Journal of Hydroscience and Hudraulic Engineering
Vol. 4, No. 2, November, 1986, pp.79-95.

A NEW APPROACH TO PARAMETER ESTIMATIONS OF GAMMA-TYPE DISTRIBUTIONS

By

Kiyoshi Hoshi

Civil Engineering Research Institute
Hokkaido Development Bureau
Sapporo 062, Japan

and

Uruya Leeyavanija

Division of Water Resources Engineering
Asian Institute of Technology
Bangkok 10501, Thailand

SYNOPSIS

The quantile and sextile methods are developed as new parameter estimation techniques for the gamma-type distributions. The applicability of the proposed methods is examined by choosing the Pearson type 3 (P3) and log Pearson type 3 (LP3) distributions as representatives of gamma-type distributions. Extensive Monte Carlo experiments are conducted to evaluate the performance of six fitting procedures for different sample sizes and different combinations of population statistics.

Of hydrologic concern is the accuracy with which the fitted distributions give rise to good estimates of various quantiles (e. g., T-year design magnitudes). It follows from the root mean square error (rmse) criterion of quantile estimates that for less variable P3 distributions, there is little difference between performances of five fitting methods examined herein. The sextile method as well as the quantile methods coupling with moment and maximum likelihood estimators dominate the widely used method of moments, when the observations are highly variable and skewed. In the application to the LP3 distribution, the method of quantiles combined with the method using the first two moments in real space yields the most accurate estimates of high extreme quantiles with probability levels larger than 0.99 for highly variable and skewed data sequences.

INTRODUCTION

Frequency analysis of hydrologic variables has a long history in the phases associated with hydrology for water resources planning, design and management. Although numerous probability density functions have been used to describe sequences of hydrologic variables, no "a priori" knowledge exists to support the assumption that the specific distribution is indeed the underlying distribution. Classical statistical tests of goodness of fit are not powerful enough to discriminate among reasonable choice of distributions. Generally, the adoption of distributions would be determined from their computational simplicity.

In most cases, frequency distributions of hydrologic data are positively skewed. To meet such practical requirements, a versatile family of gamma-type disitributions has been widely used to describe the frequency of observed hydrologic data. The gamma-type distributions are classified into two-parameter gamma (P2), Pearson type 3 (P3), log Pearson type 3 (LP3), and generalized gamma (Stacy and Mihram, 1965; Prentice, 1974;

Hoshi and Yamaoka, 1980) distributions. For example, Benson (1968) and the U. S. Water Resources Council (1977) recommended the use of a log Pearson type 3 distribution as a base technique in flood frequency methods. For methods of fitting gamma-type distributions, much is known about the methods of moments and maximum likelihood. Matalas and Wallis (1973), Bobée (1973), Bobée and Robitaille (1975), Bobée and Robitaille (1977), and Buckett and Oliver (1977) have examined the statistical precision and properties of the two methods of fitting the P3 distribution to observed or generated data. A parameter estimating technique for the LP3 distribution adopted by the U. S. Water Resources Council (WRC) is the method of moments by which a P3 distribution is fitted to the logarithmically transformed data sequences. As an alternative method to the U. S. WRC technique, Bobée (1975) and Hoshi and Burges (1981) developed a parameter estimation scheme of the LP3 distribution which could preserve the first three moments in real space. Condie (1977) applied the maximum likelihood estimator (MLE) of the LP3 distribution to the data in the Canadian rivers. Nozdryn-Plotnicki and Watt (1979) examined the performance of the three different fitting techniques for the LP3 distribution, i. e., the method of maximum likelihood, two methods of moments in real and log spaces. The comparison of additional LP3 estimation procedures was also made by Rao (1980) and Phien and Hira (1983) who used the method of mixed moments in real and log spaces.

In the past years considerable arguments existed concerning the relative merits of uses of regional skew maps in the log domain, advocated by the WRC (Landwehr et al., 1978; Wallis and Wood, 1985). Common to the discussions raised by the above studies is the limited properties of a LP3 distribution that if the skew coefficient of the logarithms of data is negative, the scale parameter of this distribution must be negative and therefore the distribution has an upper bound. It is well known that a large portion of flood data exhibits negative skews in log space. Thus the probability of the LP3 having an upper bound becomes quite large, when using the WRC procedures on such data.

Recently, Stedinger (1980) and Hoshi et al. (1984) used Monte Carlo techniques to investigate sampling properties of lower and upper quantile estimates using several fitting methods for a three-parameter log normal (LN3) distribution. The results show that Stedinger's quantile-lower bound methods perform better for highly skewed LN3 distributions. Jenkinson (1969) developed the method of sextiles for a generalized extreme value distribution which has the three parameters. The main features of this method are such that the sextile means have smaller sampling errors than individual sample observations and that the effect of outliers is significantly removed.

To date no attempts have been made to apply the quantile and sextile methods to parameter estimations of gamma-type distributions. The present study is intended to develop these two methods to be capable of fitting the Pearson type 3 (P3) and log Pearson type 3 (LP3) distributions. Monte Carlo simulations are carried out to assess the performances of the quantile and sextile methods by comparing with presently available procedures.

PEARSON TYPE 3 DISTRIBUTION AND PARAMETER ESTIMATION

The density function of the Pearson type 3 (P3) distribution used in this study is

$$f(y) = [(y-c)/a]^{b-1} \exp[-(y-c)/a]/[a\Gamma(b)] \tag{1}$$

where a, b and c are scale, shape and location parameters, respectively. The parameters a and b are positive, and $\Gamma(b)$ denotes the gamma function. The cumulative distribution function of P3 is defined by

$$p = F(y) = \int_0^w f(w)\,dw \tag{2}$$

where

$$w = (y-c)/a \tag{3}$$

$$f(w) = \exp(-w)w^{b-1}/\Gamma(b) \tag{4}$$

p is a non-exceedance probability of a P3-distributed variate y, and $F(y)$ is the cumulative distribution function. Equation 4 is the standard gamma distribution and w is a standard gamma quantile dependent on a probability level and parameter b. Several procedures employed for estimating the distribution parameters are discussed below.

Method of Moments

The mean μ_y, variance σ_y^2, and coefficient of skewness γ_y of a random variable y with a P3 distribution are given by

$$\mu_y = ab + c \quad ; \quad \sigma_y^2 = a^2 b \quad ; \quad \gamma_y = 2/b^{1/2} \tag{5}$$

Moment estimates of a, b and c for the P3 distribution can be obtained by replacing the population statistics μ_y, σ_y^2 and γ_y by the sample mean, unbiased variance and an estimate of the skew coefficient. Let $y_1, y_2, \cdots\cdots, y_N$ be a sample of N independently distributed random variables from a P3 distribution. The unbiased sample variance is computed as

$$\hat{\sigma}_y^2 = [N/(N-1)]s_y^2 \tag{6}$$

where

$$s_y^2 = \sum_{i=1}^{N} (y_i - \bar{y})^2/N \quad ; \quad \bar{y} = \sum_{i=1}^{N} y_i/N \tag{7}$$

The sample coefficient of skewness is computed as

$$S_{ky} = \sum_{i=1}^{N} (y_i - \bar{y})^3/[N s_y^3] \tag{8}$$

Two methods of moments are considered. The first procedure makes use of $\bar{y}$, $\hat{\sigma}_y^2$, and the corrected estimate $\hat{\gamma}_{y1}$ of the sample skew S_{ky} given by

$$\hat{\gamma}_{y1} = \sqrt{[N(N-1)]}\; S_{ky}/(N-2) \tag{9}$$

It is well known that, in general, $\hat{\gamma}_{y1}$ is not an unbiased estimate. Use of $\bar{y}$, $\hat{\sigma}_y^2$ and $\hat{\gamma}_{y1}$ yields moment estimators of parameters a, b and c as

$$\hat{a} = \hat{\sigma}_y/\hat{b}^{1/2} \quad ; \quad \hat{b} = (2/\hat{\gamma}_{y1})^2 \; ; \quad \hat{c} = \bar{y} - \hat{\sigma}_y \hat{b}^{1/2} \tag{10}$$

The second moment method makes use of $\bar{y}$, $\hat{\sigma}_y$ and $\hat{\gamma}_{y2}$ given by

$$\hat{\gamma}_{y2} = S_{ky}\;(A + B S_{ky}^2) \tag{11}$$

where

$$A = 1 + 6.51/N + 20.2/N^2 \quad ; \quad B = 1.48/N + 6.77/N^2 \tag{12}$$

The bias correction formula for small samples drawn from the P3 distribution, eq. 11, was proposed by Bobée and Robitaille (1975). A slight modification of the above method is to replace $\hat{\gamma}_{y1}$ in eq. 10 by $\hat{\gamma}_{y2}$. The resulting estimators are

$$\hat{a} = \hat{\sigma}_y/\hat{b}^{1/2} \quad ; \quad \hat{b} = (2/\hat{\gamma}_{y2})^2 \; ; \quad \hat{c} = \bar{y} - \hat{\sigma}_y \hat{b}^{1/2} \tag{13}$$

Method of Quantiles

Iwai (1949) and Stedinger (1980) proposed the quantile method to estimate the lower bound or location parameter for a three-parameter log normal (LN3) distribution in which the largest and smallest observations with the sample median were employed. Considerably less is known for parameter estimation techniques of gamma-type distributions via the quantile method. We introduce here the quantile-lower bound estimation procedures using a simple extension of a Wilson-Hilferty transformation.

The standard gamma quantile w in eq. 3 is approximated using a Wilson-Hilferty transformation as

$$w = b[1 - (1/9b) + (t/3b^{1/2})]^3 \tag{14}$$

where w is distributed as standard gamma distribution with shape parameter b, and t is a standard normal variate corresponding to a probability level. Hence, a P3-distributed variate y is expressed from eqs. 3 and 14 as

$$y = c + ab[1 - (1/9b) + (t/3b^{1/2})]^3 \tag{15}$$

Using the median, and the p and $(1-p)$ quantiles, the following three equations are obtained:

$$\begin{cases} y_m = c + ab[1 - (1/9b)]^3 \\ y_p = c + ab[1 - (1/9b) + (t_p/3b^{1/2})]^3 \\ y_{1-p} = c + ab[1 - (1/9b) - (t_p/3b^{1/2})]^3 \end{cases} \tag{16}$$

with $t_m = 0$ and $t_{1-p} = -t_p$ from the properties of the standard normal distribution. From a set of the above three equations, we have

$$2(y_m - c)^{1/3} = (y_{1-p} - c)^{1/3} + (y_p - c)^{1/3} \tag{17}$$

Note that the quantile estimate of c does not depend on t_p and other two parameters a and b. Solving eq. 17 for c, the quantile-lower bound estimator $\hat{c}_Q$ is calculated simply as

$$\hat{c}_Q = \frac{(G_1 - 3E^2/4) \pm \sqrt{(G_1 - 3E^2/4)^2 - 4(G_2 - 3E/2)(G_3 - E^3/8)}}{2(G_2 - 3E/2)} \tag{18}$$

where

$$\begin{cases} G_1 = y_m y_{1-p} + y_{1-p} y_p + y_p y_m \quad ; \quad G_2 = y_m + y_{1-p} + y_p \\ G_3 = y_m y_{1-p} y_p \quad ; \quad 3E = 8y_m - y_{1-p} - y_p \end{cases} \tag{19}$$

As was done by Stedinger (1980) for the LN3 distribution, the best choices of y_m, y_p and y_{1-p} would be the median, largest, and smallest values of observations, respectively. Two constraints are imposed on two roots of eq. 18. The first condition is such that if $\hat{c}_Q$ is the estimate of the location parameter for the P3 distribution, then $\hat{c}_Q$ should be less than or equal to the smallest observation $y_{(1)}$:

$$\hat{c}_Q \leq y_{(1)} \tag{20}$$

The second condition is such that when $\hat{c}_Q$ is obtained from eq. 18, $\hat{c}_Q$ has to satisfy the original cubic equation of eq. 17:

$$2(y_m - \hat{c}_Q)^{1/3} = (y_{1-p} - \hat{c}_Q)^{1/3} + (y_p - \hat{c}_Q)^{1/3} \tag{21}$$

The quantile estimate $\hat{c}_Q$ of c may be combined with either the maximum likelihood estimates (MLE) or the moment estimates of a and b for given c. Both combinations are considered in this paper. When $\hat{c}_Q$ is obtained from eq. 18, the other two parameters a and b can be estimated by solving the two likelihood equations as follows:

$$\begin{cases} \hat{b} = \sum_{i=1}^{N} (y_i - \hat{c}_Q)^{-1} / [\sum_{i=1}^{N} (y_i - \hat{c}_Q)^{-1} - N^2 / \sum_{i=1}^{N} (y_i - \hat{c}_Q)] \\ \\ \hat{a}\hat{b} = \sum_{i=1}^{N} (y_i - \hat{c}_Q)/N \end{cases} \tag{22}$$

Similarly, moment estimates of a and b combined with the quantile estimate of $\hat{c}_Q$ are given from eq. 5, using the sample moments, $\bar{y}$ and $\hat{\sigma}_y$ as

$$\hat{a} = \hat{\sigma}_y^2/(\bar{y} - \hat{c}_Q) \quad ; \quad \hat{b} = (\bar{y} - \hat{c}_Q)^2/\hat{\sigma}_y^2 \tag{23}$$

Method of Sextiles

The method of sextiles was developed by Jenkinson (1969) for the generalized extreme value distribution which has location, scale and shape parameters. The basic idea of this method is as follows: the range of variates is divided into six intervals in such a manner that the cumulative probability in each interval is one sixth. Jenkinson's sextile method can easily be extended to the P3 distribution for estimating the distribution parameters a, b and c. The derivation and procedure are detailed below.

From the cumulative distribution function of the P3 distribution defined in eq. 2, the bounding values of w_{j-1} and w_j in the jth interval are determined from

$$j/6 = F(w_j) = \int_0^{w_j} t^{b-1} exp(-t)dt / \Gamma(b) = \gamma(b, w_j) / \Gamma(b) \tag{24}$$

$$(j = 1, 2, 3 \cdots\cdots, 5)$$

or

$$1/6 = F(w_j) - F(w_{j-1}) = \int_{w_{j-1}}^{w_j} t^{b-1} exp(-t)dt/\Gamma(b) \tag{25}$$

where

$$y_j = c + aw_j \qquad (j = 1, 2, 3, \cdots\cdots, 5) \tag{26}$$

$\gamma(b, w_j)/\Gamma(b)$ denotes the incomplete gamma ratio. The standard gamma quantile w_j corresponding to probability $j/6$ $(j=1, 2, 3, \cdots\cdots, 5)$ is calculated from eq. 24 for a given value of b. Let υ_j be the mean value of variate in the jth interval for a standard gamma distribution defined in eq. 4, which is given by

$$\upsilon_j = [\int_{w_{j-1}}^{w_j} wf(w)dw]/[\int_{w_{j-1}}^{w_j} f(w)dw]$$

$$= 6[\int_0^{w_j} t^b exp(-t)dt - \int_0^{w_{j-1}} t^b exp(-t)dt]/\Gamma(b)$$

$$= 6[\gamma(b+1, w_j) - \gamma(b+1, w_{j-1})]/\Gamma(b) \qquad (j = 1, 2, 3, \cdots\cdots, 6) \tag{27}$$

The six means $\upsilon_1, \upsilon_2, \cdots\cdots, \upsilon_6$ called by Jenkinson the sextile means, characterize the parent distribution. From the statistical property of any distribution, the mean of sextile means has to be the mean of the assumed distribution. This can be achieved by

$$\mu_\upsilon = \sum_{j=1}^{6} \upsilon_j/6 \tag{28}$$

Substituting eq. 27 into eq. 28 yields

$$\mu_\upsilon = 6[\Gamma(b+1)/\Gamma(b)]/6 = b \tag{29}$$

It is clear that eq. 29 verifies the theoretical mean of the standard gamma distribution. The variance σ_υ^2 of sextile means can be expressed as

$$\sigma_\upsilon^2 = \sum_{j=1}^{6} (\upsilon_j - \mu_\upsilon)^2/6 \tag{30}$$

Jenkinson also defined the ℓ ratio as follows:

$$\ell = (\upsilon_2 - \upsilon_1)/(\upsilon_6 - \upsilon_5) \tag{31}$$

Note that the values of σ_υ and ℓ are a function of shape parameter b only. Thus if b is specified and $\upsilon_j (j = 1, 2, 3, \cdots\cdots, 6)$ can be computed from eq. 27, then σ_υ and ℓ ratio are consequently determined. The relationships between σ_υ and b as well as ℓ and b are tabulated by Leeyavanija (1983).

Now we restrict our attention to sextile estimates of a, b and c for the P3 distribution. The y_j contained in eq. 26 defines the bounding value in the jth interval for the P3 distribution. Let η_j be the mean value of the P3 distributed variate in the jth interval. When making use of a linear relationship between P3 and standard gamma distributed variates, the sextile means η_j are expressed by corresponding values of υ_j as

$$\eta_j = c + a\upsilon_j \qquad (j = 1, 2, 3 \cdots\cdots, 6) \tag{32}$$

Hence, the mean μ_η, and variance σ_η^2 of sextile means η_j are given from eq. 32 by

$$\mu_\eta = c + a\mu_\upsilon = c + ab \quad ; \quad \sigma_\eta^2 = a^2\sigma_\upsilon^2 \tag{33}$$

where

$$\mu_\eta = \sum_{j=1}^{6} \eta_j/6 \quad ; \quad \sigma_\eta^2 = \sum_{j=1}^{6} (\eta_j - \mu_\eta)^2/6 \tag{34}$$

From the relation of eq. 32, the ℓ ratio in eq. 31 is obtained from

$$\ell = (\eta_2 - \eta_1)/(\eta_6 - \eta_5) = (\upsilon_2 - \upsilon_1)/(\upsilon_6 - \upsilon_5) \tag{35}$$

As mentioned earlier, the ℓ ratio depends on shape parameter b only and hence an estimate of $\hat{\ell}$ from the observations is used to determine the corresponding shape parameter $\hat{b}$. In the next step, $\hat{\sigma}_\upsilon$ can be obtained from known $\hat{b}$. Finally, the estimates of $\hat{\mu}_\eta$ and $\hat{\sigma}_\eta$ can be computed using eq. 34. The sextile estimates of $\hat{a}$ and $\hat{c}$ are given from eq. 33 as follows:

$$\hat{a} = \hat{\sigma}_\eta / \hat{\sigma}_* \quad ; \quad \hat{c} = \hat{\mu}_\eta - \hat{a}\,\hat{b} \tag{36}$$

As a summary of how to estimate the parameters a, b and c for the P3 distribution by the sextile method, the following computational schemes are taken:

(1) Arrange the observations in ascending order and divide the ordered observations into six groups of equal size. If the sample size is N and M is an integer part of $N/6$, there should be M sample members in some groups and $(M+1)$ in others; it is advisable to assign the larger items in the lower and upper tails of the distribution to reduce the sampling variabilities of data.

(2) Compute the sextile mean in each group of samples, η_j ($j = 1, 2, 3, \cdots\cdots, 6$).

(3) Compute the mean $\hat{\mu}_\eta$, variance $\hat{\sigma}^2_\eta$ and ratio $\hat{\ell}$ of sextile means η_j, using eqs. 34 and 35.

(4) Compute the shape parameter $\hat{b}$ from $\hat{\ell}$ and then $\hat{\sigma}_*$ from the corresponding estimate $\hat{b}$, using eqs. 30 and 31. Of practical concern is how to determine the shape parameter b from the ℓ ratio in eq. 31 as well as σ_* from b in eq. 30. For computational convenience, the polynomial interpolating curves are fitted to the tabular results to yield an estimate of b from the ℓ ratio (Leeyavanija, 1983). Similarly we use polynomial fitting functions to estimate σ_* as a function of b.

(5) Compute the scale parameter $\hat{a}$ and location parameter $\hat{c}$, using eq. 36.

Five fitting techniques of the Pearson type 3 (P3) distribution tested in this study are summarized below:

Method 1 ; the method of moments with the biased skew, using eq. 10

Method 2 ; the method of moments with the unbiased skew, using eq. 13

Method 3 ; the quantile method with the maximum likelihood estimates, using eqs. 18 and 22

Method 4 ; the quantile method with the moment estimates, using eqs. 18 and 23

Method 5 ; the method of sextiles, using eqs. 35 and 36

LOG PEARSON TYPE 3 DISTRIBUTION AND PARAMETER ESTIMATION

The density function of a log Pearson type 3 (LP3) distribution is defined by

$$f(x) = [(\ell n\, x - c)/a]^{b-1} \exp[-(\ell n\, x - c)/a]/[\,|a|\,\Gamma(b)x] \tag{37}$$

By definition, the logarithmically transformed variate y ($=\ell n\, x$) is distributed as P3 whose density function is given by eq. 1 for $a > 0$. The shape of the LP3 distribution is characterized as follows:

If $a > 0$, then y has a positive skewness and $c \le y < +\infty$ and $exp(c) \le x < +\infty$.

If $a < 0$, then y has a negative skewness and $-\infty < y \le c$ and $0 < x \le exp(c)$.

The first three moments of $y = \ell n\, x$ are given by eq. 5 in which γ_s should be replaced by $2a/(\,|a|\,b^{1/2})$. The cumulative distribution function of a LP3-distributed variate x is defined in standard form as

$$p = F(x) = \begin{cases} \int_0^w f(w)dw & a > 0 \\ 1 - \int_0^w f(w)dw & a < 0 \end{cases} \tag{38}$$

where

$$w = (\ell n\, x - c)/a \tag{39}$$

$F(x)$ is the cumulative distribution function of a LP3-distributed variate x, and $f(w)$ in eq. 38 is the standard gamma distribution defined by eq. 4.

The population moments of the LP3 distribution are given by

$$\upsilon_{rr} = exp(rc)(1 - ra)^{-b} \qquad 1 - ra > 0 \qquad (r = 1, 2, 3 \cdots\cdots) \tag{40}$$

where υ_{rr} is the rth moment about the origin.

The sign of scale parameter a for the LP3 distribution is classified into two cases:

$$\text{if } \gamma_x < C^3_{vx} + 3C_{vx}, \text{ then } a < 0 \text{ and } 0 < x \le exp(c) \tag{41}$$

$$\text{if } \gamma_x > C^3_{vx} + 3C_{vx}, \text{ then } a > 0 \text{ and } exp(c) \le x < +\infty \tag{42}$$

where C_{vx} and γ_x are the coefficients of variation and skewness for the LP3 distribution, respectively.

The conditions of eqs. 41 and 42 are of practical use to check whether or not the fitted distribtions have upper bounds, when real-space moments are used to estimate the parameters of LP3 distribution.

Method of Moments

There are two ways to estimate the parameters of a log Pearson type 3 (LP3) distribution using moments; the first method fits a Pearson type 3 (P3) distribution to logarithmically transformed data (U. S. WRC, 1977), the second fits a LP3 distribution to untransformed data (Bobée, 1975; Hoshi and Burges, 1981). When using log-space moments for parameter estimation of LP3, we follow the same procedures as in eq. 13 for the P3 distribution in which all sample statistics are computed in log space ($y = \ell n\, x$). Following the procedure developed by Bobée (1975), the parameter estimation technique using real-space moments is briefly described. The sample moments about the origin are computed from

$$\hat{\theta}_{xr} = \sum_{j=1}^{N} x_j^r / N \qquad (r = 1, 2, 3) \tag{43}$$

From eqs. 40 and 43 the estimate of scale parameter $\hat{a}$ of the LP3 distribution is expressed as

$$\frac{\ell n\,[(1-\hat{a})^3/(1-3\hat{a})]}{\ell n\,[(1-\hat{a})^2/(1-2\hat{a})]} = D \tag{44}$$

where

$$D = \frac{\ell n\,\hat{\theta}_{x3} - 3\,\ell n\,\hat{\theta}_{x1}}{\ell n\,\hat{\theta}_{x2} - 2\,\ell n\,\hat{\theta}_{x1}} \tag{45}$$

Once $\hat{a}$ has been determined from eq. 44, the other two parameters can easily be computed from eqs. 40 and 43 as

$$\begin{cases} \hat{b} = \ell n(\hat{\theta}_{x2}/\hat{\theta}_{x1}^2)/\ell n[(1-\hat{a})^2/(1-2\hat{a})] \\ \hat{c} = \ell n\,\hat{\theta}_{x1} + \hat{b}\,\ell n(1-\hat{a}) \end{cases} \tag{46}$$

To solve eq. 44 for $\hat{a}$, several approximate functions are proposed by Bobée (1975) and Kite (1977). We have also attempted to fit polynomial functions to the results of the theoretical relationship between parameter a and D ratio for computational convenience (Leeyavanija, 1983).

Method of Quantiles

Previously the quantile-lower bound method was presented to fit the P3 distribution to hydrologic data. Such a method can readily be extended to estimate the parameters of the LP3 distribution. There are two cases where the LP3 distribution has a lower bound or an upper bound, depending on the constraints imposed by eq. 42 or eq. 41. For both cases, the estimate of location parameter c can be computed by eq. 18; the quantities of y_m, y_p and y_{1-p} contained in eq. 19 should be replaced by $y_m = \ell n\, x_m$, $y_p = \ell n\, x_p$ and $y_{1-p} = \ell n\, x_{1-p}$, respectively. For the LP3 distribution having a lower bound (i. e., eq. 42), two conditions to be satisfied are given by eqs. 20 and 21 in which $y_{(1)}$ should be replaced by $y_{(1)} = \ell n\, x_{(1)}$ with $x_{(1)}$ being the smallest observation. Similarly for the LP3 distribution having an upper bound (i. e., eq. 41), two constraints are imposed on the roots of eq. 18 as follows:

$$\begin{cases} \hat{c}_Q \geq y_{(N)} = \ell n\, x_{(N)} \\ 2\,(\hat{c}_Q - y_m)^{1/3} = (\hat{c}_Q - y_{1-p})^{1/3} + (\hat{c}_Q - y_p)^{1/3} \end{cases} \tag{47}$$

where $x_{(N)}$ is the largest observation; $y_m = \ell n\, x_m$, $y_{1-p} = \ell n\, x_{1-p}$, and $y_p = \ell n\, x_p$.

For both cases where the LP3 distribution has either lower bound or upper bound, the maximum likelihood estimates of a and c coupling with the quantile estimate $\hat{c}_Q$ are given by eq. 22 in which y_i should be replaced by $y_i = \ell n\, x_i$ ($i = 1, 2, 3, \cdots\cdots, N$). The quantile estimate of $\hat{c}_Q$ is also combined with the moment estimators of a and b, using the sample moment estimates $\bar{y}$ and $\hat{\sigma}_y^2$ in log space: eq. 23 can be used for both cases where the scale parametre $\hat{a}$ becomes either positive or negative.

When the estimate $\hat{c}_Q$ of location parameter c is obtained from the quantile method, it is possible to estimate the parameters a and b, using the first two moments in the real domain. For known $\hat{c}_Q$, an estimate of a can be expressed by the sample moments about the origin of eq. 43 as

$$\frac{\ell n\,(1-\hat{a})}{\ell n\,(1-2\hat{a})} = \frac{\ell n\,\hat{\upsilon}_{x1} - \hat{c}_Q}{\ell n\,\hat{\upsilon}_{x2} - 2\hat{c}_Q} = D_1 \tag{48}$$

Again to facilitate practical computations, we use polynomial interpolating functions to estimate a as a function of D_1, based on tabular results between the scale parameter a and D_1 ratio. When a has been estimated, the shape parameter b is easily estimated using eq. 40 as

$$\hat{b} = -(\ell n\,\hat{\upsilon}_{x1} - \hat{c}_Q)/\ell n(1-\hat{a}) \tag{49}$$

where $\hat{\upsilon}_{x1}$ is the sample mean of the untransformed data.

Method of Sextiles

An attempt is made to extend the methodology of the sextile method developed for the P3 distribution to parameter estimation of the LP3 distribution. If the variate x is distributed as LP3, then $y = \ell n\,x$ is distributed as P3 whose relationship is given by eq. 39, i.e.:

$$y = \ell n\,x = c + aw \tag{50}$$

When the sample skew of logarithmically transformed data is positive (i.e., $a>0$), the parameter estimating procedures for LP3 are exactly the same as those for P3; log-space statistics $\hat{\mu}_y$, $\hat{\sigma}_y$ and $\hat{\ell}$ result from using ordered sequences of $y_{(i)} = \ell n\,x_{(i)}$ $(i=1, 2, 3, \cdots\cdots, N)$ where $x_{(i)}$ is the ith smallest observation. The shape parameter b is estimated from an estimate $\hat{\ell}$ via polynomial equations, while the estimates of a and c result from eq. 36.

A slight modification is needed to accommodate parameter estimation of the LP3 distribution for which the sample skewness in log space is negative (i.e., $a < 0$). From eqs. 38 and 50, the P3 quantile y_j $(j = 1, 2, 3, \cdots\cdots, 5)$ corresponding to a non-exceedance probability level $j/6$ is related to the standard gamma quantile w_j dependent on a probability level of $(1-j/6)$ for $a < 0$. Thus for $a < 0$, eq. 26 should be replaced by the following relationship:

$$y_j = c + a\,w_{6-j} \qquad (j = 1, 2, \cdots\cdots, 5) \tag{51}$$

Accordingly, eq. 32 reduces to

$$\eta_j = c + a\,\upsilon_{7-j} \qquad (j = 1, 2, 3, \cdots\cdots, 6) \tag{52}$$

where η_j and υ_j are the sextile mean of P3 and standard gamma distributed variates y and w in the jth interval, respectively.
The substitution of eq. 52 into eq. 31 yields

$$\ell = (\eta_6 - \eta_5)/(\eta_2 - \eta_1) = (\upsilon_2 - \upsilon_1)/(\upsilon_6 - \upsilon_5) \tag{53}$$

It is of interest to compare eq. 53 with eq. 35 in that the ℓ statistic for the P3 having a negative skew is the reciprocal of the ℓ ratio for the positively skewed case. What is required to obtain the sextile estimators of parameters a and c is identical with the procedure given by eq. 36; if the skewness coefficient of transformed data is negative, then the negative sign should be assigned to $\hat{a}$.

As for the Pearson type 3 (P3) distribution, Monte Carlo experiments are conducted to evaluate the performance of fitting methods for the log Pearson type 3 (LP3) distribution. The following are the parameter estimation techniques examined for the LP3 distribution:

Method 1 ; the method of moments in log space (LS), using eq. 13
Method 2 ; the method of moments in real space (RS), using eqs. 44 and 46
Method 3 ; the quantile method with the maximum likelihood estimates (MLE), using eqs. 18 and 22
Method 4 ; the quantile method with the moment estimates in log space (LS), using eqs. 18 and 23
Method 5 ; the quantile method with the moment estimates in real space (RS), using eqs. 18, 48 and 49
Method 6 ; the method of sextiles, using eqs. 35 and 36 (or eqs. 53 and 36)

PERFORMANCE COMPARISON

To evaluate the performance of the quantile and sextile methods as new fitting procedures for the

Pearson type 3 (P3) and log Pearson type 3 (LP3) distributions, a comparison is made between the proposed and currently available methods, using Monte Carlo techniques. Performance criterion used herein is the root mean square error (rmse) of the estimate of quantile corresponding to each of six probability levels. This criterion certainly allows for an objective comparison of the methods and can aid in identification of the best fitting method to use in a particular application. The root mean square error (rmse) of quantile estimate $\hat{z}_p$ ($\hat{y}_p$ for P3 and $\hat{x}_p$ for LP3) is defined by

$$[\mathrm{rmse}(\hat{z}_p)]^2 = E[(z_p - \hat{z}_p)^2] = [z_p - E(\hat{z}_p)]^2 + \mathrm{Var}(\hat{z}_p) \tag{54}$$

where $\mathrm{Var}(\hat{z}_p)$ is the variance of quantile estimates ; $\mathrm{Var}(\hat{z}_p) = E\{[\hat{z}_p - E(\hat{z}_p)]^2\}$ and z_p is the true quantile of the parent distribution corresponding to a probability level p.
In order to eliminate the effect of scale on Monte Carlo results, the ratio $\mathrm{rmse}(\hat{z}_p)/z_p$ is used to allow for the comparison of the relative errors which are made in estimating z_p.

The Monte Carlo experiments evaluate the performance of quantile estimates of the $100p$ percentile z_p (y_p for P3 or x_p for LP3) for $p = 0.01, 0.10, 0.50, 0.90, 0.99$ and 0.998. These correspond to the 1.01, 1.11, 2, 10, 100, and 500 year events, respectively. Study of the results corresponding to $p = 0.01$, 0.10 and 0.5 allows investigation of the relative performance of the methods in a drought frequency context. Results with $p =$ 0.90, 0.99 and 0.998 provide a comparison of how well the methods might perform in flood frequency investigations.

Pearson Type 3 Distribution

The first Monte Carlo experiments are to fit the Pearson type 3 (P3) distribution to observations drawn from a two-parameter gamma (P2) distribution (i.e., $c = 0$ and $\gamma_y = 2C_v$). Two-parameter gamma distributed deviates with sample size $N = 20$, 40 and 80 are generated, using a simple extension of Kirby's (1972) algorithm (Hoshi and Burges, 1981). In all cases, the population mean μ_y is always set to 1.0. The population coefficient of variation C_v examined in this analysis is $C_v = 0.25$, 0.50 and 1.0; hence corresponding skew coefficients are 0.50, 1.0 and 2.0, respectively.

Generated sequences are sometimes rejected for several reasons, because five fitting methods fail to produce parameter estimates with some samples. The following are the reasons for rejecting a generated sequence:

(1) When the sample skew is very small ($S_{ky} \leq 0.05$), the sample is rejected, because the use of a normal distribution might be appropriate for such a sequence. A comparable problem results if the skew estimator is used when a negative sample skew is generated.

(2) In the method of quantiles, the sample sequences result in estimates $\hat{c}_Q$ of lower bound c which are imaginary in eq. 18 or violate the two constraints on $\hat{c}_Q$ given by eqs. 20 and 21. In these cases the generated sample is rejected and a new sequence is generated.

(3) Samples with values of $\hat{b} \leq 0$ are rejected as being inappropriate to fit with P3 by the quantile-maximum likelihood estimator in eq. 22.

(4) With the sextile method the rejection occurs when the ℓ ratio becomes larger than 0.85. Beyond this range,

Table 1
Number of Rejected Sequences for Various Reasons before 2500 Acceptable Samples are Obtained when Fitting the P3 Distribution to the P2 Distribution

	Coefficient of variation C_v	0.25			0.50			1.00		
Reasons	Sample size N	20	40	80	20	40	80	20	40	80
Small skewness ($S_{ky} \leq 0.05$)		133	199	62	54	15	2	3	0	0
Negative skewness ($S_{ky} < 0$)		828	470	123	209	41	0	13	0	0
Infeasible solution for $\hat{c}_Q$		82	105	70	53	12	0	421	635	1106
ℓ ratio greater than 0.85		212	296	256	86	41	19	4	0	0
Negative $\hat{b}$ in MLE		2	3	8	3	2	0	0	0	0

Table 2
Root Mean Square Error of Quantile Estimates with $C_v = 0.25$ and $\gamma_y = 0.50$

Fitting method	$rmse(\hat{y}_p)/y_p$					
p	0.01	0.10	0.50	0.90	0.99	0.998
			Sample size N = 20			
1) Moments with biased skew	0.188	0.091	0.060	0.070	0.105	0.132
2) Moments with unbiased skew	0.250	0.098	0.064	0.070	0.124	0.173
3) Quantile + MLE	0.211	0.093	0.061	0.071	0.114	0.149
4) Quantile + Moments	0.213	0.093	0.061	0.070	0.111	0.145
5) Sextiles	0.221	0.092	0.074	0.073	0.106	0.142
			Sample size N = 40			
1) Moments with biased skew	0.138	0.064	0.044	0.049	0.075	0.097
2) Moments with unbiased skew	0.163	0.065	0.045	0.049	0.083	0.113
3) Quantile + MLE	0.143	0.064	0.045	0.050	0.082	0.109
4) Quantile + Moments	0.145	0.064	0.044	0.049	0.078	0.101
5) Sextiles	0.147	0.066	0.049	0.049	0.081	0.109
			Samle size N = 80			
1) Moments with biased skew	0.104	0.046	0.033	0.035	0.055	0.072
2) Moments with unbiased skew	0.112	0.047	0.033	0.035	0.057	0.078
3) Quantile + MLE	0.101	0.046	0.034	0.036	0.060	0.080
4) Quantile + Moments	0.105	0.046	0.033	0.035	0.055	0.073
5) Sextiles	0.104	0.047	0.035	0.035	0.053	0.070

Table 3
Root Mean Square Error of Quantile Estimates with $C_v = 0.50$ and $\gamma_y = 1.00$

Fitting method	$rmse(\hat{y}_p)/y_p$					
p	0.01	0.10	0.50	0.90	0.99	0.998
			Sample size N = 20			
1) Moments with biased skew	0.790	0.223	0.125	0.128	0.180	0.213
2) Moments with unbiased skew	1.014	0.265	0.131	0.126	0.202	0.268
3) Quantile + MLE	0.672	0.219	0.122	0.128	0.180	0.212
4) Quantile + Moments	0.679	0.219	0.122	0.128	0.180	0.213
5) Sextiles	0.710	0.220	0.136	0.133	0.174	0.206
			Sample size N = 40			
1) Moments with biased skew	0.651	0.157	0.091	0.093	0.144	0.176
2) Moments with unbiased skew	0.749	0.176	0.094	0.092	0.154	0.201
3) Quantile + MLE	0.461	0.149	0.088	0.095	0.143	0.172
4) Quantile + Moments	0.473	0.149	0.087	0.093	0.137	0.162
5) Sextiles	0.537	0.154	0.092	0.093	0.136	0.165
			Sample size N = 80			
1) Moments with biased skew	0.528	0.114	0.068	0.064	0.106	0.132
2) Moments with unbiased skew	0.558	0.120	0.069	0.064	0.110	0.142
3) Quantile + MLE	0.328	0.108	0.065	0.068	0.110	0.133
4) Quantile + Moments	0.346	0.106	0.063	0.064	0.097	0.116
5) Sextiles	0.389	0.107	0.065	0.065	0.096	0.115

Table 4
Root Mean Square Error of Quantile Estimates with C_v = 1.00 and γ_y = 2.00

Fitting method	$rmse(\hat{y}_p)/y_p$					
p	0.01	0.10	0.50	0.90	0.99	0.998
			Sample size N = 20			
1) Moments with biased skew	39.83	1.269	0.360	0.230	0.295	0.330
2) Moments with unbiased skew	29.98	1.665	0.318	0.229	0.290	0.332
3) Quantile + MLE	10.96	0.841	0.308	0.225	0.265	0.283
4) Quantile + Moments	11.26	0.843	0.301	0.229	0.272	0.293
5) Sextiles	12.54	0.814	0.257	0.237	0.276	0.294
			Sample size N = 40			
1) Moments with biased skew	33.15	1.033	0.263	0.161	0.222	0.254
2) Moments with unbiased skew	25.14	1.094	0.235	0.161	0.213	0.242
3) Quantile + MLE	5.15	0.488	0.211	0.159	0.190	0.202
4) Quantile + Moments	5.20	0.541	0.210	0.161	0.195	0.209
5) Sextiles	10.00	0.552	0.184	0.162	0.194	0.209
			Sample size N = 80			
1) Moments with biased skew	27.71	0.878	0.200	0.119	0.173	0.200
2) Moments with unbiased skew	21.66	0.748	0.182	0.119	0.164	0.186
3) Quantile + MLE	2.44	0.291	0.146	0.116	0.135	0.140
4) Quantile + Moments	2.33	0.377	0.152	0.120	0.145	0.154
5) Sextiles	8.28	0.417	0.139	0.122	0.151	0.162

the fitted distribution yields an estimate of shape parameter b larger than 80, resulting in a small skew, and hence the use of the P3 distribution would probably not be appropriate.

Table 1 shows the frequency with which sample sequences are rejected before 2500 traces of length N = 20, 40 and 80 are ultimately obtained. As might be expected, the number of samples rejected for small or negative skew coefficients, and ℓ ratios larger than 0.85 is generally larger for the small coefficient of variation (C_v = 0.25 and γ_y = 0.50). A large portion of rejected samples with the quantile method for C_v = 1.0 is due to the violation of the first constraint on $\hat{c}_Q$ which should be less than the smallest value (i.e., eq. 20). Five fitting methods of the P3 distribution are applied to 2500 acceptable samples and the performance of estimates of the 100p percentile y_p is compared between the five fitting techniques for each of three sample sizes and three different population statistics of the P2 distribution.

Tables 2 to 4 report the performance of five methods of fitting the P3 distribution to samples drawn from a P2 distribution with C_v =0.25, 0.50, and 1.00, respectively. Several general conclusions can be drawn from Monte Carlo results shown in the tables. First consider the method of moments using the biased and unbiased skew coefficients. It is found that when $C_v \leq 0.5$, the use of unbiased skew coefficients is never better than the biased ones when estimating the extreme lower and upper quantiles. The opposite performance occurs for C_v = 1.0; the use of the unbiased skew perhaps yields the smaller root mean square error except for rmse($\hat{y}_{0.10}$) for all sample sizes examined. Second, a comparison is made between the methods of moments and maximum likelihood (MLE) with $\hat{c}_Q$ (i.e., quantile method). While the differences among the performance of these two estimators are generally small, the use of the moment estimate with $\hat{c}_Q$ tends to produce the better performance for $C_v \leq 0.5$ when estimating all five quantiles except for estimation of $\hat{y}_{0.01}$. On the other hand, the quantile estimator $\hat{c}_Q$ with the MLE can do appreciably better at estimating y_p for $p \geq 0.90$ and C_v = 1.0. The rmse's of the sextile method are comparable to those of moment and MLE estimators combined with $\hat{c}_Q$. The method of sextiles performs to estimate the median quantile ($\hat{y}_{0.50}$) considerably better than do the others for C_v = 1.0.

Overall, for less variable distributions with C_v = 0.25, there is little difference between performances of five fitting methods; the method which uses the biased skew estimator does very well at estimating $y_{0.01}$, $y_{0.99}$, and $y_{0.998}$. When $C_v \geq 0.50$, the quantile and sextile methods are always superior to the method of moments.

These new estimators may be used as a surrogate for the usual skew estimators when highly skewed distributions are encountered; as a consequence, these procedures yield very good estimates of the lower and upper quantiles.

It is of practical importance to check which methods of fitting the P3 distribution perform well when the observations are drawn from other than the P2 distribution. To assess the robustness of the methods, the second Monte Carlo analysis is to fit the P3 distribution to observations drawn from a two-parameter log normal (LN2) distribution (i.e., $\gamma_y = C_{vy}^3 + 3C_{vy}$). While we have not shown the results for the LN2 case, the study of robustness indicates that for highly variable and skewed distributions with $C_{vy} = 1.0$ and $\gamma_y = 4.0$, the moment and maximum likelihood methods with $\hat{c}_Q$ produce the best performance as measured by rmse($\hat{y}_{0.01}$) and rmse($\hat{y}_{0.10}$). For the quantile estimation of $y_{0.99}$ and $y_{0.998}$, the method of moments using the unbiased skew results in the best performance, when $\gamma_y \geq 1.6$.

Log Pearson Type 3 Distribution

As has been done with the Pearson type 3 distribution, Monte Carlo experiments are carried out to examine performances of six different fitting techniques for the log Pearson type 3 (LP3) distribution. The population mean (μ_x), coefficients of variation (C_{vx}) and skewness (γ_x) are specified to generate the LP3 distributed deviates of sample size $N = 20$, 40 and 80. The mean value is set to $\mu_x = 1$ for all the cases examined. Two parameter combinations are considered for C_{vx} and γ_x values; the first experiment uses $C_{vx} = 0.66$ and $\gamma_x = 3.95$, and the second set uses $C_{vx} = 0.5$ and $\gamma_x = 1.0$. The first combination of parameters was chosen from the experiment conducted by Landwehr et al. (1978) who used this set to explain the "condition of separation" by the Wakeby distribution. The second combination of population statistics is the same as that for a two-parameter gamma (P2) distribution. Table 5 contains the true values of population parameters and six quantiles of the two different LP3 distributions examined here. Note from eqs. 41 and 42 that the first case gives a positive scale parameter ($a > 0$) with the parent LP3 distribution having a lower bound, while the second the negative value ($a < 0$) with the upper bound.

In the Monte Carlo analysis, some samples are rejected for various reasons as shown in Table 6. The reasons for rejecting the sequences drawn from the LP3 distribution are almost the same as those for the P3, which have been reported in Table 1. Some specific comments can be made in reference to rejection processes

Table 5
Parameters and Quantiles of LP3 Distributions with $\mu_x = 1$

C_{vx}	γ_x	a	b	c	$x_{0.01}$	$x_{0.10}$	$x_{0.50}$	$x_{0.90}$	$x_{0.99}$	$x_{0.998}$
0.66	3.95	0.1366	14.2579	−2.0946	0.3187	0.4623	0.8253	1.7103	3.4808	5.2363
0.50	1.00	−0.1184	19.8036	2.2158	0.2187	0.4387	0.9145	1.6755	2.5118	3.0192

Table 6
Number of Rejected Sequences for Various Reasons before 2500 Acceptable Samples are Obtained when Fitting the LP3 Distribution to the LP3 Distribution

Reasons	$C_{vx} = 0.66, \gamma_x = 3.95$			$C_{vx} = 0.50, \gamma_x = 1.00$		
	$N = 20$	$N = 40$	$N = 80$	$N = 20$	$N = 40$	$N = 80$
Small skewness ($S_{kx} \leq 0.05$)	7	0	0	76	12	2
Negative skewness ($S_{kx} < 0$)	24	0	0	238	27	2
$\hat{c}_Q$ less than the largest observation	2	0	0	0	0	0
$\hat{c}_Q$ and $\hat{a}$ have conflicts	19	20	11	29	24	17
D ratio yields $\vert \hat{a} \vert > 0.005$	85	132	186	39	30	34
ℓ ratio greater than 0.85	624	548	349	1163	971	784
Infeasible $\hat{b}$ in MLE	11	16	23	15	25	21

with particular methods. If $\hat{c}_0$ would be a lower bound, using the quantile method, the estimate $\hat{a}$ resulting from the D_1 ratio of eq. 48 should be positive. Similarly when $\hat{c}_0$ is an upper bound, $\hat{a}$ should be negative. The generated sequence is rejected unless the above requirements are satisfied. The LP3 distribution degenerates to the log normal (LN2) distribution, when the parameters of a and b become zero and infinity, respectively; if the D ratio of eq. 44 would yield the absolute value of a larger than 0.005, the shape parameter b estimated from eq. 46 becomes quite large. Hence, the generated sample is rejected in such cases where the use of the LP3 distribution would probably be inappropriate. In all cases, six parameter estimating techniques are applied to ultimately accepted 2500 traces, each trace having a length of N = 20, 40 and 80.

Tables 7 and 8 report the relative root mean square error of quantile estimate, rmse$(\hat{x}_p)/x_p$ for probability levels of p = 0.01, 0.10, 0.50, 0.90, 0.99 and 0.998 where performances of fitting the LP3 distributions to observations from the same distributions are compared between the six estimating methods. First consider Monte Carlo results shown in Table 7 with C_{vx} = 0.66 and γ_x =3.95 for which the assumed distribution has the lower bound. There is no significant difference between the quantile method coupled with either maximum likelihood or log-space moment estimators, and the sextile method at estimating $x_{0.01}$ and $x_{0.10}$ in terms of the least rmse values. The quantile-moment estimator in real space (RS) achieves the minimum values of rmse$(\hat{x}_{0.99})$ and rmse$(\hat{x}_{0.998})$, while the performance of the moment method in RS is the second best at estimating the upper quantiles of highly skewed distributions. All the procedures except for the method using the first three real-space moments are of almost the same performance for quantile estimate of $x_{0.50}$ and $x_{0.90}$. Second, for the case of C_{vx} = 0.50 and γ_x = 1.0 with the LP3 distribution having the upper bound, the results of Table 8 reveal that the estimator with all real-space moments is the best with respect to the overall performance in the rmse $(\hat{x}_p)$; no single estimation performs significantly better than any other, when the sample size becomes larger (N = 80). Finally, it is of interest to compare methods which use the sample skew estimators in real or log spaces. When the skew coefficient in log space (LS) is positive in the experiments of C_{vx} = 0.66 and γ_x = 3.95, the method of moments in LS produces the smaller root mean square error (rmse) of $\hat{x}_p$ for $p \leq 0.90$. On the other

Table 7

Root Mean Square Error of Quantile Estimates with C_{vx} = 0.66 and γ_x = 3.95

Fitting method	p	0.01	0.10	0.50	0.90	0.99	0.998
		rmse $(\hat{x}_p)/x_p$					
		Sample size N = 20					
1) Moments in LS		0.308	0.148	0.127	0.193	0.642	2.042
2) Moments in RS		0.449	0.246	0.142	0.263	0.398	0.467
3) Quantile + MLE		0.245	0.135	0.124	0.198	0.537	1.103
4) Quantile + Moments in LS		0.246	0.135	0.124	0.196	0.507	0.988
5) Quantile + Moments in RS		0.298	0.181	0.134	0.187	0.331	0.464
6) Sextiles		0.257	0.130	0.132	0.185	0.459	0.949
		Sample size N = 40					
1) Moments in LS		0.199	0.095	0.090	0.140	0.386	0.883
2) Moments in RS		0.394	0.205	0.102	0.190	0.318	0.399
3) Quantile + MLE		0.161	0.090	0.088	0.144	0.359	0.633
4) Quantile + Moments in LS		0.163	0.090	0.087	0.141	0.326	0.550
5) Quantile + Moments in RS		0.201	0.123	0.094	0.137	0.263	0.384
6) Sextiles		0.167	0.089	0.092	0.137	0.342	0.619
		Sample size N = 80					
1) Moments in LS		0.129	0.063	0.064	0.099	0.245	0.450
2) Moments in RS		0.333	0.166	0.074	0.135	0.244	0.326
3) Quantile + MLE		0.111	0.063	0.063	0.103	0.255	0.433
4) Quantile + Moments in LS		0.114	0.062	0.063	0.099	0.220	0.346
5) Quantile + Moments in RS		0.140	0.085	0.066	0.100	0.205	0.307
6) Sextiles		0.112	0.062	0.063	0.097	0.210	0.336

Table 8
Root Mean Square Error of Quantile Estimates with $C_{vx} = 0.5$ and $\gamma_x = 1.00$

Fitting method	p	0.01	0.10	0.50	0.90	0.99	0.998
				$rmse\ (\hat{x}_p)\ /x_p$			
				Sample size N = 20			
1) Moments in LS		0.495	0.206	0.133	0.136	0.275	0.449
2) Moments in RS		0.383	0.204	0.126	0.133	0.190	0.242
3) Quantile + MLE		0.495	0.207	0.129	0.130	0.276	0.495
4) Quantile + Moments in LS		0.492	0.207	0.128	0.130	0.268	0.467
5) Quantile + Moments in RS		0.566	0.232	0.125	0.131	0.205	0.313
6) Sextiles		0.514	0.204	0.139	0.137	0.285	0.555
				Sample size N = 40			
1) Moments in LS		0.299	0.139	0.095	0.097	0.180	0.260
2) Moments in RS		0.275	0.141	0.093	0.096	0.146	0.192
3) Quantile + MLE		0.310	0.142	0.094	0.096	0.175	0.267
4) Quantile + Moments in LS		0.300	0.139	0.093	0.096	0.176	0.265
5) Quantile + Moments in RS		0.372	0.164	0.092	0.096	0.154	0.217
6) Sextiles		0.303	0.141	0.095	0.096	0.190	0.307
				Sample size N = 80			
1) Moments in LS		0.198	0.095	0.065	0.067	0.122	0.167
2) Moments in RS		0.204	0.099	0.064	0.066	0.106	0.143
3) Quantile + MLE		0.217	0.100	0.065	0.067	0.118	0.168
4) Quantile + Moments in LS		0.197	0.095	0.064	0.066	0.122	0.173
5) Quantile + Moments in RS		0.259	0.115	0.064	0.067	0.109	0.149
6) Sextiles		0.193	0.095	0.063	0.067	0.119	0.169

Table 9
Number of Statistic Relationships in Real and Log Spaces for 2500 Accepted LP3 Sequences

Outcome of statistic relationships	$C_{vx} = 0.66, \gamma_x = 3.95$			$C_{vx} = 0.50, \gamma_x = 1.00$		
	$N = 20$	$N = 40$	$N = 80$	$N = 20$	$N = 40$	$N = 80$
$\hat{\gamma}_x > \hat{C}_{vx}^3 + 3\hat{C}_{vx},\ \hat{\gamma}_y > 0$	388	801	1280	92	34	3
$\hat{\gamma}_x > \hat{C}_{vx}^3 + 3\hat{C}_{vx},\ \hat{\gamma}_y < 0$	7	6	2	16	40	26
$\hat{\gamma}_x < \hat{C}_{vx}^3 + 3\hat{C}_{vx},\ \hat{\gamma}_y < 0$	245	101	25	1729	2153	2399
$\hat{\gamma}_x < \hat{C}_{vx}^3 + 3\hat{C}_{vx},\ \hat{\gamma}_y > 0$	1860	1592	1193	663	273	72
Sequences with $\hat{c}_0$ as the lower bound	2181	2332	2436	734	323	106
Sequences with $\hat{c}_0$ as the upper bound	319	168	64	1766	2177	2394

hand, quantile estimates $\hat{x}_p$ for $p > 0.5$, employing the sample moments in real space (RS), yield the smaller rmse values, when the skew in LS is negative in the experiments with $C_{vx} = 0.5$ and $\gamma_x = 1.00$. The above observations bear out similar conclusions made by Nozdryn-Plotnicki and Watt (1979); the most efficient quantile estimates were obtained by using log-space moments for $\gamma_y > 0$, and real-space counterparts for $\gamma_y < 0$, when fitting the LP3 to samples generated from the same distribution. The method of moments recommended by the U.S. WRC (1977) for flood frequency analysis is never superior to use of real-space moments at estimating the upper quantiles of the distribution, especially for $\hat{x}_{0.998}$.

Recently, there are considerable arguments about the U.S. WRC flood frequency procedures, which fit a P3 distribution to the first three moments of the logarithmically transformed data sequences. For example, Landwehr et al. (1978) showed that the regional skew maps in the log domain, advocated by the WRC was incompatible with those in real space; it would be extremely difficult to infer the properties of real-space skews from those of log-space skews. Wallis and Wood (1985) pointed out, based on Monte Carlo experiments, that the WRC procedures performed very poorly for basins having negative log skews. Equations 41 and 42 impose the conditions under which the LP3 distribution fitted in RS has either lower or upper bounds. It is obvious that if the skew $\hat{\gamma}_y$ in LS is positive (or negative), the estimate of scale parameter, $\hat{a}$ becomes positive (or negative).

Table 9 reports the frequency with which 2500 accepted sequences would detect either positive or negative signs of $\hat{a}$, when using moment estimators in RS and LS, and the quantile method to estimate $\hat{c}_Q$ of location parameter c. It is noted from information contained in Table 9 that a large portion of generated sequences tends to result in the opposite sign of $\hat{a}$, estimated from real- and log-space moments in the case of C_{vx} =0.66 and γ_x =3.95 for which the parent LP3 distribution expects to yield the positive scale parameter. On the other hand, the minimum of about 70% yields the negative value of $\hat{a}$ in both real and log spaces for C_{vx} = 0.5 and γ_x = 1.0 as the parent expects to do. The last two rows of Table 9 indicate the extent to which the fitted LP3 distributions resulting from use of the quantile method have either lower or upper bounds. In contrast with statistical behaviors involved in moment estimators in RS and LS, generated sequences yield dominantly lower-bound estimates of $\hat{c}_Q$ (more than 87%) for C_{vx} = 0.66 and γ_x = 3.95, and upper-bound estimates of $\hat{c}_Q$(more than 70%) for C_{vx} = 0.5 and γ_x = 1.0 as the parent LP3 distributions expect to do.

CONCLUSIONS

New knowledge obtained from the present study is the quantile and sextile methods for estimating the parameters of gamma-type distributions. Monte Carlo simulation techniques were employed to evaluate the performance of these two methods by comparing with that of currently available procedures. Five estimators were used for the Monte Carlo analysis with the Pearson type 3 (P3) distribution, while six estimators for the log Pearson type 3 (LP3) distribution.

The full maximum likelihood estimator (MLE) was not pursued in this study for several reasons. The MLE cannot necessarily retain its desirable asymptotic properties for small samples and requires the numerical solution of a single nonlinear equation. Hoshi et al. (1984) showed that the quantile method combined with the MLE estimator of other two parameters always did better or almost as well as the full MLE estimator for the log normal distributions; moreover the use of the quantile method avoids the need to optimize the likelihood function numerically.

The accuracy of various quantile estimates is of practical concern because of the frequent use of quantile estimates in hydrologic designs. Thus the performance of several fitting procedures was evaluated using the criterion of the root mean square error (rmse) of quantile estimates. Not only are fitted the gamma-type distributions to the observations drawn from the same distributions but also to the samples drawn from other distributions for the test of robustness.

Based on the results of Monte Carlo experiments performed to determine the precision with which parameter estimating techniques estimate several quantiles, the following conclusions have been reached:

(1) For the P3 distributions, new procedures of quantile-MLE, quantile-moment, and sextile methods yielded a very good all-around quantile estimate and generally performed best or about the best, in terms of the rmse of estimates of y_p for p = 0.01, 0.99 and 0.998. When fitting the P3 distribution, there is little advantage to using unbiased skew estimates in the method of moments at estimating the upper quantiles for $\gamma_y \leq 1.0$.

(2) For the LP3 distribution having the lower bound (C_{vx} = 0.66 and γ_x = 3.95), the most accurate estimates of high extreme quantiles with $p \geq 0.99$ were obtained by using the quantile method coupling with the method of moments in real space. The quantile-MLE and quantile-moment (in LS) estimators yielded the least rmse values of $\hat{x}_{0.01}$ and $\hat{x}_{0.10}$; the sextile method also became very competitive.

(3) For the LP3 distribution having the upper bound (C_{vx} = 0.5 and γ_x = 1.0), overall performances of the method using the first three moments in the real domain were the best for quantile estimates $\hat{x}_p$ with either small or large p. The method of quantiles with use of real-space moments was also capable of estimating the quantiles for $p \geq 0.99$ with acceptable accuracy. The method of moments using the first three moments in log space performed very poorly. The principal usefulness of the quantile method is that it avoids the problems that are often introduced when the scale parameters resulting from use of real- and log-space statistics have different signs.

ACKNOWLEDGEMENTS

The work reported in this paper was supported in part by the research grant from the Japan International Cooperation Agency while the senior author was at the Asian Institute of Technology (AIT). The valuable assistance of AIT and Regional Computer Center is gratefully acknowledged.

REFERENCES

1. Benson, M.A. : Uniform flood frequency estimating methods for federal agencies, Water Resources Research, Vol. 4, No. 5, pp. 891-908,1968.
2. Bobée, B. : Sampling error of T-year events computed by fitting a Pearson type 3 distribution, Water Resources Research, Vol. 9, No. 5, pp. 1264-1270, 1973.
3. Bobée, B.: The log Pearson type 3 distribution and its application in hydrology, Water Resources Research, Vol. 11, No. 5, pp. 681-689, 1975.
4. Bobée, B. and R. Robitaille : Correction of bias in the estimation of the coefficient of skewness, Water Resources Research, Vol. 11, No. 6, pp. 851-854, 1975.
5. Bobée, B. and R. Robitaille : The use of the Pearson type 3 and log Pearson type 3 distributions revisited, Water Resources Research, Vol. 13, No. 2, pp. 427-442, 1977.
6. Buckett, J. and F.R. Oliver : Fitting the Pearson type 3 distribution in practice, Water Resources Research, Vol. 13, No. 5, pp. 851-852, 1977.
7. Condie, R.: The log Pearson type 3 distribution : The T-year event and its asymptotic standard error by maximum likelihood theory, Water Resources Research, Vol. 13, No. 6, pp. 987-991, 1977.
8. Jenkinson, A.F.: Estimation of maximum floods, World Meteorological Office, Technical Note 98, Chapter 5, pp. 183-257, 1969.
9. Hoshi, K. and I. Yamaoka : The generalized gamma probability distribution and its application in hydrology, Hydraulic Papers, No. 7, Department of Civil Engineering, Hokkaido University, 256 p., 1980.
10. Hoshi, K. and S.J. Burges : Sampling properties of parameter estimates for the log Pearson type 3 distribution, using moments in real space, Journal of Hydrology. No. 53, pp. 305-316,1981.
11. Hoshi, K. and S.J. Burges : Approximate estimation of the derivative of a standard gamma quantile for use in confidence interval estimates, Journal of Hydrology, No. 53, pp. 317-325, 1981.
12. Hoshi, K., J.R. Stedinger and S.J. Burges : Estimation of log-normal quantiles : Monte Carlo results and first-order approximation, Journal of Hydrology, No. 71, pp. 1-30, 1984.
13. Iwai, S.: Studies on nonsymmetrics of Slade-type distribution and several new solutions, Proc. of the Japan Society of Civil Engineers, No. 4, pp. 84-104,1949 (in Japanese).
14. Kirby, W.: Computer-oriented Wilson-Hilferty transformation that preserves the first three moments and the lower bound of the Pearson type 3 distribution, Water Resources Research, Vol. 8, No. 5, pp. 1251-1254, 1972.
15. Kite, G.W.: Frequency and risk analyses in hydrology, Water Resources Publications, 224 p., Fort Collins, Colorado, 1977.
16. Landwehr, J.M., N.C.Matalas and J.R. Wallis : Some comparisons of flood statistics in real and log space, Water Resources Research, Vol. 14, No. 5, pp. 902-920, 1978.
17. Leeyavanija, U.: Parameter estimates for gamma-type distributions by quantile method, Master thesis, No. WA-84-10, Asian Institute of Technology, 98 p., Bangkok, Thailand, 1983.
18. Matalas, N.C. and J.R. Wallis: Eureka! it fits a Pearson type 3 distribution, Water Resources Research, Vol. 9, No. 2, pp. 281-289, 1973.
19. Nozdryn-Plotnicki, M.J. and W.E. Watt : Assessment of fitting techniques for the log Pearson type 3 distribution using Monte Carlo simulation, Water Resources Research, Vol. 15, No. 3, pp. 714-718, 1979.
20. Phien, H.N. and M.A. Hira : Log Pearson type 3 distribution : Parameter estimation, Journal of Hydrology, No. 64, pp. 25-37,1983.
21. Prentice, R.L.: A log gamma model and its maximum likelihood estimation, Biometrika, Vol. 61, No. 3, pp. 539-544, 1974.
22. Rao, D.V. : Log Pearson type 3 distribution : Method of mixed moments, Journal of Hydraulic Division, ASCE,Vol. 106, No. HY6, pp. 999-1019, 1980.
23. Stacy, E.W. and G.A. Mihram : Parameter estimation for a generalized gamma distribution, Technometrics,

Vol. 7, No. 3, pp. 349–357, 1965.
24. Stedinger, J.R. : Fitting log normal distributions to hydrologic data, Water Resources Research, Vol. 16, No. 3, pp. 481–490,1980.
25. U.S.WRC (U.S. Water Resources Council) : Guidelines for determining flood flow frequency, U.S. WRC, Hydrology Committee, Washington, D.C., Bulletin 17A, 1977.
26. Wallis, J.R. and E.F. Wood : Relative accuracy of log Pearson III procedures, Journal of Hydraulic Engineering, ASCE, Vol. 111, No. 7, pp. 1043–1056, 1985.

APPENDIX – NOTATION

The following symbols are used in this paper :

a = scale parameter ;
A = coefficient defined in eq. 12 ;
b = shape parameter ;
B = coefficient defined in eq. 12 ;
c = location parameter ;
$\hat{c}_Q$ = estimate of lower or upper bounds with use of quantile method ;
C_{vx} = coefficient of variation of variate x ;
D = ratio defined in eq. 44 ;
D_l = ratio defined in eq. 48 ;
$f(\cdot)$ = probability density function ;
$F(\cdot)$ = cumulative distribution function ;
ℓ = ratio defined in eq. 31 ;
$LN2$ = two parameter log normal distribution ;
$LN3$ = three parameter log normal distribution ;
$LP3$ = log Pearson type 3 distribution ;
N = sample size ;
p = non-exceedance probability ;
$P2$ = two parameter gamma distribution ;
$P3$ = Pearson type 3 distribution ;
$rmse(\cdot)$ = root mean square error ;
s_y = sample standard deviation of variate y ;
S_{ky} = sample skew coefficient of variate y ;
t = standard normal variate ;
w = standard gamma variate ;
x = log Pearson type 3 distributed variate ;
y = Pearson type 3 distributed variate ;
$\bar{y}$ = sample mean of variate y ;
y_i = sample observation of variate y ;
$y_{(i)}$ = ith largest observation of variate y ;
$\hat{z}$ = estimate of variate or parameter z ;
z_p = $100p$ percentile of variate z ;
μ_y = population mean of variate y ;
σ_y = population standard deviation of variate y ;
γ_y = population skew coefficient of variate y ;
$\gamma(\cdot,\cdot)$ = incomplete gamma function ;
ν_l = sextile mean of standard gamma variate ;
ν_{xr} = rth moment about the origin for LP3 distribution ;
η_l = sextile mean of P3 distributed variate ; and
$\Gamma(\cdot)$ = gamma function.