Log Pearson type 3 quantile estimators with regional skew information and low outlier adjustments

V. W. Griffis and J. R. Stedinger

School of Civil and Environmental Engineering, Cornell University, Ithaca, New York, USA

T. A. Cohn

U.S. Geological Survey, Reston, Virginia, USA

Received 22 September 2003; revised 15 March 2004; accepted 3 May 2004; published 15 July 2004.

[1] The recently developed expected moments algorithm (EMA) [*Cohn et al.*, 1997] does as well as maximum likelihood estimations at estimating log-Pearson type 3 (LP3) flood quantiles using systematic and historical flood information. Needed extensions include use of a regional skewness estimator and its precision to be consistent with *Bulletin 17B*. Another issue addressed by *Bulletin 17B* is the treatment of low outliers. A Monte Carlo study compares the performance of *Bulletin 17B* using the entire sample with and without regional skew with estimators that use regional skew and censor low outliers, including an extended EMA estimator, the conditional probability adjustment (CPA) from *Bulletin 17B*, and an estimator that uses probability plot regression (PPR) to compute substitute values for low outliers. Estimators that neglect regional skew information do much worse than estimators that use an informative regional skewness estimator. For LP3 data the low outlier rejection procedure generally results in no loss of overall accuracy, and the differences between the MSEs of the estimators that used an informative regional skew are generally modest in the skewness range of real interest. Samples contaminated to model actual flood data demonstrate that estimators which give special treatment to low outliers significantly outperform estimators that make no such adjustment. *INDEX TERMS:* 1821 Hydrology: Floods; 1860 Hydrology: Runoff and streamflow; 1854 Hydrology: Precipitation (3354); *KEYWORDS:* Bulletin 17B, censored data, conditional probability adjustment, expected moments, floods, log Pearson type 3

Citation: Griffis, V. W., J. R. Stedinger, and T. A. Cohn (2004), Log Pearson type 3 quantile estimators with regional skew information and low outlier adjustments, *Water Resour. Res.*, *40*, W07503, doi:10.1029/2003WR002697.

1. Introduction

[2] Uniform flood-frequency techniques recommended for use by Federal agencies are presented in *Bulletin 17B* [*Interagency Committee on Water Data (IACWD)*, 1982]. The fields of hydrology and flood frequency analysis have substantially evolved since *Bulletin 17* was first published in 1976 and last updated in 1982, but new techniques have yet to become part of standard practice. This study attempts to quantify the value of regional skew information and the impacts of adjustments for low outliers in the flood-frequency techniques employed by U.S. federal agencies.

[3] The original *Bulletin 17* [*Water Resources Council*, 1976] included an algorithm for weighting the station skew and a regional skew. Introduction of such a weighting scheme was a new idea; *Bulletin 15* had employed the station skew when estimating an LP3 distribution. However, *Bulletin 17* lacked a theoretical justification for the proposed weights. *Tasker* [1978] suggested that the minimum variance skew estimator would be obtained by weighting station and regional skews by the inverse of their variances; *Bulletin 17B* recommends an inverse MSE weighting scheme to reflect estimator bias. This paper illustrates the value of the mean-square error (MSE)-skew weighting scheme as a function of the precision of the regional estimate and the sample size.

[4] *Bulletin 17B* (hereinafter referred to as B17) defines outliers as "data points which depart significantly from the trend of the remaining data." B17 uses a log-transformation of the data; therefore, "one or more unusual low-flow values can distort the entire fitted frequency distribution" [*Stedinger et al.*, 1993, p. 18.45]. If low outliers are identified and removed from the sample, B17 recommends the use of a conditional probability adjustment (CPA) to compute a frequency curve with the retained values.

[5] Methods developed to use historical data and censored samples can be extended for the treatment of low outliers. The expected moments algorithm (EMA) was originally developed by *Cohn et al.* [1997] for the incorporation of historical information in flood frequency analyses. This paper extends EMA to make use of a regional skewness estimator and considers use of EMA when low outliers are censored. Another alternative is probability plot regression (PPR), employed by *Gilliom and Helsel* [1986] and *Helsel and Cohn* [1988] as an estimation technique for distribution parameters of censored water-quality data sets. *Kroll and Stedinger* [1996] consider its use for water quality

0043-1397/04/2003WR002697$09.00

and low-flow frequency analyses. The research described here explores use of EMA and PPR as alternatives to the CPA estimator in *Bulletin 17B* for flood frequency analysis following the identification of low outliers.

2. Bulletin 17B Procedures

[6] B17 recommends fitting a log-Pearson type 3 (LP3) distribution to annual flood series. For a systematic record of length N years, the recommended technique is to use the method of moments to fit a Pearson type 3 (P3) distribution to the base 10 logarithms of the flood peaks, denoted $\{X_1, \ldots, X_N\}$. Estimates of the mean, standard deviation, and skew coefficient of the logarithms of the sample data are computed using traditional moment estimators.

2.1. Weighted Skew Estimation

[7] The data available at a given site are generally limited to less than 100 years and are often less than 30 years in length. The accuracy of the station skewness estimator should be improved by combining it with a regional skew estimator obtained by pooling data from nearby sites. B17 recommends combining the sample skew $\hat{\gamma}$ and the regional skew G to obtain a weighted skew:

$$\tilde{G} = \frac{\mathrm{MSE}_{\hat{\gamma}} G + \mathrm{MSE}_G \hat{\gamma}}{\mathrm{MSE}_{\hat{\gamma}} + \mathrm{MSE}_G}, \tag{1}$$

where $\mathrm{MSE}_{\hat{\gamma}}$ is the mean-square error (equal to the variance plus bias, squared) of the station skew, and MSE_G is the estimation error of the regional skew. This weighting scheme was adopted from *Tasker* [1978] but was extended by the B17 work group to address the bias in the sample skew estimate; this equation minimizes the MSE of the skew estimator provided that G is unbiased and independent of the station skewness estimator $\hat{\gamma}$ [*Griffis*, 2003].

[8] B17 recommends approximating $\mathrm{MSE}_{\hat{\gamma}}$ as a function of the sample skew and sample size using the equation provided therein, which was based on empirical values reported by *Wallis et al.* [1974]. This approximation yields relative errors as large as 10% within the hydrologic region of interest with log space skews $|\gamma| \leq 1.414$ [*Griffis*, 2003]. *Griffis* [2003] generated 10 million replicates for different cases to allow derivation of a more accurate and smooth approximation consistent with the asymptotic variance for $\hat{\gamma}$ provided by *Bobée* [1973]. She obtained

$$\mathrm{MSE}_{\hat{\gamma}} = \left[\frac{6}{N} + a(N)\right]\left[1 + \left(\frac{9}{6} + b(N)\right)\gamma^2 + \left(\frac{15}{48} + c(N)\right)\gamma^4\right], \tag{2}$$

where $a(N)$, $b(N)$, and $c(N)$ are correction factors for small samples:

$$a(N) = -\frac{17.75}{N^2} + \frac{50.06}{N^3}$$

$$b(N) = \frac{3.93}{N^{0.3}} - \frac{30.97}{N^{0.6}} + \frac{37.1}{N^{0.9}}$$

$$c(N) = -\frac{6.16}{N^{0.56}} + \frac{36.83}{N^{1.12}} - \frac{66.9}{N^{1.68}}.$$

This approximation was developed for systematic record lengths $N \geq 10$ and $|\gamma| \leq 1.414$. Within that range, the largest relative error is -0.62%. In practice, a reasonable estimator of the true skew γ should be employed in equation (2).

[9] The regional skew may be obtained from the skew map provided in B17, which was originally developed by *Hardison* [1974]. The standard error of the map is reported to be 0.55, indicating that MSE_G is approximately 0.302. *Tasker and Stedinger* [1986] showed that the B17 estimate of map error is most likely too large; in that study, their regional skew had a MSE of 0.11. Values of the same order are reported by *Martins and Stedinger* [2002] and *Reis et al.* [2003]. Those studies indicate that the estimate of the standard error of the regional skew is reduced when one accounts for the actual sampling error in the at-site skewness estimators used to construct a regional skewness estimator.

[10] Reducing the variance of the regional skew implicitly increases the hydrologic information represented in that skewness estimator. Given N years of record at station x, equations (1) and (2) are used to weight the station skew with the regional skew to obtain the minimum MSE weighted skewness estimator $\tilde{G}$ with precision [*Griffis*, 2003]:

$$\mathrm{MSE}_{\tilde{G}} = \frac{\mathrm{MSE}_{\hat{\gamma}}\mathrm{MSE}_G}{\mathrm{MSE}_{\hat{\gamma}} + \mathrm{MSE}_G} = \left\{\frac{1}{\mathrm{MSE}_{\hat{\gamma}}} + \frac{1}{\mathrm{MSE}_G}\right\}^{-1}. \tag{3}$$

Using $\mathrm{MSE}_{\tilde{G}}$ from equation (3), the effective number e of additional years of record provided by a regional skew estimator with a known variance is defined as the solution of

$$\mathrm{MSE}_{\tilde{G}} = \mathrm{MSE}_{\hat{\gamma}}(N + e, \gamma), \tag{4}$$

wherein γ is the true skew employed to compute the MSE of $\hat{\gamma}$. In this way, the effective record length e of the regional skew is defined as the additional number of years of record needed to provide an at-site skewness estimator $\hat{\gamma}$ with precision $\mathrm{MSE}_{\tilde{G}}$. Use of equation (4) requires knowledge of the true skew γ, and γ was known in the Monte Carlo analysis presented in this paper; however, in other applications, a reasonable estimator of γ would need to be employed.

2.2. Low Outlier Identification

[11] Prior to combining the sample skew with the regional skew, B17 recommends using the sample moments of the complete sample to determine thresholds for the identification of low and high outliers. In this paper it is assumed that historical information is unavailable, and therefore no adjustments for high outliers can be made, and only tests and adjustments for low outliers are conducted. If low outliers are identified, adjustments to the frequencies of the flood flows above the threshold should be made to capture the actual frequency of floods in the sample.

[12] Low outliers in log space are identified by specifying a "truncation level":

$$X_L = \bar{X} - K_N S, \tag{5}$$

which is defined by the one-sided 10% significance level for a P3 distribution with zero skew (i.e., a two-parameter normal distribution). The 10% frequency factors K_N for normal data as a function of sample size (for $10 \leq N \leq 149$)

are tabulated in B17. These values of K_N may be computed using the compact formula for $5 \leq N \leq 150$ [*Stedinger et al.*, 1993, p. 18.45]:

$$K_N = -0.9043 + 3.345\sqrt{\log_{10}(N)} - 0.4046\log_{10}(N). \quad (6)$$

B17 states that this procedure is appropriate for use with LP3 distributions with skews on the interval [−3, +3]. Any values below the truncation level X_L are considered to be low outliers and are censored in the analyses reported here. The selection of this outlier test by the B17 committee is reviewed by *Thomas* [1985]. *Spencer and McCuen* [1996] argue that a more appropriate frequency factor could be computed to handle different values of skew, detection of multiple outliers, and alternative significance levels. Identification and the censoring of low outliers frees a fitting procedure from the constraint that the LP3 distribution should describe both the distribution of the smallest and largest floods.

3. Conditional Probability Adjustment

[13] A conditional probability adjustment (CPA) of the frequency curve is recommended by B17 when low outliers are censored, when the record contains zero flows, or when there is a recording threshold resulting in a truncated data set. These critical events are censored from the record of size N and a conditional P3 distribution $F(x)$ is fit to the r retained logarithms of the annual maximum floods that exceeded the truncation level X_L. B17 does not recommend using CPA when more than 25% of the observations are censored. CPA was originally developed by *Jennings and Benson* [1969] to account for the removal of zero-flow events from a systematic record before fitting the LP3 distribution. For LP3 and lognormal data, *Kroll* [1996] compares the precision of low-flow quantile estimates obtained with CPA to maximum likelihood estimation (MLE), log-probability-plot regression (LPPR), and partial probability weighted moments (PPWM) estimators. For samples censored at the 5th, 20th, and 45th percentiles, *Kroll* [1996] observed that with LP3 data CPA performed poorly compared with MLE, LPPR, and PPWM when estimating quantiles just above the censoring threshold.

[14] The probability that a given event exceeds the truncation level is estimated as $p_e = r/N$. The formula for conditional probability expressed in terms of exceedance probabilities indicates that the flood flows exceeded with a probability $p \leq p_e$ in any year are obtained by solving $p = p_e[1 - F(x)]$ to obtain $F(x) = 1 - p/p_e$. The B17 CPA uses this equation to compute the logarithms of the flood flows ($Q_{0.99}$, $Q_{0.90}$, and $Q_{0.50}$) which will be exceeded with probabilities p = 0.01, 0.10, and 0.50. These three values are used to define a new P3 distribution for the logarithms of the flood flows which reflects the unconditional frequencies of the above threshold values. The new P3 distribution is defined by the "synthetic" moments

$$G_{\text{syn}} = -2.50 + 3.12[\log_{10}(Q_{0.99}/Q_{0.50})]/[\log_{10}(Q_{0.90}/Q_{0.50})]$$

$$S_{\text{syn}} = [\log_{10}(Q_{0.99}/Q_{0.50})]/[K_{0.99} - K_{0.50}] \quad (7)$$

$$M_{\text{syn}} = \log_{10}(Q_{0.50}) - K_{0.50}S_{\text{syn}},$$

where $K_{0.99}$ and $K_{0.50}$ are P3 frequency factors dependent on the synthetic skew and the exceedance probability, 0.01 or 0.50, respectively. The approximation for G is said to be appropriate for skew coefficients on the interval [−2.0, +2.5] [*IACWD*, 1982]. The absolute error in the computed skew is an unnecessarily large 0.04 for $|\gamma| \leq 0.2$, which are in the center of the hydrologic region of interest [*Griffis*, 2003].

[15] The final fitted distribution used to estimate the frequency of the r above-threshold values is given by the synthetic mean, synthetic standard deviation, and a weighted skew obtained by combining the synthetic skew with a regional skew using equation (1).

4. Probability Plot Regression

[16] Probability plot regression (PPR) is a statistical estimation method that has been employed with censored water-quality, low-flow, and flood data. PPR fills in missing observations and zeros using estimates of the missing observations obtained by a regression of the observed values against their normal scores or another appropriate variate. The method was formalized by *Gilliom and Helsel* [1986] and was later studied by *Helsel and Cohn* [1988] and *Kroll and Stedinger* [1996]. This method also appears in the statistical literature where it has been applied to normal samples [*David*, 1980]. Hydrologic applications of the method have employed a lognormal model, though other models could be adopted. Here an extension of PPR for use with P3 distributions is proposed.

[17] Cumulative plotting positions for the censored observations (low outliers) are computed employing the Blom formula as

$$p_i = (c/N)\left[\left(i - {}^3\!/_8\right)/\left(N - c + {}^1\!/_4\right)\right] \quad (8)$$

for $i = 1\ldots c$, where c is the number of censored observations and N is the total number of observations in the sample [*Hirsch and Stedinger*, 1987]. If r is the number of retained observations, then $N = c + r$. The quantity c/N is the probability a flood is below the threshold. The estimate of the probability that a flood exceeds the threshold is $r/N = 1 - c/N$ so that the plotting positions for the r retained observations beginning with a cumulative probability of c/N can be computed as

$$p_i = (c/N) + (1 - c/N)\left[\left(i - {}^3\!/_8\right)/\left(N - c + {}^1\!/_4\right)\right] \quad (9)$$

for $i = 1\ldots r$, where $i = 1$ corresponds to the smallest retained observation [*Stedinger et al.*, 1993, p. 18.42]. *Hirsch and Stedinger* [1987] and *Kottegoda and Rosso* [1997, p. 496] employ equations (8) and (9) with historical information.

[18] The standard P3 variates K_p are determined for all observations using the assigned plotting positions and the regional skew. (K_p values employed in this study were computed using the MATLAB command "gaminv.") The two moments μ and σ relating the standard P3 variates and the r above-threshold observations are determined using ordinary least squares with the simple linear model

$$x_p = \mu + \sigma K_p(\gamma). \quad (10)$$

Equation (10) is also used to estimate the values of the censored observations using the standard P3 variates based on plotting positions from equation (8) and the regional skew. These estimates are combined with the retained observations to form a completed data set for which the moments of the distribution are computed using the traditional B17 estimators.

[19] A weakness of the PPR method is that the magnitudes assigned to the low outliers are not affected by the value of the censoring threshold, and thus some critical information is lost describing the possible values of those censored observations. This concern is illustrated by the fact that censored observations can be assigned values that exceed the threshold. However, this does not appear to be a significant problem with typical samples and may make our PPR flood estimators more robust in this application.

[20] PPR should perform well when the regional skew is relatively accurate so that the use of the regional skew adds little error to the estimated moments of the distribution and when a modest number ($\leq$25% of sample) of outliers are identified. For example, *Kroll and Stedinger* [1996] show that PPR works well for such cases with censored normal samples.

5. Expected Moments Algorithm

[21] The expected moments algorithm (EMA) was proposed by *Cohn et al.* [1997] as an alternative to maximum likelihood estimation (MLE) and the B17 methodology for incorporation of historical data into flood frequency analyses. EMA employs an iterative procedure for computing parameter estimates using censored data. The process begins with an initial set of parameter estimates obtained using the systematic stream gage record and then updates the parameters using the known magnitudes of historical peaks and the expected contribution to the moment estimators of the below-threshold floods.

[22] For an LP3 distribution, *Cohn et al.* [1997] demonstrated that EMA is more efficient than the B17 method for using historical data and is nearly as efficient as MLEs with cases for which the MLE procedure converged reliably. Their results were limited to estimators of the 99th percentile; *England et al.* [2003b] further evaluated the use of EMA with historical and paleohydrologic information to estimate larger percentiles. Application of EMA to practical cases was investigated by *England et al.* [2003a]. The *National Research Council* [1999] employed EMA for flood frequency analysis on the American River in California. *Jarrett and Tomlinson* [2000] used EMA in their study on the Yampa River in Colorado.

[23] The expected moments algorithm for low outlier adjustment includes the following steps:

[24] 1. A threshold (X_L) is defined below which observations are considered outliers.

[25] 2. Using the values that exceeded the threshold ($X_L^>$), initial estimates of the sample moments ($\hat{\mu}_1$, $\hat{\sigma}_1$, $\hat{\gamma}_1$) are computed as if one had a complete sample.

[26] 3. For iteration $i = 1, 2, \ldots$, the parameters of the P3 distribution ($\hat{\alpha}_{i+1}$, $\hat{\beta}_{i+1}$, $\hat{\tau}_{i+1}$) are estimated using the previously computed sample moments:

$$\hat{\alpha}_{i+1} = 4/\hat{\gamma}_i^2; \hat{\beta}_{i+1} = \frac{1}{2}\hat{\sigma}_i\hat{\gamma}_i; \hat{\tau}_{i+1} = \hat{\mu}_i - \hat{\alpha}_{i+1}\hat{\beta}_{i+1}.$$

[27] 4. New sample moments ($\hat{\mu}_{i+1}$, $\hat{\sigma}_{i+1}$, $\hat{\gamma}_{i+1}$) are estimated using expected moments such as

$$\hat{\mu}_{i+1} = \frac{\sum X_L^> + N^< E\left[X_L^<\right]}{N}, \tag{11}$$

where $N^<$ represents the number of observations below the threshold, N is the total number of observations, and $E[X_L^<]$ is the expected value of an observation known to have a value below the low outlier threshold X_L. The expected value is a conditional expectation given that $X < X_c$, where X_c denotes the EMA censoring threshold which is defined here as the smallest retained observation. Use of the smallest retained observation rather than X_L to define the possible range of censored values made the EMA algorithm less sensitive to the distribution of low outliers [*Griffis*, 2003]. With the current parameter estimates ($\hat{\alpha}_{i+1}$, $\hat{\beta}_{i+1}$, $\hat{\tau}_{i+1}$), the conditional expectation is expressed in terms of the incomplete Gamma function [*Cohn et al.*, 1997]:

$$E\left[X_L^<\right] = \tau + \beta\frac{\Gamma\left[\frac{X_c - \tau}{\beta}, \alpha + 1\right]}{\Gamma\left[\frac{X_c - \tau}{\beta}, \alpha\right]}. \tag{12}$$

[28] The second and third moments are estimated using

$$\hat{\sigma}_{i+1}^2 = \frac{1}{N}\left\{c_2 \sum\left(X_L^> - \hat{\mu}_{i+1}\right)^2 + N^< E\left[\left(X_L^< - \mu\right)^2\right]\right\}, \tag{13}$$

$$\hat{\gamma}_{i+1} = \frac{1}{N\hat{\sigma}_{i+1}^3}\left\{c_3 \sum\left(X_L^> - \hat{\mu}_{i+1}\right)^3 + N^< E\left[\left(X_L^< - \mu\right)^3\right]\right\}, \tag{14}$$

wherein $c_2 = N^2/(N-1)$ and $c_3 = N^2/[(N-1)(N-2)]$.

[29] Equation (14) neglects regional skewness information. B17 recommends weighting the regional skew with the synthetic skew obtained after adjusting the fitted P3 distribution for low outliers using CPA. The same approach could be used here to obtain a weighted skewness estimator $\tilde{G}$ via equation (1) using $\hat{\gamma}_{i+1}$ from equation (14) to estimate $\hat{\gamma}$. However, the methodology should be improved by incorporating the regional skew into the EMA procedure to ensure that the weighted skew corresponds to the adjusted mean and standard deviation fit to the data. The suggested extension of EMA for computing the third moment with regional skew information is

$$\hat{\gamma}_{i+1} = \frac{1}{(N+n)\hat{\sigma}_{i+1}^3}\left\{c_3 \sum\left(X_L^> - \hat{\mu}_{i+1}\right)^3 + N^< E\left[\left(X_L^< - \mu\right)^3\right] + nG\hat{\sigma}_{i+1}^3\right\}, \tag{15}$$

where n is the additional years of record assigned to the regional skew. Here $\hat{\gamma}_{i+1}$ is a weighted skewness estimator. To ensure that EMA is consistent with *Bulletin 17B* when no low outliers are identified (i.e., $\hat{\gamma}_{i+1} = \tilde{G}$ in equation (1)), the required value of n is

$$n = N\frac{\text{MSE}_{\hat{\gamma}}}{\text{MSE}_G}. \tag{16}$$

In this sense, n is the regional skew weight measured in years.

[30] The expected contribution to the second and third central moments (m = 2 and 3, respectively) of the below-threshold values is [*Cohn et al.*, 1997]

$$E\left[\left(X_L^< - \mu\right)^m\right] = \sum_{j=0}^{m} \binom{m}{j} \beta^j (\tau - \mu)^{m-j} \frac{\Gamma\left[\frac{X_c - \tau}{\beta}, \alpha + j\right]}{\Gamma\left[\frac{X_c - \tau}{\beta}, \alpha\right]}. \quad (17)$$

Steps 3 and 4 are repeated until the parameter estimators for the P3 mean, standard deviation, and skew values converge.

[31] *Cohn et al.* [1997] discussed the use of EMA with historical data. Equations (11), (13), and (15) can be modified to include historical data by adding the terms $N_H^< E[(X_H^< - \mu)^m]$ and $\hat{c}_m \sum (X_H^> - \hat{\mu}_{i+1})^m$, where the subscript H denotes the historical threshold and m is the moment being evaluated. In equations (13) and (15), the latter term is multiplied by an appropriate bias-correction factor reflecting the use of a mean estimator $\hat{\mu}_{i+1}$. In this case, the correction factors $\hat{c}_m$ should be based upon the total systematic record length plus the number of observed historical floods as recommended by *Cohn et al.* [1997].

[32] Confidence intervals for flood quantiles based upon EMA flood quantile estimators were developed by *Cohn et al.* [2001]. This is a feature lacking in the B17 approach when historical data or low outliers are present, as well as proposed improvements for regular data sets [*Chowdury and Stedinger*, 1991; *Whitley and Hromadka*, 1999].

5.1. Bias-Correction Factors

[33] The Cohn EMA procedure includes bias-correction factors which ensure that the computed moments coincide with those used in B17 when no historical information is employed [*Cohn et al.*, 1997, p. 2091]. Their bias-correction factors are

$$\tilde{c}_2 = \left(N_S^< + N^>\right) / \left(N_S^< + N^> - 1\right)$$
$$\tilde{c}_3 = \left(N_S^< + N^>\right)^2 / \left[\left(N_S^< + N^> - 1\right)\left(N_S^< + N^> - 2\right)\right]. \quad (18)$$

These scale the summation terms of the observed peaks in the computation of the variance and skew, respectively. Here $N_S^<$ is the number of observed peaks in the systematic record below the historical threshold and $N^>$ is the number of observed peaks in both the systematic record and historical period which exceed the historical threshold. The bias-corrections do not include the number of censored-historical values, because if the censoring threshold is quite high, they would provide very little information pertaining to the mean.

[34] In the extension of EMA for use with low outliers, the corrections are only applied to the observed values greater than X_L. However, unlike the historical information case, additional years of information are not added to the record when adjusting for low outliers: N remains unchanged as does the relative information in the sample. When low outliers are censored from the record, the number of above-threshold values decreases and thereby increases $\tilde{c}_2$ and $\tilde{c}_3$ in equation (18); thus the relative weight placed on the above-threshold values would increase. This does not make any sense. The computed weights in equations (13) and (14) avoid these inconsistencies. These equations are consistent with traditional moment estimators and the EMA estimators currently used with historical information and reflect a reasonable bias correction for the use of the "sample" average $\hat{\mu}_{i+1}$ in the summation terms in the estimators of the second and third moments.

[35] In equations (13) and (14), the bias-corrections c_2 and c_3 are only applied to the summation terms involving the above-threshold observations. The expectation of the contribution from the low outliers to the variance and skew coefficient are computed using equation (17). Equation (17) was derived (Appendix A) assuming the true mean μ is known. Thus it truly is the expectation of $E[(X_L^< - \mu)^m]$ for m = 2 or 3. Therefore application of a bias correction to these terms is not appropriate. The correction is appropriate for sample estimators $(X_i - \bar{X})^m$ which suffer from the correlation between X_i and $\bar{X}$. Similarly, the regional skew estimate is assumed to be unbiased, so when included in the EMA algorithm, as in equation (15), this term should not be adjusted for bias.

5.2. Weighted Skew Constraints

[36] Negatively skewed P3 distributions have an upper bound but are unbounded in the lower tail. As a result, it is possible for the skew to become increasingly negative with each EMA iteration. In the literature, population skews are commonly restricted to values of ±1.0 [*Chowdury and Stedinger*, 1991; *Spencer and McCuen*, 1996; *Cohn et al.*, 1997; *Whitley and Hromadka*, 1999; *McCuen*, 2001]. *Chowdury and Stedinger* [1991] restrict generated sample skews to be less than ±1.5. It is unlikely that the population skew would ever fall below −1.4, corresponding to shape parameter α = 2 and a P3 distribution whose density function goes to zero linearly at the upper bound. Therefore it is reasonable to restrict the skew computed by EMA to be greater than or equal to −1.4. This skew constraint is imposed by performing a check at the end of each iteration to see if $\hat{\gamma}_{i+1} \geq -1.4$. If $\hat{\gamma}_{i+1} < -1.4$, then $\hat{\gamma}_{i+1}$ is set equal to −1.4 and the algorithm proceeds. Still, in some extreme cases, EMA fit a P3 distribution with an upper bound within the observed data, and this too was a concern.

[37] EMA utilizes the method of moments, which summarizes the information in the data set by the sample moments. Thus it is quite possible for the computed upper bound to be smaller than one or more of the observations. The upper bound must be at least as large as the largest observation for the fitted distribution to be valid; because of the interest in larger flood quantiles, we added this additional constraint to the estimation procedure.

[38] For $\hat{\gamma}_{i+1} < 0$, the upper bound constraint is checked at the end of each iteration by computing the upper bound ($\tilde{\tau}$) of the distribution corresponding to the updated sample moments ($\hat{\mu}_{i+1}$, $\hat{\sigma}_{i+1}$, $\hat{\gamma}_{i+1}$), where

$$\tilde{\tau} = \hat{\mu}_{i+1} - 2\left[\hat{\sigma}_{i+1} / \hat{\gamma}_{i+1}\right]. \quad (19)$$

[39] The upper bound $\tilde{\tau}$ is compared to the maximum observation $x_{\max}$. If $\tilde{\tau} < x_{\max}$, then the upper bound is within the data, and the skew is recomputed as

$$\hat{\gamma}_{i+1} = 2\hat{\sigma}_{i+1} / \left(\hat{\mu}_{i+1} - x_{\max}\right). \quad (20)$$

The next iteration of the algorithm uses this adjusted skew, which must equal or exceed both a value of −1.41 and the value in equation (20).

[40] In a few cases with positive skews, the lower bound τ exceeded the smallest observation; however, this is generally not a concern in estimating floods. This would be a concern in low-flow analyses and similar constraints could be implemented. The P3 distribution fitted using the EMA algorithm would not be used to describe the frequency of flood flows below the censoring threshold X_c because the model has not attempted to reproduce the distribution of floods in that range.

6. Monte Carlo Analysis

[41] A Monte Carlo experiment was conducted to compare the following seven P3 fitting methods: (1) MOMn: method of moments utilizing all of the sample data, with no weighting of the sample skew with the regional skew; (2) CPA: conditional probability adjustment as recommended by B17 for adjusting fitted sample parameters following the identification of low outliers; (3) CPAc: conditional probability adjustment with a lower bound of −1.4 on the fitted skew and a constraint that the upper bound must equal or exceed the largest observation; (4) EMAbc: expected moments algorithm for low outlier adjustment with the incorporation of regional skew following the B17 recommendation summarized by equation (1) with a lower bound on the fitted skew and a constraint on the upper bound; (5) PPR: probability plot regression used to fill in values of censored observations (the final P3 parameters are determined using method of moments with the completed sample); (6) MOM: method of moments utilizing all sample data with weighting of the station skew with the regional skew as recommended by B17; and (7) MOMc: method of moments utilizing all sample data with weighting of the station skew with the regional skew, with a lower bound on the skew and a constraint on the upper bound.

[42] Because the results from EMA can be improved by imposing a lower bound on the skew and a constraint on the upper bound, it is likely that the performances of CPA and MOM would also be improved by the same constraints. The methods CPAc and MOMc were used to check the affect of the constraints on the performance of CPA and MOM, respectively.

[43] The seven parameter estimation methods were compared using the mean square error (MSE) and bias of the quantile estimators of a range of quantiles. Results for the 100-year event ($X_{0.99}$) are reported here. The MSE was computed as

$$\mathrm{MSE} = \frac{1}{M}\sum_{i=1}^{M}\left(\hat{X}_p(i) - X_p\right)^2. \tag{21}$$

The MSE performance measure in log space reflects the precision with which the fitted P3 distributions approximate the true quantiles of the parent population from which the samples were generated. *Kroll and Stedinger* [1996] compare real- and log-space MSEs.

6.1. Data Generation

[44] Data for the experiment were generated from P3 populations with sample sizes of 10, 25, 50, and 100 and log space regional skews between −1.0 and +1.0. For each sample, possible population skews were randomly generated about the specified regional skew with a specified variance using the methodology proposed by *Chowdury and Stedinger* [1991]. If the regional skew was negative, then the population skews were randomly generated from a gamma distribution with a lower bound of −1.4. If the regional skew was positive, then the population skews were generated from a gamma distribution with an upper bound of 1.4. For regional skews of zero, the population skews were generated from a normal distribution. Samples generated from a P3 distribution with a mean of 3.5, a standard deviation of 0.26, and the specified population skew distributions with variances of 0.010, 0.100, and 0.302 are considered. The Monte Carlo analyses presented in this paper consider only the bias and mean square error of quantile estimators, so the choice of the mean and variance of the P3 distribution is not critical to the problem. For computation of weighted skewness estimators, the estimation error of the regional skew MSE_G is equated to the specified variance of the population skews Var[γ].

[45] This study considers regional skews in the range [−1.0, +1.0]. *Hardison* [1974] reports mean regional skewness values in the range [−0.5, +0.6], with a standard error for individual station estimators of 0.550 (corresponding to $\mathrm{MSE}_G = 0.302$). For a partition of the United States into 14 regions, *Landwehr et al.* [1978] report mean regional skew values in the range [−0.4, +0.3]. The wider interval for regional skews of [−1.0, +1.0] was adopted to allow exploration of a broader range that encompasses the most likely values.

[46] Another issue would be a realistic range for the population skews for an individual station. As noted above, −1.4 is a realistic lower bound. The range [−1.4, +1.4] is certainly within the distribution of site-to-site variability observed in *Hardison*'s [1974] estimates and substantially larger than that suggested by $\mathrm{MSE}_G \approx 0.100$ as reported by *Tasker and Stedinger* [1986] and *Reis et al.* [2003]. Thus it is certainly reasonable to place bounds of ±1.4 on generated population skews so as to restrict the analysis to reasonable values.

6.2. Results

[47] The Monte Carlo experiment was conducted for four types of samples: (1) P3 distributed data using all generated samples, (2) P3 distributed data using only samples containing low outliers, (3) contaminated samples, and (4) P3 distributed samples with censoring at the 20th percentile. The following sections discuss the four sets of results. The design of the experiment was the same for each; differences resulted from the treatment of the samples after they were generated. Only P3 distributed samples containing at least one low outlier are considered in case 2 to see if averaging results over all of the generated samples masks the value of low outlier procedures. To model real flood data and better assess the value of low outlier procedures, P3 distributed samples are contaminated in case 3 by reducing the smallest observations by a specified factor. Finally in case 4, the effect of an increased truncation level on the low outlier adjustment methods was assessed using a 20% frequency factor instead of the frequency factor recommended by B17, which would censor only one in 10 normal samples.

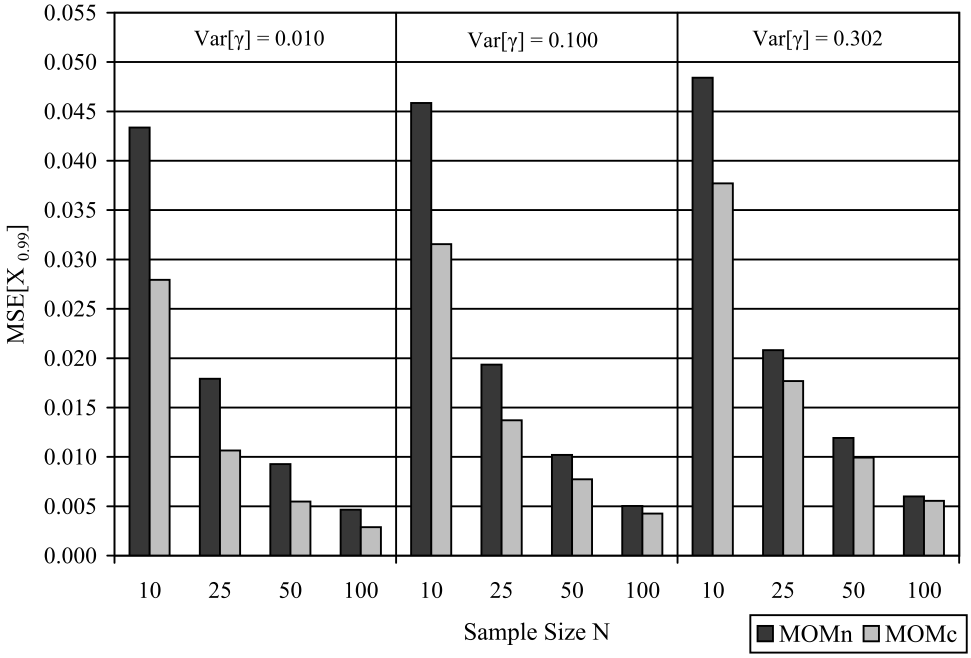

Figure 1. MSE of $X_{0.99}$ for MOMn and MOMc quantiles estimators in P3 distributed samples as a function of N and Var[γ] for $G = 0$.

6.2.1. P3 Distributed Data Using All Generated Samples

[48] Using all generated samples, regardless of whether they contain low outliers or not, illustrates the benefit of weighting with a regional skew and allows the overall need for and effect of the low outlier adjustment to be described. Comparisons of quantile estimates were made for all combinations of sample size, regional skew, and population skew variance using $M = 5000$ replicates. Figure 1 illustrates the MSE of the $X_{0.99}$ estimators using MOMn and MOMc with a regional skew of 0 as a function of sample size N and the variance of the population skew Var[γ]. Figures 2 and 3illustrate the MSE and bias, respectively, of the $X_{0.99}$ estimators using all seven fitting methods with a sample size of 25 years and Var[γ] = 0.100. (*Griffis* [2003] provides figures illustrating the MSE and bias of the $X_{0.99}$ estimators for all combinations of sample size, regional skew, and population skew variance.)

[49] In samples of size 25 with a regional skew of −1.0, roughly 42% of the samples contained at least one low outlier; the fraction of samples containing outliers increases to 55% in samples of size 50. With a regional skew of +1.0, the percentage of samples containing low outliers is less than 1%.

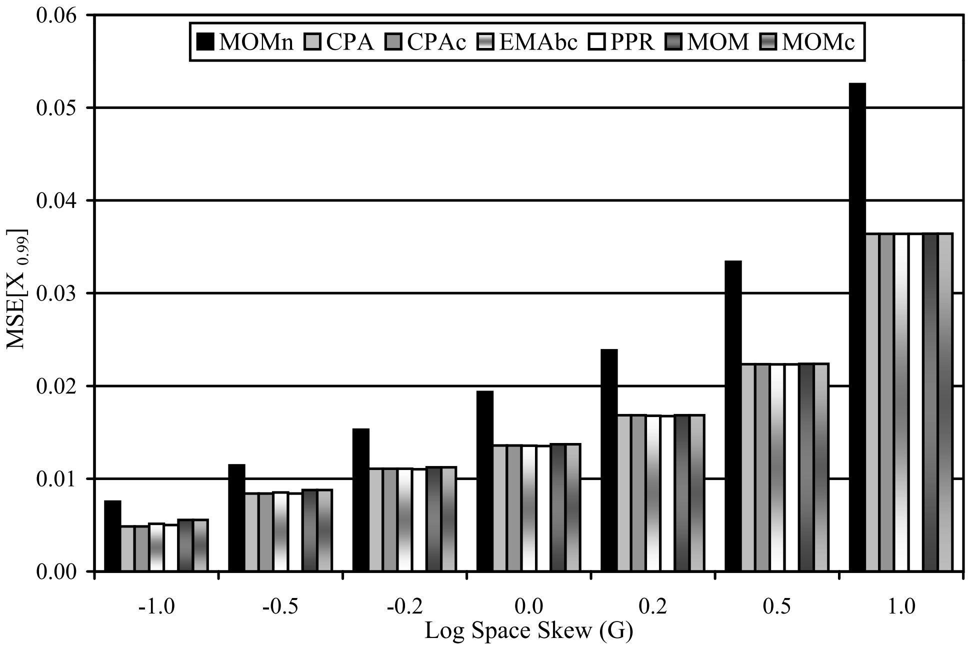

Figure 2. MSE of $X_{0.99}$ estimators for each method in P3 distributed samples ($N = 25$, Var[γ] = 0.100).

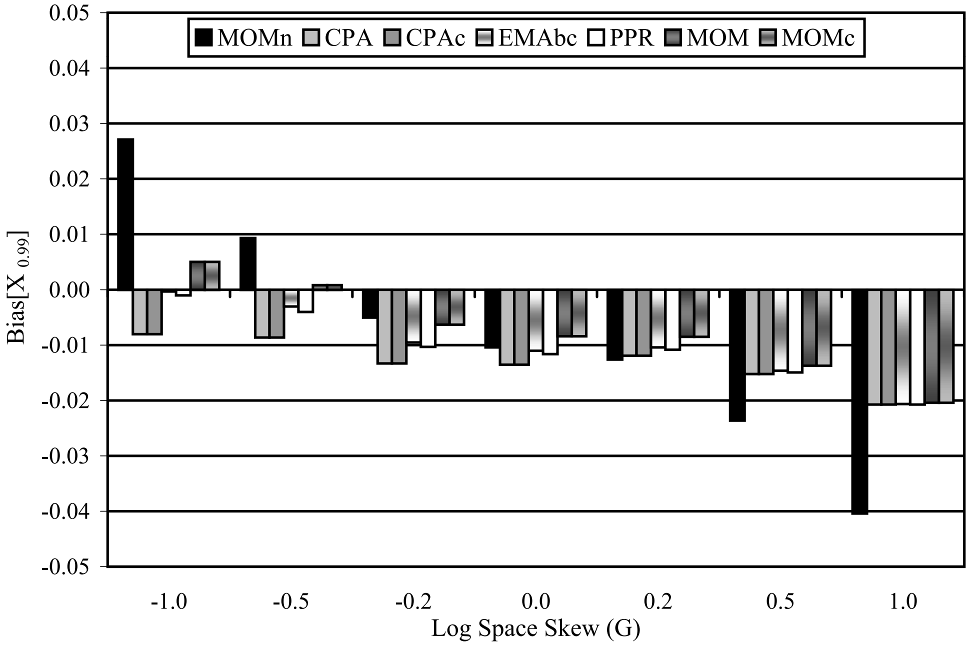

Figure 3. Bias of $X_{0.99}$ estimators for each method in P3 distributed samples (N = 25, Var[γ] = 0.100).

6.2.1.1. Impact of Constraints

[50] Table 1 reports the frequencies with which the lower bound on the weighted skew and the constraint on the upper bound are active with each fitting method for specified sample sizes, regional skew values, and population skew variances. Except for the MOMn results, the frequencies for sample size-regional skew combinations not included in the table are zero (in 5000 replicate samples); the constraints were active only in samples with a regional skew of −1.0 (an extreme case), except for N = 100, where a skew of −0.5 with a variance of 0.302 (an extreme case) also generated upper bound constraint violations. Violation of the lower bound constraint on the weighted skew only occurred with Var[γ] = 0.302 (again an extreme case). Although PPR was not constrained, the frequencies with which the computed upper bound fell within the sample data are also reported.

6.2.1.2. Value of Regional Skew

[51] Weighting the sample skew with an informative regional skew dramatically reduces the MSE and bias of the $X_{0.99}$ estimators. In Figure 1 the large differences between the MSE of the MOMn and the MOMc quantile estimators illustrate the significant benefit of weighting a station skew with an informative regional skew. The MOMn estimator does not utilize regional skew, but the MSE of the MOMn estimator increases with Var[γ], and thus MSE_G, due to the character of the generated samples. The benefit of weighting with an informative regional skew is evident as the relative difference between the MSE of the MOMn and MOMc estimators increases as the variance of the population skew decreases (i.e., the precision of the regional skew increases). Furthermore, as skew estimates associated with smaller samples have greater error, the benefit of weighting

Table 1. Frequencies (%) of Invoking Weighted Skew and Upper Bound Constraints in P3 Distributed Samples

Sample Size	Regional Skew	$\tilde{G} < -1.4$			$\hat{\tau} < x_{max}$				
		CPAc	EMAbc	MOMc	MOMn	CPAc	EMAbc	MOMc	PPR
				Population Skew Variance = 0.010					
10	−1.0	0.0	0.0	0.0	2.5	1.7	1.1	0.9	0.9
25	−1.0	0.0	0.0	0.0	8.3	2.7	1.9	1.5	1.7
50	−1.0	0.0	0.0	0.0	12.7	2.7	1.7	1.5	1.7
100	−1.0	0.0	0.0	0.0	14.1	1.8	1.3	1.2	1.3
				Population Skew Variance = 0.100					
10	−1.0	0.0	0.0	0.0	2.5	0.0	0.0	0.0	0.0
25	−1.0	0.0	0.0	0.0	8.5	0.0	0.1	0.2	0.0
50	−1.0	0.0	0.0	0.0	11.8	0.1	0.3	0.5	0.2
100	−1.0	0.0	0.0	0.0	14.5	0.1	0.8	1.2	0.5
				Population Skew Variance = 0.302					
25	−1.0	0.0	0.0	0.0	9.2	0.0	0.1	0.2	0.0
50	−1.0	0.0	0.4	3.3	13.1	0.0	0.3	1.0	0.1
100	−1.0	0.0	1.6	6.2	16.3	0.1	0.9	2.3	0.2
100	−0.5	0.0	0.0	0.0	6.4	0.0	0.0	0.3	0.0

with a regional skew is more evident in these samples. *Tasker* [1978] demonstrates the value of reasonable weighting of at-site and regional skewness estimators. However, his Monte Carlo analysis only included an approximation of the optimal weighting factors for $MSE_G = 0.302$, which is the largest value considered here.

[52] The relative differences between the MOMn and MOMc estimators shown in Figure 1 for a regional skew of 0 are typical of other values of regional skew, with the only difference being changes in the actual values of the MSE. In terms of MSE, the value of weighting decreases with sample size. In Figure 1 for a regional skew of 0 with an estimation error of 0.100 and an effective record length $e \approx 60$, the MSE is reduced approximately 31% with $N = 10$ but only 18% with $N = 100$. The value of weighting is greater in cases where $e \gg N$ and diminishes as N approaches or exceeds e. However, for a regional skew of 0 with an estimation error of 0.302 which has an effective record length $e \approx 20$, the MSE is reduced approximately 22% with $N = 10$ and 7.5% with $N = 100$. The value of weighting is smaller with a less informative regional skew.

[53] For a fixed value of Var[γ], the value of weighting also increases with the absolute value of the regional skew coefficient because the variance of the at-site skew is larger when $|\gamma|$ is larger. As a consequence, the value of e increases. For samples of size 50 with Var[γ] = 0.100, the value of e increases from roughly 60 years with a regional skew of 0 to almost 112 years with a regional skew of ± 1.0. As a result, in samples of size 50 with Var[γ] = 0.100, the MSE is reduced an average of 30% by weighting the sample skew with a regional skew of ± 1.0 ($e \approx 112$ years) and is reduced roughly 23% using a regional skew of 0 ($e \approx 60$ years). However, in Figure 2 for samples of size 25, the MSE is reduced an average of 31% by weighting the sample skew with a regional skew of ± 1.0 ($e \approx 128$ years) and is reduced roughly 30% using a regional skew of 0 ($e \approx 60$ years). The increased value of weighting with a larger regional skew coefficient is negligible in samples of size $N \leq 25$ in Figure 2, because for all values of skew considered, the information in the regional skew overwhelms the sample skew.

[54] Figure 3 shows that use of a regional skew generally reduces the bias of the $X_{0.99}$ estimators, particularly for larger regional skews. In samples of 100, the use of a regional skew generally resulted in increased bias of quantile estimators because the sample size exceeded the effective record length of the regional skew. While bias is a part of MSE, it also describes a different character of the estimators and can be worthwhile considering. However, because this analysis employs base 10 logarithms, the worst biases reported in Figure 3 correspond to an error on the order of 5% of the real space flood quantiles. Therefore none of the biases is a significant part of the estimators' MSE.

6.2.1.3. Regional Skew and Skew Constraints

[55] As reported in Table 1, averaging the sample skew with a regional skew in MOMc significantly reduces the frequency with which the computed upper bound falls within the sample data when compared with using the pure method of moments (MOMn). Using an informative regional skew coefficient with Var[γ] $\leq$ 0.100, the lower bound on the weighted skew and the upper bound constraint have little effect on the CPA and MOM quantile estimators as the lower bound constraint on the weighted skew is never binding and the constraint on the upper bound is invoked infrequently, and only for extreme cases when the regional skew has a value of -1.0. The constraint is invoked more frequently with Var[γ] = 0.010 than with Var[γ] = 0.100 for all methods utilizing a weighted skew estimate, because the weighting scheme places much more weight on this unrealistic regional skew than the sample skew. Therefore the weighted skew estimate will have a value approximately equal to the regional skew of -1.0. Cases with more realistic regional skew values resulted in no violations of the constraint on the upper bound.

6.2.2. P3 Distributed Data Using Only Samples Containing Low Outliers

[56] The overall impact of low outlier procedures and the effect of the choice of quantile estimators are assessed by comparing the performance of CPA, EMAbc, and PPR with MOM. Because MOM utilizes all of the sample data, one might suspect that it would result in the best performance and thus have the smallest MSE in this case wherein all of the data are actually from a P3 distribution. As shown in Figure 2, applying low outlier adjustment methods to P3 data results in no observable loss of overall accuracy in terms of MSE when all generated samples are considered. However, averaging the results over all generated samples masks the actual affect of the low outlier adjustment, particularly when $G > 0$.

[57] Because few samples are identified with regional skew values $G \geq +0.5$, the low outlier adjustment procedures are used infrequently and the results from all methods utilizing regional skew information should coincide. Therefore low outlier adjustments are relatively insignificant in this skew range, and regional skew values of +0.5 and +1.0 were omitted.

[58] Comparisons of quantile estimates were made for all combinations of sample size and regional skew ($G = -1.0$, -0.5, -0.2, 0.0, and +0.2) and population skew variance using $M = 1000$ replicates. For only samples containing low outliers, Figure 4 illustrates the MSE of the $X_{0.99}$ estimators of all seven fitting methods as a function of sample size with a regional skew of 0 and a population skew variance of 0.100. Figure 5 compares the MSE of the $X_{0.99}$ estimators as a function of regional skew in samples of size 25 with a population skew variance of 0.100. Figure 6 compares the MSE of the $X_{0.99}$ estimators as a function of Var[γ] with a regional skew of 0. The results for CPAc and MOMc are not included in Figure 6 because they are identical to the results of CPA and MOM, respectively, because the constraints are never binding with $G = 0$. (*Griffis* [2003] provides results for all combinations of sample size, regional skew, and population skew variance.)

[59] In Figure 4 for a regional skew of 0 with a variance of 0.100, only in very small samples $N = 10$ did MOM significantly outperform the estimators employing a low outlier adjustment procedure. EMA and PPR outperform CPA for $N \leq 25$. The differences in the MSEs of estimators that use a weighted skew are negligible with $N \geq 50$. In Figure 5 for samples of size 25 for cases with a population skew variance of 0.100, EMA and PPR consistently outperformed CPA with reasonable regional skew values $|G| \leq 0.5$.

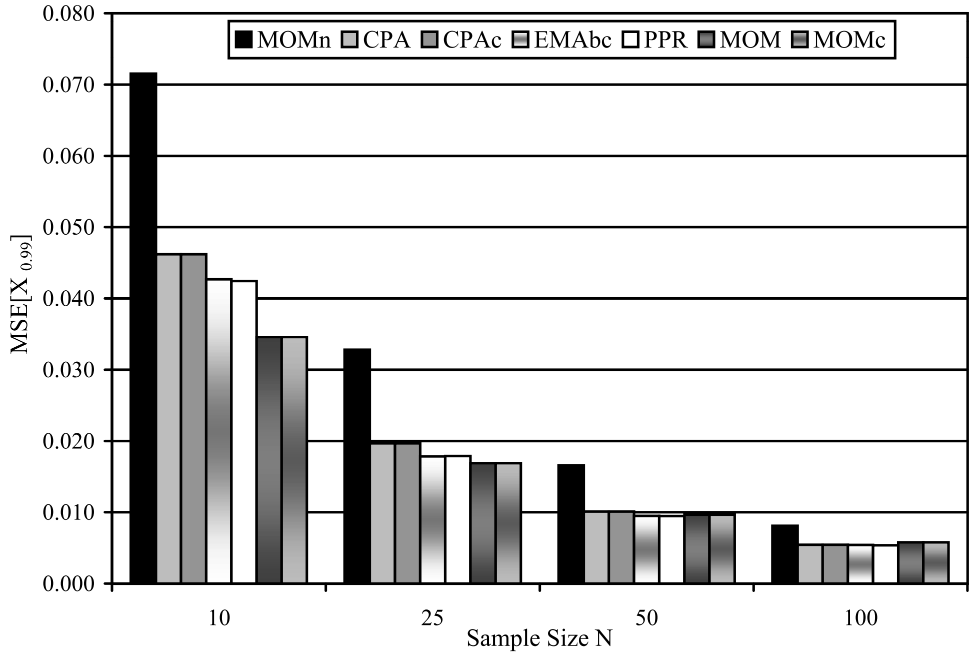

Figure 4. MSE of $X_{0.99}$ estimators for each method in P3 distributed samples containing low outliers (G = 0, Var[γ] = 0.100).

[60] In general, for the cases considered here for P3 distributed samples containing at least one low outlier, the performance of EMA and PPR were similar in terms of MSE. EMA and PPR generally did as well as or better than CPA. In Figure 6 for a regional skew of 0 with a large variance of 0.302, so that the information in the sample exceeded the information in the regional skew (i.e., $N > e \approx 20$), CPA, EMA, and PPR generally outperformed MOM. On the other hand, for a realistic population skew variance of 0.100 with smaller samples ($N \leq 25$) so that $e \approx 60 > N$, MOM had smaller MSEs. Use of EMA results in no loss of overall accuracy when outliers are identified in P3 distributed samples of typical size ($25 \leq N \leq 50$) with an informative regional skew (MSE_G = 0.100) and reasonable skew values ($|G| \leq 0.2$), as compared to CPA, or to MOM with the entire data set.

6.2.3. Contaminated P3 Distributed Samples

[61] Thus far the Monte Carlo analysis of the seven estimation methods has assumed that the sample data is truly P3 distributed, which is the underlying assumption of

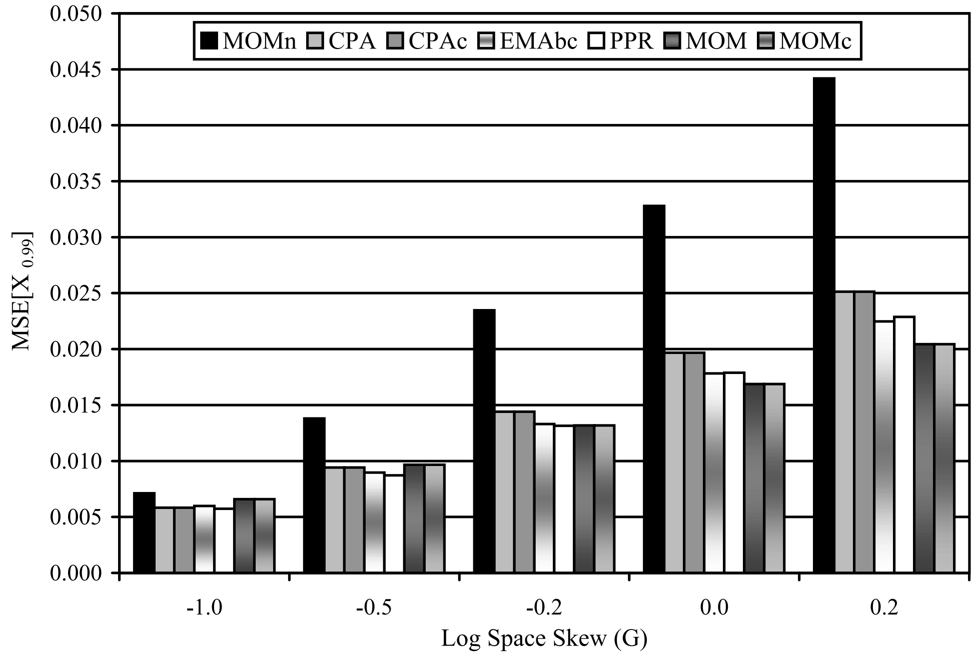

Figure 5. MSE of $X_{0.99}$ estimators for each method in P3 distributed samples containing low outliers (N = 25, Var[γ] = 0.100).

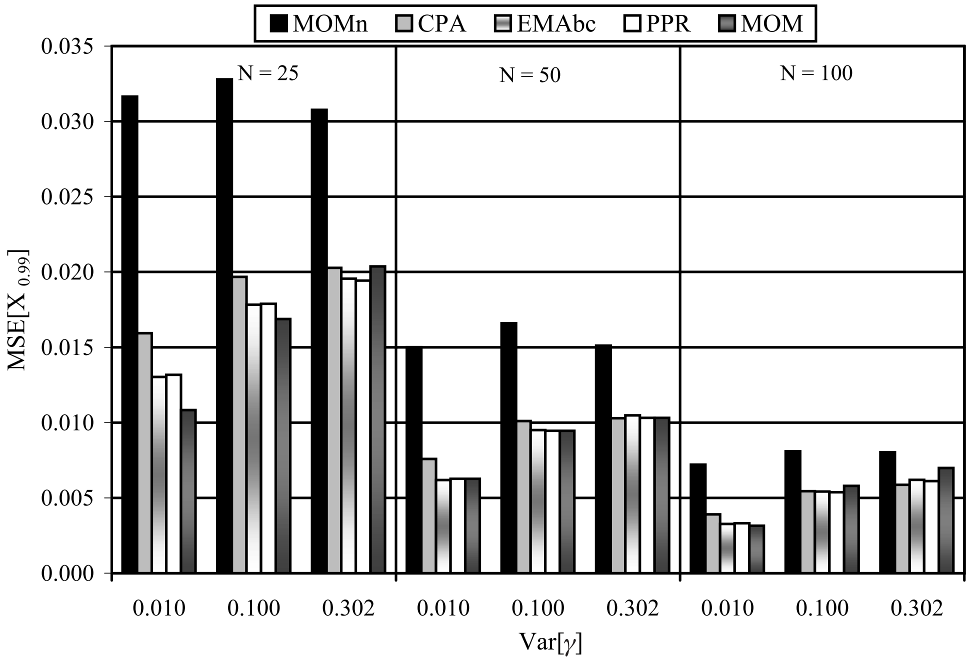

Figure 6. MSE of $X_{0.99}$ estimators for each method in P3 distributed samples containing low outliers (G = 0).

the methods recommended by B17. However, in reality, flood records are most likely not truly P3 distributed and true low outliers can depart significantly from the general trend of the data. To pursue this real concern, contaminated P3 samples were considered to illustrate the potential value of a low outlier detection step and adjustment of the fitted distribution.

6.2.3.1. Contamination of Samples

[62] Evaluation of several annual maximum flood series indicates that low outliers typically depart from the general trend of the data by factors in the range of 2 to 5. (See, for example, *Bulletin 17B*, pp. 12–24.) However, observations are generally not identified as low outliers using equation (5) unless they differ from the general trend of the data by a factor of 3 or more. Furthermore, when a sample contains more than one low outlier, the outliers often depart from the general trend of the data by the same severity [*Griffis*, 2003].

[63] For this analysis, P3 distributed samples of size 25, 50, and 100 were generated as described in section 6.1. Samples of size 10 were omitted from the analysis because even without contamination, there are insufficient data to adequately fit a three-parameter distribution to the sample (although B17 allows such small samples, as did Figures 1 and 4).

[64] To demonstrate that outlier adjustment is truly advantageous when samples contain real outliers, only samples that contained a specific number of outliers were considered. The numerical values of the smallest k = 1, 2, and 3 observations in the generated P3 distributed samples of size 25, 50, and 100, respectively, were contaminated to model real samples containing low outliers. The smallest k observations in each sample were contaminated by subtracting $\log(f)$, equivalent to dividing by a factor f in real space. The original sample value was replaced by the contaminated value resulting in a contaminated sample.

[65] To model actual flood records containing low outliers a factor f = 5 was used to contaminate the generated P3 distributed samples. The moments of the contaminated samples were used in equation (5) to estimate a low outlier threshold. P3 distributions were fit to the contaminated samples using the seven estimation methods. The use of a large f factor ensured that the contamination always provided at least one low outlier that would be identified by equation (5). The results are actually relatively insensitive to the value of f provided it is large enough to cause the value to be censored, for then its exact value is ignored.

6.2.3.2. Appropriate Regional Skew

[66] The use of contaminated distributions and the general belief that flood records are not truly P3 distributed raises concerns regarding the value of the regional skewness coefficients. To reduce the uncertainty in sample skew estimates, B17 recommends weighting the sample skew with the regional skew using equation (1). In the absence of low outliers, the regional skew is weighted with an unadjusted sample skew estimate obtained using method of moments. If low outliers are identified, then B17 recommends adjusting the sample moments of the flood record using CPA. The adjusted sample skew produced by CPA is then weighted with the regional skew to obtain a weighted skew estimate for use in the final fitted P3 distribution. Therefore, if we truly believe that low outlier adjustment is necessary and appropriate to improve quantile estimates at the upper end of the distribution, then the regional skew should be estimated using samples which have been appropriately adjusted following the identification of low outliers. For a study in South Carolina, *Feaster and Tasker* [2002, p. 14] observed that the computed regional skewness coefficients were not sensitive to high and low outliers.

[67] The assumption that regional skew estimates are obtained from adjusted samples is utilized in the Monte Carlo analysis in the application of CPA and PPR. This

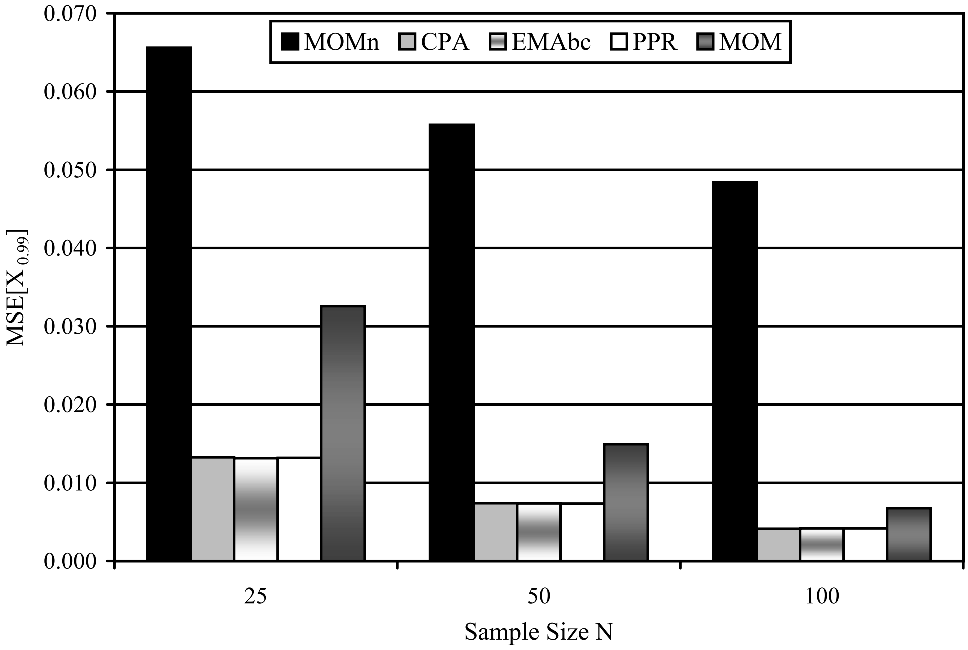

Figure 7. MSE of $X_{0.99}$ estimators for each method in samples with k = 1, 2, and 3 contaminated observations (G = 0, Var[γ] = 0.100).

same assumption is then appropriate when the regional skew is weighted with the sample skew obtained using method of moments (MOM) procedures in the absence of low outliers. In this Monte Carlo analysis, MOM is also used to illustrate the value of low outlier adjustments and uses the same regional skew as the other methods without the deletion of low outliers.

6.2.3.3. Simulation Results

[68] Comparisons of quantile estimates with the contaminated samples were made for all combinations of sample size (N = 25, 50, and 100), regional skew, and population skew variance using M = 5000 replicates. Figure 7 compares the MSE of the $X_{0.99}$ estimators as a function of sample size with a regional skew of 0 and Var[γ] = 0.100. These results are typical of the $X_{0.99}$ estimators for all combinations of sample size, regional skew, and population skew variance considered [see *Griffis*, 2003]. The results for CPAc and MOMc are not shown in Figure 7 because they are equivalent to CPA and MOM, respectively, as the constraints are never binding with a regional skew of 0 and Var[γ] ≤ 0.100.

[69] Table 2 reports the frequencies with which the lower bound on the weighted skew and the constraint on the upper

Table 2. Frequencies (%) of Invoking Weighted Skew and Upper Bound Constraints in Samples With k = 1, 2, and 3 Observations Contaminated

Sample Size	Regional Skew	$\tilde{G} < -1.4$			$\hat{\tau} < x_{max}$				
		CPAc	EMAbc	MOMc	MOMn	CPAc	EMAbc	MOMc	PPR
				Population Skew Variance = 0.010					
25	−1.0	0.0	0.0	0.0	92.7	4.0	1.7	0.0	1.1
50	−1.0	0.0	0.0	0.0	99.1	4.3	1.4	0.0	1.3
100	−1.0	0.0	0.0	0.0	100.0	2.5	1.4	0.0	1.1
				Population Skew Variance = 0.100					
25	−1.0	0.0	0.0	0.0	92.1	0.0	0.0	0.9	0.0
50	−1.0	0.0	0.0	0.0	98.9	0.0	0.3	2.7	0.1
100	−1.0	0.0	0.0	0.0	99.8	0.0	0.9	10.3	0.5
100	−0.5	0.0	0.0	0.0	99.4	0.0	0.0	1.7	0.0
				Population Skew Variance = 0.302					
25	−1.0	0.0	0.0	0.0	90.1	0.0	0.0	4.7	0.0
25	−0.5	0.0	0.0	0.0	96.1	0.0	0.0	0.1	0.0
50	−1.0	0.1	0.7	73.1	97.3	0.0	0.3	12.8	0.0
50	−0.5	0.0	0.0	0.0	94.3	0.0	0.0	5.0	0.0
50	−0.2	0.0	0.0	0.0	91.0	0.0	0.0	0.5	0.0
100	−1.0	0.0	2.0	87.7	97.9	0.1	1.2	22.7	0.2
100	−0.5	0.0	0.0	0.0	95.8	0.0	0.0	30.7	0.0
100	−0.2	0.0	0.0	0.0	91.9	0.0	0.0	16.9	0.0
100	0.0	0.0	0.0	0.0	88.4	0.0	0.0	9.8	0.0
100	0.2	0.0	0.0	0.0	80.9	0.0	0.0	3.3	0.0

bound are invoked by each fitting method. Except for the MOMn results, the frequencies for sample size-regional skew combinations not included in the tables are zero (in 5000 replicate samples). With an informative regional skew (Var[γ] $\leq$ 0.100), the lower bound on the skew is never binding and the upper bound constraint is binding only in samples with a regional skew of -1.0 (an extreme case), except for $N = 100$ where a skew of -0.5 with a variance of 0.100 also generated upper bound constraint violations for MOMc. Although PPR was not constrained, the frequencies with which the computed upper bound fell within the sample data are also reported.

[70] Low outlier adjustment dramatically reduces the MSE of the $X_{0.99}$ estimators in contaminated samples. In Figure 7 for a regional skew of 0 with a population skew variance of 0.100, the large differences between the MSE of the MOM estimator and the MSE of CPA, PPR, and EMA estimators illustrates the advantage of adjusting for true low outliers. In samples of size 25, the MSE is reduced an average of 60% by adjusting for low outliers versus using MOM. The value of adjusting for low outliers decreases with sample size, but the reduction is still tremendous; the MSE is reduced an average of 40% with outlier adjustment in samples of 100. In all cases considered, the differences in the MSEs of the $X_{0.99}$ estimators using the low outlier adjustment methods are modest. As the contaminated samples model the character of some real data sets, these results indicate that adjustment for low outliers is advantageous when fitting P3 distributions to annual maximum flood series.

6.2.4. Excess Censoring From P3 Distributed Samples

[71] Following B17 recommendations, the low outlier threshold X_L is defined by the 10% censoring level across samples of normal data. For samples of size $N = 25$, when one observation in 10 samples is censored, this corresponds to censoring the 0.4th percentile. It is likely that the differences in the performance of the estimation methods are small because low outliers were identified in few samples, and a relatively small fraction of an individual sample was censored. In section 6.2.1 for P3 distributed sample sizes of 25, 50, and 100, the maximum number of low outliers identified were 2, 3, and 4, respectively, with extreme regional skews of -1.0. These correspond to 8%, 6%, and 4% of the sample values being censored. *Kroll and Stedinger* [1996] compare several quantile estimation methods, including a lognormal MLE and a log PPR, using censored data. They found there was little difference between the MSEs of the MLE and PPR estimators when censoring at or below the 20th percentile for $N = 10$, the 40th percentile for $N = 25$, and the 60th percentile for $N = 50$.

[72] To see if increased censoring would result in more noticeable differences in the performances of the low outlier adjustment procedures, the Monte Carlo analysis was conducted using a censoring threshold at the 20th percentile of the flood distribution. A higher censoring level was not used because CPA for low outlier adjustment is not recommended by B17 when more than 25% of the sample is censored. Data were generated from P3 populations using the method described in section 6.1 for sample sizes 25, 50, and 100. The analysis is the same as that in section 6.2.1; all that is changed is the censoring threshold.

[73] Low outliers were identified using equation (5) as recommended by B17. Use of this equation requires an estimate of the frequency factor. Equation (6) assumes a normal distribution and thus is only dependent on the sample size. The value of the frequency factor increases with sample size. This results in a smaller truncation level X_L as defined by equation (5), and thus a smaller percentile for the censoring threshold as N increases. Results of *Kroll and Stedinger* [1996] indicate that this is wrong: As N increases, an increasing fraction of the sample should be censored because there are still enough observations to obtain a good estimator. It is also reasonable that more outliers would be identified in samples with larger skew values, and thus the estimate of the frequency factor should depend on the skew, as suggested by *Spencer and McCuen* [1996].

[74] In order to censor 20% of the sample, a standard P3 frequency factor defined by the 20th percentile was estimated as a function of the regional skew in place of equation (6). Thus the censoring threshold was defined to be

$$X_L = \bar{X} - S_x K_{0.20}(G). \tag{22}$$

[75] Comparisons of quantile estimates were made for all combinations of sample size ($N = 25$, 50, and 100), regional skew, and population skew variance using $M = 5000$ replicates. Figure 8 illustrates the MSE of the $X_{0.99}$ estimators of all seven fitting methods for samples of 25 and a population skew variance of 0.100. Figure 9 illustrates the MSE of the $X_{0.99}$ estimators as a function of sample size and population skew variance with a regional skew value of 0. The results for CPAc and MOMc are not shown in Figure 9 because they are equivalent to CPA and MOM, respectively, because the constraints are never binding with a regional skew of 0. Figure 10 illustrates the bias of the $X_{0.99}$ estimators in samples of size 25 with a population skew variance of 0.100. (*Griffis* [2003] provides results for all combinations of sample size, regional skew, and population skew variance.)

[76] Table 3 reports the frequencies with which the lower bound on the weighted skew and the constraint on the upper bound are invoked by each fitting method for specified sample sizes, regional skew values, and population skew variances. Except for the MOMn results, the frequencies for sample size-regional skew combinations not included in the table are zero (in 5000 replicate samples); the constraints were binding only in samples with a regional skew of -1.0 (an extreme case), except for $N = 100$ with a population skew variance of 0.302 (an extreme case) where a regional skew of -0.5 also generated constraint violations. Violation of the lower bound constraint on the weighted skew only occurred with Var[γ] = 0.302 (again an extreme case). Although PPR was not constrained, the frequency with which the computed upper bound fell within the sample data is also reported.

[77] The results presented in this section for the MOMn, MOM, and MOMc estimators are the same as those from section 6.2.1 (Figures 1, 2, and 3), because the data are P3 distributed and these methods use all of the sample data and are thus unaffected by the censoring. Therefore the same observations regarding the benefit of weighting the sample skew with an informative regional skew pertain.

[78] Because MOM utilizes all of the sample data, one might suspect that it would result in the best performance

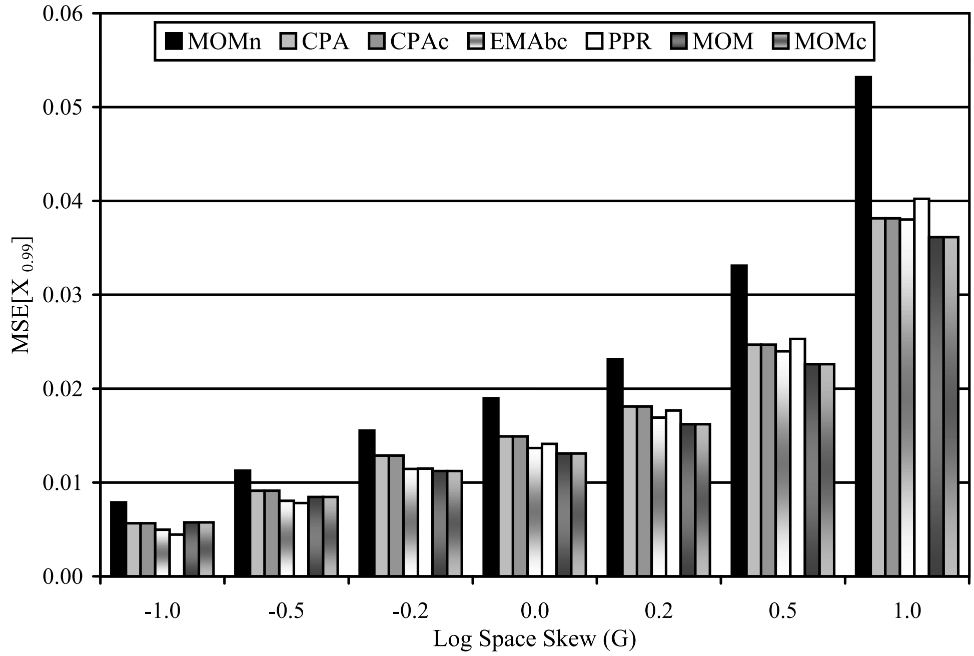

Figure 8. MSE of $X_{0.99}$ estimators for each method with 20% censoring from P3 distributed samples (N = 25, Var[γ] = 0.100).

and thus have the smallest MSE in this case, wherein all of the data are actually from a P3 distribution. In Figure 8 for samples of size 25 with a population skew variance of 0.100, the MOM estimators had the smallest MSE for $G \geq -0.2$, but the differences between the estimators that use a weighted skew decrease with the value of the regional skew. Similar results were observed in this skew range with an extreme population skew variance of 0.010 for all quantile estimators except CPA; with a highly informative regional skew G ≤ -0.5, CPA performed as poorly as MOMn, which did not employ regional skew information. For all sample sizes with Var[γ] $\leq$ 0.100, EMA outperforms CPA and PPR for $|G| \leq 0.2$; PPR generally outperforms CPA. Furthermore, as shown in Figure 10 for samples of size 25 with Var[γ] = 0.100, CPA is significantly more biased than the other methods.

[79] The differences in the estimators that use a weighted skew are generally minor in cases where the sample size is

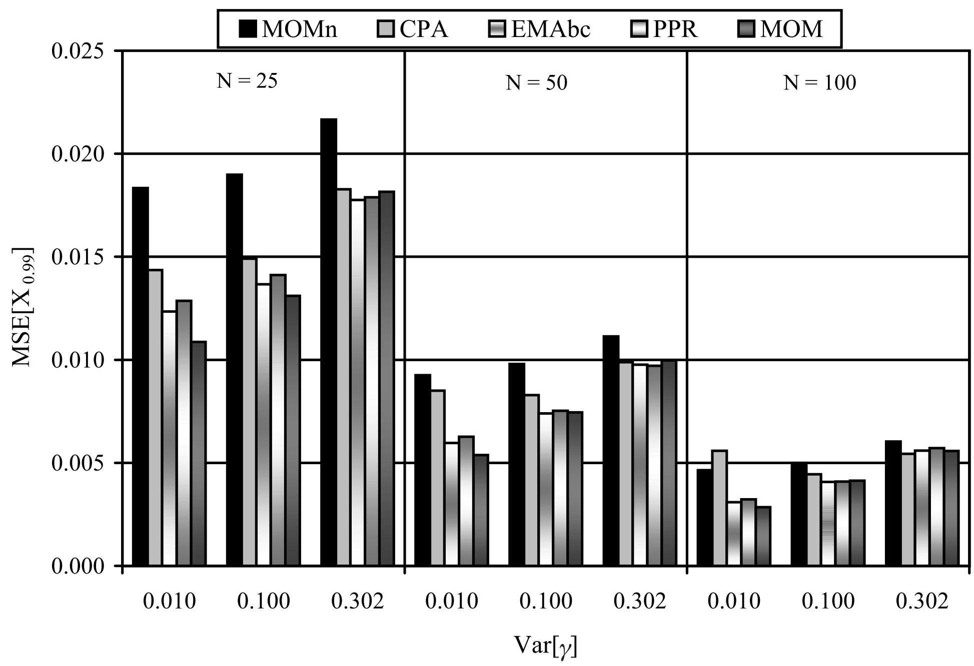

Figure 9. MSE of $X_{0.99}$ estimators for each method with 20% censoring from P3 distributed samples (G = 0).

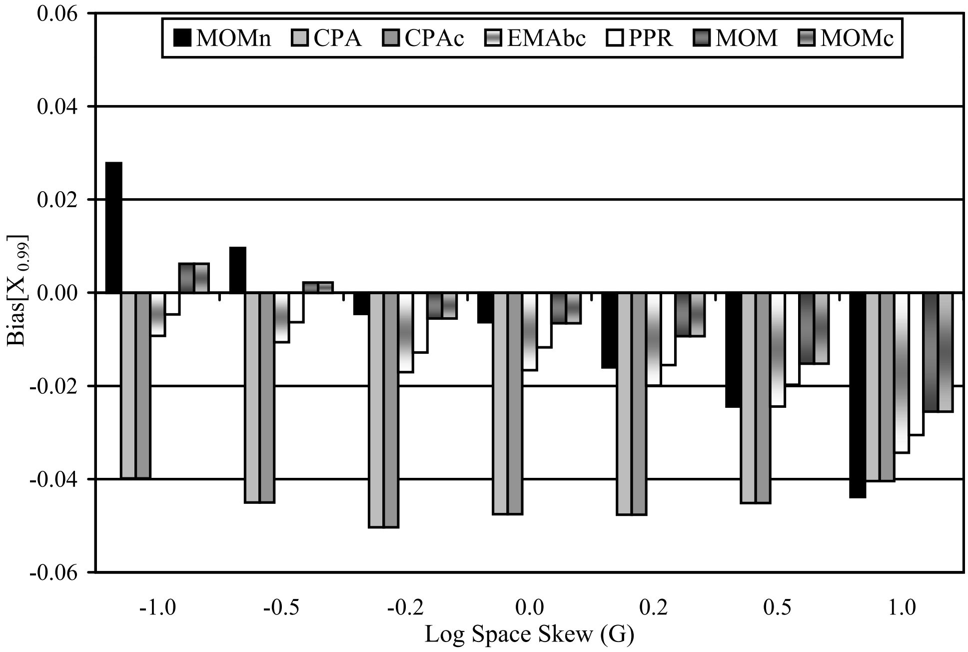

Figure 10. Bias of $X_{0.99}$ estimators for each method with 20% censoring from P3 distributed samples (N = 25, Var[γ] = 0.100).

roughly equivalent to or exceeds the effective record length of the regional skew. As shown in Figure 9 for a given Var[γ], the relative differences between the estimators decrease as the sample size increases. In Figure 9 for a regional skew of 0 with a variance of 0.100 ($e \approx 60$), the differences between the estimators are modest for N = 50 and 100. With a variance of 0.302 ($e \approx 20$), the differences are modest for all samples sizes. In general, these results are typical for $G \geq -0.5$ with these sample sizes and population skew variances.

[80] These results indicate that an appropriate outlier adjustment method such as EMA results in no loss of overall accuracy when significant amounts of data are censored from the sample, in our case 20%. For all sample sizes with informative regional skew coefficients of reasonable values, EMA provides the most accurate quantile estimates compared with the other outlier adjustment methods.

7. Conclusions

[81] Weighting the sample skew by an informative regional skew significantly reduces the MSE of flood quantile estimators. However, the differences between the MSE of the estimation techniques that either use the complete sample or which censor low outliers are generally modest, provided they employ an informative regional skew and a relatively small fraction of the sample is censored, as generally occurs with the *Bulletin 17B* outlier detection criterion. In these cases, there is no loss of efficiency when low outlier adjustments are employed in P3 distributed data. However, a higher censoring threshold would be advisable in that no loss of

Table 3. Frequencies (%) of Invoking Weighted Skew and Upper Bound Constraints With 20% Censoring From P3 Distributed Samples

Sample Size	Regional Skew	$\tilde{G} < -1.4$ CPAc	EMAbc	MOMc	$\hat{\tau} < x_{max}$ MOMn	CPAc	EMAbc	MOMc	PPR
				Population Skew Variance = 0.010					
25	−1.0	0.0	0.0	0.0	8.3	12.1	1.8	1.7	0.9
50	−1.0	0.0	0.0	0.0	12.1	14.8	1.8	1.2	1.0
100	−1.0	0.0	0.0	0.0	14.0	9.0	1.5	1.0	0.9
				Population Skew Variance = 0.100					
25	−1.0	0.0	0.0	0.0	8.8	0.0	0.0	0.1	0.0
50	−1.0	0.0	0.0	0.0	11.5	0.0	0.3	0.6	0.3
100	−1.0	0.0	0.0	0.0	14.3	0.0	0.5	1.4	1.1
				Population Skew Variance = 0.302					
25	−1.0	0.0	0.0	0.0	8.6	0.0	0.0	0.2	0.0
50	−1.0	0.0	0.0	2.9	12.4	0.0	0.0	0.7	0.0
100	−1.0	0.0	0.0	6.1	16.5	0.0	0.2	2.6	0.3
100	−0.5	0.0	0.0	0.0	6.8	0.0	0.0	0.3	0.0

precision is likely with LP3 data using an efficient and flexible estimation procedure such as EMA, whereas substantially greater protection is provided against low outliers.

[82] Significant reductions in the MSE are observed in contaminated samples when low outlier adjustment procedures are used in addition to an informative regional skew. Therefore identification and adjustment for low outliers is advisable in fitting P3 distributions to real annual maximum flood series. Regional skew estimates should also be derived using samples adjusted for low outliers.

[83] With respect to the different low outlier adjustment estimators, EMA is more intuitively appealing than CPA and PPR. CPA is a relatively ad hoc approach which includes inconsistencies in its representation of the data. EMA and PPR both provide a direct fit to the entire data set. PPR assigns specific values to the low outliers, which may be larger than the smallest retained observation, which could be difficult to justify. EMA estimates the expected contributions of the low outliers to the sample moments in a very reasonable and consistent framework. *Griffis* [2003] compares these methods using real data sets.

[84] EMA has already proven efficient for use with historical information and can provide good confidence intervals for estimated quantiles, which CPA and PPR currently do not. EMA has been successfully generalized to include regional skew information and to adjust for low outliers. EMA is more flexible than CPA and PPR and works as well as or better than CPA at low censoring levels and better than CPA at higher censoring levels. EMA is also attractive because it is a logical extension of the method of moments recommended by B17. Thus it maintains the same statistical structure and allows for easy implementation within the existing B17 framework.

[85] The authors of this paper attempted to stay within the guidelines of *Bulletin 17B* as much as possible in hopes that the EMA algorithm would be adopted by the B17 community. As currently formulated, EMA has been shown to be as good or better than B17's CPA approach for handling low outliers and more efficient than the current B17 method for using historical data. Areas of future research could include evaluation of EMA's performance when low outlier adjustments are required and both historical and regional skew information are employed, though the effort hardly seems necessary.

Appendix A: Derivation of $E[(X_L^< - \mu)^m]$

[86] To update parameter estimates, the EMA for low outlier adjustment uses the expected contribution of the below-threshold values in the computations of the second and third moments. These expected contributions are computed using equation (17). The expectation $E[(X_L^< - \mu)^m]$, where μ is the true mean and m is the moment being evaluated, can be derived using the P3 probability density function $f_x(X)$ [*Cohn et al.*, 1997]:

$$E\left[\left(X_L^< - \mu\right)^m \mid \alpha, \beta, \tau\right] = E[(X-\mu)^m \mid X < X_c, \alpha, \beta, \tau]. \quad \text{(A1)}$$

Denoting $E[(X-\mu)^m \mid X < X_c, \alpha, \beta, \tau] = E[(X-\mu)^m]$, then for $m = 1$

$$E[(X-\mu)] = E[(X-\tau) + (\tau-\mu)] = E[X-\tau] + E[\tau-\mu], \quad \text{(A2)}$$

where

$$E[\tau-\mu] = \tau-\mu. \quad \text{(A3)}$$

If $E[X - \tau] = E(y)$, then the expected value can be computed as

$$E(y) = \int_0^t y f_x(y) dy = \frac{1}{\Gamma(t,\alpha)} \int_0^t \beta^{-\alpha}(y)\left(y^{\alpha-1}\right) \exp\left(-\frac{y}{\beta}\right) dy$$
$$= \frac{\beta}{\Gamma(t,\alpha)} \int_0^t (z^\alpha) \exp(-z) dz, \quad \text{(A4)}$$

where $z = y/\beta = (X - \tau)/\beta$. By definition of the incomplete gamma function [*Cohn et al.*, 1997], equation (A4) reduces to

$$E(y) = \frac{\beta\Gamma(t, \alpha+1)}{\Gamma(t,\alpha)}, \quad \text{(A5)}$$

where $t = (X_c - \tau)/\beta$. Substitution of equation (A3) and (A5) into (A2) yields

$$E[(X-\mu) \mid X < X_c, \alpha, \beta, \tau] = (\tau-\mu) + \beta \frac{\Gamma\left[\frac{X_c-\tau}{\beta}, \alpha+1\right]}{\Gamma\left[\frac{X_c-\tau}{\beta}, \alpha\right]}. \quad \text{(A6)}$$

$E[(X_L^< - \mu)^m]$ can be similarly derived for $m = 2$ and 3 [*Griffis*, 2003].

[87] **Acknowledgments.** We greatly acknowledge support provided by a Water Resources Institute Internship Award 02HQGR0128 by the U.S. Geological Survey, U.S. Department of the Interior. Comments by David Goldman and other reviewers are greatly appreciated and substantially improved the manuscript.

References

Bobée, B. (1973), Sample error of *T*-year events computed by fitting a Pearson type 3 distribution, *Water Resour. Res.*, *9*(5), 1264–1270.
Chowdury, J. U., and J. R. Stedinger (1991), Confidence interval for design floods with estimated skew coefficient, *J. Hydraul. Eng.*, *117*(7), 811–831.
Cohn, T. A., W. L. Lane, and W. G. Baier (1997), An algorithm for computing moments-based flood quantile estimates when historical information is available, *Water Resour. Res.*, *33*(9), 2089–2096.
Cohn, T. A., W. L. Lane, and J. R. Stedinger (2001), Confidence intervals for expected moments algorithm flood quantile estimates, *Water Resour. Res.*, *37*(6), 1695–1706.
David, H. A. (1980), *Order Statistics*, 2nd ed., John Wiley, Hoboken, N.J.
England, J. F. Jr., R. D. Jarrett, and J. D. Salas (2003a), Data-based comparisons of moments estimators using historical and paleoflood data, *J. Hydrol.*, *278*(4), 172–196.
England, J. F. Jr., J. D. Salas, and R. D. Jarrett (2003b), Comparisons of two moments-based estimators that utilize historical and paleoflood data for the log Pearson type III distribution, *Water Resour. Res.*, *39*(9), 1243, doi:10.1029/2002WR001791.
Feaster, T. D., and G. D. Tasker (2002), Techniques for estimating the magnitude and frequency of floods in rural basins of South Carolina, 1999, *U.S. Geol. Surv. Water Resour. Invest. Rep., 02-4140*, 50 pp.
Gilliom, R. J., and D. R. Helsel (1986), Estimation of distributional parameters for censored trace level water quality data: 1. Estimation techniques, *Water Resour. Res.*, *22*(2), 135–146.
Griffis, V. W. (2003), Evaluation of log-Pearson type 3 flood frequency analysis methods addressing regional skew and low outliers, M.S. thesis, Dep. of Civ. and Environ. Eng., Cornell Univ., Ithaca, N. Y.

Hardison, C. H. (1974), Generalized skew coefficients of annual floods in the United States and their application, *Water Resour. Res.*, *10*(4), 745–751.

Helsel, D. R., and T. A. Cohn (1988), Estimation of descriptive statistics for multiple censored water quality data, *Water Resour. Res.*, *24*(12), 1997–2004.

Hirsch, R. M., and J. R. Stedinger (1987), Plotting positions for historical floods and their precision, *Water Resour. Res.*, *23*(4), 715–727.

Interagency Committee on Water Data (IACWD) (1982), Guidelines for determining flood flow frequency, *Bull. 17B*, 28 pp., Hydrol. Subcomm.,Washington, D. C., March.

Jarrett, R. D., and E. M. Tomlinson (2000), Regional interdisciplinary paleoflood approach to assess extreme flood potential, *Water Resour. Res.*, *36*(10), 2957–2984.

Jennings, M. E., and M. A. Benson (1969), Frequency curves for annual flood series with some zero events or incomplete data, *Water Resour. Res.*, *5*(1), 276–280.

Kottegoda, N. T., and R. Rosso (1997), *Statistics, Probability, and Reliability for Civil and Environmental Engineers*, McGraw-Hill, New York.

Kroll, C. N. (1996), Censored data analyses in water resources, Ph.D. dissertation, Dep. of Civ. and Environ. Eng., Cornell Univ., Ithaca, N. Y.

Kroll, C. N., and J. R. Stedinger (1996), Estimation of moments and quantiles with censored data, *Water Resour. Res.*, *32*(4), 1005–1012.

Landwehr, J. M., N. C. Matalas, and J. R. Wallis (1978), Some comparisons of flood statistics in real and log space, *Water Resour. Res.*, *14*(5), 902–920.

Martins, E. S. P. R., and J. R. Stedinger (2002), Efficient regional estimates of LP3 skew using GLS regression, paper presented at the 2002 Conference on Water Resources Planning and Management, Am. Soc. of Civ. Eng., Roanoke, Va., 19–22 May.

McCuen, R. H. (2001), Generalized flood skew: Map versus watershed skew, *J. Hydraul. Eng.*, *6*(4), 293–299.

National Research Council (NRC) (1999), Improving American river flood frequency analyses, report, Comm. on Am. River Flood Frequencies, Water Sci. and Technol. Board, Natl. Acad., Washington, D. C.

Reis, D. S. Jr., J. R. Stedinger, and E. S. Martins (2003), Bayesian GLS regression with application to LP3 regional skew estimation, in *Proceedings, World Water and Environmental Resources Congress 2003*, edited by P. Bizier and P. DeBarry, Am. Soc. of Civ. Eng., Reston, Va., June 23–26.

Spencer, C. S., and R. H. McCuen (1996), Detection of outliers in Pearson type 3 data, *J. Hydrol. Eng.*, *1*(1), 2–10.

Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou (1993), Frequency analysis of extreme events, in *Handbook of Hydrology*, chap. 18, pp. 18.1–18.66, McGraw-Hill, New York.

Tasker, G. D. (1978), Flood frequency analysis with a generalized skew coefficient, *Water Resour. Res.*, *14*(2), 373–376.

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

Thomas, W. O. Jr. (1985), A uniform technique for flood frequency analysis, *J. Water Resour. Plann. Manage.*, *111*(3), 321–337.

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

Water Resources Council (1976), Guidelines for determining flood flow frequency, *Bull. 17*, Hydrol. Comm., Washington, D. C.

Whitley, R., and T. V. Hromadka (1999), Approximate confidence intervals for design floods for a single site using a neural network, *Water Resour. Res.*, *35*(1), 203–209.

T. A. Cohn, U.S. Geological Survey, 12201 Sunrise Valley Drive, Reston, VA 20192, USA. (tacohn@usgs.gov)

V. W. Griffis and J. R. Stedinger, School of Civil and Environmental Engineering, Cornell University, Hollister Hall, Ithaca, NY 14853-3501, USA. (vlw7@cornell.edu; jrs5@cornell.edu)