INSTITUTE
OF
HYDROLOGY

ESTIMATION OF THE GENERALISED EXTREME-VALUE
DISTRIBUTION BY THE METHOD OF
PROBABILITY WEIGHTED MOMENTS

by

J R M HOSKING, J R WALLIS*
& E F WOOD[†]

[* IBM Research Center, New York
† Princeton University:
visiting Institute of Hydrology]

ABSTRACT

The method of probability weighted moments is used
to derive estimators of the parameters and quantiles
of the generalised extreme-value distribution. The
properties of these estimators are investigated in
large samples via asymptotic theory and in small and
moderate samples via computer simulation. Probability
weighted moment estimators have low variance and no
severe bias, and compare favourably with estimators
obtained by the methods of maximum liklihood or
sextiles. The method of probability weighted moments
also yields a convenient and powerful test of whether
an extreme-value distribution is of Fisher-Tippett
type I, II or III.

REPORT NO 89

July 1984

CONTENTS

INTRODUCTION

The generalized extreme-value distribution of Jenkinson (1955) is widely used for modelling extremes of natural phenomena, and is of considerable importance in hydrology, since it is recommended by the Flood Studies Report (NERC 1975) for modelling the distribution of annual maxima of daily streamflows of British rivers. Currently favoured methods of estimation of the parameters and quantiles of the distribution are Jenkinson's (1969) method of sextiles, and the method of maximum likelihood (Jenkinson 1969; Prescott and Walden 1980, 1983). Neither method is completely satisfactory: the justification of the maximum-likelihood approach is based on large-sample theory, and there has been little assessment of the performance of the method when applied to small or moderate samples; while the sextile method involves an inherent arbitrariness (why sextiles rather than, say, quartiles or octiles?), requires interpolation in a table of values of a function in order to estimate the shape parameter of the distribution, and has statistical properties which are not known even for large samples.

Probability weighted moments, a generalization of the usual moments of a probability distribution, were introduced by Greenwood et al (1979). There are several distributions, for example the Gumbel, logistic and Weibull, whose parameters can be conveniently estimated from their probability weighted moments. The Gumbel distribution, being a special case of the generalized extreme-value distribution, is of particular interest. Landwehr et al (1979) investigated the small-sample properties of probability weighted moment estimators (PWM estimators) of parameters and quantiles for the Gumbel distribution and found them superior in many respects to the conventional moment and maximum-likelihood estimators. The estimators used by Landwehr et al (1979) are identical to Downton's (1966b) "Linear estimates with linear coefficients", and thus share the asymptotic properties of the latter: in

particular, the asymptotic efficiencies of the PWM estimators of the Gumbel scale and location parameters are 0.756 and 0.996 respectively.

In the present work we summarize some theory for probability weighted moments and show that they can be used to obtain estimates of parameters and quantiles of the generalized extreme-value distribution. We derive the asymptotic distributions of these estimators, and compare, via computer simulation, the small-sample properties of the probability weighted moment, sextile, and maximum-likelihood estimators. The method of probability weighted moments outperforms the other methods in many cases and will usually be preferred to them. We also derive from the PWM estimator of the shape parameter of the generalized extreme-value distribution a test of whether this shape parameter is zero, and assess the performance of this test by computer simulation.

PROBABILITY WEIGHTED MOMENTS

The probability weighted moments of a random variable X with distribution function $F(x) = P(X < x)$ are the quantities

$$M_{p,r,s} = E[X^p \{F(X)\}^r \{1 - F(X)\}^s] \quad , \tag{1}$$

where p, r and s are real numbers. Probability weighted moments are likely to be most useful when the inverse distribution function $x(F)$ can be written in closed form, for then we may write

$$M_{p,r,s} = \int_0^1 \{x(F)\}^p F^r (1-F)^s dF \quad , \tag{2}$$

and this is often the most convenient way of evaluating these moments. The quantities $M_{p,0,0}$, $p = 1,2,\ldots,$ are the usual noncentral moments of X. The moments $M_{1,r,s}$ may, however, be preferable for estimating the parameters of the distribution of X, since the occurrence of only the first power of X in the expression for $M_{1,r,s}$ means that the relationship between parameters

and moments often takes a simpler form in this case than when using the conventional moments. When r and s are integers, $F^r(1-F)^s$ may be expressed as a linear combination either of powers of F or powers of $(1-F)$, so it is natural to summarize a distribution either by the moments $M_{1,r,0}$, $r = 0,1,2,\ldots,$ or by $M_{1,0,s}$, $s = 0, 1, 2, \ldots$. Greenwood et al (1979) generally favour the latter approach but here we will consider the moments

$$\beta_r = M_{1,r,0} = E[X\{F(X)\}^r], \quad r = 0,1,2,\ldots .$$

Given a random sample of size n from the distribution F, estimation of β_r is most conveniently based on the ordered sample $x_1 < x_2 < \ldots < x_n$. The statistic

$$b_r = n^{-1} \sum_{j=1}^{n} \frac{(j-1)(j-2)\ldots(j-r)}{(n-1)(n-2)\ldots(n-r)} x_j \tag{3}$$

is an unbiased estimator of β_r (Landwehr et al 1979). Instead one may estimate β_r by

$$\hat{\beta}_r [p_{j,n}] = n^{-1} \sum_{j=1}^{n} p_{j,n}^r x_j \tag{4}$$

where $p_{j,n}$ is a plotting position, i.e. a distribution-free estimate of $F(x_j)$. Reasonable choices of $p_{j,n}$, such as $p_{j,n} = (j-a)/n$, $0 < a < 1$, or $p_{j,n} = (j-a)/(n+1-2a)$, $-\frac{1}{2} < a < \frac{1}{2}$, yield estimators $\hat{\beta}_r[p_{j,n}]$ which are asymptotically equivalent to b_r and therefore are consistent estimators of β_r.

The estimators b_r are closely related to U-statistics (Hoeffding 1948), which are averages of statistics calculated from all subsamples of size $j < n$ of a given sample of size n. In particular $b_0 = n^{-1} \sum x_j$ is a trivial example of a U-statistic, and is a natural estimator of location of a distribution;

$$2b_1 - b_0 = \frac{1}{2} U_2 = \frac{1}{2} \binom{n}{2}^{-1} \sum_{i > j} (x_i - x_j) \tag{5}$$

is a U-statistic for estimating the scale of a distribution - the statistic U_2, sometimes known as Gini's mean difference, has a history going back at

least as far as von Andrae (1872), and $\frac{1}{2}\sqrt{\pi}\, U_2$ is a 98% efficient estimator of the scale parameter of a Normal distribution (Downton 1966a); and

$$6b_2 - 6b_1 + b_0 = \frac{1}{3} U_3 = \frac{1}{3} \binom{n}{3}^{-1} \sum_{i>j>k} (x_i - 2x_j + x_k) \qquad (6)$$

is a U-statistic for estimating skewness which has been used as the basis of a test for Normality by Locke and Spurrier (1976). U-statistics are widely used in nonparametric statistics (see, for example, Fraser, 1957, Chap. 4; Randles and Wolfe, 1979, Chap. 3) and their desirable properties of robustness to outliers in the sample, high efficiency and asymptotic normality may be expected to extend to the probability weighted moment estimators b_r and other quantities calculated from them.

PWM ESTIMATORS FOR THE GENERALIZED EXTREME-VALUE DISTRIBUTION

The generalized extreme-value (GEV) distribution, introduced by Jenkinson (1955), combines into a single form the three possible types of limiting distribution for extreme values, as derived by Fisher and Tippett (1928). The distribution function is

$$F(x) = \begin{cases} \exp[-\{1 - k(x-\xi)/\alpha\}^{1/k}], & k \neq 0 \\ \exp[-\exp\{-(x-\xi)/\alpha\}], & k = 0 \end{cases} \qquad (7)$$

with x bounded by $\xi + \alpha/k$ from above if $k > 0$ and from below if $k < 0$. Here ξ and α are location and scale parameters respectively, while the shape parameter k determines which extreme-value distribution is represented: Fisher-Tippett types I, II and III correspond to $k = 0$, $k < 0$ and $k > 0$ respectively. When $k = 0$ the GEV distribution reduces to the Gumbel distribution. The inverse distribution function is

$$x(F) = \begin{cases} \xi + \alpha\{1 - (-\log F)^k\}/k, & k \neq 0 \\ \xi - \alpha \log(-\log F), & k = 0 \end{cases} \qquad (8)$$

The probability weighted moments of the GEV distribution for $k \neq 0$ are given by

$$\beta_r = (r+1)^{-1}[\xi + \alpha\{1 - (r+1)^{-k}\Gamma(1+k)\}/k], \qquad k > -1 \qquad (9)$$

(for proof see Appendix 1). When $k < -1$, β_0 (the mean of the distribution) and the rest of the β_r do not exist. From (9) we have

$$\beta_0 = \xi + \alpha\{1 - \Gamma(1+k)\}/k, \qquad (10)$$

$$2\beta_1 - \beta_0 = \alpha\Gamma(1+k)(1 - 2^{-k})/k, \qquad (11)$$

$$\frac{3\beta_2 - \beta_0}{2\beta_1 - \beta_0} = \frac{1 - 3^{-k}}{1 - 2^{-k}}. \qquad (12)$$

The PWM estimators $\hat{\xi}, \hat{\alpha}, \hat{k}$ of the parameters are the solutions of (10) - (12) for ξ, α and k when the β_r are replaced by their estimators b_r or $\hat{\beta}_r[p_{j,n}]$. To obtain $\hat{k}$ we must solve the equation

$$\frac{3b_2 - b_0}{2b_1 - b_0} = \frac{1 - 3^{-\hat{k}}}{1 - 2^{-\hat{k}}}. \qquad (13)$$

The exact solution requires iterative methods, but because the function $(1 - 3^{-k})/(1 - 2^{-k})$ is almost linear over the range of values of k, $-\frac{1}{2} < k < \frac{1}{2}$, which is usually encountered in practice, low-order polynomial approximations for $\hat{k}$ are very accurate. We propose the approximate estimator

$$\hat{k} = 7.8590c + 2.9554c^2, \quad c = \frac{2b_1 - b_0}{3b_2 - b_0} - \frac{\log 2}{\log 3}, \qquad (14)$$

the error in $\hat{k}$ due to using (14) rather than (13) is less than 0.0009 throughout the range $-\frac{1}{2} < \hat{k} < \frac{1}{2}$. Given $\hat{k}$, the scale and location parameters can be estimated successively from equations (11) and (10) as

$$\hat{\alpha} = \frac{(2b_1 - b_0)\hat{k}}{\Gamma(1+\hat{k})(1 - 2^{-\hat{k}})}, \quad \hat{\xi} = b_0 + \hat{\alpha}\{\Gamma(1+\hat{k}) - 1\}/\hat{k}. \qquad (15)$$

Equations (13) and (15), or their equivalent forms with b_r replaced by

$\hat{\beta}_r$ $[p_{j,n}]$, define the PWM estimators of the parameters of the GEV distribution. Given the estimated parameters, the quantiles of the distribution are estimated using the inverse distribution function (8).

When calculated using b_r as the estimator of β_r, the PWM estimates of the GEV distribution satisfy a feasibility criterion, namely that $\hat{k} > -1$ and $\hat{\alpha} > 0$ almost surely (for proof see Appendix 2). This is clearly a desirable property, since one would like estimates calculated using a set of sample moments to yield an estimated distribution for which the corresponding population moments exist. We have not been able to prove that this feasibility criterion is satisfied when plotting-position estimators $\hat{\beta}_r$ $[p_{j,n}]$ are used, but no examples of the criterion not being satisfied have been discovered in practice.

ASYMPTOTIC DISTRIBUTION OF PWM ESTIMATORS

When modelling the properties of extremes of physical processes it rarely occurs that the available data set is large enough to ensure that asymptotic large-sample theory may be directly applied to the problem. It is nonetheless valuable to investigate the asymptotic properties of a new statistical technique, for two main reasons. First, one may seek to establish the integrity of the technique, in the sense that when a large sample is available, the new method should not be grossly inefficient compared to an established, asymptotically optimal method such as maximum likelihood. Second, asymptotic theory may provide an adequate approximation to some aspect of the distribution of a statistic even in quite small samples. In the present case we shall see that the variance of PWM estimators of parameters and quantiles of the GEV distribution is well approximated by asymptotic theory for sample sizes of 50 or larger.

We consider first the asymptotic distribution of the b_r. From (3), b_r is a linear combination of the order statistics $x_1, \ldots, x_n$, and the results of Chernoff et al (1967) may be used to prove that the vector $b = (b_0 \ b_1 \ b_2)^T$ has asymptotically a multivariate Normal distribution with mean $\beta = (\beta_0 \ \beta_1 \ \beta_2)^T$ and covariance matrix $n^{-1}V$. The elements of V and details of the proof are given in Appendix 3.

The asymptotic distribution of the PWM estimators of the GEV parameters follows from the preceding result. Let $\theta = (\xi \ \alpha \ k)^T$, $\hat{\theta} = (\hat{\xi} \ \hat{\alpha} \ \hat{k})^T$, and write (13) and (15) as the vector equation $\hat{\theta} = f(b)$. Define the 3×3 matrix $G = (g_{ij})$ by $g_{ij} = \partial f_i / \partial b_j$. Then asymptotically $\hat{\theta}$ has a multivariate Normal distribution with mean $f(\beta) = \theta$ and covariance matrix $n^{-1}GVG^T$ (Rao, 1973, p. 388). The covariance matrix has the form

$$n^{-1}GVG^T = n^{-1} \begin{bmatrix} \alpha^2 w_{11} & \alpha^2 w_{12} & \alpha w_{13} \\ \alpha^2 w_{12} & \alpha^2 w_{22} & \alpha w_{23} \\ \alpha w_{13} & \alpha w_{23} & w_{33} \end{bmatrix} . \tag{16}$$

The w_{ij} are functions of k and have a complicated algebraic form, but they can be evaluated numerically and are given in Table 1 for several values of

TABLE 1 ELEMENTS OF THE ASYMPTOTIC COVARIANCE MATRIX OF THE PWM ESTIMATORS OF THE PARAMETERS OF THE GEV DISTRIBUTION

k	w_{11}	w_{12}	w_{13}	w_{22}	w_{23}	w_{33}
-0.4	1.6637	1.3355	1.1405	1.8461	1.1628	2.9092
-0.3	1.4153	0.8912	0.5640	1.2574	0.4442	1.4090
-0.2	1.3322	0.6727	0.3926	1.0013	0.2697	0.9139
-0.1	1.2915	0.5104	0.3245	0.8440	0.2240	0.6815
0.0	1.2687	0.3705	0.2995	0.7395	0.2249	0.5635
0.1	1.2551	0.2411	0.2966	0.6708	0.2447	0.5103
0.2	1.2474	0.1177	0.3081	0.6330	0.2728	0.5021
0.3	1.2438	-0.0023	0.3297	0.6223	0.3033	0.5294
0.4	1.2433	-0.1205	0.3592	0.6368	0.3329	0.5880

k. As **k** approaches $-\frac{1}{2}$, the variance of the GEV distribution becomes infinite and the variances of the b_r and of the PWM parameter estimators are no longer of order n^{-1} asymptotically.

The asymptotic biases of the estimators are of order n^{-1} and can be evaluated by methods similar to those of Rao (1973, p. 388). The biases, graphed in Figure 1, are negligible in large samples provided that $k > -0.4$.
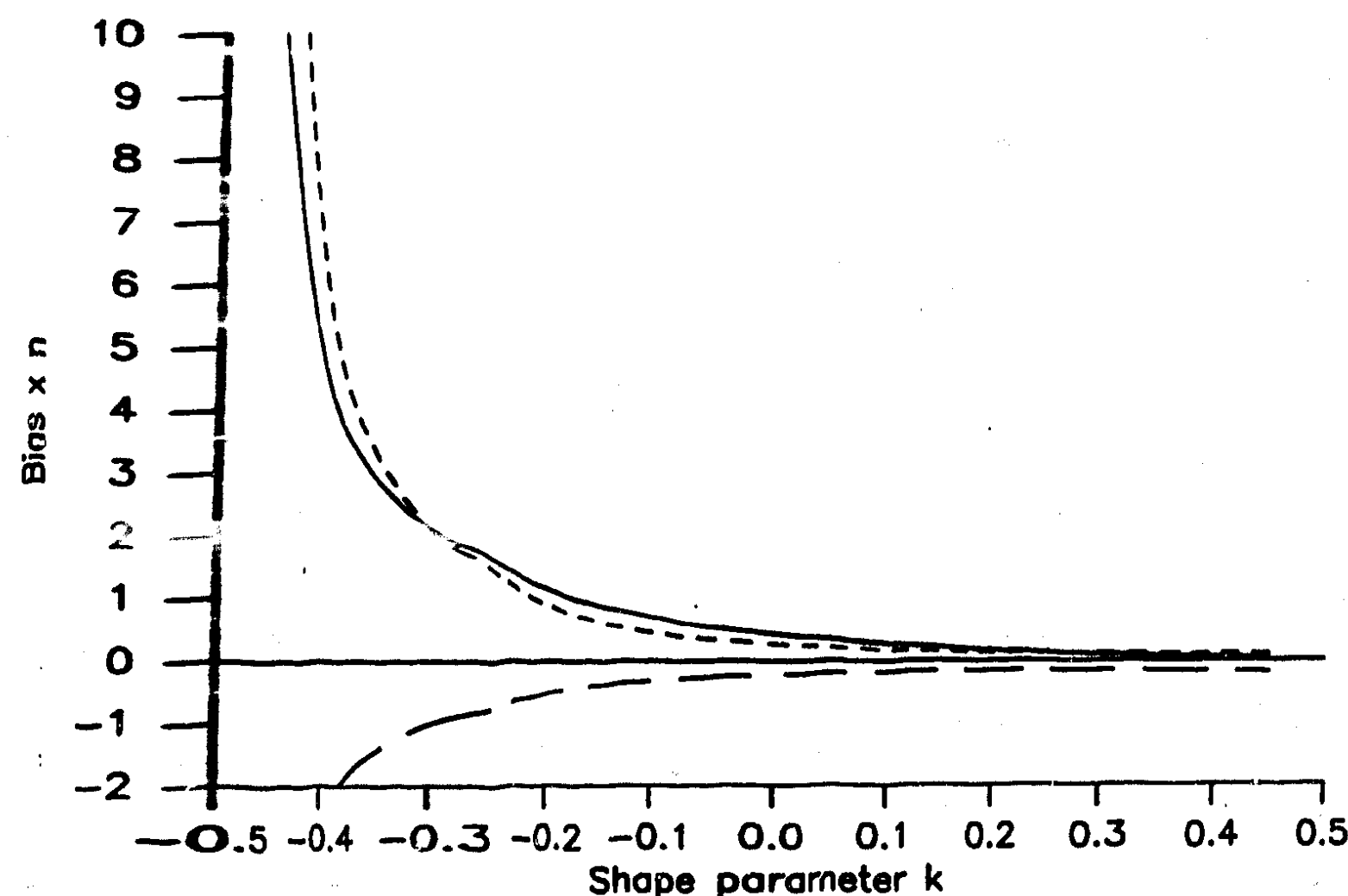

FIGURE 1   Asymptotic bias of PWM estimators of parameters of the GEV distribution $\hat{k}$, $\hat{\alpha}$ $\hat{\xi}$

For comparison, asymptotic biases of the maximum-likelihood estimators of the parameters of the GEV distribution are graphed in Figure 2. These biases were calculated using equation (3.13) of Shenton and Bowman (1977) and are functions of expected values of third derivatives of the log-likelihood function. For $k > 1/3$ these expectations do not exist and the biases of the maximum-likelihood estimators are not of order n^{-1} asymptotically.

In Figures 1 to 4, the curves on the graphs have not been identified. They correspond to parameter estimates, as follows:
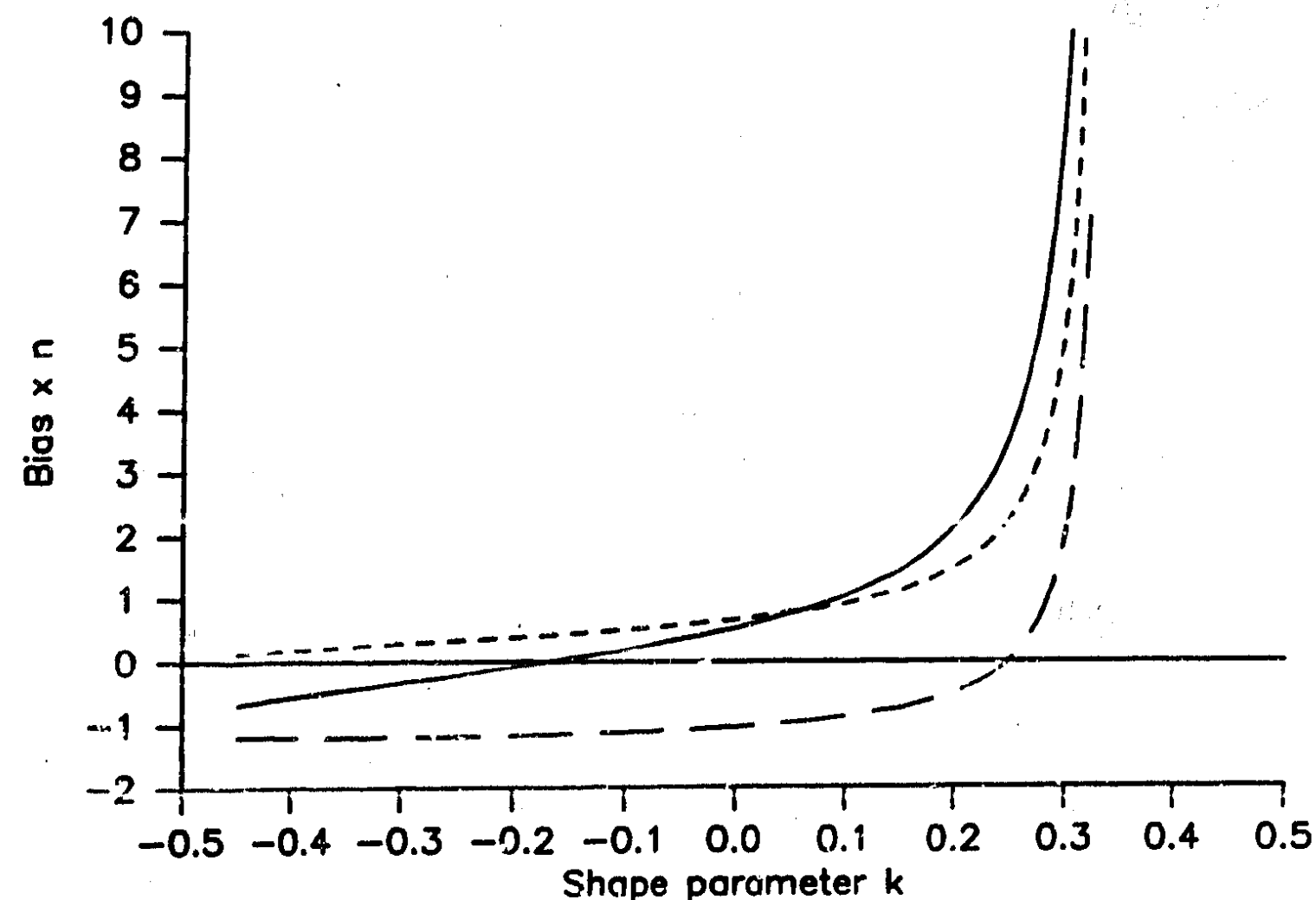——— $\hat{k}$, —— $\hat{\alpha}$, ---- $\hat{\xi}$.

FIGURE 2   Asymptotic bias of ML estimators of parameters of the GEV distribution: $\hat{k}$, $\hat{\alpha}$, $\hat{\xi}$.

The asymptotic variances of the estimators are graphed in Figure 3, and their asymptotic efficiencies in Figure 4. Asymptotic efficiency is defined as the ratio

$$eff(\hat{\theta}_i) = \lim_{n \to \infty} (var\ \tilde{\theta}_i / var\ \hat{\theta}_i)$$

for each element θ_i of the parameter vector θ, where $\tilde{\theta}_i$ is the maximum-likelihood estimator of θ_i. Overall efficiency is the ratio of the determinants of the asymptotic covariance matrices of $\tilde{\theta}$ and $\hat{\theta}$. The overall efficiency of the PWM estimators tends to zero at $k = \pm 0.5$ but for values of k not too far from zero the PWM method is reasonably efficient. Within the range $-0.2 < k < 0.2$, which is valid for many hydrological data sets, each PWM parameter estimator has efficiency of over 0.7.
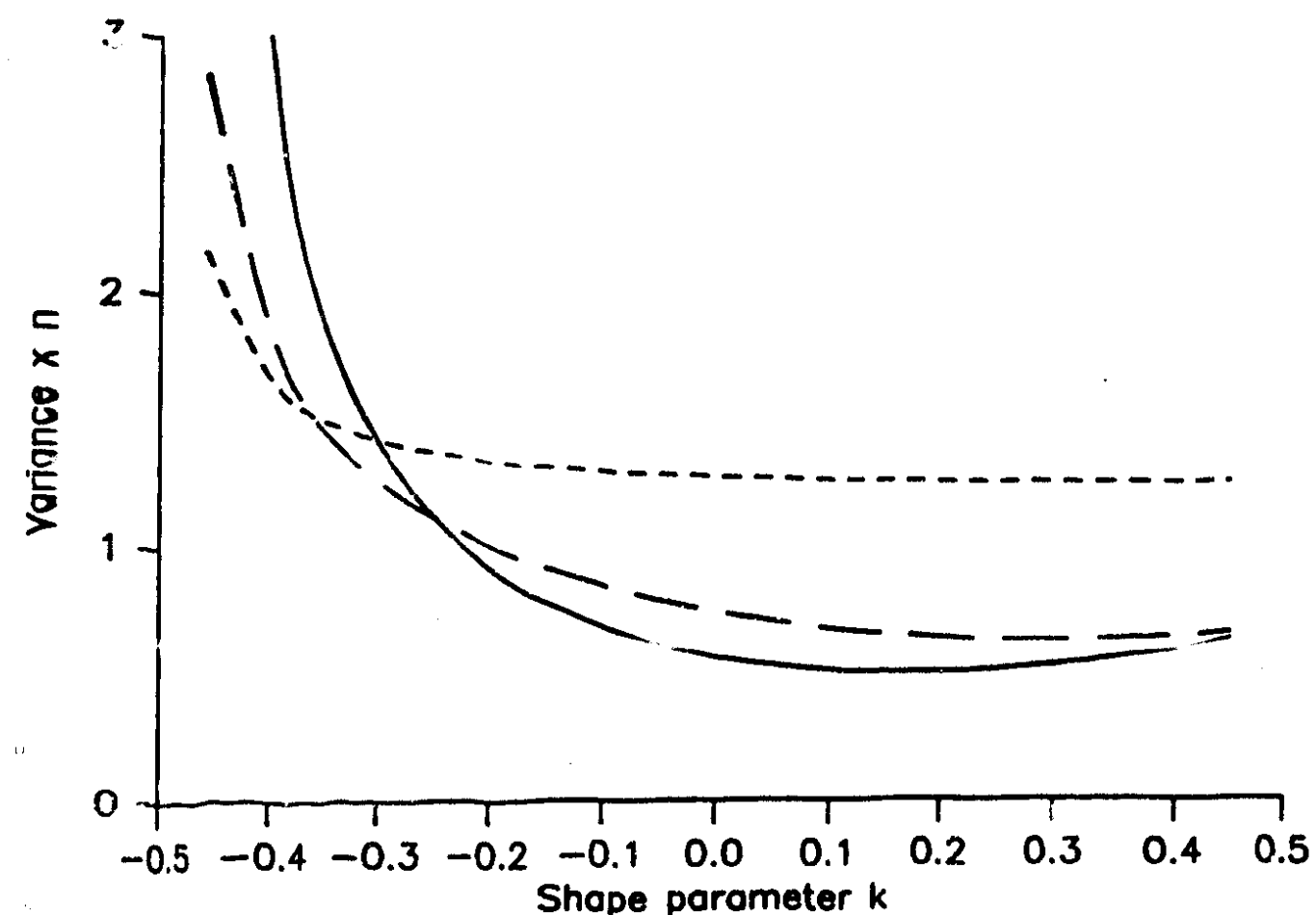
FIGURE 3   Asymptotic variance of PWM estimators of parameters of the
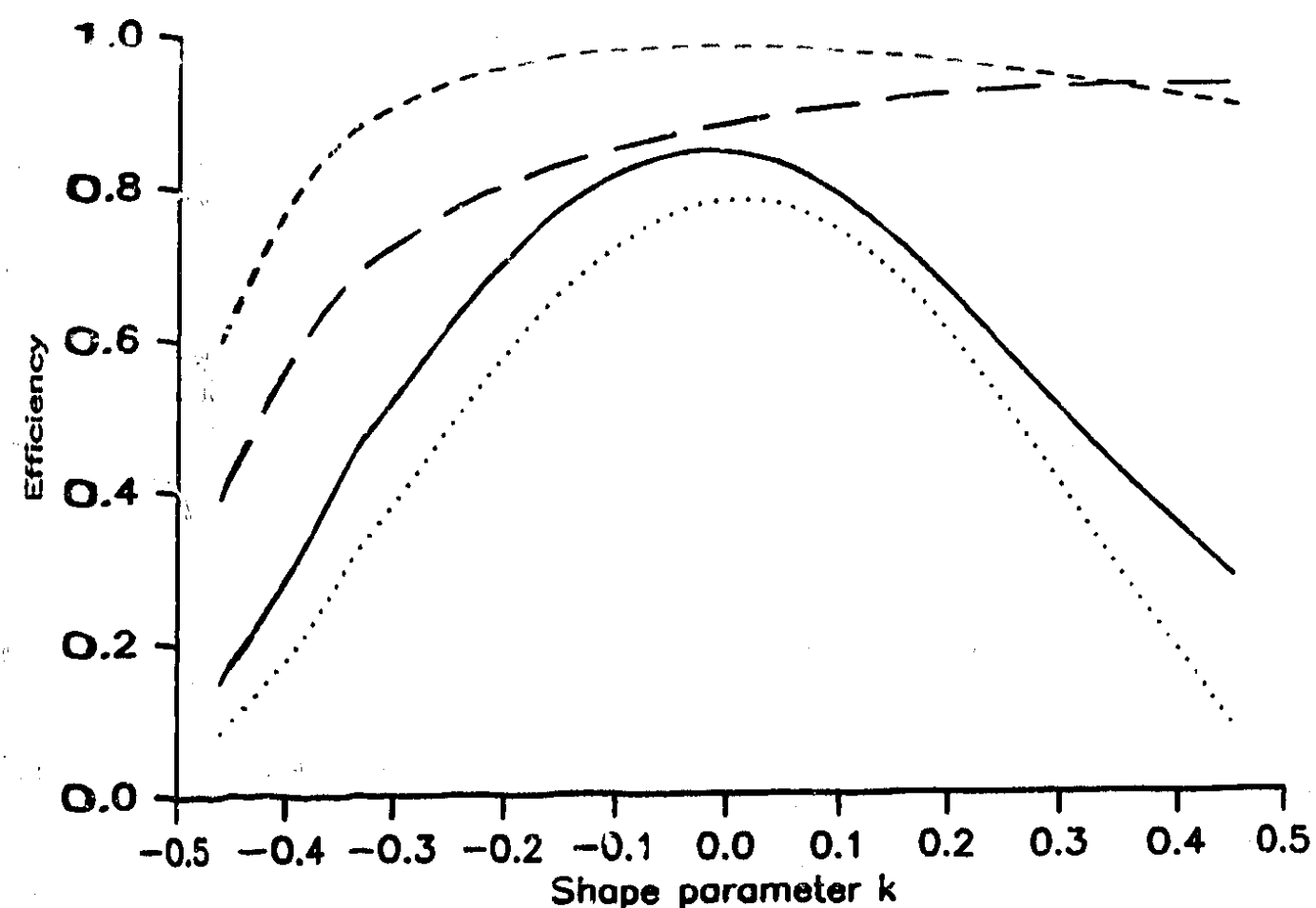GEV distribution: $\hat{k}$, $\hat{\alpha}$, $\hat{\xi}$.

FIGURE 4   Asymptotic efficiency of PWM estimators of parameters of the
GEV distribution: $\hat{k}$, $\hat{\alpha}$, $\hat{\xi}$, ... overall efficiency
i.e. ratio of determinants of asymptotic covariance matrices of
ML and PWM estimators.

Corresponding results may be obtained for PWM and maximum-likelihood
estimators of quantiles of the GEV distribution. These are not presented in
full, but Tables 2 and 3 give results for various quantiles when k = -0.2,
and for various values of k at the F = 0.98 quantile. The tables
illustrate the main characteristics of PWM quantile estimators, which are:
high positive bias in extreme upper tail, arising from positive bias in $\hat{k}$;
high variance in upper tail when k < 0; fair or high efficiency except when

TABLE 2   ASYMPTOTIC BIAS AND VARIANCE OF PWM AND ML ESTIMATORS OF GEV
QUANTILES, AND EFFICIENCY OF THE PWM ESTIMATORS COMPARED WITH
ML. Parameters $\xi = 0, \alpha = 1, k = -0.2$.

F	x(F)	n × bias		n × variance		Efficiency
		PWM	ML	PWM	ML	
0.001	-1.60	-1.2	1.7	3.78	2.29	0.61
0.01	-1.32	-0.2	1.6	2.06	1.35	0.66
0.1	-0.77	0.8	1.2	0.86	0.79	0.92
0.2	-0.45	1.0	0.9	0.88	0.88	1.00
0.5	0.38	0.6	-0.1	1.92	1.79	0.93
0.8	1.75	-1.3	-1.2	6.10	6.00	0.98
0.9	2.84	-3.1	-1.1	16.1	15.9	0.99
0.98	5.91	-4.4	5.5	147	131	0.89
0.99	7.55	-1.6	13.4	336	289	0.86
0.998	12.33	23.9	54.3	1760	1430	0.81
0.999	14.90	49.1	88.2	3310	2630	0.80

TABLE 3   ASYMPTOTIC BIAS, VARIANCE AND EFFICIENCY OF THE PWM AND ML
ESTIMATORS OF THE F = 0.98 QUANTILE OF THE GEV DISTRIBUTION,
AND THE EFFICIENCY OF THE PWM ESTIMATORS.   Parameters $\xi = 0$,
$\alpha = 1$.

k	x(F)	n × bias		n × variance		Efficiency
		PWM	ML	PWM	ML	
-0.4	9.41	-64.8	31.5	1170	574	0.49
-0.3	7.41	-15.3	15.3	369	275	0.75
-0.2	5.91	-4.4	5.5	147	131	0.89
-0.1	4.77	-1.1	-0.3	64.8	62.0	0.96
0.0	3.90	-0.1	-3.8	30.2	28.8	0.95
0.1	3.23	0.3	-6.2	14.7	13.0	0.88
0.2	2.71	0.5	-8.9	7.53	5.62	0.75
0.3	2.30	0.5	-23.2	4.04	2.28	0.56
0.4	1.98	0.6	-	2.28	0.83	0.36

k is close to ±0.5. The maximum-likelihood quantile estimators have lower variance than the PWM estimators but have some very large biases, particularly in the extreme upper tail of the distribution.

The results of this section were derived for PWM estimators which use b_r to estimate β_r. If the plotting-position estimates $\hat{\beta}_r[p_{j,n}]$ are used instead, the asymptotic variances and efficiencies remain unchanged, but the asymptotic biases are different and cannot be easily calculated, being affected by the biases in the $\hat{\beta}_r[p_{j,n}]$ themselves.

SMALL-SAMPLE PROPERTIES OF ESTIMATES OF THE GEV DISTRIBUTION

A computer simulation experiment was run to compare three methods of estimation of the parameters and quantiles of the GEV distribution. Simulations were performed for sample sizes n = 15, 25, 50, 100 with the shape parameter of the distribution taking values k = -0.4, -0.2, 0.0, 0.2, 0.4. All the methods of estimation are invariant under linear transformations of the data, so without loss of generality the location and scale parameters $\xi = 0$ and $\alpha = 1$ were used throughout. For each combination of values n and k, 10000 random samples were generated from the GEV distribution, and for each sample the parameters ξ, α and k, and the quantiles x(F), where F = 0.001, 0.01, 0.1, 0.2, 0.5, 0.8, 0.9, 0.98, 0.99, 0.998, 0.999, were estimated by each of three methods: (1) the method of probability weighted moments (PWM), described above; (2) the method of maximum likelihood (ML), using Newton-Raphson iteration to maximize the likelihood function, as recommended by Prescott and Walden (1983); and (3) Jenkinson's (1969) method of sextiles (JS). The PWM method requires a choice of a suitable estimator of β_r. Several possibilities were investigated, including the unbiased estimator b_r and a number of plotting-position estimators $\hat{\beta}_r[p_{j,n}]$. The

best overall results were given by the estimator $\hat{\beta}_r[p_{j,n}]$ with $p_{j,n} = (j-0.35)/n$, and the simulation results presented below for the PWM method refer to this version of the PWM estimators.

For some simulated samples, the maximum-likelihood and sextile estimates could not be found. The cause of this problem for the maximum-likelihood method was nonconvergence of the Newton-Raphson iteration, usually due to an extreme outlier in the sample; for the sextile method, a ratio of sextile means used to estimate the shape parameter of the GEV distribution occasionally lay outside the range of a table of values in which it was to be interpolated. Such cases were omitted from the simulations. No such problems were encountered with the PWM estimators, which could always be calculated.

The simulation results for estimation of the parameters of the GEV distribution are summarized in Table 4 and 5. Results for the estimator of k are of the greatest importance, since this parameter determines the overall shape of the GEV distribution and the rate of increase of the upper quantiles x(F) as F approaches 1. Apart from the case n = 100, when all the methods have comparable performance, the PWM estimator has consistently the lowest standard deviation of the three estimators of k, its advantage being particularly marked in small samples, n = 15 and n = 25. The PWM estimator has in general a larger bias than the other estimators but its bias is small near the important value k = 0 and is in any case relatively insignificant compared to the standard deviation in its contribution to the mean square error of $\hat{k}$. The sextile estimator of k has a large positive bias in small samples when k < 0 and its standard deviation is generally larger than that of the PWM estimator.

Similar results can be seen for estimators of ξ and α, although the differences between the variances of the estimators are less marked than is the case with estimators of k. In general, PWM estimators have smallest standard deviation, particularly for n = 15 and n = 25, and their bias,

TABLE 4 BIAS OF ESTIMATORS OF GEV PARAMETERS

		Bias($\hat\xi$)					Bias($\hat a$)					Bias($\hat k$)				
k:		-0.4	-0.2	0.0	0.2	0.4	-0.4	-0.2	0.0	0.2	0.4	-0.4	-0.2	0.0	0.2	0.4
n	Method															
	PWM	.10	-.05	-.02	.00	-.03	.06	-.06	-.10	-.11	-.12	.11	.03	-.03	-.08	-.12
15	ML	.03	-.03	.05	.05	.04	-.07	-.07	-.07	-.06	-.07	-.04	-.02	.02	.04	.03
	JS	.11	-.08	.06	.04	.02	-.05	-.06	-.07	-.08	-.08	.10	.06	.03	.01	-.01
	PWM	.06	-.03	.01	-.01	-.02	.00	-.04	-.06	-.07	-.07	.08	.02	-.02	-.05	-.07
25	ML	.01	-.02	.03	.03	.04	-.04	-.04	-.04	-.03	-.03	-.02	-.01	.02	.04	.05
	JS	.06	-.04	.03	.02	.01	-.03	-.03	-.04	-.04	-.05	.07	.04	.02	.01	-.00
	PWM	.04	-.02	.01	.00	-.01	.01	-.02	-.03	-.04	-.04	.05	.02	-.01	-.02	-.04
50	ML	.01	-.01	.02	.02	.02	-.02	-.02	-.02	-.02	-.01	-.01	.00	.01	.02	.03
	JS	.0?	-.0?	.0?	.01	.01	-.0?	-.0?	-.0?	-.0?	-.02	.04	.02	.01	.00	.00
	PWM	.02	-.01	.00	.00	-.01	.00	-.01	-.02	-.02	-.02	.03	.01	.00	-.01	-.02
100	ML	.00	-.00	.01	.01	.01	-.01	-.01	-.01	-.01	.00	-.01	.00	.00	.01	.02
	JS	.02	-.01	.01	.00	.00	-.01	-.01	-.01	-.01	-.01	.02	.01	.00	.00	.00

TABLE 5 STANDARD DEVIATION OF ESTIMATORS OF GEV PARAMETERS

		St.dev.($\hat\xi$)					St.dev($\hat a$)					St.dev($\hat k$)				
k:		-0.4	-0.2	0.0	0.2	0.4	-0.4	-0.2	0.0	0.2	0.4	-0.4	-0.2	0.0	0.2	0.4
n	Method															
	PWM	.32	.30	.29	.28	.28	.33	.25	.21	.19	.19	.20	.19	.18	.18	.19
15	ML	.32	.32	.31	.30	.28	.28	.25	.23	.22	.21	.36	.32	.29	.27	.23
	JS	.33	.31	.30	.29	.28	.32	.28	.24	.21	.21	.25	.24	.23	.23	.23
	PWM	.24	.23	.22	.22	.22	.24	.19	.17	.15	.16	.18	.16	.14	.14	.15
25	ML	.24	.24	.23	.23	.22	.21	.19	.17	.16	.17	.24	.21	.20	.18	.17
	JS	.24	.23	.23	.22	.22	.24	.21	.18	.17	.16	.19	.18	.17	.17	.17
	PWM	.17	.16	.16	.16	.16	.17	.14	.12	.11	.11	.14	.12	.11	.10	.11
50	ML	.17	.16	.16	.16	.16	.15	.13	.12	.11	.11	.15	.13	.12	.11	.11
	JS	.17	.16	.16	.16	.16	.17	.14	.13	.12	.11	.14	.13	.12	.11	.12
	PWM	.12	.12	.11	.11	.11	.12	.10	.09	.08	.08	.11	.09	.07	.07	.08
100	ML	.12	.11	.11	.11	.11	.10	.09	.08	.08	.08	.10	.09	.08	.07	.07
	JS	.12	.11	.11	.11	.11	.12	.10	.09	.08	.08	.10	.09	.08	.08	.08

though often larger than for the ML or sextile estimators, is not severe. The standard deviations of the PWM estimators for $n > 50$ are well approximated by their large-sample values given by (16) and Table 1. Maximum-likelihood estimators are the least biased but are more variable than PWM estimators in small samples. Even at sample size 100, the asymptotic inefficiency of the PWM method compared to maximum likelihood is not apparent in the simulation results. Sextile estimators in general have larger standard deviations than PWM estimators and have some significant biases in small samples when $k < 0$.

The statistical properties of estimators of quantiles of the GEV distribution were evaluated for many combinations of quantiles and values of the shape parameter k, and only a few representative simulation results are presented in Table 6. The most important aspect of quantile estimation in hydrological applications is estimation of the extreme upper quantiles, particularly for heavy-tailed GEV distributions with $k < 0$. Table 5 gives the bias and standard deviation of the estimated upper quantiles for two GEV

TABLE 6 BIAS AND STANDARD DEVIATION OF ESTIMATORS OF GEV QUANTILES.
Tabulated values are bias and standard deviation of the ratio $\hat x(F)/x(F)$ rather than of the quantile estimator $\hat x(F)$ itself.

		k=-0.2						k=0.2					
		F=0.9		F=0.99		F=0.999		F=0.9		F=0.99		F=0.999	
		x(F)=2.84		x(F)=7.55		x(F)=14.90		x(F)=1.81		x(F)=3.01		x(F)=3.74	
n	Method	bias	s.d.	bias	s.d.	bias	s.d.	bias	s.d.	bias	s.d.	bias	s.d.
	PWM	-.06	.34	-.02	.55	.15	1.12	-.04	.23	.08	.32	.25	.56
15	ML	.01	.50	*	*	*	*	-.06	.23	.02	.79	.44	8.10
	JS	-.08	.34	-.05	.56	.14	1.20	-.06	.22	-.01	.33	.11	.64
	PWM	-.04	.27	-.01	.45	.11	.88	-.02	.18	.05	.24	.14	.39
25	ML	-.01	.32	.16	.97	.74	9.59	-.04	.17	-.04	.29	.02	.86
	JS	-.05	.27	-.02	.46	.09	.89	-.04	.17	-.01	.25	.05	.42
	PWM	-.02	.19	-.01	.33	.06	.61	-.01	.12	.02	.16	.07	.25
50	ML	-.01	.22	.05	.40	.18	.86	-.03	.12	-.03	.16	-.02	.28
	JS	-.02	.19	-.01	.33	.04	.59	-.02	.12	.01	.17	.02	.27
	PWM	-.01	.14	.00	.24	.04	.42	-.01	.09	.01	.12	.03	.17
100	ML	.00	.15	.02	.25	.08	.43	-.01	.09	-.02	.11	-.01	.15
	JS	-.01	.14	.00	.24	.03	.40	-.01	.09	.00	.12	.01	.18

* indicates values which varied widely between different sets of 1000 simulations and consequently could not be estimated accurately.

distributions, one with k < 0 and one with k > 0. Results are presented for the ratios $\hat{x}(F)/x(F)$ rather than for the $\hat{x}(F)$ themselves, since the former quantities are more easily compared at different F values. For sample size 100 the three methods have comparable performance. In small samples the upper quantiles obtained by PWM estimation are rather biased, but are still preferable to the maximum-likelihood estimators since these have very large biases and standard deviations. The errors in the maximum-likelihood quantile estimators arise chiefly from a small number of simulated series which yield large negative estimates of k, and consequently give very large estimates of extreme upper quantiles.

Estimation of extreme lower quantiles tends to be less important in practice than estimation of upper quantiles, so simulation results for this case are not given in detail. All three methods give comparable results when n > 50, but for small samples the PWM estimators have smallest standard deviation and small or moderate bias, and are generally to be preferred.

All the methods of quantile estimation are very inaccurate when estimating extreme quantiles in small samples with k < 0. It is of course to be expected that a quantile x(F) cannot be estimated reliably from a sample of size n if F > 1-1/n. Nonetheless it is sometimes possible to obtain useful estimates of extreme quantiles from short data sets, by combining information from a number of independent data sets. Such a "regionalization procedure", based on the PWM estimation method for the GEV distribution, is described in Hosking et al (1984).

TESTING WHETHER THE SHAPE PARAMETER IS ZERO

The type I extreme-value distribution, or Gumbel distribution, is a particularly simple special case of the GEV distribution, and it is often

useful to test whether a given set of data is generated by a Gumbel rather than a GEV distribution. This is equivalent to testing whether the shape parameter k is zero in the GEV distribution. A test of this hypothesis may be based on the PWM estimator of k. On the null hypothesis $H_0 : k = 0$ the PWM estimator $\hat{k}$ is asymptotically distributed as $N(0, 0.5635/n)$ so the test may be performed by comparing the statistic $Z = \hat{k}(n/0.5635)^{1/2}$ with the critical values of a standard Normal distribution. Significant positive values of Z imply rejection of H_0 in favour of the alternative k > 0, and significant negative values of Z imply rejection in favour of k < 0.

The size of the test based on Z for various sample sizes and its power for sample size 50 are given in Tables 7 and 8. These results are based on computer simulations of 50000 samples for each value of n and k. The results may be compared with Hosking's (1984) survey of tests of this hypothesis: the Z-test has power almost as high as the likelihood-ratio test and for samples of size 25 or more its distribution on H_0 is adequately approximated by the standard Normal significance levels. Since the statistic Z is very simple to compute, the Z-test can be strongly recommended as a convenient and powerful indicator of the sign of the shape parameter of the GEV distribution.

TABLE 7 EMPIRICAL SIGNIFICANCE LEVELS OF THE STATISTIC Z FOR TESTING THE HYPOTHESIS H : k=0 AGAINST ONE-SIDED AND TWO-SIDED ALTERNATIVES.

Alternative:	k<0		k>0		k≠0	
Nominal level:	10%	5%	10%	5%	10%	5%
Sample size						
15	10.3	4.3	7.3	3.7	8.0	3.5
25	10.4	4.6	8.4	4.3	8.9	4.1
50	10.5	4.9	8.9	4.6	9.6	4.7
100	10.4	5.1	9.4	4.9	10.0	5.1
200	10.4	5.0	9.7	5.1	10.2	5.2
500	10.5	5.3	9.6	4.9	10.2	5.1

TABLE 8 POWER OF THE Z-TEST OF THE HYPOTHESIS k=0 AGAINST ONE-SIDED AND TWO-SIDED ALTERNATIVES. Sample size 50, nominal significance level 5%

Alternative:	k<0	k>0	k≠0
k			
-0.5	.96	-	.94
-0.4	.90	-	.85
-0.3	.77	-	.68
-0.2	.54	-	.43
-0.1	.25	-	.17
0.0	.05	.05	.05
0.1	-	.18	.11
0.2	-	.50	.37
0.3	-	.83	.73
0.4	-	.96	.93
0.5	-	1.00	.99

CONCLUSIONS

Estimators of parameters and quantiles of the GEV distribution have been derived using the method of probability weighted moments. These estimators have several advantages over existing methods of estimation. They are fast and straightforward to compute and always yield feasible values for the estimated parameters. The biases of the estimators are small, except when estimating quantiles in the extreme tails of the GEV distribution, and decrease rapidly as the sample size increases. The standard deviations of the PWM estimators are comparable with those of the maximum-likelihood estimators for moderate sample sizes (n = 50, 100) and often substantially less than those of the maximum likelihood estimators for small samples (n = 15, 25). PWM estimators of GEV parameters and quantiles have asymptotic Normal

distributions and the large-sample approximation to the variance of the estimators is adequate for sample sizes of 50 or more. Although PWM estimators are asymptotically inefficient compared to maximum-likelihood estimators, no inefficiency is detectable in samples of size 100 or less. The PWM estimator of the shape parameter k of the GEV distribution may be used as the basis of a test of the hypothesis $H_0 : k = 0$, and this test is simple to perform, powerful and accurate.

REFERENCES

Chernoff, H., Gastwirth, J. L. and Johns, M. V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. Ann. Math. Statist., 38, 52-72.

Downton, F. (1966a). Linear estimates with polynomial coefficients. Biometrika, 53, 129-141.

Downton, F. (1966b). Linear estimates of parameters in the extreme value distribution. Technometrics, 8, 3-17.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Camb. Phil. Soc., 24, 180-190.

Fraser, D. A. S. (1957). Nonparametric methods of statistics. New York: Wiley.

Gradshteyn, I. S. and Ryzhik, I. W. (1980). Table of integrals, series and products, 4th edn. New York: Academic Press.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C. and Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. Water Resour. Res., 15, 1049-1054.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. Ann. Math. Statist., 19, 293-325.

Hosking, J. R. M. (1984). Testing whether the shape parameter is zero in the generalized extreme-value distribution. Biometrika, 71, in press.

Hosking, J. R. M., Wallis, J. R. and Wood, E. F. (1984). The U.K. Flood Studies Report — nine years on. Submitted for publication.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) of meteorological elements. Quart. J. R. Met. Soc., 81, 158-171.

Jenkinson, A. F. (1969). Statistics of extremes. WMO Technical Note 98. Geneva: World Meteorological Office.

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

Locke, C. and Spurrier, J. D. (1976). The use of U-statistics for testing normality against non-symmetric alternatives. Biometrika, 63, 143-147.

N.E.R.C. (1975). Flood Studies Report, vol. 1, London: Natural Environment Research Council.

Prescott, P. and Walden, A. T. (1980). Maximum-likelihood estimation of the parameters of the generalized extreme-value distribution. Biometrika, 67, 723-724.

Prescott, P. and Walden, A. T. (1983). Maximum-likelihood estimation of the parameters of the three-parameter generalized extreme-value distribution from censored samples. J. Statist. Comput. Simul., 16, 241-250.

Randles, R. H. and Wolfe, D. A. (1979). Introduction to the theory of nonparametric statistics. New York: Wiley.

Rao, C. R. (1973). Linear statistical inference and its applications, 2nd edn. New York: Wiley.

Shenton, L. R. and Bowman, K. O. (1977). Maximum-likelihood estimation in small samples. London: Griffin.

Von Andrae (1872). Uber die Bestimmung das wahrscheinlichen Fehlers durch die gegebenen Differenzen von ein gleich genauen Beobachtungen einer Unbekannten. Astron. Nach., 79, 257-272.

APPENDIX 1   PROBABILITY WEIGHTED MOMENTS FOR THE GEV DISTRIBUTION

For the GEV distribution we have from (2) and (3)

$$\beta_r = M_{1,r,0} = \int_0^1 [\xi + \alpha\{1 - (-\log F)^k\}/k] F^r dF$$

$$= \int_0^\infty \{\xi + \alpha(1-u^k)/k\} e^{-(r+1)u} du \qquad \text{substituting } u = -\log F ,$$

$$= (\xi + \alpha/k) \int_0^\infty e^{-(r+1)u} du - (\alpha/k) \int_0^\infty u^k e^{-(r+1)u} du$$

$$= (\xi + \alpha/k)(r+1)^{-1} - (\alpha/k)(r+1)^{-1+k}\Gamma(1+k) \quad \text{provided that } k > -1 ,$$

$$= (r+1)^{-1}[\xi + \alpha\{1 - (r+1)^{-k}\Gamma(1+k)\}/k] . \tag{A1}$$

APPENDIX 2   FEASIBILITY OF PWM ESTIMATES OF THE GEV PARAMETERS

The PWM estimator $\hat{k}$ satisfies (13), and therefore $\hat{k} > -1$ provided
that

$$(2b_1 - b_0)/(3b_2 - b_0) > \frac{1}{2} . \tag{A2}$$

Now

$$2b_1 - b_0 = \frac{1}{n(n-1)} \sum_{i>j} (x_i - x_j) \tag{A3}$$

and

$$3b_2 - b_0 = \frac{2}{n(n-1)(n-2)} \sum_{i>j>k} (2x_i - x_j - x_k) \tag{A4}$$

are both positive, so (A2) reduces to $b_0 - 4b_1 + 3b_2 < 0$. But we can write

$$b_0 - 4b_1 + 3b_2 = \frac{2}{n(n-1)(n-2)} \sum_{i>j>k} (-x_j + x_k) , \tag{A5}$$

thus $b_0 - 4b_1 + 3b_2 < 0$ almost surely and therefore $\hat{k} > -1$ almost
surely. Results (A3) - (A5) above are easily proved by induction on the
sample size n.

Furthermore, since

$$\hat{\alpha} = \frac{(2b_1 - b_0)\hat{k}}{\Gamma(1+\hat{k})(1-2^{-\hat{k}})} \tag{A6}$$

and $2b_1 - b_0 > 0$ as noted above, $k/(1-2^{-k}) > 0$ for all k and
$\Gamma(1+\hat{k}) > 0$ because $\hat{k} > -1$, it follows that we must have $\hat{\alpha} > 0$.

APPENDIX 3   ASYMPTOTIC DISTRIBUTION OF THE b_r

The statistic b_r may be written as a linear combination of the order
statistics of a random sample:
we have

$$b_r = n^{-1} \sum_{j=1}^n c_{nj}^{(r)} x_j \tag{A7}$$

where $c_{nj}^{(r)} = (j-1)\ldots(j-r)/\{(n-1)\ldots(n-r)\}$ and $x_1 < x_2 < \ldots < x_n$ is the
ordered sample. As $n \to \infty$ and $j \to \infty$ with $j/n \to \theta$, $0 < \theta < 1$, $c_{nj}^{(r)}$ is
asymptotically a function of the plotting position $j/(n+1)$: in fact
$c_{nj}^{(r)} \sim \{j/(n+1)\}^r$. It is straight-forward to verify that b_r satisfies the
conditions of Theorem 1 of Chernoff et al (1967), and from that theorem it
follows that b_r is asymptotically Normally distributed with mean β_r and
variance

$$n^{-1} v_{rr} = 2n^{-1} \iint_{x<y} \{F(x)\}^r \{F(y)\}^r . F(x)\{1 - F(y)\} dxdy . \tag{A8}$$

A similar argument applies to any linear combination of the b_r,
$r = 0, 1, 2, \ldots,$ and it follows that the b_r are asymptotically jointly
Normally distributed with covariance given by

$$v_{rs} = \lim_{n\to\infty} n \, cov(b_r, b_s) = \frac{1}{2} (g_{rs} + g_{sr}) \tag{A9}$$

where

$$g_{rs} = 2 \iint_{x<y} \{F(x)\}^r \{F(y)\}^s . F(x)\{1 - F(y)\} dxdy . \tag{A10}$$

To evaluate the g_{rs} for the GEV distribution we consider first the
case $k > 0$ and let

$$I_{rs} = 2 \iint_{x<y} \{F(x)\}^r \{F(y)\}^s dxdy \tag{A11}$$

so that

$$g_{rs} = I_{r+1,s} - I_{r+1,s+1} \; . \tag{A12}$$

Substituting (7) in (A11) and making the further substitution
$u = \{1 - k(x-\xi)/\alpha\}^{1/k}, \; v = \{1 - k(y-\xi)/\alpha\}^{1/k}$, we have

$$I_{rs} = 2\alpha^2 \int_0^\infty \int_v^\infty u^{k-1} e^{-ru} du \cdot v^{k-1} e^{-sv} dv$$

$$= 2\alpha^2 r^{-k} \int_0^\infty v^{k-1} e^{-sv} \Gamma(k,rv) dv$$

$$= 2\alpha^2 r^{-2} (r+s)^{-2k} \Gamma(1+2k) \, {}_2F_1[1,2k; 1+k; s/(r+s)] \tag{A13}$$

(Gradshteyn and Ryzhik, 1980, pp. 317, 663); here $\Gamma(\cdot,\cdot)$ is the incomplete gamma function and ${}_2F_1$ is the hypergeometric function. It is convenient to transform the hypergeometric function in (A13), using results from Gradshteyn and Ryzhik (1980, p. 1043):

$${}_2F_1[1,2k; 1+k; s/(r+s)] = 2^{2k} \{\Gamma(1+k)\}^2/\Gamma(1+2k) \quad \text{if} \quad r = s \; ,$$

$$= \{r/(r+s)\}^{-2k} G(s/r) \qquad \text{if} \quad r > s \; ,$$

$$= -\{s/(r+s)\}^{-2k} G(r/s) + 2r^{-k} s^{-k} (r+s)^{2k} \{\Gamma(1+k)\}^2/\Gamma(1+2k) \quad \text{if} \quad r < s \; , \tag{A14}$$

where G denotes the hypergeometric function $G(x) = {}_2F_1(k,2k; 1+k; -x)$ – note that $G(0) = 1$. Substituting back into (A12) and (A9) we obtain the following expressions for the v_{rs}:

$$v_{rr} = \alpha^2 k^{-2} (r+1)^{-2k} [\Gamma(1+2k) G\{r/(r+1)\} - \{\Gamma(1+k)\}^2] \; , \tag{A15}$$

$$v_{r,r+1} = \tfrac{1}{2} \alpha^2 k^{-2} [(r+2)^{-2k} \Gamma(1+2k) G\{r/(r+2)\} + (r+1)^{-k} \{(r+1)^{-k} - 2 \cdot (r+2)^{-k}\} \{\Gamma(1+k)\}^2] \; , \tag{A16}$$

$$v_{r,r+s} = \tfrac{1}{2} \alpha^2 k^{-2} [(r+s+1)^{-2k} \Gamma(1+2k) G\{r/(r+s+1)\} - (r+s)^{-k} \Gamma(1+2k) G\{(r+1)/(r+s)\} + 2(r+1)^{-k} \{(r+s)^{-k} - (r+s+1)^{-k}\} \{\Gamma(1+k)\}^2], \quad s > 2 \; . \tag{A17}$$

When $k \leqslant 0$ the foregoing argument is not valid because the integral in (A11) does not converge. However, the expressions (A15) – (A17) are analytic functions of k for all $k > -\tfrac{1}{2}$ and hence by analytic continuation expressions (A15) – (A17) are valid solutions of the integral representation (A9) – (A10) throughout the domain $-\tfrac{1}{2} < k < \infty$. At the value $k = 0$, the v_{rs} are given by the limits of (A15) – (A17) as $k \to 0$; these limits are well-defined.

The results stated in this Appendix are valid for arbitrary positive integers r and s, though only the cases $r, s = 0, 1, 2$ are required for deriving the asymptotic distributions of PWM estimators.