 and 25 years indicated a declination with time except for the ARI 50 and 100 years. There was slight fluctuation for rainfall intensity for both the ARI 50 and 100 years as the rainfall intensity decreases and increased slightly at the end of duration.

The Quantile-Quantile plots being graphical representation of input data plotted against theoretical distribution quantiles show the Y-axis of Q-Q plot represents sample quantiles whereas X-axis represents theoretical quantiles. From Figure 2, rainfall intensities (30 minutes) by the Gumbel, Normal and Log Pearson Type III distribution were analyzed to compute Q-Q plots. Overall, the Q-Q plots for all distributions were acceptable as all the values concentrated along the provided linear line.

As shown in Table 2, durations of 30 minutes, 360 minutes (6 hours), 720 minutes (12 hours) and 1440 minutes (24 hours) were chosen for comparison between the empirical and the Gumbel distributions. Among all rainfall stations, the Gob station gave the least percentage difference (8%) between the empirical and Gumbel distribution, at 1440 minutes with ARI 5 years. This means that it has the least error as the rainfall intensity by empirical distribution (6.14mm/hr) was only slightly different from rainfall intensity developed by the Gumbel distribution (5.626mm/hr). Nevertheless, the JPS Kuala Krai station has the highest percentage difference, 77%. Rainfall intensity from this station showed smaller value compared to empirical distribution, thus giving in large positive value.

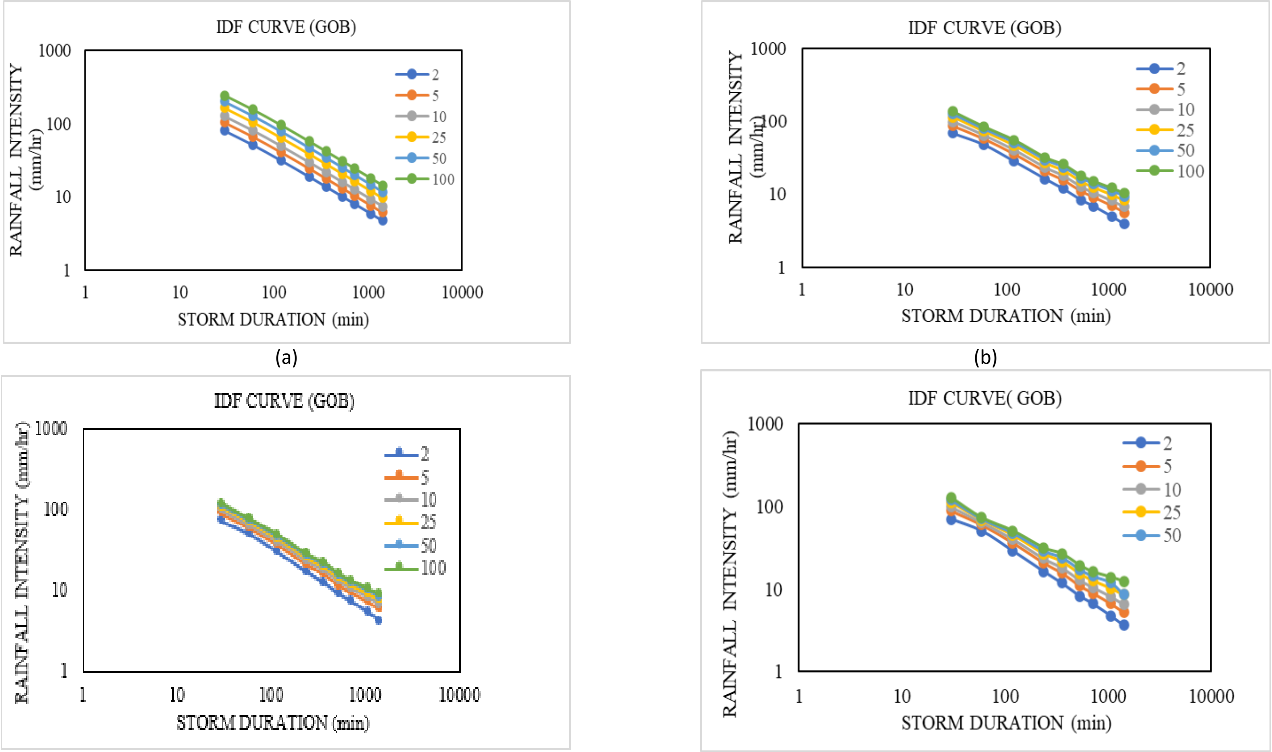
Table 3 demonstrates the fitting results of the rainfall intensities with the Gumbel, Normal and Log Pearson Type III distributions into the Kolmogorov-Smirnov (K-S) and Mann-Whitney tests. To evaluate the best distribution, a score table was tabulated to determine the scores between the Gumbel, Normal and Log Pearson Type III distributions. Distribution with 1-point score indicated that the distribution fits well in the particular of either K-S or Mann-Whitney tests. On the contrary, zero score indicated rejection by goodness of fit test. The distribution with highest total score will be indicated as the best fit.

The Log Pearson Type III distribution demonstrated moderately good fits because it showed more than 50% passing rate. The results are consistent with the findings of Kuchment and Demidov (2013) where the Log Pearson Type III distribution gave the second-best performance in fitting the annual maximum rainfall series varying from 30 minutes to 24 hours.

Meanwhile, the Normal distribution did not perform well as the percentages of passing rate for both the K-S and Mann-Whitney tests were lower than 50% and almost all fittings were unacceptable. This result may be explained by the fact that there are substantial variations of rainfall data within Kelantan River Basin and it is characterized by tropical climate and northeast monsoon (heavy rainfall season) regime throughout the entire year, thus it was not suitable for the rainfall characteristics in Kelantan River Basin. A similar finding of Alghazali and Alawadi (2014) was where the Normal distribution gave the worst performance for all thirteen stations studied.

In general, the Mann-Whitney test showed relatively low percentage of passing rate. This is because the points scored by the Gumbel, Normal and Log Pearson Type III distributions were lower than those of the K-S test. Based on Table 2, the Gumbel distribution was the best distribution when compared with the Normal distribution and Log Pearson Type III distributions, for both the K-S and Mann Whitney tests. The Gumbel distribution gave the highest score (62 and 54 points) followed by Log Pearson Type III distribution and Normal distribution. It was found out that only two rainfall stations with the Gumbel distribution gave several rejection fits (zero point). Since most of the p-values generated by K-S and Mann-Whitney tests were greater than 5% of significance level which results in acceptance of null hypothesis (that is that the the two samples follow the same distribution), the percentages of passing rate for Gumbel distribution for the tests are 86% and 75% respectively.

The most extensively used distribution in IDF analysis is the Gumbel distribution. This is due to its simplicity and propriety in only modelling extreme events such as peak rainfall and maximum values. The Gumbel distribution showed the higher percentage to fit the graph than the Log Pearson Type III distribution, thus indicating the Gumbel distribution is the better distribution in developing the IDF curve in the region of Pahang (Asma'Suhaimee, 2018). Nevertheless, Jefrin *et al.* (2017) stated that Gumbel distribution performed better than the Log Pearson Type III distribution as it depicted lower chi-square values than critical value. Thus, the Gumbel distribution is the more appropriate in developing the IDF Curve in Kelantan River Basin, Malaysia.



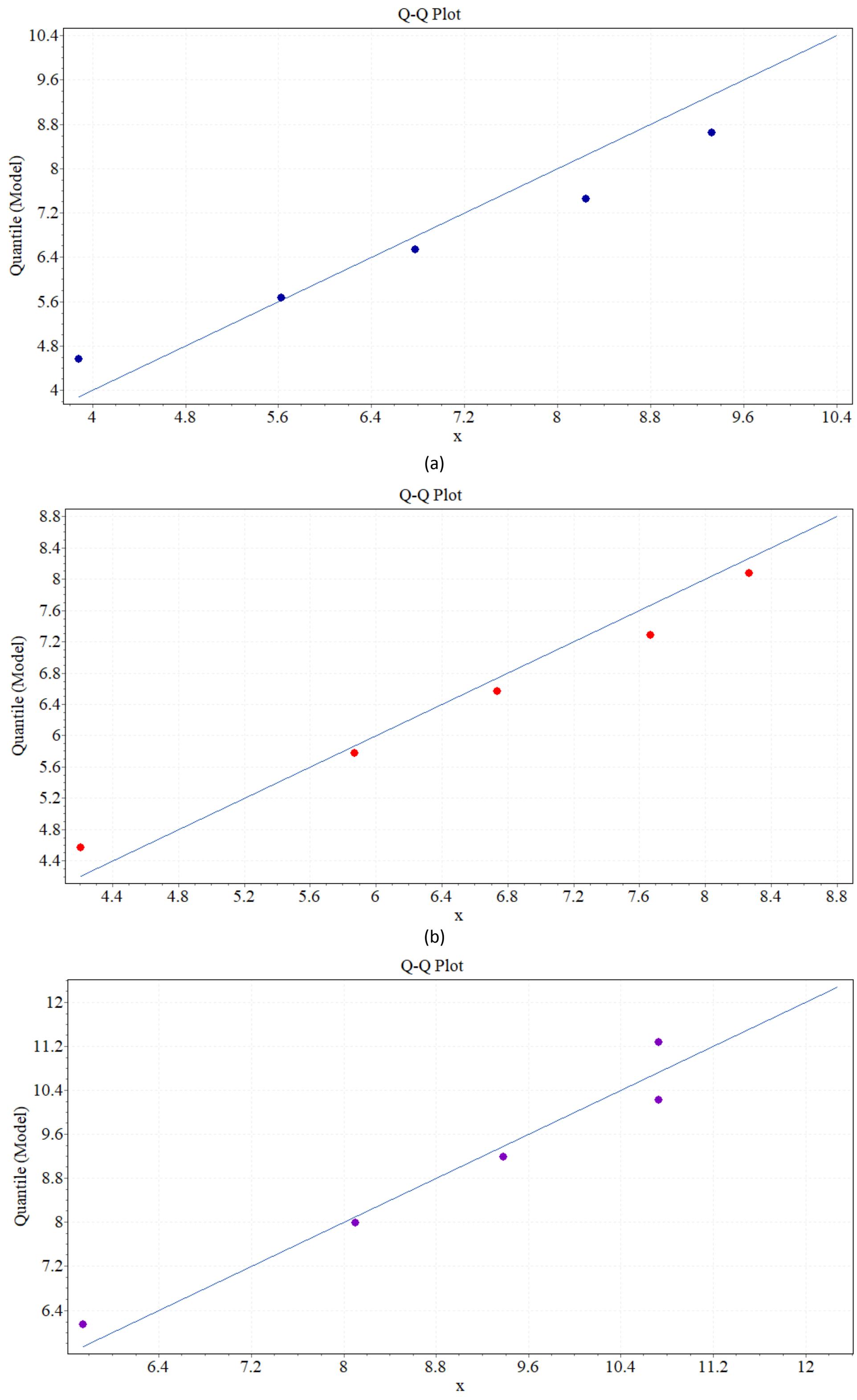
(a)

(b)

(c)

(d)

Figure 1 IDF Curve for Gob station (Stn 5216001): (a) Empirical distribution; (b) Gumbel distribution; (c) Normal distribution; (d) Log Pearson Type III distribution



(c)

Figure 2 Q-Q plots for Gob station at 30 minutes (Stn 5216001): (a) Gumbel distribution; (b) Normal distribution; (c) Log Pearson Type III distribution

Table 2 Percentage difference for Brook, Blau, Gob, Kg. Lalok and JPS Kuala Krai station

STATION	ARI PERIOD (YEARS)	EMPIRICAL DISTRIBUTION				GUMBEL DISTRIBUTION				PERCENTAGE DIFFERENCE (%)			
		RAINFALL INTENSITY (mm/hr)											
		STORM DURATION (MIN)											
		30	360	720	1440	30	360	720	1440	30	360	720	1440
BROOK	2	82.33	14.59	8.53	4.96	65.688	10.862	6.010	3.289	20%	26%	30%	34%
	5	109.97	19.48	11.39	6.62	78.555	13.509	7.675	4.080	29%	31%	33%	38%
	10	136.88	24.25	14.18	8.24	87.042	15.256	8.773	4.603	36%	37%	38%	44%
	25	182.84	32.39	18.94	11.01	97.836	17.477	10.170	5.267	46%	46%	46%	52%
	50	227.59	40.32	23.58	13.70	105.821	19.121	11.204	5.758	54%	53%	52%	58%
	100	283.31	50.19	29.35	17.06	113.747	20.752	12.229	6.245	60%	59%	58%	63%
GOB	2	80.68	13.87	8.13	4.74	68.719	11.907	6.825	3.879	15%	14%	16%	18%
	5	104.56	17.98	10.54	6.14	86.814	15.658	9.084	5.626	17%	13%	14%	8%
	10	127.21	21.87	12.82	7.48	98.752	18.133	10.574	6.779	22%	17%	17%	9%
	25	164.85	28.34	16.61	9.69	113.933	21.280	12.470	8.245	31%	25%	25%	15%
	50	200.57	34.48	20.21	11.79	125.163	23.608	13.871	9.329	38%	32%	31%	21%
	100	244.02	41.95	24.59	14.34	136.310	25.920	15.263	10.406	44%	38%	38%	27%
KG. LALOK	2	108.39	19.84	11.51	6.62	77.085	13.609	8.525	5.503	29%	31%	26%	17%
	5	153.21	28.04	16.27	9.35	94.764	17.281	11.368	8.006	38%	38%	30%	14%
	10	199.07	36.43	21.13	12.15	106.427	19.704	13.243	9.656	47%	46%	37%	21%
	25	281.38	51.50	29.87	17.18	121.259	22.785	15.628	11.756	57%	56%	48%	32%
	50	365.59	66.91	38.82	22.32	132.231	25.064	17.392	13.309	64%	63%	55%	40%
	100	475.01	86.94	50.43	29.00	143.121	27.326	19.144	14.850	70%	69%	62%	49%
JPS KUALA KRAI	2	102.94	20.61	12.21	7.17	69.776	13.351	9.213	6.032	32%	35%	25%	16%
	5	158.07	31.65	16.27	11.00	90.103	18.420	13.747	8.923	43%	42%	15%	19%
	10	218.66	43.78	25.94	15.22	103.513	21.763	16.739	10.830	53%	50%	35%	29%
	25	335.77	67.23	39.83	23.37	120.565	26.016	20.543	13.255	64%	61%	48%	43%
	50	464.47	93.00	55.10	32.33	133.180	29.161	23.357	15.049	71%	69%	58%	53%
	100	642.49	128.64	76.21	44.72	145.702	32.283	26.151	16.830	77%	75%	66%	62%

Table 3 Score Table for Goodness of Fit Test

STATION	STORM DURATION (MIN)	GOODNESS OF FIT TEST					
		K-S			MANN-WHITNEY		
		SCORE					
		GUMBEL	NORMAL	LOG PEARSON TYPE III	GUMBEL	NORMAL	LOG PEARSON TYPE III
BROOK	30	1	0	0	0	0	0
	60	1	0	1	0	0	0
	120	1	1	1	1	1	1
	240	1	0	1	1	0	1
	360	1	0	1	0	0	0
	540	0	0	0	0	0	0
	720	0	0	1	0	0	0
	1080	0	0	1	0	0	0
	1440	0	0	0	0	0	0
BLAU	30	1	0	1	1	0	1
	60	1	0	1	1	0	1
	120	1	0	1	1	0	1
	240	1	0	1	1	0	1

	360	1	0	1	1	0	1
	540	1	0	1	1	0	1
	720	1	0	1	1	0	1
	1080	1	0	1	1	0	1
	1440	1	0	1	1	0	1
GOB	30	1	1	1	1	1	1
	60	1	1	1	1	1	1
	120	1	1	1	1	1	1
	240	1	1	1	1	1	1
	360	1	1	1	1	1	1
	540	1	1	1	1	1	1
	720	1	1	1	1	1	1
	1080	1	1	1	1	1	1
	1440	1	1	1	1	1	1
KG. LALOK	30	1	1	1	1	1	1
	60	1	1	1	1	1	1
	120	1	1	1	1	1	1
	240	1	1	0	1	1	0
	360	1	1	0	1	1	0
	540	1	1	1	1	1	1
	720	1	1	0	1	1	0
	1080	1	1	1	1	1	1
	1440	1	1	1	1	1	1
JPS KUALA KRAI	30	1	1	1	1	1	1
	60	1	1	1	1	1	1
	120	1	1	1	1	1	1
	240	1	1	1	1	1	1
	360	1	1	0	1	1	0
	540	1	1	1	1	1	1
	720	1	1	0	1	1	0
	1080	1	1	1	1	1	1
	1440	1	1	1	1	1	1
KG. JELI	30	0	0	0	0	0	0
	60	0	0	0	0	0	0
	120	0	0	0	0	0	0
	240	0	0	0	0	0	0
	360	0	0	0	0	0	0
	540	0	0	0	0	0	0
	720	1	0	0	0	0	0
	1080	1	0	0	0	0	0
	1440	1	1	0	1	0	0
RUMAH KASTAM RANTAU PJG	30	1	0	0	1	0	0
	60	1	0	0	0	0	0
	120	1	0	0	1	0	0
	240	1	0	0	1	0	0
	360	1	0	0	1	0	0
	540	1	0	0	1	0	0
	720	1	0	0	1	0	0
	1080	1	0	0	1	0	0
	1440	1	0	1	1	0	1
SETOR JPS KOTA BAHRU	30	1	0	0	0	0	0
	60	1	0	0	1	0	0
	120	1	0	1	0	0	0
	240	1	1	1	1	1	1
	360	1	1	1	1	1	1
	540	1	1	1	1	1	1
	720	1	1	1	1	1	1
	1080	1	0	0	1	0	0
	1440	1	0	0	1	0	0
TOTAL SCORE		62	33	43	54	32	38
PERCENTAGE OF PASSING RATE (%)		86	46	60	75	44	53

Notes: 1 indicates pass and 0 indicates unaccepted

4.0 CONCLUSIONS

In summary, the development of the IDF curve for the Kelantan River Basin using frequency analysis was carried out successfully. The Gumbel, Normal and Log-Pearson type III distributions were fitted into annual maximum series of durations between 30 minutes and 24 hours. In addition to that, the Kolmogorov-Smirnov, Mann-Whitney and Quantile-Quantile (Q-Q) goodness of fit tests were utilized to identify the extend of the goodness of fit between the rainfall intensities developed with the empirical distribution and the theoretical distribution. Finally, the best distribution will be chosen based on the results of score table. In general, rainfall estimates for all IDF curves were increasing with increasing return period. Rainfall intensities were decreasing with rainfall durations in all return periods varying from 2 to 100 years. The Gumbel distribution showed that the majority of the p-values generated by both K-S and Mann-Whitney were greater than 5% of significance level which results in acceptance of null hypothesis (that is, the two samples follow the same distribution). Nevertheless, all Q-Q plots for the Gumbel, Normal and Log Pearson Type III distributions depicted a good fit with similar graph trend. Last but not least, the Gumbel distribution showed the highest percentage of passing rate for both the K-S and Mann-Whitney tests (86% and 75% respectively). The Gumbel distribution is recommended as the most suitable distribution in Kelantan River Basin, followed by Log Pearson Type III and Normal distributions. It is believed that Gumbel distribution would contribute to the design engineers effective design and utilization of infrastructure facilities which results in protection of public and cost saving.

This study covered only eight rainfall stations in Kelantan River Basin, thus, the results produced may not demonstrate a complete indication of the whole river basin. Further studies should be investigated by covering more rainfall stations and application of Artificial Neural Network (ANN) is recommended. ANN has a great advantage in handling large amount of data sets such as rainfall data and capable of presenting a better model of IDF curve without manual calculation, hence human errors can certainly be avoided.

Acknowledgements

The authors would like to express their sincere gratitude to the Department of Irrigation and Drainage (DID) Malaysia for providing the daily rainfall data for the Kelantan River Basin, Malaysia.

References

[1] Adnan, N. A., Atkinson, P. M., Yusoff, Z. M., & Rasam, A. R. A. 2014. Climate variability and anthropogenic impacts on a semi-distributed monsoon catchment runoff simulations. In 2014 IEEE 10th International Colloquium on Signal Processing and its Applications. 178-183. IEEE.

[2] Alghazali, N. O., & Alawadi, D. A. 2014. Fitting statistical distributions of monthly rainfallfor some Iraqi stations. *Civil and Environmental Research*. 6(6): 40-46.

[3] Asma'Suhaimee, N. 2018 The Development of Intensity-duration Frequency Curves for Pahang (Doctoral dissertation, UMP).

[4] Chang, K. B., Lai, S. H., & Othman, F. 2016. Comparison of annual maximum and partial duration series for derivation of rainfall intensity-duration-frequency relationships in Peninsular Malaysia. *Journal of Hydrologic Engineering*. 21(1): 05015013.

[5] Elsebaie, I. H. 2012. Developing rainfall intensity–duration–frequency relationship for two regions in Saudi Arabia. *Journal of King Saud University-Engineering Sciences*. 24(2): 131-140.

[6] Galoie, M., Zenz, G., & Eslamian, S. 2013. Application of L-moments for IDF determination in an Austrian basin. *International Journal of Hydrology Science and Technology*. 3(1): 30-48.

[7] Government of Irrigation and Drainage Malaysia, 2012. Urban Stormwater Management Manual for Malaysia.

[8] Gubareva, T. S. 2011. Types of probability distributions in the evaluation of extreme floods. *Water Resources*. 38(7): 962-971.

[9] Huang, Y. F., Mirzaei, M., & Amin, M. Z. M. 2016. Uncertainty quantification in rainfall intensity duration frequency curves based on historical extreme precipitation quantiles. *Procedia Engineering*. 154: 426-432.

[10] Jefrin, N., Bolong, N., Sentian, J., Abustan, I., Mohammad, T. A., & Ayog, J. L. 2017. The Development of Intensity-Duration Frequency Curve for Ulu Moyog and Kaiduan Station of Sabah. *Transactions on Science and Technology*. 4(2): 149-156.

[11] Kaboosi, K., & Jelini, R. 2017. The efficiency of detention reservoirs for flood control on the Jafar Abad River in Golestan province (Iran). *Russian Meteorology and Hydrology*. 42(2): 129-134.

[12] Kuchment, L. S., & Demidov, V. N. 2013. On the application of copula theory for determination of probabilistic characteristics of springflood. *Russian Meteorology and Hydrology*. 38(4): 263-271.

[13] Lee, C. 2005. Application of rainfall frequency analysis on studying rainfall distribution characteristics of Chia Nan plain area in Southern Taiwan. *Crop, Environment & Bioinformatics,* 2(1): 31-38

[14] Leščešen, I., Urošev, M., Dolinaj, D., Pantelić, M., Telbisz, T., Varga, G. & Milošević, D. 2019. Regional Flood Frequency Analysis Based on L-Moment Approach (Case Study Tisza River Basin). *Water Resources*. 46(6): 853-860.

[15] Liew, S. C., Raghavan, S. V., & Liong, S. Y. 2014. How to construct future IDF curves, under changing climate, for sites with scarce rainfall records?. *Hydrological Processes*. 28(8): 3276-3287.

[16] Millington, N., Das, S., & 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.

[17] Mirzaei, M., Huang, Y. F., Lee, T. S., El-Shafie, A., & Ghazali, A. H. 2014. Quantifying uncertainties associated with depth duration frequency curves. *Natural Hazards*. 71(2): 1227-1239.

[18] Mohd Ariff, N., Jemain, A. A., & Wan Zin, W. Z. 2013. Assessment of the intensity-duration-frequency (IDF) curves for storms in Peninsular Malaysia based on the generalized extreme value distribution. In *AIP Conference Proceedings*, 1522(1): 1185-1194. American Institute of Physics.

[19] National Hydraulic Research Institute of Malaysia. 2010. Reviewed and Updated Version of Hydrological Procedure No. 1.

[20] Ng, C. K., Ng, J. L., Huang, Y. F., Xun, T. Y., & Mirzaei, M. 2020. Tropical rainfall trend and stationarity analysis. *Water Supply*. 20(7): 2471-2483

[21] Ng, J. L., Abd Aziz, S., Huang, Y. F., Mirzaei, M., Wayayok, A., & Rowshon, M. K. 2019. Uncertainty analysis of rainfall depth duration frequency curves using the bootstrap resampling technique. *Journal of Earth System Science*. 128(5): 113.

[22] Ng, J. L., Abd Aziz, S., Huang, Y. F., Wayayok, A., & Rowshon, M. K. 2016. Analysis of annual maximum rainfall in Kelantan, Malaysia. In *III International Conference on Agricultural and Food Engineering*, 1152:11-18.

[23] Ng, J. L., Abd Aziz, S., Feng, H. Y., Wayayok, A., & Kamal, M. R. 2015. Homogeneity analysis of rainfall in Kelantan, Malaysia. *Jurnal Teknologi*, 76(15): 1-6.

[24] Nhat, L. M., Tachikawa, Y., & Takara, K. 2006. Establishment of intensity-duration-frequency curves for precipitation in the monsoon area of Vietnam. *Annuals of Disaster Prevention Research Institute, 93-103.*

[25] Pradhan, B., & Youssef, A. M. 2011. A 100-year maximum flood susceptibility mapping using integrated hydrological and hydrodynamic models: Kelantan River Corridor, Malaysia. *Journal of Flood Risk Management*, 4(3): 189-202.

[26] Wagesho, N., & Claire, M. 2016. Analysis of Rainfall Intensity-Duration-Frequency Relationship for Rwanda. *Journal of Water Resource and Protection*, 8(07): 706.