Water Resources Management **8:** 183–201, 1994.

Frequency Analysis of Upper Cauvery Flood Data by L-Moments

A. RAMACHANDRA RAO and KHALED H. HAMED
School of Civil Engineering, Purdue University, West Lafayette, IN 47907, U.S.A.

(Received: 12 November 1993)

Abstract. The objectives of the present study are to investigate the hydrological homogeneity of Upper Cauvery annual maximum flow data and to select a suitable distribution for the frequency analysis. The L-moments method is used in this analysis. The Upper Cauvery river basin is shown to be hydrologically heterogeneous. The 3 parameter log normal and the generalized extreme value distributions are recommended for the frequency analysis of data in this region.

Key words: hydrological homogeneity, flood data, L-moments methods

1. Introduction

The concept of probability weighted moments (PWM) was introduced by Greenwood *et al.* (1979). Since then it has received considerable attention from Landwehr *et al.* (1979a, b), Hosking *et al.* (1985), Hosking (1986), Hosking and Wallis (1987) and others. PWM estimates are robust in the presence of outliers. Parameter estimates from small samples computed by using the PWM method are sometimes more accurate than even the maximum likelihood (ML) estimates. The PWM method is less complicated than the ML method. With some distributions, such as the symmetrical Lambda and Weibull distributions, explicit expressions for the parameters are obtained by the PWM method, which cannot be done with either the ML or the method of moments (MOM).

Hosking (1986, 1990) has defined the L-moments which are analogous to the conventional moments and are estimated by linear combinations of order statistics. They can also be expressed by linear combinations of PWM. Thus, procedures based on PWM and L-moments are equivalent. However, L-moments are more convenient because they are directly interpretable as measures of the scale and of the shape of probability distributions. Hosking (1990) has used L-moment ratio diagrams to identify underlying parent distributions and L-moment ratios for testing hypotheses about forms of probability distributions. Hosking and Wallis (1991) extended the use of L-moments and developed statistics that can be used in regional frequency analysis to measure discordancy, regional homogeneity and goodness-of-fit.

The objective of the present work is to analyze the annual maximum flow data from the Cauvery River basin in south India by using the L-moment method. Both regional and at-site parameter and quantile estimates are used and the differences between them are studied. The L-moment ratios are used to identify candidate distributions and to evaluate the effectiveness of regional analysis.

2. Theoretical Background

Probability weighted moments $M_{1,r,s}$ are defined by Greenwood *et al.* (1979) as in Equations (1) and (2), where F is the cumulative probability.

$$M_{1,r,0} = \beta_r = \int_0^1 x(F)F^r \mathrm{d}F \,, \tag{1}$$

$$M_{1,o,s} = \alpha_s = \int_0^1 x(F)(1-F)^s \mathrm{d}F \,. \tag{2}$$

Both α_s and β_r are linear in x and are of sufficient generality for parameter estimation. Also, α_s and β_r are related as in Equations (3) and (4).

$$\alpha_s = \sum_{k=0}^{s} \binom{s}{k} (-1)^k \beta_k \,, \tag{3}$$

$$\beta_r = \sum_{k=0}^{r} \binom{r}{k} (-1)^k \alpha_k \,. \tag{4}$$

L-moments λ_{r+1} are defined by Hosking (1986, 1990) in terms of PWMs α_s and β_r as in Equation (5),

$$\lambda_{r+1} = (-1)^r \sum_{k=0}^{r} p_{r,k}\alpha_k = \sum_{k=0}^{r} p_{r,k}\beta_k, \tag{5}$$

where

$$p_{r,k} = (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \,. \tag{6}$$

In particular,

$$\lambda_1 = \alpha_0 = \beta_0 \,, \tag{7}$$

$$\lambda_2 = \alpha_0 - 2\alpha_1 = 2\beta_1 - \beta_0 \,, \tag{8}$$

$$\lambda_3 = \alpha_0 - 6\alpha_1 + 6\alpha_2 = 6\beta_2 - 6\beta_1 + \beta_0 \,, \tag{9}$$

$$\lambda_4 = \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 = 20\beta_3 - 30\beta_2 - 12\beta_1 + \beta_0 \,. \tag{10}$$

For a given ordered sample $x_1 \leq \cdots \leq x_n, n > r$ and $n > s$, the sample PWMs are calculated (Hosking, 1986) by Equations (11) and (12),

$$a_s = \hat{\alpha}_s = \frac{1}{n} \sum_{i=1}^{n} (1 - P_{i:n})^s x_i \,, \tag{11}$$

$$b_r = \hat{\beta}_r = \frac{1}{n}\sum_{i=1}^{n} P_{i:n}^{r} x_i \ , \tag{12}$$

where $P_{i:n}$ is a plotting position. The use of $P_{i:n} = (i - 0.35)/n$ usually gives good results for the generalized extreme value distribution (GEV) (Hosking *et al.*, 1985) and is recommended in general for analysis of hydrologic data (Cunnane, 1989). Sample L-moments (l_r) can be calculated by using Equations (7)–(10) by replacing α_s or β_r by their sample estimates a_r and b_r. L-moment ratios, which are analogous to the conventional moment ratios are defined by Hosking (1986, 1990) in Equations (13) and (14).

$$\tau = \lambda_2/\lambda_1 \ , \tag{13}$$

$$\tau_r = \lambda_r/\lambda_2 \ , \quad r \geq 3 \ , \tag{14}$$

where λ_1 is a measure of location, τ is a measure of scale and dispersion (LC_v), τ_3 is a measure of skewness (LC_s), τ_4 is a measure of kurtosis (LC_k). Sample L-moment ratios *t and* t_r are calculated by using Equations (13) and (14) and substituting l_r for their population values λ_r. The L-moment ratios (Hosking, 1990) offer an easy way to identify underlying distributions, particularly the skewed distributions. The sample L-moment ratios plot as well separated groups for different distributions. Therefore, different distributions are easily discriminated by using them. A distribution is considered to be suitable if the data spread consistently around it. Hosking (1990) also suggests a test for normality against the skew alternative based on the statistic

$$N_n = v_n^{-1/2} t_3, \tag{15}$$

where v_n is the variance of t_3 and is given by $0.1866n^{-1} + 0.8\,n^{-2}$.

The critical limits of N_n are obtained from the standard Normal tables at the required significance level. Hosking and Wallis (1991) have derived statistics to measure discordancy (D), regional heterogeneity (H) and goodness of fit (Z). A full description of these statistics is found in Hosking and Wallis (1991). A site i is considered to be unusual if D_i is large. A suitable criterion for defining largeness is that D_i should be greater than 3. A region is declared heterogeneous if H is sufficiently large. Hosking and Wallis suggest the region be regarded as *acceptably homogeneous* if H is less than 1, *possibly heterogeneous* if H is between 1 and 2, and *definitely heterogeneous* if H is greater than 2. Also, a given distribution is declared adequate if Z^{DIST} is sufficiently close to zero. An acceptable criterion being that $| Z^{DIST} |$ is less than or equal to 1.64.

As in the method of moments (MOM), parameter estimates are obtained by equating sample PWM or L-moments to the corresponding population values. Depending on data availability, parameter and quantile estimates are obtained by using either data at a site or regional data or both (Cunnane, 1989). Flood estimates

TABLE I. Details of Cauvery river and tributaries (source Ramesh *et al.*, 1987)

Station	Name of the station	Stream	Drainage area (km^2)	Annual rainfall (mm)	Data available	Q_{max} ($m^3 s^{-1}$)
1	Chunchanakatte	Cauvery	2968	680	1918–1980	2879
2	Akkihebbal	Hemavathy	5198	742	1918–1980	2979
3	Unduwadi	Lakshmanathirtha	1502	763	1918–1980	1825
4	Nugu Dam	Nugu	984	920	1918–1980	877
5	Hullahalli	Kabini	4850	920	1918–1974	4129
6	Markonahalli Dam	Shimsha	4131	764	1918–1980	1952
7	Mangala Dam	Nagini	748	764	1918–1980	707
8	Kanva Dam	Kanva	344	839	1918–1980	998
9	Suvarnavathi Dam	Suvarnavathi	1437	676	1918–1980	2084

may be based on data at a site if the record is exceptionally long, or when regional data are not available, or when a region is very heterogeneous. The advantage of joint use of at-site and regional data is that, in general, there is sufficient information in the combined set of data so that a multi-parameter distribution can be reliably used. A method of combining regional data, which is used here, is the index flood method (Cunnane, 1989). The variate Q normalized by its mean is assumed to have the same distribution at each site. The quantile Q_T at a site is estimated by Equation (16).

$$\hat{Q}_T = \mu_i q_T , \tag{16}$$

where q_T is the quantile estimate from the regional distribution and μ_i is the mean at the site. The regional distribution parameters are obtained by using regional weighted averages of dimensionless L-moments computed by dividing the moments by the mean μ_i of each station.

3. Analysis of the Cauvery River Data

The annual maximum flow data from nine gauging stations on the tributaries to the Cauvery River in the state of Karnataka, India are used in this study. The locations of the gauging stations are shown in Figure 1. Details of the rivers and stations included in this study are given in Table I.

The record length of each station is 63 years except for station 5 which has a record length of 57 years. Probability weighted moments as well as L-moments are calculated for data from each of the nine stations by using equations given in Section 2. Table II gives the standardized moments, which are the original moments computed by using the data divided by the mean at each station, as well as their weighted regional averages for these stations. The values of L-moment ratios LC_v, LC_s and LC_k are given in Table III, as well as their weighted regional averages.

TABLE II. Standardized probability weighted moments (PWMs) and L-moments

Station No.	b_0	b_1	b_2	b_3	b_4	a_0	a_1	a_2	a_3	a_4	l_1	l_2	l_3	l_4
1	1.0000	0.5952	0.4356	0.3477	0.2913	1.0000	0.4048	0.2452	0.1734	0.1330	1.0000	0.1904	0.0422	0.0296
2	1.0000	0.6144	0.4576	0.3705	0.3141	1.0000	0.3856	0.2288	0.1592	0.1204	1.0000	0.2288	0.0595	0.0534
3	1.0000	0.6926	0.5520	0.4681	0.4112	1.0000	0.3074	0.1668	0.1102	0.0806	1.0000	0.3852	0.1566	0.1122
4	1.0000	0.7343	0.6044	0.5217	0.4630	1.0000	0.2657	0.1358	0.0886	0.0654	1.0000	0.4686	0.2206	0.1133
5	1.0000	0.6227	0.4711	0.3869	0.3322	1.0000	0.3773	0.2258	0.1585	0.1208	1.0000	0.2454	0.0907	0.0757
6	1.0000	0.8042	0.6793	0.5911	0.5249	1.0000	0.1958	0.0709	0.0342	0.0195	1.0000	0.6084	0.2507	0.0935
7	1.0000	0.7397	0.6078	0.5238	0.4645	1.0000	0.2603	0.1283	0.0803	0.0569	1.0000	0.4794	0.2083	0.1205
8	1.0000	0.7414	0.5978	0.5055	0.4408	1.0000	0.2586	0.1150	0.0636	0.0398	1.0000	0.4828	0.1382	0.0736
9	1.0000	0.7065	0.5628	0.4757	0.4165	1.0000	0.2935	0.1498	0.0932	0.0643	1.0000	0.4129	0.1378	0.1093
Average	1.0000	0.6953	0.5529	0.4665	0.4073	1.0000	0.3047	0.1623	0.1062	0.0774	1.0000	0.3906	0.1455	0.0869

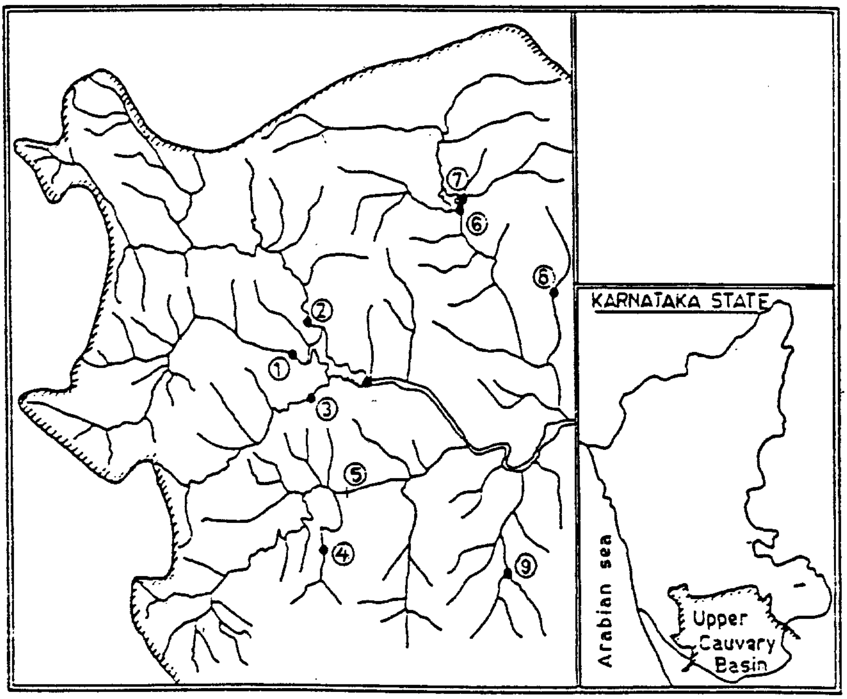

Fig. 1. The Upper Cauvery Basin (Source: Ramesh *et al.* 1987).

TABLE III. L-moment ratios and discordancy measures D_i

Site	N	Name	LC_v	LC_s	LC_k	D_i
1	63	Site 1	0.1904	0.2217	0.1556	1.93
2	63	Site 2	0.2288	0.2599	0.2332	0.62
3	63	Site 3	0.3852	0.4065	0.2913	0.44
4	63	Site 4	0.4686	0.4707	0.2417	1.18
5	57	Site 5	0.2454	0.3696	0.3085	0.89
6	63	Site 6	0.6084	0.4121	0.1537	1.25
7	63	Site 7	0.4794	0.4345	0.2513	0.33
8	63	Site 8	0.4828	0.2862	0.1524	1.17
9	63	Site 9	0.4129	0.3338	0.2646	1.19
Weighted means			0.3906	0.3548	0.2272	

The discordancy measures D_i for each of the nine sites are given in Table III. The largest value of D_i is 1.93 for site 1, which is less than 3 recommended by Hosking and Wallis (1991). Consequently, none of the nine sites may be considered to be unusual. The LC_v – LC_s moment ratio diagram for different rivers in shown in Figure 2.

The LC_s versus LC_k diagram for the data used in this study as well as for some of the common three parameter distributions is shown in Figure 3. The heterogeneity

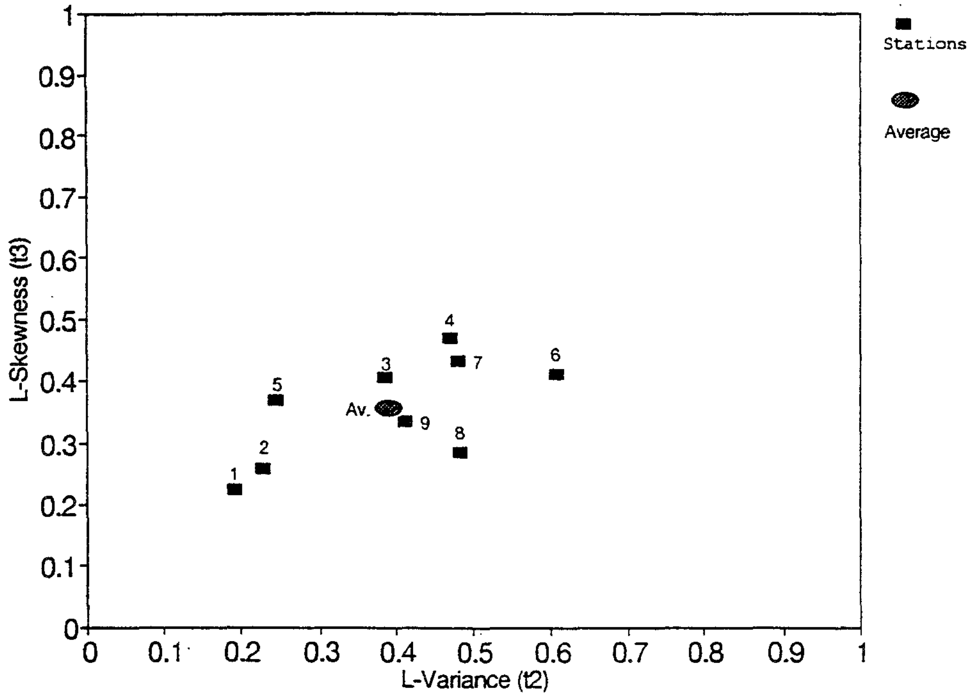

Fig. 2. LC_v vs LC_s moment ratio diagram.

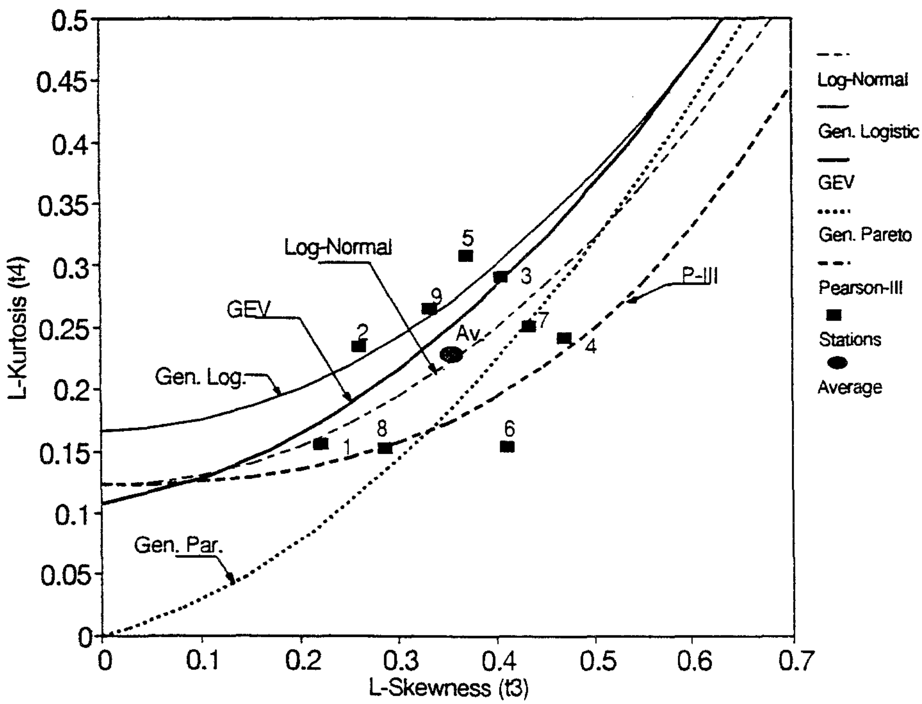

Fig. 3. The LC_s vs LC_k moment ratio diagram.

TABLE IV. Heterogeneity measures H (Number of simulations = 500)

Measure	Item	Value
H_1	Observed S.D. of group LC_v	0.1323
	Sim. mean of S.D. of group LC_v	0.0357
	Sim. S.D. of S.D. of group LC_v	0.0095
	Standardized test value	10.21
H_2	Observed ave. of LC_v/LC_s	0.1394
	Sim. mean of ave. LC_v/LC_s	0.0689
	Sim. S.D. of ave. LC_v/LC_s	0.0162
	Standardized test value	4.37
H_3	Observed ave. of LC_s/LC_s	0.0944
	Sim. mean of ave. LC_s/LC_k	0.0866
	Sim. S.D. of ave. LC_s/LC_k	0.0206
	Standardized test value	0.38

TABLE V. Goodness of fit measures Z^{DIST}

Distribution	LC_k	Z
Gen. logistic	0.272	1.34
Gen. extreme value	0.250	0.53
Log-normal	0.222	−0.53
Pearson Type III	0.174	−2.34
Gen. pareto	0.184	−1.97

measures in Table IV. The standardized test value H_1 is 10.21 which is much higher than 2, which suggests that the region is definitely heterogeneous. Consequently the upper Cauvery river basin is heterogeneous, a conclusion supported by the information in Figures 2 and 3. The data in Figure 3 do not cluster around any distribution, but are scattered around all of them. The average value of the statistic H_2 (4.37) is again larger than 2. H_2 is shown later to represent the relationship between the at-site and regional estimates. Consequently, the relationship between the regional and at-site estimates at different sites is quite diverse. The statistic H_3 (0.38), which is shown later to represent the relationship between observed and fitted data indicates that, in general, there would be good agreement between at-site estimates and observed data.

The test of normality outlined in section 2 yields ν_n = 0.00316. The critical values of N_n at 5% significance level are $\pm$ 1.96. The normality assumption can be accepted if $| t_3 | < 0.11$. Consequently, the normality hypothesis is rejected for all nine stations. Only skew distributions can be considered for these stations. In Table V the values of the goodness of fit measure Z^{DIST} for different distributions

TABLE VI. Parameter estimates for the GEV distribution

Site	1	2	3	4	5
u	0.8319	0.7906	0.6117	0.5154	0.7573
α	0.2539	0.2863	0.3615	0.3783	0.2494
K	–0.0794	–0.1359	–0.3386	–0.4209	–0.2897

Site	6	7	8	9	Regional
u	0.3852	0.5108	0.5483	0.6003	0.6129
α	0.5639	0.4222	0.5768	0.4511	0.3945
K	–0.3460	–0.3750	–0.1740	–0.2408	–0.2937

(Hosking, 1991) are given. Candidate distributions for which the $|Z^{DIST}|$ is less than 1.96 are the Generalized Logistic distribution (GLOG), the Generalized Extreme Value distribution (GEV) and the Log-normal (LN) distribution. Consequently, they are used in further analysis. The Pearson-III (P-III) distribution is included for comparison purposes although it does not provide a good fit.

3.1. GEV DISTRIBUTION

Parameter estimates of the GEV distribution are obtained from Hosking (1986 and 1990) and given in Equations (17)–(20).

$$\hat{K} = 7.8590C + 2.9554C^2, \tag{17}$$

$$\hat{\alpha} = \frac{l_2\hat{K}}{(1-2^{-\hat{K}})\Gamma(1+\hat{K})}, \tag{18}$$

$$\hat{u} = l_1 + \frac{\hat{\alpha}}{\hat{K}}[\Gamma(1+\hat{K}) - 1], \tag{19}$$

$$C = \frac{2}{3+t_3} - \frac{\log 2}{\log 3}. \tag{20}$$

Parameter estimates at each station as well as their regional averages are given in Table VI. Quantile estimates are obtained using Equation (21).

$$\hat{q}_T = \hat{u} + (\hat{\alpha}/\hat{K})[1 - (-\log F)^K]. \tag{21}$$

Figure 4 (a–d) shows at-site as well as regional quantile estimates along with the corresponding observed values for rivers Cauveri, Lakshmanathirtha, Kabini and Shimsha.

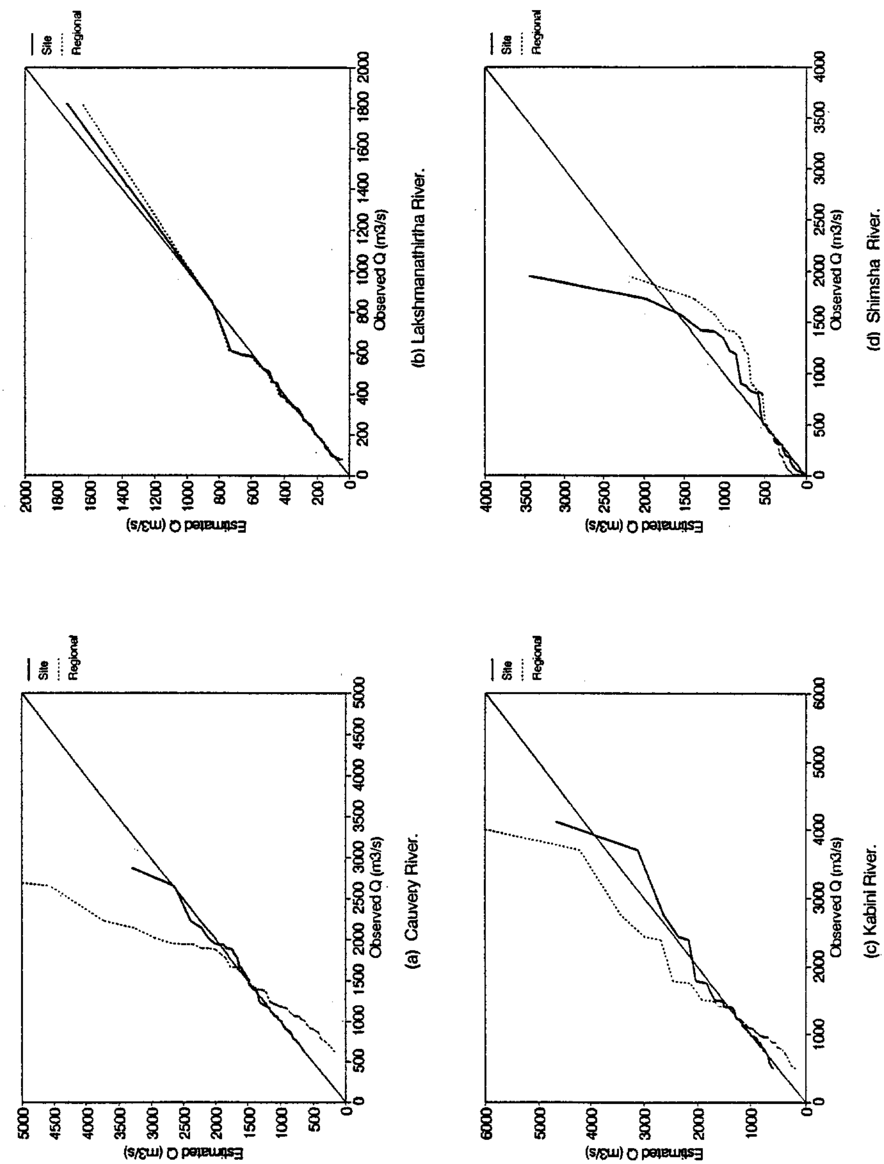

Fig. 4. The estimates and observed data for the GEV distribution.

TABLE VII. Parameter estimates for the GLOG distribution

Site	1	2	3	4	5
u	0.9322	0.9504	0.7627	0.6748	0.8606
α	0.1754	0.2042	0.2887	0.3156	0.1939
K	–0.2217	–0.2599	–0.4065	–0.4707	–0.3696
Site	**6**	**7**	**8**	**9**	**Regional**
u	0.6208	0.6879	0.7817	0.7854	0.7857
α	0.4521	0.3438	0.4203	0.3413	0.3146
K	–0.4121	–0.4345	–0.2862	–0.3338	–0.3548

3.2. GENERALIZED LOGISTIC (GLOG) DISTRIBUTION

Parameter estimates of the GLOG distribution are obtained from Equations (22)–(25) (Hosking, 1986 and 1990).

$$\hat{K} = -t_3, \tag{22}$$

$$\hat{\alpha} = \frac{l_2}{\Gamma(1+\hat{K})\Gamma(1-\hat{K})}, \tag{23}$$

$$\hat{u} = l_1 + (l_2 - \hat{\alpha})/\hat{K}. \tag{24}$$

Parameter estimates for each station as well as their regional averages are given in Table VII. Quantile estimates are computed by using Equation (25).

$$q_T = \hat{u} + \frac{\hat{\alpha}}{\hat{K}}[1 - \{(1-F)/F\}^{\hat{K}}] \tag{25}$$

The at site as well as regional quantile estimates along with the observed data for Cauveri, Lakshmanathirtha, Kabini, and Shimsha rivers are shown in Figure 5.

3.3. LOG-NORMAL DISTRIBUTION

Parameter estimates of the LN distribution are obtained from Equations (26)–(29) (Hosking, 1990),

$$\hat{\sigma} = 0.999281Z - 0.006118Z^3 + 0.000127Z^5\,, \tag{26}$$

$$\hat{\mu} = \log\,[l_2\,/\,\mathrm{erf}\,(\hat{\sigma}\,/\,2)] - \frac{\hat{\sigma}^2}{2}\,, \tag{27}$$

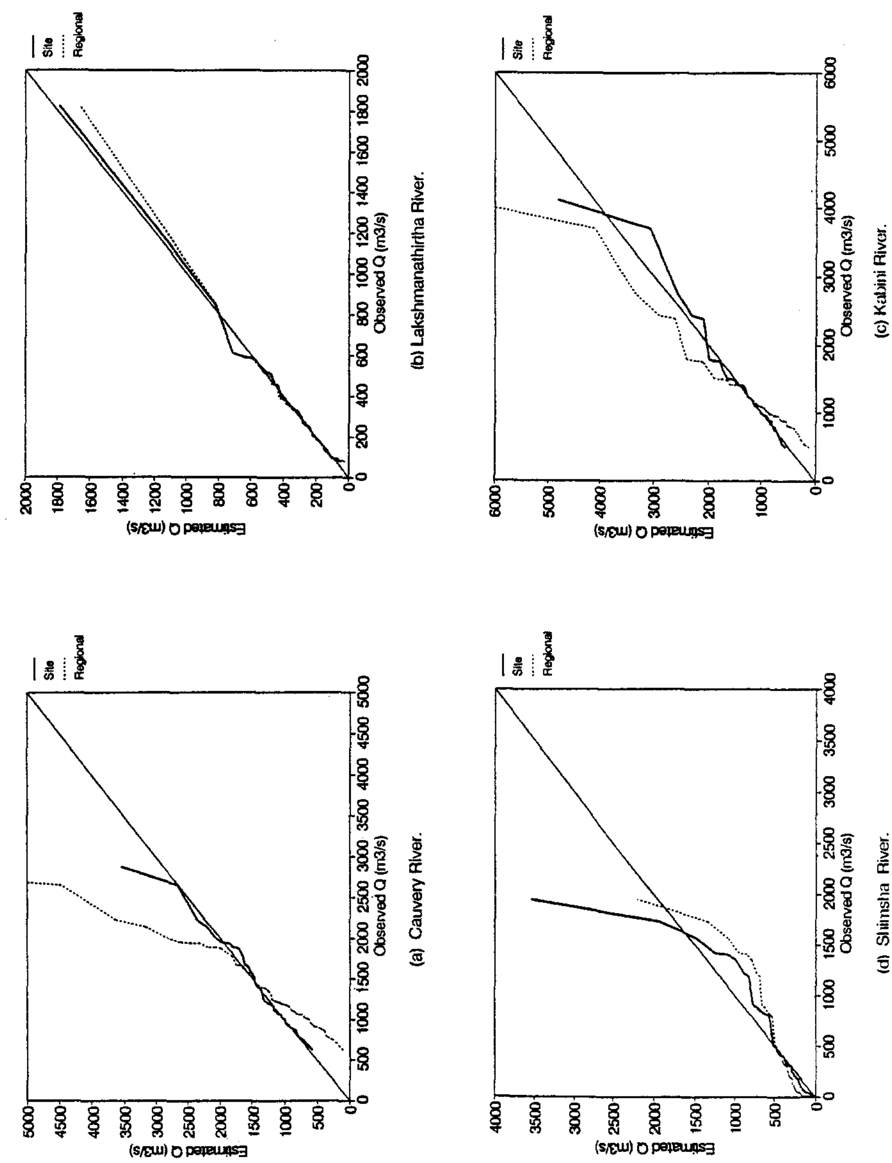

Fig. 5. The estimates and observed data for the GLOG distribution.

TABLE VIII. Parameter estimates for the LN distribution

Site	1	2	3	4	5
α	0.2504	0.2302	0.1624	0.1142	0.4150
μ	–0.3932	–0.4072	–0.5524	–0.6410	–0.8416
σ	0.4582	0.5396	0.8663	1.0195	0.7815

Site	6	7	8	9	Regional
α	–0.3056	0.0221	–0.4774	–0.0870	0.0312
μ	–0.1201	–0.4570	0.2124	–0.1623	–0.3116
σ	0.8794	0.9323	0.5965	0.7010	0.7482

$$\hat{a} = l_1 - \exp\left(\hat{\mu} + \frac{\hat{\sigma}^2}{2}\right), \tag{28}$$

$$Z = \sqrt{(8/3)}\Phi^{-1}\left(\frac{1+t_3}{2}\right), \tag{29}$$

where Φ^{-1} is the inverse standard Normal distribution function, and erf (.) is the error function, erf $(x) = 2\,\Phi\,(x\sqrt{2})$ -1. Parameter estimates for each station as well as their regional averages are given in Table VIII. Quantile estimates are obtained using Equation (30).

$$q_T = \exp[\hat{\sigma}\ \Phi^{-1}\ (F) + \hat{\mu}] + \hat{a} \tag{30}$$

The at-site as well as regional quantile estimates along with the observed data for Cauveri, Lakshmanathirtha, Kabini, and Shimsha rivers are shown in Figure 6.

3.4. PEARSON-III DISTRIBUTION

Parameter estimates for the P-III distribution are obtained from Equations (31)–(34) (Hosking, 1991).

For $t_3 \geq 1/3$ then $t = 1 - t_3$ and

$$\hat{\beta} = \frac{(0.36067t - 0.59567t^2 + 0.25361t^3)}{(1 - 2.78861t + 2.56096t^2 - 0.77045t^3)} \tag{31}$$

and for $t_3 < 1/3$ then $t = 3\pi\, t_3^2$

$$\hat{\beta} = \frac{(1 + 0.2906t)}{(t + 0.1882t^2 + 0.0442t^3)}, \tag{32}$$

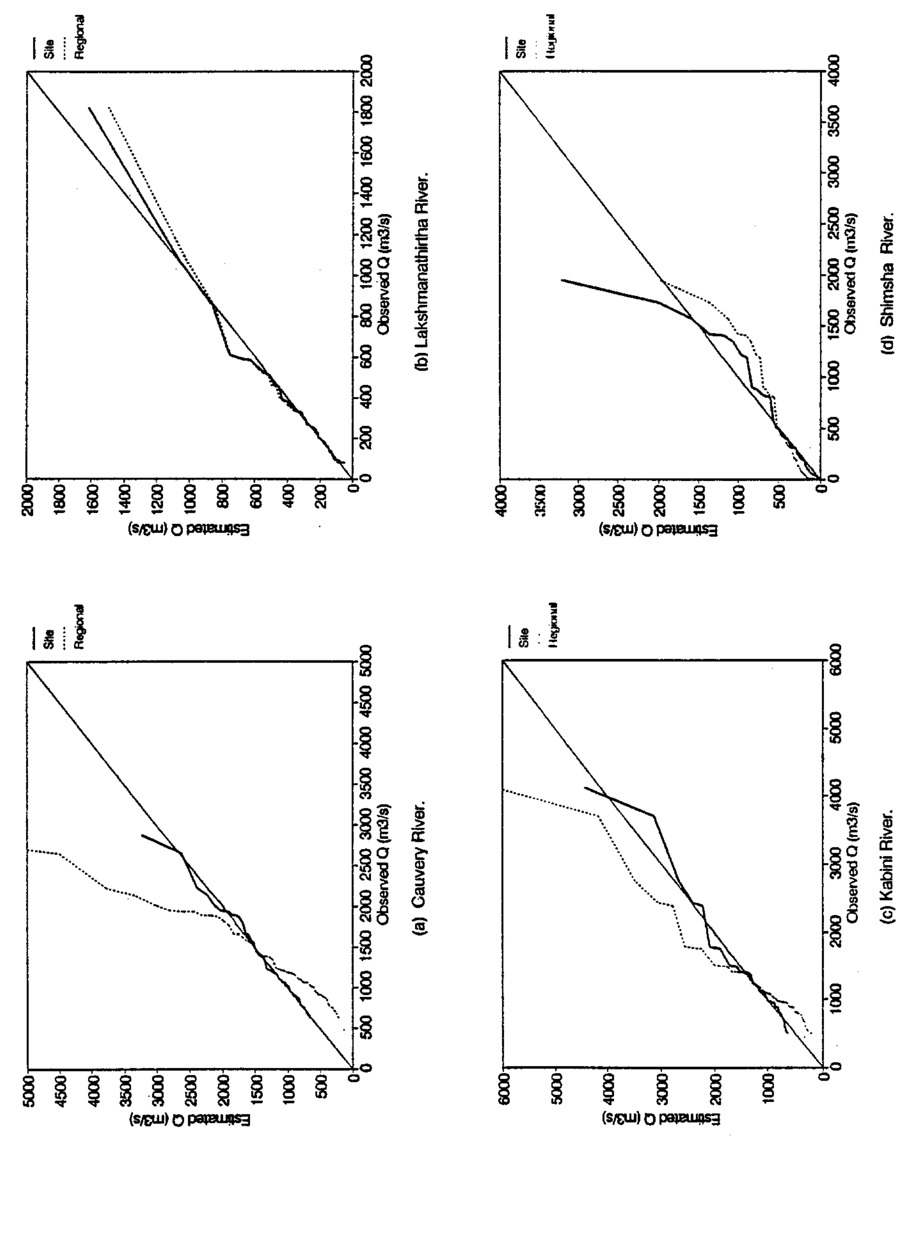

Fig. 6. The estimates and observed data for the LN distribution.

TABLE IX. Parameter estimates for the P-III distribution

Site	1	2	3	4	5
α	0.2388	0.3418	0.9970	1.5050	0.5592
β	2.2327	1.6363	0.6661	0.4844	0.8114
γ	0.4668	0.4407	0.3359	0.2709	0.5463

Site	6	7	8	9	Regional
α	1.6052	1.3644	0.8053	0.8272	0.8448
β	0.6470	0.5777	1.3535	0.9974	0.8815
γ	−0.0386	0.2117	−0.0899	0.1750	0.4668

$$\hat{\alpha} = \sqrt{\pi} l_2 \frac{\Gamma(\hat{\beta})}{\Gamma(\hat{\beta} + 1/2)}, \tag{33}$$

$$\hat{\gamma} = l_1 - \hat{\alpha}\hat{\beta} . \tag{34}$$

Parameter estimates as well as their regional averages are given in Table IX. Quantile estimates can be obtained using the frequency factor formula in Equation (35), where K_T is computed using any of the formulae given by Bobée and Ashkar (1991) corresponding to a given probability F. The at-site as well as regional quantile estimates along with the observed data for Cauveri, Lakshmanathirtha, Kabini, and Shimsha rivers are shown in Figure 7.

$$q_T = \hat{\alpha}\hat{\beta} + \hat{\gamma} + K_T\sqrt{\hat{\alpha}^2\hat{\beta}}. \tag{35}$$

4. Discussion and Results

From the results given in Figures 4–7, in general, three parameter distributions are acceptable for fitting the observed data. According to the goodness-of-fit measure (Z^{DIST}) the regional estimates obtained by the P-III distribution do not agree with the observed data as well as the other three distributions. However, it is hard to reach any conclusions by using only the graphical results. This is due to the fact that the difference in t_4 between the regional value and the fitted distribution in Figure 3, which is a measure for the goodness of fit (Hosking and Wallis, 1991), is not that much larger for the P-III distribution than it is, for example, the GLOG distribution.

Examination of the results in Figure 4 shows that the difference between regional and at-site estimates, for the same distribution, depends on the site location, being maximum for station 1 and minimum for station 3. Results in Figure 3 offer no explanation for this observation. Figure 2 explains these results better because,

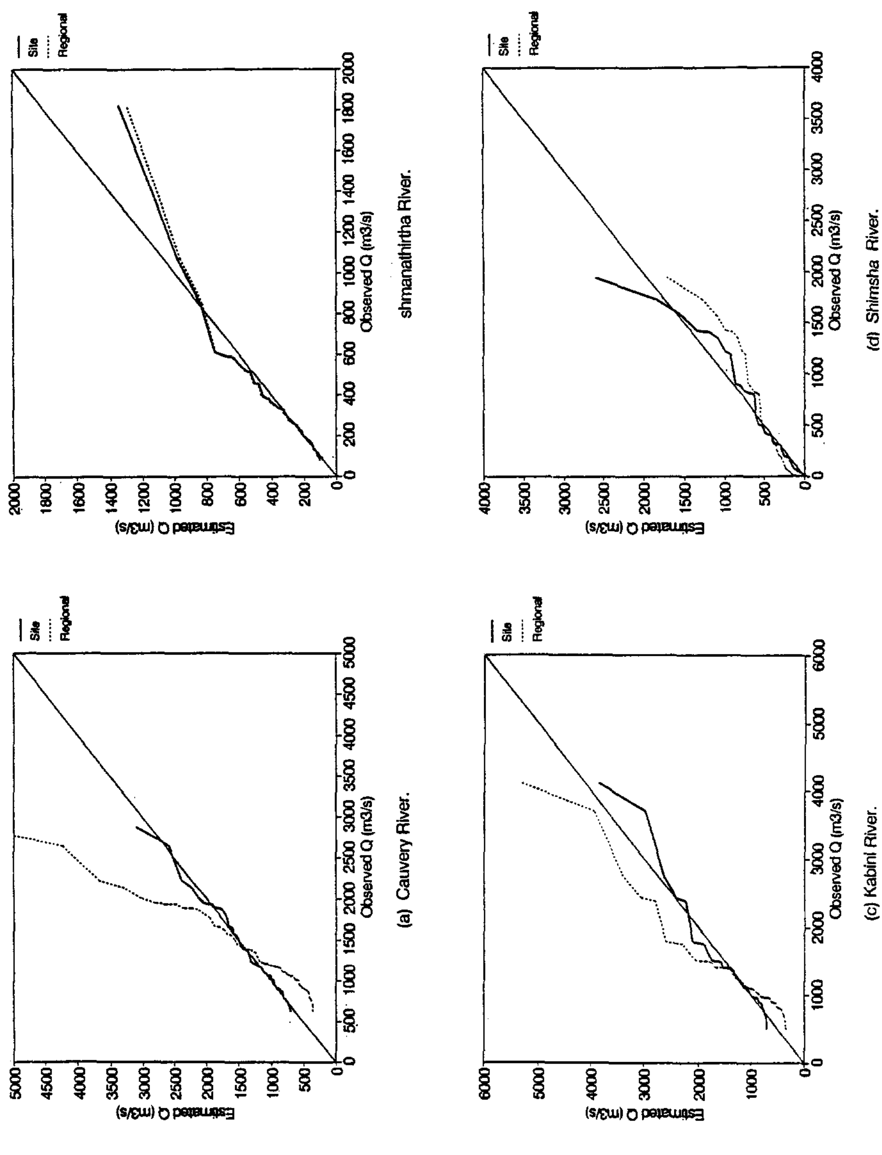

Fig. 7. The estimates and observed data for the P-III distribution.

TABLE X. Higher quantile estimates for the LN distribution.

	$F = 0.98$		$F = 0.99$		$F = 0.995$		$F = 0.998$	
Station	R	S	R	S	R	S	R	S
1	4601	2651	5633	2960	6780	3277	8490	3714
2	3782	2472	4630	2824	5573	3194	6978	3715
3	1040	1082	1274	1357	1533	1673	1920	2159
5	4395	3276	5380	3928	6476	4659	8109	5759
4	494	632	605	829	728	1064	912	1442
6	1383	2051	1693	2640	2083	3318	2553	4366
7	381	479	466	616	561	777	703	1029
8	736	800	901	959	1085	1130	1358	1373
9	1233	1257	1510	1528	1818	1826	2276	2264

R: Regional estimates; S: at-site estimates.

for three parameter distributions, the parameter estimates depend on λ_1, λ_2 and t_3 which are included in Figure 2. Figure 3 contains t_3 and t_4, but only t_3 is used in parameter estimation. Therefore, the closer the statistics of a site is to the average in Figure 2, the closer is the agreement between regional and at-site estimates. In this sense, the heterogeneity measure H_2 based on the LC_v / LC_s ratio quantifies the average difference between regional and at-site estimates. For stations with t less than the average value in Figure 2 the regional estimates of high quantiles are greater than the at-site estimates, and the opposite is true for stations with t greater than the average value of t. This conclusion is supported, in general, by the results in Table X. When at-site estimates are considered, each site is treated individually. The goodness of fit measure (Z) depends only on the difference between the regional average and fitted distribution value of t_4. The difference between at-site estimates and observed data depends on the difference between the at-site and fitted distribution value of t_4, i.e. the location of the site statistic with respect to the distribution curves on the LC_s vs LC_k diagram given in Figure 3. The data from Station 3, according to Figure 3, must fit the GEV distribution much better than the data from Station 6. The Regional heterogeneity measure H_3, based on the LC_s / LC_k distance, gives an estimate of the average deviation between at-site estimates and the observed data.

Both the heterogeneity measures H_2 and H_3 based on LC_v / LC_s and LC_s / LC_k distances respectively, are important to select a good distribution for regional analyses. The measure H_2 indicates whether at-site and regional estimates are close to each other. The measure H_3 indicates whether the at-site estimates and the observed data are in agreement. A large value of H_2 usually indicates a large deviation between regional and at-site estimates and in turn between regional estimates and the observed data. Similarly, a large H_3 value indicates, in general, a large deviation between at-site estimates and the observed data, and in turn between

TABLE XI. Quantile estimates ($F = 0.998$) at different sites

Station	Regional				At-site			
	GLOG	GEV	LN	P-III	GLOG	GEV	LN	P-III
1	10628	10176	8490	6833	4389	3845	3714	3483
2	8735	8364	6978	5617	4487	3946	3715	3393
3	2403	2301	1920	1545	2703	2513	2159	1628
4	1142	1093	912	734	1797	1712	1442	987
5	10150	9719	8109	6525	7094	6526	5759	4650
6	3195	3059	2553	2054	5523	5131	4366	3209
7	879	842	703	565	1292	1214	1029	740
8	1700	1628	1358	1093	1715	1501	1373	1201
9	2848	2727	2276	1832	2834	2545	2264	1866

regional estimates and the observed data, as in the case with large H_2. In the present study, the value of H_2 in Table IV is large (4.37) whereas H_3 is small (0.38). As a result, at-site estimates are much closer to the observed data than the regional estimates as shown in Figures 4–7. Consequently, the region may have to be further subdivided to obtain better regional estimates.

Finally, quantile estimates for a probability $F = 0.998$ are given in Table XI. For both regional and at-site estimates, the estimates for a certain station in descending order are given by GLOG, GEV, LN and P-III. These results conform to the relative position of these distributions in Figure 3. For a given station, the appropriate distribution is the one which is closest to that station in Figure 3. If the statistic for a station plots higher than all distributions such as station 5, then higher quantiles will tend to be underestimated. If a station plots lower than other distributions such as station 6, then the flood estimate tends to be overestimated. If the station plots within the distributions (station 1), the distributions below will tend to underestimate and the distributions above will tend to overestimate flood magnitudes.

5. Conclusions

On the basis of results presented herein, the following conclusions are presented.

(1) The Upper Cauvery basin is hydrologically heterogeneous. It will have to be subdivided into smaller regions to get hydrologically homogeneous regions.

(2) As a result of conclusion 1 above, the at-site estimates are better than the regional estimates for the Upper Cauvery river data.

(3) The lognormal or the generalized extreme value distributions is recommended for use in the Upper Cauvery river basin.

Acknowledgement

Dr. J. R. M. Hosking most graciously sent us his computer program and reports for our use. We are grateful to him. We would like to thank Dr R. Prasad for sending us the data used in this study.

References

Bobée, B. and Ashkar, F.: 1991, *The Gamma Family and Derived Distributions Applied in Hydrology*, Water Resour. Publications, Colorado.

Cunnane, C.: 1989, Statistical distributions for flood frequency analysis, WMO Operational Hydrology Report No. 33, WMO-No. 718.

Greenwood, J. A., Landwehr, J. M., and Matalas, N. C.: 1979, Probability weighted moments: Definition and relation to parameters of several distributions expressible in the inverse form, *Water Resour. Res.* **15**(5), 1049–1054.

Hosking, J. R. M.: 1986, The theory of probability weighted moments, IBM Research Report RC 12210 (#54860).

Hosking, J.R.M.: 1990, L-moments: Analysis and estimation of distributions using linear combination of order statistics, *J. Royal Stat. Soc. B* **52**(1), 105–124.

Hosking J. R. M.: 1991, Fortran routines for use with the method of L-moments, version 2. IBM Research Report RC 17097.

Hosking, J. R. M. and Wallis, J. R.: 1991, Some statistics useful in regional frequency analysis, IBM Research Report RC 17096 (#75863).

Hosking, J. R. M. and Wallis, J. R.: 1987, Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics* **29**, 339–349.

Hosking, J. R. M., Wallis, J. R., and Wood, E. F.: 1985, Estimation of the generalized extreme value distribution by the method of probability weighted moments, *Technometrics* **27**, 251–261.

Landwehr, J. M., Matalas, N. C., and Wallis, J. R.: 1979a, Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles, *Water Resour. Res.* **15**(5), 1055–1064.

Landwehr, J. M., Matalas, N. C., and Wallis, J. R.: 1979b, Estimation of parameters and quantiles of Wakeby distribution, Parts 1 and 2, *Water Resour. Res.* **15**(6), 1361–1379.

Ramesh, M., Murthy, M. C. S., and Prasad, R.: 1987, Analysis of flood frequencies in the Cauvery Valley, in *Hydrologic Frequency Modelling*, Proc. Int. Symp. on Flood Frequency and Risk Analysis, D. Reidel, Dordrecht, pp. 505–513.