WATER RESOURCES RESEARCH, VOL. 27, NO. 12, PAGES 3149–3158, DECEMBER 1991

Probability Plot Goodness-of-Fit and Skewness Estimation Procedures for the Pearson Type 3 Distribution

RICHARD W. VOGEL[1] AND DANIEL E. MCMARTIN

Department of Civil Engineering, Tufts University, Medford, Massachusetts

Uniform flood frequency guidelines in the United States currently recommend fitting a Pearson (P3) distribution to the logarithms of annual maximum flood flows. As a result, a plethora of procedures have been recommended for obtaining unbiased plotting positions and unbiased estimates of the skew coefficient and for inverting the cumulative distribution function of a P3 variate. These developments are precisely the ingredients required for the construction of P3 probability plots. Using Monte Carlo simulation, we develop a probability plot correlation coefficient (PPCC) hypothesis test for the P3 distribution. Power studies are performed to evaluate the ability of the test to discriminate among competing distributional alternatives and to enhance our understanding of why the P3 distribution often appears to provide such a good fit to observed flood flow data. A new estimator of the skew coefficient is presented which, unlike the biased and unbiased moment estimators, is unbounded and has significantly lower root mean square error than the moment estimators for highly skewed samples.

INTRODUCTION

Ever since the log Pearson type 3 (LP3) distribution was mandated by the *U.S. Water Resources Council* [1967] for use in fitting sequences of annual peak flood flows in the United States, its application in flood studies has become widespread. *Benson* [1968] reported the conclusions and the reasoning which led to the selection of the LP3 distribution as the "base method" in the United States (with provisions for departures from the "base method" where justified). Similar conclusions were reached elsewhere: for example, *McMahon and Srikanthan* [1981] recommend the use of the LP3 distribution for flood frequency analysis in Australia. There is by no means universal acceptance of the LP3 distribution for modeling flood flows. For example, the three-parameter lognormal distribution is recommended in Ontario [*Sangel and Kallio*, 1977] and the generalized extreme value distribution is recommended in the United Kingdom and Ireland [*Natural Environment Research Council*, 1975]. The debate among statistical hydrologists in the United States regarding the justification for employing the LP3 distribution for modeling flood flow frequencies has intensified due to recent advances in both at-site and regional flood flow frequency analysis (see for example, *Wallis and Wood* [1985] and the discussion of their work by *Beard* [1987] and by *Landwehr et al.* [1987]). It is unlikely that a consensus will ever be reached regarding the selection of a universal parent distribution of flood flows.

One of the primary impediments to reaching a consensus on the selection of an appropriate distribution to model flood flows is the lack of powerful at-site goodness-of-fit procedures for discriminating among alternative hypotheses. For example, in describing the U.S. Water Resources Council (WRC) Work Group study [*U.S. Water Resources Council*, 1967], *Benson* [1968, p. 902] argued that "no single method of testing [alternate hypotheses] . . . was acceptable to all those on the Work Group, and the statistical consultants could not offer a mathematically rigorous method" leading to the conclusion (p. 904) that "there are no rigorous statistical criteria on which to base a choice of method." *Matalas and Wallis* [1973, p. 281] continued this logic when they concluded that "classical statistical tests of goodness of fit are not powerful enough to discriminate among reasonable choices of distributions, and, more often than not, by default the choice is made by fiat."

Since the original WRC study in 1967, a variety of goodness-of-fit procedures have been introduced which allow a rigorous examination of alternative distributional hypotheses. For example, *Wallis* [1988] and *Hosking* [1990] describe the use of L moment diagrams, *Vogel and Kroll* [1989, 1991] describe the application of regional uniform probability plot goodness-of-fit procedures, and *Chowdhury et al.* [1991] compare several regional goodness-of-fit procedures for testing the fit of alternative families of distributions to regional samples. Regional goodness-of-fit procedures are appealing since they can be used to compare the fit of a wide class of distributions to regional samples of flood flow sequences. Unfortunately, most regional goodness-of-fit procedures are based upon at-site tests which usually lack the power to distinguish among very similar distributional hypotheses for the small sample sizes available in most situations. For example, *Kite* [1975] used Monte Carlo simulation experiments to show that extreme events from some commonly used probability density functions are statistically indistinguishable from the normal population for the typical sample sizes and coefficients of variation encountered in practice. Similarly, *Wallis* [1988] performed a simple Monte Carlo experiment which provides very convincing evidence of the inability of at-site goodness-of-fit tests to distinguish among similar distributional hypotheses. Recent work by *Wallis* [1988], *Hosking* [1990], and *Chowdhury et al.* [1991] indicates that L moment diagrams and associated goodness-of-fit procedures may improve our ability to discern distributional differences in regional studies. *D'Agostino and Stephens* [1986] provide a review of goodness-of-fit procedures.

Given the widespread acceptance and usage of the Pearson type 3 (P3) and the LP3 distributions in water resource investigations, it follows that a goodness-of-fit test should be

[1]On leave at Department of Civil and Agricultural Engineering, University of Melbourne, Parkville, Victoria, Australia.

Paper number 91WR02116.
0043-1397/91/91WR-02116$05.00

available. Recently, *Vogel* [1986] and *Vogel and Kroll* [1989] have introduced probability plot correlation coefficient (PPCC) hypothesis tests into the water resources literature. Such tests were originally suggested by *Filliben* [1975] for the normal distribution and later extended to the Gumbel [*Vogel*, 1986; *Kinnison*, 1989], Weibull and uniform [*Vogel and Kroll*, 1989] and the generalized extreme value [*Chowdhury et al.*, 1991] distributional hypotheses. In addition, *D'Agostino and Stephens* [1986] describe PPCC tests for the exponential, logistic and Cauchy distributions.

This study derives and evaluates the use of P3 and LP3 probability plots and associated PPCC goodness-of-fit tests. The PPCC test statistic essentially summarizes the linearity of a probability plot. In addition, the PPCC test statistic may be used to evaluate the probability of a type I error under the null hypothesis of a prespecified distribution. Hence PPCC tests combine a common tool of the practitioner, a probability plot, with a rigorous hypothesis test. Another attractive feature of the at-site hypothesis tests developed here and elsewhere is our ability to extend them to regional hypothesis tests using the procedures suggested by *Vogel and Kroll* [1989, 1991]. Power studies are performed to evaluate the ability of the derived goodness-of-fit tests to discriminate among competing distributional alternatives. These studies enhance our understanding of why so many previous investigators have recommended the LP3 distribution for modeling flood flow and other hydrologic variables. In addition, a new unbounded estimator of the skewness based on probability plots is described which has significantly lower root mean square error than the alternative at-site moment estimators for highly skewed samples.

Pearson Type 3 Distribution

The Pearson type 3 probability density function may be expressed as

$$f(y) = \frac{|\beta|}{\Gamma(\lambda)} [\beta(y - m)]^{\lambda - 1} e^{-\beta(y - m)} \tag{1}$$

where β, λ, and m are parameters. When $\beta > 0$, y has positive skewness leading to $m \le y \le +\infty$. When $\beta < 0$, y has negative skewness leading to $-\infty \le y \le m$. Hence, m is the lower bound of a positively skewed P3 random variable and m is the upper bound of a negatively skewed P3 random variable. The parameters β, λ, and m are related to the first three months of the random variable y as follows:

$$\mu = m + \frac{\lambda}{\beta} \tag{2}$$

$$\sigma^2 = \frac{\lambda}{\beta^2} \tag{3}$$

$$\gamma = \frac{2\beta}{|\beta|\lambda^{1/2}} \tag{4}$$

The (two-parameter) gamma distribution is a special case of the Pearson type 3 distribution when the location parameter m is equal to zero.

Log Pearson Type 3 Distribution

If the random variable $y = \ln(x)$ is distributed P3, then x is distributed LP3. Since the *U.S. Water Resources Council* [1967] recommended the use of the LP3 distribution in flood frequency studies, many studies have documented that this distribution appears to fit a wide range of flood flow sequences. For a more recent review of the recommended procedures for fitting an LP3 distribution see *Interagency Advisory Committee on Water Data (IACWD)* [1982]. *Tasker* [1987] and *Vogel and Kroll* [1989] also recommend the use of the LP3 distribution in low-flow frequency analysis for fitting sequences of annual minimum d day low flows. *Bobee* [1975] shows that the LP3 distribution is more flexible than the P3 distribution. Given the ability of the LP3 distribution to take on many different shapes, it should be no surprise that this distribution appears to fit a wide range of flood flow data.

However, the LP3 distribution does entail certain caveats. For example, in portions of the United States sequences of the logarithms of annual peak flood flow exhibit a negative skewness (see generalized skew map in the work by *IACWD* [1982]) which implies that under the LP3 null hypothesis, significant portions of the United States contain flood flows which are bounded above by an amount $x_{max} = \exp(m)$. *Gilroy* [1972] and *Bobee* [1975] describe this issue and *Reich* [1972] showed that in certain cases, observed streamflows from small samples were actually greater than x_{max}.

Another caveat associated with the LP3 distribution follows from its dependence upon the skew coefficient γ. Since *Wallis et al.* [1974*a*] uncovered the enormous sampling variability associated with small-sample estimates of the skew coefficient, many investigators have warned us to avoid using small-sample estimates of the skew coefficient in hydrologic design and planning.

Probability Plots for the Pearson Type 3 and the Log Pearson Type 3 Distributions

Probability plots are used widely in the field of water resources engineering. Most practitioners would not make engineering decisions regarding the frequency of observations without the use of a graphical display (i.e., a probability plot). In general, a probability plot is a plot of the ordered observations $y_{(i)}$, $i = 1, \cdots, n$ versus the inverse of the cumulative distribution function (cdf) which we term M_i and define as

$$M_i = F^{-1}(\hat{F}(y_{(i)})) \tag{5a}$$

$$M_i = F^{-1}(p_i) \tag{5b}$$

where $p_i = \hat{F}(y_{(i)})$, termed a plotting position, is an estimate of the cdf corresponding to the ith ordered observation.

D'Agostino and Stephens [1986] devote several chapters to a review of the construction of probability plots and associated goodness-of-fit procedures for a wide range of standard distributions. In discussing the gamma distribution, D'Agostino and Stephens argue that this important distribution "does not lend itself immediately to the standard probability plotting techniques." They argue that even with the aid of transformations, it cannot be put in the simple form of a distribution dependent upon a location and scale parameter.

Wilk et al. [1962] describe procedures for constructing probability plots for the P3 distribution. However, their procedures are cumbersome to implement since they require the use of tables or complex algorithms to invert the P3 cdf. In the following sections we describe some recent developments in the water resources literature which do allow us to construct probability plots for the P3 and LP3 distribution using very simple yet accurate procedures.

The Inverse of a Pearson Type 3 Distribution

The cdf of a P3 random variable is defined as

$$F(y) = \int_m^y f(y)\, dy \quad \gamma > 0$$
$$F(y) = \int_{-\infty}^y f(y)\, dy \quad \gamma < 0 \tag{6}$$

which, given the complex form of $f(y)$ in (1), is not easily inverted. Fortunately, many investigators have developed approximate inversion formulae. *Chowdhury and Stedinger* [1991] compare the accuracy of five approximations to K_i, the inverse of a standardized P3 random variable where

$$M_i = \mu + \sigma K_i \tag{7}$$

In the water resources literature, K_i is usually referred to as the frequency factor for the P3 distribution. All of the approximations yield an estimate of K_i in terms of the skew coefficient γ and $\Phi^{-1}(p_i)$, the inverse of a standard normal distribution function. *Chowdhury and Stedinger* [1991] concluded that *Bobee*'s [1979] formula is best for skew coefficients in the range $-2.0 \leq \gamma \leq 5.0$; however, Bobee's formula only provides values of K_i for 15 values of the nonexceedance probability p_i, hence interpolation would be required for constructing a probability plot. *Kirby*'s [1972] algorithm, which was ranked second best, reproduces the correct lower bound of $-2/\gamma$, performs well over a wider range of skews $-9.0 \leq \gamma \leq 9.0$, and does not require interpolation. We employed *Kirby*'s [1972] algorithm which takes the form

$$K_i = A\left\{\max\left[H,\ 1 - \left(\frac{B}{6}\right)^2 + \left(\frac{B}{6}\right)\Phi^{-1}(p_i)\right]^3 - C\right\} \tag{8}$$

where

$$A = \max\ (2/\gamma,\ 0.40)$$
$$B = \gamma - 0.063\ \max\ (0,\ \gamma - 1)^{1.85}$$
$$C = 1 + 0.0144\ \max\ (0,\ \gamma - 2.25)^2$$
$$H = \{C - [(2/\gamma)/A]\}^{1/3}$$

and $\Phi^{-1}(p_i)$ is the inverse of a standard normal distribution function. Kirby's algorithm is a modification of the *Wilson and Hilferty* [1931] transformation which was shown to perform poorly for highly skewed samples by *McMahon and Miller* [1971].

Plotting Positions for the Pearson Type 3 Distribution

Many investigators have advocated the use of quantile-unbiased plotting positions when constructing probability plots (see *Cunnane* [1978] for a review). A quantile-unbiased plotting position defined as

$$p_i = F(E[y_{(i)}]) \tag{9}$$

reproduces the expected value of the order statistics for the distribution of interest. Most such plotting positions advocated in the literature take the form

$$p_i = \frac{i - \alpha}{n + 1 - 2\alpha} \tag{10}$$

For the P3 and LP3 distributions, a quantile-unbiased plotting position depends upon the shape parameter β and hence γ in addition to n and i. *Cunnane* [1978] argues that for most situations encountered in hydrology, the distribution of P3 variates will have shapes which range from a normal to an exponential distribution. Hence Cunnane suggests that for the P3 distribution, α in (10) should range from 0.375 to 0.44, which leads to quantile-unbiased plotting positions for the normal [*Blom*, 1958], and the exponential distributions [*Gringorten*, 1963], respectively. *Sutcliffe et al.* [1975], and *Srikanthan and McMahon* [1981] suggest using $\alpha = 0.4$ as a reasonable compromise for the P3 and LP3 distributions.

More recently, *Xuewu et al.* [1984] and *Nguyen et al.* [1989] developed exact and approximate quantile-unbiased plotting positions for the P3 and LP3 distributions all of which depend upon sample estimates of γ. The approximate formula developed by *Xuewu et al.* [1984] requires the use of tables and associated interpolation schemes. The approximate unbiased plotting position developed by *Nguyen et al.* [1989] takes the form

$$p_i^* = \frac{i - 0.42}{n + 0.3\gamma + 0.05} \tag{11}$$

and is suitable for skews in the range $-3 \leq \gamma \leq 3$ and samples in the range $5 \leq n \leq 100$. *Harter* [1984], *Vogel* [1986] and *Vogel and Kroll* [1989] argue that if one's interest is in exploiting a probability plot to evaluate the goodness of fit of a particular distribution, then the choice among plotting positions is not critical.

We test two plotting positions, *Blom*'s [1958] plotting position which is given by (10) with $\alpha = 0.375$, and the approximately unbiased plotting position given in (11). When *Blom*'s [1958] plotting position is used, $\Phi^{-1}(p_i)$ approximates $E[Z_{(i)}]$, where $Z_{(i)}$ is the ith ordered value from a standard normal distribution.

Estimation of the Skew Coefficient of a Pearson Type 3 Variable

The construction of probability plots for the P3 and LP3 distributions requires estimation of the skew coefficient in order to obtain the frequency factor K_i in (8) and to estimate the unbiased plotting position in (11). An estimate of the skew coefficient is required to construct the necessary probability paper, upon which the observations are to be plotted. The method-of-moment estimator of the skew coefficient is usually defined as

$$G = \frac{1}{s^3}\left[\sum_{i=1}^{n} y_i^3/n - 3\bar{y}s^2 - \bar{y}^3\right] \tag{12}$$

where

$$s = \left[\sum_{i=1}^{n} \bar{y}_i/n - \bar{y}^2\right]^{1/2}$$

$$\bar{y} = \sum_{i=1}^{n} y_i/n$$

Craig [1929] first derived expressions for the expectation, variance and skewness of the estimator G. However, his expressions are only accurate for samples sizes, n, in excess of about 100. *Bowman and Shenton* [1988] provide analytic approximations for the moments and distributional properties of G. *Wallis et al.* [1974*a*] used Monte Carlo simulation to summarize the sampling properties of G for the small samples ($n \leq 90$) of interest in water resource applications. Ever since *Wallis et al.* [1974*a*] exposed the significant small-sample bias associated with the estimator G, a number of investigators have sought to develop [*Bobee and Robitaille*, 1975; *Lettenmaier and Burges*, 1980; *Tasker and Stedinger*, 1986] and compare [*Bobee and Robitaille*, 1977; *Lall and Beard*, 1982] unbiased alternatives. All of the unbiased estimators of γ are simply factors which when multiplied by G in (12) produce an estimator whose expectation is approximately equal to γ. The unbiasing factor recommended by *IACWD* [1982], equal to $[n(n - 1)]^{1/2}/(n - 2)$, was originally derived by *Fisher* [1950] for the normal distribution; hence its use for the P3 distribution is questionable. On the basis of empirical studies, *Hazen* [1930] suggested the use of the unbiasing factor $[1 + 8.5/n]$ for the P3 distribution. More recently, *Bobee and Robitaille* [1975] used *Wallis et al.*'s [1974*a*, *b*] results to derive the approximately unbiased estimator

$$G_u = \bar{G}\left[\left(1 + \frac{6.51}{n} + \frac{20.2}{n^2}\right) + \left(\frac{1.48}{n} + \frac{6.77}{n^2}\right)\bar{G}^2\right] \tag{13}$$

where $\bar{G}$ is the mean of the distribution of the sample skewness for a sample of size n from a P3 distribution (usually estimated using G since only one sample is typically available). In a subsequent comparison of alternative unbiased estimators, *Bobee and Robitaille* [1977] recommended (13) if γ is outside the range of (0.5; 2). Since this study develops hypothesis test procedures for the P3 distribution over a wide range of skews ($-5 \leq \gamma \leq 5$), the unbiased estimator G_u is employed here.

Construction of Probability Plots for the P3 and LP3 Distribution

A probability plot for the P3 distribution is constructed by plotting the ordered observations $y_{(i)}$, $i = 1, \cdots, n$ versus an estimate of the inverse of the fitted distribution M_i given by (7) and (8). To obtain K_i in (8), an estimate of the plotting position p_i and skewness γ, is required. We employed *Blom*'s [1958] plotting position given by (10) with $\alpha = 0.375$ and the approximately unbiased estimator of γ given by G_u in (12) and (13). For an LP3 random variable x, one plots the ordered logarithms $y_{(i)} = \ln(x_{(i)})$ versus the estimated M_i where now μ, σ, and γ are the mean, standard deviation and skewness of the logarithms of x.

PROBABILITY PLOT CORRELATION COEFFICIENT TEST

The probability plot correlation coefficient (PPCC) test is a goodness-of-fit test which measures and evaluates the linearity of the probability plot. If the sample to be tested is actually drawn from the hypothesized distribution, one expects a plot of the ordered observations $y_{(i)}$ versus the expected value of the order statistics, M_i, to appear linear with a correlation coefficient in the neighborhood of one. The PPCC test statistic is

$$r = \frac{\sum_{i=1}^{n} (y_{(i)} - \bar{y})(M_i - \bar{M})}{\left[\sum_{i=1}^{n} (y_{(i)} - \bar{y})^2 \sum_{i=1}^{n} (M_i - \bar{M})^2\right]^{1/2}} \tag{14}$$

where $\bar{y}$ and $\bar{M}$ are sample estimates of the mean of the y_i and the M_i. The PPCC test statistic, r, is an estimate of the population correlation coefficient ρ, between the y_i and the M_i defined as

$$\rho = \frac{\text{Cov}(y_i, M_i)}{[\text{Var}(y_i)\ \text{Var}(M_i)]^{1/2}} \tag{15}$$

which can be combined with (7) and simplified to yield

$$\rho = \frac{\text{Cov}(y_i, K_i)}{[\text{Var}(y_i)\ \text{Var}(K_i)]^{1/2}} \tag{16}$$

Since K_i is only a function of the plotting position p_i and the skewness γ, the PPCC test statistic only depends upon the observations y_i, the assumed plotting position p_i and the estimated skew coefficient. One attractive property of the test is that the test statistic in (16) does not depend upon the location and scale parameters, μ and σ.

Percentiles of the Probability Plot Correlation Coefficient Test Statistic

The basic idea of this study is to use Monte Carlo simulation to obtain the sampling distribution of the PPCC test statistic, r, under the null hypothesis that the observations arise from a P3 distribution. Critical values of the PPCC test statistic were obtained as follows:

1. Sequences of P3 random variables with $\mu = 1$ and $\sigma = 0.25$ were generated of length n = 10, 15, 25, 50, 75, 100, 200, and 500 with skew coefficients γ = 0.01, 1, 2, 3, 4, and 5, using the International Mathematical Subroutine Library (IMSL) subroutine RNGAM. A total of 100,000 sequences were generated for each of these 48 combinations of n and γ.

2. For each of the previously generated samples of P3 random variables the corresponding values of M_i are obtained from (7) and (8). Since the M_i are a function of both γ and p_i, three separate methods for estimating the M_i are used to evaluate the influence of both the skew coefficient and the plotting position on the resulting sampling properties of the test statistic r. Method 1 employs the unbiased estimator of the skew coefficient G_u, given by (12) and (13) and *Blom*'s [1958] plotting position (10) with $\alpha = 0.375$. Method 2 employs the true skew γ and Blom's plotting position. Method 3 employs the true skew γ and the unbiased plotting position p_i^* given in (11).

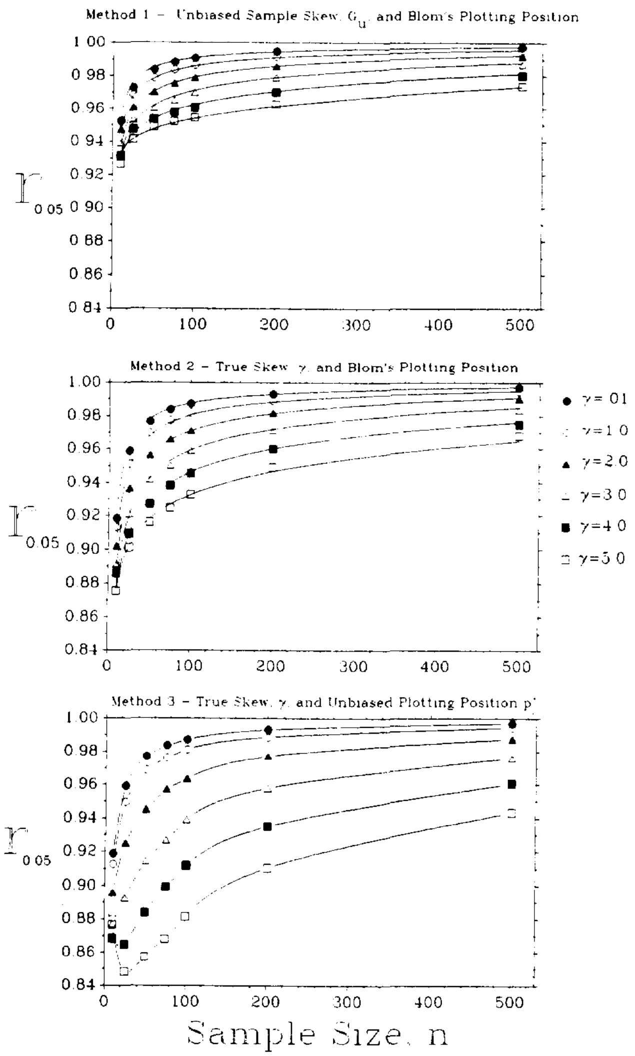

Fig. 1. The 5% level Pearson type 3 probability plot correlation coefficient test statistic $r_{0.05}$ as a function of sample size, n, and skew coefficient, γ, using three different methods to construct the hypothesis test.

3. The PPCC test statistic, r, for each sample is obtained using (14). Percentage points of the distribution of r were obtained using the empirical sampling procedure

$$r_q = r_{(100,000q)} \tag{17}$$

where r_q denotes the qth quantile of the distribution of r, and $r_{(100,000q)}$ denotes the 100,000q largest observation in the sequence of 100,000 generated values of r. Essentially q is the prespecified type I error (or significance level) associated with the P3 or LP3 hypothesis test.

Figure 1 summarizes the values of the 5% level test statistic, $r_{0.05}$ (the value of r which is exceeded 95% of the time) when observations arise from a P3 distribution, using the three different methods described above. For all three methods $r_{0.05}$ tends to increase with sample size and to decrease with skew. The 5% level test statistic is uniformly higher for all values of n and γ when one employs the approximately unbiased sample skew G_u instead of γ to construct the probability plot and estimate r. Apparently, when one uses the sample skew to construct the probability paper (or estimate the M_i), samples appear more linear than when the true skew is employed. The estimated sample skew acts to adjust the probability paper to make the sample, when plotted, appear more linear than it would if the true skew had been used to construct the plot. Methods 1 and 2 lead to significantly different percentage points of the distribution of r, indicating the importance of the skew coefficient in constructing P3 probability plots and associated hypothesis tests.

A comparison of methods 2 and 3 in Figure 1 shows that $r_{0.05}$ is insensitive to the plotting position for skews in the range (0, 2). The approximately unbiased plotting position p_i^* is only valid for $\gamma \le |3|$; hence the results for large skews corresponding to method 3 in Figure 1 are questionable. Since p_i^* is only valid for small skews and since $r_{0.05}$ is insensitive to the plotting position in this region, we no longer consider the use of the unbiased plotting position p_i^*.

In Figure 1 the symbols represent the results of the Monte Carlo simulations and the solid lines represent equations which were fit to these points. For method 1 in Figure 1 the 5% level PPCC test statistic can be approximated using

$$r_{0.05} = \exp\,[2.97 - 0.0307G_u^2 - 0.000796n]n^{(0.103G_u - 0.652)} \tag{18}$$

as long as $|G_u| \le 5$. The 5% level PPCC test statistic corresponding to method 2 in Figure 1 can be approximated using

$$r_{0.05} = \exp\,[3.77 - 0.0290\gamma^2 - 0.000670n]n^{(0.105\gamma - 0.758)} \tag{19}$$

as long as $|\gamma| \le 5$.

Power Studies

In this section we investigate the power of the q = 5% level PPCC tests developed in the previous section. Power is defined as the probability that the 5% level P3 hypothesis test will be able to detect a sample which arises from an alternative population. More formally, power is defined as the probability that the P3 null hypothesis is rejected when it is false. The power of a P3 PPCC test was investigated by generating 20,000 samples (μ = 1 and σ = 0.25) of length n = 10, 25, and 100, from alternative hypotheses including the normal, lognormal, uniform and Gumbel distributions. For each generated sample from the alternative distribution, the PPCC test statistic, r, was computed from (14) using (7) and (8) with Blom's plotting position and assumed skew coefficients γ = 0, 1, 2, and 3. The computed test statistic was compared with $r_{0.05}$ computed from (19) using the four assumed skew values. Power is defined as the percentage of the 20,000 samples which led to values of $r < r_{0.05}$, for each alternative probability. Figure 2 summarizes the power (in percent) of a 5% significance level P3 PPCC test against the normal, lognormal, Gumbel and uniform distributional hypotheses. As anticipated, the power generally increases with sample size. While the 5% level P3 PPCC test has modest power against the uniform alternative regardless of the assumed skew, it appears unable to reject alternatives which are similar to a P3 distribution. For the normal, lognormal and Gumbel alternatives, one observes that when the assumed skew used to construct the test is close to the

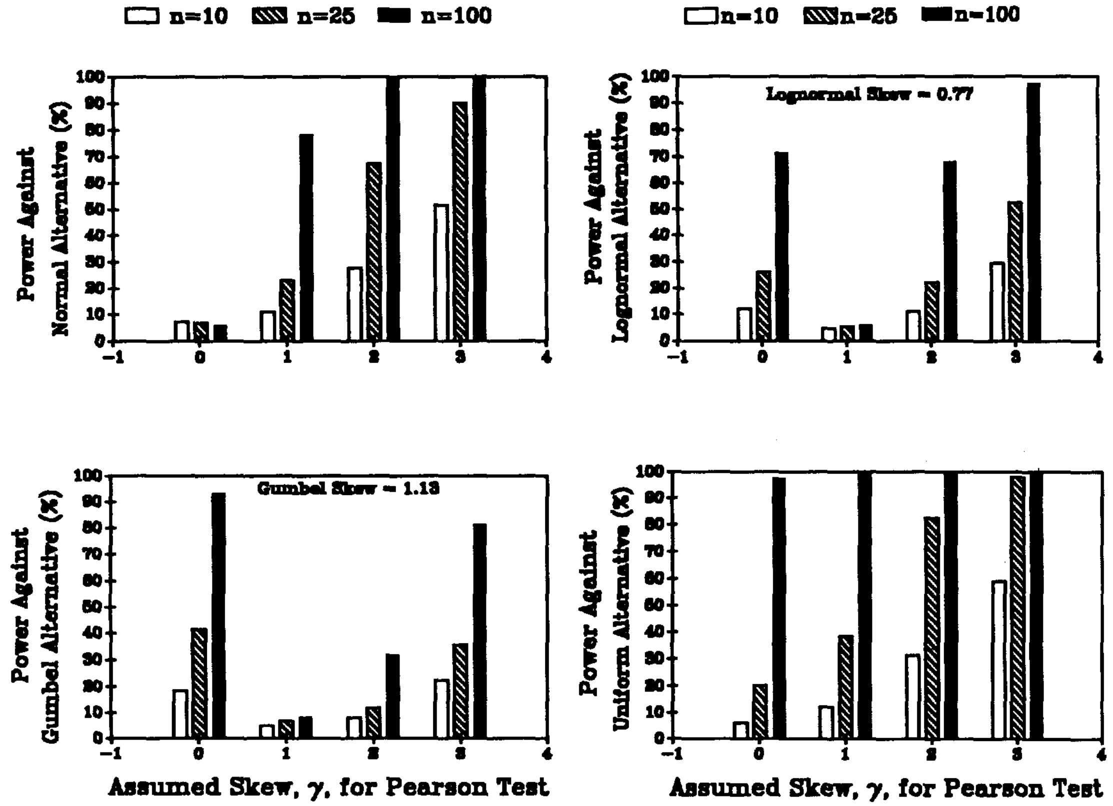

Fig. 2. Power of a 5% level Pearson type 3 PPCC test against four alternative distributional assumptions using Bloms plotting position and an assumed skew coefficient.

population skew of these alternatives, the power is extremely low, indicating the inability of this test to discriminate among distributional alternatives which are similar in shape to the P3 alternative. The more the assumed skew differs from the population skew of the distributional alternative tested, the greater the power of the P3 PPCC test.

Figure 2 implies that a 5% level P3 PPCC test could discriminate against similar distributional alternatives only if the assumed skew used to construct the test differs from the population skew of the alternatives. Hence, if one assumes a fixed skew coefficient for a particular region, the P3 PPCC tests described here should be able to detect samples which have population skews which differ from the assumed regional skew. Similarly, *Chowdhury et al.* [1991] performed power studies to evaluate PPCC tests for the generalized extreme value hypothesis assuming fixed (or regional) values of the shape parameter. They show that PPCC tests can be useful for determining whether or not the at-site estimate of the shape parameter differs from the assumed regional shape parameter.

In practice, one must use a sample estimate of γ to perform an at-site P3 PPCC hypothesis test. Figure 3 summarizes the power of a $q = 5\%$ level P3 PPCC test against the same four alternatives when method 1 (Blom's plotting position and unbiased estimator G_u) is used to estimate the test statistic and (18) is used with G_u to estimate $r_{0.05}$. Here 20,000 samples of length $n = 10$, 25, 100, and 500 are generated from the same four alternate hypotheses as in Figure 2. The power of an at-site 5% level P3 PPCC test against the normal, lognormal and Gumbel alternatives is remarkably low. An at-site test is only able to discriminate against the uniform alternative, and even then only for relatively large sample sizes. As we showed earlier in Figure 1, when the sample skew G_u is used to construct a P3 probability plot, the samples appear more linear than they should, leading to very low power.

A PROBABILITY PLOT CORRELATION ESTIMATOR OF THE SKEW COEFFICIENT

Ever since *Wallis et al.* [1974*a*] discovered the remarkably large small-sample bias and variance associated with the sample skew estimator G, investigators have discouraged the use of such estimators particularly for the estima-

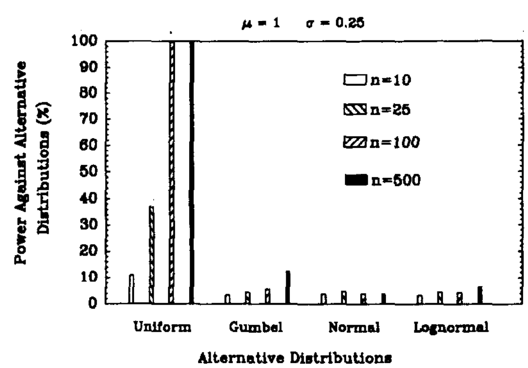

Fig. 3. Power of a 5% level Pearson type 3 PPCC test against four alternatives using Blom's plotting position and an unbiased estimator of the skew coefficient G_u.

TABLE 1. A Comparison of the Bias and Root Mean Square Error (RMSE) of the Bounded Sample Skew Estimators G and G_u With the Unbounded Probability Plot Correlation Skew Estimator G_r

	Bias $[\theta] = \gamma - E[\theta]$			RMSE $[\theta] = E[(\theta - \gamma)^2]$			
Skew γ	$\theta = G$	$\theta = G_u$	$\theta = G_r$	$\theta = G$	$\theta = G_u$	$\theta = G_r$	Percent of Rejects
			Sample Size n = 10				
0.25	0.12 (0.12)*	−0.02	0.00	0.59 (0.59)	1.24	1.15	0
1	0.49 (0.49)*	−0.11	−0.04	0.77 (0.77)	1.39	1.27	0
3	1.67 (1.66)	−0.33	−0.27	1.78 (1.77)	2.07	1.81	0.2
5	3.20 (3.18)	0.03	0.08	3.25 (3.23)	2.29	1.87	4.6†
			Sample Size n = 50				
0.25	0.03 (0.03)	0.00	0.00	0.34 (0.34)	0.39	0.39	0
1	0.14 (0.14)	−0.01	−0.03	0.45 (0.46)	0.54	0.45	0
3	0.72 (0.70)	−0.10	−0.32	1.01 (1.06)	1.35	1.06	0.6
5	1.84 (1.66)	0.20	−0.28	2.00 (1.92)	1.75	1.24	9.0†

*Values in parentheses are those reported by *Wallis et al.* [1974*b*].

†Percentage of samples rejected because $|G_r| > 9$. A total of 20,000 were generated, not including the rejected samples, except for the case $\gamma = 5$, when 30,000 samples were generated.

tion of flood flow design quantiles [*Bobee*, 1975; *Kite*, 1975]. In addition, *Kirby* [1974] showed that G is bounded so that

$$|G| \leq \frac{(n-2)}{(n-1)^{1/2}} \tag{20}$$

Interestingly, this bound does not depend upon γ; hence the bias associated with G (Bias$[G] = \gamma - E[G]$) increases rather dramatically for large γ and small n. Due to the potentially large uncertainty associated with the sample skew estimator G, *IACWD* [1982] recommend the use of a weighted skew estimator which is the weighted average of the at-site skew estimator G, and a generalized skew coefficient obtained from a map. *Chowdhury and Stedinger* [1991] and *Tasker and Stedinger* [1986] provide a review of the use of such weighted skew estimators.

The significant attention in the literature given to estimators of γ for the P3 distribution, in combination with our observations in Figure 1 regarding the strong association between the PPCC test statistic r and γ, led us to investigate the properties of a probability plot correlation estimator of the skew coefficient which we term G_r. The estimator G_r is defined as that value of γ which maximizes the PPCC test statistic r. Again the PPCC test statistic r is defined by (14) with the M_i estimated using (7), (8) and (10) with $\alpha = 0.375$. Table 1 and Figure 4 summarize Monte Carlo experiments which were performed to compare the bias and root mean square error (RMSE) associated with the three estimators G, G_u, and G_r for sample sizes $n = 10$ and 50, and for skews $\gamma = 0.25$, 1, 3, and 5. Samples which led to estimates of G_r outside the interval $[-9, 9]$ were rejected since Kirby's inversion formula is only accurate inside that interval. A total of 20,000 samples were generated for each combination of n and γ, not including the rejected samples. For comparison, the bias and RMSE associated with G obtained from *Wallis et al.* [1974*a*, *b*] are provided in parentheses in Table 1. For the cases when $\gamma = 3$ and 5, Table 1 documents that a small percentage of samples had to be rejected; hence our estimates of the bias and RMSE associated with G differ slightly from those of *Wallis et al.* [1974*a*, *b*] in those situations.

Overall one observes from Table 1 and Figure 4 that G_r has lower RMSE than G_u for all cases considered and that both G_r and G_u are nearly unbiased when compared with the biased estimator G. The dramatic increase in RMSE associated with G for large values of γ results from the large bias introduced by its bound given in (20). Since G_r is unbounded (except for the bounds introduced by Kirby's approximation to K_i) it is an especially attractive estimator of γ for highly skewed samples as evidenced by its low RMSE for $\gamma = 5$. For example, G_r may be an attractive estimator for samples which appear to exhibit outliers. Furthermore, in terms of RMSE, G_r is always a more attractive estimator than its unbiased competitor G_u.

Figure 5 displays the relationship between the probability plot correlation coefficient, r, and the assumed skew coefficient for four different population skews $\gamma = -1$, 1, 3, and 6. Each curve in Figure 5 is based on an individual sample generated from a P3 distribution with population skew γ and $n = 50$. The particular samples used in the construction of Figure 5 had estimates of $G_r \approx \gamma$, that is, the maximum value of r for each sample is based on an assumed skew coefficient which is approximately equal to the population skew γ. Figure 5 shows how sensitive the linearity of a P3 probability plot is to the value of the assumed skew.

CONCLUSIONS

On the basis of power studies, *Filliben* [1975] showed that a probability plot correlation coefficient (PPCC) test for the normal distribution compares favorably with seven other standard hypothesis tests. As a result, PPCC hypothesis tests have been extended to a variety of other two-parameter and three-parameter distributional hypotheses (see introduc-

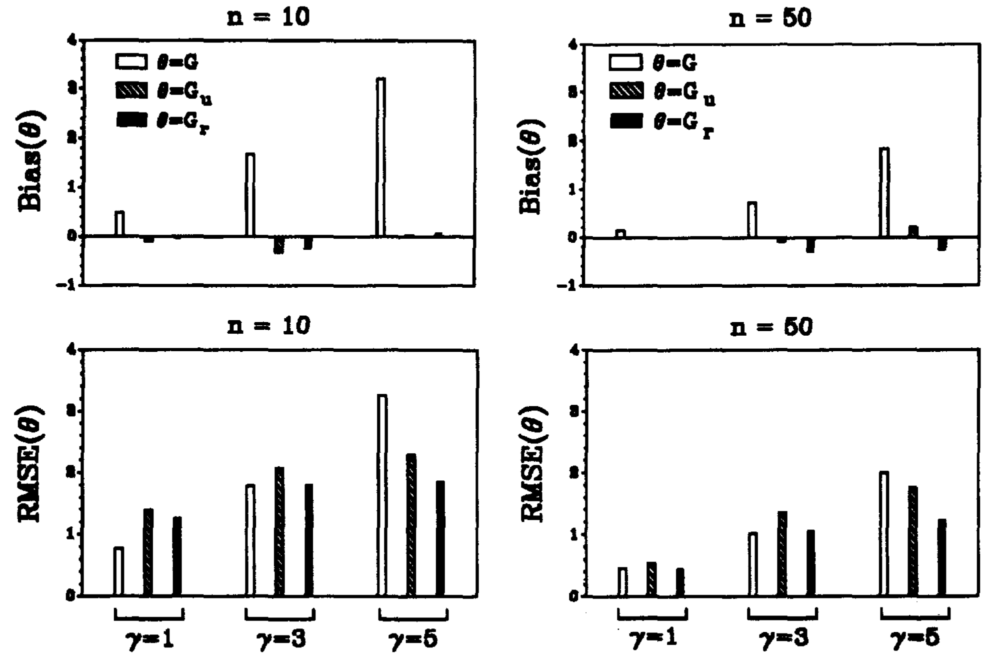

Fig. 4. A comparison of the bias and root mean square error (RMSE) associated with the biased moment skew estimator G, the unbiased moment skew estimator G_u, and the unbounded probability plot correlation skew estimator G_r.

tion for citations). Since the Pearson (P3) and log Pearson type 3 (LP3) distributions do not exhibit a fixed shape parameter, the construction of a probability plot and associated hypothesis test requires an estimate of the skew coefficient to invert the cumulative distribution function. Using Monte Carlo simulation we developed a relationship between a measure of the linearity of the probability plot, the skewness and the sample size for P3 samples. On the basis of that relationship we reached the following conclusions:

1. Probability plots for the P3 and LP3 distribution based on an estimate of the sample skew will, in general, appear more linear then they should. Essentially, the estimated sample skew acts to adjust the probability scale to make the sample, when plotted, appear more linear than it would if the true skew had been used to construct the plot. Since the true skew is never known in practice, and use of the sample skew to construct P3 or LP3 probability paper leads to plots which appear more linear than they should, probability plotting procedures for these distributions are not recommended for evaluating whether an individual sample arises from an alternative distribution. In effect, samples from any distribution which resembles the P3 (or LP3) alternative will appear P3 (or LP3) when the sample skew is used to construct the plot. This conclusion is supported by the relationship between $r_{0.05}$, γ and n, given in Figure 1 and the power studies summarized in Figures 2 and 3.

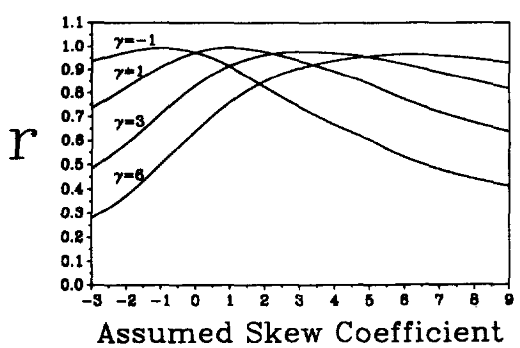

Fig. 5. Relationship between the probability plot correlation coefficient, r, and the assumed skew coefficient used to construct the probability plot over a range of population skews, γ.

2. The power studies summarized in Figure 2 document that PPCC hypothesis testing procedures could be useful in regional studies which seek to determine whether an at-site estimate of the skew coefficient is significantly different from an assumed regional skew coefficient. Similarly, *Chowdhury et al.* [1991] found PPCC tests compared favorably with other goodness-of-fit procedures for testing whether or not the shape parameter of at-site data differs from the assumed regional shape parameter associated with the regional generalized extreme value hypothesis.

3. In terms of the linearity of a P3 or LP3 probability plot, there does not appear to be a significant difference between the use of a biased or an unbiased plotting position, even when the population skew is known. We have not investigated the impact of using the sample skew to estimate an unbiased plotting position for the P3 distribution, and we do not recommend future investigations along these lines.

4. If a sample which is assumed to be P3 or LP3 exhibits an outlier and/or an estimate of G near the bound given in (20), then we recommend the use of our unbounded probability plot correlation skew estimator G_r instead of either G or G_u.

Construction of probability paper for the P3 and LP3 distributions requires an estimate of the skew coefficient; hence general probability paper is unavailable as it is for the uniform, exponential, normal, lognormal, Weibull, and Gumbel distributions, all of which can be transformed to exhibit a fixed shape parameter. Since probability paper for the P3 and LP3 distribution is generally unavailable, practitioners often construct probability plots for the P3 and LP3 distribution using other readily available probability paper for the normal, lognormal or Gumbel distributions. Such procedures generally lead to fitted P3 or LP3 curves which, again, appear to fit the data better than the two-parameter alternatives. If goodness of fit is the only criterion by which to select a specified distribution, then distributions with three or more parameters will usually provide a much better fit than two-parameter alternatives.

In spite of the flexibility of the P3 and LP3 distributions to fit a wide range of at-site flood samples, recent studies have indicated that use of regional procedures such as the index flood method can lead to significantly lower root mean square errors associated with design quantiles than at-site procedures [*Potter and Lettenmaier*, 1990]. Hence future work related to hypothesis testing for P3, LP3 and other distributions in hydrology should focus on the primary assumption associated with such index flood type procedures, that a normalized regional index flood distribution provides an adequate description for the region of interest.

Acknowledgments. The authors gratefully acknowledge the financial support of the following federal agencies who jointly sponsored this research: the U.S. Geological Survey, the U.S. Army Corps of Engineers Hydrologic Engineering Center, and the Federal Emergency Management Agency. This manuscript is the project completion report to the Bulletin 17B Work Group of the Hydrology Subcommittee of the Interagency Advisory Committee on Water Data chaired by Wilbert O. Thomas, Jr., from the U.S. Geological Survey.

REFERENCES

Beard, L. R., Discussion on "Relative accuracy of log Pearson III procedures," *J. of Hydraul. Eng., 113*(7), 1205–1206, 1987.

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

Blom, G., *Statistical Estimates and Transformed Beta Variables*, pp. 143–146, John Wiley, New York, 1958.

Bobee, B. B., The log Pearson type 3 distribution and its application in hydrology, *Water Resour. Res., 11*(5), 681–689, 1975.

Bobee, B. B., Comment on "Fitting the Pearson type 3 distribution in practice," by J. Buckett and F. R. Oliver, *Water Resour. Res., 15*(3), 730, 1979.

Bobee, B. B., and R. Robitaille, Correction of bias in the estimation of the coefficient of skewness, *Water Resour. Res., 11*(6), 851–854, 1975.

Bobee, B. B., and R. Robitaille, The use of the Pearson type 3 and log Pearson type 3 distributions revisited, *Water Resour. Res., 13*(2), 427–443, 1977.

Bowman, K. O., and L. R. Shenton, *Properties of Estimators for the Gamma Distribution*, 268 pp., Marcel Dekker, New York, 1988.

Chowdhury, J. U., and J. R. Stedinger, Confidence intervals for design flood with estimated skew coefficient, *J. Hydraul. Eng., 117*(7), 811–831, 1991.

Chowdhury, J. U., J. R. Stedinger, and L. Lu, Goodness-of-fit tests for regional generalized extreme value flood distributions, *Water Resour. Res., 27*(7), 1765–1776, 1991.

Craig, C. C., Sampling when the parent population is of Pearson's type III, *Biometrika, 21*, 287–293, 1929.

Cunnane, C., Unbiased plotting positions—A review, *J. Hydrol., 37*, 205–222, 1978.

D'Agostino, R. B., and M. A. Stephens, *Goodness-of-Fit Techniques*, 560 pp., Marcel Dekker, New York, 1986.

Filliben, J. J., The probability plot correlation coefficient test for normality, *Technometrics, 17*(1), 111–117, 1975.

Fisher, R. A., Moments and product moments of sampling distributions, in *Contributions to Mathematical Statistics*, John Wiley, New York, 1950.

Gilroy, E. J., The upper bound of a log Pearson type 3 variable with negatively skewed logarithms, *U.S. Geol. Surv. Prof. Pap., 800-B*, 273–275, 1972.

Gringorten, I. I., A plotting rule for extreme probability paper, *J. Geophys. Res., 68*(3), 813–814, 1963.

Harter, H. L., Another look at plotting positions, *Commun. Stat. Theor. Methods, 13*(13), 1613–1633, 1984.

Hazen, A., *Flood Flows*, p. 46, John Wiley, New York, 1930.

Hosking, J. R. M., *L*-moments: Analysis and estimation of distributions using linear combinations of order statistics, *J. R. Stat. Soc., Ser. B, 52*(2), 105–124, 1990.

Interagency Advisory Committee on Water Data (IACWD), Guidelines for determining flood flow frequency, *Bull. 17B*, Hydrol. Subcomm., Off. of Water Data Collect., U.S. Geol. Surv., Reston, Va., 1982.

Kinnison, R., Correlation coefficient goodness-of-fit test for the extreme-value distribution, *Am. Stat., 43*(2), 98–100, 1989.

Kirby, W., Computer-oriented Wilson-Hilferty transformation that preserves the first three moments and the lower bound of the Pearson type 3 distribution, *Water Resour. Res., 8*(5), 1251–1254, 1972.

Kirby, W., Algebraic boundedness of sample statistics, *Water Resour. Res., 10*(2), 220–222, 1974.

Kite, G. W., Confidence limits for design events, *Water Resour. Res., 11*(1), 46–53, 1975.

Lall, U., and L. R. Beard, Estimation of Pearson type 3 moments, *Water Resour. Res., 18*(5), 1563–1569, 1982.

Landwehr, J. M., G. D. Tasker, and R. D. Jarrett, Discussion on "Relative accuracy of log Pearson III procedures," *J. Hydraul. Eng., 113*(7), 1206–1210, 1987.

Lettenmaier, D. P., and S. J. Burges, Correction for bias in estimation of the standard deviation and coefficient of skewness of the log Pearson 3 distribution, *Water Resour. Res., 16*(4), 762–766, 1980.

Matalas, N. C., and J. R. Wallis, Eureka! It fits a Pearson type 3 distribution, *Water Resour. Res., 9*(2), 281–289, 1973.

McMahon, T. A., and A. J. Miller, Application of the Thomas and Fiering model to skewed hydrologic data, *Water Resour. Res., 7*(5), 1338–1340, 1971.

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

Natural Environment Research Council (NERC), Flood studies report, London, 1975.

Nguyen, V. T. V., N. In-na, and B. Bobee, New plotting-position formula for Pearson type-III distribution, *J. Hydraul. Eng., 115*(6), 709–730, 1989.

Potter, K. W., and D. P. Lettenmaier, A comparison of regional flood frequency estimation methods using a resampling method, *Water Resour. Res., 26*(3), 415–424, 1990.

Reich, B. M., Log-Pearson type 3 and Gumbel analysis of floods, in *Second International Symposium in Hydrology*, pp. 290–301, Water Resources Publications, Fort Collins, Colo., 1972.

Sangel, B. P., and R. W. Kallio, Magnitude and frequency of floods in southern Ontario, *Tech. Bull., 99*, Fish. and Environ., Ottawa, 1977.

Srikanthan, R., and T. A. McMahon, Log Pearson III distribution—An empirically-derived plotting position, *J. Hydrol., 52*, 161–163, 1981.

Sutcliffe, J. V., et al., Flood studies report, vol. I, Hydrological studies, Nat. Environ. Res. Counc., London, 1975.

Tasker, G. D., A comparison of methods for estimating low flow characteristics of streams, *Water Resour. Bull., 23*(6), 1077–1083, 1987.

Tasker, G. D., and J. R. Stedinger, Regional skew with weighted LS regression, *J. Water Resour. Plann. Manage., 112*(2), 225–237, 1986.

U.S. Water Resources Council, Guidelines for determining flood flow frequency, *Bull. 15*, Hydrol. Comm., Washington, D. C., 1967.

Vogel, R. M., The probability plot correlation coefficient test for the normal, lognormal and Gumbel distributional hypothesis, *Water Resour. Res., 22*(4), 587–590, 1986.

Vogel, R. M., and C. N. Kroll, Low-flow frequency analysis using probability-plot correlation coefficients, *J. Water Resour. Plann. Manage., 115*(3), 338–357, 1989.

Vogel, R. M., and C. N. Kroll, The value of streamflow record augmentation procedures in floodflow and lowflow frequency analysis, *J. Hydrol.*, *125*, 259–276, 1991.

Wallis, J. R., Catastrophes, computing and containment: Living in our restless habitat, *Speculations Sci. Technol.*, *11*(4), 294–315, 1988.

Wallis, J. R., and E. F. Wood, Relative accuracy of log Pearson III procedures, *J. Hydraul. Eng.*, *111*(7), 1043–1056, 1985.

Wallis, J. R., N. C. Matalas, and J. R. Slack, Just a moment, *Water Resour. Res.*, *10*(2), 211–219, 1974*a*.

Wallis, J. R., N. C. Matalas, and J. R. Slack, Just a moment!, appendix, *NTIS PB-231 816*, Natl. Tech. Inf. Serv., Springfield, Va., 1974*b*.

Wilk, M. B., R. Gnanadesikan, and M. J. Huyett, Probability plots for the gamma distribution, *Technometrics*, *4*, 1–15, 1962.

Wilson, E. B., and M. M. Hilferty, The distribution of chi-square, *Proc. Natl. Acad. Sci. U. S. A.*, *17*, 684–688, 1931.

Xuewu, J., D. Jing, H. W. Shen, and J. D. Salas, Plotting positions for Pearson type III distribution, *J. Hydrol.*, *74*, 1–29, 1984.

D. E. McMartin, Department of Civil Engineering, Tufts University, Medford, MA 02155.

R. M. Vogel, Department of Civil and Agricultural Engineering, University of Melbourne, Parkville, Victoria 3052, Australia.

(Received May 31, 1991;
revised August 6, 1991;
accepted August 12, 1991.)