Turkish J. Eng. Env. Sci.
34 (2010) , 171 – 188.

doi:10.3906/muh-1002-2

Flood frequency analysis using *Mathematica*

Mohammad Ali GHORBANI[1], **Heikki RUSKEEPÄÄ**[2],
Vijay P. SINGH[3], **Bellie SIVAKUMAR**[4,5]
[1]*Department of Water Engineering, Tabriz University, Tabriz-IRAN*
e-mail: cusp2004@yahoo.com
[2]*Department of Mathematics, University of Turku, Turku-FINLAND*
[3]*Department of Biological & Agricultural Engineering, Texas A & M University, College Station, USA*
[4]*School of Civil and Environmental Engineering, The University of New South Wales, Sydney-AUSTRALIA*
[5]*Department of Land, Air and Water Resources, University of California, Davis, USA*

Received 06.02.2010

Abstract

This study analyzes flood frequencies using discharge data from 6 gaging stations in the Aji River basin in Iran. Eighteen different distributions are fitted to the maximum annual discharges from each of these stations, and parameters of these distributions are estimated using the method of maximum likelihood and the method of moments. Calculations are performed with *Mathematica*, a computer algebra system developed by Wolfram Research. The advantage of using this software is that the symbolic, numerical, and graphical computations can be combined and all quantities can be accurately calculated; in particular, there is no need to resort to any approximate methods for the calculation of quantiles. There is a ready-to-use command for calculating quantiles from distributions that are built in *Mathematica*, while for other distributions they can be easily and accurately calculated by inverting the cumulative distribution functions or by solving nonlinear equations where the inversion is not possible. The best distribution is selected based on the root mean square error (RMSE), the coefficient of determination (R^2), and the probability plot correlation coefficient (PPCC). Relations between the distributions' parameters and the area, average discharge, and time of concentration are explored. The complete *Mathematica* code and sample data files are included in http://users.utu.fi/ruskeepa/.

Key Words: Flood frequency analysis, Probability distribution, Annual discharge, Aji River basin, *Mathematica*

Introduction

Planning, design, and management of water resources systems often require knowledge of flood characteristics, such as peak, volume, and duration. Of fundamental importance in many design problems is the determination of the probability distribution of maximum annual discharge. Based on an assumed probability distribution, one can compute statistics of flows of various magnitudes, which can then be used for planning, design, and management of water resources projects.

Among the many probability distributions, the ones that are commonly used in stochastic hydrology are the normal, log-normal, gamma, Weibull, Pearson type III, log-Pearson type III, and extreme value distributions (e.g., Hromadka and Whitley, 1989; Moughamian et al., 1987; Robert, 1987; Opere et al., 2006). The log-normal and Pearson distributions seem to adequately fit peak rainfall and stream-flow, while the Weibull and extreme value distributions are commonly used for these and other extremes of hydrologic variables (Aksoy, 2000; Burn and Goel, 2001). The selection of the best fitting distribution(s) among the many available ones has always been a great challenge. An excellent review of the issues involved in the selection of the most appropriate distribution (for a region or country) was made by Cunnane (1989) in an operational hydrology report for the World Meteorological Organization (WMO). Nevertheless, it is all too common to employ a host of distributions, as the following examples for different regions around the world reflect.

Benson (1968), Wallis (1988), and Vogel et al. (1993) used several distributions for describing flood flows in the USA. The Natural Environment Research Council undertook a flood study for UK conditions (NERC, 1975). McMahon and Srikanthan (1981) evaluated flood distributions for Australian conditions. Rossi et al. (1984) and Ahmed et al. (1988) studied flood distributions for Italy and Scotland, respectively. The flood distributions for Turkey were investigated by Haktanır (1991, 1992) and Haktanır and Horlacher (1993), while Mutua (1994) compared several frequency distributions for floods in Kenya. An extensive study on the selection of the probability distribution function of annual maximum, mean, and minimum stream-flows in the USA was performed by Vogel and Wilson (1996), who analyzed flow data observed from a large network of 1455 stations.

In a similar vein, an attempt is made in the present study to perform flood frequency analysis for Iranian conditions. To this end, flow series from 6 stations (AharChai, Hervy, Lighvan, Moshiran, SofiChai, Vanyar) in the Aji River basin in eastern Azerbaijan are studied. Eighteen different distributions are fitted to the maximum annual discharges from each of these stations. These distributions include: exponential, Frechet, gamma, generalized Pareto, inverse gamma, inverse Gaussian, Kumaraswamy, log-normal, log-Pearson type III, Maxwell, Rayleigh, truncated Cauchy, truncated extreme value, truncated Gumbel, truncated logistic, truncated normal, truncated Pearson type III, and Weibull. Parameters of these distributions are obtained using maximum likelihood estimation (MLE) and the method of moments (MOM). The performances of these distributions are evaluated using 3 statistical criteria: the root mean square error (RMSE), the coefficient of determination (R^2), and the probability plot correlation coefficient (PPCC).

In addition to the contribution to the regional hydrology of Iran, the novelty of this study is the use of the software *Mathematica* (www.wolfram.com) for flood frequency calculations. This software has extensive symbolic and numerical capabilities and, thus, enables us to do calculations in a simpler, faster, and more accurate way. It has several statistical distributions already built-in, and there is also a ready-to-use command for calculation of quantiles. Even for distributions that are not embedded in *Mathematica* (and thus where calculations are not possible explicitly), quantiles can be calculated by solving nonlinear equations. In this study, *Mathematica* version 7 (released in 2008) is used. For interactive estimation of densities with *Mathematica*, see Ruskeepää and Ghorbani (2010).

Probability Density Functions

Eighteen different probability distributions are considered in this study, some of which are very widely used in hydrologic frequency analysis. These distributions and their probability density functions are presented in

Table 1. Probability distributions and their density functions.

Distribution	PDF	Assumption	Domain
Exponential	$\lambda e^{-\lambda(x-\gamma)}$	$\lambda > 0$	$x > \gamma$
Frechet	$\frac{c\alpha^c}{x^{c+1}}e^{-(\frac{\alpha}{x})^c}$	$c > 0, \alpha > 0$	$x > 0$
Gamma	$\frac{x^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)}e^{-\frac{x}{\beta}}$	$\alpha > 0, \beta > 0$	$x > 0$
Generalized Pareto	$\frac{1}{\sigma}(1+k\frac{x-\mu}{\sigma})^{-\frac{1}{k}-1}$	$k > 0, \sigma > 0\ 0 < \mu < \frac{\sigma}{k}$	$x > \mu$
Inverse gamma	$\frac{\beta^\alpha}{\Gamma(\alpha)x^{\alpha+1}}e^{-\frac{\beta}{x}}$	$\alpha > 0, \beta > 0$	$x > 0$
Inverse Gaussian	$\sqrt{\frac{\lambda}{2\pi x^3}}e^{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}}$	$\lambda > 0, \mu > 0$	$x > 0$
Kumaraswamy	$\frac{pq}{b}(\frac{x}{b})^{p-1}(1-(\frac{x}{b})^p)^{q-1}$	$p > 0, q > 0, b > 0$	$0 < x < b$
Log-normal	$\frac{1}{\sqrt{2\pi}\sigma x}e^{-\frac{1}{2}(\frac{\log(x)-\mu}{\sigma})^2}$	$\sigma > 0$	$x > 0$
Log-Pearson type III	$\frac{1}{\Gamma(\alpha)\beta}(\frac{x-\varepsilon}{\beta})^{\alpha-1}e^{-\frac{x-\varepsilon}{\beta}}$	$\alpha > 0, \beta > 0$	$x > e^\varepsilon$
Maxwell	$\frac{\sqrt{2}x^2}{\sqrt{\pi}\sigma^3}e^{-\frac{1}{2}(\frac{x}{\sigma})^2}$	$\sigma > 0$	$x > 0$
Rayleigh	$\frac{x}{\sigma^2}e^{-\frac{1}{2}(\frac{x}{\sigma})^2}$	$\sigma > 0$	$x > 0$
Truncated Cauchy	$\frac{1}{b\pi}(1+(\frac{x-a}{b})^2)^{-1}(\frac{1}{2}+\frac{1}{\pi}\tan^{-1}(\frac{a}{b}))^{-1}$	$b > 0$	$x > 0$
Truncated extreme value	$\frac{1}{\beta}(1-e^{-e^{\frac{\alpha}{\beta}}})^{-1}e^{-\frac{x-\alpha}{\beta}-e^{-\frac{x-\alpha}{\beta}}}$	$\beta > 0$	$x > 0$
Truncated Gumbel	$\frac{1}{\beta}e^{e^{-\frac{\alpha}{\beta}}}e^{\frac{x-\alpha}{\beta}-e^{\frac{x-\alpha}{\beta}}}$	$\beta > 0$	$x > 0$
Truncated logistic	$\frac{1+e^{-\frac{\mu}{\beta}}}{\beta}e^{-\frac{x-\mu}{\beta}}(1+e^{-\frac{x-\mu}{\beta}})^{-2}$	$\beta > 0$	$x > 0$
Truncated normal	$\frac{\sqrt{2}}{\sigma\sqrt{\pi}}(1+erf(\frac{\mu}{\sigma\sqrt{2}}))^{-1}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$	$\sigma > 0$	$x > 0$
Truncated Pearson type III	$\frac{1}{\beta\Gamma(\alpha,-\frac{\varepsilon}{\beta})}(\frac{x-\varepsilon}{\beta})^{\alpha-1}e^{-\frac{x-\varepsilon}{\beta}}$	$\alpha > 0, \beta > 0, \varepsilon < 0$	$x > 0$
Weibull	$\frac{\alpha x^{\alpha-1}}{\beta^\alpha}e^{-(\frac{x}{\beta})^\alpha}$	$\alpha > 0, \beta > 0$	$x > 0$

Table 1. Only truncated (rather than whole) versions of Cauchy, extreme value, Gumbel, logistic, normal, and Pearson type III distributions are used. The usual domains for these 6 distributions are the whole real line for the extreme value as well as normal density functions and values larger than ε for the Pearson density function, which can, in principle, be any real number, but would be negative in this study. Since discharge is always non-negative, it is more realistic to truncate the density functions so that they yield a domain that consists of only non-negative values. The truncation is done by simply dividing the original density function by a suitable constant, to make the integral of the truncated density function equal to one; the constant is given by $P(X \geq 0) = 1 - F(0)$, where $F(x)$ is the cumulative distribution function of the original distribution. The CDF of the truncated distribution is then calculated by integrating the truncated density from 0 to x. [Note: In the truncated Pearson type III density function, the term $\Gamma\left(\alpha, -\frac{\varepsilon}{\beta}\right)$ is the value of the incomplete gamma function $\Gamma(a, z) = \int_z^\infty t^{a-1} e^{-t} dt$].

The Pearson type III distribution has been adopted in some countries as the standard distribution for flood frequency analysis because of its better performance (Sumioka et al., 1997). If $\varepsilon = 0$, then this distribution reduces to the gamma distribution. The extreme value distribution is the limiting distribution for the largest values in large samples drawn from a variety of distributions, including normal, exponential, and Weibull distributions.

Estimation of Parameters and Comparison of Probability Density Functions

Many methods are available for estimating the parameters of the above distributions, such as least-squares, maximum likelihood, moments, weighted moments, linear moments, and entropy. Extensive details of these methods are already available in the literature (e.g., Rao and Hamed, 2000; Singh, 1996) and, therefore, are not reported here. In this study, only 2 of these methods are employed: maximum likelihood estimation and the method of moments.

There is no specific reason for preferring these 2 methods against the others, except that they are simple and also sufficient for the purpose of this study. They are neither treated as superior to the other methods nor any effort is made compare with them.

The probability density functions thus fitted are compared using quantiles. Assuming that there are n number of observations, Cunnane's plotting positions are first calculated as: $p_i = (i - 0.4)/(n + 0.2)$ for $i = 1, \ldots, n$, where i is the order of the i^{th} observation arranged in ascending order and p_i is the probability of non-exceedance of the i^{th} observation estimated by the Cunnane's plotting position formula. For each of the density functions, the p_i-quantiles, given by Q_{p_i}, $i = 1, \ldots, n$, are calculated. These quantiles are then compared with the observed values, denoted as x_i, the i^{th} ordered value. Three statistical indicators are used to compare the computed and the observed quantiles (e.g., O'Donnell 1985): (1) the root mean square error (RMSE), which is the square root of $\frac{1}{n} \sum_{i=1}^{n} (Q_{p_i} - x_i)^2$; (2) the coefficient of determination (R^2), which is the square of the coefficient of correlation between the computed and the observed quantiles; and (3) the probability plot correlation coefficient (PPCC). The probability plot correlation coefficient (PPCC) test was developed by Filliben (1975), and it is a simple but powerful goodness-of-fit test. The test uses the correlation r between the ordered observations and the corresponding fitted quantiles Q_{p_i}, determined by plotting position p_i for each x_i. The PPCC test is a measure of linearity of a probability plot. If the sample to be tested is actually drawn from the hypothesized distribution, the curve formed by the fitted quantiles against observed quantiles is expected to be nearly linear and the correlation coefficient will be near to one.

Mathematica contains distribution functions, density functions, and also quantile functions for exponential, gamma, inverse gamma, inverse Gaussian, log-normal, Maxwell, Rayleigh, and Weibull distributions, and also for some others that are not employed in this study such as Laplace and Levy distributions. Further, for the Frechet, generalized Pareto, Kumaraswamy, truncated Cauchy, truncated extreme value, truncated Gumbel, truncated logistic, and truncated normal distributions, one can easily calculate the density distribution, and quantile functions. However, for the truncated Pearson type III and log-Pearson type III distributions, one can explicitly calculate only the density and distribution functions, but not the quantile functions. Therefore, for these distributions, the needed quantiles are calculated numerically by solving the corresponding nonlinear equations.

Data Analysis and Results

In this study, flood frequency analysis is performed for the Aji River basin in Iran. The Aji River basin is approximately 13700 km^2 and is situated in the eastern part of the Lake Urmieh in the north west of Iran. For the present analysis, 6 gaging stations within a sub-basin of the Aji River basin are considered: AharChai, Hervy, Lighvan, Moshiran, SofiChai, and Vanyar. Figure 1 presents a map of the Aji River basin, wherein the locations of these 6 stations are also indicated. The data considered for the flood frequency analysis are the annual maximum discharge values. Figure 2 shows the variations of these discharge values for the 6 stations (in the above order), and Table 2 presents their annual flood data. In Table 3, some basic characteristics of these stations and of the associated flows (i.e. area, mean flow, time of concentration) are presented. For all these stations, the time of concentration (T_c) is computed using the Bransby Williams (Institution of Engineers Australia, 1987) method.

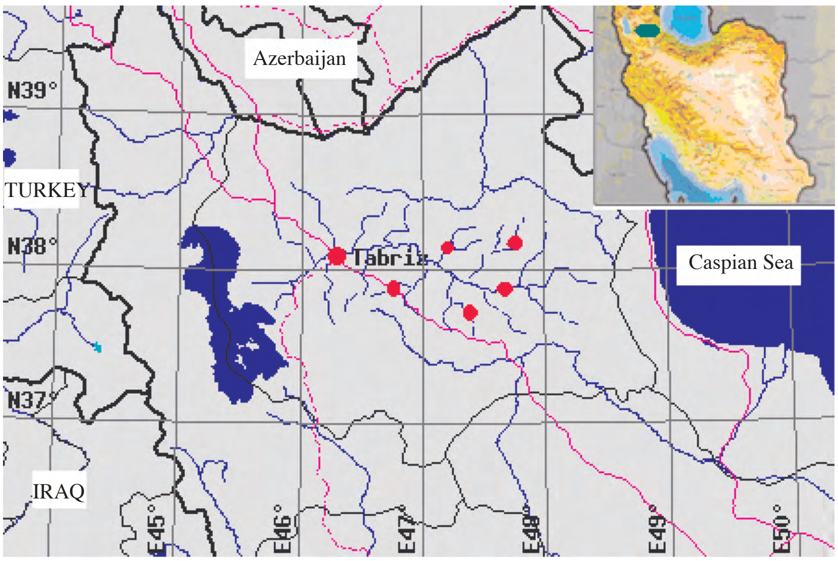

Figure 1. Map of the Aji River basin and locations of the 6 gaging stations.

Table 2. Annual flood data for hydrometric stations.

Year	AharChai	Hervy	Lighvan	Moshiran	SofiChai	Vanyar
1950	-	-	-	-	-	51
1951	-	-	-	-	-	8.07
1952	-	-	-	-	-	52
1953	-	-	-	-	-	27.8
1954	-	-	-	-	-	72.9
1955	-	-	-	-	-	20.8
1956	-	-	-	-	-	87.6
1957	-	-	-	-	-	62
1958	-	-	-	-	-	18.06
1959	-	-	-	-	-	56
1960	-	-	-	-	-	27.8
1961	-	-	-	-	-	28.6
1962	-	-	-	-	-	16.56
1963	-	-	-	-	-	71.38
1964	-	-	-	-	-	109.41
1965	-	-	-	-	-	47.3
1966	-	-	-	-	-	50.27
1967	-	-	-	-	-	123.3
1968	-	-	-	-	-	53.9
1969	-	-	-	-	-	178.8
1970	-	-	-	-	-	40.8
1971	-	-	-	-	-	30.8
1972	11	4.5	-	32.8	35.9	62.42
1973	11.9	4.38	-	77.4	31.93	44.2
1974	8.43	5.44	-	46.2	19.5	83.6
1975	3.41	2.9	-	23.33	9.3	50.1
1976	3.65	4.44	-	45.19	11.88	95.07
1977	8.77	30.04	-	16.37	23.68	36.31
1978	6.94	43	-	54.05	40.9	34.31
1979	48.5	5.35	-	67.6	56.25	47.56
1980	44.7	2.05	-	118.55	35.41	55.81
1981	29.36	5.25	-	57.5	30.43	70.82
1982	53.7	5.96	-	231	40.18	54.65
1983	48.12	8.18	-	127	47.77	54.4
1984	24.4	5.58	-	110.3	55.05	30.52
1985	41.3	2.87	4.04	87.1	14.86	114.22
1986	48.5	4.72	3.75	126	34.5	43.82
1987	24.6	1.67	1.98	70.4	9.11	43.82
1988	33.6	6.5	2.37	128	33.92	107.7
1989	55.5	7.93	3.44	72.8	35.9	24.59
1990	45.6	6.75	8.65	140	22.65	32.56
1991	63.5	4.8	4.27	133	33.9	31.72
1992	15.7	3.05	3.66	118	15.3	74.4
1993	67.6	6.26	4.07	139.9	22.74	73.41
1994	24.84	4.17	5.42	86.4	45.6	84.6
1995	33.8	4.33	1.65	171	34.9	51.73
1996	30.5	4.66	3.48	156	54.6	53.66
1997	31.6	5.59	2.12	119	21.7	32.4
1998	33.4	3.63	4.2	105	25.4	21.26
1999	51.4	5.77	6.18	112.6	31.6	6.78
2000	55.61	3.74	3.66	197.1	30.7	16.4
2001	18.4	2.36	2.44	158	34.7	3.15
2002	29.11	4.65	3.9	191.69	21.28	34.93
2003	56.87	3.6	5.38	75.95	39.25	65.3
2004	31.79	14.73	4.01	468.2	37.54	32.8
2005	48.89	2.78	3.84	235.2	27.16	39.1

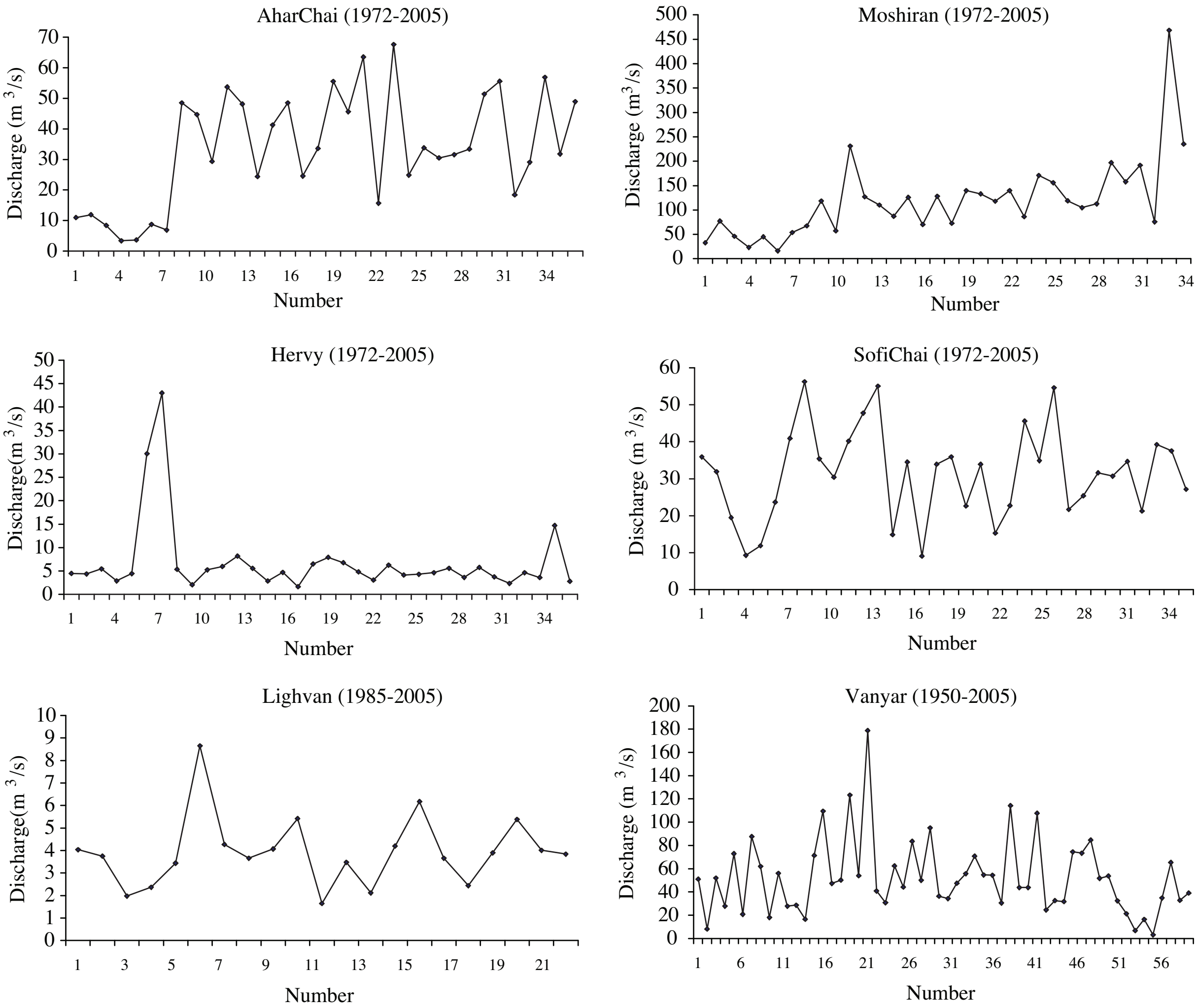

Figure 2. Maximum annual discharge at the 6 gaging stations.

Table 3. Some elementary characteristics of river basin and data.

Characteristics	AharChai	Hervy	Lighvan	Moshiran	SofiChai	Vanyar
Area (km^2)	2058.8	135	76	11454	240.5	7586
Q_{mean}(m^3/s)	33.68	6.81	3.93	120.55	31.34	52.56
Time of Concentration (h)	17.02	5.6	3.38	34.6	6.91	29.83

A station-wise flood frequency analysis was carried out using all the above 18 distribution functions and with the 2 parameter estimation methods (i.e. MLE and MOM). Tables 4 and 5 show the magnitudes of the distribution parameters estimated by the maximum likelihood method and the method of moments, respectively. Among the 18 considered, some distributions have no solution due to calculation difficulties and, thus, are not referred to in Tables 4 and 5 (exponential and generalized Pareto distributions for MLE; Kumaraswamy, log-Pearson Type III, and truncated Cauchy distributions for MOM).

Table 4. Distribution parameter values estimated using the MLE method.

Distributions	AharChai	Hervy	Lighvan	Moshiran	SofiChai	Vanyar
Frechet	$c = 1.08$ $\alpha = 17.39$	$c = 2.03$ $\alpha = 3.87$	$c = 2.62$ $\alpha = 3.01$	$c = 1.35$ $\alpha = 69.37$	$c = 1.91$ $\alpha = 22.31$	$c = 1.09$ $\alpha = 28.95$
Gamma	$\alpha = 2.33$ $\beta = 14.44$	$\alpha = 1.89$ $\beta = 3.61$	$\alpha = 7.10$ $\beta = 0.55$	$\alpha = 2.62$ $\beta = 45.98$	$\alpha = 5.58$ $\beta = 5.61$	$\alpha = 2.59$ $\beta = 20.30$
Inverse gamma	$\alpha = 1.41$ $\beta = 25.40$	$\alpha = 3.30$ $\beta = 14.40$	$\alpha = 6.92$ $\beta = 23.50$	$\alpha = 2.15$ $\beta = 164.48$	$\alpha = 4.33$ $\beta = 109.71$	$\alpha = 1.58$ $\beta = 47.78$
Inverse Gaussian	$\lambda = 38.81$ $\mu = 33.68$	$\lambda = 12.09$ $\mu = 6.81$	$\lambda = 24.99$ $\mu = 3.93$	$\lambda = 210.21$ $\mu = 120.55$	$\lambda = 132.58$ $\mu = 31.34$	$\lambda = 71.06$ $\mu = 52.56$
Kumaraswamy	$b = 68.36$ $p = 1.25$ $q = 1.29$	$b = 507462.64$ $p = 1.18$ $q = 527549.09$	$b = 636.83$ $p = 2.66$ $q = 542382.91$	$b = 31477.04$ $p = 1.61$ $q = 6575.52$	$b = 97.92$ $p = 2.70$ $q = 15.08$	$b = 29684.79$ $p = 1.72$ $q = 43751.44$
Log-normal	$\mu = 3.29$ $\sigma = 0.79$	$\mu = 1.63$ $\sigma = 0.64$	$\mu = 1.30$ $\sigma = 0.38$	$\mu = 4.59$ $\sigma = 0.67$	$\mu = 3.35$ $\sigma = 0.46$	$\mu = 3.76$ $\sigma = 0.71$
Log-Pearson type III	$\alpha = 328.34$ $\beta = 0.05$ $\varepsilon = -11.56$	$\alpha = 228.79$ $\beta = 0.04$ $\varepsilon = -7.64$	$\alpha = 1442.80$ $\beta = 0.01$ $\varepsilon = -13.25$	$\alpha = 6266.39$ $\beta = 0.01$ $\varepsilon = -48.32$	$\alpha = 7457.84$ $\beta = 0.01$ $\varepsilon = -36.55$	$\alpha = 262.15$ $\beta = 0.05$ $\varepsilon = -8.31$
Maxwell	$\sigma = 22.06$	$\sigma = 6.03$	$\sigma = 2.44$	$\sigma = 83.97$	$\sigma = 19.41$	$\sigma = 35.53$
Rayleigh	$\sigma = 27.02$	$\sigma = 7.39$	$\sigma = 2.98$	$\sigma = 102.84$	$\sigma = 23.78$	$\sigma = 43.52$
Truncated Cauchy	$a = 31.46$ $b = 16.08$	$a = 4.60$ $b = 1.29$	$a = 3.82$ $b = 0.57$	$a = 101.65$ $b = 45.41$	$a = 31.87$ $b = 8.09$	$a = 43.12$ $b = 19.90$
Truncated extreme value	$\alpha = 23.42$ $\beta = 17.58$	$\alpha = 4.26$ $\beta = 3.23$	$\alpha = 3.24$ $\beta = 1.20$	$\alpha = 85.96$ $\beta = 55.85$	$\alpha = 25.30$ $\beta = 11.33$	$\alpha = 37.76$ $\beta = 24.55$
Truncated Gumbel	$\alpha = 36.99$ $\beta = 21.81$	$\alpha = -2427294$ $\beta = 232544$	$\alpha = 4.18$ $\beta = 2.41$	$\alpha = 0.00$ $\beta = 213.38$	$\alpha = 35.73$ $\beta = 13.87$	$\alpha = 12.29$ $\beta = 78.26$
Truncated logistic	$\beta = 12.67$ $\mu = 30.87$	$\beta = 4.77$ $\mu = -0.53$	$\beta = 0.82$ $\mu = 3.78$	$\beta = 47.10$ $\mu = 98.25$	$\beta = 7.32$ $\mu = 30.94$	$\beta = 20.20$ $\mu = 44.01$
Truncated normal	$\mu = 30.82$ $\sigma = 20.55$	$\mu = -454.05$ $\sigma = 56.61$	$\mu = 3.90$ $\sigma = 1.57$	$\mu = 69.22$ $\sigma = 113.18$	$\mu = 31.11$ $\sigma = 12.48$	$\mu = 41.72$ $\sigma = 39.94$
Truncated Pearson type III	$\alpha = 1364.91$ $\beta = 0.55$ $\varepsilon = -725.12$	$\alpha = 0.00$ $\beta = 8.09$ $\varepsilon = -31.07$	$\alpha = 7.10$ $\beta = 0.55$ $\varepsilon = 0.00$	$\alpha = 2.00$ $\beta = 120.38$ $\varepsilon = -85411.15$	$\alpha = 305.26$ $\beta = 0.71$ $\varepsilon = -185.34$	$\alpha = 3.87$ $\beta = 15.87$ $\varepsilon = -9.11$
Weibull	$\alpha = 1.87$ $\beta = 37.71$	$\alpha = 1.18$ $\beta = 7.32$	$\alpha = 2.66$ $\beta = 4.42$	$\alpha = 1.61$ $\beta = 135.28$	$\alpha = 2.81$ $\beta = 35.21$	$\alpha = 1.72$ $\beta = 59.01$

Distribution Parameters and Their Relations to Basin Characteristics

Using regression analysis, distribution parameters estimated by the maximum likelihood method and the method of moments are related to watershed area (A), mean of discharge (Q_{mean}), and time of concentration (T_c). The results of this regression analysis are given in Tables 6 and 7 for MLE and MOM, respectively. This analysis could be an alternative for estimating flood peaks of various return periods for ungaged stream in the same basin. Based on the results in Tables 6 and 7, relations between the selected distribution parameters and the basin characteristics may be discussed as follows. For brevity, only some are discussed.

For the Weibull distribution, parameter β has good correlation with A, Q_{mean}, and T_c. In other words, the value of β increases with increasing area, discharge, and concentration time, with the largest β corresponding to the largest area, discharge, and concentration time.

For the truncated Pearson type III distribution, parameter β has a meaningful relationship with A, Q_{mean}, and T_c, but the other parameters have no such relationship (except parameter ε with Q_{mean}). For the truncated Gumbel distribution, there are no meaningful relationships between the parameters (α and β) and A and Q_{mean} (and even T_c). For the gamma distribution, parameter β has a direct and meaningful relationship with A, Q_{mean}, and T_c, while parameter α has an inverse relationship with all the basin characteristics. For the log-normal distribution, parameter μ has a meaningful relationship with Q_{mean} and T_c, while parameter σ seems to have, in general but not always, no meaningful relationship with A, Q_{mean}, and T_c. For the truncated normal distribution, the parameter σ has meaningful relationships with basin characteristics (i.e. an increase in the parameter value is observed for an increase in A and Q_{mean}) in the case of MOM.

Table 5. Distribution parameter values estimated using the MOM method.

Distributions	AharChai	Hervy	Lighvan	Moshiran	SofiChai	Vanyar
Exponential	$\gamma = 15.61$ $\lambda = 0.06$	$\gamma = -1.11$ $\lambda = 0.13$	$\gamma = 2.39$ $\lambda = 0.65$	$\gamma = 39.18$ $\lambda = 0.01$	$\gamma = 19.14$ $\lambda = 0.08$	$\gamma = 20.54$ $\lambda = 0.03$
Frechet	$\alpha = 26.21$ $c = 3.43$	$\alpha = 4.47$ $c = 2.41$	$\alpha = 3.26$ $c = 4.24$	$\alpha = 89.14$ $c = 3.01$	$\alpha = 25.99$ $c = 4.26$	$\alpha = 39.78$ $c = 3.18$
Gamma	$\alpha = 3.48$ $\beta = 9.69$	$\alpha = 0.74$ $\beta = 9.21$	$\alpha = 6.56$ $\beta = 0.60$	$\alpha = 2.19$ $\beta = 54.93$	$\alpha = 6.60$ $\beta = 4.75$	$\alpha = 2.69$ $\beta = 19.51$
Generalized Pareto	$\mu = 1.77$ $\sigma = 65.72$ $k = -1.06$	$\mu = 0.20$ $\sigma = 5.61$ $k = 0.15$	$\mu = 2.12$ $\sigma = 2.17$ $k = -0.20$	$\mu = 42.28$ $\sigma = 75.34$ $k = 0.04$	$\mu = 11.25$ $\sigma = 37.29$ $k = -0.86$	$\mu = 16.40$ $\sigma = 41.14$ $k = -0.14$
Inverse gamma	$\alpha = 5.48$ $\beta = 150.75$	$\alpha = 2.74$ $\beta = 11.85$	$\alpha = 8.56$ $\beta = 29.69$	$\alpha = 4.19$ $\beta = 385.11$	$\alpha = 8.60$ $\beta = 238.25$	$\alpha = 4.69$ $\beta = 194.15$
Inverse Gaussian	$\lambda = 117.08$ $\mu = 33.68$	$\lambda = 5.04$ $\mu = 6.81$	$\lambda = 25.76$ $\mu = 3.93$	$\lambda = 264.56$ $\mu = 120.55$	$\lambda = 206.91$ $\mu = 31.34$	$\lambda = 141.59$ $\mu = 52.56$
Log-normal	$\mu = 3.39$ $\sigma = 0.50$	$\mu = 1.49$ $\sigma = 0.92$	$\mu = 1.30$ $\sigma = 0.38$	$\mu = 4.60$ $\sigma = 0.61$	$\mu = 3.37$ $\sigma = 0.38$	$\mu = 3.80$ $\sigma = 0.56$
Maxwell	$\sigma = 21.10$	$\sigma = 4.27$	$\sigma = 2.46$	$\sigma = 75.54$	$\sigma = 19.64$	$\sigma = 32.94$
Rayleigh	$\sigma = 26.87$	$\sigma = 5.44$	$\sigma = 3.13$	$\sigma = 96.18$	$\sigma = 25.00$	$\sigma = 41.94$
Truncated extreme value	$\alpha = 25.44$ $\beta = 14.13$	-	$\alpha = 3.24$ $\beta = 1.20$	$\alpha = 77.51$ $\beta = 65.52$	$\alpha = 25.85$ $\beta = 9.51$	$\alpha = 37.20$ $\beta = 25.32$
Truncated Gumbel	$\alpha = 35.35$ $\beta = 24.25$	-	$\alpha = 4.54$ $\beta = 1.49$	$\alpha = 42.09$ $\beta = 171.98$	$\alpha = 36.18$ $\beta = 11.85$	$\alpha = 42.94$ $\beta = 53.44$
Truncated logistic	$\beta = 11.62$ $\mu = 30.62$	-	$\beta = 0.90$ $\mu = 3.87$	$\beta = 59.05$ $\mu = 84.77$	$\beta = 7.11$ $\mu = 30.84$	$\beta = 21.90$ $\mu = 43.35$
Truncated normal	$\mu = 30.82$ $\sigma = 20.55$	-	$\mu = 3.90$ $\sigma = 1.57$	$\mu = 69.22$ $\sigma = 113.18$	$\mu = 31.11$ $\sigma = 12.48$	$\mu = 41.72$ $\sigma = 39.94$
Truncated Pearson type III	$\alpha = 3.48$ $\beta = 9.69$	$\alpha = 0.74$ $\beta = 9.21$	$\alpha = 6.56$ $\beta = 0.60$	$\alpha = 2.19$ $\beta = 54.93$	$\alpha = 6.60$ $\beta = 4.75$	$\alpha = 2.69$ $\beta = 19.51$
Weibull	$\alpha = 1.94$ $\beta = 37.98$	$\alpha = 0.86$ $\beta = 6.32$	$\alpha = 2.77$ $\beta = 4.41$	$\alpha = 1.51$ $\beta = 133.64$	$\alpha = 2.78$ $\beta = 35.20$	$\alpha = 1.69$ $\beta = 58.88$

For the truncated logistic distribution, both the parameters β and μ have meaningful relationships with all the basin characteristics. For the inverse gamma distribution, parameter β has a meaningful relationship with Q_{mean}; parameter α has an inverse relationship with all the basin characteristics. Finally, for the inverse Gaussian distribution, parameter μ has a meaningful relationship with all the basin characteristics, and parameter λ has a direct relationship with Q_{mean} but no meaningful relationship with the other 2 basin characteristics.

On the basis of these observations, it is fair to conclude that (in a majority of the cases considered) the distribution parameters have meaningful relationships with the basin area (A) and the average discharge (Q_{mean}), and they also have (at least in some cases) a direct or inverse relationship with the time of concentration (T_c), as the case may be. In general, the parameters have greater correlations with A and Q_{mean} than they do with T_c.

Table 6. Distribution parameters using the MLE method and relations to basin characteristics.

Distributions	Relation with A	Relation with Q_{mean}	Relation with T_c
Frechet	$\alpha = 0.004A + 7.515$ ($R^2 = 0.833$)	$\alpha = 0.568Q_{mean} + 0.592$ ($R^2 = 0.992$)	$\alpha = 1.570T_c - 1.335$ ($R^2 = 0.736$)
Gamma	$\beta = 0.003A + 3.056$ ($R^2 = 0.916$)	$\beta = 0.387Q_{mean} - 0.989$ ($R^2 = 0.970$)	$\beta = 1.165T_c - 3.827$ ($R^2 = 0.852$)
Inverse gamma	—	$\beta = 1.185Q_{mean} + 15.06$ ($R^2 = 0.712$)	—
Inverse Gaussian	$\mu = 0.008A + 11.55$ ($R^2 = 0.877$)	$\lambda = 1.574Q_{mean} + 16.32$ ($R^2 = 0.778$) $\mu = Q_{mean}$ ($R^2 = 1.000$)	$\mu = 2.863T_c - 4.972$ ($R^2 = 0.795$)
Kumaraswamy	—	—	—
Log-normal	—	$\mu = 0.026Q_{mean} + 1.900$ ($R^2 = 0.777$)	$\mu = 0.082T_c + 1.654$ ($R^2 = 0.741$)
Log-Pearson type III	—	—	—
Maxwell	$\sigma = 0.005A + 7.314$ ($R^2 = 0.885$)	$\sigma = 0.695Q_{mean} - 0.606$ ($R^2 = 0.998$)	$\sigma = 1.990T_c - 4.058$ ($R^2 = 0.793$)
Rayleigh	$\sigma = 0.007A + 8.960$ ($R^2 = 0.885$)	$\sigma = 0.851Q_{mean} - 0.741$ ($R^2 = 0.998$)	$\sigma = 2.438T_c - 4.969$ ($R^2 = 0.793$)
Truncated Cauchy	$a = 0.006A + 11.64$ ($R^2 = 0.834$) $b = 0.003A + 3.475$ ($R^2 = 0.887$)	$a = 0.835\ Q_{mean} + 1.436$ ($R^2 = 0.994$) $b = 0.386Q_{mean} - 0.812$ ($R^2 = 0.981$)	$a = 2.347T_c - 2.004$ ($R^2 = 0.762$) $b = 1.150T_c - 3.442$ ($R^2 = 0.843$)
Truncated extreme value	$\alpha = 0.005A + 8.877$ ($R^2 = 0.862$) $\beta = 0.004A + 4.769$ ($R^2 = 0.895$)	$\alpha = 0.711Q_{mean} + 0.494$ ($R^2 = 0.998$) $\beta = 0.467Q_{mean} - 0.440$ ($R^2 = 0.993$)	$\alpha = 2.015T_c - 2.704$ ($R^2 = 0.778$) $\beta = 1.372T_c - 3.315$ ($R^2 = 0.830$)
Truncated Gumbel	—	—	—
Truncated logistic	$\beta = 0.003A + 3.441$ ($R^2 = 0.911$) $\mu = 0.006A + 10.15$ ($R^2 = 0.840$)	$\beta = 0.391Q_{mean} - 0.743$ ($R^2 = 0.981$) $\mu = 0.828Q_{mean} + 0.200$ ($R^2 = 0.988$)	$\beta = 1.149T_c - 3.174$ ($R^2 = 0.823$) $\mu = 2.362T_c - 3.772$ ($R^2 = 0.780$)
Truncated normal	—	—	—
Truncated Pearson type III	$\beta = 0.008A - 5.973$ ($R^2 = 0.733$)	$\beta = 1.025Q_{mean} - 18.16$ ($R^2 = 0.855$) $\varepsilon = 738.9Q_{mean} - 16575$ ($R^2 = 0.818$)	—
Weibull	$\beta = 0.009A + 12.87$ ($R^2 = 0.877$)	$\beta = 1.123Q_{mean} - 0.106$ ($R^2 = 1$)	$\beta = 3.216T_c - 5.694$ ($R^2 = 0.795$)

Three goodness-of-fit methods including RMSE, R^2, and PPCC are considered to select the best distribution. For the 6 stations, the results for selected distributions are presented in Tables 8 and 9 for the maximum likelihood estimation and for the method of moments, respectively. Figure 3 presents the best estimated density functions (chosen based on RMSE, R^2, and PPCC) for the 6 stations obtained with these 2 parameter estimation methods; for each station, the best curve(s) is presented. Tables 10 and 11 show distribution parameters and discharges exceeding a given value (i.e. quantile) with the given probability for MLE and MOM, respectively.

Table 7. Distribution parameters using the MOM method and relations to basin characteristics.

Distributions	Relation with A	Relation with Q_{mean}	Relation with T_c
Exponential	$\gamma = 0.002A + 6.721$ ($R^2 = 0.731$)	$\gamma = 0.323Q_{mean} + 2.547$ ($R^2 = 0.914$)	$\gamma = 0.917T_c + 1.073$ ($R^2 = 0.714$)
Frechet	$\alpha = 0.006A + 9.564$ ($R^2 = 0.864$)	$\alpha = 0.737Q_{mean} + 0.905$ ($R^2 = 0.998$)	$\alpha = 2.104T_c - 2.667$ ($R^2 = 0.790$)
Gamma	$\beta = 0.003A + 2.626$ ($R^2 = 0.866$)	$\beta = 0.447Q_{mean} - 2.109$ ($R^2 = 0.926$)	$\beta = 1.275T_c - 4.247$ ($R^2 = 0.731$)
Generalized Pareto	$\mu = 0.003A + 1.487$ ($R^2 = 0.826$)	$\mu = 0.361Q_{mean} - 2.669$ ($R^2 = 0.937$)	—
Inverse gamma	—	$\beta = 3.003Q_{mean} + 43.70$ ($R^2 = 0.855$)	—
Inverse Gaussian	$\mu = 0.008A + 11.55$ ($R^2 = 0.877$)	$\lambda = 2.003Q_{mean} + 43.71$ ($R^2 = 0.724$) $\mu = Q_{mean}$ ($R^2 = 1$)	$\mu = 2.863T_c - 4.972$ ($R^2 = 0.795$)
Log-normal	—	$\mu = 0.026Q_{mean} + 1.882$ ($R^2 = 0.756$)	$\mu = 0.084T_c + 1.621$ ($R^2 = 0.732$)
Maxwell	$\sigma = 0.005A + 7.237$ ($R^2 = 0.877$)	$\sigma = 0.626Q_{mean} - 0.000$ ($R^2 = 1$)	$\sigma = 1.794T_c - 3.116$ ($R^2 = 0.795$)
Rayleigh	$\sigma = 0.006A + 9.214$ ($R^2 = 0.877$)	$\sigma = 0.797Q_{mean} - 0.000$ ($R^2 = 1$)	$\sigma = 2.284T_c - 3.969$ ($R^2 = 0.795$)
Truncated extreme value	$\alpha = 0.006A + 1.425$ ($R^2 = 0.724$) $\beta = 0.004A + 3.963$ ($R^2 = 0.891$)	$\alpha = 0.754Q_{mean} - 7.664$ ($R^2 = 0.853$) $\beta = 0.541Q_{mean} - 1.938$ ($R^2 = 0.973$)	$\alpha = 2.206T_c - 12.14$ ($R^2 = 0.707$) $\beta = 1.550T_c - 4.647$ ($R^2 = 0.775$)
Truncated Gumbel	—	—	—
Truncated logistic	$\beta = 0.004A + 3.114$ ($R^2 = 0.883$) $\mu = 0.006A + 4.617$ ($R^2 = 0.735$)	$\beta = 0.486Q_{mean} - 2.137$ ($R^2 = 0.958$) $\mu = 0.800Q_{mean} - 5.007$ ($R^2 = 0.864$)	$\beta = 1.389T_c - 4.497$ ($R^2 = 0.758$) $\mu = 2.356T_c - 10.020$ ($R^2 = 0.726$)
Truncated normal	$\sigma = 0.007A + 13.79$ ($R^2 = 0.711$)	$\sigma = 0.813Q_{mean} + 5.221$ ($R^2 = 0.757$)	—
Truncated Pearson type III	$\beta = 0.003A + 2.626$ ($R^2 = 0.866$)	$\beta = 0.447Q_{mean} - 2.109$ ($R^2 = 0.926$)	$\beta = 1.275T_c - 4.247$ ($R^2 = 0.731$)
Weibull	$\beta = 0.009A + 12.770$ ($R^2 = 0.877$)	$\beta = 1.112Q_{mean} - 0.081$ ($R^2 = 0.999$)	$\beta = 3.193T_c - 5.733$ ($R^2 = 0.799$)

Comparison of Distributions

With the maximum likelihood estimation method, the best distributions are as follows:

AharChai: Kumaraswamy (PPCC, R^2, and RMSE);

Hervy: truncated Cauchy (PPCC and R^2) and Frechet (RMSE);

Lighvan: inverse gamma (PPCC, R^2, and RMSE);

Moshiran: inverse gamma (PPCC and R^2) and log-Pearson type III (RMSE);

Table 8. Performance evaluation for selected distributions using the MLE method.

Distribution	AharChai			Hervy			Lighvan			Moshiran			SofiChai			Vanyar		
	PPCC	R^2	RMSE	PPCC	R^2	RMSE	PPCC	R^2	RMSE	PPCC	R^2	RMSE	PPCC	R^2	RMSE	PPCC	R^2	RMSE
Frechet	0.60	0.36	123.58	0.95	0.91	**3.47**	0.97	0.93	0.90	0.95	0.89	181.85	0.81	0.66	25.47	0.80	0.64	248.56
Gamma	0.95	0.89	6.99	0.82	0.68	4.86	0.96	0.92	0.44	0.95	0.90	26.70	0.98	0.96	2.62	0.99	0.98	**4.40**
Inverse gamma	0.68	0.46	58.09	0.92	0.84	4.57	**0.97**	**0.95**	**0.36**	**0.97**	**0.95**	52.80	0.92	0.84	8.40	0.89	0.79	75.61
Inverse Gaussian	0.86	0.75	15.25	0.87	0.76	4.48	0.97	0.94	0.40	0.97	0.93	21.34	0.96	0.92	4.35	0.98	0.97	11.73
Kumaraswamy	**0.99**	**0.99**	**2.10**	0.83	0.69	4.52	0.94	0.89	0.52	0.94	0.88	28.60	0.99	0.98	1.84	0.98	0.97	5.97
Log-normal	0.87	0.76	15.85	0.87	0.76	4.92	0.97	0.94	0.39	0.97	0.94	19.90	0.96	0.92	4.35	0.99	0.98	10.27
Log-Pearson type III	0.85	0.72	19.33	0.88	0.77	4.95	0.97	0.94	0.38	0.97	0.94	**19.80**	0.96	0.92	4.50	0.98	0.96	15.69
Maxwell	0.98	0.96	5.09	0.73	0.53	6.36	0.95	0.90	0.49	0.92	0.84	40.35	0.99	0.97	2.17	0.97	0.94	11.53
Rayleigh	0.98	0.95	3.93	0.75	0.56	5.94	0.95	0.91	0.65	0.92	0.85	33.82	0.98	0.97	4.18	0.98	0.95	7.92
Truncated Cauchy	0.66	0.44	56.45	**0.96**	**0.93**	3.60	0.96	0.93	0.53	0.95	0.90	109.07	0.75	0.57	24.33	0.83	0.69	78.77
Truncated extreme value	0.95	0.91	6.58	0.80	0.65	5.38	0.97	0.94	0.39	0.95	0.90	27.86	0.97	0.94	3.44	**0.99**	**0.98**	4.66
Truncated Gumbel	0.99	0.99	2.40	0.86	0.75	4.04	0.92	0.85	0.76	0.92	0.85	34.84	0.98	0.97	2.43	0.97	0.94	8.55
Truncated logistic	0.98	0.95	4.18	0.83	0.69	4.66	0.95	0.90	0.52	0.93	0.87	30.94	0.99	0.98	1.90	0.98	0.96	6.58
Truncated normal	0.99	0.97	3.12	0.85	0.73	4.22	0.94	0.88	0.54	0.93	0.86	31.16	0.99	0.98	**1.83**	0.97	0.95	7.54
Truncated Pearson type III	0.98	0.97	3.20	0.87	0.75	3.96	0.96	0.92	0.44	0.96	0.92	40.62	**0.99**	**0.98**	1.83	0.99	0.98	5.20
Weibull	0.97	0.95	4.22	0.83	0.69	4.52	0.94	0.89	0.52	0.94	0.88	28.60	0.99	0.98	1.85	0.98	0.97	5.98

Table 9. Performance evaluation for selected distributions using the MOM method.

Distribution	AharChai			Hervy			Lighvan			Moshiran			SofiChai			Vanyar		
	PPCC	R^2	RMSE	PPCC	R^2	RMSE	PPCC	R^2	RMSE	PPCC	R^2	RMSE	PPCC	R^2	RMSE	PPCC	R^2	RMSE
Exponential	0.90	0.81	7.98	0.86	0.74	4.04	0.96	0.92	0.44	0.96	0.92	23.19	0.92	0.85	4.75	0.99	0.97	5.44
Frechet	0.87	0.75	9.16	**0.94**	**0.88**	4.24	**0.97**	**0.95**	0.46	**0.98**	**0.96**	27.70	0.92	0.84	4.96	0.98	0.96	9.23
Gamma	0.96	0.92	5.18	0.88	0.78	3.73	0.96	0.92	0.43	0.95	0.90	25.20	0.98	0.96	2.42	0.99	0.98	4.52
Generalized Pareto	**0.99**	**0.99**	**2.17**	0.90	0.82	3.44	0.96	0.91	0.45	0.96	0.93	22.63	0.98	0.95	2.63	0.99	0.97	5.35
Inve se gamma	0.91	0.82	7.63	0.93	0.87	3.69	0.97	0.95	**0.38**	0.98	0.95	21.86	0.96	0.91	3.63	0.99	0.98	5.69
Inverse Gaussian	0.93	0.87	6.53	-	-	-	0.97	0.94	0.40	0.96	0.93	21.89	0.97	0.94	3.01	0.99	0.99	4.10
Log-normal	0.93	0.87	6.53	0.92	0.85	**3.22**	0.97	0.94	0.40	0.97	0.94	**21.36**	0.97	0.94	3.02	**0.99**	**0.99**	**3.99**
Maxwell	0.98	0.96	5.33	0.73	0.53	6.19	0.95	0.90	0.49	0.92	0.84	41.29	0.99	0.97	2.17	0.97	0.94	12.07
Rayleigh	0.98	0.95	3.94	0.74	0.55	5.83	0.95	0.91	0.68	0.92	0.85	34.41	0.98	0.97	4.49	0.98	0.95	8.19
Truncated extreme value	0.96	0.91	5.30	-	-	-	0.97	0.94	0.39	0.95	0.90	25.41	0.97	0.94	2.90	0.99	0.98	4.38
Truncated Gumbel	0.99	0.98	2.31	-	-	-	0.90	0.81	0.67	0.91	0.84	33.41	0.98	0.96	2.40	0.96	0.92	9.01
Truncated logistic	0.98	0.95	3.91	-	-	-	0.95	0.90	0.49	0.94	0.89	27.45	0.99	0.98	1.94	0.98	0.97	5.79
Truncated normal	0.99	0.97	3.12	-	-	-	0.94	0.88	0.54	0.93	0.86	31.16	**0.99**	**0.98**	**1.83**	0.97	0.95	7.54
Truncated Pearson type III	0.96	0.92	5.18	0.88	0.78	3.73	0.96	0.92	0.43	0.95	0.90	25.20	0.98	0.96	2.42	0.99	0.98	4.52
Weibull	0.98	0.95	4.01	0.89	0.79	3.63	0.94	0.89	0.52	0.94	0.89	27.67	0.99	0.98	1.84	0.98	0.97	5.84

SofiChai: truncated Pearson type III (PPCC and R^2) and truncated normal (RMSE);

Vanyar: truncated extreme value (PPCC and R^2) and gamma (RMSE).

With the method of moments, the best distributions are as follows:

AharChai: generalized Pareto (PPCC, R^2, and RMSE);

Hervy: Frechet (PPCC and R^2) and log-normal (RMSE);

Lighvan: Frechet (PPCC and R^2) and inverse gamma (RMSE);

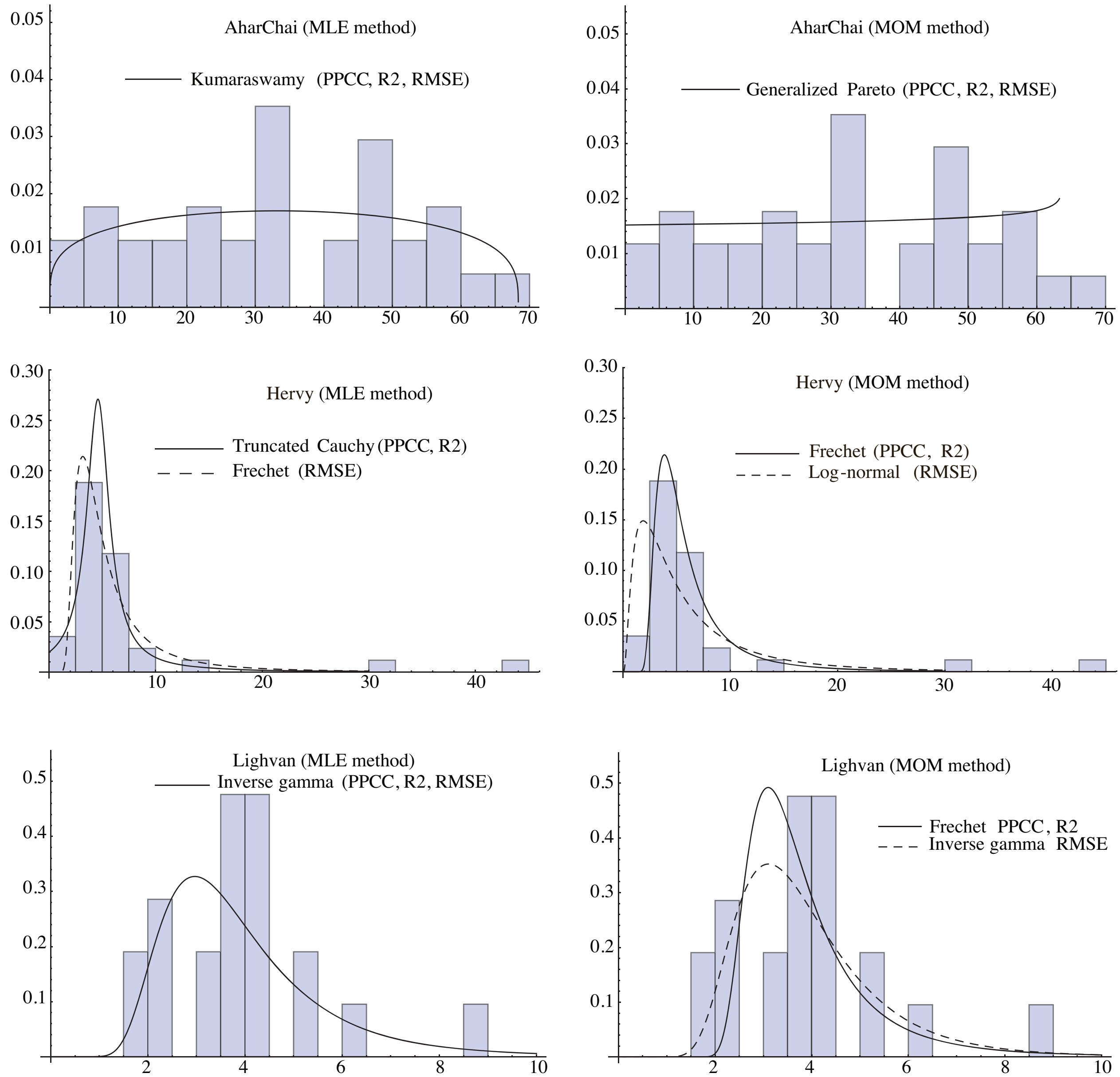

Figure 3. Histograms and the best estimated probability density functions.

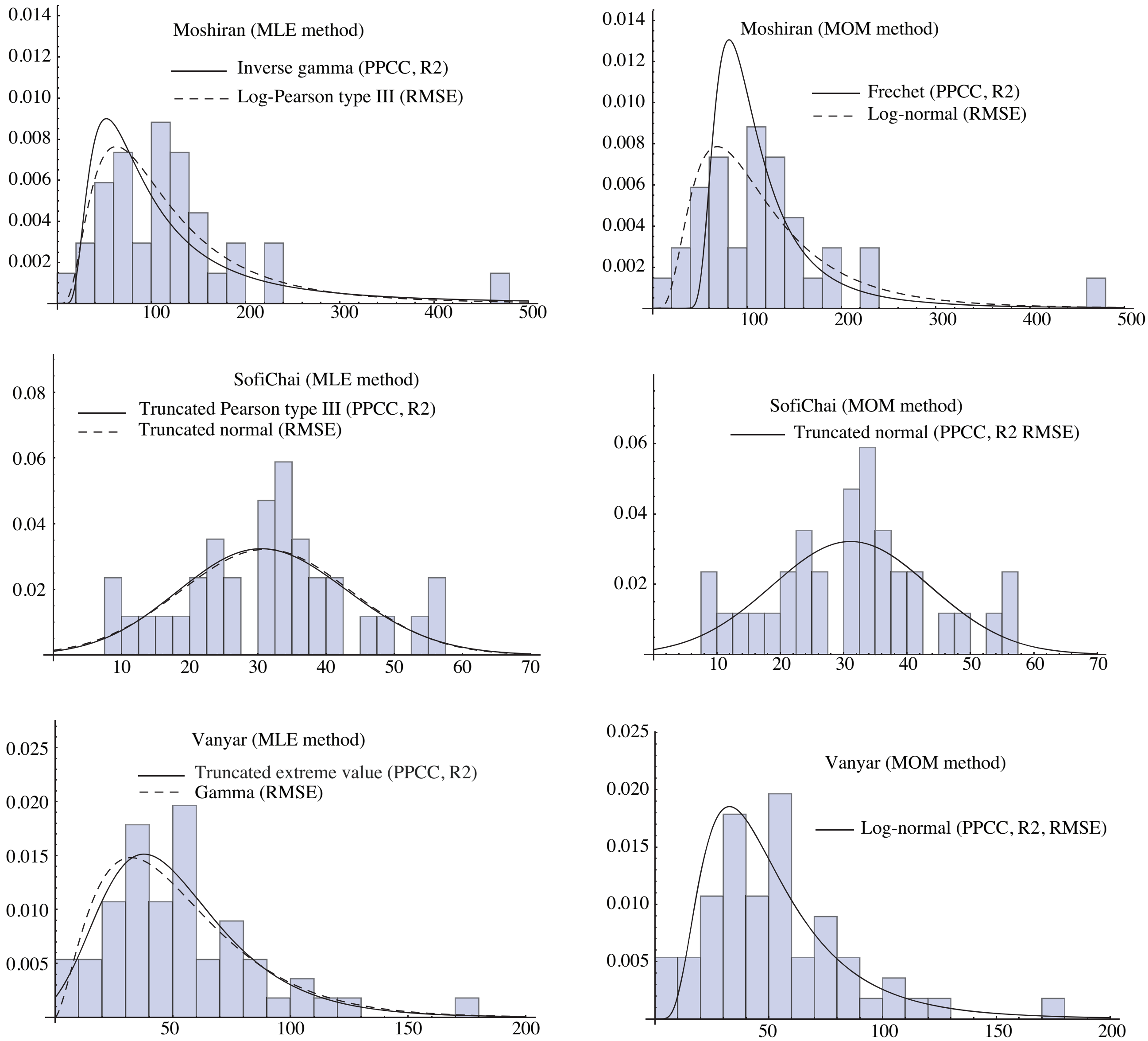

Figure 3. Continued.

Moshiran: Frechet (PPCC and R^2) and log-normal (RMSE);

SofiChai : truncated normal (PPCC, R^2, and RMSE);

Vanyar: log-normal (PPCC, R^2, and RMSE).

Based on these results, it may be inferred that the inverse gamma, log-Pearson type III, and log-normal distributions are generally suitable for both large and small basins, when the maximum likelihood method used for parameter estimation. With the method of moments, however, the Frechet, inverse gamma, and the log-normal distributions seem more suitable. Taking these collectively, the inverse gamma distribution may be suggested as an appropriate distribution for the Aji River basin, and possibly for other Iranian basins, although caution needs to be exercised in making such a generalization.

Table 10. Distribution parameters and discharges exceeding a given value with a given probability (MLE method).

		Discharge quantiles							
Site	Best Distribution	$p = 0.5$	$p = 0.6$	$p = 0.7$	$p = 0.8$	$p = 0.9$	$p = 0.95$	$p = 0.99$	$p = 0.999$
AharChai	Kumaraswamy	33.8298	39.7491	45.7856	52.0904	58.9738	62.9092	66.8007	68.0982
Hervy	Truncated Cauchy	4.77509	5.17804	5.7045	6.58882	8.95855	13.5043	49.4229	452.963
	Frechet	4.63752	5.39131	6.43684	8.11303	11.749	16.7596	37.4602	116.948
Lighvan	Inverse gamma	3.56532	3.93981	4.40044	5.03416	6.12649	7.27171	10.2787	15.809
Moshiran	Inverse gamma	90.1586	109.054	135.463	177.976	269.998	395.525	899.821	2733.43
	Log-Pearson type III	98.1566	116.287	139.463	172.609	232.217	296.922	471.804	795.342
SofiChai	Truncated Pearson type III	31.0194	34.1613	37.5596	41.5838	47.2475	51.9966	61.0802	71.5337
	Truncated normal	31.213	34.3577	37.727	41.6743	47.1535	51.6809	60.1776	69.7051
Vanyar	Truncated extreme value	47.0928	54.554	63.3486	74.8441	93.2521	110.918	150.929	207.569
	Gamma	45.9679	54.0447	63.6767	76.3024	96.3299	115.16	156.308	211.822

Table 11. Distribution parameters and discharges exceeding a given value with a given probability (MOM method).

		Discharge quantiles							
Site	Best Distribution	$p = 0.5$	$p = 0.6$	$p = 0.7$	$p = 0.8$	$p = 0.9$	$p = 0.95$	$p = 0.99$	$p = 0.999$
AharChai	Generalized Pareto	34.0359	40.2996	46.4693	52.5145	58.3733	61.1834	63.3032	63.7326
Hervy	Frechet	5.20021	5.90305	6.85284	8.32668	11.372	15.3349	30.1784	78.6791
	Log-normal	4.44161	5.61449	7.21426	9.67429	14.5325	20.3365	38.1973	77.4256
Lighvan	Frechet	3.54966	3.81429	4.15112	4.63609	5.53262	6.5551	9.6238	16.5728
	Inverse gamma	3.60943	3.94673	4.3555	4.908	5.8371	6.78443	9.16984	13.283
Moshiran	Frechet	100.689	111.437	125.566	146.744	188.307	239.196	411.146	885.05
	Log-normal	99.9149	116.694	137.778	167.339	219.112	273.745	415.623	663.711
SofiChai	Truncated normal	31.213	34.3577	37.727	41.6743	47.1535	51.6809	60.1776	69.7051
Vanyar	Log-normal	44.8838	51.75	60.2632	72.0209	92.2165	113.099	165.865	254.774

Conclusions

In this study, flood frequency analysis was performed for Iranian conditions. Maximum annual discharge values observed at each of 6 gaging stations in the Aji River basin were studied. Eighteen different probability distributions were fitted, and the method of maximum likelihood and the method of moments were used for parameter estimation. The performances of these distributions for different quantiles were compared using root mean square error (RMSE), coefficient of determination (R^2), and probability plot correlation coefficient (PPCC). A regression analysis was carried out to establish relations between the distribution parameters and

3 basin characteristics: area (A), mean discharge (Q_{mean}), and time of concentration (T_c). This study is also the first one to use the software *Mathematica* for performing any type of flood frequency analysis.

The results generally suggest meaningful relationships between the distribution parameters and all the 3 basin characteristics considered, but having far greater correlations with A and Q_{mean} than with T_c. The results also indicate that, among the 18 different distributions, the inverse gamma distribution is the most appropriate for the Aji River basin, followed by the inverse Gaussian distribution.

The present study has important implications for flood frequency analysis for Iran in particular, and for regional hydrology in general. Further, the use of *Mathematica* provides a new dimension to the flood frequency analysis. With the many challenges faced in using the existing methods (often due to difficulties in calculations) for the selection of the most appropriate probability distribution for a given region, the symbolic, numerical, and graphical capabilities of *Mathematica*, together with its flexibility, can go a long way. Future work will focus on advancing the use of *Mathematica* towards developing a more generalized and flexible framework for flood frequency analysis, details of which will be reported elsewhere.

References

Ahmad, M.I., Sinclair, C.D. and Werritty, A., "Log-Logistic Flood Frequency Analysis", J. Hydrology, 98, 205-224, 1988.

Aksoy, H., "Use of Gamma Distribution in Hydrological Analysis", Turk. J. Engin. Environ. Sci., 24, 419-428, 2000.

Benson, M.A., "Uniform Flood Frequency Estimating Methods for Federal Agencies", Water Resour. Res, 4, 891-908, 1968.

Burn, D.H. and Goel, N.K., "Flood Frequency Analysis for the Red River at Winnipeg", Canadian J., Eng, 28, 355-362, 2001.

Cunnane, C., "Statistical Distributions for Flood Frequency Analysis", World Meteorological Organization, Operational Hydrology Report No. 33, WMO, No.718, Secretariat of the World Meteorological Organization, Geneva, Switzerland, 1989.

Filliben, J.J., "The Probability Plot Correlation Coefficient Test for Normality", Technometrics, 17, 111-117, 1975.

Haktanır, T., "Statistical Modeling of Maximum Flows in Turkish Rivers", Hydrol. Sci. J., 36, 367-389, 1991.

Haktanır, T., "Comparison of Various Flood Frequency Distributions Using Annual Flood Peaks Data of Rivers in Anatolia", J. Hydrology, 136, 1-31, 1992.

Haktanır, T. and Horlacher, H.B., "Evaluation of Various Distributions for Flood Frequency Analysis", Hydrol. Sci. J., 38, 15-32, 1993.

Hromadka, T.V. and Whitley, R.J., "Checking Flood Frequency Curves Using Rainfall Data", J. Hydraul. Engg, ASCE, 115, 544-548, 1989.

Institution of Engineers Australia, "Australian Rainfall and Runoff, a Guide to Flood Estimation", (ed. by D.H. Pilgrim), Canberra, 1987.

McMahon, T.A. and Srikanthan, R., "Log-Pearson Type 3 Distribution – Is It Applicable to Flood Frequency Analysis of Australian Streams?", J. Hydrology, 52, 139-147, 1981.

Moughamiam, M.S., McLaughlin, D.B. and Bras, R.E., "Estimation of Flood Frequency: An Evaluation of Two Derived Distribution Procedures", Water Resour. Res, 23, 1309-1319, 1987.

Mutua, F.M., "The Use of the Akaike Information Criterion in the Identification of an Optimum Flood Frequency Model", Hydrol. Sci. J., 39, 235-244, 1994.

NERC, "Flood Studies Reports. Natural Environment Research Council", London, Vols. 1-5, 1975.

O'Donnell, T., "A Direct Three-Parameter Muskingum Procedure Incorporating Lateral Inflow", Hydrol. Sci. J., 30, 494-495, 1985.

Opere, A.O., Mkhandi, S. and Willems, P., "At Site Flood Frequency Analysis for the Nile Equatorial Basins", Phys. Chem. Earth, 31, 919-927, 2006.

Rao, A.R. and Hamed, K.H., "Flood Frequency Analysis", CRC Press, Boca Raton, Florida, USA, 2000.

Robert, M.H., "Plotting Positions for Historic Floods and Their Precision", Eos, AGU Trans. 23, 715-719, 1987.

Rossi, F., Fiorentino, M. and Versaece, P., "Two-Component Extreme Value Distribution for Flood Frequency Analysis", Water. Resour. Res, 20, 847-856, 1984.

Ruskeeṕáá, H. and Ghorbani, M.A., "Maximum Likelihood Estimation of Ordinary and Finite Mixture Distributions", from The Wolfram Demonstrations Project, http://demonstrations.wolfram.com/MaximumLikelihoodEstimationOfOrdinaryAndFiniteMixtureDistrib/, 2010.

Singh, V.P., "Entropy-Based Parameter Estimation in Hydrology", Kluwer Academic Publishers, Dordrecht, the Netherlands, 1996.

Sumioka, S.S., Kresch, D.L. and Kasnick, K.D., "Magnitude and Frequency of Floods in Washington", USGS Water Investigations Report 97-4277, Washington, DC, USA, 1997.

Vogel, R.M. and Wilson, I., "The Probability Distribution of Annual Maximum, Minimum and Average Streamflow in the United States", J. Hydrol. Engg, ASCE 1, 69-76, 1996.

Vogel, R.M., Thomas, W.O. and McMahon, T.A., "Flood Flow Frequency Model Selection in the Southwestern United States", J. Wat. Resour. Plan. Manage, ASCE, 119, 353-366, 1993.

Wanielista, M., Kersten, R. and Eaglin R., "Hydrology Water Quantity and Quality Control", 2^{nd} ed, John Wiley & Sons, 1997.