**Hydrology and
Earth System
Sciences**

Probability distribution of flood flows in Tunisia

H. Abida and M. Ellouze

Department of Earth Sciences, Faculty of Science, Sfax, BP 802, 3018 Tunisia

Abstract. L (Linear) moments are used in identifying regional flood frequency distributions for different zones Tunisia wide. 1134 site-years of annual maximum stream flow data from a total of 42 stations with an average record length of 27 years are considered. The country is divided into two homogeneous regions (northern and central/southern Tunisia) using a heterogeneity measure, based on the spread of the sample L-moments among the sites in a given region. Then, selection of the corresponding distribution is achieved through goodness-of-fit comparisons in L-moment diagrams and verified using an L moment based regional test that compares observed to theoretical values of L-skewness and L-kurtosis for various candidate distributions. The distributions used, which represent five of the most frequently used distributions in the analysis of hydrologic extreme variables are: (i) Generalized Extreme Value (GEV), (ii) Pearson Type III (P3), (iii) Generalized Logistic (GLO), (iv) Generalized Normal (GN), and (v) Generalized Pareto (GPA) distributions. Spatial trends, with respect to the best-fit flood frequency distribution, are distinguished: Northern Tunisia was shown to be represented by the GNO distribution while the GNO and GEV distributions give the best fit in central/southern Tunisia.

1 Introduction

Peak or flood flow is an important hydrologic parameter in the determination of flood risk, management of water resources and design of hydraulic structures such as dams, spillways, culverts and irrigation ditches. The estimate of the design event must be fairly accurate to avoid excessive costs in the case of overestimation of the flood magnitude or excessive damage and even loss of human lives while under-

estimating the flood potential. There is a need, therefore, to estimate how often a specific flood event will occur, or how large a flood will be for a particular probability of exceedence or recurrence interval. This might be achieved through flood frequency analysis procedures, which involve the estimation of distributional parameters and the extrapolation of cumulative distribution functions to generate extreme flood values.

Flood frequency analysis is performed either for a single site, when extensive historic peak flow data are available, or on a regional basis, when there is little or no historic flow data at a particular site. In this latter case, all data from other local basins within the same region are pooled to get an efficient estimate of parameters of a chosen distribution and hence a more robust quantile estimate. Bobee and Rasmussen (1995) reported that the use of regional information allows a reduction of sampling uncertainty by introducing more data, as well as a reduction of model uncertainty by facilitating a better choice of distribution.

Recently, research efforts focused more on regional rather than the conventional at-site flood frequency analysis. Hosking and Wallis (1993) organized regional flood frequency analysis into 4 stages: (i) screening of the data, (ii) identification of homogeneous regions, (iii) choice of a regional probability distribution, and (iv) estimation of the regional flood frequency distribution. Recent research efforts also focused on the use of L-moment diagrams for the identification of flood frequency distributions, such as the studies performed in Bangladesh (Abdul karim and Chowdhury, 1995), New Zealand (Pearson, 1991), Australia (Nathan and Weinmann, 1991), Canada (Pilon and Adamowski, 1992, Van-Thanh and Van Nguyen, 2006), United States (Wallis, 1988; Vogel and Wilson, 1996), China (Jingyi and Hall, 2004), India (Rakesh and Chandranath, 2006) and the globe (Onoz and Bayazit, 1995). In fact, there appear to be a general world-wide agreement among agencies and governments to re-evaluate their flood frequency standard procedures using L-moment based techniques.

Correspondence to: H. Abida
(habib.abida@voila.fr)

In this context, this study uses L-moment diagrams to select the flood frequency distribution that best fits the annual maximum flood flows in Tunisia. The paper first presents a survey of similar previous L-moment-based studies all over the world. Then, the study area and the data used in the numerical analysis are described. Next, the flood frequency identification procedure is presented. Finally, the results of the analysis are discussed and summarized.

2 Literature review

2.1 Global survey of flood frequency models

Many statistical distributions for flood-frequency analysis have been investigated in hydrology. Annual flood series were found to be often skewed, which led to the development and use of many skewed distributions, with the most commonly applied distributions now being the Gumbel (EV1), the Generalized Extreme Value (GEV), the Log Pearson Type III (LP3), and the 3 parameter Lognormal (LN3) (Pilon and Harvey, 1994). The proponents of each distribution have been able to show some degree of confirmation for their particular distribution by comparing theoretical results and measured values. However, there is no theoretical basis for justifying the use of one specific distribution for modeling flood data and long term flood records show no justification for the adoption of a single type of distribution (Benson, 1968).

Different studies were undertaken on distribution selection for flood data in different countries all over the world. Beard (1974) estimated the 1000 year floods at 300 stations in USA with 14 200 station-years of data by eight different models and concluded, based on split sample experiments, that the two parameter lognormal (LN2) and the log Pearson Type III (LP3) were the best. Gunasekara and Cunnane (1992) repeated the split sample experiments of Beard (1974) with synthetic data consisting of samples of 40 events. They concluded that the GEV distribution with probability weighted moments (PWMs) estimated parameters was the best at-site method to estimate the 100 and 1000 year floods and the LP3 with regional skew yielded comparable results.

McMahon and Srikanthan (1981) used the moment ratio diagrams to compare various distributions with the data from 172 streams in Australia and concluded that LP3 was the only one suitable. Farquharson et al. (1987) fit a GEV distribution to annual flood flow data at 1121 gauging stations in 70 different countries using probability weighted moments. McMahon et al. (1992) and Finlayson and McMahon (1992) analyzed annual maximum flood flow data at 974 stations around the world using ordinary product moment diagrams. The authors tested several probability distributions and concluded that the LP3 distribution provided the best fit to observed flood flow data. However, other testing methods should have been used in this study because the estimates of ordinary product moment ratios such as the coefficient of variation and skewness contain significant bias (Vogel and Fennessey, 1993), especially for small and highly skewed samples.

2.2 Standard distributions adopted by national institutions in the world

Based on large scale studies of their own flood data, many countries adopted standard methods to be used by governments or private agencies to achieve uniformity in flood frequency analysis and estimation. A working group in the USA (US Water Resources Council (Benson, 1968) recommended the LP3 distribution whereas a similar study in the United Kingdom (NERC, 1975) proposed the GEV distribution as a standard. The generalized gamma distribution was recommended in the former USSR (Kritsky and Menkel, 1969) while the P3 and the LP3 distributions were generally recommended in West Germany. The LP3 distribution was also advocated by the Institution of Engineers in Australia (Institution of Engineers, Australia (IEA, 1977).

More recently, a worldwide survey of flood frequency methods, prepared for the World Meteorological Organization in 1984 and involving 55 agencies from 28 countries, reported the use of six distributions namely EV1, EV2, GEV, LN2, P3, and LP3. The survey, which was summarized by Cunnane (1989), revealed that EV1, LN2, P3, and LP3 were the most common distributions while only one country used the GEV distribution in spite of its recent popularity.

2.3 L-moments and flood frequency analysis

In the last century, probably one of the most significant scientific contributions to statistical hydrology is the L-moments of Hosking (1990). The advantages of the L-moments are that (i) they characterize a wider range of distributions than conventional moments, (ii) they are less sensitive to outliers in the data, (iii) they approximate their asymptotic normal distribution more closely, and (iv) they are nearly unbiased for all combinations of sample sizes and populations (Hosking and Wallis, 1990).

Wallis (1988), Cunnane (1989) and Hosking (1990) illustrated that, compared to the product moment ratio diagram, the L-moment ratio diagrams possess a better ability to discriminate between distributions. Vogel and Fennessey (1993) reported that conventional product moment estimators should be replaced by L-moment estimators for most goodness-of-fit applications in hydrology. They showed that L-moment diagrams perform always better than ordinary product moment diagrams, regardless of the sample sizes, probability distributions, or skews involved. Cong et al. (1993) reported that L-moment goodness-of-fit tests are more robust than classical single-site goodness-of-fit tests since they use regional rather than single-site data to discriminate between alternative distributions.

Table 1. Previous L-moment based flood frequency studies.

Location	Number of stations	Recommended distribution	Reference
Eastern United States	55	GEV	Wallis (1988)
Central Victoria, Australia	53	GEV	Nathan and Weinmann (1991)
South Island, New Zealand	275	EV1, EV2, and GEV.	Pearson (1991)
New Brunswick, Canada	53	GEV	Gingras and Adamowski (1992)
Nova Scotia, Canada	25	GEV	Pilon and Adamowski (1992)
Southwestern United States	383	LN2, LN3, GEV, and LP3	Vogel et al. (1993a)
Australia	61	GEV, GAP, LP3, and LN3	Vogel et al. (1993b)
United States	1455	LN3, GEV, and LP3	Vogel and Wilson (1996)
Wabash Basin, Indiana	93	P3 and GEV	Rao and Hamed (1997)

Numerous studies have used L-moment diagrams in regional flood frequency analysis, most of which are summarized in Table 1. In spite of this recent tendency of using L-moments world-wide, Klemes (2000a, b) articulated some cautionary notes about their use in hydrological frequency analysis. He claimed that L-moments artificially impose a structure upon a data set and de-emphasize the importance of observed extremes, which leads to the underestimation of extreme design events. However, Alila and Mtiraoui (2002) argued that if the annual floods in a sample are distributed identically and the outliers are caused by sampling variability (for instance, a 100-year event in a 10-year sample) they should not be given an undue weight. If any historic information can be found for any high outlier, a reasonably well-established method, referred to as "flood frequency analysis with historic information", could be used (Pilon and Harvey, 1994). Unfortunately, however, in the absence of any historic information, such high outliers are often either removed from the sample or simply ignored and, consequently, the use of conventional moments would either over- or underestimates the T-year flood event. Therefore, in this case, it is more rational to use a method that is less sensitive to outliers in the data, such as L-moments.

As a conclusion, L-moments provide undeniable advantages over conventional moments in using flood frequency analysis for the estimation of flood quantiles. This is particularly true when considering regional trends in higher order moment statistics. The use of L-moments permits the delineation of regional trends that otherwise might be obscured by biases and sampling variability [variability (Cathcart, 2001).

3 Study area

Tunisia is a relatively small (162 155 km^2) North African country, located at the northeastern tip of Africa at the center of the Mediterranean Sea. Linked on the west to the rest of North Africa by the mighty ridges of the Atlas Mountains, it stretches out to the south into the Sahara, of which it occupies a small part.

Opening on its northern and eastern fronts to the Mediterranean Sea, Tunisia enjoys a clement and mild although notoriously capricious climate. By its latitude it is situated halfway between the temperate zone and the tropical zone, forming thus a meeting place at which cold air masses are confronted by the masses of warm air coming from the tropical regions. It has a rather unstable climate. When it is swept at the equinoxes by tides of opposing depressions, the result is severe cold fronts along with violent storms and frequent downpours. With a general profile stretching lengthwise from north to south, Tunisia shows some climatic variations accentuated by its diversified geographical aspect. The Atlas Mountains stretching from east to west create a variety of large climatic areas distinct from each other mainly by their rainfall.

Rainfall in Tunisia might be crudely characterized by its shortage, irregularity and erratic distribution, all leading either to periods of drought or intensive rainy periods, with storms causing disastrous flooding. The North receives an adequate and fairly reliable rainfall of approximately 600 mm. The central area has an annual mean of 200 mm, whilst the south receives less than 100 mm of rainfall. These averages are subject to very wide annual fluctuations of 1 to 6 for central Tunisia and 1 to 12 for the South.

Generally the country is divided into three main climatic and geographical regions (Fig. 1):

– The High Tell or Northern Tunisia, characterized by its fertile soil and its high degree of moisture. It is an area of high mountains surrounding plains irrigated by the Medjerda River and its tributaries. The western Tell is continued by the northeastern Tell, a maritime area on account of its being deeply penetrated by the Gulf of Tunis and its climatic influences. This is an area of plains and hills crossed by large rivers.

– Central Tunisia is the region that covers the high and low steppes stretching out to the eastern coast. The high steppes represent a region of lofty mountains and wide hollow dips, cut across by large creeks (wadis).
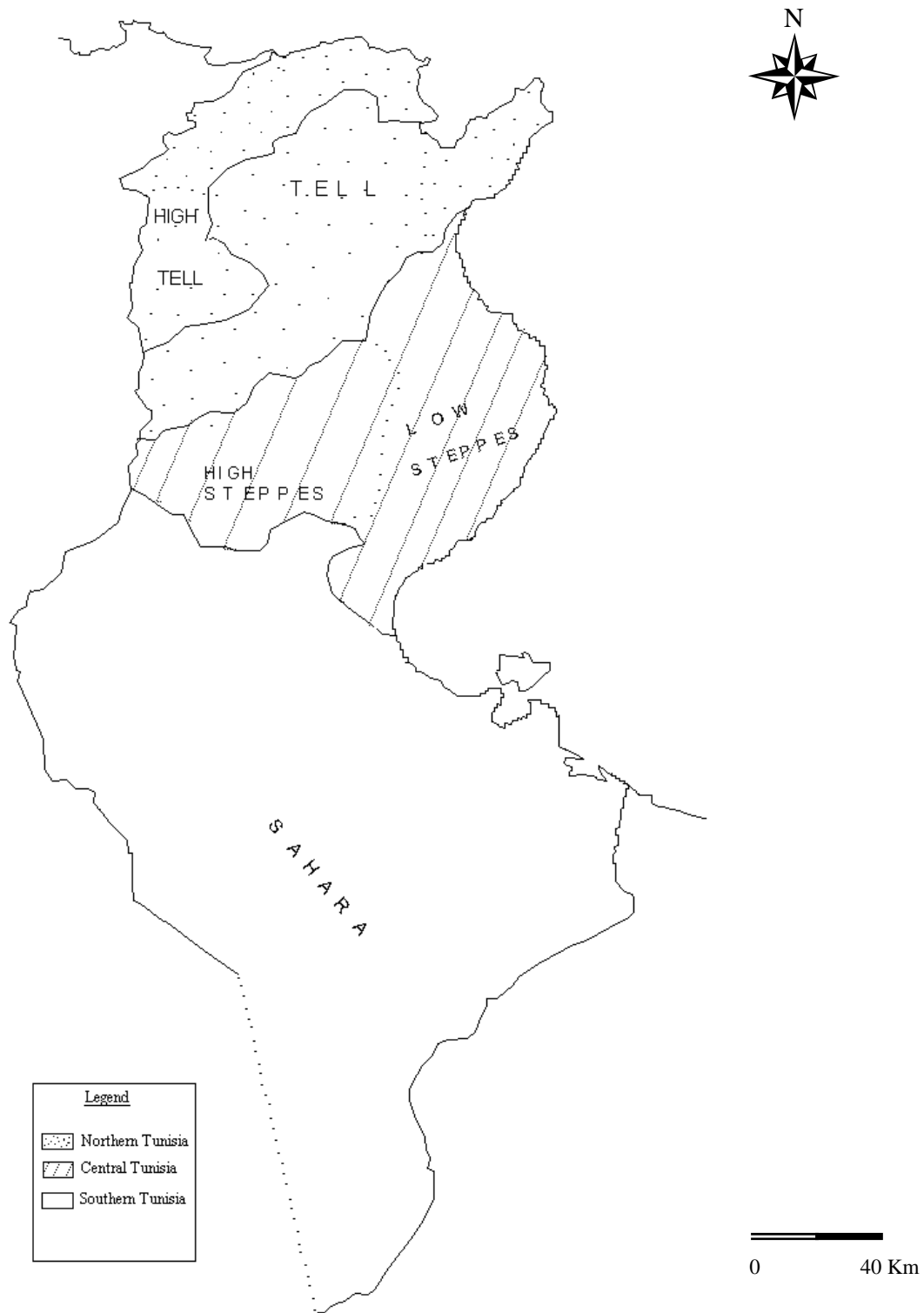
Fig. 1. Physiographic zones of Tunisia.

The vegetation is made up of forests, often stunted, and fields of alfalfa grass. The continental climate contributes to the barrenness of the region. Further to the east the low steppe stretches over wide alluvial plains and hills cut across by large creeks running down the Atlas Mountains.

– Southern Tunisia, bordered on the west by Algeria and on the east by Libya, is jutting out into the Sahara, of which it occupies a part.

4 Data used

A total of 49 annual flood series representing natural hydrologic regimes, obtained from the publications of the Tunisian Ministry of Agriculture and Water Resources, were used for the identification of the appropriate flood frequency distribution.

Discharges were estimated by observing water levels and employing pre-calibrated rating curves to convert measured stages to observed flow rates. Rating curves were determined through velocity measurements using a current meter and graphic integration of the velocity distribution over the entire cross-section. Regulated stations, influenced by the existence of hydraulic structures, were eliminated.

The annual flood data series need to be independent, random, homogeneous, and without trends. These properties were verified by four nonparametric tests using the Consolidated Frequency Analysis (CFA) package of Environment Canada (Pilon and Harvey, 1994). This package allows the user to perform nonparametric tests of homogeneity, trend, independence and randomness; determine T-year events for samples and adjust the appropriate probability distribution. Tests considered include the Spearman tests of independence and trend, the run test for general randomness, and the Main-Whitney split sample test for homogeneity.

Only 37 gauged stations met the screening criteria of having a minimum record length of 10 years, representing unregulated natural flow regimes, and passing all of the nonparametric tests at the 5% level of significance. Nevertheless, all data series were preserved in the analysis by lowering the level of significance to 1% and eliminating inconsistent data values.

5 Procedure used to select a distribution

Initially, the study area was divided into the three main physiographic regions (Fig. 1), for which separate flood frequency analysis procedures were performed. The procedure adopted to select appropriate flood frequency distributions, first, uses the three statistical measures for regional flood frequency analysis of Hosking and Wallis (1993): (i) a discordance measure for identifying unusual sites in a region, (ii) a heterogeneity measure, for assessing whether a proposed region is homogeneous, and (iii) a goodness of fit measure, for assessing whether a given distribution provides an adequate fit to the regional annual maximum flood flow data. Then, flood frequency distributions are selected from L-moment diagrams that compare observed to theoretical values of L-skewness and L-kurtosis for various candidate distributions. In the selection process, either the weighted sample average

or the line of best fit through the data points is used in the comparison with theoretical curves, depending on the outcome of the hetetrogeneity test, as was recommended by Peel et al. (2001).

5.1 Discordance and heterogeneity tests

First, data screening was performed using the discordance measure of Hosking and Wallis (1993) to eliminate gross errors and inconsistencies. For each flood series, sites with $D_i > 3$, believed to be grossly discordant with the group as a whole, were dropped. This resulted in decreasing the stations' number from 49 to 42 (Table 2). Homogeneity testing was performed then, through the heterogeneity measure H, which is based on the spread of the sample L-moments among the sites in a given region. Hosking and Wallis (1993) proposed a homogeneity test (H test) based on the fact that time series at all sites within a homogeneous region have the same population L-moment statistics (L-coefficient of variation, L-skewness and L-kurtosis). The assessment of heterogeneity is obtained by comparing L-moments of observed data and those of Monte Carlo simulation (Hosking and Wallis, 1993).

The weighted average distance (V) from the site to the network weighted mean on L-skewness and L-kurtosis diagram is used as a measure of the between-site variability of sample L-moments, and a homogeneity measure given by Hosking and Wallis (1993) is adopted:

$$H = \frac{V_{\text{obs}} - \mu_v}{\sigma_v} \tag{1}$$

Here, μ_v and σ_v are respectively the mean and standard deviation of the 500 values (large number) of their synthetic counterpart and V_{obs} is the real network value of V. This latter, which represents V_3 in Hosking and Wallis (1993), is expressed in terms of L-skewness and L-kurtosis.

Hosking and Wallis (1993) suggested, through Monte Carlo experiments, that the region under testing should be regarded as "acceptably" homogeneous if $H < 1$, "possibly" heterogeneous if $1 \leq H < 2$, and "definitely" heterogeneous if $H \geq 2$.

Homogeneity was investigated only with respect to skewness and kurtosis because these dimensionless statistical characteristics are commonly used to identify candidate regional flood frequency distributions. Homogeneity in the coefficient of variation was not considered because this statistic was shown to vary, among other things, with the size of the catchments and therefore constancy can not be achieved in any geographical region (Gupta et al., 1994).

6 Goodness-of-fit test

The goodness-of-fit test used compares the observed regional L-kurtosis to the theoretical values of various candidate

Table 2. Characteristics of the data used.

No.	Wadi	Station name	Basin drainage area (km^2)	Record Period (years)	Discordancy measure Di	mean discharge (m^3/s)	maximum discharge (m^3/s)
1	Tine	Cassis	418	1968–1978 1982–1998	0.37	125.7	478
2	Joumine	Dj. Antra	234	1975–1985 1982–1998	1	113.3	337
3	Mellila	H. Bourguiba	83	1989–1998	9.24*	48.9	331
4	Melah	Ouchtata	315	1975–1983 1988–1997	0.84	262.1	531
5	Joumine	Mateur	1096	1954–1998	1.21	133.4	239
6	Maden	Boubrima	145	1959–1989 1992–1998	2.56	99.8	295
7	Barbar	Joueouda	108.6	1982–1998	0.12	286.2	849
8	Zouaraa	S. Barrak	875	1979–1991	3.28*	390.5	1230
9	Rarai	Rarai supérieur	91,3	1975–1998	0.31	69.9	250
10	Rarai	Rarai plaine	370	1975–1998	1.1	178.9	422
11	Medjerda	Ghardimaou	1480	1949–1998	0.19	322.7	2370
12	Mellègue	K13	9014	1923–1998	1.71	768.6	4770
13	Rmel	Pt. Rte. Sakiet	407	1995–1977	0.11	96.6	429
14	Sarrath	Pt. Rte. Tejerouine	1520	1975–1998	0.28	302.9	945
15	Haidra	S. Abdelhak	328	1970–1998	1.4	126.3	564
16	Tessa	Sidi Mediene	1952	1975–1986 1993–1998	1.29	1057.8	3800
17	Souani	Pt. Rte. Souani	278	1985–1994	0.37	394.4	1350
18	Izid	Barrage	61	1986–1997	0.34	32.6	92
19	Ghezalah	Fernana	137	1946–1998	0.28	124.6	373
20	Medjerda	Jendouba	2410	1953–1998	0.54	343.4	2420
21	Siliana	Dj. Laouej	2066	1975–1998	4.41*	752.9	2060
22	El Bey	Pt. Rte. No. 41	464	1975–1998	0.29	182.2	917
23	Mkhachbia	Amont	18,5	1978–1979 1983–1995	0.34	8.6	34
24	Mkhachbia	Aval	106	1983–1998	0.73	83	600
25	Bazina	Bazina	18,5	1983–1995	3.83*	10.1	34.6
26	Medjerda	Slouguia	20 995	1975–1998	0.88	269.2	735
27	El Abid	Pt. Rte. No. 9	81	1975–1998	0.09	71.6	380
28	Miliane	Tuburbo Majus	748	1975–1981 1978–1998	0.39	179.3	1500
29	El Hamma	Aval	221	1967–1998	0.29	57	205
30	Bouarada	Bouarada	103	1967–1996	0.33	69.5	630
31	Miliane	Cheylus	1147	1975–1977 1977–1997	0.09	259.1	1800
32	Miliane	Madleine	1946	1975–1991 1997–1998	0.21	258.6	1300
33	Skhira	Kef Labiod	188	1969–1970 1977–1998	0.36	74.2	300
34	El Bey	S. Said Aval	31,7	1988–1998	0.2	8.3	37.1
35	El Baten	Bir Bouregba	70	1988–1998	2.01	124.4	586
36	Rmel	Si Abdallah	684	1945–1970	1.29	298	1500
37	El Oudiane	Oudiane	68	1958–1968	5.64*	71.67	149
38	Merguellil	Haffouz	675	1975–1998	4.87*	1197.6	2240
39	Hatob	Ain Saboun	813	1975–1998	0.59	279.0	1510
40	Hatab	Khanget Zazia	2200	1957–1998	0.76	412.5	1530
41	Negada	Bled Lassoued	5290	1975–1998	0.99	230.8	1860
42	Negada	Lassoud	16330	1975–1998	6.08*	246.8	548
43	Chaffar	Chaffar	236	1927–1998	0.16	176.8	1130
44	Sidi Salah	Sidi Salah	211	1986–1998	2.67	37.9	346
45	Zita	Telemen	3.2	1985–1998	0.33	16.8	103
46	Oglat Hamma	Hamma	735	1975–1996	1.25	50.6	460
47	Dijir	Dijir	146	1975–1998	0.16	75.1	381
48	Oum Zessar	Oum Zessar	278	1975–1980 1985–1998	0.77	146.5	1470
49	Sidi Aich	Sidi Aich	1780	1964–1966 1971–1995	0.45	357.9	2640

* Data series did not pass the discordancy measure.

Table 3. Properties of the stream flow data.

Region	Number of stations	Maximum record length (years)	Number of site years	Average record length (years)
Tunisia	42	76	1134	27
Northern Tunisia	14	50	361	25.8
Central Tunisia	21	76	561	26.7
Southern Tunisia	7	72	205	29.3

Table 4. Homogeneity and goodness-of-fit tests results.

Simulation experiment	Region	Average L-skewness (L-C$_s$)	Average L-kurtosis (L-C$_k$)	H-value	Z-value GLO	GEV	GNO	GPA	GAM
1	Tunisia	0.392	0.258	2.27	0.71	1.21[+]	−1.85[*+]	−4.65	−3.33
	Northern Tunisia	0.330	0.216	2.45	1.12	0.23	−2	−1.91	−1.76
2	Central Tunisia	0.382	0.249	1.82	0.58	−1.07	−1.13	−2.95	−2.21
	Southern Tunisia	0.487	0.331	0.27	−0.25	−1.46	−1.33	−2.81	−1.48
3	Northern Tunisia	0.328	0.209	1.73	1.21	0.32[+]	−1.51[*+]	−1.99	−1.92
	Central/Southern Tunisia	0.422	0.280	1.05	0.20	−1.42[+]	−1.83[*+]	−4.24	−2.70

* Distribution selected based on Lowess in L-moment diagram.
[+] Distribution selected based on weighted average in L-moment diagram.

distributions (Hosking and Wallis, 1993):

$$Z^{DIST} = \frac{\{\bar{t}_4 - \tau_4^{DIST}\}}{\sigma_{\bar{t}_4}} \qquad (2)$$

Where $\bar{t}_4$ is the regional average L-kurtosis of the observed network in the homogeneous region and τ_4^{DIST} is the theoretical L-kurtosis, and $\sigma_{\bar{t}_4}$ is the standard deviation of $\bar{t}_4$ obtained by repeated simulations of the homogeneous region with the DIST frequency distribution as a parent. Based on Monte Carlo simulation performed by Hosking and Wallis (1993), the goodness-of-fit of a particular distribution should be considered acceptable at the 90% confidence level if $|Z| \leq 1.64$. The Z-test uses regional data as opposed to single-site information. Therefore, it is more reliable than single-site goodness-of-fit testing. The Z-test discriminates between five of the most frequently used distributions in the analysis of hydrologic extreme variables, namely: (i) Generalized Extreme Value (GEV) Pearson type III (P3), (iii) Generalized Logistic (GLO), (iv) Generalized Normal (GN), and (v) Generalized Pareto (GPA) distributions.

6.1 Graphical goodness-of-fit from L-moment diagrams

An L-moment ratio diagram of L-kurtosis versus L-skewness compares sample estimates of the dimensionless ratios with their population counterparts for a range of statistical distributions. It has the advantage of comparing the fit of several statistical distributions with observed data using a single graphical instrument. L-moment diagrams are useful for discerning groupings of sites with similar flood frequency behavior, and identifying the statistical distribution likely to adequately describe this behavior. The distances separating sample points from the curve for a certain distribution can be taken as a measure of the goodness of fit. Peel et al. (2001) demonstrated that the graphical selection process of a distribution from the L-moment ratio diagram depends on the homogeneity of regional data. If the regional data are homogeneous, the selection should be based on comparison of theoretical curves with the weighted sample average. On the other hand, for very heterogeneous regional data the line of best fit through the data points known as LOWESS (LOcally WEighted Scatterplot Smoothing) should rather be considered. In this study, the delineated regions correspond to the three chosen areas described above, which cannot be claimed to be homogeneous. Therefore, similarity between theoretical distribution curves and LOWESS was adopted in the selection of the most suitable flood frequency distribution from the L-moment diagram for any particular region.

7 Results and discussions

The regional weighted average L-skewness and L-kurtosis were computed for the three considered regions, based on flood data series for only 42 stations, and the corresponding results are shown in Table 4. Computations of L-skewness and L-kurtosis values for the entire territory were also performed. The values obtained were very close to those of
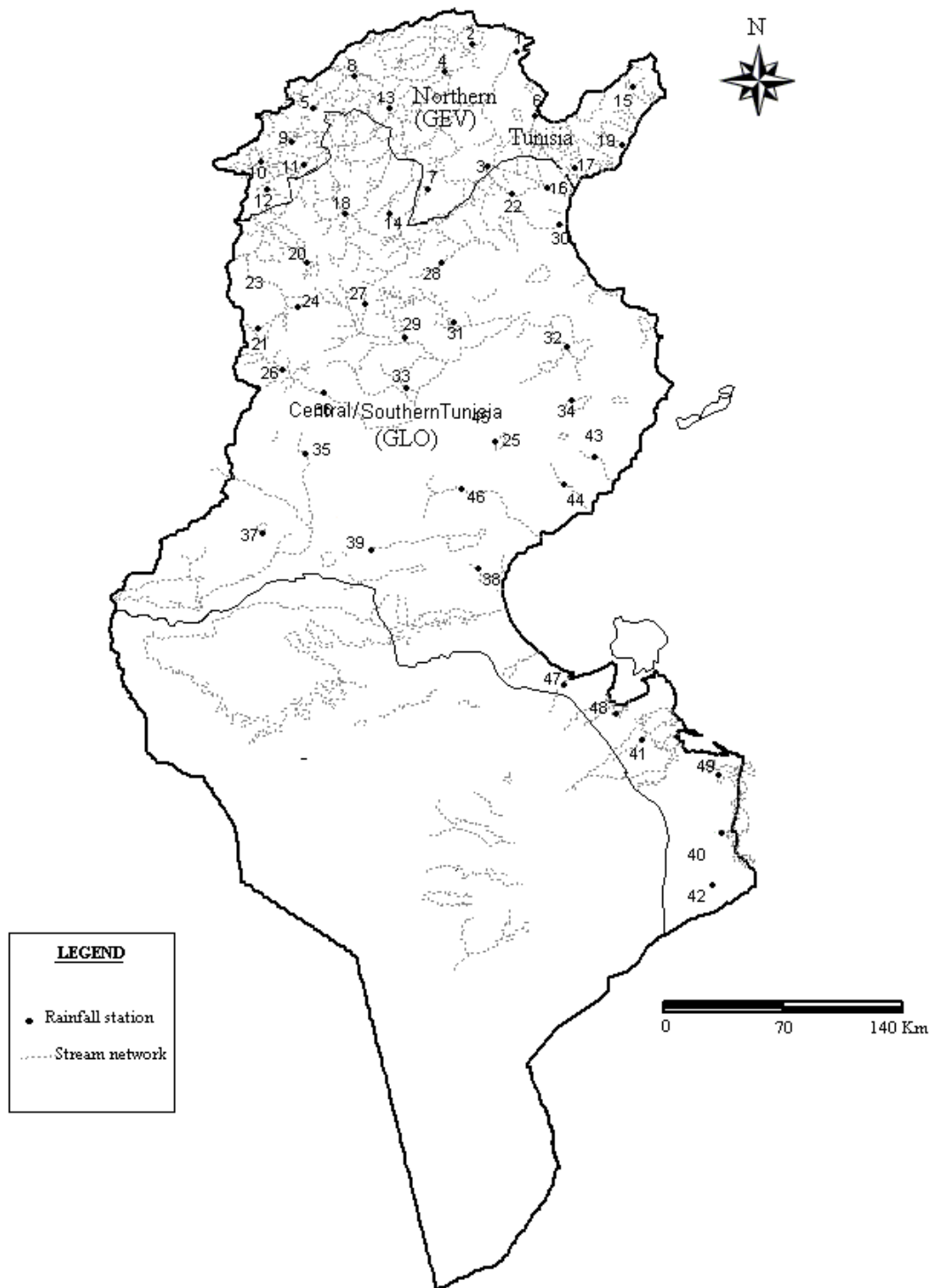
Fig. 2. The final classification of homogeneous regions adjusted for L-skewness and L-kurtosis.

northern Tunisia, since 75% of the stations were located in the north. The smallest L-moment values were obtained for the north while the highest were associated with the south. These differences are generally small in spite of the contrasting climatic and physiographic differences, which affect flood flows from a region to another in Tunisia. However, small differences in L-skewness and L-kurtosis values usually result in substantial differences in the tail characteristics of flood frequency distributions and therefore in different flood quantile estimates.
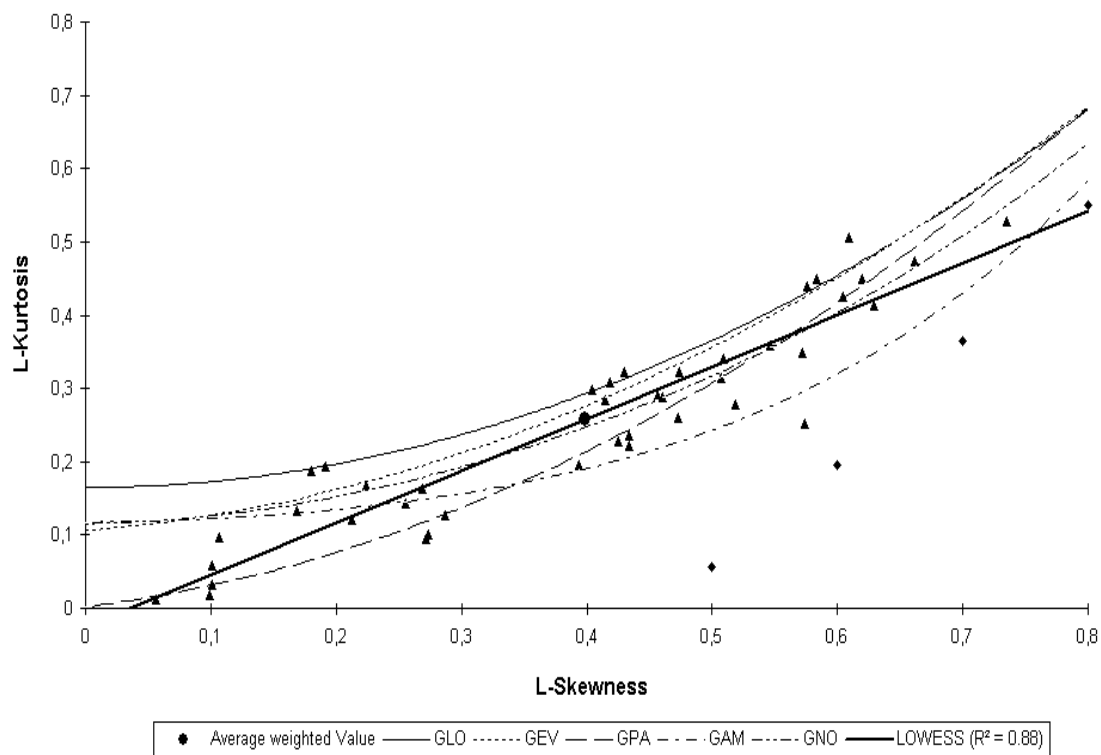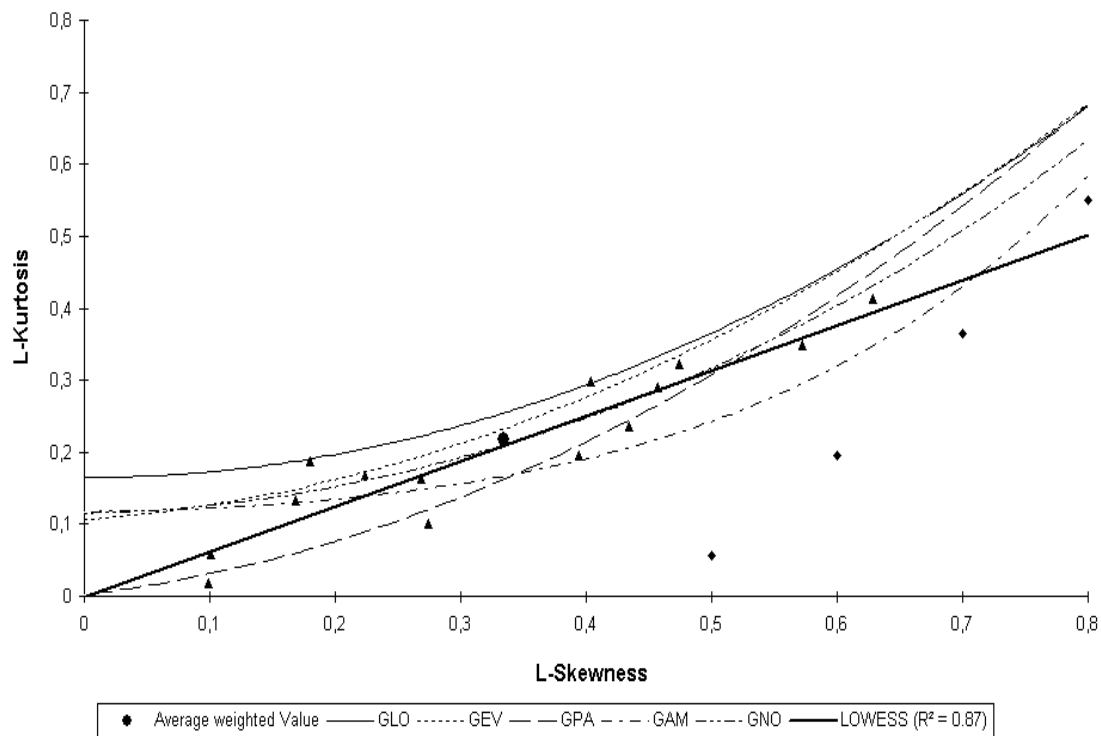
Fig. 3. L-moment ratio diagram for Tunisia.

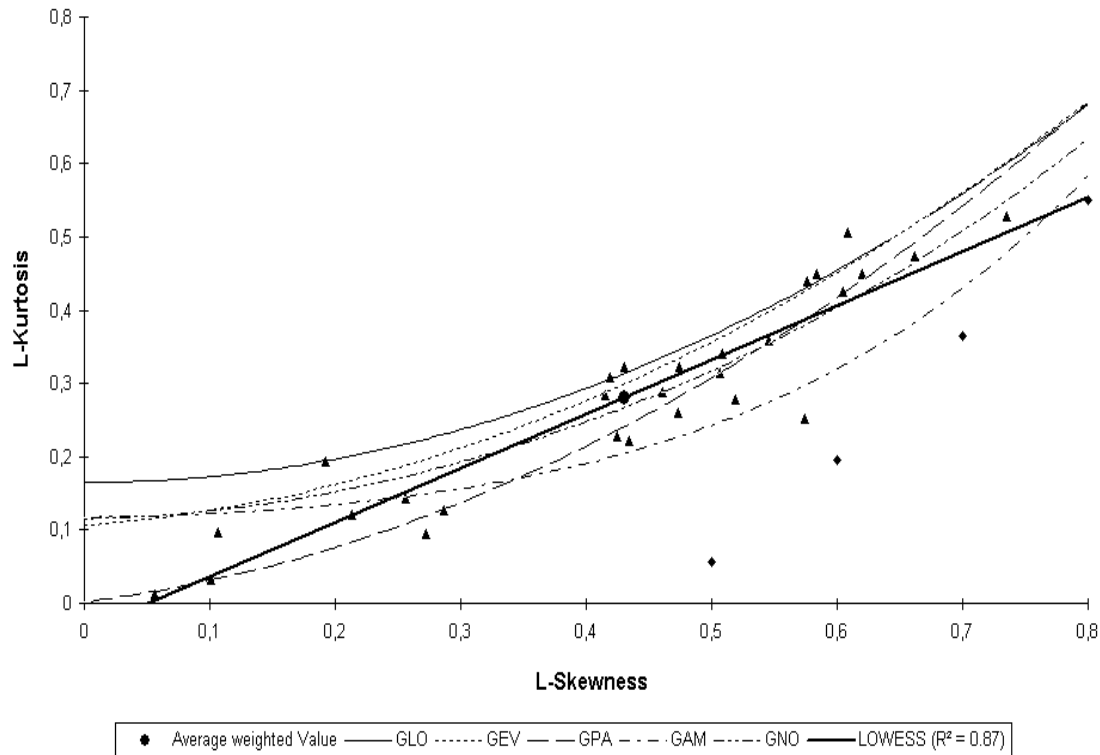**Fig. 4.** L-moment ratio diagram for northern Tunisia.

Fig. 5. L-moment ratio diagram for central/southern Tunisia.

Table 4 also presents H-values and recommended best-fit distributions for all regions considered, based on L-moment ratio diagrams on one hand, and the Z-test of Hosking and Wallis (1993) on the other. H-values, reflecting homogeneity increased from south to north. Only Southern Tunisia was found to be definitely homogeneous (Table 4, experiment 2).

However, based on stream flow data (Table 3 and Fig. 2), it can be shown that, compared to other regions, southern Tunisia has fewer gauging stations. Thus, central and southern Tunisia were treated as a single unit (Central/Southern Tunisia). This resulted in reducing the number of study regions from 3 to 2. Simulations were repeated for the two identified regions.

The obtained H values (Table 4) were 1.73 and 1.05 for northern and central/southern Tunisia, respectively, implying possibly homogeneous zones in terms of L-skewness and L-kurtosis. Heterogeneity was avoided by a trial and error procedure through regrouping stations and modifying the boundaries of the previously delineated regions. The outcome of this exercise that is the final delineation is presented in Fig. 2.

Figures 3 to 5 compare the observed relationships between L-kurtosis and L-skewness of annual maximum flood flows with the theoretical probability distributions: GLO, GEV, GPA, P III, and GNO Shown on the same figures are the locally weighted scatterplot smoothings (LOWESS) of L-skewness/L-kurtosis data, with the corresponding correla-

tion coefficients. The average Weighted L-Skewness and L-Kurtosis for the region was also represented in the diagram.

As was recommended by Peel (2001), selection of the appropriate probability distribution from the L-moment diagram was based on weighted average and lowess for homogeneous and heterogeneous zones respectively. Lowess was used for the first numerical experiment, dealing with Tunisia as a whole which was shown to be definitely heterogeneous. Comparing Lowess to the theoretical curves (Fig. 3), GNO distribution was found to give the best fit.

Considering the third experiment, the two regions were found to be possibly homogeneous. Therefore, the distribution identification was based on both Lowess and weighted average. Figure 4, which displays results for Northern Tunisia, shows that GNO distribution gives the best agreement with Lowess, while two potential distributions (GNO and GEV) can be selected based on the weighted average. The same conclusions were found to be valid for Central/Southern Tunisia (Fig. 5). Results of the Z-test (Table 4) confirmed the choice of GNO and GEV, especially for Northern Tunisia. The final outcome of both L-moment diagram and Z statistical test is therefore the GNO distribution for Northern Tunisia and both GEV and GNO distributions for Central/Southern Tunisia.

Flood quantiles for stations 11, 19 (Northern Tunisia), and 20, and 43 (Central/Southern Tunisia) were simulated based on both selected distributions (GNO and GEV) and

Table 5. Regional and single-site flood estimates for four selected stations.

Station	Return period	Observed	Regional flood (GEV) Q (m3/s)	Ei (%)	Regional flood (GNO) Q (m3/s)	Ei (%)	Single site flood Q (m3/s)	Ei (%)
11	5	455.0	448.7	−1.4	455.6	0.1	470.0	3.3
	10	635.0	636.1	0.2	640.9	0.9	647.6	2.0
	20	699.0	819.1	17.2	854.1	22.2	890.8	27.4
	50	2110.0	1304.8	−38.2	1455.7	−31.0	1095.8	−48.1
19	5	181.0	183.7	1.5	198.9	9.9	141.4	−21.9
	10	221.0	187.0	−15.4	165.4	−25.1	252.4	14.2
	20	262.0	263.1	0.4	279.7	6.7	232.7	−11.2
	50	306.0	308.7	0.9	320.1	4.6	294.5	−3.7
20	5	484.0	511.3	5.6	534.3	10.4	437.1	−9.7
	10	669.0	652.8	−2.4	643.2	−3.9	749.1	12.0
	20	826.0	916.1	10.9	1021.4	23.7	1058.5	28.1
	50	2300.0	1631.6	−29.1	1476,5	−35.8	1370.1	−40.4
43	5	271.0	259.0	−4.4	234.7	−13.4	190.8	−29.6
	10	392.0	295.7	−24.6	274.9	−29.9	402.4	2.6
	20	661.0	583.0	−11.8	435.3	−34.1	429.8	−35.0
	50	1090.0	905.6	−16.9	877.5	−19.5	672.9	−38.3

the corresponding results are displayed in Table 5. Shown on the same table are observed and single site floods. The non- parametric method was used to compute the observed quantiles in order to ovoid any commitment to a particular parent distribution for frequency analysis (Pilon and Harvey, 1994). Table 5 also shows relative errors between observed and simulated floods given as:

$$E_i(\%) = \frac{(Q_T)_{\text{observed}} - (Q_T)_{\text{simulated}}}{(Q_T)_{\text{observed}}} \times 100 \qquad (3)$$

Compared to single-site floods, those obtained by the regional approach, were shown to be in much better agreement with observed data. Relative errors were found to be 12%, 15% and 21% using the GEV, GNO models and the single-site method respectively. Relative errors were found to be relatively high for return periods exceeding 20 years, especially for stations 11 and 20. This may be due to their relatively short record lengths (less than 50 years).

8 Conclusions

Flood frequency analysis procedure was adopted to identify appropriate distributions for delineated zones within Tunisia. Hosking and Wallis (1993) statistical measures were used to eliminate grossly discordant sites from the analysis, determine the extent of heterogeneity for a given study region, and test the goodness-of-fit of a particular flood frequency distribution to observed data. Distributions were selected from L-moment diagrams, based on a comparison between the line of best fit through the data points and theoretical distribution curves.

Unlike similar previous work, such as the studies by Vogel et al. (1993b) and Vogel and Wilson (1996), in which flood frequency distributions were determined for all of Australia and the United States respectively, Tunisia, in this study, was divided into two sub-regions, for which separate flood frequency analysis procedures were applied.

H-values for the study regions, describing homogeneity, generally increase from south to north. Northern Tunisia was found to be represented by GNO distribution, while both GNO and GEV distributions were shown to be most suitable for the center and the south of the country.

Finally, compared to single-site analysis, the proposed regional flood frequency model was shown to yield better flood estimates.

Edited by: M. Mikos

References

Abdul, K. M. and Chowdhury, J. U.: A comparison of four distributions used in flood frequency analysis in Bangladesh, Hydro. Sci., 40(2), 55–66, 1995.

Alila, Y. and Mtiraoui, A.: Implications of heterogeneous flood-frequency distributions on traditional stream-discharge prediction techniques, Hydrol. Processes, 16, 1065–1084, 2002.

Beard, L. R.: Flood flow frequency techniques, Center for research in water Resources, The University of Texas, Austin, 1974.

Benson, M. A.: Uniform flood frequency estimation methods for federal agencies, Water Resour. Res., 4(5), 891–908, 1968.

Bobee, B. and Rasmussen, P. F.: Reviews of Geophysics, US National Report, twenty-first General assembly, International Union of Geodesy and Geophysiscs, Boulder, Colorado, 1111–1116, 1995.

Cathcart, J. G.: The effects of scale and storm severity on the linearity of watershed response revealed through the regional L-moment analysis of annual peak flows, Ph.D thesis, Institute of Resources and Environment, University of British Columbia, Canada, 2001.

Cong, S., Li, Y., Vogel, J. L., and Schaake, J. C.: Identification of the underlying distribution form of precipitation by using regional data, Water Resour. Res., 29(4), 1103–1111, 1993.

Cunnane, C.: Statistical distributions for flood frequency analysis, Operational hydro., World Meteorological Org. (WMO), Geneva, Switzerland, Rep. 33, 1989.

Farquharson, F. A. K., Green, C. S., Meigh, J. R., and Sutcliffe, J. V.: Comparison of flood frequency curves for many different regions in the world, Regional flood frequency analysis, edited by: Singh, V. P., D. Reidel Publishing Co., Dordrecht, Holland, 223–256, 1987.

Finlayson, B. L. and McMahon, T. A.: Global runoff Encyclopedia of earth system science, Academic Press Inc., San Diego, Calif, vol. 2, 1992.

Gingras, D. and Adamowski, K.: Coupling of non-parametric frequency and L-moment analysis for mixed distribution identification, Water Resour. Bull., 28(2), 263–272, 1992.

Gupta, V. K., Mesa, O. J., and Dawdy, D. R.: Multiscaling theory of flood peaks: Regional quantile analysis, Water Resour. Res., 30(12), 3405–3421, 1994.

Gunasekara, T. A. G. and Cunnane, C.: Split sampling technique for selecting a flood frequency analysis procedure, J. Hydrol., 130, 189–200, 1992.

Hosking, J. R. M.: L-moments: analysis and estimation of distributions using linear combinations of order statistics, J. Royal Statis. Soc., Series B, 52, 105–124, 1990.

Hosking, J. R. M. and Wallis, J. R.: Regional flood-frequency analysis using L-moments, IBM Research, Yorktown Heights, New York, Rep. 15658, 1990.

Hosking, J. R. M. and Wallis, J. R.: Some statistics useful in regional frequency analysis, Water Resour. Res., 29(2), 271–281, 1993.

IEA (Institution of Engineers, Australia): Australia rainfall and runoff: Flood analysis and design, IEA, Canberra, 1977.

Jingyi, Z. and Hall, M. J.: Regional flood frequency analysis for the Gan-Ming River basin in China, J. Hydrol., 296(1–4), 98–117, 2004.

Klemes, V.: Tall tables about tails of hydrological distributions I, J. Hydrol. Eng., 5(3), 227–231, 2000a.

Klemes, V.: Tall tables about tails of hydrological distributions II, J. Hydrol. Eng., 5(3), 232–239, 2000b.

Kritsky, S. N. and Menkel, F. M.: On principles of estimation methods of maximum discharge, IAHS Publ., 84, 1969.

McMahon, T. A. and Srikanthan, R.: Log Pearson III distribution – is it applicable to flood frequency analysis of Australian streams?, J. Hydrol., 52, 139–147, 1981.

McMahon, T. A., Finlayson, B. L., Haines, A. T., and Srikanthan, R.: Global runoff – continental comparisons of annual flows and peak discharges, Catena Verlag, Cremlingen-Destedt, Germany, 1992.

Nathan, R. J. and Weinmann, P. E.: Application of at-site and regional flood frequency analysis, Proc. Int. Hydrology and Water Resources Symposium, 769–774, 1991.

NERC: Natural Environment Research Council, Flood studies rep. I, hydro. studies, London, England, 1975.

Onoz, B. and Bayazit, M.: Best fit distributions of largest available flood samples, J. Hydrol., Amsterdam, The Netherlands, 167, 195–208, 1995.

Pearson, C. P.: New Zealand regional flood frequency analysis using L-moments, J. Hydrol., 30(2), 53–63, 1991.

Peel, M. C., Wang, Q. J., Vogel, R. M., and McMahon, T. A.: The utility of L-moment ratio diagrams for selecting a regional probability distribution, Hydrol. Sci., 46(1), 147–155, 2001.

Pilon, P. J. and Adamowski, K.: The value of regional information to flood frequency analysis using the method of L-moments, Can. J. Civil Eng., Ottawa, Canada, 19, 137–147, 1992.

Pilon, P. J. and Harvey, K. D.: Consolidated frequency analysis, Reference manual, Environment Canada, Ottawa, Canada, 1994.

Rakesh, K. and Chandranath, C.: Regional Flood Frequency Analysis Using L-Moments for North Brahmaputra Region of India, J. Hydrol. Eng., 11(4), 380–382, 2006.

Rao, A. R. and Hamed, K. H.: Regional frequency analysis of Wabash River flood data by L-moments, J. Hydrol. Eng., 2(4), 169–179, 1997.

Thanh, V. and Nguyen, V.: On Regional Estimation of Floods for Engaged Sites, Asia Oceania Geosciences Society, McGill University, Singapore, 2006.

Vogel, R. M. and Fennessey, N. M.: L-moment diagrams should replace product moment diagrams, Water Resour. Res., 29(6), 1745–1752, 1993.

Vogel, R. M., Thomas Jr., W. O., and McMahon, T. A.: Flood flow frequency model selection in Southwestern United States, J. W. Res. Plan. M., ASCE, 119(3), 353–366, 1993a.

Vogel, R. M., McMahon, T. A., and Chiew, F. H. S.: Flood flow frequency model selection in Australia, J. Hydrol., Amsterdam, The Netherlands, 146, 421–449, 1993b.

Vogel, R. M. and Wilson, I.: Probability distribution of annual maximum, mean, and minimum streamflows in the United States, J. Hydrol. Eng., 69–76, 1996.

Wallis, J. R.: Catastrophes, computing and containment: living with our restless habitat, Speculation in Science and Technology, 11(4), 295–324, 1988.