Applied Mechanics and Materials
ISSN: 1662-7482, Vols. 773-774, pp 1266-1270
doi:10.4028/www.scientific.net/AMM.773-774.1266

Submitted: 2014-09-15
Revised: 2014-11-19
Accepted: 2014-11-30
Online: 2015-07-15

Comparison of Distribution Methods of Low Flow Analysis for Bandar Segamat, Johor

Yuliarahmadila Erfen[1,a], Mohd Shalahuddin Adnan[1,b], Noorfahiah Che Ali[1,c], Nurul Farehah Amat[1,d], Zawani Mohd Zahudi[1,e]

[1]Department of Water and Environmental Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Johor, Malaysia.

[a]yuliaerfen@gmail.com, [b]mohdshalahuddin@gmail.com, [c]fathiazu@gmail.com, [d]nurulfarehahamat@gmail.com and [e]wanizahudi88@gmail.com

Keywords: Stream flow, statistical analysis, Chi-Square test, Segamat

Abstract. During the monsoon season, certain areas in Malaysia are experiencing a flood. While during the transition period Malaysia is experiencing a drought. This phenomenon could lead to severe disaster and precaution monitoring is needed to avoid this occurrences. Low flow during the dry season could lead to several negative effects on the river ecosystem. Thus, this study was conducted to determine the low flow frequency and intensity for the Segamat city. The duration for 2 years to 100 years based on the previous 20 years of stream flow data were used to calculated. Stream flow data were obtained from the Department of Irrigation and Drainage (DID). Two prominent distribution analyses methods known as Gumbel Distribution and Log pearson Type III Distribution were applied. The distribution results were validated using Root Mean Square Error (RMSE) and California method and Weibull method are selected. Based on the analyses results, it clearly shows that the distibution of low flow are between 1 m^3/s to 10 m^3/s. The flow are significantly correlate with the rainfall intensity. RMSE was selected based on the lowest value of 0.721 for the Gumble Distribution and 1.831 for Log Pearson Type III Distribution. Chi-square test shows a good agreement for Gumble Distribution (7.615<12.59) and Log Pearson Type III(5.201<11.07) using 5% significant level. The confident level form both tests are valid (p>0.05), thus, this results could be used for further analyses to alleviate the low flow in the study area.

Introduction

In general, Malaysia is a country with a uniform temperature, high humidity and annually abundant rainfall. Therefore, Malaysia has never run away from the drought and heavy rainfall problems. The season variation of rainfall in Peninsular Malaysia has three main types [1]. Usually, the season of maximum rainfall (winter) in the east coast state of peninsular Malaysia is usually in November, December, and January, while June, July and August are the driest months (summer) in most districts [1, 9,10]. Continuous heavy rainfall may cause flooding.

In hydrological cycle, precipitation was one of the important component and parameter for various natural and socio-economic systems. The study of consequences of global climate change on these systems requires scenario of future precipitation change as input in hydrological cycle [2]. Hydrological data can be analyzing statistical method base on frequency analysis such as statistical distribution. Frequency analysis is related to magnitude of extreme events to their frequency of occurrence using the probability of distribution [3].

Flow statistics can be used to characterize the magnitude of the flow of a specific river location selected. Flow statistic is very important for the federal, state and local agencies in regulating water quality and water supply planning and management [4]. These statistics usually used as a benchmark when setting wastewater treatment plant effluent limits and allowable pollutant loads to meet water quality standards. Stream flow monitoring provides critical information about the potential dangers of floods and high water warnings to local communities. In summary, rainfall, water level and stream flow data is very important in analysis of hydrological process. This paper will discuss the result of statistical analysis to compute period of record 7 day low-flow frequency statistic at Bandar Segamat (7Q20). These was aimed at analyzing the statistic distribution to compare the best analysis using Gumble Distribution and Log Pearson Type III based on chi-squared test for Segamat River and involved only one station study area.

Description of Study Area

Segamat is one districts which located in the state of Johor Darul Ta'zim. It is bordered by four district namely Muar in south, Tampin in northwest, Rompin in northeast and Batu Pahat in Southeast. Segamat District has an area of 283534.5 hectares with percentage of population with 8.6% of Johor population in 2010. Flooding in Bandar Segamat in December 2006 has caused severe damage to property and negative impacts on the population live in that area. 70% of Segamat River catchment consists of highland area (max=1000 m) and the rest 30% are rolling hills with a little swamp [5]. Water level, stream flow and rainfall data have been retrieved from DID database from year 1993 until 2012.

Methodology

Flow data collected at continuous streamflow-gauging station. Based on the daily mean flow records at streamflow-gauging station operated by DID used to compute flow statistics.

Low-flow Frequency

Determination of low flow conditions in rivers and streams is essential because it could influenced the ecological status of watercourse [6]. The 7-day, 20-years (7Q20) flow statistic based on an annual series of the smallest values of mean discharge was computed over seven consecutive days during the annual period. The low flow frequency was calculated for monthly basis by limiting the daily data used in annual series on the month of interest. The probability distribution fit to the annual 7day minimum flow with 20 years recurrence interval. The application of type 3 generalized extreme value (GEV) or Weibull distribution allows the quantiles of the flow distribution to be determined and return periods of any design events estimated as mentioned in [6].

Weibull Method

Weibull methods are considered as type III extremal and suitable for the low flow analysis. The distribution also suitable for flood analysis if ɣ=0 in the type III extreme distribution. It has three parameters which are location, scale and shape. The variate low flow of a random hydrological series with a return period can be estimate using equation 1.

$$Q_T = Q_{ave} - K_T \sigma \tag{1}$$

Gumble Method

Gumble distribution (EV1) is commonly used for low flow analysis because it allows a linear curve fitting from double exponential distribution as describe in Gumble (1958) [7,11]. Equation 2 will be applied in calculation with considered the location (w) and scale parameter (α). Gumble's methods has been simplified in such manner that one can obtained the magnitude of a given period flood without the value of the coefficient of variation of the given data.

$$f(x) = \left(\frac{1}{\alpha}\right) e^{\left\{-\left[\frac{(w-x)}{\alpha}\right]\right\}} e^{-e^{\left\{-\left[\frac{(w-x)}{\alpha}\right]\right\}}} \tag{2}$$

Log Pearson Type III Methods

Log Pearson Type III distribution is a 3 parameter distribution operation upon the logarithms of variable, the Log Pearson Type III would appear to be an extremely versatile distribution. However, its applicability in hydrology is strictly limited. The probability density function is as follow in equation 3 where α, β and γ are the scale, shape and location parameters respectively. This distribution is computed using formula in equation 3.

$$p(x) = \frac{1}{x\,\Gamma\,\beta} \left(\frac{\ln x - y}{\alpha}\right)^{(\beta-1)} e^{-\left(\frac{\ln x - y}{\alpha}\right)} \tag{3}$$

Chi-Square (χ^2) Test

The Chi-Square test for goodness of fit provides the solution in comparing the observed frequency distribution of a random variable with a population distribution based on a theoretical probability density function [7]. If the assumption being tested is true, then the chi-square test statistic is approximately distribution as a random variable having a chi-square distribution with degree of freedom v = k – m – 1, where m is the number of parameters estimated from the data I order to obtained the theoretical frequency for each of k categories in which the data have been classified. Large value of chi-square test [12] statistic lead to rejection of the assumption being tested.

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - E_i)^2}{E_i} \quad (4)$$

Result and Discussion

From the data collected, analyzed using Gumble Distribution and Log Pearson Type III distribution was carried out and the Chi-square test was performed to determine which analysis is more suitable. The calculation of Q_T for EV1and LPT3 was derived from equation 2 and equation 3. The value of Q_T was derived from the study conducted by [7]. Table 1 shows the result of Q_T value for both methods include four other methods namely California, Hazen, Weibull and Gringorten methods.

Cumulative probability versus low flow graph was produced to identified the most suitable methods among the applied statistical method. Referred to figure 1, the California method was less appropriate because it shows a big gap compared to others method. While, Table 2 shows the cumulative probability and Root Mean Square Error (RMSE). RMSE is used to identify the most appropriate methods used in subsequent analyzes by comparing two selected methods that have been identified earlier. The lowest value of RMSE was used to determine the suitability of results as well as the accuracy of produced results. For EV1, California was used and Weibull Methods was chosen for LPT3. A study conducted by [5,10] has concluded that the California and Weibull method are the suitable method that can be applied with suit with local climate in Malaysia.

Table 1: The value of Q_T for Gumble and Log Pearson Distribution for Sungai Segamat for 20 years period from year 1993 to 2012

Rank,m	Gumble				Log Pearson Type III			
	California	Hazen	Weibull	Gringorten	California	Hazen	Weibull	Gringorten
1	-0.415	-1.813	-0.514	-1.598	1.097	0.968	1.087	0.986
2	1.010	0.414	0.908	0.482	1.296	1.201	1.277	1.213
3	1.868	1.479	1.763	1.517	1.473	1.386	1.448	1.393
4	2.496	2.202	2.388	2.226	1.646	1.559	1.611	1.566
5	2.999	2.759	2.887	2.776	1.822	1.733	1.779	1.740
6	3.424	3.219	3.309	3.231	2.008	1.914	1.955	1.917
7	3.798	3.617	3.678	3.625	2.207	2.106	2.139	2.110
8	4.136	3.971	4.011	3.976	2.425	2.314	2.340	2.318
9	4.447	4.294	4.316	4.297	2.667	2.543	2.562	2.543
10	4.740	4.596	4.603	4.596	2.941	2.800	2.805	2.800
11	5.020	4.881	4.875	4.881	3.255	3.092	3.086	3.092
12	5.293	5.157	5.138	5.155	3.623	3.431	3.402	3.431
13	5.562	5.427	5.395	5.423	4.064	3.833	3.781	3.824
14	5.833	5.234	5.652	5.691	4.608	4.321	4.236	4.310
15	6.113	5.971	5.912	5.964	5.304	4.933	4.785	4.919
16	6.408	6.258	6.181	6.248	6.237	5.733	5.501	5.696
17	6.660	6.566	6.467	6.553	7.582	6.841	6.465	6.788
18	7.117	6.915	6.784	6.898	9.762	8.521	7.821	8.436
19	7.638	7.350	7.159	7.325	14.246	11.512	10.062	11.264
20	-	9.917	9.537	9.858	55.334	19.511	14.537	18.610

Chi-square tests have been performed and the analysis shows that both methods are suitable for analyze low flow for Segamat River. The calculation of test statistic χ^2 using equation 4 and the value of χ^2 for EV1 is 7.615 while 5.201 for LPT3. By referring to the table of critical value of χ^2 at 5% significance level, the value of χ^2 is 12.59for EV1 and 11.07 for LPT3. Because the typical value obtained is larger, the distributions are suitable in the low flow analysis at 5% significance level. The prediction of return period using formula and graph method has been calculated. Figure 2 show the magnitude of low flow at return period and table 3 shows simplify of prediction methods using formula and graph. Based on table 3 it clearly shows that two selected method can be used to analyzed the low flow for this river.

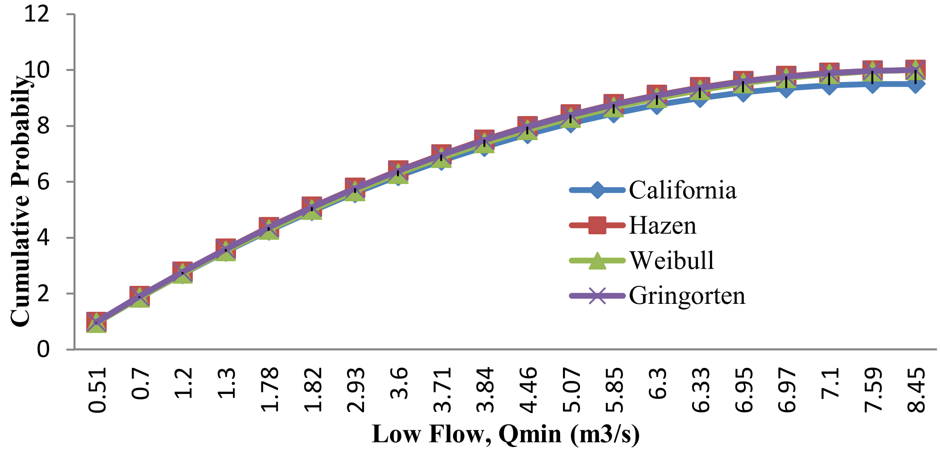

Figure 1: Graph of Cumulative Probability versus Low Flow

Table 2: Distribution probability, return period and cumulative probability for rank 16 and Root Mean Square Error (RMSE)

Methods	Formula	Probability, p'	Return period, T_r	(1 – p')	Cumulative probability	EV1	LPT3
						$RMSE = \sqrt{\frac{\sum(o_i - P_i)^2}{n}}$	
California	$p' = \frac{m}{n}$	0.800	1.250	0.200	9.200	0.721	1.839
Hazen	$p' = \frac{2m-1}{2n}$	0.775	1.290	0.225	9.600	0.875	2.838
Weibull	$p' = \frac{m}{n+1}$	0.762	1.313	0.238	9.524	0.742	1.831
Gringorten	$p' = \frac{m-0.44}{n+0.12}$	0.773	1.293	0.227	9.590	0.825	2.647

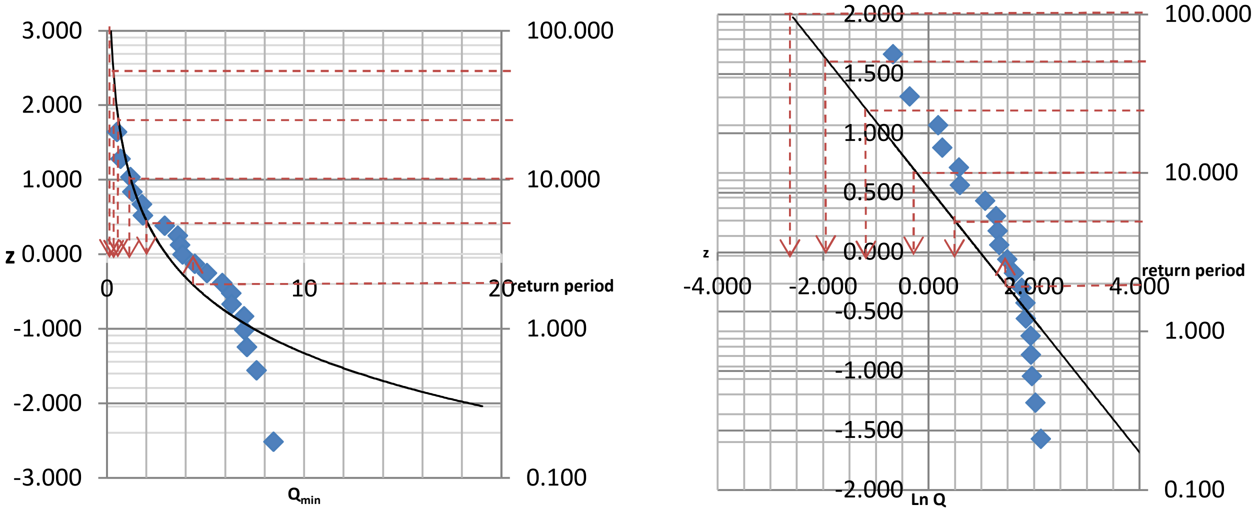

Figure 2: Prediction of magnitude of low flow at return period

Table 3: Prediction method using graph and formula

	Gumble Method			Log-Pearson Type III	
Year	Q_T (Theory)		Q_T (graph)	Q_T (Theory)	Q_T (Graph)
2	4.740		4.4	2.935	4.4817
5	2.496		2	1.644	1.6487
10	1.010		1.2	1.295	0.7788
25	-0.867	≈ 0	0.65	1.049	0.3012
50	-2.260	≈ 0	0.4	0.936	0.1653
100	-3.643	≈ 0	0.2	0.856	0.1225

Conclusion

As a conclusion, the selected methods to analyze the distribution of low flow for the Segamat River were successfully discussed. Based on the results of low flow, Q_{min} of Segamat River, the prediction of magnitude of low flow used Gumble and Log-Pearson Type III methods was conducted. The prediction data using theoretical method show a better result because it close to the real values. In addition, the LPT3 result are more suitable because there is no negative values. Chi-square test show the typical value of EV1 equal to 12.59 > 7.615 while typical value of LPT3 equal to 11.07 > 5.201. It can be concluded that the selected method are suitable.

References

[1] Malaysia Meteorological Department. 2006. Report on heavy rainfall that caused flood in Johor, Melaka, Negeri Sembilan and Pahang during the period 17th -20th December 2006.

[2] Mohammad Mahdavi, Khaled Osati, Sayed Ali Naghi Sadeghi, Bakhtiar Karimi & Jalil Mobaraki. Determining suitable Probability Distribution Models for annual Precipitation Data (A Case study of Mazandaran and Golestan Provinces), Journal of Sustainable Development, Vol. 3, No.1 March 2010, pp159-168

[3] Chow, V.T., Maidment, D.R. & Mays, L. W. (1988). Applied Hydrology. McGraw-Hill, New York, NY.

[4] John Risley, Adam Stonewall & Tana Haluska.Estimate Flow Duration and Low Flow Frequency statistics for Unregulated Streams in Oregon, USGS-science for changing world, Scientific Investigation Report 2008-5126, Version 1.1, June 2009

[5] Muhammad Barzani Gasim, Salmijah Surif, Mazlin Mokhtar, Mohd. Ekhwan Hj Toriman, Sahibin Abd Rahim & Chong Huei Bee. Analisis Banjir Disember 2006: Tumpuan di Kawasan Bandar Segamat, Johor, Sains Malaysiana 39(3)(2010), pp: 353-361

[6] The Environment Agency Publication, Flood and Coastal Erosion Risk Management Research and Development, Fluvia Design Guide-Chapter 2, access on January 2014 at website www.evidence.environment-agency.gov.uk

[7] Soh Bing Zheng. Low Flow Analysis-A Comparison between Gumble and Log Pearson Type III Methods, Tesis Universiti Teknologi Malaysia, April 2010

[8] Gerald Keller. Statistic for Management and Economics, Critical Value of Chi-Square Distribution Table, Appendix B-11, 2012.

[9] Diong Jeong Yik, Subramaniam Moten, Munirah Arrifin and Siva Shangari Govidan. Trends in Intensity and Frequency of Precipitation Extremes in Malaysia from 1951 to 2009. Malaysia Meteorology Department Research Paper, 2013

[10] Ahmad Syukri Mohd Nasir. Kajian Banjir di Sungai Muar, Tesis Universiti Teknologi Malaysia, April 2008.

[11] J.M.Fallot. Efficiency of Gumble Analyses for Determining Extreme Daily Prepatation in Switzerland. Quartier Doriginy, CH-1015, 2012

[12] Stacey J. Shaefer Louis Theodore. Probability and Statistics Applications for Environmental Science, 2007