	$\gamma \leq x < \infty$, γ = location parameter α = shape parameter, $\alpha > 0$ β = scale parameter, $\beta > 0$ Γ = Gamma function
Log-Pearson Type III	$f(x) = \frac{1}{x\lvert\beta\rvert\Gamma(\alpha)}(\frac{\ln(x)-\gamma}{\beta})^{\alpha-1}\exp(-\frac{\ln(x)-\gamma}{\beta})$ $0 < x \leq e^{\gamma}$, $\beta < 0$ $e^{\gamma} \leq x < \infty$, $\beta > 0$ α = shape parameter, $\alpha > 0$ β = scale parameter, $\beta \neq 0$ γ = location parameter Γ = Gamma function

3.4 Checking the goodness of fit

The two most commonly used tests of goodness of fit namely Chi-Square (CS) and Kolmogorov-Smirnov (KS) tests are applied to the data series for checking the fit of probability distributions used in this study. The test statistics of each test are computed and tested at level of significance (α = 0.05). If the computed statistic is smaller than the critical value, it indicates that the distribution fits the data well and the distribution can be accepted. Based on the test results, probability distributions are ranked from 1 (the highest rank with minimum value of test statistic) to 3 (the lowest). The highest to lowest ranked probability distributions are given the score of three, two, and one respectively.

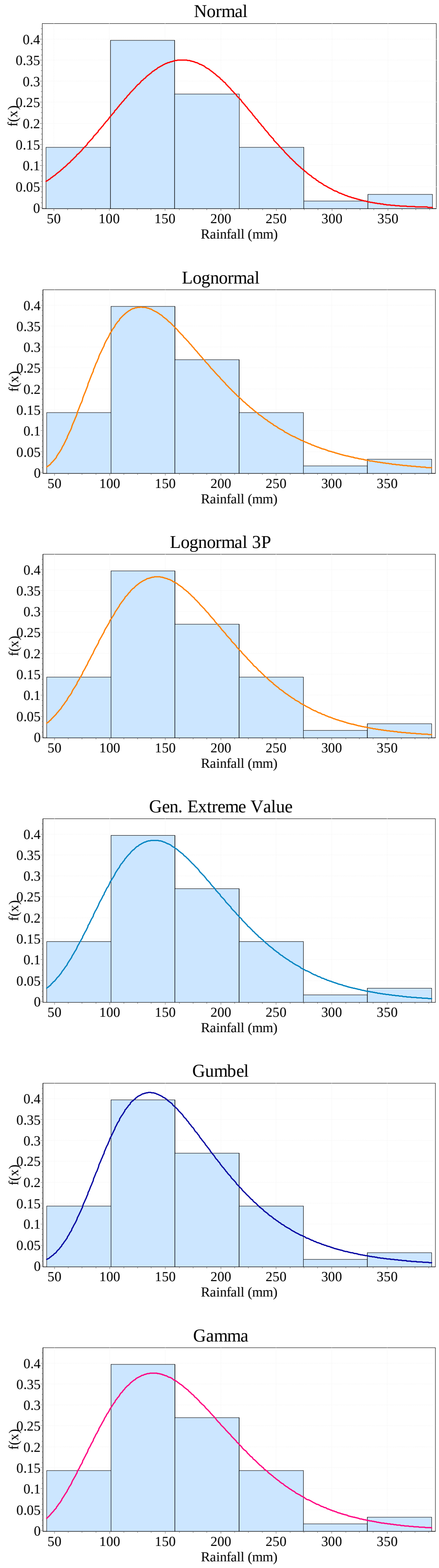
4. Results and discussion

Annual maximum daily rainfall for Kuantan river at four stations are checked whether outliers exist in the data series before using them. Based on the number of data in the series used, outliers are calculated using Eqs. (1) and (2) and given in Table 3. K_N value based on the number of data in the series is obtained from the table for outlier test in [17]. It is observed that there is no high and low outlier in all data series.

Table 3: Outliers for each data series

Station	No. of data	K_N	High outlier (mm)	Low outlier (mm)
3832015	45	2.727	724.40	25.76
3833002	24	2.467	529.60	55.70
3931013	63	2.854	506.77	45.93
3931014	65	2.866	436.70	57.82

The probability density functions of eight distributions chosen in this study for each series are plotted to observe how well the distribution fit with the data series in terms of visual comparison. As an example, the probability density functions for the distributions for Kuantan river at Station 3931013 are given in Figure 3. It is observed from Fig. 3 that all density functions fit quite well with the series except the N2 distribution.



International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

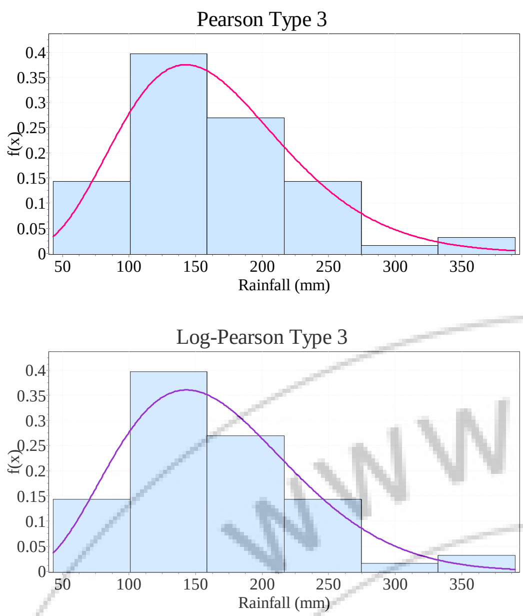


Figure 3: Probability density function of Kuantan river basin at Station 3931013

The parameters of the eight distributions under study are estimated for three data series using EasyFit software [18] and shown in Table 4. The method of moments is used to estimate the parameters for distributions with available moment estimates while the maximum likelihood estimate and least square estimate are used for other estimations. Annual maximum daily rainfall for the desired return periods are calculated using these parameters.

Table 4: Estimated parameters for distributions under study

Distribution	Stations			
	3832015	3833002	3931013	3931014
Normal	σ=91.75 μ=161.6	σ=90.64 μ=190.3	σ=65.84 μ=165.5	σ=58.15 μ=168.7
Lognormal 2P	σ=0.61 μ=4.92	σ=0.45 μ=5.15	σ=0.42 μ=5.03	σ=0.35 μ=5.07
Lognormal 3P	σ=0.43 μ=5.25 γ=-46.58	σ=0.89 μ=4.40 γ=74.92	σ=0.25 μ=5.5, γ=-91.88	σ=0.28 μ=5.27 γ= -34.23
Gamma	α= 3.10 β= 52.09	α= 4.41 β= 43.17	α=6.31 β=26.20	α=8.41 β=20.05
Pearson Type III	α= 1.94 β=70.19 γ= 25.27	α= 0.87 β=101.3 γ= 87.0	α=8.03 β=22.92 γ=-18.62	α=5.12 β=26.12 γ=34.81
Log-Pearson Type III	α=20.04 β= -0.14 γ= 7.66	α=48.99 β=0.065 γ=1.95	α=12.12 β=-0.12 γ=6.49	α=98.12 β=-0.035 γ=8.56
Generalised Extreme Value	k=-0.028 σ= 74.03 μ=120.9	k=0.113 σ=64.83 μ=144.7	k=-0.06 σ=55.32 μ=136.7	k=-0.076 σ=50.79 μ=142.91
Gumbel	σ=71.54 μ=120.3	σ=70.67 μ=149.5	σ=51.34 μ=135.8	σ=45.34 μ=142.49

The goodness of fit tests is performed to check which distribution fits the best to the data series. The values of Chi-square statistics (χ^2) and Kolmogorov-Smirnov (D) are calculated and shown in Tables 5 and 6 respectively for all data series. It is noted that the superscript number refers to the ranking of the best distribution for each station from 1 (the best) to 3 (the worst). Then the scores of points 3, 2 and 1 are given to the ranks 1, 2 and 3 respectively.

Table 5: Chi-square test for all data series

Distribution	Stations			
	3832015	3833002	3931013	3931014
N2	2.671	3.333	5.705	2.731[1]
LN 2P	1.315	3.286	1.296[2]	4.611[2]
LN 3P	1.315	3.252	3.715	6.089
G2	0.393[1]	1.582[3]	1.737[3]	6.510
P3	0.530[3]	N/A	3.702	5.592
LP3	0.409[2]	1.382[1]	3.700	6.092
GEV	0.919	1.396[2]	1.784	6.570
EVI	0.907	2.638	1.001[1]	5.392[3]

It can be seen from Table 5 that all distributions are acceptable to fit to all the data series at the significant level, α of 0.05 except P3 for Station 3833002. Based on the Chi-square test, G2 and LP3 obtained the highest scores of five.

Table 6: Kolmogorov-Smirnov test for all data series

Distribution	Stations			
	3832015	3833002	3931013	3931014
N2	0.084	0.197	0.105	0.095
LN 2P	0.099	0.160	0.067	0.063[1]
LN 3P	0.073	0.119[1]	0.060[1]	0.073
G2	0.064[2]	0.157	0.064	0.075
P3	0.089	0.165	0.063[3]	0.065[2]
LP3	0.070	0.144[3]	0.069	0.070[3]
GEV	0.060[1]	0.142[2]	0.0603[2]	0.071
EVI	0.067[3]	0.163	0.073	0.078

It can be observed from Table 6 that all distributions are acceptable to fit to the data at the significant level, α of 0.05. Based on the Kolmogorov-Smirnov test, GEV distribution obtained the highest scores of seven followed by LN 3P with the score of six.

It can be concluded from Tables 5 and 6 that G2 distribution for Station 3832015, LP3 and GEV for Station 3833002, G2 and EVI for Station 3931013 and LN 2P for Station 3931014 are considered as the best fit distributions.

In overall, based on both tests, GEV distribution obtained the highest score of nine and it is considered as the best fit for all data series of Kuantan river basin. GEV could be more suitable since it chooses EVI, EVII, and EVIII according to the characteristics of each data series. LN 2P, G2 and LP3 distributions are equally good since each distribution obtained the score of seven. The findings from this study are consistent with the findings obtained for Kuala Lumpur, Selangor and Klang river basins by Alias and Takara [12], Shabri et al. [13] and Amir et al. [14] respectively. Feng et al. [19] also used GEV distribution to model the annual extreme precipitation in China.

Annual maximum daily rainfall with recurrence intervals of 2, 10, 25, 50 and 100 years are calculated using Eqs. (3) and (4) for four probability distributions: GEV, G2, LN 2P, and LP3 distributions. The results are given in Table 7.

Table 7: Annual maximum daily rainfall (mm) obtained by different distributions for all stations

Distribution	Return period (years)				
	2	10	25	50	100
Station 3832015					
Gen. Extreme Value	147.8	282.3	347.3	394.4	440.3
Gamma	144.6	284.6	351.9	400.1	446.9
Lognormal 2P	136.8	297.2	394.9	474.4	559.5
Log-Pearson Type III	143.2	289.5	361.8	413.8	464.0
Station 3833002					
Gen. Extreme Value	169.0	310.9	394.7	462.9	536.2
Gamma	176.1	311.7	374.4	418.8	461.4
Lognormal 2P	171.8	304.7	375.7	430.2	486.0
Log-Pearson Type III	168.1	312.2	398.8	469.7	546.2
Station 3931013					
Gen. Extreme Value	156.7	253.1	297.6	329.0	358.9
Gamma	156.8	253.5	296.6	326.9	355.7
Lognormal 2P	152.6	260.4	316.8	359.5	402.8
Log-Pearson Type III	158.8	263.2	291.4	316.5	339.1
Station 3931014					
Gen. Extreme Value	161.3	248.0	287.1	314.4	340.1
Gamma	162.0	246.1	282.8	308.3	332.4
Lognormal 2P	158.9	248.8	293.2	326.1	358.7
Log-Pearson Type III	160.8	247.7	287.4	315.5	342.5

It can be seen from Table 7 that annual maximum daily rainfall estimated by LN 2P is the highest at larger return periods for the all data series except Station 3833002. Estimated values obtained by other three distributions do not vary significantly for all stations except Station 3833002. There is no trend observed for Station 3833002 since the data length is short as compared to other stations.

In overall, GEV and LP3 distributions are recommended for estimation of annual maximum daily rainfall for the Kuantan river basin. This is consistent with the findings obtained by Svensson and Jones [5] and other researchers.

5. Conclusions

A total of eight probability distributions are applied to the series of annual maximum daily rainfall of four stations for Kuantan river basin. The conclusions obtained from this study are as below.

- Based on the analysis of statistical tests, Generalised Extreme Value distribution proves to be the most appropriate distribution for annual maximum daily rainfall at all stations under study for Kuantan river basin.
- In overall, Generalised Extreme Value and Log-Pearson Type III distributions are recommended for estimation of annual maximum daily rainfall for the Kuantan river basin.
- The estimated extreme rainfall with various frequencies and durations can be used as the basic inputs in hydrologic design such as in the design of storm sewers, culverts and many other structures as well as inputs to rainfall runoff models.
- Future research can be carried out by using other rainfall stations in Klang river basin to verify that Generalized Extreme Value and Log-Pearson Type III distributions are the recommended distributions.

References

[1] K. Subramanya, Engineering Hydrology. 3rd ed., Tata McGraw-Hill Inc., 2009.

[2] V.M. Yevjevich, Probability and Statistics in Hydrology. Colorado: Water Resources Publications, 1972.

[3] L.S. Hanson, and R. Vogel, “The probability distribution of daily rainfall in the United States”. World Environmental and Water Resources Congress 2008, pp. 1-10, 2008.

[4] A.R. Rao, and Shih-Chieh Kao, “Statistical analysis of Indiana rainfall data Final Report”. FHWA/IN/JTRP-2006/8, Purdue University, USA, 2006.

[5] C. Svensson, and D.A. Jones, “Review of rainfall frequency estimation methods”. 2010. Available: www.blackwell-synergy.com, http://onlinelibrary. wiley .com/doi/10.1111/j.1753318X.2010. 01079.x/abstract. [Accessed: July 23, 2014].

[6] P. Jordan, E. Weinmann, P. Hill, and C. Wiesenfeld, “Australian rainfall and runoff revision project 2: Collection and review of areal reduction factors collation and review of areal reduction factors from applications of the CRC-FORGE method in Australia”. Final report, Engineers Australia, 2013.

[7] R.P. Canterford, N.R. Pescod, H.J. Pearce, L.H. Turner, and R.J. Atkinson, “Frequency analysis of Australian rainfall data as used for flood analysis and design”. Hydrologic Frequency Modeling, pp. 293-302, 1987.

[8] S.R. Bhakar, A.N. Bansal, N. Chhajed, R.C. Purohit, “Frequency analysis of consecutive days maximum rainfall at Banswara, Rajasthan, India”. ARPN Journal of Engineering and Applied Sciences. Vol. 1, No.3, pp. 64-67, 2006.

[9] X.S. Kwaku, and O. Duke, “Characterization and frequency analysis of one day annual maximum and two to five consecutive days’ maximum rainfall of Accra, Ghana”. APRN Journal of Engineering and Applied Sciences, Vol. 2, No. 5, pp. 27-31, 2007.

[10] B. Singh, D. Rajpurohit, A. Vasishth, and J. Singh. “Probability analysis for estimation of annual one day maximum rainfall of Jhalarapatan area of Rajasthan, India”. Plant Archives, Vol. 12, No. 2, pp. 1093-1100, 2012.

[11] S. Nadarajah, and D. Choi. “Maximum daily rainfall in South Korea”. J. Earth Syst. sci. 116, No. 4, pp. 311-320, 2007.

[12] N.E. Alias, and K. Takara, “Probability Distribution for Extreme Hydrological Value-Series in the Yodo River Basin, Japan and Kuala Lumpur, Malaysia”. In Proceedings of the 2nd International Conference on Water Resources, Langkawi, Malaysia, 2012.

[13] A.B. Shabri, Z.M. Daud and N.M. Ariff, “Regional analysis of annual maximum rainfall using TL-moments method”. Theoretical & Applied Climatology, Vol. 104, Issue 3/4, pp. 561-570, 2011.

[14] H.M.K. Amir, and K.W. Loke, “Frequency analysis of annual maximum rainfall for the Klang River basin

using GEV distribution", Malaysia Science and Technology congress, Kuala Lumpur, Malaysia, 1993.

[15] M.F.M. Nasir, M.A. Zali, H. Juahir, H. Hussain, S.M. Zain, and N. Ramli, N. "Application of receptor models on water quality data in source apportionment in Kuantan River Basin", Iranian Journal of Environmental Health Science & Engineering, pp. 9-18, 2012.

[16] V.T. Chow, D.R. Maidment, and L.W. Mays, Applied Hydrology, McGraw-Hill, Inc., 1988.

[17] Hydrology Subcommittee Bulletin #17B. "Guidelines for determining flood flow frequency", Interagency Advisory Committee on Water Data,, available from Office of Water data Coordination, U.S. Department of the Interior Geological Survey, Reston, VA, 1982.

[18] "EasyFit 3.0," computer software, MathWave Technologies, Available: <www.mathwave.com> [Accessed: July 3, 2014].

[19] S. Feng, S. Nadarajah, and Q. Hu, "Modelling annual extreme precipitation in China using the Generalized Extreme Value distribution", Journal of the Meteorological Society of Japan, Vol. 85, No. 5, pp. 599-613, 2007.

Author Profile

Ni Lar Win obtained BE (Civil) from Rangoon Institute of Technology, Myanmar in 1981 and MSc and PhD in Hydrology from Free University Brussels, Belgium in 1989 and 1994 respectively. She is currently working as Associate Professor at Faculty of Science, Technology, Engineering and Mathematics, INTI International University, Malaysia.

Khin Maung Win obtained BE (Civil) from Rangoon Institute of Technology, Myanmar in 1969, MEngSc from University of Melbourne, Australia in 1984 and PhD in Civil and Environmental Engineering from Utah State University, USA in 1993. He is currently working as Professor at Faculty of Science, Technology, Engineering and Mathematics, INTI International University, Malaysia.