Applied Mathematics and Computational Intelligence
Volume 11, No. 2, Dec 2022 [534 − 549]

At-Site and Regional Frequency Analysis of Extreme Rainfall Modelling in Peninsular Malaysia

Siti Nur Farah Atiqah Bakri[1], Norhaslinda Ali[2*], Syafrina Abdul Halim[3], Iszuanie Syafidza Che Ilias[4], Nazihah Mohamed Ali[5]

[1,2,3,5]Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
[4]Institute for Mathematical Research, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.

* Corresponding author: norhaslinda@upm.edu.my

Received: 27 October 2022; Accepted: 1 December 2022; Available online (in press): 15 December 2022

ABSTRACT

The objective of this study is to determine the appropriate probability distribution and to estimate the rainfall quantiles for monthly maxima daily rainfall data from the year 2005 to 2019 for 30 rain gauge stations in Peninsular Malaysia based on at-site and regional hydrological frequency analysis. Five three-parameters probability distributions were considered in this study i.e generalized extreme value (GEV), generalized Pareto (GPA), generalized logistic (GLO), generalized normal (GNO) and Pearson Type III (PE3). Cluster analysis based on Ward's method was used to identify the homogeneous region which is further confirmed by discordancy and heterogeneity measures. The L-moment method of estimation is used to estimate the parameter of a model. The L-moment ratio diagrams and Monte Carlo simulation based on Z^{DIST} *were used to assess the goodness of the fitted model. Results obtained by traditional at-site frequency analysis are compared with those obtained by regional frequency analysis. The results showed that the best probability distribution for monthly maxima daily rainfall data at each station and the ones of corresponding homogeneous regions obtained by regional frequency analysis were not necessarily consistent. Although the optimal probability distribution may vary according to the stations, in most cases, data for most of the stations are found to follow the generalized logistic distribution while for the regional study, rainfall data for most of the regions are well fitted by the generalized extreme value distribution. Meanwhile, the uncertainty due to quantile estimates for at-site and regional data is considerably low for less than 100 years return period but high for more than 100 years.*

Keywords: At-site frequency analysis, quantile estimates, L-moments, probability distribution, regional frequency analysis.

1 INTRODUCTION

Malaysia has an equatorial climate with high temperatures and relative humidity and luckily being exempted from many severe natural disasters. However, Malaysia is not excluded in experiencing flood, droughts, and rainstorms in every consecutive year. Due to the influence of global warming, the magnitude and pattern of precipitation extremes are expected to change. In particular, the frequency of heavy precipitation events increased over most areas. Extreme precipitation has the potential to trigger floods and droughts, which is expected to put considerable pressure on water resources.

Sustainable water resources management, planning for weather-related emergencies, and design of hydraulic structures requires knowledge on the magnitudes and frequency of extreme precipitation [1]. Statistical methods are always applied to past events to predict the exceedance probability of future events in an attempt to reduce the risk and maximize the efficiency in hydraulic design [2]. However, the estimation of the frequencies of extreme events is difficult as their events were rare, and their records were often short. Reliable estimations require very long station records if single station data are to be used [1].

Regional frequency analysis (RFA) has been frequently used for estimating design quantiles for extreme hydrological events like floods and heavy rainfalls. RFA was proposed by [3] for pooling various data sample. In pooling approach, the data samples analyzed are typically observations of the same variable at several measuring sites within a suitably defined region. The principles of regional frequency analysis, however, apply whenever multiple samples of similar data are available. This study uses extreme rainfall homogeneous regions which refer to regions that contain sites with similar characteristics of extreme rainfall data such as means, skewness and kurtosis. This indicates that the areas within the homogeneous regions have similar conditions, climatic exposure, and source of extreme rainfalls.

The choice of an appropriate probability distribution for at-site and regional frequency analysis remains of immense importance. Several probability distributions are recommended for at-site and regional frequency analysis in various countries [4–6]. Zalina et. al [7] found that GEV distribution is the most appropriate distribution for hourly rainfall of seventeen automatic stations in Malaysia. On the other hand, [8] stated that generalized extreme value and generalized logistic were identified as the best distributions for modelling daily annual maximum rainfall in Selangor, Malaysia. In addition, [9] stated that generalized extreme value, generalized Pareto and generalized logistic distributions have been extensively used in extreme value estimation of the annual flood peaks. In Southern Africa, [10] conducted a regional study using annual maximum flood data comprising 407 gauging stations from eleven countries. The Pearson Type III with probability weighted moment was selected as the optimal distribution for almost half of the 56 identified regions followed by Log-Pearson Type III using method of moments. Saf [11] conducted the study on annual maximum flood flows from 47 sites in the West Mediterranean Rivers basins of Turkey. Based on goodness-of-fit test statistic, Z, they identified Pearson Type III as the best-fitted model for Antalya and Lower-West Mediterranean while generalized logistic for Upper-West Mediterranean. While, in Turkey, [12] has found that generalized extreme value and lognormal distributions as the candidates for regional parent distributions for the maximum daily rainfall. [13] analyzed the data of the annual maximum stream flow of five gauging sites of Torne River in Sweden, where five distributions were compared using various goodness of fit and accuracy measures with maximum likelihood and L-moments method of estimations. It was found that three parameters log normal distribution with L-moments method of estimation is the most appropriate distribution compared to the other three.

In the present study, an attempt has been made to determine the appropriate probability distribution for monthly maxima daily rainfall data from the year 2005 to 2019 for 30 rain gauge stations in Peninsular Malaysia based on at-site and regional hydrological frequency analysis. Then, the at-site and regional quantile rainfall were estimated using the optimal distribution obtained by L-moments method of estimation.

2 MATERIALS AND METHODS

2.1 Data and Study Area

The study is carried out for the catchments of Peninsular Malaysia watershed. A monthly maxima daily rainfall data from 30 rain gauge stations in Peninsular Malaysia with record lengths of 15 years was obtained from the Department of Irrigation and Drainage, Malaysia. The data contains measurements of daily rainfalls in millimetres from year 2005 until 2019. Figure 1 shows the all of 30 rain gauge stations that are located at various places throughout the Peninsular Malaysia.

Figure 1 : Location of rain gauge stations used in this study.

2.2 Probability Distribution

In order to describe the behaviour of extreme rainfall at a particular site or region, it is necessary to identify the distribution that best fits the data. Previous study have showed that generalized extreme value (GEV), generalized Pareto (GPA), generalized logistic (GLO), generalized normal

(GNO) and Pearson Type III (PE3) are among the best fitted distribution for at-site and regional frequency study [7–12]. Therefore, these five three-parameters probability distributions will be used for a comparison between at-site and regional frequency analysis study.

2.3 L-moments

L-moments method of estimation is a modification of a probability weighted moments proposed by [14]. They are a linear combinations of first order statistics and are hence more robust to measurement errors or sampling uncertainty than conventional moments [15]. The rth population L-moments of a random variable X is defined as

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} E(X_{r-k:r}); \qquad r = 1, 2, ... \tag{1}$$

where $X_{r-k:r}$ is the random variable for $(r-k)$th order statistics and E denotes the expected value. The first four population L-moments for a random variable X are given by

$$\lambda_1 = E(X) \tag{2}$$

$$\lambda_2 = \frac{1}{2}E(X_{2:2} - X_{1:2}) \tag{3}$$

$$\lambda_3 = \frac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3}) \tag{4}$$

$$\lambda_4 = \frac{1}{3}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) \tag{5}$$

The higher L-moments λ_r, $r \geq 3$ usually defined by a ratio of L-moments given by

$$r_r = \frac{\lambda_r}{\lambda_2}, \qquad r = 3, 4, ... \tag{6}$$

The location and dispersion of a random variable X is measured by the first two L-moments, λ_1 and λ_2 whereas the third and forth ratio L-moments, r_3 and r_4 measure the skewness and kurtosis of a random variable X respectively.

The sample estimates of L-moments are obtained by using the unbiased sample estimator of β_r [15] given by

$$b_r = \frac{1}{n}\frac{\sum_{i=1}^{n}\binom{i-1}{r} x_{i:n}}{\binom{n-1}{r}} \tag{7}$$

where $x_{i:n}$ is an ordered n observations of a random variable X. Therefore, the first four sample L-moments are given by

$$\ell_1 = b_0 \tag{8}$$

$$\ell_2 = 2b_1 - b_0 \tag{9}$$

$$\ell_3 = 6b_2 - 6b_1 + b_0 \tag{10}$$

$$\ell_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \tag{11}$$

2.4 Regional Frequency Analysis

Regional frequency analysis using L-moments as outlined by [16] involves four steps: (1) Data screening, (2) Identification of homogeneous region, (3) Choice of optimal frequency distribution and (4) Regional growth curve development.

2.4.1 Data screening

As with any statistical analysis, the first stage in RFA is a close inspection of the data. Gross errors and inconsistencies should be eliminated and a check made that the data are homogeneous over time. The objective is to check the suitability of the data for carrying out the regional analysis. To determine an unusual station for each region in this study, the discordancy measure, D_i was used and defined as

$$D_i = \frac{1}{3}N(c_i - \bar{c})^T S^{-1}(c_i - \bar{c}) \tag{12}$$

where $c_i = (t^{(i)}, t_3^{(i)}, t_4^{(i)})^T$ represents the sample L-moment ratios for site i, N is the total sites in the region, $\bar{c} = N^{-1}\sum_{i=1}^{N} u_i$ denotes the region's unweighted average L-moments ratio, and $S = \sum (c_i - \bar{c})^T S^{-1}(c_i - \bar{c})^T$ is the sample covariance matrix. The site is identified as discordant and discarded from the analysis if D_i for site i exceeded the critical value. $D_i \geq 3$ was used for $N \geq 15$ sites.

2.4.2 Identification of homogeneous region

The formation of the homogeneous regions was initiated by choosing the candidate homogeneous regions based on cluster analysis using Ward's method which originally presented by [17]. The information of the stations such as latitude, longitude, mean annual precipitation and the average number of rainy days were used in the cluster analysis. Two statistics, discordancy (as discussed in Section 2.4.1) and heterogeneity measures were used to test the homogeneity of the candidate homogeneous region. The heterogeneity measure proposed by [16] was aims to assess the amount of heterogeneity in a set of hydrological sites and to judge whether regions might be considered as homogeneous regions. The heterogeneity measure was calculated as

$$H = \frac{(V - \mu_v)}{(\sigma_v)} \tag{13}$$

where μ_v and σ_v are respectively the mean and standard deviation of simulated V values. The statistics V is given by

$$V = \left\{ \frac{\sum_{i=1}^{N} n_i (t^{(i)} - t^R)^2}{\sum_{i=1}^{N} n_i} \right\}^{1/2} \tag{14}$$

where t^R, t_3^R and t_4^R are L-moments ratios for a given region. A region is acceptably homogeneous if $H < 1$, possibly heterogeneous if $1 \leq H < 2$ and definitely heterogeneous if $H \geq 2$ [18].

2.4.3 Choice of optimal frequency distribution

The L-moment ratio diagram and goodness of fit test based on Z^{DIST} statistic value are used to determine the optimal frequency distribution for rainfall data in specific region [15, 19–22]. The latter method also has been used to choose the at-site optimal frequency distribution. The L-moment diagram is a graphical plot between L-skewness and L-kurtosis which compare the sample L-moment ratios to the theoretical values from GLO, GEV, GNO, PE3 and GPA distributions. The closest the theoretical curve to the sample values indicate that the distribution can fits the data well.

The Z^{DIST} statistic measures how well the regional average L-kurtosis of the observed data matches with the theoretical L-kurtosis of the candidate distribution i.e GLO, GEV, GNO, PE3 or GPA. For each candidate distribution, the statistic Z^{DIST} is calculated by

$$Z^{DIST} = \frac{t_4^R - \tau_4^{DIST}}{\sigma_4} \tag{15}$$

where t_4^R is an average L-kurtosis value computed from regional data, τ_4^{DIST} is a theoretical L-kurtosis value computed from the simulation of the fitted distribution and σ_4 is the standard deviation of L-kurtosis value obtained from simulated data. The best fitted distribution is chosen when the value of $|Z^{DIST}| \leq 1.64$ or choose the smallest one when more than one distribution qualifies for the goodness of fit measure.

2.4.4 Quantile Estimation

The most interest part in RFA is the estimation of regional quantile of nonexceedance probability F, of various return periods for each homogeneous region. For at-site rainfall quantiles estimate, $\hat{Q}_i(F)$, the index flood method [16] is used to estimate the quantile using

$$\hat{Q}_i(F) = l_i \hat{q}(F) \tag{16}$$

where l_i is the average monthly maxima rainfall at site i and $\hat{q}(F)$ is the regional quantiles estimate.

3 RESULTS AND DISCUSSION

Table 1 gives the descriptive statistics for monthly maximum daily rainfall for the 30 rain gauge stations. The mean amount of monthly maximum rainfall is ranging from 39.82 mm in station Chaah, Johor to 108.02 mm in station Bukit Tandak, Kelantan with standard deviations are respectively 26.361 mm and 132.243 mm. It can be noted that Station Bukit Tandak received the greatest amount of monthly maximum rainfall which is 943.0 mm while station Ulu Kinta received the least with only 99 mm.

Table 1 : Descriptive statistics of monthly maximum daily rainfall data

No	Station	Mean	Std. dev	Max.	No	Station	Mean	Std. dev	Max.
1	Ngolang	40.36	27.417	189.5	16	Pekan Merlimau	44.64	27.855	162.0
2	Padang Besar	42.63	30.464	220.0	17	Chin Chin	44.04	23.801	155.5
3	Padang Katong	41.79	29.182	123.5	18	Johor Bahru	57.46	32.437	216.0
4	Baling	43.14	24.473	194.5	19	Parit Madirono	50.51	26.554	143.5
5	Air Itam	59.05	39.794	338.5	20	Simpang Mawai	57.23	39.560	226.0
6	Bukit Bendera	59.42	37.608	253.0	21	Chaah	39.82	26.361	214.0
7	Tanjung Malim	53.97	29.589	160.5	22	Benta	40.17	22.913	120.5
8	Ulu Kinta	46.99	23.631	99	23	Bukit Betong	46.63	27.648	146.0
9	Kalong Tengah	60.25	24.627	146.0	24	Chalok	75.74	59.998	350.0
10	Semenyih	49.74	23.564	138.0	25	Ulu Setiu	90.66	70.776	474.0
11	Edinburgh	62.80	27.881	151.5	26	Ulu Dungun	75.81	69.444	360.5
12	Genting Klang	56.47	25.983	138.0	27	Kuala Terengganu	61.32	48.263	262.0
13	Politeknik PD	48.95	31.186	256.0	28	Lalok	52.88	39.893	222.0
14	Chengkau	48.31	25.529	109.5	29	Bukit Tandak	108.02	132.243	943.0
15	Mantin	46.509	25.437	148.0	30	Upper Chiku	61.36	40.055	370.5

The first step in RFA is to perform the discordant test for data screening. By analyzing monthly maximum data, it is observed that the value of statistics D for station Bukit Tandak and Upper Chiku are exceeding three. It can be concluded that data from these two stations are discordant from the rest of the regional data. Data from these stations are excluded and the discordant test is performed again using the remaining 28 stations. The result shows that none of these 28 stations have value D exceed three. Therefore, all the 28 stations will be used to identify the homogeneous region using cluster analysis with ward's method.

Figure 2 shows the result of Ward's clustering with 11 clusters which represent regions. These preliminary regions have to be investigated in terms of their discordant and heterogeneity measure. The analysis showed that 4 out of 11 regions (G4, G5, G6 and G7) are heterogeneous even though all of them are not discordant with the others in their regions. Stations under these regions are merged with other stations in other regions and the cluster analysis is repeated until the homogeneous region is obtained. As a result, six regions have fulfilled the discordant and heterogeneity measure where all the stations within each region have the discordant measure less than three and H value less than two. The resulting regions with corresponding stations are shown in Table 2. As shown in Table 2, regions 1 and 6 respectively consist of three stations located in the northern part, while region 4 has 8 stations where most of which are in the southern and middle parts of Peninsular Malaysia.

The optimal frequency distribution for at-site and regional frequency analysis for monthly maxima daily rainfall based on Z^{DIST} values are presented in Table 3. As we can see from Table 3, the optimal frequency distribution for at-site and regional part are based on the minimum value of $|Z^{DIST}|$ among five considered distributions i.e GEV, GPA, GLO, GNO and PE3. Table 3 gives the summary of number of stations with respect to the optimal frequency distribution found based

Table 2 : Discordance (D_i), critical values of statistics D_{cri} and heterogeneity measure (H) values for monthly maximum daily rainfall for each stations.

Region	Station (D_i)	Critical value of D statistics, D_{cri}	Heterogeneity measure, H
1	Ngolang (1.00) Padang Besar (1.00) Pdg Katong (1.00)	3.00	-0.23
2	Baling (1.22) Air Itam (0.69) Bkt Bendera (0.44) Tjg Malim (1.32) Ulu Kinta (1.32)	1.33	0.69
3	Semenyih (1.00) Politeknik PD (1.00) Chengkau (1.00) Mantin (1.00)	3.00	1.54
4	Benta (0.44) Bkt Betong (0.66) Pkn Merlimau (0.93) Chin Chin (0.69) Johor Bahru (2.06) Parit Madirono (0.41) Simpang Mawai (1.07) Chaah (1.75)	2.14	1.01
5	Chalok (1.33) Ulu Setiu (0.69) Ulu Dungun (0.44) Kuala Terengganu (1.32) Lalok (1.32)	-0.27	0.69
6	Kalong Tgh (1.00) Edinburgh (1.00) Genting Klang (1.00)	3.00	0.85

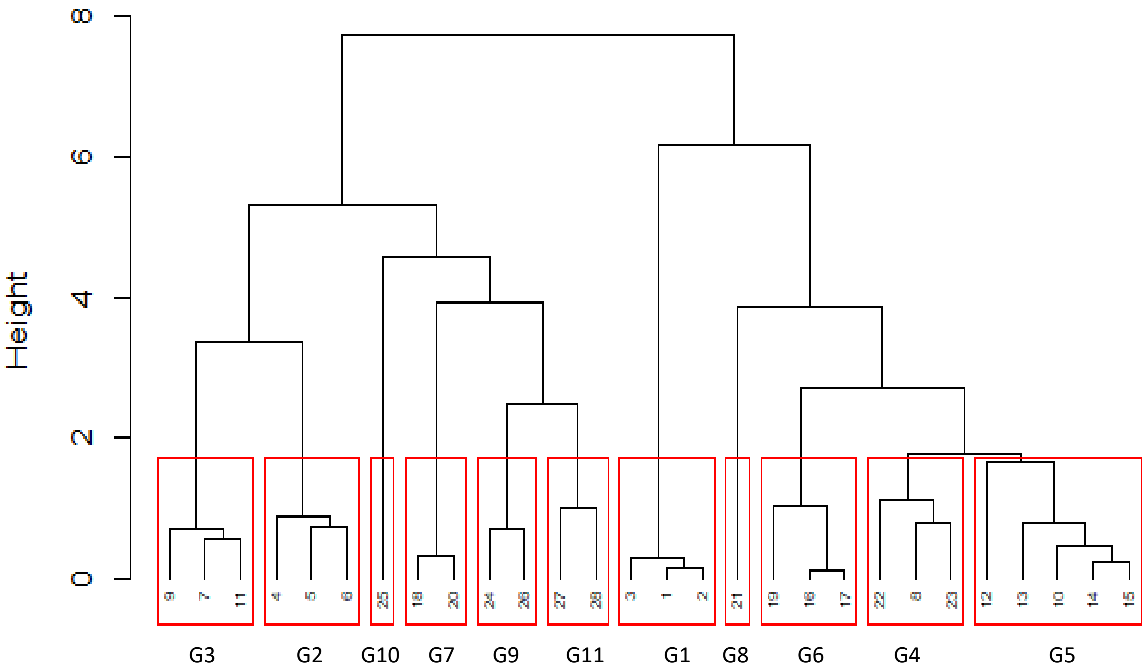

Figure 2 : Dendrogram of clustered stations (in square) by the Ward's method.

on value of Z^{DIST}. It can be noted that the optimal frequency distribution for at-site and it respective region is not necessarily consistent for e.g, in Region 1, at-site frequency distribution for station Ngolang is GLO, Padang Besar is GEV and Padang Katong si PE3. However, the optimal frequency distribution is GEV.

Table 3 : The $|Z^{DIST}|$ values for monthly maximum daily rainfall for each stations. The coloured values are minimum value of $|Z^{DIST}|$ for each station. *The optimal frequency distribution for at-site and regional.

Region	Station	$\|Z^{DIST}\|$					min($\|Z^{DIST}\|$)	At-site*	Regional*
		GLO	GEV	GNO	PE3	GPA			
1	Ngolang	0.24	0.88	1.04	1.45	3.40	0.24	GLO	GEV
	Padang Besar	0.90	0.53	0.63	1.00	3.61	0.53	GEV	
	Pdg Katong	2.86	1.02	0.99	0.62	2.90	0.62	PE3	
2	Baling	1.41	0.38	0.31	0.55	4.06	0.31	GNO	GEV
	Air Itam	0.78	0.04	0.41	1.20	1.90	0.04	GEV	
	Bkt Bendera	0.86	0.00	0.41	1.14	2.17	0.00	GEV	
	Tjg Malim	0.13	0.99	1.17	1.60	3.52	0.13	GLO	
	Ulu Kinta	3.04	0.61	1.08	1.04	3.99	0.61	GEV	
3	Semenyih	2.53	0.52	0.66	0.46	3.56	0.46	PE3	GNO
	Politeknik PD	0.31	1.64	1.66	1.91	4.43	0.31	GLO	
	Chengkau	3.03	0.84	1.00	0.78	3.60	0.78	PE3	
	Mantin	0.92	0.68	0.68	0.96	4.04	0.68	GNO	
4	Benta	0.55	2.03	1.92	2.07	5.02	0.55	GLO	GLO
	Bkt Betong	1.60	0.02	0.09	0.49	3.40	0.02	GEV	
	Pkn Merlimau	0.36	1.65	1.70	1.98	4.41	0.36	GLO	
	Chin Chin	0.48	0.74	0.92	1.36	3.46	0.48	GLO	
	Johor Bahru	0.99	1.96	2.16	2.58	4.19	0.99	GLO	
	Parit Madirono	0.44	0.76	0.94	1.38	3.45	0.44	GLO	
	Simpang Mawai	0.60	1.23	1.57	2.18	2.84	0.60	GLO	
	Chaah	0.38	1.51	1.53	1.76	3.89	0.38	GLO	
5	Chalok	0.48	0.06	0.46	1.37	1.23	0.06	GEV	GEV
	Ulu Setiu	1.05	0.58	0.01	1.04	0.84	0.01	GNO	
	Ulu Dungun	0.23	0.08	0.59	1.48	1.10	0.08	GEV	
	Kuala Terengganu	2.08	1.28	0.69	0.34	0.87	0.34	PE3	
	Lalok	0.27	1.03	1.36	1.97	2.92	0.27	GLO	
6	Kalong Tgh	1.16	0.42	0.50	0.87	3.82	0.42	GEV	PE3
	Edinburgh	2.87	0.94	0.96	0.65	3.10	0.65	PE3	
	Genting Klang	3.01	1.36	1.18	0.67	2.27	0.67	PE3	

Table 4 shows the number of stations with respect to the optimal frequency distribution between at-site and regional based on $|Z^{DIST}|$ values. In Table 4, most of the at-site optimal frequency distribution follow the GLO while the regional optimal frequency distribution follow the GEV distribution. It also can be noted that none of the at-site and regional rainfall data follow a GPA distribution.

Table 4 : The number of stations with respect to the optimal frequency distribution found based on value of Z^{DIST}

Distribution	At-site optimal frequency analysis	Regional optimal frequency analysis
GLO	11	1
GEV	8	3
GNO	3	1
PE3	6	1
GPA	0	0

The L-moments ratio diagrams for monthly maxima daily rainfall data in Figure 3 shows the location of the regional average L-moment with theoretical L-skewness and L-kurtosis relationships for the GLO, GEV, GNO, PE3 and GPA distributions. Figure 3 shows that the point are around GEV distribution for region 1, 2 and 5 while points are scattered around GNO, GLO and PE3 for region 3, 4 and 6 respectively. This graphical inspection for the choice of optimal regional frequency distribution function is consistent with the result found using Z^{DIST}.

The last step in RFA is to estimate the regional quantiles and then estimate the monthly maximum rainfall data for the different return periods. Table 5 shows the regional quantiles estimates associated with the uncertainty measured by root mean square error (RMSE). The value of the quantile estimate can be interpreted as for e.g in region 1, the quantile of GEV with a 0.99 value is 3.101. This is the amount of rainfall that will happen once in 100 years and is 3.101 times larger than the average rainfall for all three rain gauge stations in the homogeneous region 1. It also can be noted that the RMSE values increase when the return period increase which suggest that unrealiability of quantiles for the high return period.

At-site quantile estimates for e.g 50 years return period is obtained by multiplying the stations' average rainfall (l_1) with the regional quantiles of optimal frequency distribution. The result of at-site quantile estimates up to 200 years return period are shown in Table 6. From Table 6, it can be noted that station Ulu Setiu in Terengganu has the highest estimated monthly maximum rainfall for all return periods where the values are almost twice than station Ngolang.

4 CONCLUSION

The present study considers the at-site and regional frequency analysis of extreme rainfall based on monthly rainfall records at 28 rain gauge stations over Peninsular Malaysia. The rain gauge stations were grouped into six homogeneous regions after careful inspection of the discordant and homogeneity tests. Based on the L-moment ratio diagrams and goodness of fit test, Z^{DIST}, GEV was found to be the appropriate distribution that represents the rainfall characteristics in three out of six regions while GLO in almost half of the rain gauge stations. The uncertainty due to quanile estimates for both at-site and regional data are increase as the return period increased suggested that unrealiability estimate especially for more than 100 years return period.

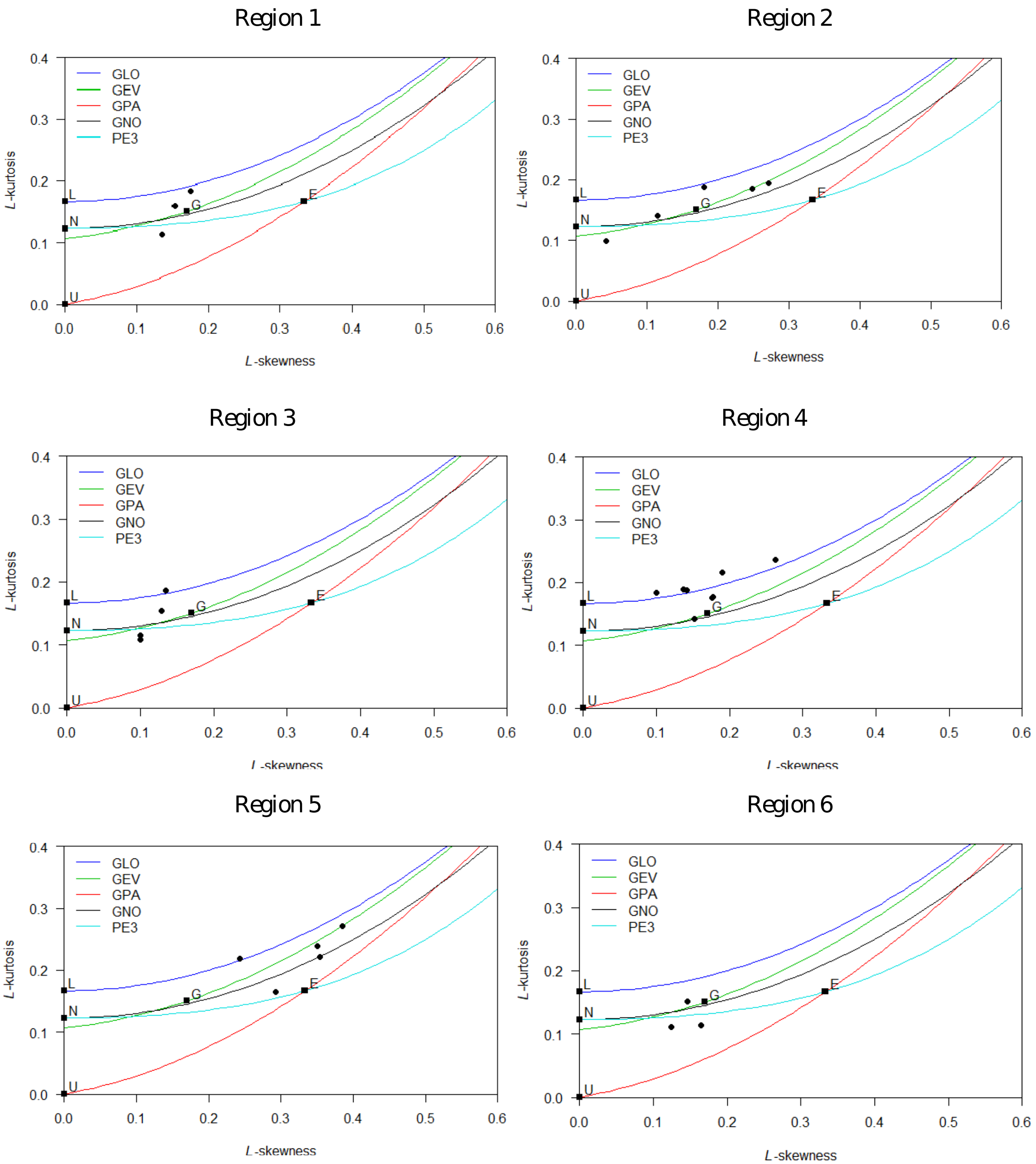

Figure 3 : L-moment ratio diagrams for monthly maxima daily rainfall data at six homogeneous region.

Table 5 : The regional quantiles and accuracy measure using RMSE for the six homogeneous regions.

Region	Distribution	Return Period	f	$q(f)$	RMSE	Error bounds	
						Lower	Upper
1	GEV	5	0.8	1.5085048	0.03012517	1.4613664	1.5583282
		10	0.9	1.9050991	0.05379169	1.8265554	2.0011653
		20	0.95	2.2787675	0.08438805	2.13619174	2.4268872
		50	0.98	2.7528116	0.13623799	2.5304931	2.9937423
		100	0.99	3.1010529	0.18452392	2.8112018	3.4170913
		200	0.995	3.4421808	0.24081493	3.0510029	3.8345819
2	GEV	5	0.8	1.4098955	0.10906062	1.2535704	1.5691820
		10	0.9	1.7470055	0.20182941	1.4624206	2.0467884
		20	0.95	2.0719108	0.29432772	1.6599165	2.5105317
		50	0.98	2.4947202	0.41874394	1.9179481	3.1340676
		100	0.99	2.8132308	0.51536383	2.1058441	3.6202630
		200	0.995	3.1320132	0.61458832	2.2877518	4.1101526
3	GNO	5	0.8	1.4196639	0.02168238	1.3858231	1.4565434
		10	0.9	1.7130937	0.03966286	1.653401	1.7805929
		20	0.95	1.9797577	0.05955888	1.8878063	2.0807846
		50	0.98	2.3088938	0.08831741	2.1702145	2.4649972
		100	0.99	2.5468629	0.11185295	2.3705369	2.7472904
		200	0.995	2.7786487	0.13689008	2.5657423	3.0265843
4	GLO	5	0.8	1.3809989	0.06553572	1.2788782	1.4813701
		10	0.9	1.7093289	0.12248915	1.5220516	1.9000495
		20	0.95	2.0543878	0.18371925	1.778605	2.3411421
		50	0.98	2.5594292	0.2755389	2.1530718	2.992908
		100	0.99	2.9902554	0.35580665	2.4701341	3.5570342
		200	0.995	3.4717069	0.44753996	2.8274092	4.201155
5	GEV	5	0.8	1.4140384	0.0566431	1.3260655	1.5025719
		10	0.9	1.9312344	0.1200819	1.7495069	2.1169772
		20	0.95	2.5180326	0.20587469	2.2105838	2.8515508
		50	0.98	3.4360276	0.36665965	2.9014406	4.0511047
		100	0.99	4.264068	0.53853773	3.4644525	5.1768238
		200	0.995	5.2315137	0.76950752	4.09377432	6.6278756
6	PE3	5	0.8	1.3407923	0.02510296	1.3036534	1.3804737
		10	0.9	1.5919685	0.04596604	1.5168917	1.6711284
		20	0.95	1.8207176	0.06713241	1.7100775	1.9322086
		50	0.98	2.1019161	0.0952996	1.9476422	2.2593396
		100	0.99	2.3036604	0.11673137	2.1160214	2.5031632
		200	0.995	2.4984826	0.13826604	2.2787949	2.7334472

Table 6 : Estimated monthly maximum rainfall (mm) corresponding to different return periods

No.	Station	Non-exceedance probability					
		0.8	0.9	0.95	0.98	0.99	0.995
		5	10	20	50	100	200
1	Ngolang	60.89	76.89	91.98	111.11	125.16	138.93
2	Pdg Besar	64.30	81.21	97.13	117.34	132.18	146.73
3	Pdg Katong	63.04	79.61	95.22	115.03	129.59	143.84
4	Baling	60.83	75.37	89.39	107.63	121.37	135.12
5	Air Itam	83.26	103.17	122.35	147.32	166.13	184.95
6	Bkt Bendera	83.78	103.81	123.11	148.24	167.16	186.11
7	Tjg Malim	76.09	94.28	111.82	134.64	151.83	169.03
8	Ulu Kinta	66.26	82.10	97.37	117.24	132.21	147.19
9	Semenyih	70.62	85.22	98.48	114.85	126.69	138.22
10	Politeknik PD	69.49	83.85	96.90	113.01	124.66	136.01
11	Chengkau	68.59	82.77	95.65	111.55	123.05	134.25
12	Mantin	66.03	79.67	92.08	107.38	118.45	129.23
13	Benta	55.48	68.67	82.53	102.82	120.13	139.47
14	Bkt Betong	64.40	79.71	95.80	119.35	139.44	161.89
15	Pekan Merlimau	61.65	76.31	91.71	114.26	133.49	154.99
16	Chin Chin	60.82	75.28	90.48	112.72	131.70	152.90
17	Johor Bahru	79.35	98.22	118.04	147.06	171.82	199.48
18	Parit Madirono	69.75	86.33	103.76	129.26	151.02	175.34
19	Simpang Mawai	79.03	97.82	117.57	146.47	171.13	198.68
20	Chaah	54.99	68.07	81.81	101.92	119.07	138.25
21	Chalok	107.10	146.27	190.71	260.24	322.95	396.22
22	Ulu Setiu	128.20	175.09	228.29	311.52	386.60	474.31
23	Ulu Dungun	107.19	146.40	190.88	260.47	323.24	396.58
24	Kuala Terengganu	86.71	118.43	154.42	210.71	261.49	320.82
25	Lalok	74.78	102.13	133.16	181.71	225.50	276.66
26	Kalong Tengah	80.79	95.92	109.70	126.65	138.80	150.54
27	Edinburgh	84.20	99.98	114.34	132.00	144.67	156.91
28	Genting Klang	75.71	89.90	102.81	118.69	130.08	141.08

ACKNOWLEDGEMENT

The authors are indebted to the staff of the Department of Irrigation and Drainage, Malaysia for providing daily rainfall data to make this paper possible.

REFERENCES

[1] H. Malekinezhad and A. Zare-Garizi, "Regional frequency analysis of daily rainfall extremes using l-moments approach," *Atmósfera*, vol. 27, no. 4, pp. 411–427, 2014.

[2] J. Smithers and R. Schulze, "A methodology for the estimation of short duration design storms in south africa using a regional approach based on l-moments," *Journal of Hydrology*, vol. 241, no. 1-2, pp. 42–52, 2001.

[3] T. Dalrymple, "Flood-frequency analyses, manual of hydrology: Part 3," USGPO,, Tech. Rep., 1960.

[4] B. Sevruk and H. Geiger, "Selection of distribution types for extremes of precipitation. operational hydrology report 15," *World Meteorological Organization. Geneva. Switzerland. WMO*, no. 560, 1981.

[5] C. Cunnane, "Statistical distributions for flood frequency analysis," *Operational hydrology report (WMO)*, 1989.

[6] A. Castellarin, S. Kohnová, L. Gaál, A. Fleig, J. Salinas, A. Toumazis, T. Kjeldsen, and N. Macdonald, "Review of applied statistical methods for flood frequency analysis in europe," 2012.

[7] M. D. Zalina, M. N. M. Desa, V. Nguyen, and A. H. M. Kassim, "Selecting a probability distribution for extreme rainfall series in malaysia," *Water science and technology*, vol. 45, no. 2, pp. 63–68, 2002.

[8] A. Shabri, U. N. Ahmad, and Z. A. Zakaria, "Tl-moments and l-moments estimation of the generalized logistic distribution," *Journal of Mathematics Research*, vol. 3, no. 1, p. 97, 2011.

[9] Z. A. Zakaria, A. Shabri, and U. N. Ahmad, "Regional frequency analysis of extreme rainfalls in the west coast of peninsular malaysia using partial l-moments," *Water resources management*, vol. 26, no. 15, pp. 4417–4433, 2012.

[10] S. Mkhandi, R. Kachroo, and T. Gunasekara, "Flood frequency analysis of southern africa: Ii. identification of regional distributions," *Hydrological Sciences Journal*, vol. 45, no. 3, pp. 449–464, 2000.

[11] B. Saf, "Regional flood frequency analysis using l-moments for the west mediterranean region of turkey," *Water Resources Management*, vol. 23, no. 3, pp. 531–551, 2009.

[12] K. Yurekli, R. Modarres, and F. Ozturk, "Regional daily maximum rainfall estimation for

cekerek watershed by l-moments," *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling*, vol. 16, no. 4, pp. 435–444, 2009.

[13] M. Ul Hassan, O. Hayat, and Z. Noreen, "Selecting the best probability distribution for at-site flood frequency analysis; a study of torne river," *SN Applied Sciences*, vol. 1, no. 12, pp. 1–10, 2019.

[14] J. A. Greenwood, J. M. Landwehr, N. C. Matalas, and J. R. Wallis, "Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form," *Water resources research*, vol. 15, no. 5, pp. 1049–1054, 1979.

[15] J. R. Hosking, "L-moments: Analysis and estimation of distributions using linear combinations of order statistics," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 52, no. 1, pp. 105–124, 1990.

[16] J. R. M. Hosking and J. R. Wallis, *Regional frequency analysis*, 1997.

[17] J. H. Ward Jr, "Hierarchical grouping to optimize an objective function," *Journal of the American statistical association*, vol. 58, no. 301, pp. 236–244, 1963.

[18] J. Hosking and J. Wallis, "Some statistics useful in regional frequency analysis," *Water resources research*, vol. 29, no. 2, pp. 271–281, 1993.

[19] C. Pearson, "Application of l-moments to maximum river flows," *The New Zealand Statistician*, vol. 28, no. 1, pp. 2–10, 1993.

[20] R. M. Vogel, T. A. McMahon, and F. H. Chiew, "Floodflow frequency model selection in australia," *Journal of Hydrology*, vol. 146, pp. 421–449, 1993.

[21] R. M. Vogel and I. Wilson, "Probability distribution of annual maximum, mean, and minimum streamflows in the united states," *Journal of hydrologic Engineering*, vol. 1, no. 2, pp. 69–76, 1996.

[22] M. C. Peel, Q. Wang, R. M. Vogel, and T. A. McMAHON, "The utility of l-moment ratio diagrams for selecting a regional probability distribution," *Hydrological Sciences Journal*, vol. 46, no. 1, pp. 147–155, 2001.