www.cafetinnova.org

Indexed in
Scopus Compendex and Geobase Elsevier, Chemical Abstract Services-USA, Geo-Ref Information Services-USA, List B of Scientific Journals, Poland, Directory of Research Journals

ISSN 0974-5904, Volume 07, No. 04

International Journal of Earth Sciences and Engineering

August 2014, P.P.1315-1321

Low Flow Frequency Analysis on Selected Rivers in Malaysia

NI LAR WIN AND KHIN MAUNG WIN
Faculty of Science, Technology, Engineering and Mathematics, INTI International University, Putra Nilai, Malaysia
Email: larwin.ni@newinti.edu.my, maungwin.khin@newinti.edu.my

Abstract: Prediction of low flow of a river in magnitude as well as in frequency is necessary for the planning and design of water resource projects since low flow affects significantly in water supply, water quality and river ecological status. A number of probability distributions are used to estimate low flow discharges with various return periods of extreme hydrologic events. Among others, Gamma, Gumbel, Lognormal with 2 and 3 parameters, Log-Pearson Type III and Weibull distributions are used in this study. The objectives of this study are: (i) to carry out the low flow frequency analysis with the above mentioned probability distributions to selected rivers in Malaysia, (ii) to identify the most appropriate probability distribution for the basins under study and (iii) to determine the low flow for selected return periods. Frequency analysis is carried out using frequency factor method to determine annual minimum low flow. 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 to check the fit of probability distributions used in this study. Based on the analysis of statistical tests, Gamma, 2P and 3P Lognormal and Log-Pearson Type III distributions fit well to the Kelantan river. Gamma and Lognormal 3P distributions fit well to the Johor river. Lognormal 2P distribution is the best fit to the Muda river followed by Lognormal 3P distribution. In overall, 2P and 3P Lognormal distributions are recommended in estimation of low flow discharges for all the rivers under study which can be used in water quality and quantity management at gauged and ungagged sites.

Keywords: *frequency analysis, probability distributions, frequency factor, Kelantan river, Johor river, Muda river*

1. Introduction

Prediction of low flow of a river in magnitude as well as in frequency is necessary for the planning and design of water resource projects since low flow affects significantly in water supply, water quality and river ecological status. Numerous studies have been carried out to estimate the floods (high flows) for the rivers. However, only a limited number of researches have been carried out on the estimation of low flow for the rivers.

Matalas [1] analysed low flow data for 34 sites in the United States using Pearson Type III, Pearson Type V, three-parameter (3P) Weibull, and 3P lognormal distributions. He concluded that the Weibull and Pearson Type III distributions performed equally well and tended to outperform the other two distributions. Vogel and Kroll [2] used 2P and 3P Lognormal, 2P and 3P Weibull, Log-Pearson Type III distributions to fit low flow data from 23 sites in Massachusetts, United States and concluded that Log-Pearson Type III slightly outperformed the other 2P and 3P distributions. They also indicated that the preferred frequency distribution varies by region and there is no one frequency distribution that clearly outperforms all others. Log-Pearson Type III distribution which is employed by the United States Geological Survey (USGS) for annual minimum streamflow series in the United States was used by many researchers [3], [4], [5] and [6].

Grandry et al. [7] reported that the most common distributions that fit low flow data in Wallonia, Belgium were 2P Lognormal and Gamma. For annual minimum flow series for Ireland, Extreme Value Type I and 2P and 3P Lognormal distributions are generally satisfactory [8].

In Malaysia, 2P or 3P Lognormal distributions and Weibull distribution for the catchments in Peninsular Malaysia for low flow studies were recommended by Department of Irrigation and Drainage (DID) [9]. Three parameter Lognormal distribution were selected for Dongjiang basin, South China since it provides the best fit, outperforming Generalized logistic, Generalized extreme value, Pearson Type III and Generalized Pareto distributions [10].

Nathan and McMahon [11] considered some practical aspects concerning the application of the Weibull distribution to low-flow frequency analysis. In their study, two- and three-parameter forms of the distribution are fitted to a total of 987 distributions derived from the daily flow data of 134 catchments located in southeastern Australia. Data for 21 rivers in the Otago region of the South Island of New Zealand were analysed using the Lognormal, Weibull, Extreme Value Type I and the Gringorton plotting positions and it was found that the best distribution for the group was the Generalized Extreme Value distribution but 3P Lognormal and Extreme Value Type I distributions also fit the data well [12].

The objectives of this study are: (i) to carry out the low flow frequency analysis with some commonly used probability distributions to selected rivers in Malaysia, (ii) to identify the most appropriate probability distribution for the basins under study and (iii) to determine the low flow for selected return periods.

2. Data used in this study

Three rivers namely Kelantan, Johor and Muda rivers are selected to determine the low flow in this study. The Kelantan River basin is located in the north eastern part of Peninsular Malaysia between latitudes 4° 40' and 6° 12' North, and longitudes 101° 20' and 102° 20' East. The maximum length and breadth of the catchment are 150 km and 140 km respectively. The river is about 248 km long and drains an area of 13,100 km^2, occupying more than 85% of the State of Kelantan. It divides into the Galas and Lebir Rivers near Kuala Krai, about 100 km from the river mouth [13]. Johor river, 122.7 km long, drains an area of 2636 km^2. It originates from Mt. Gemuruh and flows through the southeastern part of Johor and finally into the Straits of Johor. The catchment is in irregular shape. The maximum length and breadth of the catchment are 80 km and 45 km respectively [14]. Muda river is the longest river in the state of Kedah and it is situated in northern Peninsular Malaysia, with its upstream flow coming from the northern mountainous area of the state. The river which has a length of 180 km, flows towards the southern area of the state and has a catchment area of 4210 km^2 [15].

Mandal and Cunnane [16] stated that some of the commonly used low flow indices are annual minimum mean daily flow, annual minimum m-day sustained low-flow, annual minimum m-day moving average flow and 95 percentile flow. In our study, the low flow index chosen is mean annual minimum flow on a 7-day average basis (MAM7). Series of annual minimum 7-day low flow from 1980 to 2010 of three rivers under study are shown in Figure 1. In the series, data for the year is excluded from further analysis if more than 18 days of missing data (thresholds of 5% of daily data) occur during a low flow period for that given year, as guided by World Meteorological Organization [17].

The data of each series are tested with the outlier test to check whether outliers exist in the data series as described in [18] before using them. Based on the skewness coefficient of the data series used, outliers are calculated. It is observed that the discharge in the year 1995 is larger than the high outlier for Johor river. However, useful historical information is not available to adjust for high outlier and therefore it is retained in the series. There is no low outlier observed in all series. The statistical characteristics of the data series of each river are shown in Table 1.

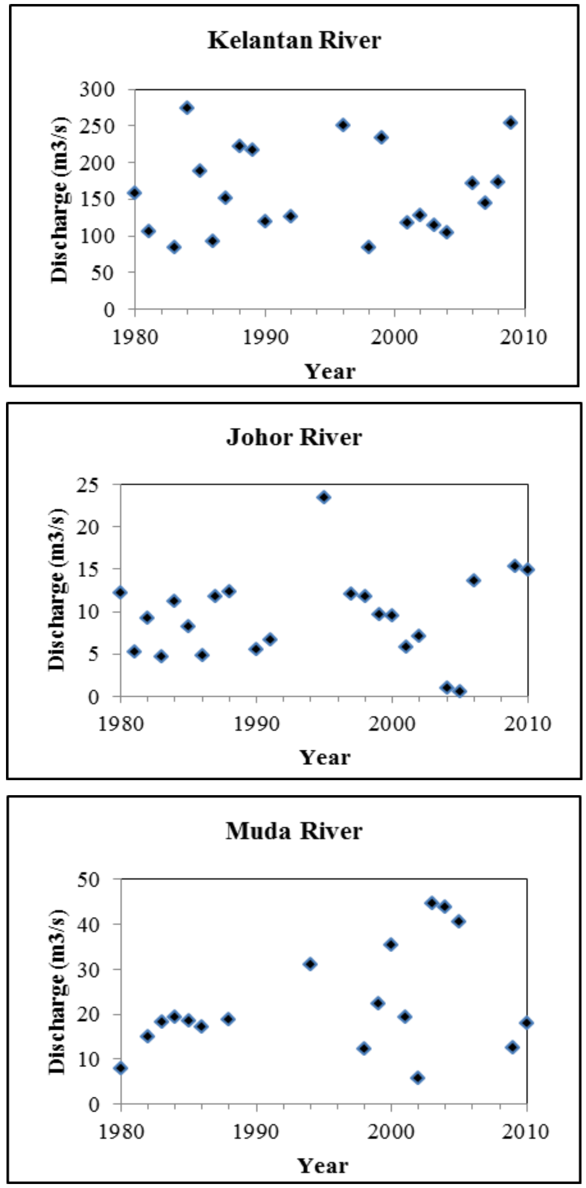


Figure 1. *Series of annual minimum 7-day low flow of Kelantan, Johor and Muda rivers*

Table 1. *Statistical characteristics of the data series of each river*

River	Discharge station	Mean (m^3/s)	Std. dev.	Skew
Kelantan	Jam Guillemard	159.7	59.33	0.54
Johor	Rantau Panjang	9.41	5.09	0.60
Muda	Jam Syed Omar	22.30	11.88	10.8

3. Methodology

3.1. Frequency analysis using frequency factors

The frequency factor equation shown in (1) which is applicable to many probability distributions used in hydrologic frequency analysis was proposed by Chow [19].

$$x_T = \mu + K_T\sigma \qquad (1)$$

which may be approximated by

$$x_T = \bar{x} + K_T s \qquad (2)$$

Where

x_T = value of the variate x of a random hydrologic series with a return period T,

$\bar{x}$ = mean of the variate,

s = standard deviation of the variate,

K_T = frequency factor which depends upon the return period T and assumed frequency distribution.

In the event that the variable analyzed is y = log x, then the same method is applied to the statistics for the logarithms of the data using

$$y_T = \bar{y} + K_T s \qquad (3)$$

and the required value of x_T is found by taking the antilog of y_T.

3.2. Fitting the probability distributions

Smakhtin [20] mentioned that different forms of Weibull, Gumbel, Pearson Type III, Lognormal distributions are the most frequently referred distribution functions in the literature for low flows. In our study, 2P and 3P Lognormal, Gamma, Gumbel, Weibull and Log-Pearson Type III are identified to evaluate the best fit probability for low flow for the Kelantan, Johor and Muda rivers. The probability density function, the range of the variable, and the parameter involved are summarized in Table 2.

Table 2. *Summary of probability density function, the range of the variable, and the distribution's parameters*

Distribution	Probability density function
Lognormal 2P	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}}\exp(-\frac{(\ln(x)-\mu)^2}{2\sigma^2})$ $0 < x < \infty$ σ = scale parameter, σ > 0 μ = shape parameter, μ > 0 γ = location parameter
Lognormal 3P	$f(x) = \frac{1}{(x-\gamma)\sigma\sqrt{2\pi}}\exp(-\frac{(\ln(x-\gamma)-\mu)^2}{2\sigma^2})$ $\gamma < x < \infty$ σ = scale parameter, σ > 0 μ = shape parameter, μ > 0 γ = location parameter
Log-Pearson Type III	$f(x) = \frac{1}{x\lvert\beta\rvert\Gamma(\alpha)}(\frac{\ln(x)-\gamma}{\beta})^{\alpha-1}\exp(-\frac{\ln(x)-\gamma}{\beta})$ $0 < x \le e^{\gamma}, \beta < 0$; $e^{\gamma} \le x < \infty, \beta > 0$ α = shape parameter, α > 0 β = scale parameter, β ≠ 0 γ = location parameter Γ = Gamma function
Gamma (2P)	$f(x) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}\exp(-\frac{x}{\beta})$ $0 \le x < \infty$ α = shape parameter, α > 0 β = scale parameter, β > 0 Γ = Gamma function
Gumbel (Extreme Value Type I)	$f(x) = \frac{1}{\sigma}\exp(-z-\exp(-z))$ $-\infty < x < \infty$ σ = scale parameter, σ > 0 μ = location parameter where $z = \frac{x-\mu}{\sigma}$
Weibull (Extreme Value Type III)	$f(x) = \frac{\alpha}{\beta}(\frac{x}{\beta})^{\alpha-1}\exp(-(\frac{x}{\beta})^{\alpha})$ $0 \le x < \infty$ α = shape parameter, α > 0 β = scale parameter, β > 0

3.3. Checking the goodness of fit

To check the fit of probability distributions used in this study, 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. The test statistics of each test are computed and tested at (α = 0.05) level of significance.

1) Chi-square test

The statistic is calculated by

$$\chi^2 = \sum_{j=1}^{k}\frac{(O_j - E_j)^2}{E_j} \qquad (4)$$

Where O_j is the observed number of events in the j^{th} class interval and E_j is the number of events that would be expected from the theoretical distribution.

2) Kolmogorov-Smirnov (K-S) test

The statistic D_n is evaluated by observing the deviation of the sample distribution function P(x) from the completely specified continuous hypothetical distribution function $P_0(x)$, such that

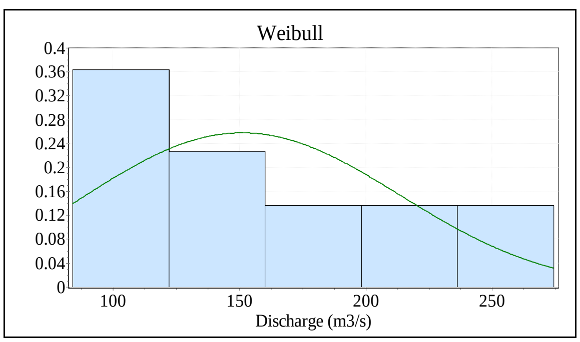
$$D_n = Max\,|\,P(x) - P_0(x)\,| \quad (5)$$

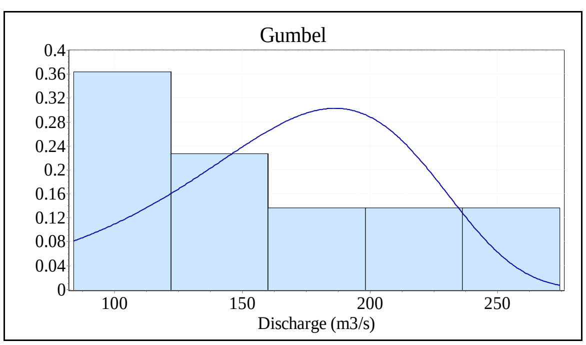
If the computed statistic is smaller than the critical value it indicates that the distribution fits the data well and the distribution can be accepted. Based on the test results, probability distributions are ranked from 1 (the best) to 6 (the worst). Then scores are given to the first three distributions. The best probability distribution is given the highest score of three points.

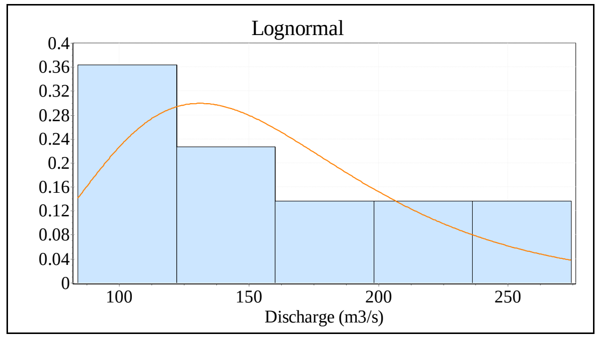
4. Results

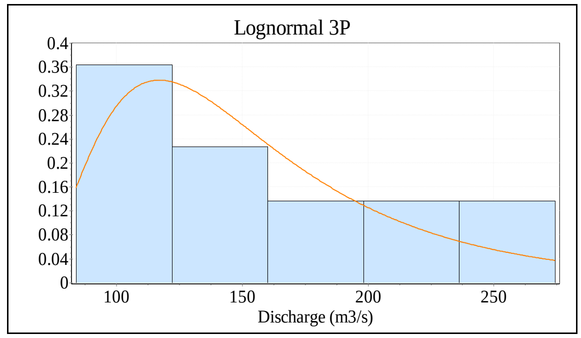
The probability density functions of six distributions chosen in this study for each series are plotted to observe how well the distribution fit with the data series. As an example, the probability density functions for the distributions for Kelantan river are given in Figure 2. It is observed from Figure 2 that all density functions fit quite well with the series except Gumbel distribution.

The parameters of the six distributions are estimated for three data series using EasyFit [21] and shown in Table 3. These parameters are used in calculating low flow discharges at the desired return periods.









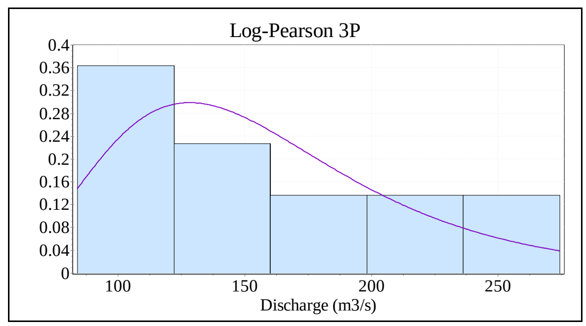


Figure 2. *Probability density function of Kelantan river at Jam Guillemard*

Table 3. *Estimated parameters for distributions under study*

Distribution	Parameters		
	Kelantan	**Johor**	**Muda**
Gamma	α=7.244	α=3.417	α=3.523
	β=22.043	β=2.754	β=6.33
Gumbel	σ=46.258	σ=3.969	σ=9.264
	μ=186.38	μ=11.701	μ=27.649
Lognormal 2P	σ=0.362	σ=0.832	σ=0.545
	μ=5.008	μ=2.015	μ=2.966
Lognormal 3P	σ=0.589	σ=0.1623	σ=0.442
	μ=4.508	μ=3.401	μ=3.168
	γ=52.979	γ=-20.995	γ=-3.849
Log-Pearson Type III	α=429.13	α=1.03	α=32.504
	β=0.0179	β=-.838	β=-0.098
	γ=-2.669	γ=2.879	γ=6.16
Weibull	α=2.994	α=1.217	α=1.942
	β=173.18	β=10.974	β=24.115

Tables 4 and 5 list the values of Chi-Square (CS) and Kolmogorov-Smirnov (KS) statistics for Kelantan river.

Table 4. *Chi-Square tests for Kelantan river*

Distribution	CS test statistics	Accepted	Score/ Rank
Gamma	0.1187	Yes	3 (1)
Gumbel	3.4485	Yes	0 (6)
Lognormal 2P	0.2329	Yes	2 (2)
Lognormal 3P	0.8239	Yes	1 (3)
Log-Pearson 3P	0.8958	Yes	0 (5)
Weibull	0.8712	Yes	0 (4)

Table 5. *Kolmogorov-Smirnov tests for Kelantan river*

Distribution	KS test statistics	Accepted	Score/Ra nk
Gamma	0.1233	Yes	0 (4)
Gumbel	0.2054	Yes	0 (6)
Lognormal 2P	0.1183	Yes	1 (3)
Lognormal 3P	0.1133	Yes	2 (2)
Log-Pearson 3P	0.1128	Yes	3 (1)
Weibull	0.1295	Yes	0 (5)

It can be seen from Tables 4 and 5 that all distributions are acceptable to fit to the data at the significant level, α of 0.05. The total score is obtained by adding up the scores from both tests. Gamma, 2P and 3P Lognormal and Log-Pearson Type III fit well to the Kelantan river since they give the highest total score of three.

Tables 6 and 7 list the values of Chi-Square (CS) and Kolmogorov-Smirnov (KS) statistics for Johor river.

Table 6. *Chi-Square test for Johor river*

Distribution	CS test statistics	Accepted	Score/Ra nk
Gamma	0.0006	Yes	3 (1)
Gumbel	1.7959	Yes	0 (5)
Lognormal 2P	1.0988	Yes	0 (4)
Lognormal 3P	0.1847	Yes	2 (2)
Log-Pearson 3P	N/A	-	-
Weibull	0.5857	Yes	1 (3)

Table 7. *Kolmogorov-Smirnov test for Johor river*

Distribution	KS test statistics	Accepted	Score/R ank
Gamma	0.1299	Yes	2 (2)
Gumbel	0.1350	Yes	1 (3)
Lognormal 2P	0.1934	Yes	0 (5)
Lognormal 3P	0.1025	Yes	3 (1)
Log-Pearson 3P	0.1651	Yes	0 (4)
Weibull	0.2076	Yes	0 (6)

It can be observed from Tables 6 and 7 that all distributions expect Log-Pearson Type III are acceptable to fit to the data at the significant level, α of 0.05. Gamma and Lognormal 3P fit well to the Johor river since they give the highest total score of five.

Tables 8 and 9 list the values of Chi-Square (CS) and Kolmogorov-Smirnov (KS) statistics for Muda river.

Table 8. *Chi-Square test for Muda river*

Distribution	CS test statistics	Accepted	Score/Ra nk
Gamma	5.1076	Yes	0 (4)
Gumbel	9.2117	No	-
Lognormal 2P	4.9481	Yes	2 (2)
Lognormal 3P	4.9264	Yes	3 (1)
Log-Pearson 3P	5.008	Yes	1 (3)
Weibull	5.2507	Yes	0 (5)

Table 9. *Kolmogorov-Smirnov test for Muda river*

Distribution	KS test statistics	Accepted	Score/Ra nk
Gamma	0.1962	Yes	0 (5)
Gumbel	0.3297	No	-
Lognormal 2P	0.1665	Yes	3 (1)
Lognormal 3P	0.1858	Yes	1 (3)
Log-Pearson 3P	0.1898	Yes	0 (4)
Weibull	0.1856	Yes	2 (2)

It can be seen from Tables 8 and 9 that all distributions except Gumbel distribution are acceptable to fit to the data at the significant level, α of 0.05. Lognormal 2P is the best fit to the Muda river since they give the highest total score of five and followed by Lognormal 3P with the total score of four.

In overall, Gamma, 2P and 3P Lognormal and Log-Pearson Type III distributions are suitable to estimate low flow discharges for all the rivers under study. Among them Lognormal 3P is considered to be the best fit and followed by Lognormal 2P and Gamma distributions based on their total scores for all rivers under study.

Low flow discharges with recurrence intervals of 2, 10, 25, 50 and 100 years are calculated using four probability distributions: Gamma, Lognormal 2P, Lognormal 3P and Log-Pearson Type III distributions. The results for Kelantan, Johor and Muda are given in Tables 10 to 12 respectively.

Table 10. *Low flow discharges X_T (cumec) obtained by different distributions for Kelantan river*

Return period (years)	Gamma	Lognor-mal 2P	Lognor-mal 3P	Log-Pearson Type III
2	152.079	149.573	143.726	150.518
10	88.895	93.010	94.891	92.666

25	72.400	78.172	84.564	77.193
50	62.462	69.870	79.290	68.510
100	54.405	63.170	75.303	61.461

It can be seen from Table 10 that estimated discharges show small differences in results obtained by all distributions. Lognormal 3P gives the higher estimates while Gamma distribution gives lower estimates at larger return periods.

Table 11. *Low flow discharges X_T (cumec) obtained by different distributions for Johor river*

Return period (years)	Gamma	Lognor-mal 2P	Lognor-mal 3P	Log-Pearson Type III
2	8.510	7.499	9.012	5.780
10	3.746	2.519	3.260	3.503
25	2.631	1.690	1.443	3.309
50	2.052	1.306	0.343	3.245
100	1.618	1.036	0.000	3.213

It can be observed from Table 11 that the estimated discharges by Lognormal 3P is the lowest while the estimated values given by Log-Pearson Type III is the highest at larger return periods. Estimated discharges by Log-Pearson Type III have a small variation for all return periods.

Table 12. *Low flow discharges X_T (cumec) obtained by different distributions for Muda river*

Return period (years)	Gamma	Lognorma l 2P	Lognorma l 3P	Log-Pearson Type III
2	20.20	19.406	19.909	18.785
10	9.079	9.462	9.414	9.698
25	6.390	7.275	6.867	7.809
50	5.020	6.139	5.488	6.836
100	3.652	5.271	4.401	6.096

The results in Table 12 show that the discharges estimated by Gamma distribution are lower than the other distributions at the larger return periods. Log-Pearson type III gives the highest estimated discharges at all return periods.

In overall, Lognormal distribution with 2P and 3P are recommended for estimation of low flow discharges for the rivers under study. This is consistent with the findings obtained by DID. Estimation of 7-day 10-year low flow by Lognormal 2P are 93.09 m^3/s, 2.519 m^3/s and 9.462 m^3/s for Kelantan, Johor and Muda rivers respectively.

5.0 Conclusions

Based on the analysis of statistical tests, Gamma, 2P and 3P Lognormal and Log-Pearson Type III distribution fit well to the Kelantan river. Gamma and Lognormal 3P distributions fit well to the Johor river. Lognormal 2P distribution is the best fit to the Muda river follow by Lognormal 3P distribution.

In overall, 2P and 3P Lognormal distributions are recommended for estimation of low flow discharges for all the rivers under study which can be used in water quality and quantity management at gauged and ungagged sites.

References

[1] Matalas, N., Probability Distributions of Low Flows, Professional Paper 434-A, USGS, Washington D.C., 1963.

[2] Vogel, R.M. and C.N. Kroll, Low-Flow Frequency Analysis Using Probability-Plot Correlation Coefficients, *ASCE Journal of Water Resources Planning and Management*, 115(3), pp. 338–357. 1989.

[3] Risley, J., Stonewall, A. and Haluska, T., Estimating Flow-Duration and Low-Flow Frequency Statistics for Unregulated Streams in Oregon, U.S. GEOLOGICAL SURVEY Scientific Investigations Report 2008–5126, 2008.

[4] Reilly, C.F. and Kroll, C.N., Estimation of 7-day, 10-year low-streamflow statistics using baseflow correlation, *Water Resources Research*, Vol., 39, No. 9, SWC 3-1 – 3-10, 2003.

[5] Waltemeyer, S.D., Analysis of the magnitude and frequency of the 4-day annual low flow and regression equations for estimating the 4-day, 3-year low-flow frequency at ungaged sites on unregulated streams in new Mexico, Water Resources Investigation Report, 01-4271, U.S. Geological Survey, 2002.

[6] Weaver, J.C., Low-Flow Characteristics and Discharge Profiles for Selected Streams in the Neuse River Basin, North Carolina, U.S. Geological Survey Water-Resources Investigations Report 98-4135, 1998.

[7] Grandry M., Gailliez S., Sohier, C., Verstraete A., and Degr'e A., A method for low-flow estimation at ungauged sites: a case study in Wallonia (Belgium), *Hydrol. Earth Syst. Sci.*, 17, 1319–1330, 2013.

[8] Brogan, L. and Cunnane, C., Low flows and low flow distributions for Ireland, National Hydrology Seminar, pp. 85 – 92, 2005.

[9] Department of Irrigation and Drainage (DID), Hydrology and water resources, Vol. 4, Government of Malaysia, Jabatan Pengairan dan Saliran Malaysia, 2009.

[10] Chen, Y.D., Huang, G., Shao, Quanxi and Xu, C.Y., Regional analysis of low flow using L-moments for Dongjiang basin, South China,

Hydrological Sciences, 51 (6), pp. 1051-1064, 2006.
[11] Nathan, R.J. and McMahon T.A., Practical aspects of low-flow frequency analysis, *Water Resources Research*, Vol. 26, Issue 9, pp. 2135-2141, 1990.
[12] Caruso, B.S., Evaluation of low-flow frequency analysis methods, *Journal of Hydrology* (NZ), New Zealand Hydrological Society, 39(1), pp. 19-47, 2000.
[13] Ibbitt, R., Takara, Kaoru, Mohd. Nor bin Mohd. Desa and Pawitan, Catalogue of rivers for Southeast Asia and the Pacific, Vol. 4, The UNESCO-IHP Regional Steering Committee for Southeast Asia and Pacific, 2002.
[14] Jayawardena, A.W., Takeuchi, K. and Machbub, B., Catalogue of rivers for Southeast Asia and the Pacific, Vol. 2, The UNESCO-IHP Regional Steering Committee for Southeast Asia and Pacific, 1997.
[15] Aminuddin Ab Ghani, Rabie Ali, Nor Azazi Zakaria, Zorkeflee Abu Hasan, Chun, K.C. and Mohd Sanusi S. Ahamed, A temporal change study of the Muda river system for 22 years, *Inti. J. River Basin Management*, International Association for Hydro Environment Engineering and research, Vol. 8, No. 1, pp. 25-37, 2010.
[16] Mandal, U. and Cunnane, C., Low-flow prediction for ungauged river catchments in Ireland, Irish National Hydrology Seminar, pp. 33–48, 2009.
[17] World Meteorological Organization (WMO), Manual on low flow estimation and prediction, Operational Hydrology Report No. 50, WMO-No. 1029, Switzerland, 2008.
[18] McCuen, H.R., *Hydrologic Analysis and Design*, 3rd ed. Prentice Hall, 2004.
[19] Chow, V.T., Maidment, D.R. and Mays, L.W., *Applied Hydrology*, McGraw-Hill, Inc., USA, 1988.
[20] Smakhtin, V.U., Low flow hydrology: a review, Journal of Hydrology, 240, pp. 147–186, 2001.
[21] EasyFit 3.0, Computer Software, MathWave Technologies, Available at www.mathwave.com, Accessed on 8th July 2014.