Comparative Evaluation of Three Estimators of Log Pearson Type 3 Distribution

BABAK NAGHAVI, JAMES F. CRUISE, AND KISHORE ARORA

Three moment-based estimation procedures for log Pearson Type 3 (LP3) parameters were compared using observed stream data samples from Louisiana and its neighboring states. The methods of direct moments, log-transformed moments, and mixed moments were compared in descriptive capabilities on the basis of computed root mean square deviation (RMSD) and mean absolute deviation (MAD) of the standardized variate. Using these performance indices, the most robust estimation method was sought. In many cases, depending on sample skewness, significant differences existed between the descriptive capabilities of these methods. However, no method performed in a clearly superior manner across the entire range of data. These results can be used in conjunction with previous Monte Carlo studies focusing on the predictive ability of these procedures to determine the most reliable moment-based estimation procedures.

The log Pearson Type 3 (LP3) distribution is one of the most widely used distributions in hydrology, particularly in flood frequency analysis, as recommended by many governmental agencies in the United States. Many highway drainage structures are designed under the assumption that flood discharges follow this distribution. The LP3 distribution was first recommended by the U.S. Water Resources Council (WRC) in 1967 as the base method of flood frequency analysis in the United States. Since then, a great deal of interest has been generated in this distribution. The LP3 distribution has been extensively discussed by Bobee (*1*), Bobee and Robitaille (*2*), Condie (*3*), Rao (*4*), and many others.

Much attention has been focused on parameter estimation. Bobee (*1*) suggested an estimation method that was based on the first three moments of raw data, the method of direct moments (MDM), whereas the WRC (*5*) recommended an estimation method that was based on the corresponding moments of the log-transformed data. Condie (*3*) proposed an estimation method that was based on maximum likelihood estimation (MLE) theory. Rao (*4*) proposed the method of mixed moments (MIX), which uses sample estimates of the means of the raw and log-transformed data and the standard deviation of the raw data in estimating the parameters. Singh and Singh (*6*) used the principle of maximum entropy (POME) to estimate the parameters for the LP3 distribution. Arora and Singh (*7*) and Ashkar and Bobee (*8*), among others, compared performance of various methods of parameter estimation via Monte Carlo simulation. In terms of root mean square error (RMSE) and bias, Arora and Singh (*7*) found that the MIX and MDM methods were clearly superior to the WRC method for simulated LP3 samples. Ashkar and Bobee (*8*) compared four versions of the method of moments and observed that the MDM method performed better than the other three. More recently, Bobee and Ashkar (*9*) studied several variations of the method of moments and concluded that different versions of the methodology could result in significantly different fits to the data series. They also concluded that no one version of moment-based methods can be considered best for all applications.

MLE of the LP3 parameters has been found to be computationally difficult and results in multiple roots of the location parameter (*10*). Arora and Singh (*10*) also found that MLE performed poorly in terms of RMSE and bias in comparison with moment-based methods on the basis of small Monte Carlo–generated samples. A large amount of CPU time was required by the search routines for MLE estimation of LP3 parameters. Thus, MLE techniques do not appear to be well suited to the estimation of parameters of the LP3 distribution. In the case of the LP3 distribution, the maximum entropy procedure and the MIX method lead to the same parameter estimation equations (*11*). Therefore, moment-based methods may be the most practical and computationally efficient procedures for estimating LP3 parameters and quantiles.

The MDM estimates the parameters of the LP3 distribution directly from the untransformed data. In this method, the observed data are equally weighted in the estimation of the parameters. Thus, this procedure maintains the significance of the larger sample values because the spatial relationship among the real data is preserved.

Conversely, the WRC method weights the logarithms of the observed data equally in parameter estimation. Therefore, in this method, the larger sample values are given less significance because of the transformation into log space before the sample statistics were computed. The WRC method has been criticized because of the sampling properties of the coefficient of skewness. This statistic has been shown to be significantly downward biased (*12*) and algebraically bounded (*13*) and possesses a large sampling variability (*14*). Studies by Wallis and Wood (*15*), Arora and Singh (*7*), and Ashkar and Bobee (*8*) have reported the poor performance of the WRC method on the basis of Monte Carlo analyses. The method of direct moments also requires the estimation of the third moment from the data sample.

B. Naghavi, Louisiana Transportation Research Center, 4101 Gourrier Avenue, Baton Rouge, La. 70808. J. F. Cruise and K. Arora, Department of Civil Engineering, Louisiana State University, Baton Rouge, La. 70803.

The MIX method combines the moment equations in real and log-transformed space in parameter estimation. MIX avoids the use of sample statistics (such as skew) based on third-moment estimates that are susceptible to large sampling errors.

Previous comparisons were based on Monte Carlo simulations in which the data were generated from known distributions, typically the LP3. However, in real-world situations, the population distributions are unknown. Therefore, a comparison of the most popular parameter estimation procedures using real data seems timely. Although it is recognized that a good fit to observed data is not a sufficient reason for accepting a particular method, an adequate fit to the observed data is a necessary condition for the acceptance of a procedure. Cunnane (*16*) compares the relative importance of predictive and descriptive abilities of flood prediction techniques. He concludes that neither attribute is more important than the other; indeed, the two characteristics are complementary. In order for a particular technique to be useful, it must possess both predictive and descriptive abilities. Predictive capabilities are usually determined from Monte Carlo studies of a particular method or distribution, whereas descriptive capabilities can be determined from analyses based on real-world data, with the added advantage of unknown population distributions. Thus, studies such as the one reported here can be used in conjunction with the Monte Carlo studies previously reported to aid in the selection of the most reliable estimation technique for LP3 parameters and quantiles. If events of small recurrence intervals ($T \leq 25$ years) are to be estimated, the method with the superior descriptive capability may be preferred, because events of this magnitude will usually already be recorded in the systematic record. In highway drainage work, many times structures are designed for small recurrence intervals whose quantiles may already have been recorded. In these cases, the interpretive ability of the method may be of paramount importance. However, if events of larger recurrence intervals are to be estimated, then some descriptive ability may be sacrificed to obtain improved predictive ability. In this study, three moment-based methods (WRC, MIX, and MDM) are compared, using gauge stations in Louisiana (87 stations) and its neighboring states (6 stations) with unknown flood distributions.

PROPERTIES OF LOG PEARSON TYPE 3 (LP3) DISTRIBUTION

The probability density of the LP3 is

$$f(x) = \frac{1}{|a|x\Gamma(b)}\left[\frac{\ln(x-c)}{a}\right]^{b-1} \exp\left[-\frac{\ln(x-c)}{a}\right] \quad (1)$$

where

x = raw (untransformed) flood data,
a = LP3 scale parameter,
b = LP3 shape parameter, and
c = LP3 location parameter.

The parameter b is always positive and Γ is the gamma function. LP3 density function is flexible and can take many different forms. The mean, variance, and skewness coefficient of the variate $y = \ln(x)$ are given by

Mean

$$\mu = c + ab \quad (2)$$

Variance

$$\sigma^2 = ba^2 \quad (3)$$

Skew

$$\gamma = \frac{|a|}{a}\frac{2}{b^{1/2}} \quad (4)$$

The moments of x about the origin are given by (*1*)

$$\mu_r' = \frac{\exp(rc)}{(1-ra)^b} \qquad 1 - ra > 0,\ r = 1, 2, \text{ and } 3 \quad (5)$$

If $a > 0$, then $\gamma_y > 0$; therefore y must be positively skewed such that $f(y)$ is lower bounded ($c \leq y < +\infty$). In this case, x must also be positively skewed, thus x also possesses a lower bound [$\exp(c) \leq x < +\infty$] (*4*). When $a < 0$, then $\gamma_y < 0$ such that y is negatively skewed and upper bounded, that is, $-\infty < y \leq c$. In this case, x either can be positively or negatively skewed, depending on the values of the parameters a and b, but x is upper bounded [$0 < x < \exp(c)$]. For this case, the density function $f(x)$ may be defined as zero at $x = 0$ (*4*).

The overall geometric shape of the LP3 distribution is governed by the parameters a and b (*1,4*). The LP3 distribution degenerates to the log normal distribution when the parameters a and b approach zero and infinity, respectively.

Fitting the LP3 Distribution by the Method of Logarithmic Moments (WRC)

This method estimates a, b, and c by applying the method of moments to the log-transformed data. Equations 2–4 are used for estimating the parameters where μ, σ^2, and γ are substituted by the mean, variance, and skewness coefficient estimates of the log-transformed sample.

Fitting the LP3 Distribution by the MIX Method

Rao (*4*) proposed the MIX method for LP3 with the objective of avoiding use of the sample skewness coefficient in parameter estimation. The MIX method preserves the sample mean and variance of raw data ($\bar{x}$, S_x^2) and sample mean of the log-transformed data ($\bar{y}$). The MIX parameter estimation equations are

$$\bar{y} = c + ab \quad (6)$$

$$\bar{x} = \frac{\exp(c)}{(1-a)^b} \quad (7)$$

$$S_x^2 = \exp(2c)\left[\frac{1}{(1-2a)^b} - \frac{1}{(1-a)^{2b}}\right] \quad (8)$$

A method of solution of Equations 6–8 has been devised by Arora and Singh (*7*).

Eliminating c by combining Equations 6 and 7,

$$\bar{y} - \ln(\bar{x}) = b[a + \ln(1 - a)] \tag{9}$$

Again, c can be eliminated by combining Equations 7 and 8.

$$\ln\left(\frac{S_x^2 + \bar{x}^2}{\bar{x}^2}\right) = 2b\ln(1 - a) - b\ln(1 - 2a) \tag{10}$$

Combining Equations 9 and 10,

$$\frac{2\ln(1 - a) - \ln(1 - 2a)}{\ln(1 - a) + a} = P \tag{11}$$

where P can be found from sample estimates of $\bar{x}$, $\bar{y}$, and S_x^2:

$$P = \frac{\ln[(S_x^2 + \bar{x}^2)/\bar{x}^2]}{(\bar{y} - \ln\bar{x})} \tag{12}$$

The left-hand side of Equation 11 is defined for $a < ½$. Therefore, values of a can be found using a trial-and-error search method or by the Newton-Raphson iteration. Alternatively, a can be determined from interpolation of the a–P table given by Arora and Singh (*7*). Parameters b and c can then be estimated from Equations 9 and 6, respectively.

Fitting the LP3 Distribution by the MDM

The MDM applies the method of moments directly to the raw data to determine parameters a, b, and c. Substituting the first three sample moment estimates in Equation 5 yields three simultaneous equations:

$$\ln\mu_1' = c - b\ln(1 - a) \tag{13}$$

$$\ln\mu_2' = 2c - b\ln(1 - 2a) \tag{14}$$

$$\ln\mu_3' = 3c - b\ln(1 - 3a) \tag{15}$$

These equations are solved in a manner given in Arora and Singh (*7*) that is similar to the method proposed by Bobee (*1*). Equations 13–15 can be rearranged to give

$$\frac{\ln\mu_3' - 3\ln\mu_1'}{\ln\mu_2' - 2\ln\mu_1'} = \frac{3\ln(1 - a) - \ln(1 - 3a)}{2\ln(1 - a) - \ln(1 - 2a)} \tag{16}$$

The right-hand side of Equation 16 is defined for $a < ⅓$. In practice, B is obtained from the sample estimates of the first three moments about the origin:

$$B = \frac{\ln\mu_3' - 3\ln\mu_1'}{\ln\mu_2' - 2\ln\mu_1'}$$

With B calculated, the value of a follows from Equation 16 using a trial-and-error search method, Newton-Raphson iteration, or interpolation of the a–B table given by Bobee (*1*) or Arora and Singh (*7*). Parameters b and c can then be estimated from Equations 13 and 14.

COMPARATIVE EVALUATION OF THE THREE METHODS

A total of 114 gauge stations with 20 years or more of record were initially available for use in this comparative study. Analysis of these data revealed that the records of 10 gauge stations were contaminated by diversions, regulation, backwater, etc., and thus were eliminated from further analysis. The pertinent data for the 93 remaining stations are presented in Table 1. Sample skew of the untransformed data varied from −0.40 to 6.18, and sample coefficient of variation varied from 0.29 to 1.75. This range represents a fairly broad range in the statistical characteristics of the available data samples. Although all of the data are drawn from one region of the United States, because of the range in skewness of the data base the results may hold significance for other regions.

Grubbs and Beck outlier analysis (*17*) at 10 percent significance level ($\alpha = 0.10$) was conducted, and 10 stations with single outliers were identified. Data with and without the outlier were analyzed for these sites. The performance of the three methods was evaluated using performance indices similar to those used by Singh and Singh (*6*), among others. These are standardized root mean square deviation (SRMSD) given by

$$\text{SRMSD} = \left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{x}_i - x_i}{\bar{x}}\right)^2\right]^{1/2} \tag{17}$$

and standardized mean absolute deviation (SMAD) given by

$$\text{SMAD} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{x}_i - x_i}{\bar{x}}\right| \tag{18}$$

where N is the sample size $(x_1, x_2, \ldots, x_N)$,

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i,$$

and $\hat{x}_i$ is an estimate of x_i obtained from $F^{-1}[p(x_i)]$; $p(x_i)$ is approximated by the Weibull plotting position: $p(x_i) = m_i/(N + 1)$, where m_i is the rank of x_i in descending order. The Weibull is an unbiased empirical estimate of the quantile probability and is the most widely used plotting position formula in hydrology. In this analysis, F^{-1} values were approximated numerically, and LP3 quantiles were obtained using a routine that translates the LP3 to a chi-squared random variate.

The performance indices used here are different from those used in the previous references in that the deviations between the predicted and observed variate are standardized by the sample mean rather than the observed value itself. In this way, every observed value of the variate is given equal weight in the computation of the performance index. Results of performance evaluation using SRMSD and SMAD are presented in Tables 2 and 3, respectively. The maximum percent differences were obtained from the differences between the

TABLE 1 PERTINENT DATA OF WATERSHEDS

Gage Station	Area km^2	No. of Obs.	Skew Coefficient	Coefficient of Variation
02491500	2,564	66	2.12	0.75
02492000	3,139	50	3.13	0.85
07344450	207	31	2.38	1.05
07348700	1,566	30	2.23	0.93
07349500	1,413	49	1.53	0.68
07351500	171	49	2.28	0.91
07352000	399	47	1.60	0.85
07351000	205	43	1.17	0.72
07366200	538	32	2.86	1.07
07371500	919	49	1.61	0.77
07372200	4,914	30	2.33	0.88
07373000	132	46	1.92	1.11
07375000	267	44	1.92	0.96
07375500	1,672	49	2.15	0.82
07376000	639	47	1.39	0.75
07376500	207	44	1.20	0.59
07377000	1,501	39	0.80	0.73
07377500	375	45	0.70	0.70
07378000	735	44	1.22	0.57
07378500	3,313	49	1.57	0.71
07381800	176	33	1.26	0.78
07382000	621	50	6.18	1.75
*07382000	621	49	2.89	0.70
08010000	339	49	0.77	0.49
08012000	1,364	49	2.53	0.72
*08012000	1,364	48	2.47	0.63
08013000	1,291	44	1.62	0.75
08013500	1,949	49	2.21	0.72
08014500	1,320	48	4.83	1.30
*08014500	1,320	47	1.30	0.74
08015500	4,399	49	3.43	0.84
*08015500	4,399	48	1.34	0.59
08014000	443	27	2.10	0.90
08025500	383	31	2.03	1.16
08028000	945	36	1.91	1.15
02490105	188	22	1.30	0.82
07375222	119	22	-0.18	0.56
07380160	52	33	0.40	0.48
07375170	228	20	1.18	0.63
07377300	2,290	35	1.27	0.61
07376600	36	32	0.07	0.37
07375480	236	20	1.61	0.85
02491700	114	20	1.56	0.95
02491350	109	21	1.39	0.85
07375800	232	32	2.62	1.11
07375307	135	22	1.61	1.06
07373500	91	21	0.73	0.61
07364300	702	24	2.18	0.94
07369500	800	52	0.05	0.29
07386500	49	28	0.80	0.43
08011800	114	24	1.21	0.67
08014200	244	37	3.39	1.06
07353500	122	26	1.73	1.11
07372500	238	31	4.16	1.20
*07372500	238	30	1.46	0.64
07370750	123	30	2.28	0.86
07372110	62	23	2.24	1.22
07372000	1,694	42	0.92	0.67
07370500	702	30	1.32	0.80
07370000	2,025	60	0.49	0.40
07367250	23	20	1.88	1.10
07366403	1	22	1.95	1.03
07366420	293	22	3.16	1.20
07365000	919	28	1.68	0.80
07364870	122	22	1.47	0.74
07365500	461	30	4.02	1.26

TABLE 1 (*continued on next page*)

TABLE 1 (continued)

Gage Station	Area km²	No. of Obs.	Skew Coefficient	Coefficient of Variation
*07365500	461	29	1.47	0.70
07366000	1,197	43	3.37	1.11
07364700	365	22	2.59	1.43
08016600	213	38	1.03	0.63
08028700	34	26	2.87	0.70
*08028700	34	25	1.26	0.46
08014600	68	20	1.63	0.88
08013800	27	21	1.46	0.77
08013610	1	22	-0.40	0.30
07354000	55	30	0.39	0.49
07353990	97	22	1.54	1.01
08016800	458	31	2.69	0.81
08016400	383	39	1.62	0.73
08015000	616	31	1.66	0.97
07352500	1,096	43	1.08	0.70
02490000	31	20	1.90	1.01
07348725	86	22	0.70	0.63
07348800	173	24	2.02	0.88
07347000	300	25	2.37	0.50
*07347000	300	24	0.79	0.34
07362100	997	49	3.01	1.03
07364190	3,030	45	-0.36	0.35
07365800	466	29	3.89	1.68
*07365800	466	28	4.02	1.17
07373550	1	30	0.61	0.41
08014800	311	24	1.50	0.74
08025850	25	20	2.30	0.98
08024060	8	24	0.09	0.43
08023000	250	28	1.39	0.74
07351700	50	26	4.60	1.53
*07351700	50	25	0.02	0.54
07368500	109	28	0.42	0.34
07364500	4,260	52	0.06	0.36
02492360	453	21	0.95	0.61
08031000	216	34	1.31	0.68
08030000	179	32	1.55	0.64
08029500	332	36	2.55	1.15

*Run without the outliers.

methods with the largest and smallest SRMSD (Table 2) and SMAD (Table 3).

DISCUSSION OF RESULTS

The most robust estimation technique in terms of descriptive ability was determined from among the moment-based procedures. The robust procedure is that which performs best across all variation in sample statistics. Kuczera (*18*) discusses two possible measures of robustness: minimax RMSD and minimum average RMSD. Based on the minimax criterion, the preferred estimator is the one whose maximum RMSD for all cases is minimum. The minimum average criterion is to select the estimator whose RMSD average over the test cases is minimum.

Table 4 presents the results of robustness studies using the two performance indices SRMSD and SMAD from Equations 17 and 18. The table shows the minimum, average, and maximum values of each performance index both for with and without (parenthesis) outlier cases. The results indicate that if the SRMSD index is preferred, the MDM is superior both under the average SRMSD and minimax SRMSD criteria. If the SMAD index is used, however, then the WRC method is the most robust estimator under both criteria. The table also shows that removal of the outliers has a large effect on the maximum values of the two indices, some effect on the averages, and of course, no effect on the minimum values.

From Table 2, in most cases when outliers were removed, the WRC method performed better than MDM and MIX by the SRMSD index. This result implies that the WRC method may be more sensitive to the presence of outliers than the other methods within the range of skewness characteristic of the data base.

Further analyses were performed by examining the performance of the different methods within particular ranges of

TABLE 2 SRMSD TEST RESULTS FOR THE THREE LP3 FITTING METHODS

Gage Station	SRMSD			Method(s) with Min. SRMSD	Max. Diff. %
	WRC	MIX	MDM		
02491500	0.208	0.210	0.211	WRC	1.4
02492000	0.307	0.296	0.296	MIX/MDM	3.7
07344450	0.328	0.354	0.326	MDM	8.6
07348700	0.250	0.281	0.272	WRC	12.4
07349500	0.135	0.141	0.144	WRC	6.7
07351500	0.200	0.137	0.137	MIX/MDM	46.0
07352000	0.100	0.138	0.135	WRC	38.0
07351000	0.142	0.124	0.126	MIX	14.5
07366200	0.446	0.437	0.423	MDM	5.4
07371500	0.233	0.198	0.200	MIX	17.7
07372200	0.253	0.260	0.261	WRC	3.2
07373000	0.178	0.208	0.187	WRC	16.9
07375000	0.155	0.202	0.198	WRC	30.3
07375500	0.182	0.199	0.201	WRC	10.4
07376000	0.107	0.126	0.126	WRC	17.8
07376500	0.080	0.093	0.095	WRC	18.8
07377000	0.148	0.113	0.104	MDM	42.3
07377500	0.178	0.115	0.089	MDM	100.0
07378000	0.104	0.104	0.107	WRC/MIX	2.9
07378500	0.122	0.139	0.141	WRC	15.6
07381800	0.106	0.130	0.121	WRC	22.6
07382000	1.096	1.150	1.004	MDM	14.5
*07382000	0.199	0.223	0.226	WRC	13.6
08010000	0.087	0.065	0.067	MIX	33.8
08012000	0.208	0.235	0.235	WRC	13.0
*08012000	0.185	0.199	0.203	WRC	9.7
08013000	0.144	0.151	0.156	WRC	8.3
08013500	0.181	0.178	0.182	MIX	2.2
08014500	0.664	0.640	0.613	MDM	8.3
*08014500	0.086	0.106	0.110	WRC	27.9
08015500	0.336	0.317	0.319	MIX	6.0
*08015500	0.121	0.126	0.127	WRC	5.0
08014000	0.278	0.302	0.292	WRC	8.6
08025500	0.328	0.383	0.293	MDM	30.7
08028000	0.252	0.288	0.214	MDM	34.6
02490105	0.197	0.218	0.188	MDM	16.0
07375222	0.179	0.137	0.095	MDM	88.4
07380160	0.075	0.069	0.064	MDM	17.2
07375170	0.145	0.164	0.155	WRC	13.1
07377300	0.103	0.120	0.120	WRC	16.5
07376600	0.046	0.048	0.048	WRC	4.3
07375480	0.168	0.218	0.218	WRC	29.8
02491700	0.153	0.219	0.229	WRC	49.7
02491350	0.159	0.198	0.179	WRC	24.5
07375800	0.259	0.349	0.326	WRC	34.7
07375307	0.247	0.283	0.232	MDM	22.0
07373500	0.101	0.109	0.102	WRC	7.9
07364300	0.263	0.282	0.281	WRC	7.2
07369500	0.030	0.030	0.028	MDM	7.1
07386500	0.132	0.108	0.104	MDM	26.9
08011800	0.119	0.141	0.146	WRC	22.7
08014200	0.412	0.410	0.400	MDM	3.0
07353500	0.164	0.234	0.214	WRC	42.7
07372500	0.604	0.643	0.589	MDM	9.2
*07372500	0.196	0.205	0.199	WRC	4.6
07370750	0.233	0.274	0.258	WRC	17.6
07372110	0.378	0.451	0.346	MDM	30.3
07372000	0.089	0.090	0.089	WRC/MDM	1.1
07370500	0.149	0.141	0.147	MIX	5.7
07370000	0.090	0.090	0.086	MDM	4.7
07367250	0.430	0.453	0.357	MDM	26.9
07366403	0.263	0.326	0.265	WRC	24.0
07366420	0.488	0.537	0.502	WRC	10.0
07365000	0.180	0.191	0.189	WRC	6.1

TABLE 2 (*continued on next page*)

TABLE 2 (*continued*)

Gage Station	SRMSD			Method(s) with Min. SRMSD	Max. Diff. %
	WRC	MIX	MDM		
07364870	0.223	0.193	0.195	MIX	15.5
07365500	0.596	0.651	0.593	MDM	9.8
*07365500	0.157	0.181	0.174	WRC	15.3
07366000	0.485	0.471	0.453	MDM	7.1
07364700	0.706	0.732	0.566	MDM	29.3
08016600	0.166	0.157	0.145	MDM	14.5
08028700	0.268	0.279	0.276	WRC	4.1
*08028700	0.115	0.116	0.120	WRC	4.4
08014600	0.252	0.284	0.256	WRC	12.7
08013800	0.161	0.173	0.171	WRC	7.5
08013610	0.084	0.081	0.064	MDM	31.3
07354000	0.063	0.066	0.063	WRC/MDM	4.8
07353990	0.216	0.259	0.207	MDM	25.1
08016800	0.308	0.300	0.298	MDM	3.4
08016400	0.139	0.164	0.160	WRC	18.0
08015000	0.178	0.220	0.201	WRC	23.6
07352500	0.159	0.143	0.125	MDM	27.2
02490000	0.235	0.290	0.295	WRC	25.5
07348725	0.157	0.136	0.128	MDM	22.7
07348800	0.205	0.245	0.235	WRC	19.5
07347000	0.178	0.174	0.177	MIX	2.3
*07347000	0.078	0.072	0.075	MIX	8.3
07362100	0.298	0.311	0.306	WRC	4.4
07364190	0.089	0.085	0.081	MDM	9.9
07365800	0.915	0.937	0.802	MDM	16.8
*07365800	0.641	0.623	0.612	MDM	4.7
07373550	0.067	0.068	0.071	WRC	6.0
08014800	0.145	0.173	0.178	WRC	22.8
08025850	0.334	0.384	0.334	WRC/MDM	15.0
08024060	0.092	0.088	0.089	MIX	4.5
08023000	0.130	0.158	0.161	WRC	23.8
07351700	0.976	0.967	0.895	MDM	9.1
*07351700	0.146	0.107	0.102	MDM	43.1
07368500	0.054	0.056	0.057	WRC	5.6
07364500	0.108	0.095	0.083	MDM	30.1
02492360	0.104	0.122	0.114	WRC	17.3
08031000	0.147	0.157	0.151	WRC	6.8
08030000	0.155	0.157	0.160	WRC	3.2
08029500	0.337	0.414	0.350	WRC	22.8

*Run without the outliers.

the skewness of the data samples. Tables 2 and 3 indicate that for the three samples that exhibit a negative skew coefficient, MDM is the preferred estimator by a significant amount in two of the three cases where MIX and WRC performed comparably. For the 19 cases that exhibit skew coefficients in the range 0 to 1.0, MDM is superior in 7 cases (by SRMSD) by a significant margin; WRC and MIX are preferred by a significant amount in one case each; whereas the methods perform about equally well in the other cases. These results are approximately mirrored in the SMAD cases as well. Most of the data samples in this study exhibit skew coefficients that lie in the range 1 to 3. Of the 68 samples in this moderate range of skewness, over two-thirds were better fitted by the WRC method. Of the samples that were better fitted by the WRC method, two-thirds exhibited maximum SRMSD and SMAD difference $\geq$ 10 percent. On the other hand, of the 13 samples that exhibited skew coefficients greater than 3.0, MDM and MIX were the preferred estimators in terms of SRMSD in 11 cases. However, in these cases the SRMSD differences were less than 10 percent in all but two instances. Thus, it appears that the WRC method is superior in descriptive ability when the data samples exhibit moderate skewness ($1 \leq \gamma \leq 3$), whereas for the samples of small skewness ($\gamma < 1.0$), MDM or MIX may be superior. This result is particularly evident in the small number of cases that exhibited negative skewness in the raw data. For the cases of large skewness ($\gamma > 3.0$) there was no significant difference in the performance of the methods. However, the performance of

TABLE 3 SMAD TEST RESULTS FOR THE THREE LP3 FITTING METHODS

Gage Station	SMAD			Method(s) with Min. SMAD	Max. Diff. %
	WRC	MIX	MDM		
02491500	0.101	0.102	0.107	WRC	5.9
02492000	0.109	0.109	0.109	WRC/MIX/MDM	0.0
07344450	0.150	0.159	0.190	WRC	26.7
07348700	0.116	0.122	0.128	WRC	10.3
07349500	0.067	0.067	0.069	WRC/MIX	3.0
07351500	0.086	0.062	0.063	MIX	38.7
07352000	0.067	0.081	0.075	WRC	20.9
07351000	0.064	0.065	0.069	WRC	7.8
07366200	0.190	0.186	0.217	MIX	16.7
07371500	0.090	0.078	0.081	MIX	15.4
07372200	0.087	0.090	0.089	WRC	3.4
07373000	0.112	0.119	0.099	MDM	20.2
07375000	0.091	0.094	0.083	MDM	13.2
07375500	0.084	0.086	0.084	WRC/MDM	2.4
07376000	0.067	0.078	0.081	WRC	20.9
07376500	0.050	0.055	0.059	WRC	18.0
07377000	0.103	0.086	0.080	MDM	28.8
07377500	0.121	0.096	0.072	MDM	68.1
07378000	0.052	0.052	0.053	WRC/MIX	1.9
07378500	0.064	0.069	0.067	WRC	7.8
07381800	0.073	0.090	0.085	WRC	23.3
07382000	0.211	0.281	0.464	WRC	119.9
*07382000	0.085	0.092	0.121	WRC	42.4
08010000	0.057	0.044	0.044	MIX/MDM	29.5
08012000	0.098	0.115	0.159	WRC	62.2
*08012000	0.079	0.086	0.117	WRC	48.1
08013000	0.092	0.093	0.094	WRC	2.2
08013500	0.079	0.081	0.080	WRC	2.5
08014500	0.156	0.154	0.210	MIX	36.4
*08014500	0.057	0.070	0.073	WRC	28.1
08015500	0.117	0.122	0.122	WRC	4.3
*08015500	0.069	0.069	0.069	WRC/MIX/MDM	0.0
08014000	0.136	0.146	0.160	WRC	17.6
08025500	0.184	0.246	0.230	WRC	33.7
08028000	0.154	0.168	0.132	MDM	27.3
02490105	0.096	0.122	0.128	WRC	33.3
07375222	0.134	0.116	0.083	MDM	61.4
07380160	0.067	0.059	0.049	MDM	36.7
07375170	0.111	0.123	0.115	WRC	10.8
07377300	0.080	0.085	0.080	WRC/MDM	6.3
07376600	0.041	0.042	0.039	MDM	7.7
07375480	0.112	0.116	0.110	MDM	5.5
02491700	0.126	0.153	0.155	WRC	23.0
02491350	0.125	0.128	0.104	MDM	23.1
07375800	0.121	0.136	0.116	MDM	17.2
07375307	0.138	0.140	0.123	MDM	13.8
07373500	0.071	0.076	0.070	MDM	8.6
07364300	0.119	0.125	0.121	WRC	5.0
07369500	0.025	0.026	0.023	MDM	13.0
07386500	0.075	0.068	0.062	MDM	21.0
08011800	0.095	0.097	0.096	WRC	2.1
08014200	0.119	0.119	0.133	WRC/MIX	11.8
07353500	0.125	0.136	0.116	MDM	17.2
07372500	0.180	0.203	0.268	WRC	48.9
*07372500	0.111	0.120	0.132	WRC	18.9
07370750	0.095	0.114	0.135	WRC	42.1
07372110	0.192	0.227	0.196	WRC	18.2
07372000	0.055	0.053	0.055	MIX	3.8
07370500	0.097	0.091	0.090	MDM	7.8
07370000	0.066	0.066	0.068	WRC/MIX	3.0
07367250	0.202	0.262	0.268	WRC	32.7
07366403	0.140	0.173	0.154	WRC	23.6
07366420	0.181	0.188	0.200	WRC	10.5
07365000	0.093	0.095	0.100	WRC	7.5

TABLE 3 (*continued on next page*)

TABLE 3 (*continued*)

Gage Station	SMAD WRC	SMAD MIX	SMAD MDM	Method(s) with Min. SMAD	Max. Diff. %
07364870	0.151	0.131	0.121	MDM	24.8
07365500	0.188	0.196	0.227	WRC	17.2
*07365500	0.124	0.130	0.118	MDM	10.2
07366000	0.186	0.181	0.231	MIX	27.6
07364700	0.294	0.369	0.366	WRC	25.5
08016600	0.114	0.120	0.120	WRC	5.0
08028700	0.116	0.118	0.136	WRC	17.2
*08028700	0.080	0.077	0.082	MIX	6.5
08014600	0.160	0.181	0.188	WRC	17.5
08013800	0.085	0.090	0.092	WRC	8.2
08013610	0.072	0.071	0.053	MDM	35.8
07354000	0.051	0.053	0.053	WRC	3.9
07353990	0.110	0.142	0.141	WRC	29.1
08016800	0.145	0.140	0.155	MIX	10.7
08016400	0.080	0.093	0.102	WRC	27.5
08015000	0.109	0.130	0.114	WRC	19.3
07352500	0.094	0.092	0.097	MIX	5.4
02490000	0.159	0.175	0.177	WRC	11.3
07348725	0.124	0.113	0.105	MDM	18.1
07348800	0.084	0.100	0.091	WRC	19.0
07347000	0.081	0.076	0.081	MIX	6.6
*07347000	0.054	0.047	0.049	MIX	14.9
07362100	0.097	0.100	0.111	WRC	14.4
07364190	0.067	0.065	0.063	MDM	6.3
07365800	0.344	0.353	0.427	WRC	24.1
*07365800	0.246	0.254	0.264	WRC	7.3
07373550	0.053	0.053	0.055	WRC/MIX	3.8
08014800	0.109	0.119	0.115	WRC	9.2
08025850	0.172	0.199	0.210	WRC	22.1
08024060	0.079	0.073	0.073	MIX/MDM	8.2
08023000	0.098	0.105	0.099	WRC	7.1
07351700	0.355	0.355	0.446	WRC/MIX	25.6
*07351700	0.124	0.088	0.076	MDM	63.2
07368500	0.038	0.039	0.039	WRC	2.6
07364500	0.068	0.062	0.054	MDM	25.9
02492360	0.077	0.091	0.083	WRC	18.2
08031000	0.089	0.093	0.103	WRC	15.7
08030000	0.087	0.086	0.093	MIX	8.1
08029500	0.162	0.197	0.208	WRC	28.4

*Run without the outliers.

all three methods decreased significantly as the skew coefficient increased. The average SRMSD for the MDM of the five samples with the smallest skew coefficients is 0.074, whereas the average SRMSD for the MDM of the five samples with the largest skew coefficients is 0.739. This result represents a deterioration in SRMSD performance of 892 percent. The MDM resulted in the better fit in all 10 of these extreme cases.

CONCLUSION

The results of this study demonstrate that in many cases there is a significant difference, depending on sample skewness, between the descriptive capability of these three moment-based methods. However, no method demonstrated clear superiority across all samples. For samples that exhibit skew coefficients greater than 1.0, the WRC method performs comparatively well in terms of both performance indices. For samples that exhibit skew coefficients of less than 1.0, the WRC method is clearly inferior to MDM and MIX, on the basis of the limited number of samples in this range. Previous Monte Carlo studies (*7,8*) that compared the relative predictive ability of these methods were based on samples generated from known populations and generally concluded that WRC did not perform as well as MDM and MIX in this regard. However, the results of the Monte Carlo studies may not translate to the real-world situations wherein the populations are unknown.

TABLE 4 COMPARISON OF ROBUSTNESS

Method	Min	Average	Max
		SRMSD	
WRC	.030	.238 (.196)	1.096 (.706)
MIX	.030	.252 (.209)	1.150 (.732)
MDM	.028	.233 (.195)	1.004 (.612)
		SMAD	
WRC	.025	.114 (.105)	.355 (.294)
MIX	.026	.121 (.111)	.369 (.369)
MDM	.023	.126 (.111)	.464 (.366)

Note: Values in parenthesis denote performance indices without the outliers.

The results may be of particular significance to engineers working in the area of highway drainage design. These structures are frequently designed for small recurrence intervals. The results demonstrate that for data with skew coefficients greater than 1.0, but particularly in the range $1 \leq \gamma \leq 3$, the WRC method possesses superior interpretive ability. Thus, it appears that this method may continue to be used confidently by engineers engaged in the design of small drainage structures.

ACKNOWLEDGMENT

This study was supported by the Louisiana Transportation Research Center. Stream data were provided by the U.S. Geological Survey. The authors wish to express their thanks to Glenn Chutz for his assistance in this work.

REFERENCES

1. B. Bobee. The Log Pearson Type 3 Distribution and Its Application in Hydrology. *Water Resources Research*, Vol. 11, No. 5, 1975, pp. 681–689.
2. B. Bobee and R. Robitaille. Correction of Bias in the Estimation of the Coefficient of Skewness. *Water Resources Research*, Vol. 11, No. 6, 1975, pp. 851–854.
3. R. Condie. 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, 1977, pp. 987–991.
4. D. V. Rao. Log Pearson Type 3 Distribution: Method of Mixed Moments. *Journal of Hydraulics Division, ASCE*, Vol. 106, No. 6, 1980, pp. 999–1019.
5. *Guidelines for Determining Flood Flow Frequency*. U.S. Water Resources Council, Bulletin 17B, Washington, D.C., 1982.
6. V. P. Singh and K. Singh. Parameter Estimation for Log-Pearson Type III Distribution by POME. *Journal of Hydraulic Engineering, ASCE*, Vol. 114, No. 1, 1988, pp. 112–122.
7. K. Arora and V. P. Singh. A Comparative Evaluation of the Estimators of Log Pearson Type 3 Distribution. *Journal of Hydrology*, Vol. 105, 1989, pp. 19–37.
8. F. Ashkar and B. Bobee. The Generalized Method of Moments as Applied to Problems of Flood Frequency Analysis: Some Practical Results for the Log-Pearson Type 3 Distribution. *Journal of Hydrology*, Vol. 90, 1987, pp. 199–217.
9. B. Bobee and F. Ashkar. Generalized Method of Moments Applied to LP3 Distribution. *Journal of Hydraulic Engineering, ASCE*, Vol. 114, No. 8, 1988, pp. 899–909.
10. K. Arora and V. P. Singh. On the Method of Maximum Likelihood Estimation for the Log-Pearson Type 3 Distribution. *Journal of Stochastic Hydrology and Hydraulics*, Vol. 2, No. 2, 1988, pp. 155–160.
11. H. N. Phien and V. Nguyen. Discussion on Derivation of the Pearson (PT) III Distribution by Using the Principle of Maximum Entropy (POME). *Journal of Hydrology*, Vol. 90, 1987, pp. 351–357.
12. J. R. Wallis, N. C. Matalas, and J. R. Slack. Just a Moment! *Water Resources Research*, Vol. 10, No. 2, 1974, pp. 211–219.
13. W. Kirby. Algebraic Boundedness of Sample Statistics. *Water Resources Research*, Vol. 10, No. 2, 1974, pp. 220–222.
14. N. C. Matalas, J. R. Slack, and J. R. Wallis. Regional Skew in Search of a Parent. *Water Resources Research*, Vol. 11, No. 6, 1975, pp. 815–826.
15. J. R. Wallis and E. F. Wood. Relative Accuracy of Log Pearson III Procedures. *Journal of Hydraulic Division, ASCE*, Vol. 111, No. 7, 1985, pp. 1043–1056.
16. C. Cunnane. Review of Statistical Models for Flood Frequency Estimation. In *Hydrologic Frequency Modeling*, V. P. Singh, ed., Reidel, Dordrecht, Holland, 1987, pp. 49–95.
17. F. E. Grubbs. Procedures for Detecting Outlying Observations in Samples. *Technometrics*, Vol. 11, No. 1, 1969, pp. 1–7.
18. G. Kuczera. Robust Flood Frequency Models. *Water Resources Research*, Vol. 18, No. 2, 1982, pp. 315–324.

Publication of this paper sponsored by Committee on Hydrology, Hydraulics, and Water Quality.