Illinois Institute of Natural Resources

State Water Survey Division

SURFACE WATER SECTION

SWS Contract Report 272

A NEW METHODOLOGY FOR FLOOD FREQUENCY ANALYSIS WITH OBJECTIVE DETECTION AND MODIFICATION OF OUTLIERS/INLIERS

By

Krishan P. Singh, Ph.D., Principal Scientist

Masahiro Nakashima, Graduate Research Assistant

Champaign, Illinois

August 1981

A NEW METHODOLOGY FOR FLOOD FREQUENCY ANALYSIS WITH OBJECTIVE DETECTION AND MODIFICATION OF OUTLIERS/INLIERS

by

Krishan P. Singh, Ph.D., Principal Scientist

Masahiro Nakashima, Graduate Research Assistant

Champaign, Illinois

August, 1981

ABSTRACT: Prerequisites for deriving satisfactory estimates of design floods are the objective detection and modification of outliers/inliers at desired levels of significance and a versatile technique of flood frequency analysis. Objective detection and modification of outliers/inliers has been accomplished by developing statistics for outliers/inliers at both the high and low end of the flood spectrum. An inlier at the high end is defined as a variate (generated or observed) which is lower than indicated by the trend of the rest of the data, and an inlier at the low end is higher than indicated by the rest of the data. The tests developed in this study can be used to check for outliers/inliers at different levels of significance, such as 0.01, 0.05, 0.1, 0.2, 0.3, and 0.4.

Three transformations for converting an observed flood series to an approximately normally distributed series were tested on flood series at 28 gaging stations in Illinois. The transformations considered were the power, Wilson-Hilferty, and 3-parameter log transformation. Analyses of the transformed series indicate that power transformation is the best of the three tested. The observed flood series is converted to a quasi normally-distributed series with the power transformation and, then, the statistical tests are applied for detection and modification of any outliers/inliers at various levels of significance.

Flood frequency methodologies (normal distribution after power transformation, log Pearson type III distribution, and mixed distribution) were tested on flood series observed at 37 gaging stations in Illinois. These analyses indicate that 1) regionalization of skew values alone, as recommended by the Water Resources Council in their Bulletin 17, is not a satisfactory solution to flood frequency problems, 2) the outlier criterion as recommended in Bulletin 17 is too severe, 3) an observed flood series needs to be checked for both inliers and outliers, 4) the power transformed series can have kurtosis lower or higher than 3 (it is 3 for a normal distribution) and the kurtosis correction can be applied if the transformed series is symmetrical, 5) the 37 power-transformed series exhibited asymmetry insofar as 5th and higher order odd moments were not zero, 6) only the mixed distribution can account for asymmetry displayed by the transformed series, and 7) the mixed distribution applied to series after detection and modification of outliers yields design flood estimates which exhibit regional consistency.

The versatile flood frequency analysis with the mixed distribution, coupled with objective detection and modification of outliers at various levels, provides a very satisfactory solution to the flood frequency problem. The method has been written as an efficient computer program.

TABLE OF CONTENTS

INTRODUCTION

"An accurate estimate of the flood potential is a key element to an effective nationwide flood damage abatement program. To obtain both a consistent and accurate estimate of flood losses requires development, acceptance and widespread application of a uniform, consistent and accurate technique for determining flood-flow frequencies."

The above excerpt from the Foreword of Bulletin 17A of the Hydrology Committee, U.S. Water Resources Council, stresses the need of a uniform, consistent and accurate technique for flood frequency analyses. From their analyses and research in the last two decades, the Council published bulletins (1967, 1976, and 1977) containing guidelines for determining flood frequency. They have recommended the fitting of log-Pearson type III, or LP3, distribution to observed annual flood peaks. The method of moments is used to determine the statistical parameters of the distribution from station data. Generalized relationships are used to define the skew coefficient for short record stations. Methods are proposed for treatment of some flood record problems encountered. The problem of outliers is recognized and it is dealt with in respect to outliers at both the low and high end. For the existence of low outliers, the criterion is

$$\left|\frac{x - \bar{x}}{s}\right| > [2.5 + 1.2 \log ({}^{n}/_{10})](1 - 0.4\, g_r) \qquad (1)$$

in which x = log Q, Q is an annual flood, x and s are the mean and standard deviation of log-transformed floods, n is the sample size, and g_r is the regional skew coefficient. The generalized skew coefficient for Illinois (with the exception of the lower portion of southern Illinois) is -0.4

(figure 1). The test statistic or the left-hand side of equation 1 needs to be greater than 3.45 and 3.87 for n = 25 and 50, respectively. Use of this criterion is equivalent to rejection at the 1 percent level of significance. When one or more outliers are identified, they are deleted from the record and the remaining record is treated as an incomplete record. If a high outlier is suspected, a comparison with historical flood data and flood information at nearby sites is made. If such information is available, a plotting position is assigned to each outlier and the procedure for historic floods is used; otherwise the outlier is retained as it is in the basic computations.

Previous Study

A study on the regional and sample skew values in flood-frequency analyses and the effect of outliers on the distribution parameters was conducted at the State Water Survey (Singh, 1980). In this study, the storm, basin, stream, soil, floodplain, and other relevant factors were investigated for 62 basins in areas drained by the Sangamon, Rock, and Little Wabash Rivers in Illinois (figure 1), to understand the variation in skew values from a number of annual flood series. Various flood frequency analyses were conducted for the annual flood series at each of the 62 study basins. The main conclusions drawn from this study were: 1) the criterion for a low outlier, as given in Bulletin 17A of the Water Resources Council, is too severe and needs modification — it yielded no outliers in any of the 62 series; 2) one or two very low floods greatly decrease the skew value and distort the fitted distribution curve — such low floods were found in about 30 percent of the flood series analyzed; 3) a high outlier can be confirmed

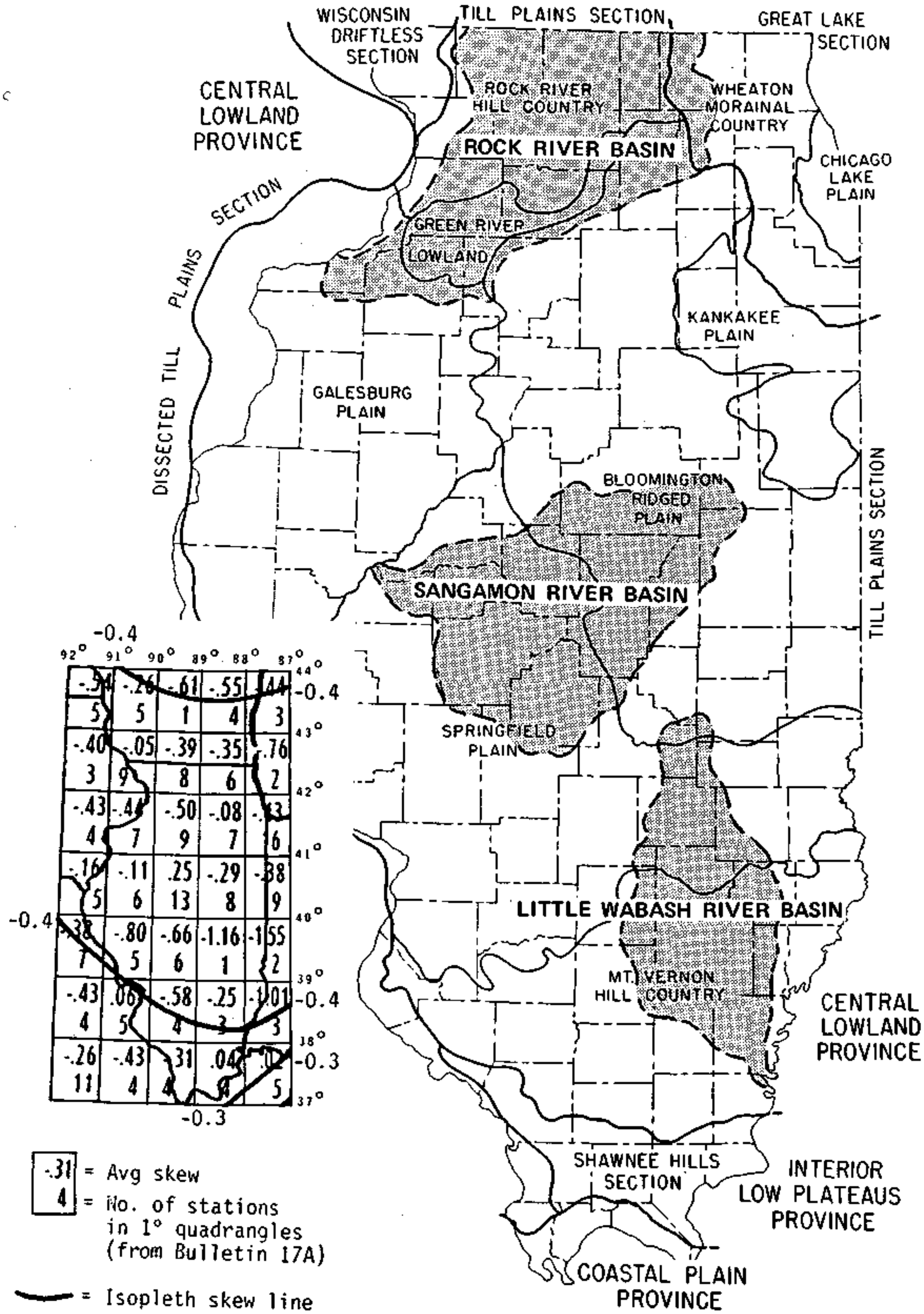


Figure 1. Study basins and physiographic divisions in Illinois, and skew coefficients from Bulletin 17A

by statistics of the storm producing it; 4) some better tests need to be developed for identifying or perceiving low and high outliers; 5) the methodology developed in this study for modifying outliers is generally satisfactory — this methodology depends on specifying the floods perceived as outliers by the analyst; 6) the modification of outliers developed for the LP3 changes both the standard deviation and skew, and regionalization of both the parameters may be needed instead of the skew alone; 7) the observed floods should be plotted using the best statistical plotting position instead of the commonly used Weibull plotting position, for checking the fit of the derived distribution curve with the observed data; 8) the standard deviation and skew appear to be correlated with basin and stream characteristics; 9) a change in the flow-section characteristics when river discharge begins inundating the floodplain introduces a new storage element that can affect the distribution shape of the observed floods; and 10) the distribution parameters (mean, standard deviation, and skew) below the junction of two major tributaries are affected greatly by the degree of concurrency of tributary flood peaks in time and flow magnitude.

The most important conclusions from this study were the need for developing tests to detect outliers/inliers at various levels of significance and better flood-frequency methods, and the inappropriateness of regionalizing skew values and using them in flood-frequency analyses without consideration of atypical hydrologic and hydraulic conditions.

Present Study

The main objectives of the study presented in this report are:

1. Development of statistical tests for outliers and inliers: An

inlier at the high end is a flood which is lower than indicated by the trend of the rest of the data and an inlier at the low end is a flood which is higher than indicated by the rest of the data (figure 2). These tests should check for outliers/inliers at different levels of significance, such as 0.01, 0.05, 0.1, 0.2, 0.3 and 0.4.

2. The statistical tests will be developed for the normally distributed series because of fewer number of distribution parameters. The observed flood series will be transformed to a series distributed as N (μ, σ^2). The available tranformations will be tested to choose the best.

3. Various methods of analyzing floods will be reviewed and their advantages and disadvantages examined. Their theoretical development, practical use, and any basic assumptions will be considered.

4. The flood-frequency methods will be computerized in a general package which will include testing for inliers/outliers and modification of any inliers/outliers detected at various significance levels. The results obtained with the use of 30 or more annual flood series from the Rock, Sangamon, and Little Wabash River sub-basins will be compared to determine the best method.

All of the objectives of this study have been met. Statistical tests for outliers/inliers at various levels of significance have been developed from extensive use of random number generators. The transformation technique that consistently and efficiently converts an observed flood series to a normally distributed series has been found. A new flood frequency methodology has been developed. It is much better than the others tested. The methodology yields flood estimates at various recurrence intervals with outliers/inliers detected and modified at various levels of significance.

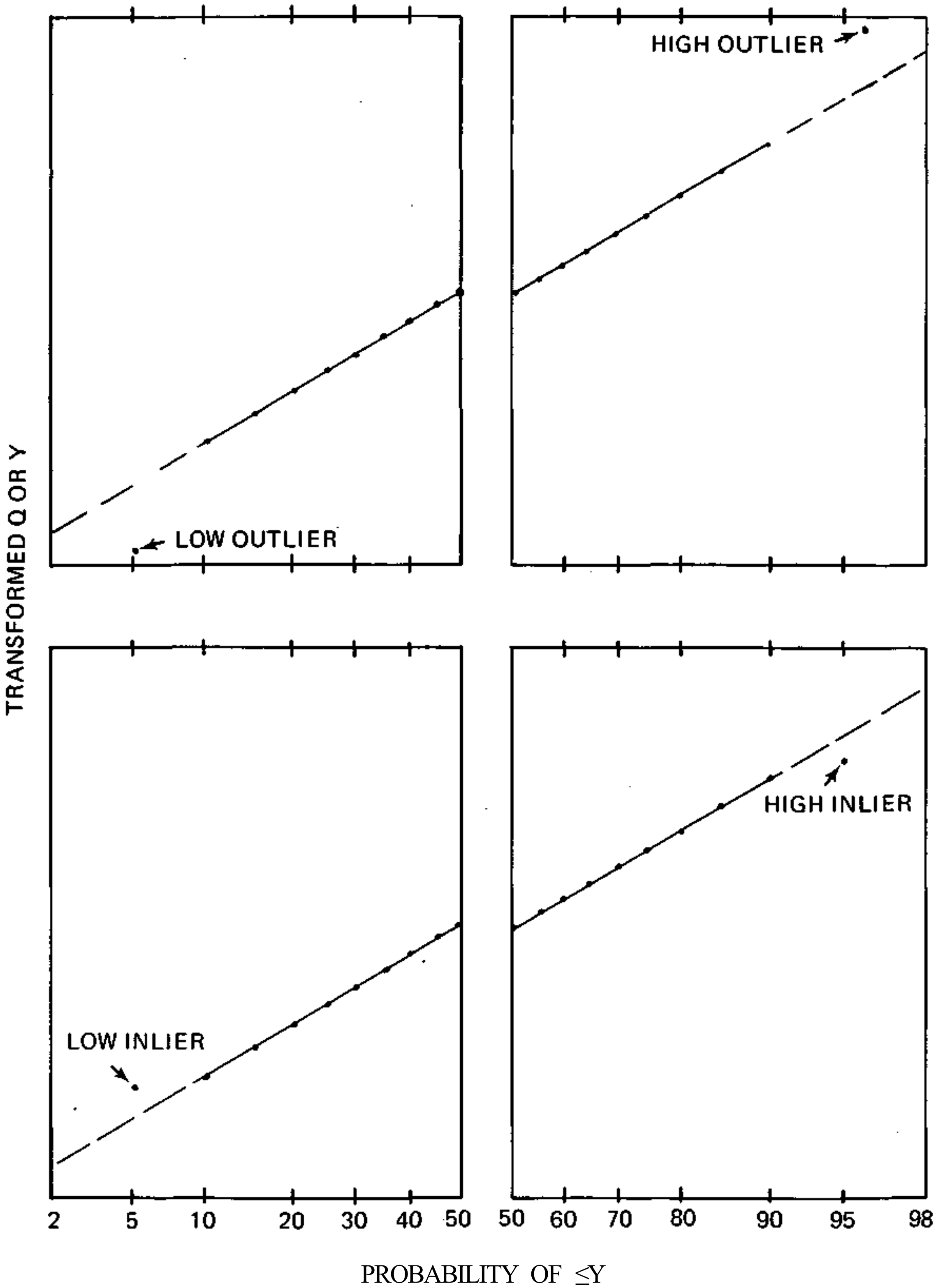


Figure 2. Definition sketch for low and high outliers and inliers

Acknowledgments

This study was jointly supported by the Division of Water Resources of the Illinois Department of Transportation and the Illinois State Water Survey. Bruce Barker of the Division of Water Resources served in a liaison capacity during the course of this study.

Ganapathi Ramamurthy, part-time graduate research assistant, helped in analyses of various transformation methods. John W. Brother, Jr. supervised the preparation of illustrations.

STATISTICAL TESTS FOR OUTLIERS AND INLIERS

An outlier in a set of data is defined as an observation or subset or observations, which appears to be inconsistent with the remainder of that set of data (Barnett and Lewis, 1978). The inconsistency can be interpreted as the observation being either significantly higher or lower at the high end (or lower or higher at the low end) than the value indicated by the rest of the data; the observation will be termed as an outlier or an inlier, respectively. The outlier can depart considerably from the assumed underlying distribution curve but the inlier departs by a lesser amount because the next observation can replace an inlier. In conventional flood-frequency analyses, it has been a matter of subjective judgment on the part of the analyst whether or not he picks up some observation for scrutiny. As stated earlier in the text, the criterion given for outlier detection in Bulletin 17 of the U.S. Water Resources Council is too severe. Literature search did not show the existence of statistical tests for checking outliers (at higher than 5% level) and inliers at different probability levels of their occurrence. The development of suitable statistical tests, detailed hereafter in this section, was an important part of this study.

Generation of Normally Distributed Random Numbers

Four methods or algorithms for generating normally distributed random numbers were tested extensively regarding their suitability, stability, and effectiveness in generating such numbers. A brief background of these methods is given here.

Box and Muller Method (BAMM)

Box and Muller (1958) presented a method of generating normally distributed random numbers, X_1 and X_2, with zero mean and unit variance:

$$X_1 = (-2 \ln U_1)^{\frac{1}{2}} \cos 2\pi U_2 \quad (1)$$

$$X_2 = (-2 \ln U_1)^{\frac{1}{2}} \sin 2\pi U_2 \quad (2)$$

in which U_1 and U_2 are random numbers drawn from a uniform or rectangular distribution function, U (0, 1), and ln is the natural logarithm. X_1 and X_2 are a pair of independent random variables such that

$$f(X_1,X_2) = f(X_1)\, f(X_2) \quad (3)$$

According to Box and Muller, this scheme should generate normal random numbers which are more reliable in the two extreme tails of the distribution.

The Polar Method (PLRM)

Box, Muller, and Marsaglia (Knuth, 1969) presented a method, commonly known as the polar method, for calculating two independent normally distributed variables, X_1 and X_2, from two independent random numbers from a uniform distribution, U (0,1). Computation of these variables follows the procedure (Knuth, 1969) given below.

1) Generate two independent random variables, U_1 and U_2, uniformly distributed between 0 and 1. Set $V_1 \leftarrow 2U_1-1$ and $V_2 \leftarrow 2U_2-1$. Then, V_1 and V_2 are uniformly distributed between -1 and +1.
2) Set $S \leftarrow V_1^2 + V_2^2$
3) If $S \geq 1$, return to step 1.
4) Set X_1 and X_2 according to the following two equations:

$$X_1 = V_1 \sqrt{(-2 \ln S)/S} \qquad (4)$$

$$X_2 = V_2 \sqrt{(-2 \ln S)/S} \qquad (5)$$

According to Knuth, the polar method is easy to computerize and has essentially perfect accuracy.

Inverse Normal Function Method (INFM)

International Mathematical Statistics Library (IMSL) has a normal or Gaussian random deviate generator which interprets the random numbers distributed as U(0, 1) to be cummulative probabilities and computes the corresponding normal deviates through an inverse normal function subroutine. The subroutine computes X. so that:

$$U_i = \text{Gauss}\ (X_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{X_i} \exp\ (-\frac{t^2}{2})\ dt \qquad (6)$$

Central Limit Theorem Method (CLTM)

The normally distributed random numbers can also be generated by the application of the central limit theorem. It states that the sum of a large number of components tends to the normal distribution as the number of components (regardless of their initial distribution) increases without limit (Ang and Tang, 1975). Therefore, the sum of a fixed number of uniform deviates on the interval (0, 1) should be distributed as gaussian. According to Cramer (1946), the mean and standard deviation of $f_n = \sum_{i=1}^{n} U_i$ are n/2 and $\sqrt{n/12}$, and f approaches the normal deviate rapidly as n increases. To generate standard normal deviates distributed as N (0, 1), n must be 12 for unit variance, and then $f_n = \sum_{i=1}^{12} U_i - 6$ for zero mean. Then, the deviate f_n constitutes normal deviate X.

Evaluation of Random Number Generators

Suitability of the random number generators was evaluated by two methods: the consistency of the statistics derived from 10 samples of generated normal random deviates of size 1,000 to 50,000 and the consistency of the statistics derived from 20 to 1000 samples of size 15 to 100.

Random sampling distribution theory aids in finding distribution parameters of the 3 statistics (mean, variance or standard deviation, and skew — their population values are 0, 1, and 0) being used in comparative evaluation of the 4 algorithms. Denoting mean, standard deviation and skew by $\bar{X}$, s, and g, the expected value and variance of the 3 statistics are given by the following equations (Cramer, 1946):

$$E(\bar{X}) = 0 \tag{7}$$

$$Var(\bar{X}) = \frac{1}{n} \tag{8}$$

$$E(s) = 1 \tag{9}$$

$$Var(s) = \frac{1}{2n} \tag{10}$$

$$E(s^2) = 1 \tag{11}$$

$$Var(s^2) = \frac{2}{n-1} \text{ or } \frac{2}{n} \text{ as n becomes large} \tag{12}$$

$$E(g) = 0 \tag{13}$$

$$Var(g) = \frac{6n\ (n-1)}{(n-2)(n+1)(n+3)} \text{ or } \frac{6}{n} \text{ as n becomes large} \tag{14}$$

A. Consistency of Statistics with Sample Size 1000 to 50,000

The intent was to investigate the variation of the mean and standard deviation of some statistics from the respective population values with respect to the length of the generated sequences from each of the 4 algorithms. The procedure, applied to each algorithm, can be considered in 4 steps.

1) Generate a sequence of 50,000 deviates.

2) Compute statistics: mean, standard deviation, and skewness for each of the 12 sample sizes of 1,000, 2,000,. . . . ., and 50,000 deviates from the beginning of the sequence.

3) Repeat steps 1 and 2 10 times, giving 10 values of the 3 statistics for each of the 12 sample sizes.

4) Compute the mean and standard deviation from the 10 values of each statistic for each of the 12 sample sizes.

$$AV(statistic) = \sum_{i=1}^{10} (statistic)_i/10 \qquad (15)$$

$$STD(statistic) = [\sum_{i=1}^{10} ((statistic)_i - AV(statistic))^2/9]^{0.5} \qquad (16)$$

The AV in equation 15 can be compared with expected values from equations 7, 9, 11, and 13 which are 0, 1, 1, and 0. The STD in equation 16 corresponds to $\sqrt{Var}$. The STD values for the 3 statistics and 12 sample sizes are given below.

Sample size, n	$\sqrt{\frac{1}{n}}$ STD($\bar{X}$)	$\sqrt{\frac{1}{2n}}$ STD(s)	$\sqrt{\frac{6}{n}}$ STD(g)
1,000	0.03162	0.02236	0.07745
2,000	0.02236	0.01581	0.05477
3,000	0.01826	0.01291	0.04472
5,000	0.01414	0.01000	0.03464
7,000	0.01195	0.00845	0.02928
10,000	0.01000	0.00707	0.02449
15,000	0.00817	0.00577	0.02000
20,000	0.00707	0.00500	0.01732
25,000	0.00632	0.00447	0.01549
30,000	0.00517	0.00408	0.01414
40,000	0.00500	0.00354	0.01224
50,000	0.00447	0.00316	0.01095

Evaluation of Statistics

Mean: The expected value of the mean of the random deviates, N (0, 1), is zero according to equation 7. The values of AV of the mean from the 4 algorithms are plotted with respect to sample size in figure 3a. It is evident that PLRM and CLTM yield AV vs n curves that are closer to zero than the other two. The values of STD of the mean are graphed in figure 3b together with the curve corresponding to equation 8, i.e., $STD(\bar{X}) = \sqrt{1/n}$. The curves from PLRM and INFM lie below the equation 10 curve, practically for the whole range of n.

Standard Deviation: The expected value of the standard deviation, s, of the random deviates, N(0, 1), is 1 according to equation 9. The values of AV of standard deviation from the 4 algorithms are plotted with respect to sample size in figure 4a. The curves show that PLRM is the best, closely followed by CLTM and BAMM. However, the STD curves together with the $\sqrt{\frac{1}{2n}}$ curve (figure 4b) show that INFM is the best, PLRM and CLTM are equally good, and BAMM is the worst. The overall rating, considering both AV and STD, in the decreasing order of preference is PLRM, CLTM, INFM, and BAMM.

Skewness: The expected value of the skew for deviates, N (0, 1), is zero according to equation 13. The values of AV of the skew from the 4 algorithms are plotted with respect to sample size in figure 5a. It is evident that CLTM and BAMM are better than PLRM which is better than INFM. The comparison of STD curves with $\sqrt{6/n}$ curve (figure 5b) shows that all algorithms are similar for n larger than 10,000.

B. Consistency of Statistics with Sample Size 15 to 100

The aim was to analyze the variation in the mean and standard deviation of $\bar{X}$, s^2, and g for small sample sizes but with the number of samples

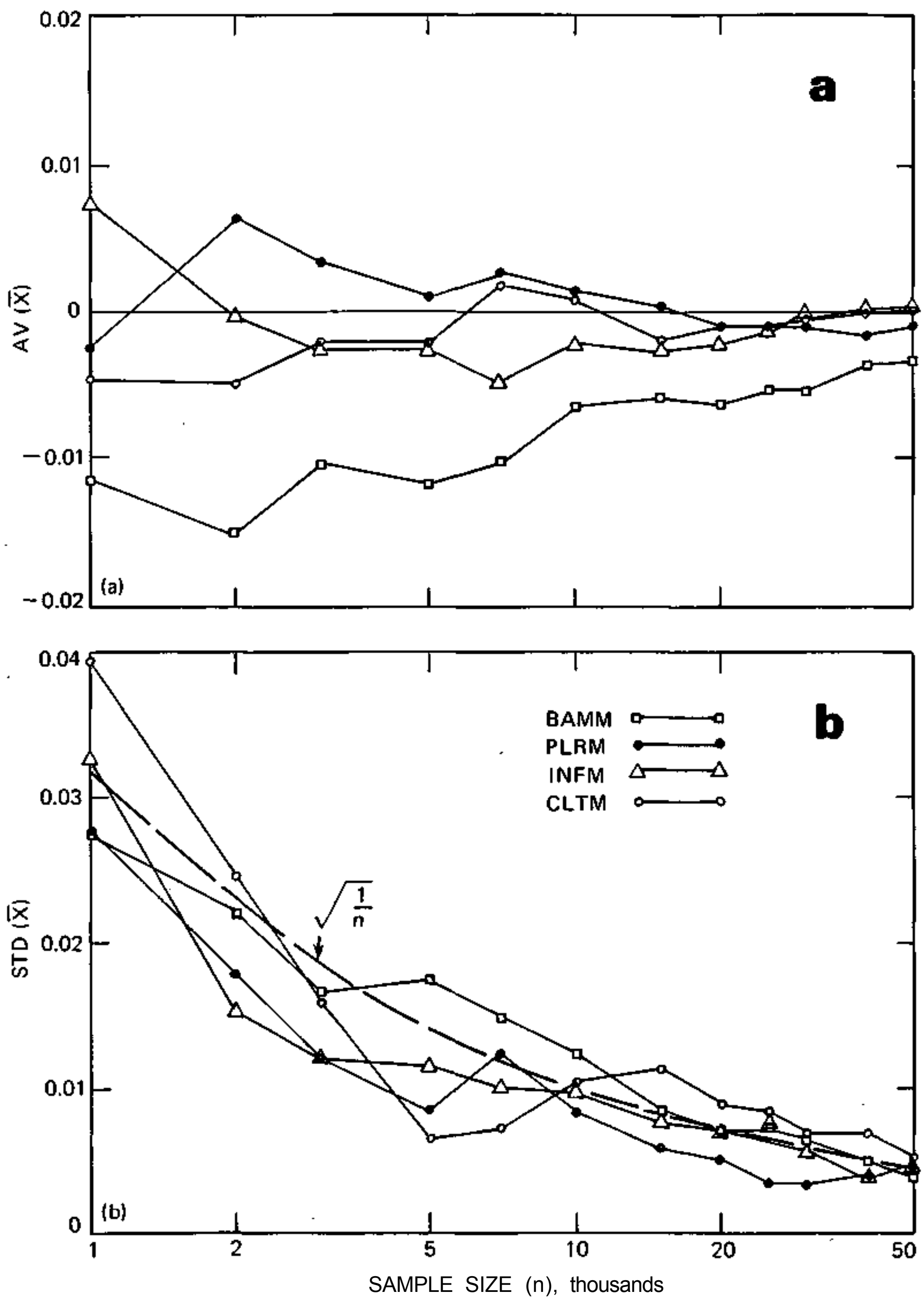


Figure 3. AV(X) and STD(X) versus sample size

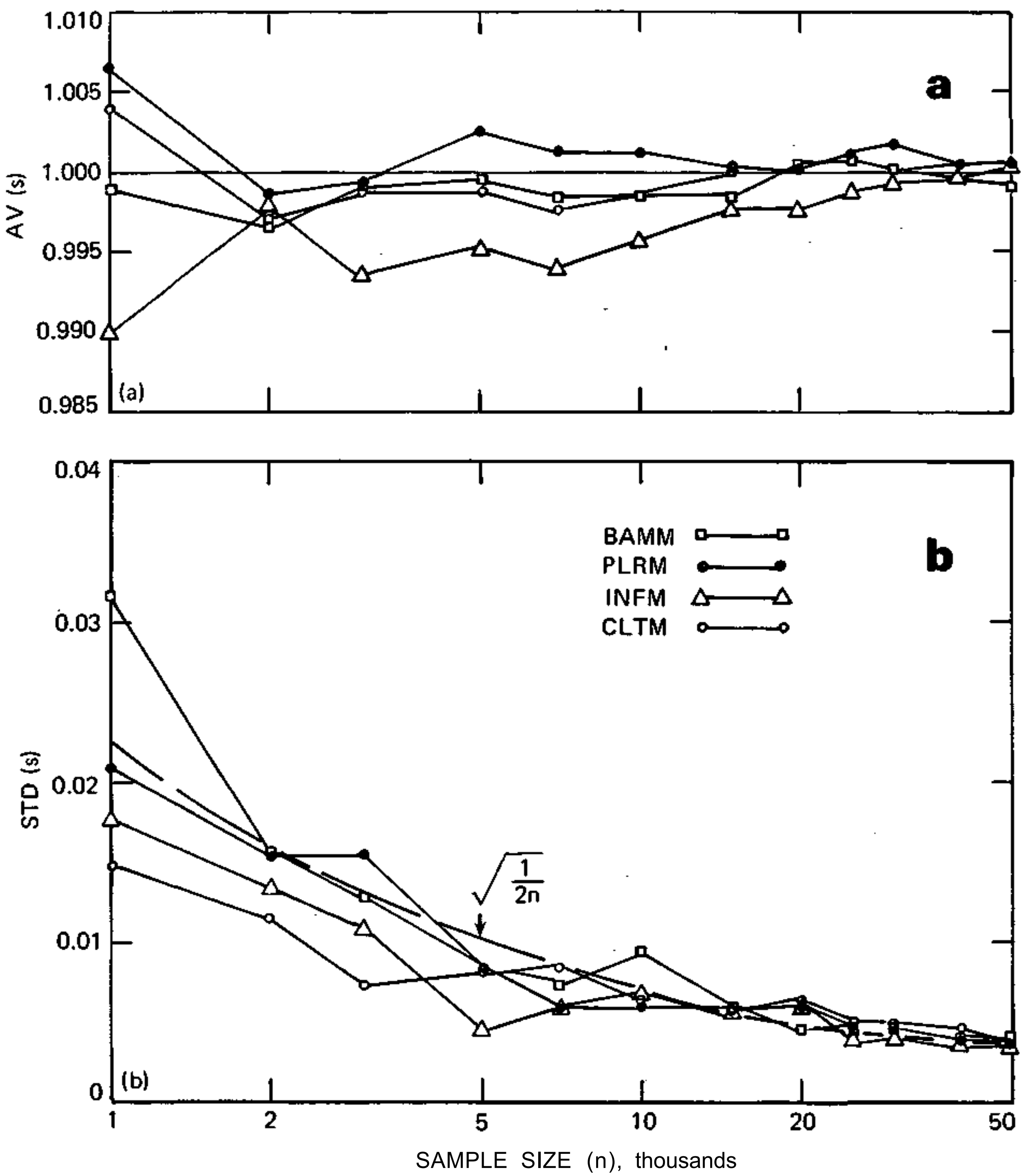


Figure 4. AV(s) and STD(s) versus sample size

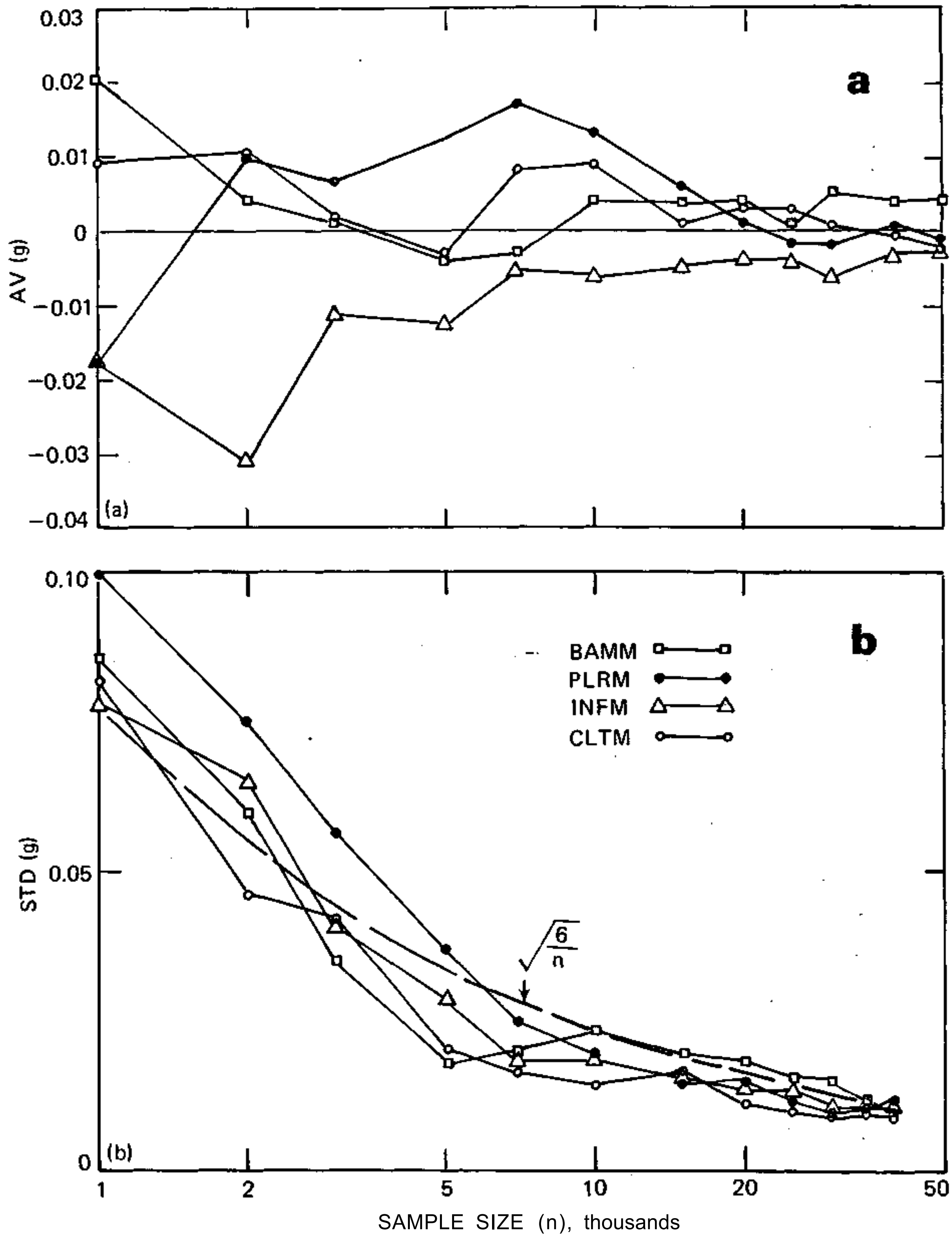


Figure 5. AV(g) and STD(g) versus sample size

varying from 20 to 1000. The procedure applied to each of the 4 algorithms can be considered in 4 steps.

1) Generate a sequence of 50,000 to 75,000 deviates.
2) Dissect the sequence into sizes of 15, 25, 40, 50, 75, and 100 resulting in 20, 50, 100, 300, 500, and 1000 samples of all sizes with some exceptions.
3) Compute 3 statistics, mean $\bar{X}$, variance s^2, and skewness g for each sample.
4) Compute the mean and standard deviation or variance of each of the 3 statistics for each of the 6 number of samples of size 15, 25, 40, 50, 75, and 100.

The values of AV, STD, and Var of the three statistics for different sample sizes and number of samples were calculated from the 4 algorithms. The expected values and standard deviations or variances of the 3 statistics were also computed from equations 7, 8, and 11 to 14.

Evaluation of Statistics

The evaluation of AV and STD or VAR of mean, variance, and skew for 20, 50, 100, 300, 500, and 1000 samples of sizes 15, 25, 40, 50, 75, and 100 is explained by the example of sample size 15 in Table 1. The best rating is 4 assigned to AV, STD, OR VAR from samples closest to the E, $\sqrt{\text{Var}}$ or Var from equation 7, 8 and 11 to 14 for the value of n under consideration. The ratings for 6 values of N (where N is the number of samples of size n) are added to give the total. It is evident from Table 1 that BAMM consistently underestimates statistics AV(X), AV(s^2), and AV(g). The combined overall ratings (sum of the 6 totals) for the BAMM, PLRM, INFM, and CLTM are 63, 109, 95, and 93, respectively. The BAMM algorithm does not perform as well as others.

Table 1. Evaluation of AV and STD or VAR of $\bar{X}$, s^2, and g for Sample Size 15

N	BAMM	*	PLRM	*	INFM	*	CLTM	*	E, VAR, or $\sqrt{VAR}$
AV($\bar{X}$)									
20	-0.0290	2	-0.0257	3	0.0681	1	-0.0108	4	0.0000
50	-0.0128	2	0.0158	1	-0.0008	4	-0.0058	3	0.0000
100	-0.0371	1	-0.0112	3	0.0032	4	-0.0122	2	0.0000
300	-0.0150	1	0.0117	3	-0.0069	4	-0.0149	2	0.0000
500	-0.0176	1	-0.0006	4	0.0022	2	0.0014	3	0.0000
1000	-0.0059	1	-0.0013	2	-0.0003	4	0.0011	3	0.0000
		8		16		19		17	
STD($\bar{X}$)									
20	0.2945	2	0.2821	3	0.2142	1	0.2531	4	0.2582
50	0.2705	4	0.2939	2	0.2312	3	0.2092	1	0.2582
100	0.2772	2	0.2809	1	0.2545	4	0.2486	3	0.2582
300	0.2711	1	0.2690	2	0.2515	3	0.2517	4	0.2582
500	0.2654	3	0.2658	2	0.2477	1	0.2625	4	0.2582
1000	0.2680	1	0.2640	3	0.2501	2	0.2534	4	0.2582
		13		13		14		20	
AV(s^2)									
20	0.9159	2	1.0017	4	1.1513	1	1.0162	3	1.0000
50	0.9123	3	0.9841	4	1.1398	1	0.9053	2	1.0000
100	0.9221	1	1.0090	4	1.0463	3	0.9304	2	1.0000
300	0.9693	2	1.0324	1	0.9890	3	1.0067	4	1.0000
500	0.9694	1	1.0019	3	1.0014	4	0.9903	2	1.0000
1000	0.9891	3	0.9937	4	1.0122	2	0.9783	1	1.0000
		12		20		14		14	
VAR(s^2)									
20	0.0831	1	0.1468	4	0.1696	2	0.1312	3	0.1429
50	0.0990	1	0.1055	3	0.1843	2	0.1163	4	0.1429
100	0.1061	1	0.1163	4	0.1714	3	0.1117	2	0.1429
300	0.1278	1	0.1368	3	0.1411	4	0.1330	2	0.1429
500	0.1210	1	0.1369	2	0.1423	4	0.1375	3	0.1429
1000	0.1326	1	0.1385	3	0.1422	4	0.1347	2	0.1429
		6		19		19		16	
AV(g)									
20	-0.1248	2	0.0149	4	-0.1695	1	-0.0485	3	0.0000
50	-0.2157	1	-0.1029	2	0.0109	4	0.1018	3	0.0000
100	-0.0910	1	-0.0163	4	0.0307	3	0.0784	2	0.0000
300	-0.0489	1	0.0050	4	-0.0171	2	0.0083	3	0.0000
500	-0.0213	1	0.0201	2	-0.0001	4	0.0125	3	0.0000
1000	-0.0163	1	0.0159	2	-0.0055	4	0.0126	3	0.0000
		7		18		18		17	
STD(g)									
20	0.5534	3	0.5612	4	0.5304	1	0.6147	2	0.5801
50	0.6015	3	0.5608	4	0.6580	2	0.7215	1	0.5801
100	0.6226	3	0.5938	4	0.6327	2	0.6729	1	0.5801
300	0.5929	1	0.5770	4	0.5733	2	0.5743	3	0.5801
500	0.5849	4	0.5872	3	0.5913	2	0.5547	1	0.5801
1000	0.5880	3	0.5791	4	0.5901	2	0.5531	1	0.5801
		17		23		11		9	

*refers to the rating, 4 is the highest and 1 is the lowest
N denotes the number of samples of size n which is 15 in this table

The information contained in Table 1 for sample size 15 is given for all the sample sizes 15, 25, 40, 50, 75 and 100 for the evaluation of AV, STD or VAR of the mean, variance, and skew in Table 2. The overall ratings for the 4 algorithms are:

	BAMM	PLRM	INFM	CLTM
$AV(\bar{X})$	49	$\breve{105}$	96	100
$STD(\bar{X})$	84	64	83	89
$AV(s^2)$	71	113	92	74
$VAR(s^2)$	70	114	81	85
$AV(g)$	59	100	85	76
$STD(g)$	82	90	97	51
	415	586	534	475

Thus, the PLRM algorithm seems to be the best in generating normal random numbers, with statistical attributes closely resembling samples drawn from a population distributed as N (0,1). This algorithm was used in developing departure distribution tables for detection of any outliers and/or inliers at the low and/or high end of the observed flood series.

Determination of Departure Distributions

Departure is defined here as the standard normal deviate corresponding to the plotting position of the high or low point of the series under consideration, minus the sample standard deviate for that point. The magnitude of these departures for outliers and inliers at the two extreme ends of the series needs to be determined at various probability levels. The theoretical departure distribution depends on m, n, and α in the general plotting position formula:

$$p = \frac{m - \alpha}{n + 1 - 2\alpha} \qquad (17)$$

Table 2. Evaluation of AV, STD or VAR of X, s , and g for Sample Sizes 15 to 100

n	Rating for BAMM	PLRM	INFM	CLTM	Rating for BAMM	PLRM	INFM	CLTM
	AV($\bar{X}$)				STD($\bar{X}$)			
15	8	16	19	17	13	13	14	20
25	11	16	16	17	16	10	15	19
40	9	18	15	18	12	10	14	14
50	7	18	17	18	15	7	18	10
75	8	19	17	16	15	10	10	15
100	6	18	12	14	13	14	12	11
	49	105	96	100	84	64	83	89
	AV(s^2)				VAR(s^2)			
15	12	20	14	14	6	19	19	16
25	11	22	16	11	9	21	13	17
40	13	20	15	12	16	17	12	15
50	13	19	16	12	12	21	12	15
75	13	18	15	14	13	18	15	14
100	9	14	16	11	14	18	10	8
	71	113	92	74	70	114	81	85
	AV(g)				STD(g)			
15	7	18	18	17	17	23	11	9
25	12	15	21	12	16	20	16	8
40	10	19	11	10	16	11	17	6
50	11	15	12	12	9	14	19	8
75	9	17	13	11	10	11	17	12
100	10	16	10	14	14	11	17	8
	59	100	85	76	82	90	97	51

Note: 1000 samples of certain sample sizes were not generated or processed.

in which a is a shift parameter, m is the rank, and n is the sample size. Such distributions have been derived in this study by generating large samples of departures for different m, n, and a>ranking them in an ascending order of magnitude, and determining the magnitude of departure at various probability levels and rank of outlier or inlier.

Generation of Departures

The following procedure was used in generating departures:

1) Generate 100,000 standard normal deviates with each of the four algorithms and dissect them into 1000 samples of 100 deviates each.
2) Pick one sample for each n (i.e., 10, 15, 20, 25, 30, 40, 50, 60, 75, and 100) starting from the beginning of each sample of size 100; this gives 1000 samples of 10 different sizes from 10 to 100.
3) Rank each of the 1000 samples of n size in an ascending order of magnitude and store 1000 values of the 5 lowest and 5 highest deviates in 10 series; each series corresponds to a high or low point. There are 10 series for each of the 10 sizes of n, and the size of each series is 1000.
4) Normalize each series by subtracting the mean from each deviate and dividing by the standard deviation.
5) Compute departures from 1000 normalized deviates in a series:

   Departure, $\Delta_{k,i}$ = ith theoretical standard normal deviate
   -kth normalized deviate for high/low location i (18)

   in which ith theoretical standard normal deviate equals that which corresponds to the probability from equation 17 with m = i (i=1, 2, ..., 5 and n-4, n-3,..., n) and n = 10, 15,..., 100;

and for the kth normalized deviate k = 1, 2,..., 1000; and with proper a values determined in steps 6 through 8.

6) In order to determine the appropriate value of α for different i and n values, generate 1000 departures by the step 5 with each of the following 6 values of α — 0.00, 0.25, 0.32, 0.38, 0.43, and 0.50. Compute the means of 1000 departures for each of the 6 values of a.

7) Interpolate the α values which make the means zero. These values of a were practically the same for the generating algorithms PLRM, INFM, and BAMM (values from CLTM were consistently lower) for the first highest and first lowest (similarly for the second highest and second lowest, and so on) rank for a given value of n. The results at the end of this step are given below:

n	a values for the highest and lowest rank 1 to 5				
	1	2	3	4	5
10	0.425	0.474	0.493	0.506	
15	0.414	0.465	0.484	0.494	0.505
20	0.408	0.456	0.482	0.490	0.505
25	0.405	0.448	0.477	0.494	0.500
30	0.406	0.443	0.464	0.482	0.491
40	0.404	0.439	0.460	0.472	0.484
50	0.403	0.439	0.455	0.469	0.474
60	0.402	0.439	0.452	0.469	0.472
75	0.403	0.441	0.450	0.456	0.461
100	0.402	0.440	0.451	0.456	0.460

8) Plot the a values versus n for the 5 ranks and draw smooth curves. The values from the smooth curves are as shown on the next page.

n	Smoothed a values for the highest and lowest rank				
	1	2	3	4	5
10	0.425	0.474	0.492	0.506	0.511
15	0.414	0.464	0.485	0.498	0.506
20	0.408	0.455	0.478	0.491	0.501
25	0.406	0.448	0.472	0.486	0.496
30	0.404	0.443	0.467	0.481	0.491
40	0.403	0.440	0.459	0.473	0.482
50	0.403	0.440	0.454	0.467	0.475
60	0.403	0.440	0.451	0.462	0.469
75	0.403	0.440	0.450	0.458	0.463
100	0.403	0.440	0.450	0.456	0.460

Compute departures from 1000 normalized deviates for each of the 100 series(5 series for low and 5 series for high ranks for each of the 10 sample sizes) developed from the PLRM algorithm in step 4 and use the a values derived in step 8 to compute $\Delta_{k,i}$ in step 5, then go to step 10.

10) Rank the 1000 departures in each of the 100 series, and obtain values of departures corresponding to probability, or rank/1000, equal to 0.01, 0.02, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 0.98 and 0.99. These departures are defined as $(\Delta)_{\underline{1},\underline{m},\underline{n}}$ where the subscript 1 denotes the number of probability levels 1 to 23; m refers to rank of low values 1 to 5 and high values 6 to 10, 1 is the lowest and 10 is the highest; and n denotes the sample size number 1 to 10, 1 for size 10 and 10 for size 100.

11) Generate 40 samples of departures $(\Delta)_{\underline{1},\underline{m},\underline{n}}$ from 40 generated sequences of 100,000 standard normal deviates each with the PLRM

algorithm. The mean of each of the 2300 departures ($\underline{l} \times \underline{m} \times \underline{n}$ = 23 × 10 × 10 or 2300) was obtained from the corresponding 40 values for each $(\Delta)_{\underline{l},\underline{m},\underline{n}}$. The mean departures are designated as $(\overline{\Delta})_{\underline{l},\underline{m},\underline{n}}$.

Development of Compact Departure Table

Barnett and Lewis (1978) give critical values for 1% and 5% tests of discordancy for a single outlier in a normal sample, using the deviation from the sample mean divided by the sample standard deviation as the test statistic. The corresponding test statistic is deviate corresponding to the plotting position of the higher outlier minus the departure $(\overline{\Delta})_{\underline{l},\underline{m},\underline{n}}$. The comparison of the test statistics is given below:

Sample size, n	p = 0.01		p = 0.05	
	Barnett & Lewis	This study	Barnett & Lewis	This study
20	2.88	2.89	2.56	2.55
30	3.10	3.11	2.74	2.74
40	3.24	3.24	2.87	2.86
50	3.34	3.35	2.96	2.96
60	3.41	3.43	3.03	3.03
100	3.60	3.59	3.21	3.21

Test statistics developed in this study for p = 0.01 and p = 0.05 are practically the same as given by Barnett and Lewis. However, the test statistics for inliers are not available in the literature. The test statistics for outliers of rank 2 to 5 and at other than 0.01 and 0.05 values of p are also not available in the literature.

The table of 2300 departures, $(\bar{\Delta})_{\underline{l},\underline{m},\underline{n}}$, was reduced to a compact table containing only 230 values. The following procedure was used in developing the compact table.

1. It was considered desirable to restrict the number, NO, of inliers/outliers at both the low and high end of flow spectrum, in relation to the sample size.

Sample size, n	NO
40 - 100	5
30-39	4
25-29	3
20-24	2
15-19	1

2. For the 5th outlier/inlier, the mean departure for n in the range of 40 to 100 was obtained by calculating the mean of $(\bar{\Delta})_{\underline{l},\underline{m},\underline{n}}$ for $\underline{n}$ = 6, 7, 8, 9, and 10 (or for n = 40, 50, 60, 75, and 100). Similarly, for the 1st outlier/inlier the mean departure for n in the range of 15 to 100 was obtained with n = 2, 3, 4, 5, 6, 7, 8, 9, and 10 (or for n = 15, 20, 25, 30, 40, 50, 60, 75, and 100). This reduces the departure table to $(\bar{\Delta})_{\underline{l},\underline{m}}$ in which l = 1, 2,..., 23 and m = 1, 2,..., 10. The resulting compact table of departures is shown in Table 3.

The standard deviation of departures for a particular range of n was calculated in a similar manner as the mean departure for that range in step 2. The standard deviations are given in Table 4. For p = 0.3 or 0.7, recommended for detection and modification of outliers/inliers in the later part of this report, the standard deviation varies from 0.0008 to 0.0061 and 0.0013 to 0.0050 for low and high outliers, respectively, and from

TABLE 3. DEPARTURES AT DIFFERENT PROBABILITY LEVELS FOR LOW AND HIGH OUTLIERS

P	Low 1 15-100	Low 2 20-100	Low 3 25-100	Low 4 30-100	Low 5 40-100	High 5 40-100	High 4 30-100	High 3 25-100	High 2 20-100	High 1 15-100
.01	-.6893	-.4950	-.4122	-.3633	-.3270	-.3769	-.4289	-.5109	-.6541	-1.0538
.02	-.6297	-.4452	-.3703	-.3242	-.2923	-.3271	-.3706	-.4423	-.5614	-.8993
.05	-.5325	-.3694	-.3033	-.2641	-.2373	-.2563	-.2900	-.3414	-.4328	-.6835
.10	-.4410	-.2991	-.2425	-.2107	-.1880	-.1951	-.2205	-.2580	-.3225	-.5003
.15	-.3747	-.2495	-.2008	-.1734	-.1547	-.1551	-.1746	-.2030	-.2514	-.3826
.20	-.3184	-.2091	-.1670	-.1432	-.1270	-.1236	-.1387	-.1614	-.1966	-.2951
.25	-.2687	-.1734	-.1372	-.1169	-.1040	-.0970	-.1083	-.1257	-.1514	-.2215
.30	-.2210	-.1408	-.1103	-.0933	-.0819	-.0737	-.0818	-.0942	-.1120	-.1586
.35	-.1766	-.1097	-.0846	-.0710	-.0622	-.0524	-.0573	-.0655	-.0762	-.1024
.40	-.1319	-.0795	-.0598	-.0501	-.0432	-.0322	-.0348	-.0387	-.0425	-.0511
.45	-.0879	-.0493	-.0357	-.0290	-.0247	-.0128	-.0131	-.0131	-.0109	-.0030
.50	-.0435	-.0193	-.0118	-.0085	-.0059	.0063	.0081	.0111	.0196	.0428
.55	.0028	.0114	.0122	.0123	.0128	.0247	.0288	.0356	.0495	.0878
.60	.0515	.0428	.0375	.0339	.0318	.0432	.0497	.0599	.0790	.1321
.65	.1036	.0764	.0638	.0563	.0517	.0621	.0711	.0843	.1087	.1761
.70	.1606	.1123	.0922	.0807	.0732	.0817	.0934	.1096	.1400	.2212
.75	.2243	.1519	.1234	.1070	.0966	.1030	.1168	.1364	.1731	.2675
.80	.2969	.1974	.1586	.1369	.1228	.1265	.1426	.1663	.2089	.3166
.85	.3868	.2509	.2005	.1724	.1540	.1535	.1720	.1999	.2496	.3718
.90	.5030	.3213	.2541	.2172	.1933	.1865	.2088	.2411	.2989	.4376
.95	.6809	.4309	.3366	.2853	.2534	.2348	.2627	.3000	.3686	.5285
.98	.8906	.5566	.4318	.3646	.3215	.2887	.3204	.3655	.4412	.6220
.99	1.0292	.6433	.4979	.4182	.3682	.3233	.3583	.4073	.4882	.6787

Note: 15-100,. . . . ., 40-100 denote the range of sample size, n

TABLE 4. STANDARD DEVIATION OF DEPARTURES AT DIFFERENT PROBABILITY LEVELS

P	Low 1 15-100	Low 2 20-100	Low 3 25-100	Low 4 30-100	Low 5 40-100	High 5 40-100	High 4 30-100	High 3 25-100	High 2 20-100	High 1 15-100
.01	.0062	.0041	.0112	.0142	.0115	.0071	.0134	.0193	.0296	.0588
.02	.0076	.0034	.0082	.0101	.0088	.0056	.0094	.0100	.0197	.0396
.05	.0073	.0031	.0037	.0050	.0048	.0020	.0030	.0079	.0106	.0212
.10	.0070	.0048	.0013	.0021	.0025	.0003	.0008	.0035	.0060	.0099
.15	.0073	.0052	.0019	.0010	.0013	.0016	.0009	.0012	.0026	.0067
.20	.0068	.0059	.0022	.0008	.0008	.0016	.0014	.0017	.0014	.0036
.25	.0064	.0055	.0025	.0011	.0009	.0018	.0020	.0024	.0008	.0019
.30	.0061	.0052	.0027	.0015	.0008	.0017	.0025	.0024	.0014	.0016
.35	.0062	.0047	.0032	.0016	.0009	.0019	.0025	.0029	.0019	.0022
.40	.0056	.0043	.0034	.0021	.0013	.0018	.0025	.0036	.0024	.0032
.45	.0054	.0045	.0033	.0023	.0017	.0017	.0026	.0039	.0027	.0034
.50	.0047	.0040	.0030	.0021	.0018	.0017	.0027	.0039	.0036	.0039
.55	.0043	.0040	.0029	.0023	.0017	.0018	.0026	.0040	.0042	.0041
.60	.0041	.0034	.0025	.0024	.0016	.0019	.0024	.0037	.0046	.0039
.65	.0025	.0027	.0020	.0020	.0020	.0015	.0021	.0036	.0048	.0037
.70	.0015	.0021	.0015	.0020	.0020	.0013	.0017	.0033	.0050	.0042
.75	.0030	.0016	.0015	.0020	.0023	.0011	.0017	.0029	.0047	.0047
.80	.0048	.0016	.0011	.0018	.0021	.0008	.0014	.0025	.0047	.0049
.85	.0067	.0035	.0017	.0018	.0018	.0017	.0011	.0021	.0040	.0058
.90	.0098	.0066	.0038	.0020	.0014	.0026	.0024	.0022	.0029	.0061
.95	.0198	.0118	.0068	.0039	.0028	.0043	.0053	.0050	.0023	.0058
.98	.0357	.0200	.0113	.0076	.0046	.0073	.0107	.0095	.0045	.0055
.99	.0493	.0265	.0180	.0116	.0076	.0116	.0136	.0118	.0068	.0053

Note: 15-100, ...40-100 denote the range of sample size, n

0.0015 to 0.0021 and 0.0014 to 0.0025 for low and high inliers, respectively. These small values of standard deviation justify the use of a compact table. However, the table of 2300 departures can be as easily used in the computer program, if so desired.

The distributions of departures for the lowest 5 events are graphed in figure 6 and for the highest 5 events in figure 7.

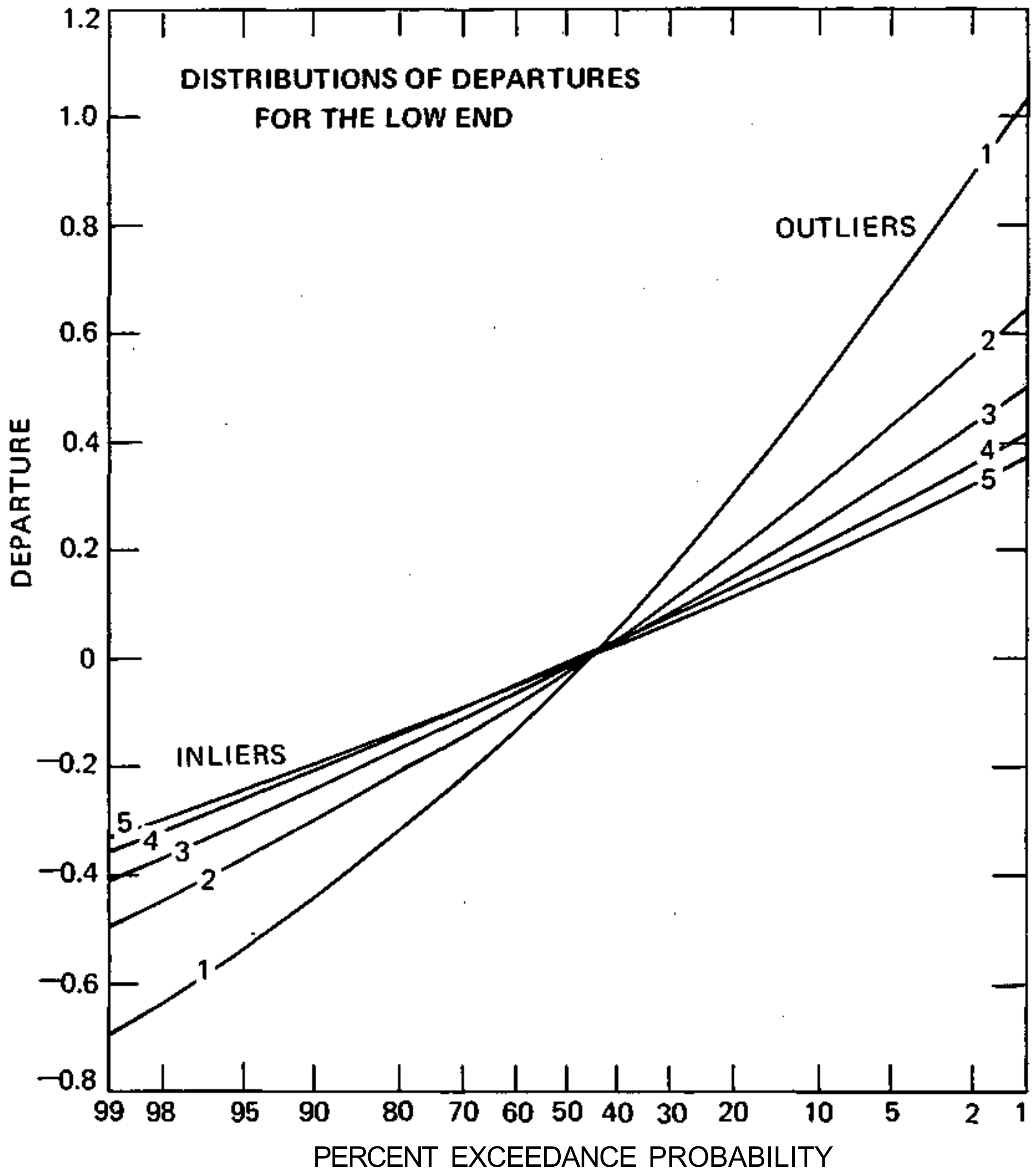


Figure 6. Distributions of departures for the low end

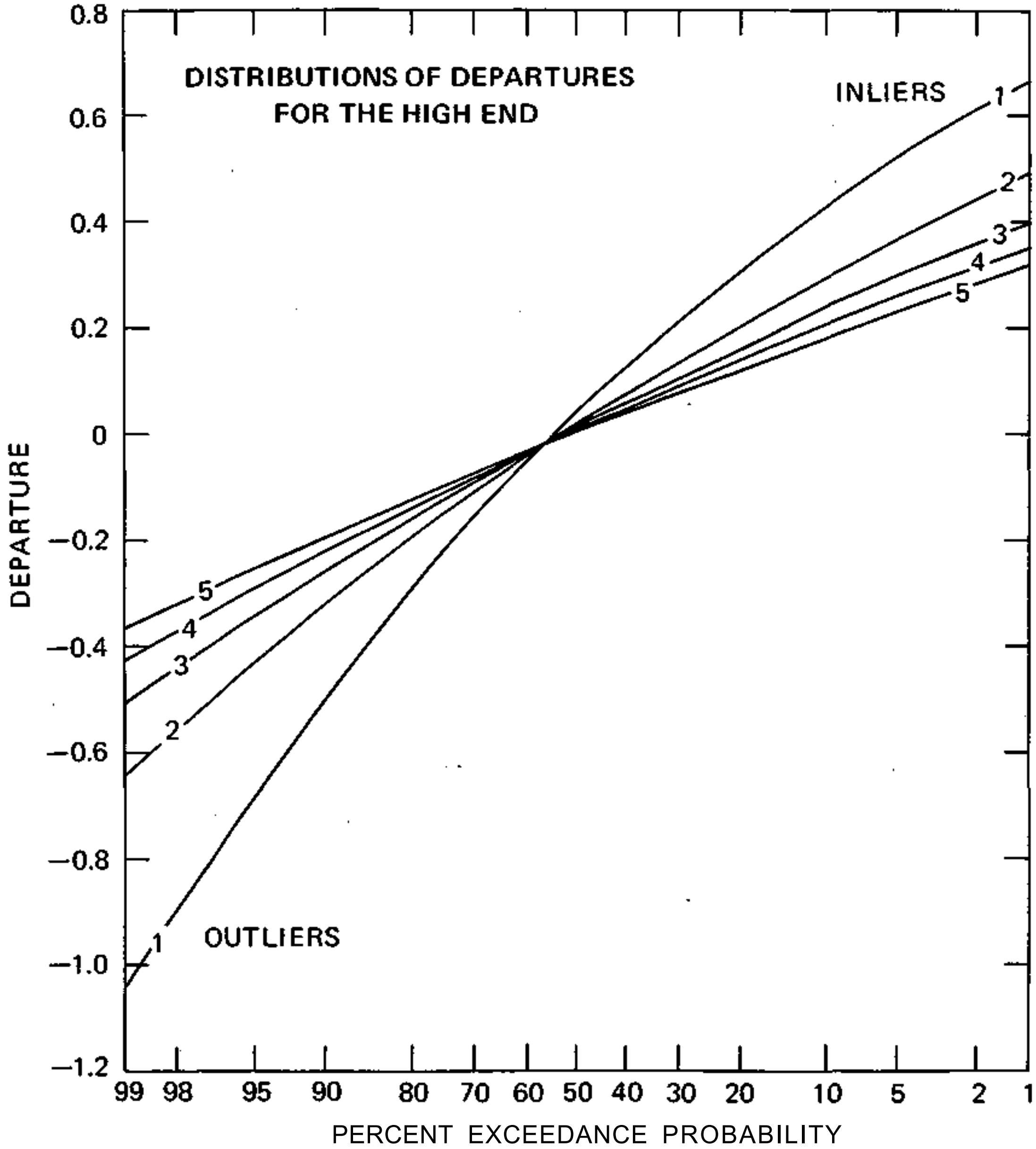


Figure 7. Distributions of departures for the high end

METHODS OF NORMALIZING DATA

The tests for determining outliers/inliers can be easily developed and applied for normally distributed samples because of the minimum number of distribution parameters, i.e., the mean and the standard deviation. Three methods of transforming an observed flood series to a series distributed as N (μ, σ^2), where u is the mean and σ^2 is the variance, were tested extensively on flood series observed at 28 gaging stations in Illinois. The methods are: power transformation, Wilson-Hilferty transformation, and 3-parameter log-normal transformation.

Power Transformation

Box and Cox (1964) suggested a transformation for normalizing a data set:

$$y_i = \frac{Q_i^{\lambda} - 1}{\lambda}, \quad \lambda \neq 0 \tag{1}$$

and

$$y_i = \log Q_i, \quad \lambda = 0 \tag{2}$$

in which Q_i is the annual flood from a sample of size n, λ is a constant of transformation, and i = 1, 2,..., n. It is a general power transformation and the logarithmic, reciprocal and square-root transformations can be considered as its special cases. The constant λ can be obtained with one of the following three criteria:

1. Maximum log-likelihood (ML) estimator of λ, when $\lambda \neq 0$, can be obtained from (Singh, 1980)

$$L_{max}(\lambda) = -1/2\ n \log \hat{\sigma}_y^{\,2}(\lambda) + \log J(\lambda;Q) \tag{3}$$

and

$$\log J(\lambda;Q) = (\lambda - 1) \sum_{i=1}^{n} \log Q_i \tag{4}$$

A plot of L_{max} (λ) versus λ can indicate the ML estimate of λ. A computer algorithm was developed for determining $L_{max}(\lambda)$.

2. Zero coefficient of skew criterion can be met by computing the skew g for y series with different values of λ,

$$g = n \sum_{i=1}^{n} (y_i - \bar{y})^3 \; / \; [(n - 1)(n - 2) \; s_y^{\,3}] \qquad (5)$$

in which $\bar{y}$ and s_y are the mean and standard deviation of the y series. Value of λ which makes g = 0 can be interpolated from the λ values giving a little positive and a little negative g. A computer program for calculating λ which yields g equal to zero was added to the ML algorithm.

3. Minimization of |g| + |5th| criterion is based on the premise that 3rd and higher order odd moments are zero in the case of a theoretical normal distribution. A computer program was added to the ML algorithm for calculating λ value which minimized the sum of the absolute values of the skew and the 5th,

$$5th = n^3 \sum_{i=1}^{n} (y_i - \bar{y})^5 \; / \; [(n - 1)(n - 2)(n - 3)(n - 4) \; s_y^{\,5}] \qquad (6)$$

Wilson-Hilferty Transformation

A standard deviate, x, can be calculated from Q, $\bar{Q}$, and s_Q :

$$x_i = \frac{\log Q_i - (\overline{\log Q})}{s\ (\log Q \text{ series})} \qquad (7)$$

If the underlying distribution is log-Pearson type III, or LP3, x is the gamma standard deviate that can be converted to the normal standard deviate by the

Wilson and Hilferty (1931) transformation

$$y_i = \frac{6}{g}\left[\left(\frac{gx_i}{2} + 1\right)^{1/3} + \frac{g^2}{36} - 1\right] \qquad (8)$$

A computer program was developed for converting the Q series to x series, and for calculating the skew of the x series. Two subprograms were added to obtain values of g (with the first estimate equal to sample g for the x series) so that the y series has zero skew and to obtain a value of g that minimized the sum of the absolute values of g and 5th of the y series. These subprograms used a reiterative procedure to obtain satisfactory values of g to meet zero skew and min [|g| + |5th|] criteria.

Three-Parameter Lognormal Transformation

The following transformation was considered for normalizing the data,

$$y_i = \log (Q_i + a) \qquad (9)$$

in which a is a constant, positive, negative or zero. By a fast-converging reiterative process, the value of a was determined for the following three criteria:

1. skew g = 0
2. minimize [|g| + |5th|]
3. kurtosis = 3.0

A computer program was developed for calculating the values of a to meet the above criteria.

Test Data and Results

Annual flood series at 28 gaging stations from a previous report (Singh, 1980) were used in testing the suitability of the three transformations.

These stations were selected and arranged in three categories:

I. 14 gaging stations with flood series having no significant high or low outliers/inliers.

II. 7 gaging stations with flood series having outliers/inliers at the high end.

III. 7 gaging stations with flood series having outliers/inliers at the low end.

The 28 stations are listed in Table 5 together with observed high and low floods and their modified values as determined in a previous study (Singh, 1980). Category I flood series has no significant high or low outliers/inliers but one outlier/inlier at either low or high end was considered for checking any effect of minor modification in values of these outliers/inliers. For both categories II and III, one and two outliers/inliers were considered separately.

Power Transformation Results

The results are presented in Table 6. Criteria A, B, and C denote g = 0, ML estimate, and min [|g| + |5th|], respectively. The TS1 and TS2 are test statistics, given by

$$TS1 = \sum_{i=1}^{n} (\Delta Z_i)^2 \text{ with } \alpha = 0 \tag{10}$$

$$TS2 = \sum_{i=1}^{n} (\Delta Z_i)^2 \text{ with } \alpha = 0.38 \tag{11}$$

$$\Delta Z_i = (Z_o)_i - (Z_c)_i \tag{12}$$

$$(Z_c)_i = (y_i - \bar{y}) / y_s \tag{13}$$

TABLE 5. Basins Used in Evaluation of 3 Transformations

No.	Stream and gaging station	USGS No.	n	Observed flood in cfs L1	L2	H1	H2	Modified flood in cfs L1	L2	H1	H2
	A. Well Behaved Flood Series										
1	Sangamon River at Mahomet	05 571000	29	1020		14600		1060		13682	
2	Sangamon River at Monticello	05 572000	68	704		19000		1110		19567	
3	Salt Creek near Greenview	05 582000	35	3440		41200		3224		46989	
4	Pecatonica River at Freeport	05 435500	63	1910		18400		1704		20082	
5	Pecatonica River at Shirland	05 437000	32	3490		16600		2946		18282	
6	Rock River at Rockton	05 437500	37	6340		30000		5402		29249	
7	Kishwaukee River near Perryville	05 440000	37	2020		16400		1579		17754	
8	Green River near Geneseo	05 447500	40	1340		12100		1323		11533	
9	Kishwaukee River at Belvidere	05 438500	37	935		10300		716		11994	
10	S. Br. Kishwaukee R. near Fairdale	05 439500	37	1010		8460		819		7559	
11	Elkhorn Creek near Penrose	05 444000	37	530		6770		563		6872	
12	Rock Creek near Morrison	05 445500	32	765		5770		720		5429	
13	Green River at Araboy	05 447000	37	480		6120		340		5327	
14	Mill Creek at Milan	05 448000	37	450		9300		399		11253	
	B. Flood Series with High Outliers/Inliers										
1	S.F. Sangamon River near Nokomis	05 574000	26			8600				9340	
						8600	6000			6439	4440
2	S.F. Sangamon River near Kincaid	05 575500	33			21500				17082	
						21500	13700			20079	15808
3	Sangamon River at Riverton	05 576500	64			68700				42826	
						68700	41000			36170	34880
4	Salt Creek near Rowell	05 578500	34			24500				16258	
						24500	12400			19949	14911
5	Kickapoo Creek near Lincoln	05 580500	32			14800				17558	
						14800	13800			12230	10066
6	Sangamon River at Oakford	05 583000	59			123000				81230	
						123000	46300			58656	52683
7	Skillet Fork near Wayne City	03 380500	48			51000				26591	
						51000	22800			28673	24441

TABLE 5. (concluded)

No.	Stream and gaging station	USGS No.	n	Observed flood in cfs L1	L2	H1	H2	Modified flood in cfs L1	L2	H1	H2
	C. Flood Series with Low Outliers/Inliers										
1	Flat Branch near Taylorville	05 574500	27	457				452			
				457	660			1088	1360		
2	S.F. Sangamon River near Rochester	05 576000	27	971				910			
				971	1230			1211	1568		
3	Spring Creek at Springfield	05 577500	29	217				154			
				217	225			377	479		
4	Lake Fork near Cornland	05 579500	29	152				432			
				152	548			394	507		
5	Leaf River at Leaf River	05 441000	37	233				309			
				233	459			215	335		
6	Little Wabash River below Clay City	03 379500	62	1440				2319			
				1440	2920			1899	2439		
7	Little Wabash River at Carmi	03 381500	37	3320				3516			
				3320	4180			5953	6567		

TABLE 6. Evaluation of Normalization of a Flood Series by Power Transformation

#	USGS No.	*	With no modification					With H1 modified					With L1 modified				
			λ	g	5th	TS1	TS2	λ	g	5th	TS1	TS2	λ	g	5th	TS1	TS2
I.	Flood series with no significant high or low outliers/inliers																
1	05571000	A	0.075	0.000	-0.204	0.445	0.321	0.103	0.000	-0.282	0.459	0.354	0.060	-0.001	-0.176	0.447	0.330
		B	0.066	-0.011	-0.295	0.445	0.320	0.090	-0.016	-0.401	0.459	0.353	0.053	-0.010	-0.244	0.447	0.330
		C	0.095	0.026	-0.002	C.447	0.324	0.134	0.036	0.001	0.463	0.360	0.078	0.022	-0.002	0.448	0.333
2	05572000	A	0.180	-0.001	-0.330	0.517	0.420	0.174	0.000	-0.306	0.511	0.401	0.105	0.000	0.098	0.464	0.472
		B	0.171	-0.015	-0.456	0.519	0.421	0.165	-0.015	-0.436	0.513	0.402	0.095	-0.013	0.004	0.464	0.473
		C	0.204	0.039	0.000	0.320	0.426	0.196	0.036	0.005	0.513	0.406	0.095	-0.013	0.004	0.464	0.473
3	05582000	A	0.060	0.000	0.717	0.884	0.834	0.005	0.000	0.864	0.889	0.807	0.078	0.000	0.690	0.848	0.778
		B	0.053	-0.009	0.653	0.863	0.834	0.004	-0.001	0.853	0.889	0.807	0.069	-0.011	0.603	0.847	0.778
		C	-0.021	-0.097	-0.002	0.G99	0.857	-0.080	-0.110	0.004	0.907	0.836	0.005	-0.091	-0.001	0.858	0.793
4	05435500	A	0.050	0.000	-0.022	0.665	0.763	0.027	0.000	0.057	0.603	0.664	0.075	0.000	-0.099	0.577	0.633
		B	0.044	-0.007	-0.061	0.666	0.763	0.024	-0.003	0.035	0.603	0.665	0.067	-0.010	-0.156	0.577	0.633
		C	0.053	0.003	-0.002	0.665	0.763	0.019	-0.009	0.000	0.604	0.666	0.089	0.016	0.000	0.579	0.635
5	05437000	A	0.532	0.000	0.200	0.635	0.995	0.430	0.000	0.036	0.770	0.813	0.589	0.000	0.137	0.740	0.794
		B	0.414	-0.077	-0.209	0.916	1.018	0.346	-0.061	0.004	0.790	0.838	0.469	-0.085	-0.337	0.750	0.801
		C	0.474	-0.038	-0.001	0.901	1.002	0.345	-0.062	-0.001	0.790	0.838	0.554	-0.025	0.000	0.740	0.792
6	05437500	A	0.194	0.000	0.332	0.927	1.009	0.218	0.000	0.286	0.946	1.041	0.272	0.000	0.233	0.717	0.731
		B	0.158	-0.124	0.198	0.937	1.021	0.176	-0.028	0.138	0.958	1.055	0.227	-0.033	0.033	0.725	0.740
		C	0.104	-0.060	0.001	0.958	1.046	0.136	-0.054	-0.001	0.973	1.072	0.220	-0.038	0.002	0.726	0.741
7	05440000	A	0.611	0.000	0.051	1.340	1.546	0.555	0.000	0.151	1.194	1.351	0.646	0.000	0.015	1.193	1.360
		B	0.437	-0.123	0.557	1.368	1.572	0.408	-0.112	-0.440	1.229	1.389	0.471	-0.132	-0.675	1.199	1.355
		C	0.596	-0.011	0.000	1.338	1.544	0.517	-0.029	-0.002	1.195	1.353	0.642	-0.003	0.000	1.192	1.359
8	05447500	A	0.743	0.000	0.414	0.626	0.367	0.798	0.000	0.224	0.549	0.325	0.745	0.000	0.403	0.628	0.366
		B	0.731	-0.016	0.267	0.629	0.369	0.774	-0.030	-0.036	0.555	0.329	0.734	-0.015	0.267	0.630	0.369
		C	0.709	-0.045	-0.004	0.637	0.377	0.777	-0.026	-0.003	0.554	0.328	0.713	-0.043	0.006	0.638	0.375
9	05438500	A	0.332	0.000	0.063	1.1-14	1.306	0.266	0.000	0.213	0.983	1.082	0.371	0.000	0.001	0.973	1.085
		B	0.248	-0.074	-0.322	1.157	1.320	0.206	-0.059	-0.125	0.998	1.101	0.284	-0.084	-0.460	0.974	1.081
		C	0.318	-0.012	-0.001	1.143	1.306	0.228	-0.037	-0.002	0.990	1.092	0.371	0.000	0,001	0.973	1.085
10	05439500	A	0.976	0.000	0.803	1.066	1.107	1.173	0.000	0.329	1.134	1.267	0.996	0.000	0.795	0.987	1.005
		B	0.763	-0.171	-0.343	1.137	1.189	0.849	-0.212	-0.815	1.264	1.381	0.785	-0.177	-0.418	1.044	1.066
		C	0.829	-0.119	0.002	1.097	1.145	1.079	-0.061	0.002	1.137	1.270	0.859	-0.115	0.000	1.004	1.024

TABLE 6 . continued

#	USGS No.	*	With no modification λ	g	5th	TS1	TS2	With H1 modified λ	g	5th	TS1	Ts2	With L1 modi-vied λ	g	5th	TS1	TS2
I. Flood series with no significant high or low outliers/inIiers continued																	
11	05444000	A	0.842	0.000	0.119	0.447	0.376	0.830	0.000	0.152	0.434	0.351	0.836	0.000	0.130	0.455	0.389
		B	0.742	-0.113	-0.671	0.470	0.391	0.734	-0.112	-0.627	0.457	0.367	0.736	-0.114	-0.643	0.480	0.405
		C	0.327	-0.017	0.003	0.447	0.374	0.811	-0.022	0.002	0.434	0.349	0.819	-0.020	0.002	0.455	0.387
12	05445500	A	0.142	-0.001	0.497	0.846	0.564	0.190	-0.001	0.337	0.784	0.523	0.172	0.000	0.416	0.875	0.577
		B	0.144	0.002	0.525	0.846	0.564	0.191	0.001	0.350	0.784	0.523	0.175	0.004	0.461	0.875	0.577
		C	0.107	-0.053	0.002	0.843	0.567	0.167	-0.037	0.005	0.786	0.525	0.145	-0.042	0.007	0.877	0.579
13	05447000	A	0.872	0.000	0.726	0.90S	0.763	1.030	0.000	0.000	0.746	0.701	0.892	-0.001	0.703	0.847	0.679
		B	0.803	-0.099	-0.082	0.953	0.813	0.897	-0.155	-1.024	0.836	0.790	0.822	-0.106	-0.167	0.892	0.724
		C	0.810	-0.089	-0.001	0.946	0.806	1.030	0.000	0.000	0.746	0.701	0.836	-0.085	0.007	0.879	0.711
14	05148000	A	3.272	0.000	0.313	0.865	0.905	0.220	0.000	0.458	0.744	0.718	0.285	0.000	0.268	0.822	0.844
		B	0.229	-0.060	-0.043	0.872	0.912	0.190	-0.046	0.163	0.752	0.728	0.242	-0.063	-0.115	0.828	0.848
		C	0.235	-0.052	0.004	0.870	0.909	0.173	-0.071	-0.003	0.762	0.738	0.255	-0.044	0.001	0.824	0.844
			With no modification					With H1 modified					With H1 and H2 modified				
II. Flood series with high outliers/inliers																	
1	05574000	A	-0.587	0.000	-0.036	0.443	0.351	-0.594	0.000	-0.024	0.436	0.339	-0.529	0.000	-0.090	0.463	0.387
		B	-0.493	0.130	0.913	0.466	0.366	-0.501	0.130	0.942	0.459	0.353	-0.436	0.114	0.713	0.476	0.395
		C	-0.583	0.005	0.002	0.443	0.350	-0.592	0.002	-0.004	0.436	0.339	-0.518	0.013	0.002	0.462	0.385
2	05575500	A	0.058	0.000	0.340	0.476	0.288	0.130	0.000	0.037	0.426	0.294	0.057	-0.001	0.157	0.450	0.276
		B	0.055	-0.006	0.291	0.476	0.286	0.118	-0.019	-0.110	0.427	0.295	0.054	-0.006	0.112	0.450	0.276
		C	0.037	-0.039	-0.002	0.478	0.291	0.127	-0.005	0.000	0.427	0.294	0.047	-0.019	0.006	0.450	0.276
3	05576500	A	0.380	0.001	3.460	3.132	2.636	0.593	0.000	0.753	1.158	0.880	0.727	0.000	0.189	0.618	0.478
		B	0.411	0.107	5.089	3.154	2.656	0.593	0.000	0.753	1.158	0.880	0.691	-0.061	-0.303	0.643	0.498
		C	0.311	-0.235	0.001	3.266	2.776	0.557	-0.078	0.000	1.197	0.916	0.713	-0.024	-0.001	0.624	0.483
4	05578500	A	-0.079	0.000	0.309	1.069	1.097	0.035	0.000	-0.007	1.216	1.330	-0.047	-0.001	0.108	1.019	1.077
		B	-0.063	0.020	0.452	1.072	1.099	0.025	-0.010	-0.061	1.214	1.329	-0.035	0.013	0.192	1.019	1.077
		C	-0.114	-0.043	0.005	1.070	1.101	0.036	0.001	-0.002	1.216	1.330	-0.062	-0.017	0.004	1.020	1.079
5	05580500	A	-0.095	0.000	0.482	0.851	0.359	-0.156	0.000	0.634	0.823	0.797	0.101	0.000	0.245	0.928	0.988
		B	-0.080	0.016	0.590	0.852	0.859	-0.135	0.025	0.814	0.825	0.796	0.078	-0.019	0.131	0.929	0.990
		C	-0.164	-0.073	0.004	0.863	0.878	-0.235	-0.091	-0.001	0.842	0.826	0.051	-0.041	-0.001	0.933	0.995

TABLE 6. (concluded)

#	USGS No.	*	With no modification					With H1 modified					With H1 and H2 modified				
			λ	g	5th	TS1	TS2	λ	g	5th	TS1	TS2	λ	g	5th	TS1	TS2
II. Flood series with high outliers/inliers continued																	
6	05583000	A	0.228	0.000	3.733	2.542	2.337	0.741	0.000	0.059	1.009	1.195	0.625	0.000	0.361	0.731	0.795
		B	0.235	0.017	3.989	2.555	2.349	0.597	-0.143	-0.735	1.044	1.219	0.531	-0.109	-0.352	0.767	0.831
		C	0.107	-0.290	-0.002	2.574	2.399	0.730	-0.011	0.000	1.005	1.190	0.577	-0.056	-0.002	0.736	0.800
7	03380500	A	0.113	-0.001	0.649	1.364	1.093	0.317	0.000	-0.476	0.858	0.815	0.281	0.000	-0.419	0.811	0.725
		B	0.115	0.004	0.712	1.365	1.094	0.284	-0.049	-0.869	0.863	0.813	0.257	-0.039	-0.745	0.816	0.726
		C	0.093	-0.053	0.015	1.364	1.092	0.359	0.060	-0.006	0.875	0.840	0.313	0.050	-0.002	0.820	0.739
			With no modification					With L1 modified					With L1 and L2 modified				
III. Flood series with low outliers/inliers																	
1	05574500	A	0.367	0.000	0.061	0.673	0.444	0.369	0.001	0.067	0.674	0.444	0.027	-0.001	0.355	0.571	0.461
		B	0.356	-0.023	-0.156	0.678	0.448	0.357	-0.024	-0.170	0.679	0.448	0.024	-0.004	0.328	0.571	0.461
		C	0.364	-0.006	0.002	0.674	0.445	0.366	-0.005	0.008	0.675	0.445	-0.013	-0.048	-0.002	0.580	0.472
2	05576000	A	0.215	0.001	0.433	0.679	0.527	0.229	0.000	0.393	0.673	0.511	0.115	0.000	0.516	0.720	0.609
		B	0.194	-0.032	0.164	0.683	0.531	0.208	-0.033	0.110	0.677	0.515	0.100	-0.019	0.366	0.722	0.612
		C	0.181	-0.052	-0.003	0.687	0.535	0.200	-0.046	0.003	0.679	0.517	0.063	-0.067	0.000	0.733	0.624
3	05577500	A	0.143	0.000	-0.706	0.832	0.720	0.195	0.000	-0.851	0.835	0.685	-0.159	0.001	-0.352	0.688	0.657
		B	0.132	-0.021	-0.880	0.837	0.723	0.184	-0.024	-1.056	0.840	0.689	-0.129	0.036	-0.109	0.692	0.662
		C	0.189	0.087	-0.004	0.831	0.723	0.243	0.100	0.000	0.836	0.693	-0.115	0.053	0.003	0.695	0.667
4	05579500	A	0.322	-0.001	-1.969	1.074	0.853	0.063	0.001	-0.790	0.668	0.531	0.110	0.000	-0.869	0.708	0.559
		B	0.325	0.006	-1.893	1.072	0.851	0.058	-0.007	-0.848	0.668	0.531	0.103	-0.011	-0.958	0.708	0.559
		C	0.409	0.194	-0.005	1.073	0.872	0.134	0.106	0.002	0.690	0.561	0.182	1.113	0.003	0.729	0.588
5	05441000	A	0.576	0.000	0.151	0.978	1.088	0.561	0.000	0.185	1.042	1.172	0.594	0.001	0.136	0.920	1.022
		B	0.439	-0.159	-0.736	1.013	1.108	0.420	-0.155	-0.639	1.086	1.208	0.456	-0.166	-0.810	0.952	1.037
		C	0.552	-0.027	0.003	0.972	1.080	0.529	-0.035	0.003	1.038	1.167	0.573	-0.024	0.000	0.914	1.013
6	03379500	A	0.163	0.001	-0.523	1.024	0.958	0.087	-0.001	-0.065	0.941	0.968	0.138	0.000	-0.328	0.934	0.910
		B	0.154	-0.015	-0.660	1.027	0.959	0.079	-0.012	-0.146	0.941	0.967	0.127	-0.018	-0.374	0.935	0.909
		C	0.199	0.064	0.006	1.032	0.976	0.093	0.008	-0.004	0.941	0.969	0.158	0.033	0.055	0.938	0.918
7	03381500	A	0.024	-0.001	-1.450	1.773	1.442	0.002	0.001	-1.322	1.709	1.388	-0.537	0.000	-0.355	0.513	0.340
		B	0.026	0.004	-1.401	1.722	1.441	0.002	0.001	-1.322	1.709	1.388	-0.507	0.036	-0.066	0.518	0.345
		C	0.084	0.133	0.007	1.767	1.440	0.059	0.124	-0.007	1.704	1.387	-0.500	0.044	0.001	0.520	0.347

* Criterion used for deriving λ

$$(p_o)_i = \frac{i - \alpha}{n + 1 - 2\alpha} \qquad (14)$$

and $(Z_o)_i$ is obtained from $(p_o)_i$ by a $p \rightarrow Z$ subroutine for normal distribution.

Category I: The following inferences can be made from the results in Table 6 for flood series without significant high or low outliers/inliers.

1. TS2 is lower than TS1 for about 50% of the basins. A lower value of the test statistic shows an overall better fit.

2. For a basin, the TS1 or TS2 values for the three criteria A, B, and C are quite close to each other with (or without) modification of any outliers/inliers.

3. Minimum values of the 5th are obtained with the criterion that minimizes [|g| + |5th|].

4. For a given basin, the values of λ for the three criteria are not much different from each other, but the ML estimate of λ is generally somewhat smaller than those with g = 0 and min [|g| + |5th|].

5. The values of λ for the three cases: with no modification of any outlier/inlier, with modification of highest inlier/outlier, and with modification of lowest outlier/inlier, are not much different from each other when the flood series are well behaved, i.e., they do not have significant high and low outliers/inliers.

Category II: The following inferences can be made from the results in Table 6 for flood series with high outliers/inliers.

1. TS2 is lower than TS1 for about two-thirds of the basins. A lower value of the test statistic shows an overall better fit. Thus, the use of $\alpha = 0.38$ seems better than $\alpha = 0.00$.

2. For a basin, the TS1 or TS2 values for the three criteria A, B, and C are rather close to each other with (or without) modification of

outliers/inliers. However, for a given criterion, these values vary considerably from each other when obtained with and without modification of outliers/inliers for three basins, each with a rather severe high outlier; e.g., in the case of basin 3 and criterion A, TS2 decreases from 2.636 with no modification of outliers, to 0.880 with modification of H1, and to 0.478 with modification of H1 and H2. The TS1 or TS2 values, after modification of H1 and H2, lie in the general range of 0.4 to 1.0, the same as for category I.

3. Minimum values of the 5th are obtained with the criterion that minimizes [|g| + |5th|].

4. For a given basin, the values of X for the three criteria, after modification of outliers/inliers, are not much different from each other but the ML estimate of X is more often somewhat smaller than those with the other criteria.

5. The value of X changes with modification of any outlier/inlier and the magnitude of change depends on the severity of the outlier/inlier.

<u>Category III</u>: The following inferences can be made from the results in Table 6 for flood series with low outliers/inliers.

1. TS2 is generally less than TS1 with and without modification of any outliers/inliers. A lower value of the test statistic indicates an overall better fit. Thus, the use of a = 0.38 seems better than a = 0.00.

2. For a basin, the TS1 or TS2 values for the three criteria A, B, and C are rather close to each other with (or without) modification of outliers/inliers. However, for a given criterion, these values vary considerably from each other when obtained with and without modification of outliers/inliers for three basins, each with 1 or 2 rather severe low outliers. The TS1

or TS2 values, after modification of L1 and L2, lie in the general range of 0.4 to 1.0, the same as for categories I and II.

3. Minimum values of the 5th are obtained with the criterion that minimizes $[|g| + |5th|]$.

4. For a given basin, the values of λ for the three criteria, after modification of outliers/inliers, are not much different from each other.

5. The value of λ changes with modification of any outlier/inlier and the magnitude of change depends on the severity of the outlier/inlier.

Wilson-Hilferty Transformation Results

The results are presented in Table 7. Criteria A, B, and C denote transformation as expressed by equation 8 (g_s = skew of x series in equation 7 and g = skew of y series in equation 8); iterative modification of g_s so that g becomes zero (g_s equals the value of g used in equation 8 so that skew of y series becomes zero); and iterative modification of g_s so that $[|g| + |5th|]$ of y series becomes minimum, respectively. The 5th, TS1, and TS2 are the same as defined under power transformation or earlier.

Category I. The following inferences can be made from the results in Table 7 for flood series without significant high or low outliers/inliers.

1. TS1 is lower than TS2 for about 50% of the basins. A lower value of the test statistic shows an overall better fit.

2. The TS1 or TS2 values for the three criteria A, B, and C are quite close to each other with (or without) modification of any outliers/inliers in the case of 11 basins, but for 3 basins these values are considerably higher with A than with B or C.

TABLE 7. Evaluation of Normalization of a Flood Series by Wilson-Hilferty Transformation

#	USGS No.	*	With no modification					With H1 modified					With L1 modified				
			g_s	g	5th	TS1	TS2	g_s	g	5th	TS1	TS2	g_s	g	5th	TS1	TS2
I.	Flood series with no significant high or low outliers/inliers																
1	05571000	A	-0.098	-0.034	-0.968	0.444	0.320	-0.129	-0.040	-0.638	0.457	0.352	-0.077	-0.028	-0.382	0.446	0.330
		B	-0.149	0.000	-0.207	0.446	0.319	-0.205	0.000	-0.284	0.460	0.351	-0.120	0.000	-0.171	0.447	0.329,
		C	-0.190	0.027	0.001	0.45?	0.323	-0.265	0.038	0.001	0.473	0.358	-0.155	0.023	0.001	0.452	0.332
2	05572000	A	-0.308	-0.049	-0.781	0.529	0.422	-0.300	-0.045	-0.730	0.524	0.404	-0.143	-0.044	-0.209	0.463	0.472
		B	-0.365	0.000	-0.364	0.535	0.415	-0.353	0.000	-0.339	0.530	0.398	-0.205	0.000	0.097	0.461	0.473
		C	-0.415	0.043	-0.001	0.563	0.428	-0.398	0.039	0.000	0.554	0.409	-0.186	-0.014	0.001	0.459	0.464
3	05582000	A	-0.073	-0.027	0.817	0.832	0.834	-0.007	-0.002	0.847	0.889	0.807	-0.097	-0.034	0.426	0.844	0.777
		B	-0.114	0.000	0.720	0.884	0.833	-0.010	0.000	0.864	0.889	0.807	-0.148	0.000	0.688	0.849	0.777
		C	-0.340	-0.097	0.001	0.903	0.858	0.158	-0.111	-0.001	0.925	0.842	-0.010	-0.091	0.001	0.857	0.793
4	05435500	A	-0.055	-0.020	-0.141	0.667	0.764	-0.031	-0.011	-0.010	0.604	0.666	-0.086	-0.029	-0.280	0.577	0.632
		B	-0.086	0.000	-0.002	0.665	0.761	-0.046	0.000	0.054	0.603	0.664	-0.131	0.000	-0.097	0.576	0.629
		C	-0.092	0.004	-0.001	0.665	0.761	-0.033	-0.009	0.000	0.604	0.666	-0.155	0.016	0.001	0.580	0.631
5	05437000	A	-0.315	-0.190	-0.809	0.955	1.080	-0.304	-0.157	-0.543	0.825	0.893	-0.420	-0.215	-1.101	0.767	0.840
		B	-0.7i6	0.000	-0.208	0.91:1	0.957	-0.509	0.000	0.384	0.782	0.786	-0.806	0.000	0.131	0.752	0.731
		C	-0.648	-0.030	0.000	0.892	0.960	-0.481	-0.063	0.001	0.772	0.809	-0.770	-0.023	0.000	0.732	0.726
6	05437500	A	-0.128	-0.064	-0.023	0.094	1.044	-0.141	-0.072	-0.100	0.977	1.081	-0.198	-0.091	-0.319	0.736	0.757
		B	-0.253	0.000	0.337	0.927	1.002	-0.283	0.000	0.292	0.947	1.033	-0.359	0.000	0.236	0.717	0.714
		C	-0.136	-0.060	0.000	0.951	1.040	-0.178	-0.054	0.000	0.964	1.064	-0.293	-0.038	0.001	0.716	0.725
7	05440000	A	-0.449	-0.276	-1.342	1.431	1.669	-0.431	-0.257	-1.217	1.295	1.491	-0.525	-0.302	-1.652	1.225	1.412
		B	-1.056	0.000	0.071	1.316	1.369	-0.963	0.000	0.215	1.165	1.193	-1.132	0.000	0.014	1.133	1.116
		C	-1.033	-0.013	-0.001	1.297	1.365	-0.903	-0.036	0.001	1.134	1.198	-1.129	-0.002	0.000	1.129	1.115
8	05447500	A	-1.080	0.344	4.905	1.831	1.301	-1.109	0.217	2.625	1.338	0.906	-1.091	0.373	5.336	1.946	1.395
		B	-0.939	0.000	0.591	0.985	0.673	-1.006	0.000	0.287	0.886	0.597	-0.944	0.000	0.584	0.993	0.679
		C	-0.908	-0.054	-0.001	0.923	0.641	-0.989	-0.029	-0.003	0.851	0.579	-0.913	-0.054	-0.004	0.931	0.647
9	05438500	A	-0.298	-0.175	-0.845	1.183	1.362	-0.262	-0.145	-0.605	1.023	1.143	-0.370	-0.202	-1.144	0.978	1.099
		B	-0.690	0.000	0.075	1.128	1.229	-0.560	0.000	0.241	0.973	1.031	-0.784	0.000	0.004	0.940	0.969
		C	-0.662	-0.014	0.000	1.119	1.230	-0.483	-0.040	0.001	0.964	1.046	-0.782	-0.001	0.000	0.940	0.969
10	05439500	A	-0.778	-0.418	-1.951	1.283	1.436	-0.812	-0.479	-2.410	1.457	1.702	-0.866	-0.427	-2.195	1.136	1.251
		B	-1.246	0.000	1.495	1.198	1.046	-1.470	0.001	0.726	1.087	0.912	-1.288	-0.001	1.549	1.119	0.926
		C	-1.138	-1.154	0.001	1.013	0.990	-1.410	-0.090	-0.001	0.941	0.868	-1.198	-1.151	0.002	0.915	0.845

TABLE 7. continued

#	USGS No.	*	With no modification g_s	g	5th	TS1	TS2	With H1 modified g_s	g	5th	TS1	TS2	With L1 modified g_s	g	5th	TS1	TS2
I.	Flood series with no significant high or low outliers/inliers continued																
11	05444000	A	-1.106	-0.243	-1.815	0.475	0.418	-1.101	-0.231	-1.696	0.468	0.394	-1.070	-0.256	-1.849	0.486	0.443
		B	-1.296	0.000	0.096	0.597	0.366	-1.277	0.000	0.171	0.592	0.354	-1.279	0.000	0.128	0.596	0.373
		C	-1.290	-0.011	0.004	0.581	0.359	-1.264	-0.020	-0.031	0.564	0.342	-1.269	-0.015	0.002	0.574	0.364
12	05445500	A	-0.210	0.021	0.712	0.364	0.577	-0.237	0.015	0.479	0.807	0.540	-0.265	0.038	0.795	0.909	0.602
		B	-0.191	0.000	0.508	0.857	0.573	-0.253	0.000	0.346	0.801	0.537	-0.233	0.000	0.420	0.893	0.592
		C	-0.144	0.053	-0.003	0 348	0.570	-0.218	-0.038	-0.001	0.791	0.532	-0.197	-0.043	0.002	0.832	0.587
13	05447000	A	-1.279	-0.012	2.004	1.612	1.402	-1.329	-0.353	-2.373	1.032	1.089	-1.420	1.627	40.15	10.43	1.293
		B	-1.282	-0.001	2.150	1.632	1.416	-1.521	0.000	0.774	1.154	0.937	-1.344	0.000	2.325	1.682	1.435
		C	-1.218	-1.176	-0.002	1.396	1.288	-1.495	-0.076	-0.005	1.041	0.887	-1.291	-0.780	-0.008	1.420	1.280
14	05'!48000	A	-0.315	-0.167	-0.670	0.883	0.939	-0.328	-0.133	-0.402	0.763	0.753	-0.412	-0.174	-0.331	0.833	0.868
		B	-0.639	0.000	0.308	0.878	0.872	-0.526	0.000	0.471	0.752	0.694	-0.675	0.000	0.255	0.837	0.808
		C	-0.564	-0.051	0.000	0.852	0.871	-0.424	-0.071	-0.001	0.734	0.711	0.618	-0.041	-0.001	0.811	0.802
			With no modification					With H1 modified					With H1 and H2 modified				
II.	Flood series with high outliers/inliers																
1	05574000	A	0.956	0.303	2.417	0.436	0.379	1.000	0.298	2.445	0.423	0.358	0.723	0.287	2.070	0.459	0.414
		B	1.267	0.000	-0.059	0.508	0.264	1.292	0.000	-0.040	0.507	0.255	1.001	0.000	-0.097	0.487	0.284
		C	1.262	0.007	-0.002	0.500	0.261	1.289	0.004	-0.001	0.502	0.252	1.088	0.012	-0.002	0.474	0.280
2	05575500	A	-0.108	-0.021	0.160	0.476	0.288	-0.210	-0.060	-0.429	0.427	0.296	-0.105	-0.022	-0.021	0.450	0.276
		B	-0.134	0.000	0.344	0.478	0.289	-0.294	0.000	0.035	0.432	0.290	-0.130	0.000	0.163	0.452	0.276
		C	-0.086	-0.039	0.000	0.477	0.290	-0.287	-0.004	-0.001	0.430	0.290	-0.108	-0.020	-0.001	0.450	0.276
3	05576500	A	-1.257	3.179	122.3	40.28	37.89	-1.442	2.453	55.64	37.92	35.58	-1.527	1.948	32.12	32.07	29.86
		B	-0.699	0.000	4.703	4.313	3.833	-1.065	-0.001	1.193	2.345	1.997	-1.280	0.000	0.328	1.527	1.238
		C	-0.605	-0.261	-0.001	3.882	3.497	-1.029	-0.094	0.003	2.166	1.878	-1.270	-0.029	0.009	1.471	1.206
4	05578500	A	0.099	0.051	0.687	1.076	1.102	-0.033	-0.021	-0.125	1.213	1.328	0.052	0.030	0.299	1.020	1.078
		B	0.207	0.000	0.315	1.068	1.090	-0.089	0.000	-0.006	1.215	1.328	0.121	0.000	0.111	1.018	1.075
		C	0.300	-0.045	0.001	1.080	1.092	-0.092	0.001	0.000	1.215	1.328	0.162	0.018	0.001	1.021	1.075
5	05580500	A	0.104	0.045	0.789	0.853	0.859	0.192	0.072	1.167	0.826	0.796	-0.084	-0.046	-0.028	0.931	0.994
		B	0.184	0.000	0.489	0.852	0.857	0.311	0.000	0.648	0.828	0.791	-0.184	0.000	0.247	0.928	0.983
		C	0.320	-0.075	0.001	0.889	0.880	0.468	-0.094	0.000	0.894	0.831	-0.094	-0.041	0.001	0.929	0.992

TABLE 7. (concluded)

#	USGS No.	*	With no modification g_s	g	5th	TS1	TS2	With H1 modified g_s	g	5th	TS1	TS2	With H1 and H2 modified g_s	g	5th	TS1	TS2
II. Flood series with high outliers/inliers continued																	
6	05583000	A	-0.520	0.165	6.928	3.162	2.873	-0.819	-0.361	-2.127	1.128	1.366	-0.775	-0.293	-1.584	0.825	0.936
		B	-0.434	0.000	4.078	2.743	2.523	-1.275	0.000	0.211	0.763	0.675	-1.087	0.000	0.695	0.711	0.581
		C	-0.220	-0.293	0.001	2.538	2.399	-1.251	-0.029	0.001	0.718	0.665	-1.021	-0.080	0.003	0.623	0.571
7	03380500	A	-0.293	0.039	1.207	1.439	1.150	-0.538	-0.131	-1.635	0.841	0.781	-0.505	-0.104	-1.372	0.808	0.707
		B	-0.261	-0.001	0.698	1.405	1.260	-0.711	0.000	-0.538	0.854	0.734	-0.633	0.000	-0.459	0.829	0.683
		C	-0.214	-0.055	0.003	1.379	1.108	-0.787	0.064	0.000	0.927	0.763	-0.692	0.051	0.001	0.880	0.704
			With no modification					With L1 modified					With L1 and L2 modified				
III. Flood series with low outliers/inliers																	
1	05574500	A	-0.822	0.035	0.534	0.845	0.565	-0.829	0.040	0.882	0.853	0.569	-0.032	-0.012	0.265	0.572	0.462
		B	-0.796	0.000	0.150	0.814	0.549	-0.800	0.000	0.146	0.817	0.551	-0.052	0.000	0.360	0.571	0.460
		C	-0.785	-0.014	0.001	0.803	0.544	-0.789	-0.015	-0.004	0.806	0.546	0.025	-0.047	0.000	0.581	0.472
2	05576000	A	-0.331	-0.096	-0.371	0.683	0.537	-0.367	-0.097	-0.443	0.678	0.520	-0.147	-0.057	0.075	0.724	0.617
		B	-0.454	0.000	0.432	0.696	0.524	-0.487	0.000	0.392	0.694	0.510	-0.236	0.000	0.516	0.722	0.606
		C	-0.390	-0.051	0.002	0.682	0.526	-0.432	-0.046	0.001	0.680	0.509	-0.131	-0.067	0.000	0.726	0.620
3	05577500	A	-0.291	-0.065	-1.246	0.844	0.731	-0.458	0.063	-1.427	0.854	0.697	0.187	0.093	0.276	0.696	0.673
		B	-0.377	0.000	-0.715	0.847	0.722	-0.533	0.000	-0.872	0.869	0.696	0.363	0.000	-0.367	0.689	0.642
		C	-0.491	0.088	0.000	0.895	0.743	-0.647	0.100	0.002	0.946	0.733	0.265	0.053	0.000	0.685	0.654
4	05579500	A	-0.961	0.235	0.338	1.535	1.172	-0.096	-0.022	-0.966	0.668	0.530	-0.182	-0.035	-1.145	0.709	0.558
		B	-0.779	0.000	-2.193	1.211	0.968	-0.126	-0.001	-0.797	0.670	0.532	-0.226	0.000	-0.876	0.713	0.561
		C	-0.939	0.204	-0.001	1.471	1.127	-0.269	0.106	0.001	0.721	0.571	-0.372	0.113	0.002	0.780	0.607
5	05441000	A	-0.746	-0.376	-2.231	1.053	1.201	-0.651	-0.363	-1.890	1.164	1.339	-0.819	-0.395	-2.427	0.981	1.127
		B	-1.310	0.000	0.118	0.904	0.775	-1.253	0.000	0.204	0.985	0.902	-1.369	0.000	0.104	0.836	0.681
		C	-1.292	-0.018	0.001	0.872	0.764	-1.214	-0.034	0.000	0.936	0.886	-1.355	-0.015	0.001	0.807	0.670
6	03379500	A	-0.305	-0.04S	-0.998	1.035	0.959	-0.129	-0.041	-0.348	0.940	0.965	-0.230	-0.055	-0.677	0.934	0.905
		B	-0.364	0.000	-0.582	1.041	0.956	-0.188	0.000	-0.066	0.939	0.961	-0.303	0.000	-0.258	0.936	0.899
		C	-0.448	0.071	0.001	1.099	0.991	-0.202	0.009	0.000	0.942	0.962	-0.348	0.035	0.001	0.956	0.910
7	03381500	A	-0.055	0.021	-1.208	1.772	1.440	-0.004	0.001	-1.314	1.709	1.388	0.651	0.084	0.294	0.562	0.381
		B	-0.040	0.000	-1.444	1.774	1.442	-0.003	0.000	-1.332	1.709	1.388	0.729	0.000	-0.448	0.612	0.391
		C	-0.137	0.132	0.002	1.797	1.459	-0.096	0.125	0.001	1.722	1.398	0.683	0.051	0.001	0.576	0.380

* Criterion used for deriving g_s.

3. The values of the 5th with criterion A are generally much higher than with ML estimate of λ. Though the criterion C minimized values of the 5th, it decreases skew very significantly (-1.154 and -1.176) below zero for two basins.

4. The values of g for the three criteria are significantly different from each other for 4 out of the 14 basins.

5. The absolute values of g with Wilson-Hilferty transformation are considerably higher than with the ML estimate of λ. Thus, the power transformation brings a flood series closer to the normal distribution than the Wilson-Hilferty transformation.

Category II. The following inferences can be made from the results in Table 7 for flood series with high outliers/inliers.

1. TS2 is lower than TS1 for about two-thirds of the basins.

2. For USGS No. 05 576500 with a significant high outlier, the Wilson-Hilferty transformed series yields g = 3.179 compared to 0.107 with the power transformation and ML estimate of λ. The absolute value of g for other basins is also somewhat higher than with the power transformation.

3. Values of the 5th with Wilson-Hilferty transformation are higher than with the power transformation.

4. The values of g for a given criterion but considering no modification and modification of H1 or H1 and H2 differ significantly for two basins out of seven. With power transformation, there are no significant differences.

Category III. The following inferences can be made from the results in Table 7 for flood series with low outliers/inliers.

1. TS2 is generally less than TS1. A lower value of test statistic indicates a better overall fit. Thus, the use of a = 0.38 seems better than

$\alpha = 0.00$.

2. The absolute value of g for the y series with the Wilson-Hilferty transformation is generally higher than with the power transformation.

3. Minimum values of the 5th are obtained with the criterion that minimizes $[|g| + |5th|]$.

Three-Parameter Lognormal Transformation

The results are presented in Table 8. Criterion A corresponds to a value of a which reduces skew of y series (equation 9) to zero; criterion B denotes the value of a that minimizes $[|g| + |5th|]$; and criterion C refers to the value of a which makes kurtosis equal to 3.0. The 5th, TS1, and TS2 have been defined earlier. It is evident from Table 8 that no results are obtained for nine basins out of a total of 28 basins analyzed. This transformation is not suitable for converting a flood series to approximately a normal distribution.

Category I: Some inferences of interest are:

1. Making the kurtosis = 3.0 (criterion C) increases tremendously the absolute values of g and 5th, and to some extent TS1 and TS2.

2. Out of 14 basins with flood series having no significant low or high outliers/inliers, results for all the criteria and modification were obtained for only seven basins.

Category II: Inferences for Category I apply to this category also.

Category III. The same remarks as for Category II apply to this category. Complete results are obtained, however, for five out of seven basins.

Selection of Transformation

The power transformation is considered the best of the tested

TABLE 8. Evaluation of Normalization of a Flood Series by 3-Parameter Lognormal Transformation

			With no modification					With H1 modified					With L1 modified				
#	USGS No.	*	a	g	5th	TS1	TS2	a	g	5th	TS1	TS2	a	g	5th	TS1	TS2
I. Flood series with no significant high or low outliers/inliers																	
1	05571000	A	284.0	0.000	-0.198	0.452	0.337	402.0	0.000	-0.265	0.471	0.379	228.0	0.000	-0.166	0.453	0.345
		B	373.0	0.026	0.000	0.456	0.344	548.0	0.037	0.003	0.480	0.392	303.0	0.022	0.000	0.457	0.351
		C	313.0	-0.590	-4.576	0.971	0.374	361.0	-0.261	-4.780	1.068	0.990	328.0	-0.605	-4.651	1.010	0.917
2	05572000	A	900.0	0.000	-0.180	C.502	0.458	862.0	0.000	-0.163	0.491	0.433	526.0	0.000	0.107	0.508	0.543
		B	1007.0	0.024	0.010	0.506	0.467	950.0	0.020	0.001	0.493	0.440	458.0	-0.016	0.000	0.504	0.536
		C	-57.8	.0.571	3.967	2.524	2.476	-56.3	0.554	3.916	2.376	2.312	-67.2	0.708	5.057	3.490	3.566
3	05582000	A	722.0	0.000	0.709	0.897	0.355	56.1	0.000	0.862	0.890	0.809	94.0	0.000	0.679	0.863	0.805
		B	219.0	-0.097	0.000	0.894	0.848	-798.0	-0.111	0.000	0.893	0.807	69.3	-0.089	0.013	0.858	0.795
		C	-239.0	0.494l	4.809	2.087	2.090	-206.0	0.417	4.445	1.745	1.698	-229.0	0.472	4.627	1.921	1.896
4	05435500	A	289.0	0.000	-0.021	0.682	0.786	152.0	0.000	0.054	0.611	0.677	441.0	0.000	-0.088	0.560	0.667
		B	311.0	0.004	0.002	0.683	0.788	107.0	-0.009	0.000	0.610	0.675	525.0	0.014	0.000	0.605	0.675
		C	378.0	-0.799	-5.264	3.383	3.514	-81.5	0.775	5.045	4.220	4.346	356.0	-0.763	-5.256	3.206	3.319
5	05437000	A	9040.0	-0.00S	0.166	0.932	1.043										
		B	7413.0	-0.037	0.004	0.938	1.049	4345.0	-0.060	0.000	0.823	0.882					
		C	-263.0	0.702	4.437	3.716	3.956	_227.0	0.571	4.580	2.337	2.419					
6	05437500	A	3256.0	0.000	0.325	0.951	1.041	3765.0	0.000	0.281	0.973	1.076	4947.0	0.000	0.232	0.750	0.777
		B	1619.0	-0.057	0.015	0.971	1.064	210S.0	-0.053	0.000	0.992	1.096	3657.0	-0.039	0.000	0.756	0.782
		C	-283.0	0.654	4.934	3.283	3.428	-298.0	0.687	4.905	3.652	3.829	-278.0	0.631	4.821	2.626	2.654
7	05444000	A															
		B															
		C															
8	05447500	A															
		B															
		C	276.0	-0.707	-5.183	1.627	1.584	268.0	-0.714	-5.237	1.670	1.636	283.0	-0.713	-5.240	1.651	1.609
9	05438500	A	1614.0	0.000	0.056	1.279	1.467	1213.0	0.000	0.185	1.099	1.226	1908.0	0.000	0.014	1.128	1.275
		B	1562.0	-0.009	0.011	1.276	1.464	989.0	-0.034	0.000	1.095	1.221	1882.0	-0.003	0.000	1.127	1.274
		C															
10	05439500	A															
		B															
		C	-115.0	0.006	3.111	4.104	4.161										

I. Flood series with no significant high or low outliers/inliers continued

			With no modification					With H1 modified					With L1 modified				
	USGS No.	*	a	g	5th	TS1	TS2	a	g	5th	TS1	TS2	a	g	5th	TS1	TS2
11	05444000	A															
		B															
		C	-32.4	-0.437	-3.068	1.020	0.913	-31.2	-0.429	-2.994	0.989	0.874	-40.4	-0.396	-2.730	1.020	0.909
12	05445500	A	342.0	0.000	0.500	0.804	0.531	403.0	0.000	0.351	0.740	0.489	424.0	0.000	0.426	0.821	0.534
		B	248.0	-0.050	0.031	0.815	0.541	398.0	-0.038	0.014	0.746	0.495	338.0	-0.044	0.000	0.830	0.542
		C	88.5	-0.524	-3.828	0.061	0.710	87.3	-0.498	3.580	0.807	0.674	95.9	-0.542	-4.051	0.879	0.726
13	05447000	A															
		B															
		C	120.0	-0.785	-5.265	2.082	2.039	112.0	-0.811	-5.450	2.261	2.302	120.0	-0.823	-5.771	2.141	2.123
14	05448000	A	917.0	0.000	0.333	0.934	1.042	6S3.0	0.000	0.457	0.827	0.840	965.0	0.000	0.309	0.924	0.988
		B	718.0	-0.055	0.015	0.964	1.038	483.0	-0.075	-0.003	0.836	0.845	770.0	-0.054	-0.007	0.922	0.982
		C	-91.3	0.712	4.859	3.878	4.112	-84.6	0.651	4.907	3.179	3.311	-90.7	0.714	4.882	3.727	3.986

II. Flood series with high outliers/inliers

			With no modification					With H1 modified					With H1 and H2 modified				
	USGS No.	*	a	g	5th	TS1	TS2	a	g	5th	TS1	TS2	a	g	5th	TS1	TS2
1	05574000	A											-274.0	0.000	-0.193	0.390	0.193
		B											-271.0	0.022	0.001	0.387	0.194
		C											10.6	-0.118	-1.147	0.385	0.173
2	05575500	A	218.0	0.000	0.350	0.467	0.293	527.0	0.000	0.079	0.439	0.331	213.0	0.000	0.187	0.444	0.283
		B	130.0	-0.041	0.000	0.473	0.294	493.0	-0.011	0.000	0.438	0.329	190.0	-0.010	0.103	0.445	0.282
		C	-70.8	0.337	3.055	0.856	0.700	336.0	-0.625	-4.644	1.157	1.071	198.0	-0.485	-3.645	0.797	0.652
3	05576500	A	7719.0	0.000	2.771	2.258	1.863										
		B	5248.0	-0.216	0.061	2.451	2.058										
		C	564.0	-0.667	-4.654	2.445	2.355										
4	05578500	A	-236.0	0.000	0.342	1.001	1.006	116.0	0.000	-0.005	1.248	1.369	-141.0	0.000	0.114	0.975	1.022
		B	-326.0	-0.048	-0.004	0.970	0.965	119.0	0.001	0.000	1.649	1.370	-171.0	-0.012	0.035	0.965	1.009
		C	329.0	0.628	-4.664	1.359	1.273	506.0	-0.736	-5.428	1.775	1.769	387.0	-0.674	-5.029	1.455	1.401
5	05580500	A	-336.0	0.000	0.507	0.827	0.823	-528.0	0.000	0.694	0.790	0.740	431.0	0.000	0.231	0.959	1.027
		B	-538.0	-0.077	0.000	0.818	0.810	-700.0	-0.085	0.097	0.785	0.731	208.0	-0.040	0.000	0.949	1.016
		C	434.0	-0.716	-4.997	1.569	1.531	359.0	-0.676	-4.705	1.410	1.336	590.0	-0.728	-4.899	1.752	1.767

TABLE 8. (concluded)

#	USGS No.	*	With no modification					With H1 modified					With H1 and H2 modified				
			a	g	5th	TS1	TS2	a	g	5th	TS1	TS2	a	g	5th	TS1	TS2
II. Flood series with high outliers/inliers continued																	
6	05583000	A	6292.0	0.000	3.162	2.247	2.103										
		B	2301.0	-0.278	-0.001	2.415	2.275										
		C	295.0	-0.331	-0.875	2.105	2.025										
7	03380500	A	758.0	0.000	0.604	1.204	0.979	2723.0	0.000	-0.345	0.944	0.963	2291.0	0.000	-0.294	0.860	0.837
		B	590.0	-0.050	0.037	1.225	0.991	3309.0	0.049	0.007	0.973	1.002	2679.0	0.040	0.011	0.878	0.863
		C	82.0	-0.226	-1.583	1.125	0.910	699.0	-0.725	-5.397	2.307	2.347	588.0	-0.697	-5.270	2.085	2.097
			With no modification					With L1 modified					With L1 and L2 modified				
III. Flood series with low outliers/inliers																	
1	05574500	A	1736.0	0.000	0.089	0.588	0.398	1745.0	0.000	0.086	0.588	0.397	97.6	0.000	0.361	0.574	0.468
		B	1690.0	-0.010	0.000	0.590	0.400	1700.0	-0.010	0.000	0.590	0.399	-42.0	-0.047	0.000	0.578	0.468
		C	10.2	-0.033	-0.1S0	0.565	0.377	11.2	-0.037	-0.219	0.563	0.375	378.0	-0.641	-4.787	1.074	0.980
2	05576000	A	1203.0	0.000	0.451	0.696	0.571	1291.0	0.000	0.425	0.686	0.554	598.0	0.000	0.525	0.744	0.649
		B	930.0	-0.054	0.017	0.701	0.574	1025.0	-0.053	0.000	0.692	0.557	289.0	-0.069	0.000	0.747	0.648
		C	-128.0	0.409	3.900	1.521	1.426	-122.0	0.387	3.706	1.427	1.319	-162.0	0.524	4.866	2.348	2.346
3	05577500	A	178.0	0.000	-0.653	0.807	0.716	246.0	0.000	-0.743	0.790	0.674	-164.0	0.000	-0.438	0.637	0.575
		B	263.0	0.087	0.028	0.808	0.727	343.0	0.095	0.023	0.794	0.691	-124.0	0.057	-0.003	0.659	0.609
		C	83.6	-0.458	-3.988	0.805	0.690	62.7	-0.380	-3.562	0.717	0.584	97.5	-0.538	-4.569	0.875	0.765
4	05579500	A	570.0	0.000	-1.584	0.692	0.702	98.4	0.000	-0.774	0.660	0.530	179.0	0.000	-0.330	0.689	0.552
		B	877.0	0.171	0.011	0.898	0.731	234.0	0.103	0.000	0.677	0.560	333.0	0.110	0.006	0.705	0.581
		C	-7.7	0.067	-1.113	1.005	0.816	62.2	-0.393	-3.467	0.672	0.544	61.2	-0.384	-3.462	0.669	0.536
5	05441000	A															
		B															
		C															
6	03379500	A	1834.0	0.000	-0.313	1.006	0.992	1005.0	0.000	-0.019	0.983	1.034	1603.0	0.000	-0.145	0.964	0.980
		B	2205.0	0.039	0.000	1.015	1.011	1034.0	0.003	0.002	0.985	1.036	1792.0	0.020	0.000	0.970	0.990
		C	503.0	-0.637	-4.978	2.143	2.151	604.0	-0.697	-5.083	2.556	2.606	553.0	-0.667	-4.988	2.334	2.366
7	03381500	A	277.0	0.000	-1.425	1.746	1.417	19.5	0.000	-1.328	1.707	1.387					
		B	1102.0	0.139	0.107	1.680	1.362	723.0	0.123	0.000	1.645	1.335					
		C	468.0	-0.588	-6.319	1.238	0.995	425.0	-0.559	-5.978	1.207	0.964					

* Criterion used for deriving a

transformations for converting an observed flood series so that it approximates a normally distributed series. The pertinent reasons are:

1. Power transformed series are more stable and consistent even when some outliers/inliers are present.

2. Power transformed series derived with λ from any of the three criteria have similar statistical properties. The maximum log-likelihood method of determining λ. can be used and it is free from bias that may be attributed to $g = 0$ and min $[|g| + |5th|]$ criteria.

3. Overall results obtained from the flood series analyzed with the power transformation are much better than from the Wilson-Hilferty transformation. The 3-parameter lognormal transformation is unsuitable for general use.

METHODS OF ESTIMATING DESIGN FLOODS

Estimation of various recurrence-interval floods was performed basically with three methods: power transformation, log-Pearson type III, and mixed distribution. The background and rationale of these methods are investigated.

Power Transformation Method

1. The observed annual flood series, Q_i, is normalized using the power transformation to y_i series

$$y_i = (Q_i^{\lambda} - 1)/\lambda \qquad \lambda \neq 0 \tag{1}$$

in which the parameter λ is determined by the maximum log-likelihood method.

2. The mean, $\bar{y}$, and standard deviation, s_y, of the normalized series are calculated from

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \tag{2}$$

$$s_y = \sqrt{\left(\sum_{i=1}^{n} (y_i - \bar{y})^2\right)/(n-1)} \tag{3}$$

3. For a desired recurrence interval of T years, the probability of nonexceedance is (1 - 1/T). A standard normal deviate z_T corresponding to this probability is obtained from a p-to-z subroutine (or it can be interpolated from a normal probability table).

4. The T-yr flood is computed from

$$Q_T = (\lambda y_T + 1)^{1/\lambda}, \text{ where } y_T = \bar{y} + z_T\, s_y \tag{4}$$

Effect of Kurtosis ≠ 3

The y. transformed series has a skew very close to zero but the kurtosis, kt, may not equal 3 as for a normal distribution:

$$kt = \frac{n^2 \sum_{i=1}^{n} (y_i - \overline{y})^4}{(n-1)(n-2)(n-3)\ s_y^{\,4}} \tag{5}$$

The normal distribution is compared with the symmetric platykurtic (kt<3) and symmetric leptokurtic (kt>3) distributions in figure 8. The normal distribution function can be modified to express these variations. The following description is based on Box and Tiao (1973).

The standard normal distribution function may be written as

$$p(x) = k \exp\left(-\frac{1}{2}\,|x|^{q}\right) \quad \text{with } q=2 \tag{6}$$

By allowing q to take values other than 2 with the following expression

$$q = 2/(1+\beta); \qquad -1 < \beta \leq 1, \tag{7}$$

the class of exponential power distribution functions can be written in the general form

$$p\,(y|\theta,\phi,\beta) = k\,\phi^{-1} \exp\left(-\frac{1}{2}\left|\frac{y-\theta}{\phi}\right|^{2/(1+\beta)}\right) \tag{8}$$

$$k^{-1} = \Gamma\left(1 + \frac{1+\beta}{2}\right)\, 2^{1+\frac{1}{2}(1+\beta)} \tag{9}$$

in which $-\infty<y<\infty$, $\phi>0$, $-\infty<\theta<\infty$, and $-1<\beta\leq 1$. In equation 9, 0 is a location parameter and ø is a scale parameter. It can be shown that

$$E(y) = \theta \tag{10}$$

$$Var(y) = \sigma^2 = 2^{(1+\beta)} \left\{\frac{\Gamma[3/2\ (1+\beta)]}{\Gamma[1/2\ (1+\beta)]}\right\} \phi^2 \tag{11}$$

The parameter 3 can be regarded as a measure of kurtosis indicating the extent of variation from the normal distribution. In particular, the distribution is normal and double exponential when $\beta=0$ and $\beta=1$, respectively, and the distribution tends to the rectangular distribution as

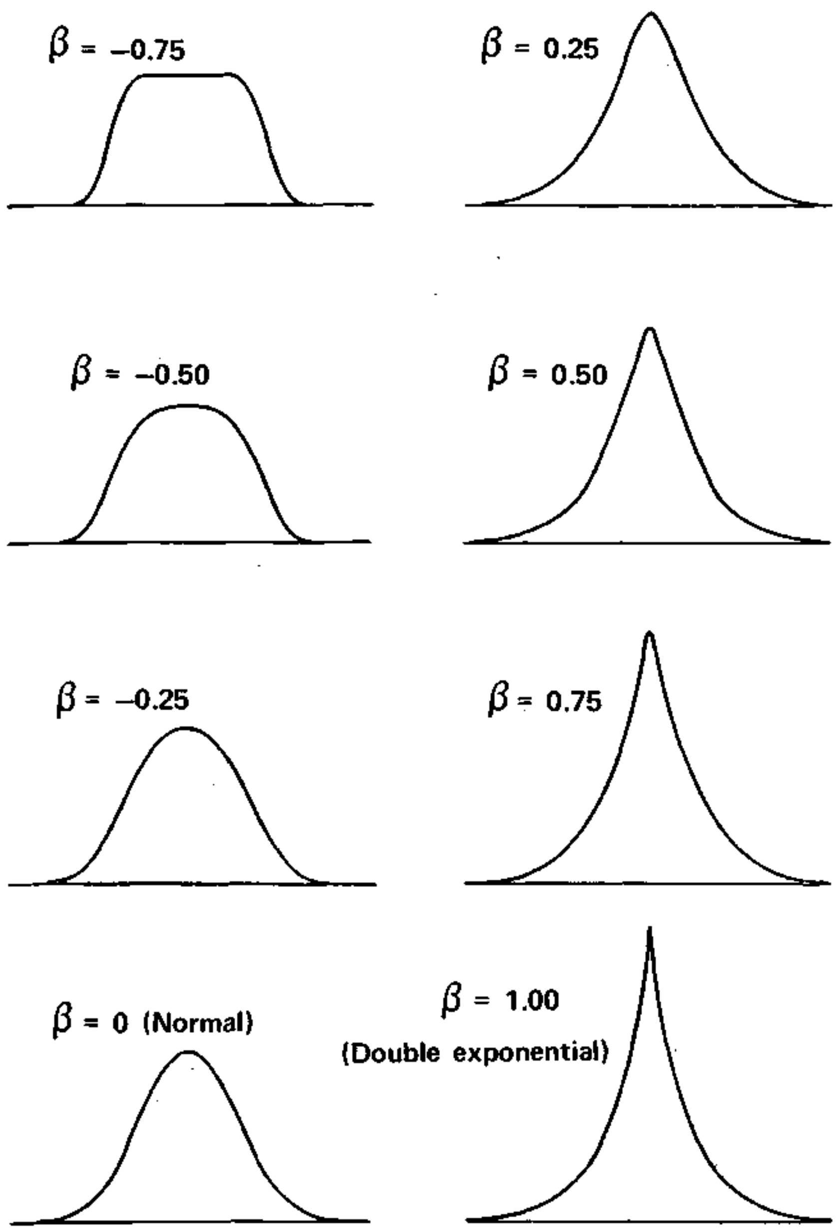


Figure 8. Platykurtic ($\beta < 0$) and leptokurtic ($\beta > 0$) distributions

β tends to -1. The kurtosis, kt, and β are related by the following expression

$$kt = \frac{\Gamma[5/2\ (1+\beta)]\ \Gamma[1/2\ (1+\beta)]}{\{\Gamma[3/2\ (1+\beta)]\}^2} \tag{12}$$

Values of kt corresponding to various 3 values, as obtained from equation 12 are:

β	-1.000*	-0.950	-0.900	-0.850	-0.800	-0.750	-0.700	-0.650
kt	1.800	1.807	1.824	1.851	1.884	1.923	1.968	2.017
β	-0.600	-0.550	-0.500	-0.450	-0.400	-0.350	-0.300	-0.250
kt	2.070	2.127	2.188	2.253	2.322	2.394	2.469	2.548
β	-0.200	-0.150	-0.100	-0.050	0.000	0.050	0.100	0.150
kt	2.631	2.718	2.808	2.902	3.000	3.102	3.208	3.319
β	0.200	0.250	0.300	0.350	0.400	0.450	0.500	0.550
kt	3.433	3.553	3.677	3.805	3.939	4.078	4.222	4.372
β	0.600	0.650	0.700	0.750	0.800	0.850	0.900	0.950
kt	4.527	4.688	4.856	5.029	5.209	5.396	5.590	5.791
β	1.000							
kt	6.000	*is the limiting case						

The kurtosis effect correction can be made in the Q_T values by modifying the z_T values. These z_T values were computed by numerical integration of equation 8 with θ=0 and ø=1. These are given in Table 9 for 41 values of β lying in the range -1* to +1 and 6 values of T: 10, 25, 50, 100, 500, and 1000 years (or corresponding p values of 0.90, 0.96, 0.98, 0.99, 0.998, and 0.999). The various recurrence-interval floods can, thus, be computed with and without correction for kurtosis. For Q_T without correction for kurtosis, the z_T values are taken for kt=3.0. In the case of correction for kurtosis, the β value is interpolated for the sample kt, and the corresponding z_T are taken from Table 9 and used in equation 4.

Table 9. Values of z_T for Various Values of β and T

g	Values of z_T for Recurrence Interval, T, of					
	10	25	50	100	500	1000
-1.00	1.386	1.593	1.663	1.697	1.725	1.729
-0.95	1.384	1.592	1.665	1.708	1.762	1.777
-0.90	1.378	1.594	1.679	1.736	1.817	1.841
-0.85	1.372	1.600	1.699	1.769	1.875	1.908
-0.80	1.366	1.608	1.721	1.803	1.935	1.977
-0.75	1.360	1.618	1.744	1.839	1.996	2.047
-0.70	1.355	1.629	1.768	1.875	2.056	2.117
-0.65	1.350	1.640	1.792	1.911	2.117	2.187
-0.60	1.345	1.651	1.815	1.946	2.178	2.257
-0.55	1.340	1.661	1.838	1.982	2.238	2.328
-0.50	1.335	1.672	1.861	2.016	2.298	2.398
-0.45	1.330	1.682	1.883	2.050	2.358	2.468
-0.40	1.326	1.691	1.904	2.083	2.418	2.538
-0.35	1.321	1.700	1.925	2.116	2.477	2.608
-0.30	1.315	1.709	1.945	2.148	2.535	2.677
-0.25	1.310	1.717	1.965	2.179	2.594	2.747
-0.20	1.305	1.725	1.984	2.210	2.651	2.816
-0.15	1.299	1.732	2.002	2.240	2.709	2.885
-0.10	1.293	1.739	2.020	2.269	2.766	2.954
-0.05	1.288	1.745	2.037	2.298	2.822	3.022
0.00	1.282	1.751	2.054	2.326	2.878	3.090
0.05	1.275	1.756	2.070	2.354	2.934	3.158
0.10	1.269	1.761	2.085	2.381	2.989	3.226
0.15	1.263	1.765	2.100	2.407	3.044	3.293
0.20	1.256	1.770	2.114	2.433	3.098	3.361
0.25	1.249	1.773	2.128	2.458	3.152	3.428
0.30	1.243	1.776	2.141	2.482	3.205	3.494
0.35	1.236	1.779	3.154	2.506	3.258	3.561
0.40	1.229	1.782	2.166	2.529	3.311	3.627
0.45	1.222	1.784	2.178	2.552	3.363	3.692
0.50	1.214	1.786	2.189	2.574	3.414	3.758
0.55	1.207	1.787	2.200	2.596	3.465	3.823
0.60	1.200	1.788	2.210	2.617	3.516	3.888
0.65	1.192	1.789	2.220	2.637	3.566	3.952
0.70	1.185	1.789	2.229	2.657	3.616	4.016
0.75	1.177	1.790	2.238	2.677	3.665	4.080
0.80	1.169	1.789	2.247	2.695	3.714	4.143
0.85	1.162	1.789	2.255	2.714	3.762	4.206
0.90	1.154	1.788	2.262	2.732	3.810	4.269
0.95	1.146	1.787	2.269	2.749	3.857	4.331
1.00	1.138	1.786	2.276	3.766	3.904	4.393

Log-Pearson Type III Distribution Method

The Water Resources Council (1976, 1977) has recommended the following technique for fitting the log-Pearson type III, LP3, distribution to an observed annual flood series, Q_i, and for computing floods at desired recurrence intervals.

1. Compute mean $\overline{x}$

$$\overline{x} = \sum_{i=1}^{n} x_i/n \tag{13}$$

in which $x = \log_{10}Q$ and n = number of years or sample size.

2. Compute standard deviation, s

$$s = [\sum_{i=1}^{n} (x_i-\overline{x})^2/(n-1)]^{0.5} \tag{14}$$

3. Compute skew coefficient, g

$$g = n \sum_{i=1}^{n} (x_i-\overline{x})^3/[(n-1)(n-2)\ s^3] \tag{15}$$

4. Compute flood of recurrence interval T years, Q_T

$$\log Q_T = \overline{x} + ks \tag{16}$$

in which k is a factor that is a function of g and the selected recurrence interval (or exceedance probability). Values of k can be obtained from a table.

Because of the errors inherent in estimating the third moment from a small sample, a regional analysis is recommended for deriving a suitable value of regional skew coefficient, g_r. The weighted skew, g_w,

$$g_w = g\ w + (1 - w)\ g_r \tag{17}$$

is used with sample $\overline{x}$ and s; the weight w equals (n-25)/75. When n equals or exceeds 100, w equals unity.

Mixed Distribution Method

This is based on the mixed distribution concept and considers the observed floods (or their logarithms) to belong to two populations with means μ_1 and μ_2, variances σ_1^2 and σ_2^2 and relative weights a and $1-a$ of the two component distributions which may be both lognormal, normal, or any other distribution type, or a mixture of two types. The mixed distribution method developed from various studies (Singh, 1968; Singh and Sinclair, 1972; and Singh, 1974) is based on the following equations:

$$p\{x\} = a p_1\{x\} + (1-a)\; p_2\{x\} \qquad 0 \le a \le 1 \tag{18}$$

$$p_1\{x\} = \frac{1}{\sigma_1 \sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{(x' - \mu_1)^2}{2\sigma_1^2}\right] dx' \tag{19}$$

$$p_2\{x\} = \frac{1}{\sigma_2 \sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{(x' - \mu_2)^2}{2\sigma_2^2}\right] dx' \tag{20}$$

in which p is the probability of being equal to or less than x. The component distributions are taken as log-normal (x = log Q) in the above equations. The mixed distribution parameters are linked to sample μ, σ^2, and g values according to the following equations (Cohen, 1967):

$$\mu = a\mu_1 + (1-a)\; \mu_2 \tag{21}$$

$$\sigma^2 = a\sigma_1^2 + (1-a)\; \sigma_2^2 + a(1-a)(\mu_2 - \mu_1)^2 \tag{22}$$

$$g = [3a(1-a)(\mu_1 - \mu_2)(\sigma_1^2 - \sigma_2^2) + a(1-a)(1-2a)(\mu_1 - \mu_2)^3]/\sigma^3 \tag{23}$$

Evaluation of Parameters

The distribution parameters for the mixed distribution were obtained by using the Generalized Reduced Gradient Method, a nonlinear programming algorithm; the computer program for which was available from the University of Illinois. Two nonlinear objective functions, i.e., minimization of $\Sigma(\Delta z)^2$ and $\Sigma|\Delta z|$, Δz equals the difference between the standard deviate corresponding to the observed probability equal to (m - 0.38)/(n + 0.24) and that fitted corresponding to p from equation 18, were considered subject to the following constraints:

$$1 = a_1 + a_2 \tag{24}$$

$$\mu = a_1\mu_1 + a_2\mu_2 \tag{25}$$

$$\sigma^2 = a_1\sigma_1^2 + a_2\sigma_2^2 + a_1a_2 (\mu_2 - \mu_1)^2 \tag{26}$$

$$g\sigma^3 = a_1m_1(3\sigma_1^2 + m_1^2) + a_2m_2(3\sigma_2^2 + m_2^2) \tag{27}$$

$$kt\sigma^4 = a_1(3\sigma_1^4 + 6m_1^2\sigma_1^2 + m_1^4) + a_2(3\sigma_2^4 + 6m_2^2\sigma_2^2 + m_2^4) \tag{28}$$

in which $m_1 = \mu_1-\mu$, $m_2 = \mu_2-\mu$, $a_1 = \alpha$, and $a_2 = 1-\alpha$. Use of $\Sigma|\Delta z|$ was found to give more consistent solutions than $\Sigma(\Delta z)^2$.

The asymmetry of the power-transformed or log-transformed flood series, as evidenced by the kt being lower or higher than 3 and by the 5th moment being significantly different from zero, is accommodated easily by the mixed distribution concept.

Various Recurrence-Interval Floods

These floods are calculated by a reiterative process. For the desired recurrence interval, value of p is obtained from (1 - 1/T). Starting from a given or assumed value of x, z_1 and z_2 are calculated from:

$$z_1 = \frac{x - \mu_1}{\sigma_1} \quad \text{and} \quad z_2 = \frac{x - \mu_2}{\sigma_2} \tag{29}$$

The corresponding p_1 and p_2 are obtained from the z-to-p subroutine. The p is calculated from p_1 and p_2 with equation 18. If this p is equal to, or within a specified tolerance of, the p corresponding to the desired recurrence interval, the value of x yields the logarithm of Q_T. Otherwise, by an iterative process, a value of x is determined that meets the p criterion.

NEW FLOOD-FREQUENCY METHODOLOGY

A new flood frequency methodology has been developed and computerized. It detects objectively the outliers and inliers at various probability levels and modifies them if needed. The computer program prints 2- to 1000-yr floods from power transformation, both with and without kurtosis correction, from log-Pearson type III distribution, both with sample skew and weighted skew, and from the mixed distribution, for levels 0, 1, 2, 3, 4, 5, and 6. The level 0 corresponds to processing of data without any testing for outliers and/or inliers. Levels 1 through 6 correspond to outlier-inlier probability pairs of .01, .99; .05, .95; .10, .90; .20, .80; .30, .70; and .40, .60. The relevant information on statistics of the three methods and the given and modified values of outliers/inliers are also printed at all the levels. The salient features of this new methodology are described in the rest of this section.

Compact Departure Table, Probability Levels and Windows

A compact departure table, containing the test value, at 6 probability levels is given in Table 10. The low 1 through 5 denote the lowest to the 5th low value from the low end and the high 1 through 5 denote the highest to the 5th high value from the high end. Considering level 1 and low 1, a departure Δ of -0.689 or less (in the algebraic sense) will indicate an inlier at 1% or less nonexceedance probability (or 99% or higher exceedance probability), and a departure of 1.029 or more will indicate an outlier at 99% or higher nonexceedance probability (or 1% or lower exceedance probability) . Similarly, for level 1 and high 1, a departure of 0.679 or more

TABLE 10. Test Values of Outlier and Inlier Departures

Window	p	Outlier/ Inlier	Test values of departures Low 1 15-100	Low 2 20-100	Low 3 25-100	Low 4 30-100	Low 5 40-100
1	<.01	Inlier	<-.689	<-.495	<-.412	<-.363	<-.327
	>.99	Outlier	>1.029	>0.643	>0.498	>0.418	>0.368
2	<.05	Inlier	<-.532	<-.369	<-.303	<-.264	<-.237
	>.95	Outlier	>0.681	>0.421	>0.337	>0.285	>0.253
3	<.10	Inlier	<-.441	<-.299	<-.243	<-.211	<-.188
	>.90	Outlier	>0.503	>0.321	>0.254	>0.217	>0.193
4	<.20	Inlier	<-.318	<-.209	<-.167	<-.143	<-.127
	>.80	Outlier	>0.297	>0.197	>0.159	>0.137	>0.123
5	<.30	Inlier	<-.221	<-.141	<-.110	<-.093	<-.082
	>.70	Outlier	>0.161	>0.112	>0.092	>0.081	>0.073
6	<.40	Inlier	<-.132	<-.080	<-.060	<-.050	<-.043
	>.60	Outlier	>0.052	>0.043	>0.037	>0.034	>0.032
			High 1 15-100	High 2 20-100	High 3 25-100	High 4 30-100	High 5 40-100
1	<.01	Outlier	<-1.054	<-.654	<-.511	<-.429	<-.377
	>.99	Inlier	>0.679	>0.488	>0.407	>0.358	>0.323
2	<.05	Outlier	<-.683	<-.433	<-.341	<-.290	<-.256
	>.95	Inlier	>0.529	>0.369	>0.300	>0.263	>0.235
3	<.10	Outlier	<-.500	<-.322	<-.258	<-.221	<-.195
	>.90	Inlier	>0.438	>0.299	>0.241	>0.209	>0.186
4	<.20	Outlier	<-.295	<-.197	<-.161	<-.139	<-.124
	>.80	Inlier	>0.317	>0.209	>0.166	>0.143	>0.126
5	<.30	Outlier	<-.159	<-.112	<-.094	<-.082	<-.074
	>.70	Inlier	>0.221	>0.140	>0.110	>0.093	>0.082
6	<.40	Outlier	<-.051	<-.043	<-.039	<-.035	<-.032
	>.60	Inlier	>0.132	>0.079	>0.060	>0.050	>0.043

Note: 15 - 100,....., and 40 - 100 denote the range of sample size n in years.

indicates an inlier at 99% or higher nonexceedance probability (or 1% or lower exceedance probability), and a departure of -1.054 or lower (in the algebraic sense) indicates an outlier at 1% or lower nonexceedance probability (or 99% or higher exceedance probability). Thus, the probability pairs for outliers and inliers have the connotation of the same relative severity.

The concept of the levels and windows is clarified in figure 9 in which departures for the high 1 are plotted on normal probability paper. For the outliers, window 1 contains Δ values ≤ -1.054, window 2 contains Δ values such that $-1.054 < \Delta \leq -0.683$, and so on for windows 3 through 6. For the inliers, window 1 contains Δ values ≥ 0.679, window 2 contains Δ values such that $0.679 > \Delta \geq 0.529$, and so on for windows 3 through 6. The departures for the low 1 are plotted on normal probability paper in figure 10. If some outliers and/or inliers are found in window 1, the same are modified to respective values at level 1, and the procedure is followed sequentially from one window to the other. If no outliers and/or inliers are detected in a particular window, no modifications are done, and the program moves to the next window after developing and printing distribution statistics and flood estimates.

Plotting Position

The plotting position for the observed floods has been a matter of considerable controversy. A general formula for computing plotting position (Harter, 1971) is:

$$p = (m - a)/(n - a - b + 1) \qquad (1)$$

in which m is the rank order of flood values arranged in an ascending order

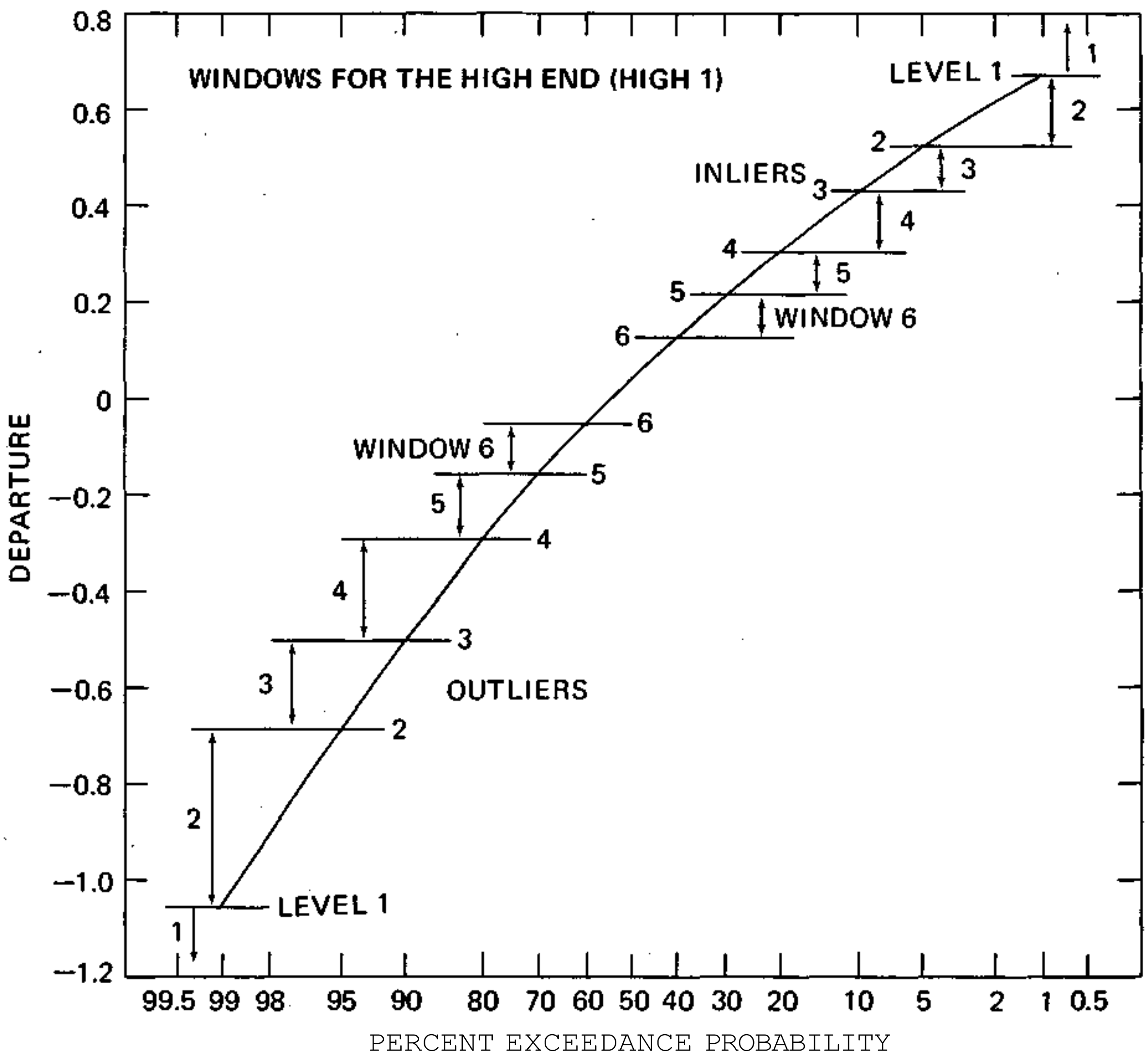


Figure 9. Levels and windows for the outliers/inliers at the high end

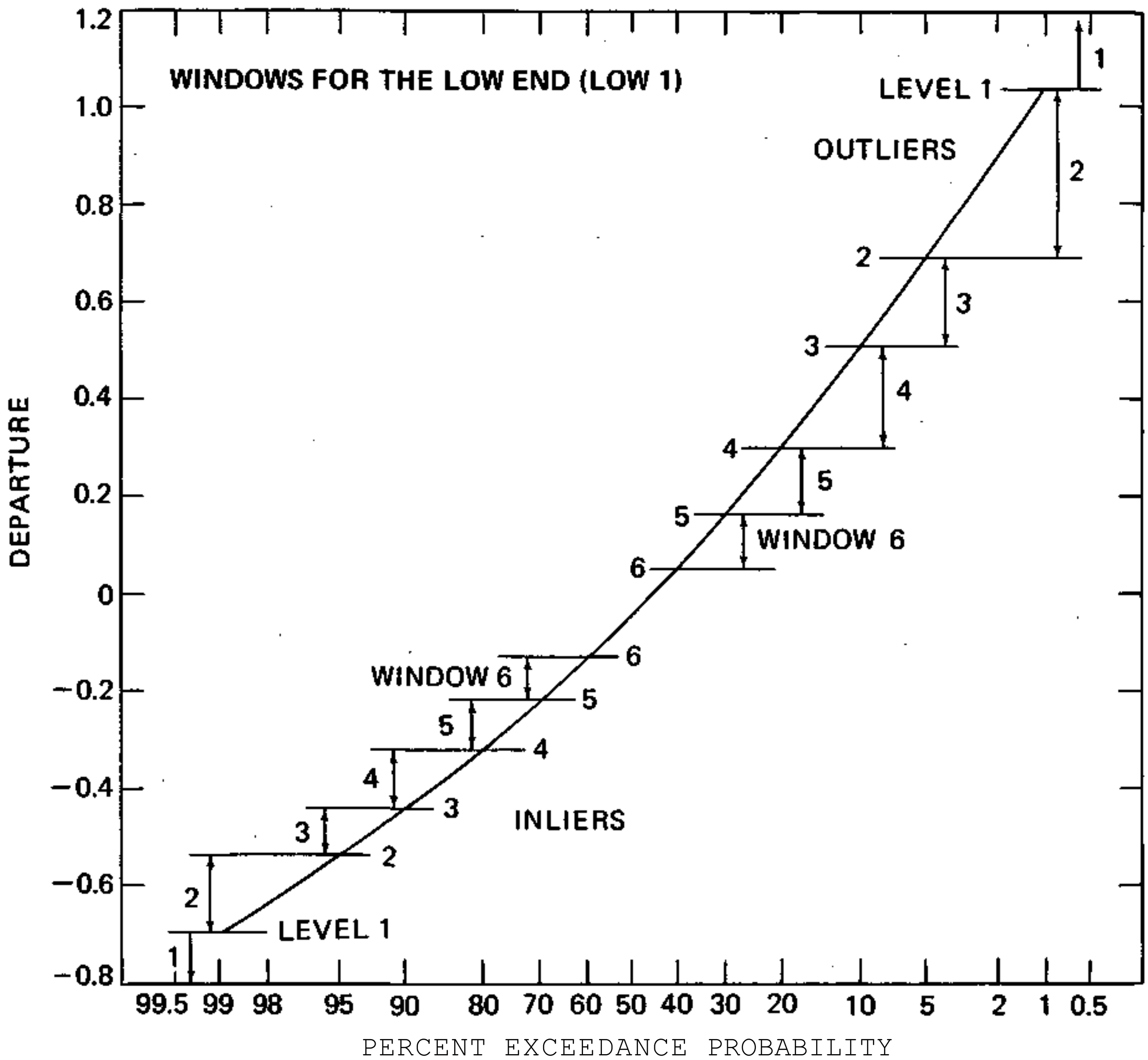


Figure 10. Levels and windows for the outliers/inliers at the low end

of magnitude, p is the probability of nonexceedance, and a and b depend on the distribution. For a symmetrical distribution, a equals b, and equation 1 can be rewritten as:

$$p = (m - \alpha)/(n + 1 - 2\alpha) \tag{2}$$

The commonly used Weibull plotting position

$$p = m/(n+1) \tag{3}$$

is obtained by putting $\alpha = 0$. However, an α value of about 0.38 (Cunnane, 1978; Blom, 1958) is the best for the normal distribution. Cunnane states that the Weibull plotting formula is exact when the distribution is uniform and that the Gringorten formula, with $\alpha = 0.44$, is satisfactory for exponential distributions.

In calculating $\sum|\Delta z|$ for evaluating the mixed distribution parameters by the nonlinear programming algorithm, the Δz is obtained from

Δz =(z corresponding to standard normal deviate for observed probability p)-(z fitted from the mixed distribution with

$$p = a p_1 + (1 - a)\ p_2) \tag{4}$$

The observed probability p is obtained from

$$p = \frac{m - 0.38}{n + 0.24} \tag{5}$$

in which $a = b = \alpha = 0.38$.

The Flow Chart

The detection and modification of outliers and/or inliers as well as flood frequency analysis follow the flow chart given in figure 11. Some relevant explanations to clarify the methodology and the computer program are given in the following few pages. The sequence numbers correspond to the numbers attached to various boxes in the flow chart.

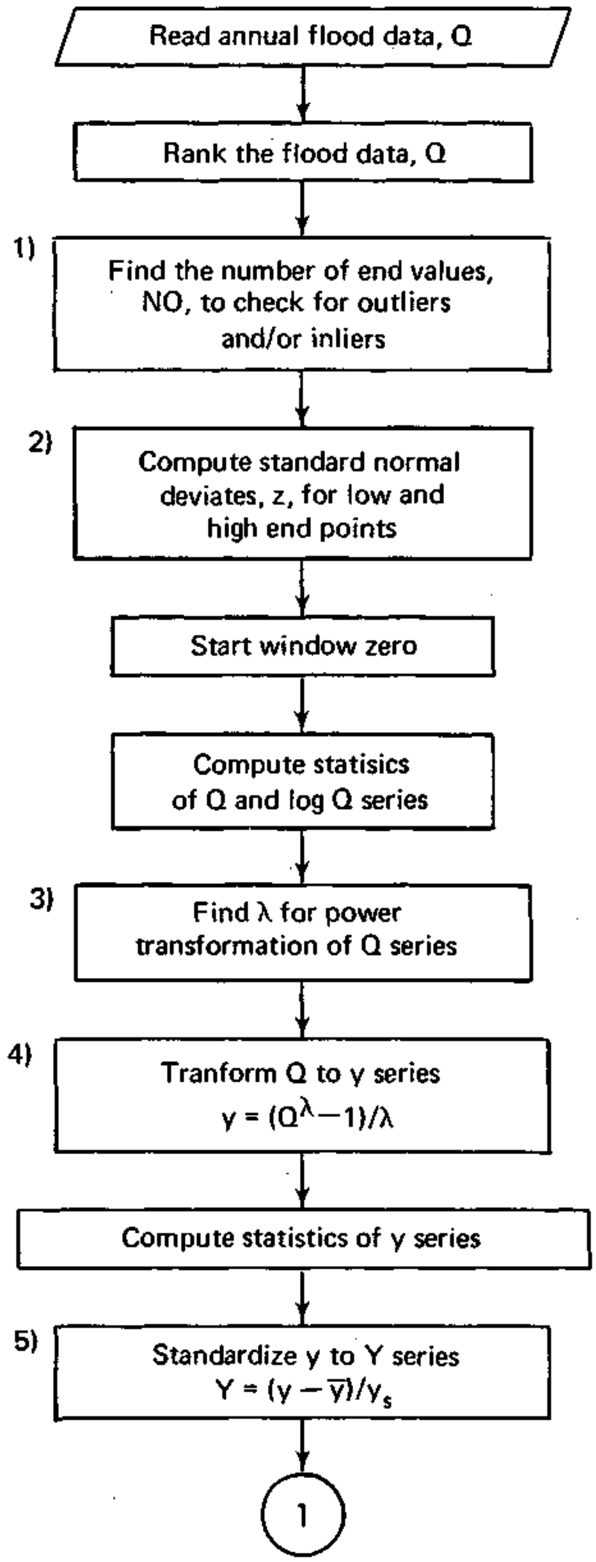


Figure 11. Flow chart for the computer program for flood frequency analyses

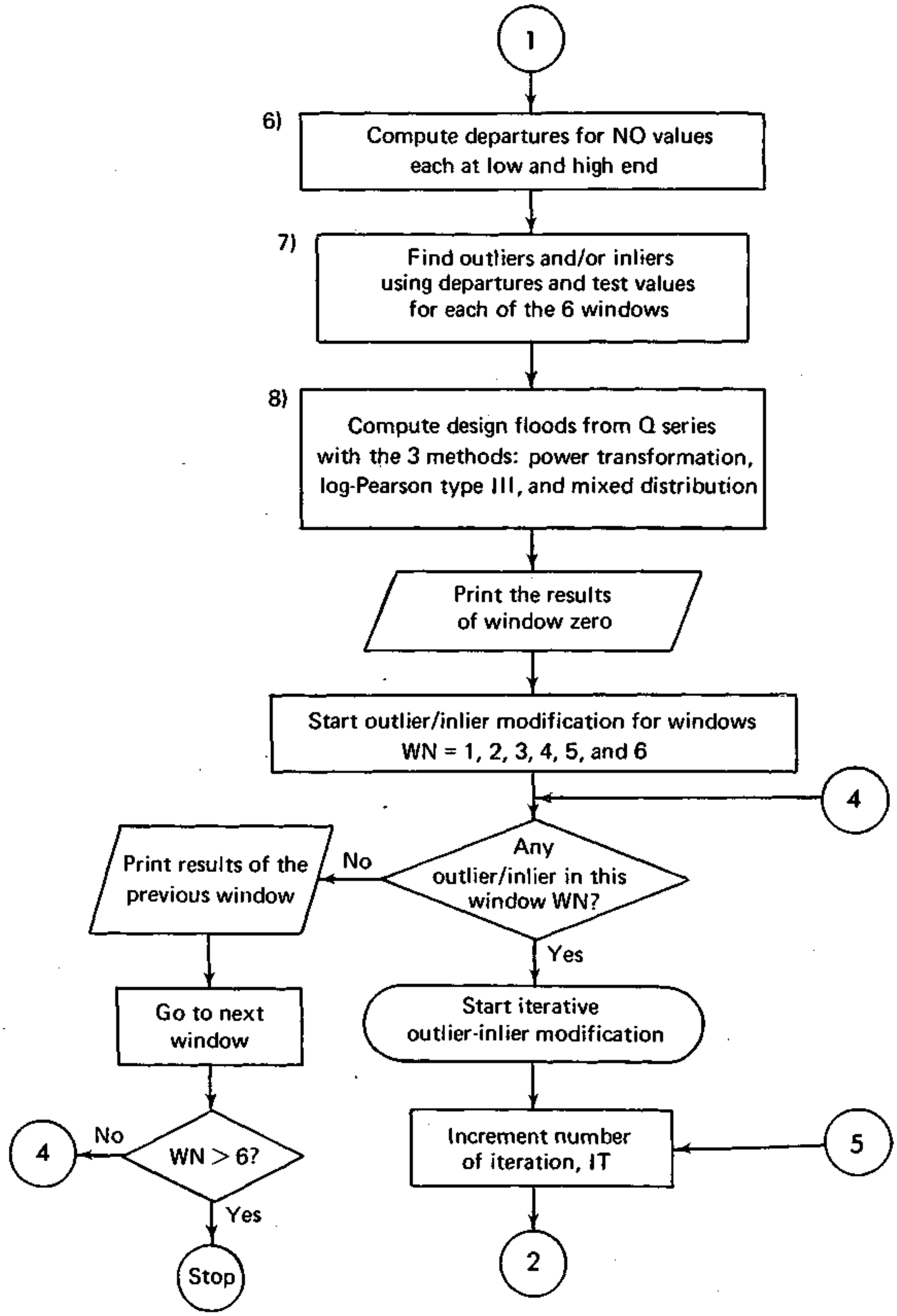


Figure 11. —Continued

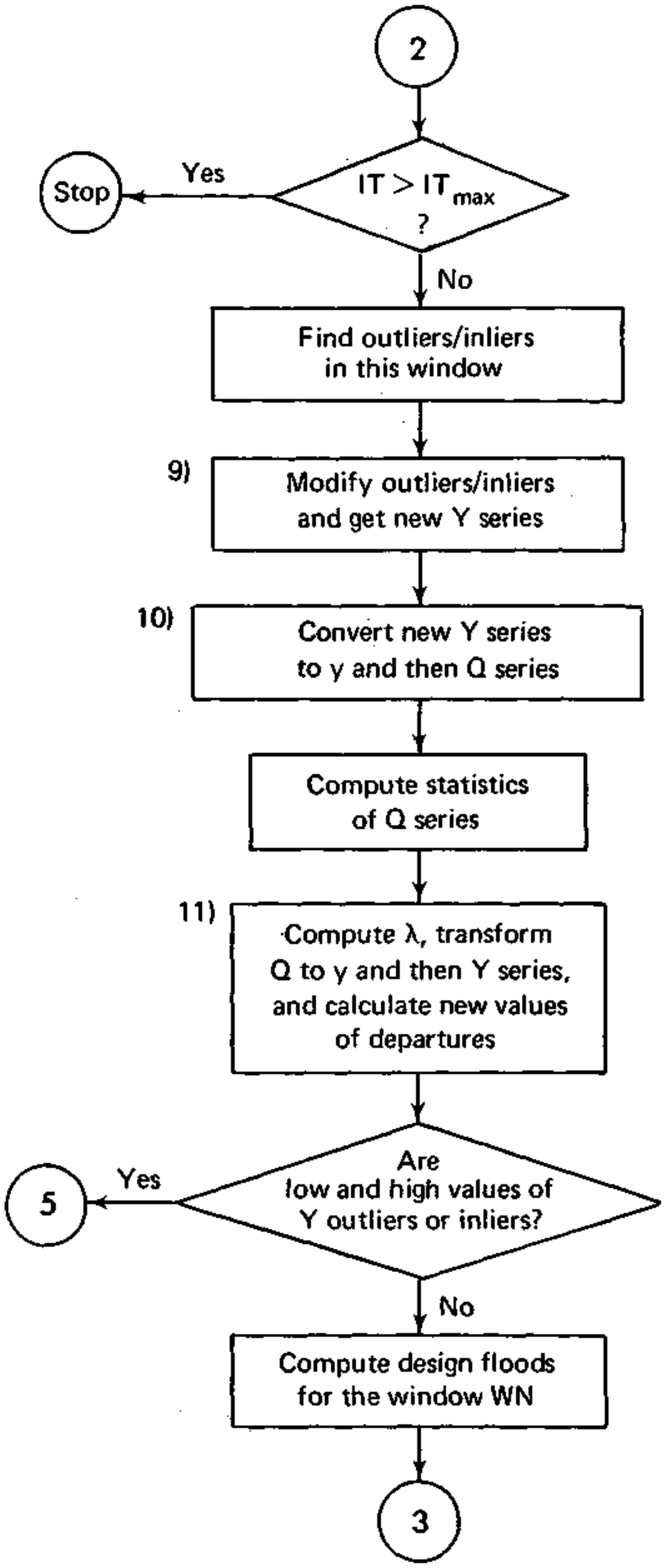


Figure 11. —Continued

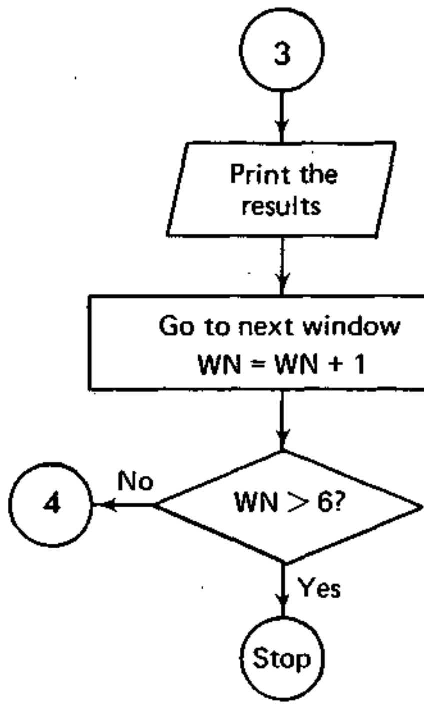


Figure 11. —Concluded

1.) Number of low as well as high floods, NO, can be provided as an input information or computed from some expression such as NO = [n/10] where n equals the number of floods in the sample series and NO = 5 for $n \geq 50$.

2.) Standard normal deviate of rank m, z_m , is computed by converting probability p , obtained from

$$p_m = (m - \alpha)/(n + 1 - 2\alpha) \tag{2}$$

with a interpolated from the smoothed a values for the rank m and sample size n (see step 8 of Generation of Departures, p. 22 to 23) to z with the p-to-z subroutine, assuming standard normal distribution.

3.) The parameter λ is computed by the maximum log-likelihood method from the given flood series.

4.) The given flood series is transformed to y series by

$$y_i = \frac{Q_i^{\lambda} - 1}{\lambda} \quad , i = 1, 2, \ldots, n \tag{6}$$

and the process is termed normalization by power transformation.

5.) The y series is standardized to Y series with

$$Y_i = \frac{y_i - \bar{y}}{y_s} \tag{7}$$

in which $\bar{y}$ and y_s are the mean and standard deviation of the transformed series, y.

6.) The departures, Δ_m for NO values at the low end, as well as at the high end, are obtained from

$$\Delta_m = z_m - Y_m \quad \text{for } m = 1, \ldots, NO \tag{8}$$

7.) Outliers and inliers, if any, are detected in each of the 6 windows according to 6 probability levels using Δ_m and test values in Table 10.

8.) The floods corresponding to 2-, 10-, 25-, 50-, 100-, 500-, and 1000-year recurrence intervals are computed with three methods: 1) power transformation with and without kurtosis correction; 2) log-Pearson type III distribution, with sample as well as with weighted skew; and 3) mixed distribution; without any modification of outliers and/or inliers, i.e. with window 0.

9.) Modification process is illustrated in figure 12 as an example. Consider 2 low and 2 high values (N0=2) as candidates for outliers. No outliers are detected in window 1; 1 high and 1 low outliers are detected in window 2; and 2 high and 1 low outliers are detected in windows 3 through 6. The values of Y_m for the detected outliers and/or inliers in a window are changed to (z_m minus departure) values for that window. This gives the new Y series in which the Y outlier/inlier values have been replaced by the corresponding threshold values.

10.) The new Y series is destandardized with $\bar{y}$ and y_s from equation 7, to get new y series.

$$y = \bar{y} + y_s Y \tag{9}$$

The new y series is detransformed with A from step 3 for the previous Q series, to obtain the new Q series (with values of Q changed only for the outliers/inliers).

11.) The new value of X is computed for the new or modified Q series. With this X, the modified Q series is transformed to y series, which in turn is standardized to Y series, and values of departures are obtained for the NO points at both low and high end.

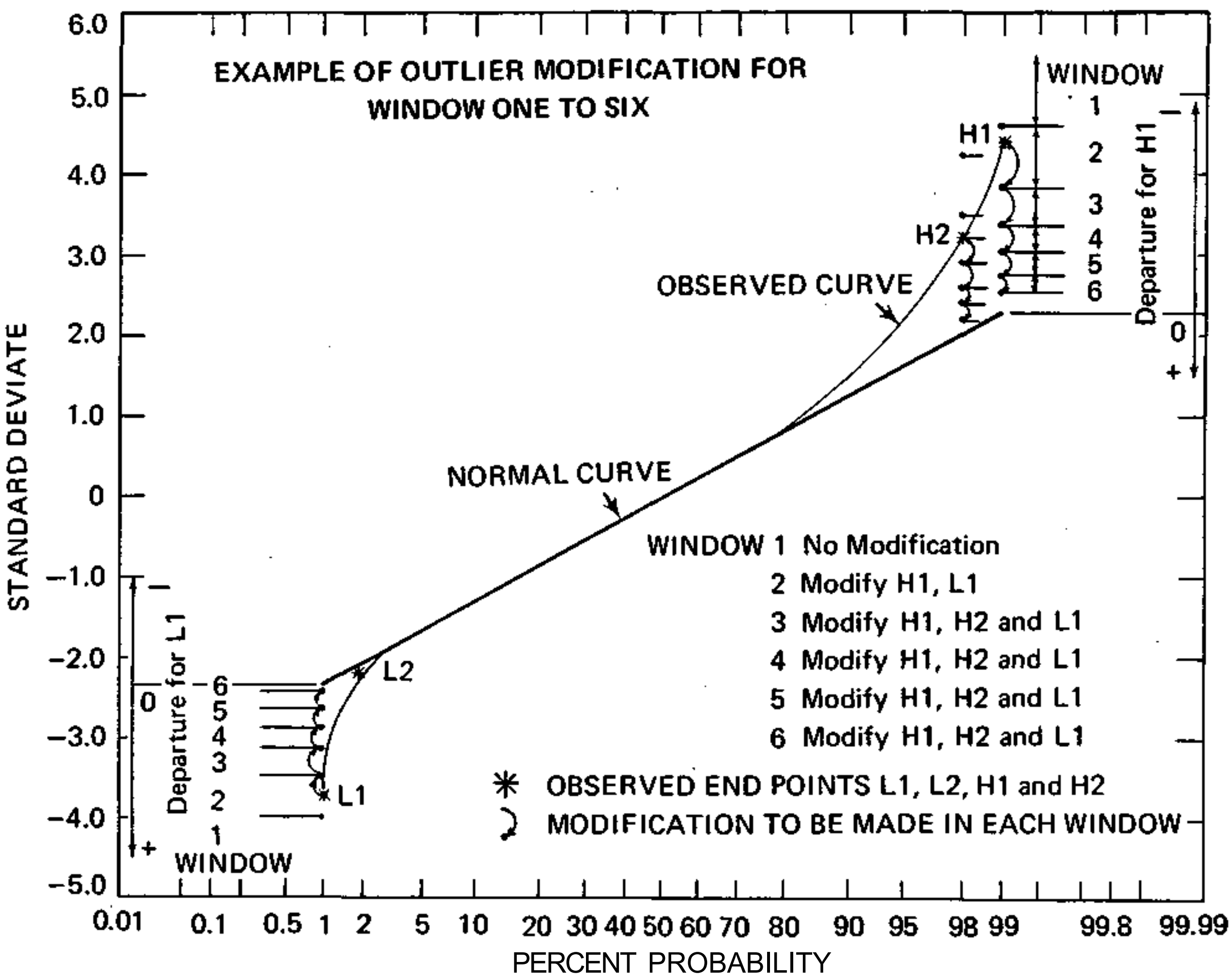


Figure 12. An example of outlier/inlier modification

An Example

The new methodology of detection and modification of any outliers/inliers and flood frequency analysis is explained by an actual example of observed floods for the Sangamon River at Oakford (USGS No. 05 583000, drainage area 5093 sq miles, n = 62 years).

Ranked discharge, Q, data in cfs

3,480	3,800	4,630	5,670	5,960	6,430	6,720	8,400
9,100	10,000	10,500	10,500	11,000	11,300	12,000	13,300
15,100	15,700	16,500	16,500	17,200	19,000	19,400	20,000
20,300	20,800	20,800	21,200	21,400	21,500	22,200	23,700
24,100	24,700	25,000	25,200	25,600	26,200	28,400	29,100
30,100	30,200	31,200	31,400	33,300	33,800	34,600	34,700
36,300	36,300	37,600	37,900	38,000	38,300	42,300	42,800
42,900	44,700	45,800	46,300	55,900	123,000		

According to Singh (1980), the maximum flood of 123,000 cfs which occurred on May 20, 1943 was caused by a 50-year storm, covering most of the basin, over very wet antecedent soil moisture conditions, giving a runoff factor of 2.2 times that for the next 4 high floods caused by 10- to 25-year storms.

This gaging station also suffers from junction problem caused by two major tributaries. Salt Creek and Sangamon River join 9 miles upstream of Oakford. Drainage areas above the gaging stations on these tributaries are 1804 square miles (5 miles upstream of confluence) and 2618 square miles (49 miles upstream of confluence), respectively. The relevant statistics for the first 5 top floods for the concurrent record of 1942-1979 at Oakford and corresponding floods at Greenview (Salt Creek) and Riverton (Sangamon River) are given below.

Flood at Oakford gage			Salt Creek near Greenview			Sangamon at Riverton		
Rank	Date	Peak	Rank	Date	Peak	Rank	Date	Peak
1	5-20-43	123,000	1	5-19-43	41,200	1	5-19-43	68,700
2	4-15-79	55,900	3	4-13-79	30,500	2	4-12-79	44,200
3	4-25-73	45,800	6	4-25-73	21,000	6	4-24-73	27,000
4	4-26-44	44,700	5	4-25-44	22,000	3	4-25-44	30,600
5	6-25-74	42,900	2	6-24-74	38,100		6-24-74	

For the flood peaks from Greenview and Riverton to coincide at Oakford, the peak at Riverton should occur a day before that at Greenview and the peak at Greenview should occur about 6 hours before that at Oakford. Concurrent maximum floods at Greenview and Riverton produce the maximum flood at Oakford. An analysis of all the floods at the 3 stations indicates that for floods exceeding a 2-year flood, there is only one chance out of 5 that the tributary floods will be in phase to produce a high flood at Oakford.

Statistics of Q data

Series	Mean	s	g	kt	5th moment
Q	25480	17821	2.739	16.573	93.480
log Q	4.311	0.307	-0.562	3.421	-3.027

It is evident that log transformation makes the series much closer to normal.

Parameter for power transformation (determined by the maximum log-likelihood method), $\lambda = 0.254$

Power transformed series, $y = (Q^{\lambda} - 1)/\lambda$

Series	Mean	s	g	kt	5th moment
y	45.816	8.580	0.018	3.820	3.900

y values (L1 to L5): 27.304, 28.010, 29.654, 31.428, 31.879

y values (H5 to H1): 55.814, 56.184, 56.350, 59.305, 73.331

Standardized series, $Y = (y - \overline{y})/y_s$

Y values (L1 to L5): -2.158, -2.075, -1.884, -1.677, -1.624

Y values (H5 to H1): 1.165, 1.208, 1.228, 1.572, 3.207

Theoretical standard normal deviates, corresponding to 5 low and 5 high values of Y.

z values (Ll to L5): -2.342, -1.958, -1.739, -1.580, -1.454

z values (H5 to H1): 1.454, 1.580, 1.739, 1.958, 2.342

A comparison of Y and z values indicates that Y(Ll) is an inlier, Y(L2 to L5) are outliers, Y(H5 to H2) are inliers, and Y(Hl) is an outlier.

Outlier/Inlier detection: Outliers and inliers are denoted by 0 and I.

Window	L1	L2	L3	L4	L5	H5	H4	H3	H2	H1
1							I	I		
2						I	I	I	I	0
3						I	I	I	I	0
4					0	I	I	I	I	0
5		0	0	0	0	I	I	I	I	0
6	I	0	0	0	0	I	I	I	I	0

Design floods with original discharge data

Method	Floods in cfs for recurrence interval (years)						
	2	10	25	50	100	500	1000
PT, kt=3.0	21,738	47,712	61,422	71,717	82,029	106,247	116,843
PT, sample kt	21,738	46,495	62,345	75,439	89,513	126,090	143,606
LP3, sample g	21,857	48,001	60,903	70,090	78,851	97,730	105,279
LP3, weighted g	21,649	48,439	62,293	72,434	82,331	104,456	113,630
Mixed distrib.	23,411	43,791	57,244	71,571	89,279	140,073	166,295

Fifty- to 1000-year floods with correction for sample kurtosis (3.820, which is greater than 3.00) become progressively higher than those with no kurtosis correction as the recurrence interval increases.

Window 1

Inliers detected: 3rd and 4th high points

After 2 iterations, the modified values of standardized series, Y, are

Y values (L1 to L5): -2.157, -2.074, -1.883, -1.676, -1.623

Y values (H5 to H1): 1.160, 1.226*, 1.336*, 1.565, 3.191

The H3 or Y_{60} is modified as explained below.

Y_{60} is increased from 1.228 so that it just gets into the next (or the second) window by making it equal to 1.739 (theoretical standard normal deviate) - 0.407 (departure for H3, window 1) × (1-0.01); or 1.336. Factor 0.01 reduces the number of iterations and just caries H3 or Y_{60} into the next lower window.

Modified data after destandardization and detransformation:

Q (L1 to L5): 3,480, 3,800, 4,630, 5,670, 5,960

Q (H5 to H1): 44,700, 46,403 49,331, 55,900, 123,000

⇅ ⇅

previous values (44,700)(45,500)

New X value = 0.252

Power transformed data:

y (L1 to L5): 27.011, 27.706, 29.323, 31.067, 31.510

y (H5 to H1): 54.981, 55.539, 56.464, 58.398, 72.109

Statistics of y series:

Mean	s	g	kt	5th moment
45.201	8.435	0.017	3.788	3.763

Standardized series, Y

Y (L1 to L5): -2.156, -2.074, -1.882, -1.676, -1.623

Y (H5 to H1): 1.159, 1.226, 1.335, 1.564, 3.190

New departures are:

Δ (L1 to L5): -0.185, 0.116, 0.144, 0.095, 0.169

Δ (H5 to H1): 0.294, 0.355, 0.403, 0.394, -0.848

A check with test departures shows no inliers/outliers in window 1.

Design floods with modified Q series (window 1)

	Floods in cfs for recurrence interval (years)						
Method	2	10	25	50	100	500	1000
PT, kt=3.0	21,761	47,882	61,700	72,086	82,498	106,973	117,692
PT, sample kt	21,761	46,701	62,607	75,729	89,814	126,341	143,806
LP3, sample g	21,872	48,179	61,219	70,526	79,420	98,644	106,354
LP3, weighted g	21,671	48,602	62,567	72,802	82,803	105,198	114,498
Mixed distrib.	23,391	44,008	57,588	71,987	89,772	140,791	167,152

Windows 2 through 6

Outliers are modified similarly for the successive windows 2 through 6. The results of this analyses are presented in Table 11 (contains

TABLE 11. Sample Computer Output with the New Methodology

STATION NO. 5583000 SANGAMON RIVER AT 0AKF0RD
DRAINAGE AREA 5093.0 Sq Mi Years of Record 62 (1910-1979)

LEVEL NO.		0	1	2	3	4	5	6
METHOD		100-Year Flood in cfs						
Power Transform, PT								
With kt = 3.0		82,029	82,498	78,988	75,264	72,347	70,357	69,059
With sample kt		89,513	89,814	83,010	76,541	71,464	68,673	67,080
Log Transform								
LP3, Sample skew		78,851	79,420	77,082	74,442	72,444	70,037	67,484
LP3, Weighted skew		82,331	82,803	81,409	79,623	78,185	77,061	76,063
Mixed Distrib., MD		89,279	89,772	82,699	75,229	69,725	69,535	70,023
Type	No.	Observed and Modified Floods in cfs						
Low	1*	3,480	3,480	3,480	3,480	3,480	2,927	2,353
	2*	3,800	3,800	3,800	3,800	3,800	3,800	3,800
	3*	4,630	4,630	4,630	4,630	4,630	4,630	4,748
	4*	5,670	5,670	5,670	5,670	5,670	5,670	5,763
	5*	5,960	5,960	5,960	5,960	6,142	6,448	6,654
High	5*	44,700	44,700	46,285	46,909	47,461	47,649	47,927
	4*	45,800	46,403	48,920	49,564	50,127	50,322	50,591
	3*	46,300	49,331	52,305	52,933	53,542	53,703	53,999
	2«	55,900	55,900	56,753	57,476	58,223	58,451	58,810
	1*	123,000	123,000	106,954	92,964	81,678	75,145	70,771
METHOD	STATISTICS	Values of Statistics						
PT	mean	45.816	45.201	63.589	93.389	131.737	180.224	230.869
	std dev	8.580	8.435	13.449	22.218	34.276	50.703	68.820
	skew	.018	.017	-.001	-.022	-.043	-.055	-.062
	kurtosis,kt	3.820	3.788	3.447	3.149	2.895	2.792	2.746
	5th moment	3.900	3.763	2.612	1.651	.856	.416	.115
	lambda	.254	.252	.300	.352	.397	.437	.468
LP3	mean	4.311	4.312	4.312	4.311	4.311	4.309	4.308
	std dev	.307	.308	.307	.305	.303	.305	.308
	sample skew	-.562	-.556	-.605	-.655	-.691	-.764	-.854
	kurtosis,kt	3.421	3.410	3.282	3.202	3.153	3.317	3.632
	5th moment	-3.027	-3.006	-3.682	-4.227	-4.606	-5.653	-7.327
MD	weight 'a'	.648	.652	.640	.605	.515	.380	.347
	mu1	4.221	4.222	4.207	4.185	4.135	4.046	4.032
	mu2	4.477	4.479	4.497	4.505	4.498	4.470	4.455
	sigma1	.338	.338	.330	.321	.306	.294	.315
	sigma2	.121	.122	.118	.123	.148	.169	.175
	Test Stat	4.718	4.568	3.925	3.634	3.396	2.906	3.000

* High & low floods considered for outlier detection and modification

100-year floods for windows 0 through 6, successive modification of low and high values and sample statistics) and in Table 12 (contains 2-, 10-, 25-, 50-, 100-, 500-, and 1000-year floods for the various methods for windows 0 through 6; 0 window corresponds to no modification of Q values). The observed and modified floods (5th window) as well as the fitted mixed distribution curve are shown in figure 13.

TABLE 12. Sample Computer Output with the New Methodology

STATION NO. 5583000 SANGAMON RIVER AT OAKFORD
DRAINAGE AREA 5093.0 Sq Mi Years of Record 62 (1910-1979)

VARIOUS RECURRENCE-INTERVAL FLOODS

METHOD	#	2	10	25	50	100	500	1000
		Flood in cfs for Recurrence Intervals (Years)						
PT, kt=3.0	0	21,738	47,712	61,422	71,717	82,029	106,247	116,843
PT, sample kt		21,738	46,495	62,345	75,439	89,513	126,090	143,606
LP3, sample skew		21,857	48,001	60,903	70,090	78,851	97,730	105,279
weighted skew		21,649	48,439	62,293	72,434	82,331	104,456	113,630
MD, mixed dist.		23,411	43,791	57,244	71,571	89,279	140,073	166,295
PT, kt=3.0	1	21,761	47,882	61,700	72,086	82,498	106,973	117,692
PT, sample kt		21,761	46,701	62,607	75,729	89,814	126,341	143,806
LP3, sample skew		21,872	48,179	61,219	70,526	79,420	98,644	106,354
weighted skew		21,671	48,602	62,567	72,802	82,803	105,198	114,498
MD, mixed dist.		23,391	44,008	57,588	71,987	89,772	140,791	167,152
PT, kt=3.0	2	21,994	47,211	60,075	69,581	78,988	100,698	110,049
PT, sample kt		21,994	46,551	60,663	71,644	83,010	110,694	123,242
LP3, sample skew		21,999	47,795	60,198	68,894	77,082	94,367	101,138
weighted skew		21,736	48,358	61,952	71,829	81,409	102,612	111,318
MD, mixed dist.		23,750	43,789	55,334	67,236	82,699	128,214	151,591
PT, kt=3.0	3	22,191	46,341	58,211	66,835	75,264	94,370	102,471
PT, sample kt		22,191	46,126	58,416	67,517	76,541	97,377	106,350
LP3, sample skew		22,084	47,217	58,918	66,972	74,442	89,841	95,733
weighted skew		21,759	47,918	61,058	70,517	79,623	99,545	107,633
MD, mixed dist.		23,879	43,775	53,865	63,160	75,229	114,125	134,446
PT, kt=3.0	4	22,358	45,612	56,702	64,652	72,347	89,552	96,759
PT, sample kt		22,358	45,762	56,548	64,164	71,464	87,571	94,242
LP3, sample skew		22,146	46,743	57,914	65,495	72,444	86,507	91,789
weighted skew		21,777	47,540	60,314	69,445	78,185	97,135	104,762
MD, mixed dist.		23,612	44,352	54,078	61,641	69,725	93,166	106,451
PT, kt=3.0	5	22,484	45,139	55,687	63,169	70,357	86,256	92,855
PT, sample kt		22,484	45,404	55,380	62,230	68,673	82,580	88,212
LP3, sample skew		22,271	46,392	56,882	63,817	70,037	82,195	86,602
weighted skew		21,807	47,424	59,895	68,707	77,061	94,894	101,959
MD, mixed dist.		23,451	44,336	54,384-61,915		69,535	88,063	96,621
PT, kt=3.0	6	22,588	44,850	55,035	62,207	69,059	84,100	90,301
PT, sample kt		22,588	45,170	54,659	61,087	67,080	79,822	84,922
LP3, sample skew		22,454	46,096	55,831	62,056	67,484	77,632	81,145
weighted skew		21,869	47,438	59,618	68,108	76,063	92,726	99,202
MD, mixed dist.		23,351	44,054	54,303	62,069	70,023	89,838	99,332

= level number

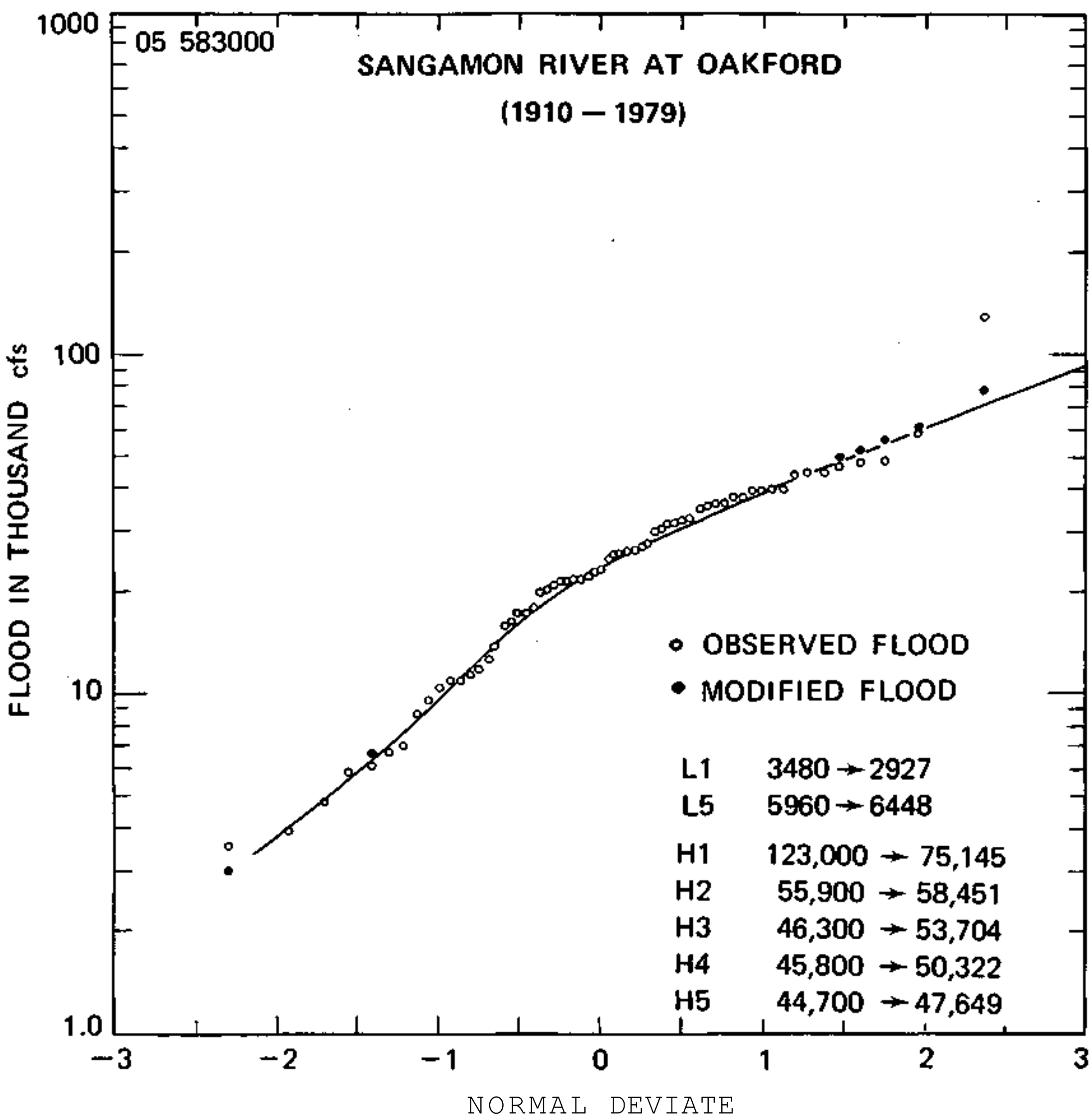


Figure 13. Observed and modified floods and the fitted mixed distribution curve (window 5) for the Sangamon River at Oakford

FLOOD FREQUENCY ANALYSES

The developed flood frequency methods were applied to 37 observed annual flood series for drainage basins with area varying from 11 to 9551 square miles and with records of about 20 to 67 years. The gaging stations and drainage basins above these stations lie in the major river basins of the Sangamon, Rock, and Little Wabash Rivers. The information on USGS gaging station number; the name of the stream and the gaging station; length of record, n, in years; the drainage area, A, in square miles; the main channel length, L, in miles; and the main channel slope, S, in ft/mi are given in Table 13 for each of the 37 basins.

With the computer program developed in this study, flood frequency analyses were carried out with the power transformation (with and without correction for kurtosis), with the log-Pearson type III, or LP3, distribution (with sample as well as with weighted skew), and with the mixed distribution, MD, method. These analyses indicated that the MD flood estimates derived in window 5, after detection and modification of any outliers/inliers, were generally satisfactory. Therefore, some results of analyses are presented only for windows 0 (in which no outlier/inlier detection and modification is attempted) and 5. Window 5 implies that outliers/inliers occurring on the average more often than in 3 samples out of 10 are not modified. Even with the small-sample bias, derivation of distribution parameters from the observed flood series, and acceptance of these parameters as representative of population parameters, the window 5 is believed to yield not only satisfactory flood estimates but also satisfactory distribution parameters.

Table 13. Study Basins and Pertinent Data

No.	USGS No.	Stream and station name	n	A	L	S
		SANGAMON RIVER BASIN				
1	05 571000	Sangamon River at Mahomet	31	362	56.41.	3.59
2	05 572000	Sangamon River at Monticello	71	550	80.04	2.75
3	05 572500	Sangamon River near Oakley	27	774	97.98	2.21
4	05 574000	S.F. Sangamon River near Nokomis	29	11.0	4.89	18.80
5	05 574500	Flat Branch near Taylorville	30	276	47.49	2.01
6	05 575500	S.F. Sangamon River near Kincaid	36	562	51.07	2.01
7	05 576000	S.F. Sangamon River near Rochester	30	867	84.82	1.32
8	05 576500	Sangamon River at Riverton	67	2618	164.83	1.48
9	05 577500	Spring Creek at Springfield	32	107	29.37	5.39
10	05 578500	Salt Creek near Rowell	37	335	53.80	2.59
11	05 579500	Lake Fork near Cornland	32	214	37.00	4.65
12	05 580000	Kickapoo Creek at Waynesville	32	227	36.08	6.23
13	05 580500	Kickapoo Creek near Lincoln	35	306	54.48	5.12
14	05 581500	Sugar Creek near Hartsburg	35	333	42.77	5.76
15	05 582000	Salt Creek near Greenview	38	1804	114.68	2.22
16	05 582500	Crane Creek near Easton	30	26.5	4.30	2.16
17	05 583000	Sangamon River at Oakford	62	5093	222.33	1.27
		ROCK RIVER BASIN (in Illinois)				
1	05 435500	Pecatonica River at Freeport	66	1326	99.14	2.01
2	05 437000	Pecatonica River at Shirland	32	2550	118.50	2.01
3	05 437500	Rock River at Rockton	40	6363	178.14	.84
4	05 438250	Coon Creek at Riley	18	85.1	16.45	5.72
5	05 438500	Kishwaukee River at Belvidere	40	538	41.31	4.59
6	05 439500	S. Br. Kishwaukee R. near Fairdale	40	387.	40.29	2.27
7	05 440000	Kishwaukee River near Perryville	40	1099	52.97	4.07
8	05 440500	Killbuck Creek near Monroe Center	40	117	26.80	6.34
9	05 441000	Leaf River at Leaf River	40	103	18.27	10.45
10	05 443500	Rock River at Como	65	8755	266.76	1.00
11	05 444000	Elkhorn Creek near Penrose	40	146	38.97	4.28
12	05 445500	Rock Creek near Morrison	32	158	38.68	3.91
13	05 446500	Rock River at Joslin	40	9551	309.23	1.11
14	05 447000	Green River at Amboy	40	201	23.63	3.85
15	05 447500	Green River near Geneseo	43	1003	80.41	2.53
16	05 448000	Mill Creek at Milan	40	62.4	22.62	7.44
		LITTLE WABASH RIVER BASIN				
1	03 379500	Little Wabash River below Clay City	65	1131	114.10	2.01
2	03 380475	Horse Creek near Keenes	19	97.2	26.38	4.07
3	03 380500	Skillet Fork near Wayne City	51	464	59.52	1.90
4	03 381500	Little Wabash River at Carmi	40	3102	207.10	1.16

Sensitivity of NO

The number of outliers/inliers, NO, that may be checked at each end of the ranked flood series was obtained from [n/10]; NO = 5 for n > 50. Frequency analyses were made with this NO as well as with other higher or lower numbers of outliers/inliers, designated as NO_1 and NO_2. The 100-year floods obtained in window 5 from LP3 with sample skew and from MD for 2 or 3 values of NO are given in Table 14 for 28 basins. The 100-year floods for window 0 as well as window 5 with NO are given for all the 37 basins. It is evident from the flood values for different values of NO in window 5 that these floods do not differ from each other very much, except in some cases where the observed flood series indicates more outliers/inliers than given by NO.. and NO . The NO can be used as a limiting guide in general. If the number of outliers/inliers is less than NO, the floods which are not outliers/inliers will not be detected or modified.

In the case of Salt Creek near Rowell in the Sangamon Basin (No. 10 and USGS No. 05 578500 in Table 14), NO = [37/10] or 3. The 4 lowest and 4 highest floods together with any modification of these floods with NO = 3 and NO_1 = 1 are given below:

	L1	L2	L3	L4	H4	H3	H2	H1
Observed flood, cfs	829	1040	1090	1310	10,300	10,600	12,400	24,500
Modified, NO_1 =1	762							
Modified, NO = 3	754	982				12,677	15,481	

Because of the detection of 2 higher inliers, H2 and H3, the 100-year flood of 24,849 cfs with NO = 3 is higher than 23,068 cfs with NO_1 = 1 with the MD method. Similar results are obtained for a 1000-year flood — 41,562 cfs with NO = 3 and 37,849 cfs with NO_1 = 1.

Table 14. 100-Year Floods with Different Values of NO

		100-year floods in cfs										
		Window 0		Window 5			Window 5			Window 5		
No.	USGS No.	LP3	MO	NO	LP3	MD	NO_1	LP3	MD	NO_2	LP3	MD
	SANGAMON RIVER BASIN											
1	05 571000	17,954	16,291	3	16,490	16,369	2	17,954	16,291	1	17.954	16,291
2	05 572000	20,277	20,066	5	20,075	19,874	1	20,277	20,066			
3	05 572500	25,630	21,888	2	28,136	24,557						
4	05 574000	11,499	10,419	2	11,273	10,224						
5	05 574500	13,468	16,192	3	13,871	16,294	4	14,016	16,253	2	14,521	16,186
6	05 575500	25,318	25,039	3	25,240	24,962						
7	05 576000	22,345	20,785	3	22,203	20,775						
8	05 576500	38,917	53,173	5	33,931	44,041	3	33,915	43,601	1	35,686	45,887
9	05 577500	10,125	9,866	3	10,943	10,037	4	10,820	9,796	2	10,853	10,181
10	05 578500	28,363	23,058	3	30,311	24,849	1	28,202	23,068			
11	05 579500	7,531	9,860	3	10,188	9,458	2	10,126	9,525	1	10,019	9,637
12	05 580000	25,793	24,284	3	25,312	24,137						
13	05 580500	25,438	23,678	3	25,013	23,562	4	25,673	23,771	2	25,266	23,762
14	05 581500	26,343	29,663	3	25,286	26,829						
15	05 582000	49,963	45,259	3	49,251	44,884						
16	05 582500	863	692	3	903	752	2	902	754	1	878	745
17	05 583000	78,851	89,279	5	70,037	69,535	3	67,647	67,266	1	65,702	65,362
	ROCK RIVER BASIN (in Illinois)											
1	05 435500	22,209	18,266	5	22,394	18,996	3	22,370	19,025	1	22,429	19,026
2	05 437000	21,981	19,893	3	22,234	20,542	1	22,195	20,494			
3	05 437500	35,881	32,527	4	35,900	32,411	2	35,900	32,411			
4	05 438250	4,795	5,927	1	3,943	4,720	2	3,655	4,483	3	4,110	4,629
5	05 438500	16,647	11,505	4	18,027	14,240	2	17,606	13,831	1	17,075	12,708
6	05 439500	10,186	8,823	4	9,914	9,643	2	9,668	9,510			
7	05 440000	24,980	18,665	4	26,412	22,608	2	25,325	21,890	1	24,786	21,618
8	05 440500	7,748	7,325	4	7,688	7,647	2	7,692	7,644			
9	05 441000	11,459	10,368	4	12,024	10,898	2	12,018	10,908			
10	05 443500	58,567	58,555	5	58,584	59,051	3	58,478	58,982			

Table 14. Concluded

No.	USGS No.	100-year floods in cfs: Window 0 LP3	Window 0 MO	Window 5 NO	Window 5 LP3	Window 5 MD	Window 5 NO	Window 5 LP3	Window 5 MD	Window 5 NO	Window 5 LP3	Window 5 MD
	ROCK RIVER BASIN (in Illinois) (Continued)											
11	05 444000	7,300	7,829	4	7,232	7,787						
12	05 445500	5,741	6,348	3	5,183	5,687	1	5,589	6,192			
13	05 446500	60,315	49,036	4	62,690	54,700	2	61,857	53,703			
14	05 447000	6,425	7,156	4	6,445	7,063	3	6,354	7,180	2	6,318	7,068
15	05 447500	12,189	13,305	4	11,907	12,838						
16	05 448000	13,431	12,428	4	13,198	12,165	2	13,861	12,943	1	13,861	12,943
	LITTLE WABASH RIVER BASIN											
1	03 379500	55,586	56,494	5	55,863	55,863	3	57,101	56,445			
2	03 380475	17,873	19,893	1	8,858	8,476	2	11,078	11,388			
3	03 380500	36,719	39,943	5	35,593	37,723	3	33,695	35,498	1	31,902	33,194
4	03 381500	46,614	54,678	4	50,145	50,593	2	52,576	51,232			

Another example is the Kishwaukee River near Belvidere in the Rock River Basin (No. 5 and USGS No. 05 438500). The relevant data with NO = 4, $NO_1 = 2$, and $NO_2 = 1$ are:

	L1	L2	L3	L4	H4	H3	H2	H1
Observed flood, cfs	935	955	1070	1090	9040	9200	9830	10,300
Modified, NO = 4	647	921		1134		9965	11,294	13,694
Modified, $NO_1 = 2$	635	911					11,115	13,397
Modified, $NO_2 = 1$	630							13,018

The 100-year floods with number of outliers/inliers equal to 4, 2, and 1 are 14,240, 13,831, and 12,708 cfs, respectively, with the MD method. It is evident that NO = 4 includes practically all outliers/inliers, and that NO = 3 would have given similar results as NO = 4.

LP3 and MD Statistics

The distribution parameters for LP3 and MD, before and after modification of outliers/inliers (at level or window 0 and 5, respectively) are given in Table 15 for the 37 basins. The LP3 distribution can simulate 3 shapes on lognormal probability paper — convex, straight, and concave for positive, zero, and negative skew, respectively — as shown in figure 14. It cannot simulate symmetrical distributions with kurtosis $\neq$ 3 but the PT method with kurtosis correction is satisfactory in such cases. However, if the distribution is not symmetrical and if the cumulative distribution exhibits an S-curve shape, the mixed distribution method, MD, provides satisfactory flood estimates. The MD becomes a normal distribution when $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$. Some of the diverse shapes that can be simulated or fitted by the MD are shown in figure 14. The dotted lines indicate the two component

Table 15. LP3 and MD Statistics

			ZERO WINDOW								FIFTH WINDOW							
			LP3			MD					LP3			MD				
No.	USGS No.	Outlier*	$\bar{x}$	s	g	a	μ_1	μ_2	σ_1	σ_2	$\bar{x}$	s	g	a	μ_1	μ_2	σ_1	σ_2
	SANGAMON RIVER BASIN																	
1	05 571000		3.607	.285	-.073	.751	3.513	3.891	.249	.174	3.608	.285	-.070	.743	3.510	3.891	.248	.175
2	05 572000	LI	3.697	.291	-.312	.651	3.607	3.867	.295	.192	3.696	.292	-.336	.561	3.573	3.855	.291	.204
3	05 572500	HI	3.755	.250	.398	.378	3.523	3.896	.089	.207	3.758	.258	.486	.362	3.536	3.885	.097	.235
4	05 574000		3.026	.359	.788	.555	2.830	3.270	.186	.373	3.025	.358	.771	.536	2.827	3.254	.182	.374
5	05 574500	LO,LI	3.560	.329	-.803	.060	2.739	3.612	.196	.259	3.560	.325	-.717	.051	2.712	3.606	.135	.263
6	05 575500		3.630	.338	-.051	.729	3.559	3.822	.327	.290	3.630	.338	-.054	.664	3.541	3.807	.324	.293
7	05 576000	LI	3.712	.320	-.453	.557	3.541	3.928	.302	.182	3.711	.321	-.463	.439	3.464	3.904	.278	.195
8	05 576500	LO,HO	4.144	.312	-1.227	.281	3.904	4.238	.442	.166	4.144	.291	-1.386	.159	3.719	4.225	.379	.182
9	05 577500	LO,LI	3.201	.386	-.325	.548	3.024	3.417	.376	.270	3.203	.384	-.203	.870	3.145	3.592	.376	.129
10	05 578500	HI	3.571	.367	.104	.418	3.240	3.809	.196	.262	3.574	.377	.110	.409	3.234	3.809	.205	.276
11	05 579500	LO	3.269	.341	-.730	.013	1.900	3.287	.140	.305	3.285	.298	.140	.000	1.849	3.285	.273	.297
12	05 580000	LI	3.589	.310	.456	.501	3.416	3.763	.200	.303	3.587	.311	.417	.558	3.429	3.786	.215	.298
13	05 580500	LI,HO	3.626	.303	.345	.338	3.413	3.735	.179	.295	3.624	.304	.299	.323	3.415	3.724	.190	.298
14	05 581500	LO	3.682	.277	.465	.966	3.656	4.428	.244	.078	3.682	.273	.438	.964	3.656	4.396	.239	.051
15	05 582000		4.071	.278	-.099	.308	3.772	4.205	.163	.207	4.071	.279	-.128	.298	3.765	4.200	.165	.206
16	05 582500	LI,HI	2.299	.329	-.525	.418	1.977	2.530	.226	.147	2.297	.347	-.579	.583	2.103	2.569	.320	.138
17	05 583000	HO,HI	4.311	.307	-.562	.648	4.221	4.477	.338	.121	4.309	.305	-.764	.380	4.046	4.470	.294	.169
	ROCK RIVER BASIN (in Illinois)																	
1	05 435500	LI,HI	3.767	.251	-.024	.810	3.691	4.094	.211	.094	3.767	.255	-.047	.868	3.714	4.115	.229	.088
2	05 437000	LI,HI	3.912	.208	-.345	.284	3.642	4.019	.096	.127	3.910	.216	-.410	.332	3.670	4.030	.140	.130
3	05 437500	LI,HI	4.128	.200	-.258	.291	3.877	4.231	.091	.128	4.126	.206	-.329	.460	3.955	4.272	.150	.114
4	05 438250	HO	3.026	.399	-.926	.414	2.793	3.191	.484	.201	3.018	.386	-1.133	.120	2.193	3.130	.234	.239
5	05 438500	LI,HI	3.553	.319	-.310	.610	3.351	3.868	.237	.097	3.554	.336	-.318	.660	3.390	3.873	.285	.139
6	05 439500	LI,HI	3.555	.269	-.869	.404	3.296	3.731	.227	.099	3.554	.281	-1.018	.438	3.344	3.717	.296	.106
7	05 440000	LI,HI	3.855	.282	-.541	.524	3.638	4.094	.215	.086	3.856	.301	-.601	.609	3.702	4.094	.285	.108
8	05 440500	HI	3.337	.349	-1.011	.386	3.018	3.527	.343	.145	3.336	.356	-1.062	.398	3.042	3.531	.376	.147
9	05 441000	HI	3.385	.382	-.755	.440	3.081	3.624	.354	.180	3.388	.385	-.712	.513	3.147	3.641	.377	.173
10	05 443500	HI	4.362	.233	-.789	.477	4.217	4.493	.245	.113	4.362	.234	-.800	.462	4.212	4.490	.248	.117
11	05 444000	LO,HO	3.468	.255	-1.049	.300	3.208	3.579	.274	.139	3.468	.252	-1.048	.292	3.208	3.576	.271	.139
12	05 445500	LO	3.333	.196	-.210	.796	3.318	3.392	.217	.046	3.330	.186	-.348	.785	3.305	3.423	.201	.040
13	05 446500	LI,HI	4.341	.218	-.421	.527	4.174	4.528	.158	.079	4.341	.229	-.453	.544	4.188	4.524	.189	.104
14	05 447000	LO	3.387	.283	-1.147	.113	2.732	3.470	.077	.169	3.398	.261	-1.016	.128	2.864	3.477	.146	.164
15	05 447500	HO	3.740	.212	-.940	.148	3.375	3.803	.180	.142	3.740	.207	-.956	.250	3.507	3.818	.222	.129
16	05 448000	HI,HO	3.401	.348	-.317	.206	2.902	3.530	.172	.252	3.399	.347	-.336	.240	2.942	3.543	.199	.244
	LITTLE WABASH RIVER BASIN																	
1	03 379500	LI,HI,HO	4.088	.315	-.324	.435	3.924	4.214	.322	.244	4.086	.316	-.318	.490	3.932	4.235	.315	.236
2	03 380475	HO	3.596	.231	.729	.489	3.533	3.657	.005	.311	3.579	.188	-.493	.303	3.726	3.514	.069	.187
3	03 380500	LI,HO,HI	3.886	.334	-.394	.108	3.454	3.939	.357	.290	3.887	.336	-.467	.612	3.781	4.053	.364	.190
4	03 381500	LO,L1,HO	4.147	.234	-.126	.519	4.131	4.164	.293	.143	4.151	.216	.306	.633	4.199	4.067	.235	.143

* LI, LO, HI, and HO denote low inlier, low outlier, high inlier, and high outlier respectively.

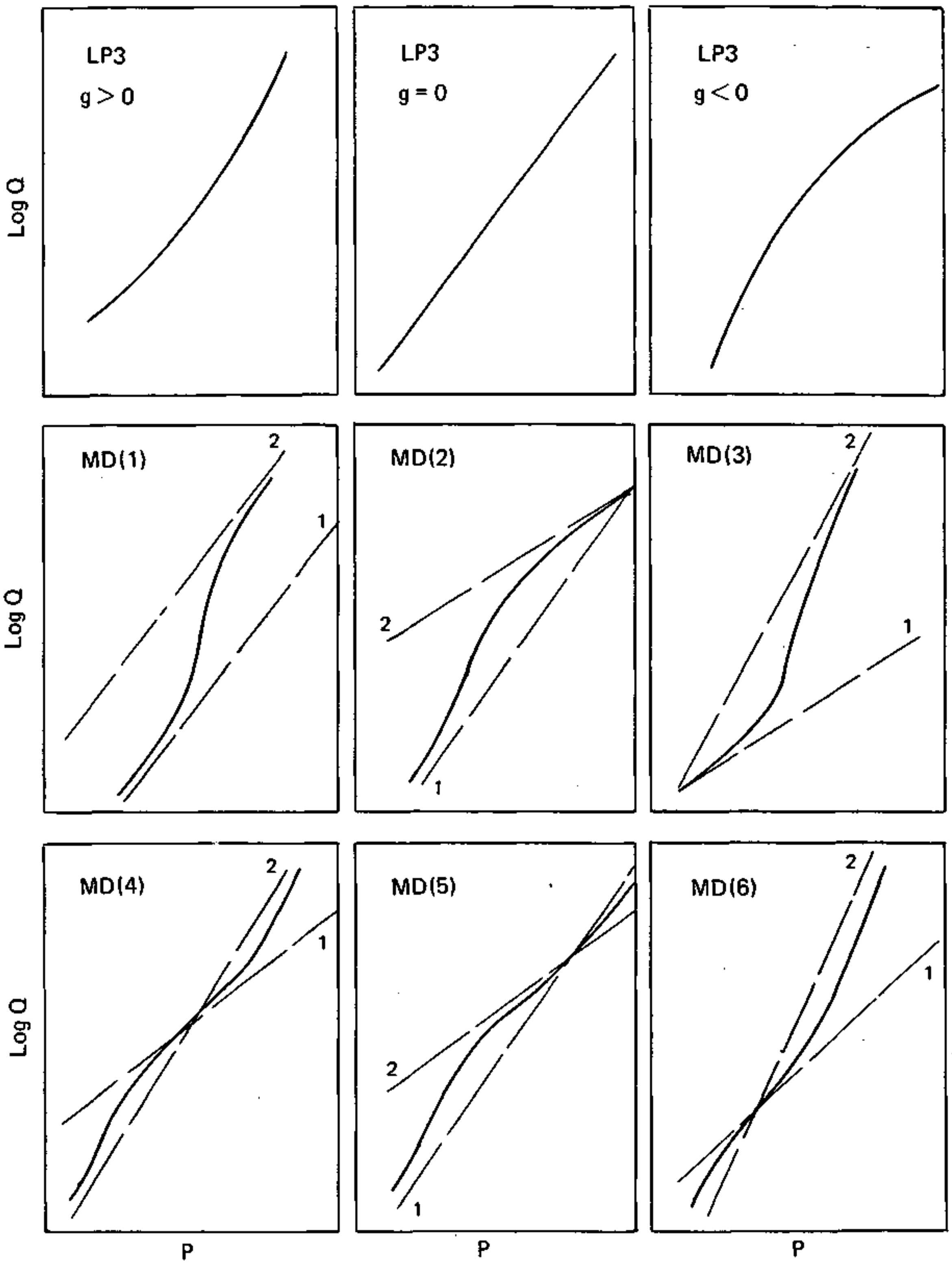


Figure 14. Probability curves for LP3 and MD

distributions and the solid curve the mixed distribution. The S-curve shapes can, to some extent, be caused or accentuated by the existence of outliers/inliers.

LP3 Statistics

It is generally felt that modification of any outliers/inliers in an observed flood series would change skew significantly. However, with the exception of 3 basins, the values of skew obtained without any modification of outliers/inliers and with modification of outliers/inliers detected up to level 5 or in window 5 are not much different from each other for the remaining 34 basins. The exceptions are: 05 579500, Lake Fork near Cornland, with a very low outlier, g changes from -0.730 in window 0 to 0.140 in window 5; 03 380475, Horse Creek near Keenes, a very high outlier, g changes from 0.729 in window 0 to -0.493 in window 5; and 03 381500, Little Wabash River at Carmi, with low outlier and inlier and a high outlier, g changes from -0.126 in window 0 to 0.306 in window 5.

Modification of high outlier(s) and/or low inlier(s) has the effect of making the skew value smaller in the algebraic sense, and the modification of high inlier(s) and/or low outlier(s) makes the skew value larger. The change in skew from window 0 to 5 can be explained generally by the type of outliers/inliers.

According to the U.S. Water Resources Council Bulletin 17, the regional skew for the 37 basins analyzed is about -0.4. The number of basins with lower and higher skew and the minimum and maximum values of skew in a major river basin are given on the next page (from information for window 5 in Table 15).

River Basin	Number of sub-basins with g < -0.4	g > -0.4	Min g	Max g
Sangamon	5	12	-1.386	0.771
Rock	11	5	-1.133	-0.047
Little Wabash	2	2	-0.493	0.306

Even after modification of outliers/inliers, the derived values of skew do not indicate a regional skew of -0.4. Values of skew for basins with more than 60 years of record are:

USGS No.	n	g(window 0)	g(window 5)
05 572000	71	-0.312	-0.336
05 576500	67	-1.227	-1.386
05 583000	62	-0.562	-0.764
05 435500	66	-0.024	-0.047
05 443500	65	-0.789	-0.800
03 379500	65	-0.324	-0.318

Again, a regional skew value of -0.4 is not indicated by the above 6 long-term stations.

MD Statistics

The mixed distribution has two component distributions, the parameters of mean and standard deviation carry subscripts 1 and 2, and the weight of the first distribution is given by *a*. The mean, μ_1 , of the first component distribution is smaller than μ_2 for the second distribution. The general shape of mixed distribution can be categorized from figure 14 with the relative values of σ_1 and σ_2. Some distortions in these shapes can be caused by unequal weight of the two component distributions. A brief summary of the MD statistics is given below:

Item	Number of sub-basins in the basin of		
	Sangamon	Rock	Little Wabash
$a > .5$	*9(8)	5(6)	1(2)
$a < .5$	8(9)	11(10)	3(2)
$\mu_1 < \mu_2$	17(17)	16(16)	4(2)
$\mu_1 > \mu_2$	0(0)	0(0)	0(2)
$\sigma_1 > \sigma_2$	9(9)	12(13)	3(3)
$\sigma_1 < \sigma_2$	8(8)	4(3)	1(1)

* for window 0 and in parentheses for window 5.

The flood series with $\sigma_1 > \sigma_2$ will have shapes similar to MD(2) and MD(5) and with $\sigma_1 < \sigma_2$ will resemble MD(3) and MD(6) in figure 14. The probability curves are affected largely by $\Delta\mu$ (or $\mu_1 - \mu_2$), $\Delta\sigma$ (or $\sigma_1 - \sigma_2$), and a. A few points of interest regarding the mixed distribution and kurtosis are

1. With $\mu_1 = \mu_2$ and $a = 0.5$

$$\mu = 0.5\ (\mu_1 + \mu_2) = \mu_1 = \mu_2 \quad (1)$$

$$\sigma^2 = 0.5\ (\sigma_1^{\ 2} + \sigma_2^{\ 2}) \quad (2)$$

$$g = 0 \quad (3)$$

$$kt = 1.5\ (\sigma_1^{\ 4} + \sigma_2^{\ 4})\ /\ \sigma^4 \quad (4)$$

in which kt is the kurtosis. The simulated distributions are symmetrical with kurtosis = 3 with σ_1/σ_2 or $\sigma_2/\sigma_1 = 1$, kt = 4.92 for a ratio of 3 or 1/3, and kt =6.0 for a ratio of infinity or zero.

2. With $\sigma_1 = \sigma_2$ and $a = 0.5$

$$\mu = 0.5\ (\mu_1 + \mu_2) \quad (5)$$

$$\sigma^2 = 0.5(\sigma_1{}^2 + \sigma_2{}^2) + 0.25(\Delta\mu)^2 = \sigma_1{}^2 + 0.25(\Delta\mu)^2 \tag{6}$$

$$g = 0 \tag{7}$$

$$k = [12\ \sigma_1{}^4 + 6\ \sigma_1{}^2\ (\Delta\mu)^4]\ /\ 4\sigma^4 \tag{8}$$

in which σ_2 is replaced by σ_1. The simulated distributions are symmetrical and yield kurtosis $= \leq 3$ depending on the ratio of $\Delta\mu\ /\ \sigma_1$, keeping the mixed distribution unimodal.

The probability distributions for cases 1 and 2 correspond to MD(4) and MD(1), respectively, in figure 14. Only in these special cases, the kurtosis correction with the PT method may yield flood estimates comparable to those from the MD. However, the asymmetry of the observed flood distributions for the 37 study basins, in terms of mixed distribution parameters varying from those for cases 1 and 2, indicates that flood estimates from the MD will be better than from the PT with kurtosis correction (based on the assumption of symmetrical distribution).

Ratios of Q_{100}/Q_2 and Q_{1000}/Q_2

Ratios of a 100-year flood, Q_{100}, to a 2-year flood, Q_2, and 1000-year flood, Q_{1000}, to Q_2 both for windows 0 and 5 and with LP3 sample skew and MD methods are given in Table 16. These ratios are plotted with respect to drainage area on a log-log paper in figures 15 and 16 for the Sangamon and in figures 17 and 18 for the Rock River basins for drainage areas exceeding 100 square miles.

Ratios for the Sangamon Basin

The ratios Q_{100}/Q_2 from LP3 in figure 15 indicate considerable scatter whereas they lie along two trend curves (one for the Salt Creek basins and

Table 16. Ratios Q_{100}/Q_2 and Q_{1000}/Q_2

			Q_{100}/Q_2				Q_{1000}/Q_2			
			Window 0		Window 5		Window 0		Window 5	
No.	USGS No.	A	LP3	MD	LP3	MD	LP3	MD	LP3	MD
SANGAMON RIVER BASIN										
1	05 571000	362	4.47	4.01	4.42	4.03	7.03	5.74	7.07	5.76
2	05 572000	550	3.93	3.83	3.89	3.79	5.71	5.99	5.60	5.74
3	05 572500	774	4.68	4.13	5.15	4.74	8.56	6.09	9.99	7.35
4	05 574000	11	12.07	11.65	11.82	11.46	36.65	23.90	35.37	23.40
5	05 574500	276	3.35	4.15	3.50	4.21	4.16	6.57	4.46	6.71
6	05 475500	562	5.89	5.83	5.88	5.80	10.39	10.13	10.34	10.07
7	05 576000	867	4.10	3.65	4.09	3.65	5.81	5.36	5.76	5.25
8	05 576500	2618	2.42	3.37	2.10	2.84	2.64	7.85	2.21	4.06
9	05 577500	107	6.07	5.79	6.65	6.10	9.89	9.85	11.58	11.88
10	05 578500	335	7.73	6.28	8.21	6.73	15.66	10.23	17.04	11.26
11	05 579500	214	3.69	5.16	5.37	4.90	4.74	8.82	9.70	8.27
12	05 580000	227	7.02	6.80	6.89	6.71	15.31	12.09	14.74	11.92
13	05 580500	306	6.26	5.97	6.15	5.90	12.65	10.27	12.20	10.20
14	05 581500	333	5.75	6.39	5.50	5.77	11.61	8.14	10.83	6.78
15	05 582000	1804	4.19	3.68	4.13	3.64	6.55	5.38	6.38	5.30
16	05 582500	26.5	4.05	2.84	4.22	3.17	5.61	3.73	5.81	4.74
17	05 583000	5093	3.61	3.81	3.14	2.97	4.82	7.10	3.89	4.12
ROCK RIVER BASIN										
1	05 435500	1326	3.79	3.22	3.82	3.32	5.83	4.07	5.86	4.53
2	05 437000	2550	2.62	2.22	2.64	2.32	3.39	2.79	3.39	2.95
3	05 437500	6363	2.62	2.24	2.61	2.27	3.44	2.82	3.39	2.81
4	05 438250	85.1	3.93	4.67	3.21	3.84	4.84	11.33	3.66	5.87
5	05 438500	538	4.49	3.12	4.83	3.74	6.76	3.81	7.41	5.12
6	05 439500	387	2.60	1.98	2.49	2.21	3.04	2.41	2.81	3.49
7	05 440000	1099	3.29	2.14	3.44	2.68	4.33	2.55	4.50	4.13
8	05 440500	117	3.12	2.65	3.08	2.78	3.65	3.75	3.55	4.60
9	05 441000	103	4.23	3.44	4.44	3.63	5.55	5.00	5.94	5.89
10	05 443500	8755	2.38	2.27	2.37	2.29	2.78	3.23	2.77	3.26
11	05 444000	146	2.25	2.35	2.23	2.35	2.50	3.14	2.47	3.12
12	05 445500	158	2.63	2.79	2.36	2.46	3.48	4.14	2.97	3.54
13	05 446500	9551	2.65	2.00	2.75	2.27	3.39	2.33	3.52	2.78
14	05 447000	201	2.33	2.58	2.33	2.53	2.56	3.49	2.62	3.38
15	05 447500	1003	2.06	2.24	2.00	2.15	2.29	2.89	2.23	2.74
16	05 448000	62.4	5.12	4.44	5.03	4.36	7.98	7.00	7.77	6.79

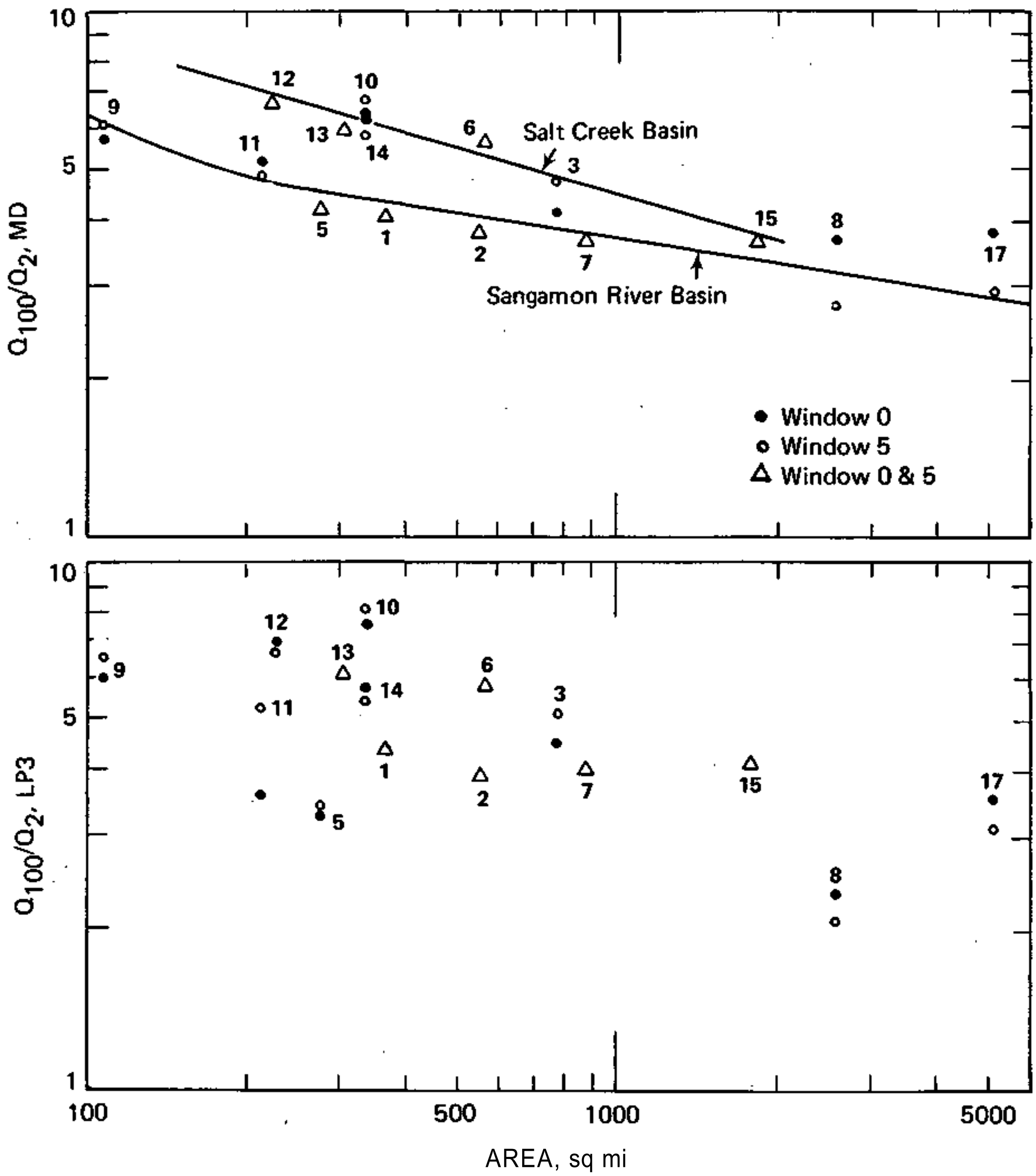


Figure 15. Q_{100}/Q_2 versus drainage area, Sangamon River Basin

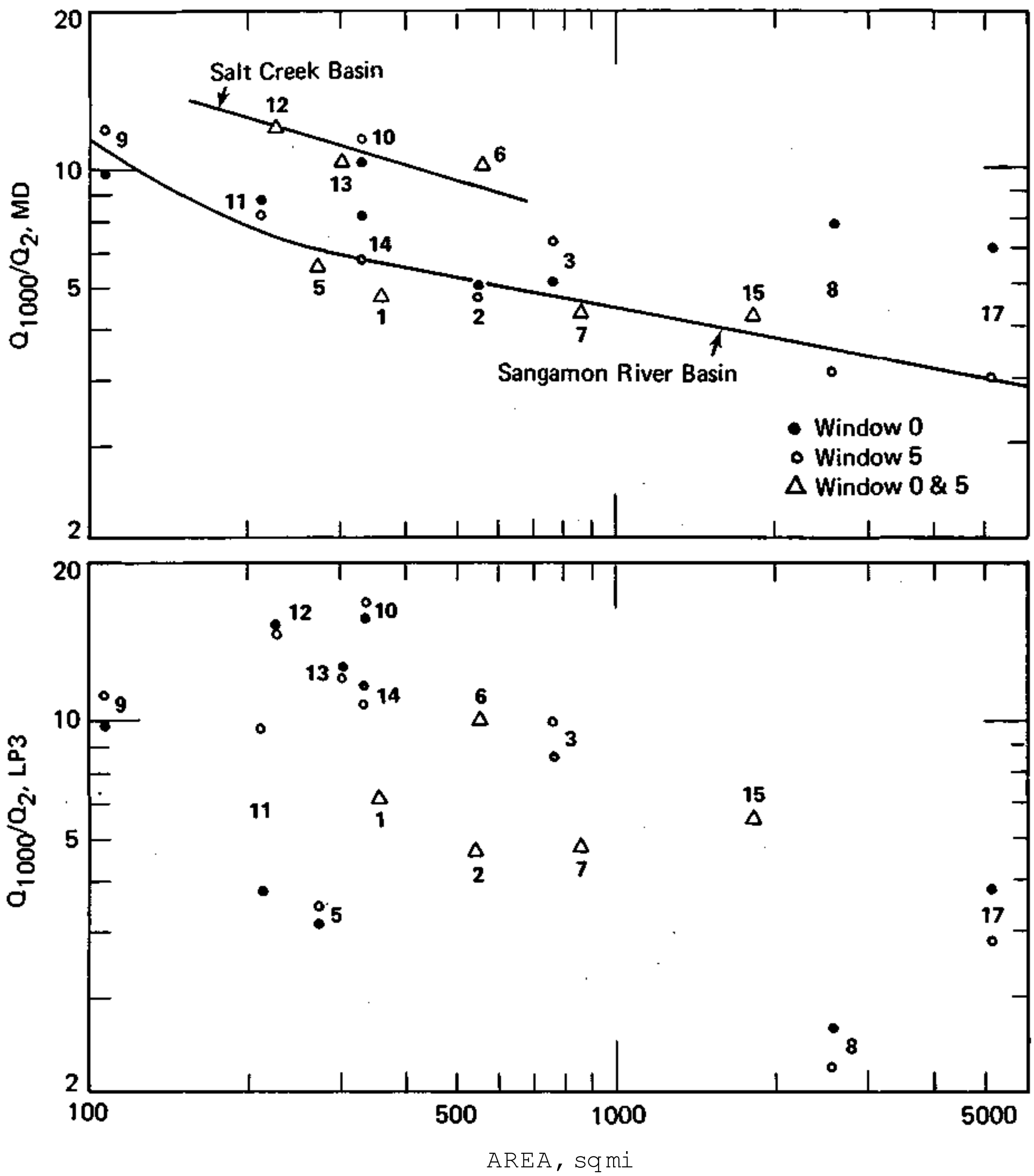


Figure 16. Q_{1000}/Q_2 versus drainage area, Sangamon River Basin

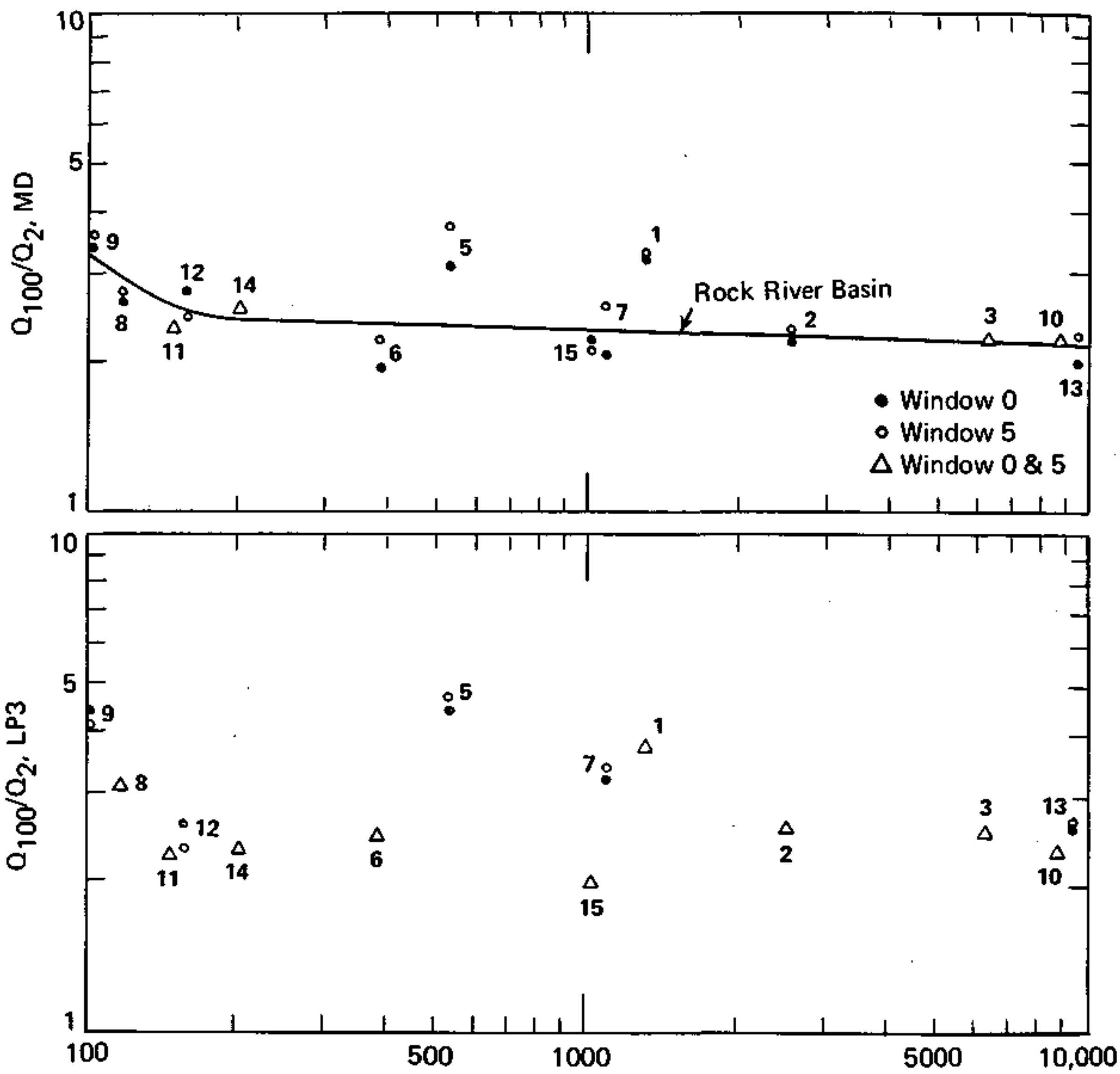


Figure 17. Q_{100}/Q_2 versus drainage area, Rock River Basin

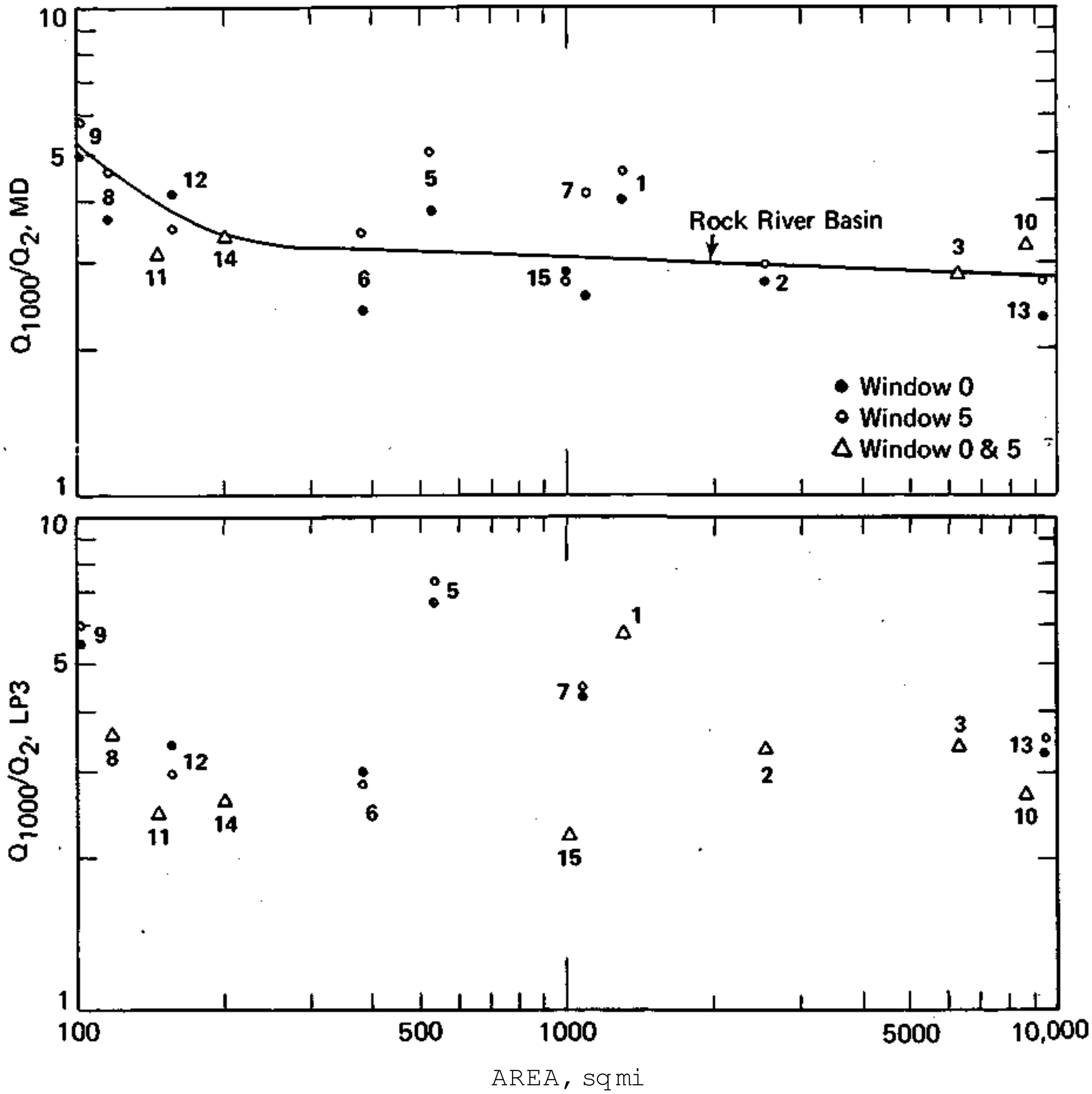


Figure 18. Q_{1000}/Q_2 versus drainage area, Rock River Basin

the other for the Sangamon River basins) for the MD and window 5. The curves represent a decrease in the value of the ratio with increase in drainage area. Similar results are shown by the Q_{1000}/Q_2 plots in figure 16. The trend curves steepen as the drainage area becomes less than 200 square miles.

Ratios for the Book Basin

The ratios of Q_{100}/Q_2 and Q_{1000}/Q_2, in figures 17 and 18, show that the MD method with window 5 indicates satisfactorily the decrease in these ratios with drainage area. However, the decrease is much smaller than in the Sangamon Basin. The trend curves steepen as area decreases below 200 square miles. The ratios with LP3 show considerable scatter.

Some Specific Examples

Examples of various types of outliers/inliers are discussed. The fitted probability curves with the MD and window 5, and computer output tables showing the modification of outliers/inliers from one window to the next are used to explain each example.

1. Lake Fork near Cornland: Low Outlier

The results obtained with the computer program are given in Table 17. It contains the 100-year flood estimates from five methods; observed and modified floods for NO points; statistics with PT, LP3, and MD methods for all the windows 0 through 6; and 2- to 1,000-year floods with the five methods for all the windows. The three high floods are not perceived as outliers or inliers up to window 5. The lowest flood is perceived as a significant low outlier but the next two floods as rather insignificant low outliers. The

Table 17. Flood Frequency Analyses: Lake Fork near Cornland

STATION NO. 5579500 LAKE FORK NEAR CORNLAND
DRAINAGE AREA 214.0 Sq Mi Years of Record 32 (1948-1979)

LEVEL NO.	0		1 2	3	4	5	6
METHOD	100-Year Flood in cfs						
Power Transform, PT							
With kt = 3.0	8,736	8,736	8,736	8,878	9,639	10,277	11,510
With sample kt	9,838	9,838	9,838	9,942	10,361	10,624	11,035
Log Transform							
LP3, Sample skew	7,531	7,531	7,531	8,068	9,625	10,188	10,666
LP3, Weighted skew	8,964	8,964	8,964	8,763	8,209	7,950	7,641
Mixed Distrib., MD	9,860	9,860	9,860	9,878	9,820	9,458	8,951

Type	No.	Observed and Modified Floods in cfs						
Low	1*	152	152	152	183	336	443	574
	2*	548	548	548	548	548	594	712
	3*	680	680	680	680	680	704	812
	4	1,000						
	5	1,010						
High	5	4,570						
	4	4,700						
	3*	4,900	4,900	4,900	4,900	4,900	4,900	4,936
	2*	6,120	6,120	6,120	6,120	6,120	6,120	6,120
	1*	8,930	8,930	8,930	8,930	8,930	8,930	8,876

METHOD	STATISTICS	Values of Statistics						
PT	mean	19.126	19.126	19.126	16.273	8.519	5.803	3.560
	std dev	3.848	3.848	3.848	2.941	.899	.394	.111
	skew	.036	.036	.036	.026	.001	.005	.028
	kurtosis,kt	4.045	4.045	4.045	3.932	3.433	3.162	2.852
	5th moment	-1.494	-1.494	-1.494	-1.448	-.919	-.593	-.292
	lambda	.215	.215	.215	.181	.031	-.073	-.232
LP3	mean	3.269	3.269	3.269	3.271	3.280	3.285	3.293
	std dev	.341	.341	.341	.333	.309	.298	.283
	sample skew	-.730	-.730	-.730	-.560	-.068	.140	.370
	kurtosis,kt	5.438	5.438	5.438	4.852	3.477	3.136	2.933
	5th moment	-12.793	-12.793	-12.793	-9.442	-1.618	.614	2.507
MD	weight 'a'	.013	.013	.013	.010	.001	.000	0.000
	mu1	1.900	1.900	1.900	1.967	1.849	1.849	1.849
	mu2	3.287	3.287	3.287	3.285	3.281	3.285	3.293
	sigma 1	.140	.140	.140	.171	.273	.273	.273
	sigma 2	.305	.305	.305	.305	.306	.297	.283
	Test Stat	3.258	3.258	3.258	3.342	.001	.004	.009

* High & low floods considered for outlier detection and modification

Table 17. Concluded

STATION NO. 5579500 LAKE FORK NEAR CORNLAND
DRAINAGE AREA 214.0 Sq Mi Years of Record 32 (1948-1979)

VARIOUS RECURRENCE-INTERVAL FLOODS

METHOD	#	Flood in cfs for Recurrence Intervals (Years) 2	10	25	50	100	500	1000
PT, kt=3.0	0	1,976	4,748	6,306	7,508	8,736	11,703	13,033
PT, sample kt		1,976	4,579	6,428	8,036	9,838	14,824	17,334
LP3, sample skew		2,042	4,681	5,915	6,758	7,531	9,092	9,676
weighted skew		1,965	4,862	6,486	7,722	8,964	11,864	13,114
MD, mixed dist.		1,912	4,731	6,580	8,139	9,860	14,527	16,866
PT, kt=3.0	1							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	2							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	3	1,965	4,742	6,340	7,589	8,878	12,039	13,475
PT, sample kt		1,965	4,587	6,460	8,097	9,942	15,096	17,719
LP3, sample skew		2,006	4,707	6,095	7,099	8,068	10,187	11,046
weighted skew		1,970	4,788	6,362	7,559	8,763	11,578	12,796
MD, mixed dist.		1,911	4,729	6,582	8,148	9,878	14,567	16,916
PT, kt=3.0	4	1,919	4,716	6,514	8,011	9,639	13,974	16,099
PT, sample kt		1,919	4,633	6,600	8,347	10,361	16,184	19,274
LP3, sample skew		1,920	4,716	6,511	8,004	9,625	13,932	16,040
weighted skew		1,989	4,589	6,023	7,112	8,209	10,787	11,908
MD, mixed dist.		1,911	4,707	6,547	8,106	9,820	14,478	16,810
PT, kt=3.0	5	1,896	4,688	6,632	8,337	10,277	15,854	18,798
PT, sample kt		1,896	4,654	6,672	8,496	10,624	16,994	20,485
LP3, sample skew		1,897	4,685	6,611	8,290	10,188	15,583	18,397
weighted skew		2,006	4,504	5,871	6,907	7,950	10,400	11,468
MD, mixed dist.		1,929	4,632	6,383	7,852	9,458	13,790	15,945
PT, kt=3.0	6	1,873	4,630	6,806	8,903	11,510	20,391	25,966
PT, sample kt		1,873	4,663	6,755	8,697	11,035	18,563	23,023
LP3, sample skew		1,886	4,628	6,655	8,500	10,666	17,232	20,872
weighted skew		2,034	4,412	5,696	6,666	7,641	9,927	10,923
MD, mixed dist.		1,963	4,528	6,149	7,493	8,951	12,826	14,730

= level number

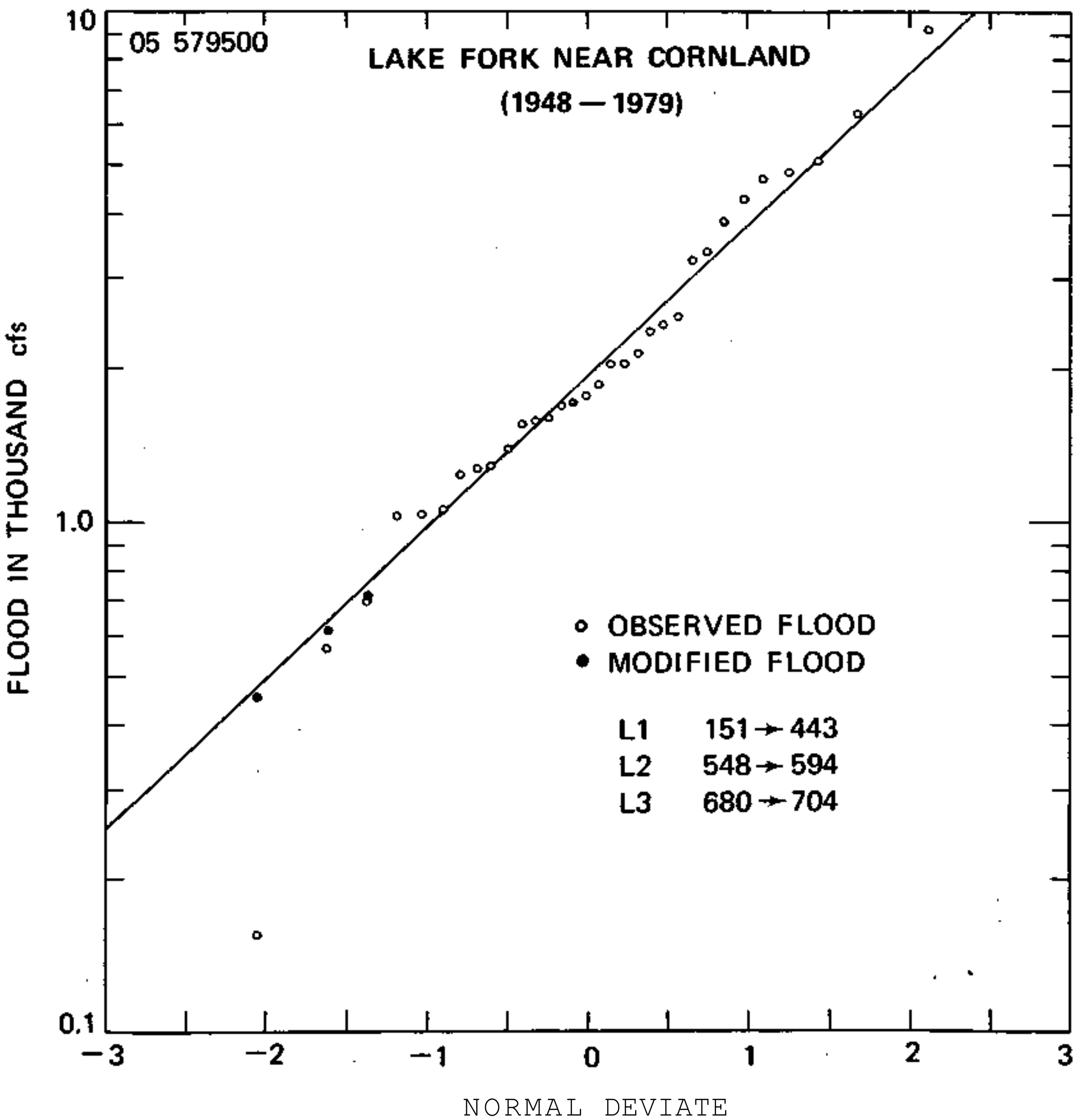


Figure 19. Observed and modified floods and the fitted mixed distribution curve (window 5) for the Lake Fork near Cornland

three lowest floods of 152, 548, and 680 cfs are modified to 443, 594, and 704 cfs, respectively, in window 5. The modification of low outliers changes Q_{100} and Q_{1000} estimates of 7,531 and 9,676 cfs in window 0 to 10,188 and 18,397 in window 5 (because of sample skew changing from -0.730 to 0.140) with the LP3 and sample skew. The mean and standard deviation change from 3.269 and 0.341 to 3.285 and 0.298.

The PT statistics show that the power transformation reduces skew close to zero and the kurtosis decreases from 4.045 in window 0 to 3.162 in window 5; the kurtosis for a theoretical normal distribution is 3.0. Flood estimates with the PT are higher/lower with sample kurtosis than with kt=3.0 if sample kurtosis is higher/lower than 3.0. The PT 100-year flood estimate with kt=3.0 increases from 8,736 in window 0 to 10,277 in window 5. With sample kt, it increases from 9,838 cfs to 10,624 cfs.

The MD flood estimates are rather insensitive to modification of low outliers. The 100-year flood changes from 9,860 cfs in window 0 to 9,458 cfs in window 5 and a 1000-year flood changes from 16,866 cfs to 15,945 cfs. The MD statistics show that the effect of the first component distribution is negligible, the weight being a maximum of 0.013, and that the distribution is practically normal (which is indicated by the LP3 in between windows 4 and 5). The MD method seems to be the best for analyzing observed flood series with low outliers. The observed floods as well as the modified low floods in the 5th window and the probability curve fitted by the MD method are shown in figure 19.

2. Rock River at Rockton: Low and High Inliers

The results obtained with the computer program are given in Table 18.

Table 18. Flood Frequency Analyses: Rock River at Rockton

STATION NO. 5437500 ROCK RIVER AT ROCKTON
DRAINAGE AREA 6363.0 Sq Mi Years of Record 40 (1940-1979)

LEVEL NO.		0	1	2	3	4	5	6
METHOD		100-Year Flood in cfs						
Power Transform, PT								
With kt = 3.0		34,432	34,432	34,432	34,432	34,465	34,544	35,249
With sample kt		30,933	30,933	30,933	30,933	31,137	31,740	33,100
Log Transform								
LP3, Sample skew		35,881	35,881	35,881	35,881	35,910	35,900	36,330
LP3, Weighted skew		34,512	34,512	34,512	34,512	34,698	35,185	35,997
Mixed Distrib., MD		32,527	32,527	32,527	32,527	32,461	32,411	33,772
Type	No.	Observed and Modified Floods in cfs						
Low	1*	5,400	5,400	5,400	5,400	5,183	4,692	4,258
	2*	6,340	6,340	6,340	6,340	6,222	5,821	5,451
	3*	6,340	6,340'	6,340	6,340	6,340	6,340	6,267
	4*	6,880	6,880	6,880	6,880	6,880	6,880	6,880
	5	7,450						
High	5	23,800						
	4*	24,300	24,300	24,300	24,300	24,300	24,300	24,300
	3*	25,400	25,400	25,400	25,400	25,400	25,400	25,874
	2*	25,700	25,700	25,700	25,700	26,247	27,048	28,076
	1*	30,000	30,000	30,000	30,000	30,000	30,434	31,950
METHOD	STATISTICS	Values of Statistics						
PT	mean	58.809	58.809	58.809	58.809	63.875	79.332	79.342
	std dev	8.686	8.686	8.686	8.686	9.801	13.388	13.733
	skew	-.051	-.051	-.051	-.051	-.052	-.054	-.047
	kurtosis,kt	2.190	2.190	2.190	2.190	2.218	2.310	2.459
	5th moment	-.007	-.007	-.007	-.007	-.047	-.120	-.095
	lambda	.310	.310	.310	.310	.322	.353	.353
LP3	mean	4.128	4.128	4.128	4.128	4.128	4.126	4.126
	std dev	.200	.200	.200	.200	.201	.206	.212
	sample skew	-.258	-.258	-.258	-.258	-.276	-.329	-.368
	kurtosis,kt	2.211	2.211	2.211	2.211	2.259	2.414	2.617
	5th moment	-1.187	-1.187	-1.187	-1.187	-1.365	-1.921	-2.460
MD	weight 'a'	.291	.291	.291	.291	.329	.460	.485
	mu1	3.877	3.877	3.877	3.877	3.896	3.955	3.972
	mu2	4.231	4.231	4.231	4.231	4.242	4.272	4.270
	sigma1	.091	.091	.091	.091	.108	.150	.173
	sigma2	.128	.128	.128	.128	.124	.114	.124
	Test Stat	2.827	2.827	2.827	2.827	2.688	2.575	2.370

* High & low floods considered for outlier detection and modification

Table 18. Concluded

STATION NO. 5437500 ROCK RIVER AT ROCKTON
DRAINAGE AREA 6363.0 Sq Mi Years of Record 40 (1940-1979)

VARIOUS RECURRENCE-INTERVAL FLOODS

METHOD	#	Flood in cfs for Recurrence Intervals (Years) 2	10	25	50	100	500	1000
PT, kt=3.0	0	13,866	23,613	28,125	31,335	34,432	41,335	44,224
PT, sample kt		13,866	24,099	27,333	29,272	30,933	34,121	35,304
LP3, sample skew		13,703	23,886	28,818	32,392	35,881	43,774	47,115
weighted skew		13,822	23,716	28,264	31,464	34,512	41,152	43,859
MD, mixed dist.		14,543	23,392	27,184	29,892	32,527	38,517	41,082
PT, kt=3.0	1							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	2							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	3							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	4	13,876	23,662	28,174	31,379	34,465	41,331	44,199
PT, sample kt		13,876	24,128	27,426	29,419	31,137	34,456	35,692
LP3, sample skew		13,710	23,947	28,880	32,443	35,910	43,720	47,010
weighted skew		13,816	23,796	28,388	31,619	34,698	41,402	44,134
MD, mixed dist.		14,487	23,489	27,231	29,890	32,461	38,286	40,768
PT, kt=3.0	5	13,898	23,783	28,295	31,484	34,544	41,312	44,125
PT, sample kt		13,898	24,195	27,677	29,842	31,740	35,490	36,899
LP3, sample skew		13,732	24,096	29,006	32,515	35,900	43,412	46,530
weighted skew		13,794	24,005	28,714	32,028	35,185	42,054	44,850
MD, mixed dist.		14,296	23,747	27,415	29,968	32,411	37,869	40,167
PT, kt=3.0	6	13,903	24,091	28,763	32,071	35,249	42,291	45,221
PT, sample kt		13,903	24,413	28,314	30,829	33,100	37,720	39,512
LP3, sample skew		13,762	24,392	29,385	32,930	36,330	43,803	46,875
weighted skew		13,791.	24,350	29,249	32,703	35,997	43,170	46,091
MD, mixed dist.		14,189	24,067	28,137	31,002	33,772	40,026	42,688

= level number

The NO equals [40/10] or 4. However, in going from window 0 to 5, only two low inliers, L1 and L2, are detected and these are modified from their original values of 5,400 and 6,340 cfs to 4,692 and 5,821 cfs, respectively. Two high inliers, H1 and H2, are detected and are modified from 30,000 and 25,700 cfs to 30,434 and 27,048 cfs. Generally, the modification of low inliers should decrease the skew and of high inliers should increase the skew. With the LP3, the skew decreases from -0.258 in window 0 to -0.329 in window 5. Because of a small change in skew (as well as mean and standard deviation), the 100- and 1000-year floods of 35,881 and 47,115 cfs in window 0 change to 35,900 and 46,530 cfs in window 5, with LP3 and sample skew.

The PT statistics show that the power transformation reduces skew close to zero but the kurtosis changes from 2.190 in window 0 to 2.310 in window 5. Thus, the flood estimates with sample kurtosis are considerably smaller than with kt=3.0. However, the 100- and 1000-year flood estimates with kt=3.0 are close to those with LP3 and sample skew.

The MD flood estimates are rather insensitive to modification of inliers as are the estimates with the PT and LP3. The 100- and 1000-year floods change from 32,527 and 41,082 cfs in window 0 to 32,411 and 40,167 cfs in window 5. Flood estimates from the five methods are summarized below:

	100-year flood, cfs		1000-year flood, cfs	
	window 0	window 5	window 0	window 5
PT, kt=3.0	34,432	34,544	44,224	44,125
sample kt	30,933	31,740	35,304	36,899
LP3, sample g	35,881	35,900	47,115	46,530
weighted *g*	34,512	35,185	43,859	44,850
MD	32,527	32,411	41,082	40,167

The MD statistics in window 5 show that neither $\mu_1 = \mu_2$ nor $\sigma_1 = \sigma_2$.

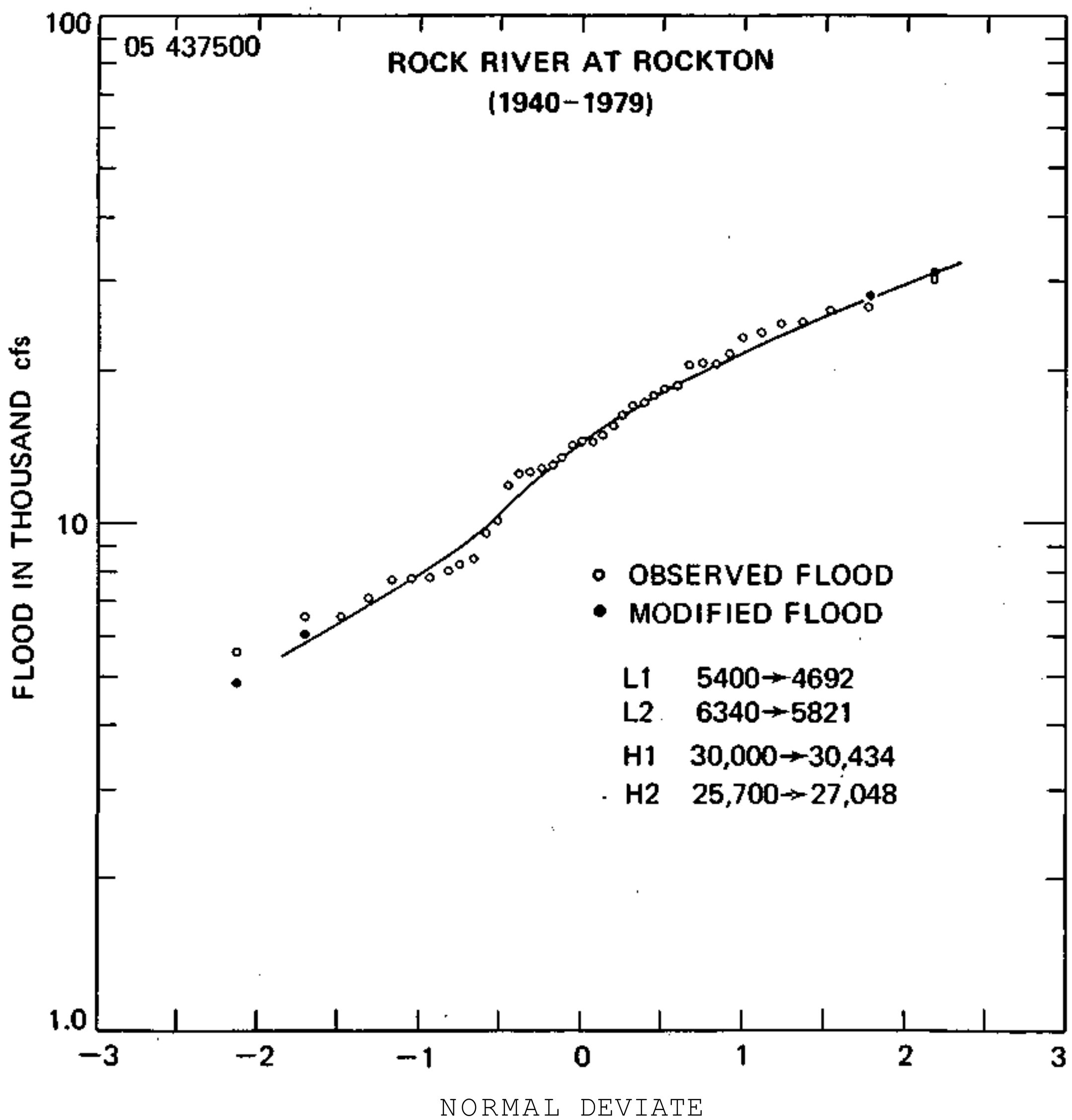


Figure 20. Observed and modified floods and the fitted mixed distribution curve (window 5) for the Rock River at Rockton

Thus, the flood series is asymmetrical. Plots in figures 17 and 18 show that MD estimates for this basin lie on the well-defined regional curve for Q_{100}/Q_2 and Q_{1000}/Q_2. Thus, the flood estimates with the MD are considered better than with other methods. The observed floods as well as the modified floods in the 5th window and the probability curve fitted by the MD method are shown in figure 20.

3. Flat Branch near Taylorville: Low Outliers and Inlier

The results obtained with the computer program are given in Table 19. The NO equals [30/10] or 3. In going from window 0 to 5, two low ouliers, L1 and L2, and a low inlier, L3, are detected and modified from their original values of 457, 660, and 1,770 cfs to 460, 841, and 1,370 cfs, respectively. Only one high oulier, H2, is detected in window 5 and it is modified from 11,300 cfs to 11,032 cfs which is relatively an insignificant modification. The LP3 sample skew changes from -0.803 in window 0 to -0.717 in window 5, the standard deviation from 0.329 to 0.325, and the mean remains unchanged. Because of a slight increase in skew (in the algebraic sense), the 100- and 1000-year floods are 13,468 and 16,691 cfs in window 0 and 13,871 and 17,697 cfs in window 5. In this example, the effects of low outlier and inlier practically balance each other.

The PT statistics indicate a skew very close to zero and a kurtosis of 3.256 in window 0 and 3.154 in window 5. Accordingly, the flood estimates with the sample kt are somewhat higher than with kt = 3.0, and the high flood estimates are higher than those with the LP3 and sample skew.

The MD statistics indicate a weight of only 0.06 to 0.05 for the first component distribution with a mean of 2.739 which is much smaller than

Table 19. Flood Frequency Analyses: Flat Branch near Taylorville

STATION NO. 5574500 FLAT BRANCH NEAR TAYLORVILLE
DRAINAGE AREA 276.0 Sq Mi Years of Record 30 (1950-1979)

LEVEL NO.		0	1	2	3	4	5	6
METHOD		100-Year Flood in cfs						
Power Transform, PT								
With kt = 3.0		14,564	14,564	14,599	14,623	14,688	14,643	15,015
With sample kt		15,002	15,002	15,007	15,005	15,020	14,921	14,816
Log Transform								
LP3, Sample skew		13,468	13,468	13,549	13,616	13,821	13,871	15,148
LP3, Weighted skew		16,633	16,633	16,679	16,725	16,575	16,351	15,151
Mixed Distrib., MD		16,192	16,192	16,219	16,228	16,379	16,294	15,948
Type	No.	Observed and Modified Floods in cfs						
Low	1*	457	457	457	457	457	460	695
	2*	660	660	660	660	745	841	1,054
	3*	1,770	1,770	1,649	1,551	1,440	1,370	1,324
	4	1,850						
	5	1,860						
High	5	7,540						
	4	8,620						
	3*	9,400	9,400	9,400	9,400	9,400	9,400	9,267
	2*	11,300	11,300	11,300	11,300	11,300	11,032	10,612
	1*	13,000	13,000	13,000	13,000	13,000	13,000	13,000
METHOD	STATISTICS	Values of Statistics						
PT	mean	45.450	45.450	45.152	45.127	42.291	40.590	23.845
	std dev	11.452	11.452	11.389	11.417	10.400	9.748	4.281
	skew	-.016	-.016	-.019	-.020	-.020	-.022	-.025
	kurtosis,kt	3.256	3.256	3.235	3.218	3.185	3.154	2.908
	5th moment	-.558	-.558	-.543	-.519	-.500	-.513	-.237
	lambda	.337	.337	.336	.336	.325	.318	.223
LP3	mean	3.560	3.560	3.559	3.558	3.559	3.560	3.568
	std dev	.329	.329	.330	.331	.328	.325	.302
	sample skew	-.803	-.803	-.796	-.790	-.747	-.717	-.400
	kurtosis,kt	4.302	4.302	4.252	4.208	4.108	4.046	3.159
	5th moment	-8.841	-8.841	-8.647	-8.479	-8.072	-7.872	-3.496
MD	weight 'a'	.060	.060	.061	.063	.053	.051	.055
	mu1	2.739	2.739	2.743	2.747	2.699	2.712	2.868
	mu2	3.612	3.612	3.612	3.612	3.607	3.606	3.608
	sigma 1	.196	.196	.194	.192	.137	.135	.005
	sigma2	.259	.259	.260	.260	.263	.263	.258
	Test Stat	2.483	2.483	2.387	2.310	2.187	2.150	3.145

* High & low floods considered for outlier detection and modification

Table 19. Concluded

STATION NO. 5574500 FLAT BRANCH NEAR TAYLORVILLE
DRAINAGE AREA 276.0 Sq Mi. Years of Record 30 (1950-1979)

VARIOUS RECURRENCE-INTERVAL FLOODS

METHOD	#	2	10	25	50	100	500	1000
		Flood in cfs for Recurrence Intervals (Years)						
PT, kt=3.0	0	3,966	8,700	11,087	12,839	14,564	18,512	20,200
PT, sample kt		3,966	8,629	11,153	13,069	15,002	19,570	21,577
LP3, sample skew		4,015	8,777	10,865	12,240	13,468	15,842	16,691
weighted skew		3,833	9,199	12,158	14,395	16,633	21,827	24,057
MD, mixed dist.		3,902	8,617	11,448	13,743	16,192	22,570	25,627
PT, kt=3.0	1	SAME AS ABOVE						
PT, sample kt								
LP3, sample skew								
weighted skew								
MD, mixed dist.								
PT, kt=3.0	2	3,958	8,706	11,103	12,865	14,599	18,571	20,270
PT, sample kt		3,958	8,640	11,166	13,079	15,007	19,555	21,551
LP3, sample skew		4,004	8,790	10,902	12,299	13,549	15,976	16,848
weighted skew		3,824	9,205	12,178	14,427	16,679	21,910	24,158
MD, mixed dist.		3,901	8,618	11,454	13,759	16,219	22,605	25,683
PT, kt=3.0	3	3,952	8,710	11,115	12,882	14,623	18,609	20,315
PT, sample kt		3,952	8,649	11,175	13,082	15,005	19,527	21,510
LP3, sample skew		3,994	8,802	10,934	12,348	13,616	16,087	16,978
weighted skew		3,817	9,212	12,199	14,460	16,725	21,991	24,255
MD, mixed dist.		3,897	8,623	11,463	13,769	16,228	22,630	25,703
PT, kt=3.0	4	3,943	8,710	11,135	12,923	14,688	18,745	20,487
PT, sample kt		3,943	8,656	11,187	13,097	15,020	19,544	21,527
LP3, sample skew		3,976	8,805	10,999	12,477	13,821	16,496	17,483
weighted skew		3,820	9,161	12,110	14,341	16,575	21,766	23,998
MD, mixed dist.		3,874	8,631	11,514	13,864	16,379	22,938	26,119
PT, kt=3.0	5	3,937	8,675	11,094	12,879	14,643	18,707	20,454
PT, sample kt		3,937	8,631	11,137	13,025	14,921	19,376	21,324
LP3, sample skew		3,965	8,770	10,985	12,491	13,871	16,656	17,697
weighted skew		3,823	9,088	11,981	14,166	16,351	21,422	23,599
MD, mixed dist.		3,868	8,604	11,467	13,798	16,294	22,810	25,962
PT, kt=3.0	6	3,889	8,602	11,142	13,069	15,015	19,639	21,683
PT, sample kt		3,889	8,632	11,110	12,964	14,816	19,156	21,051
LP3, sample skew		3,871	8,698	11,285	13,221	15,148	19,590	21,490
weighted skew		3,871	8,698	11,286	13,223	15,151	19,597	21,498
MD, mixed dist.		3,886	8,518	11,293	13,549	15,948	22,183	25,193

= level number

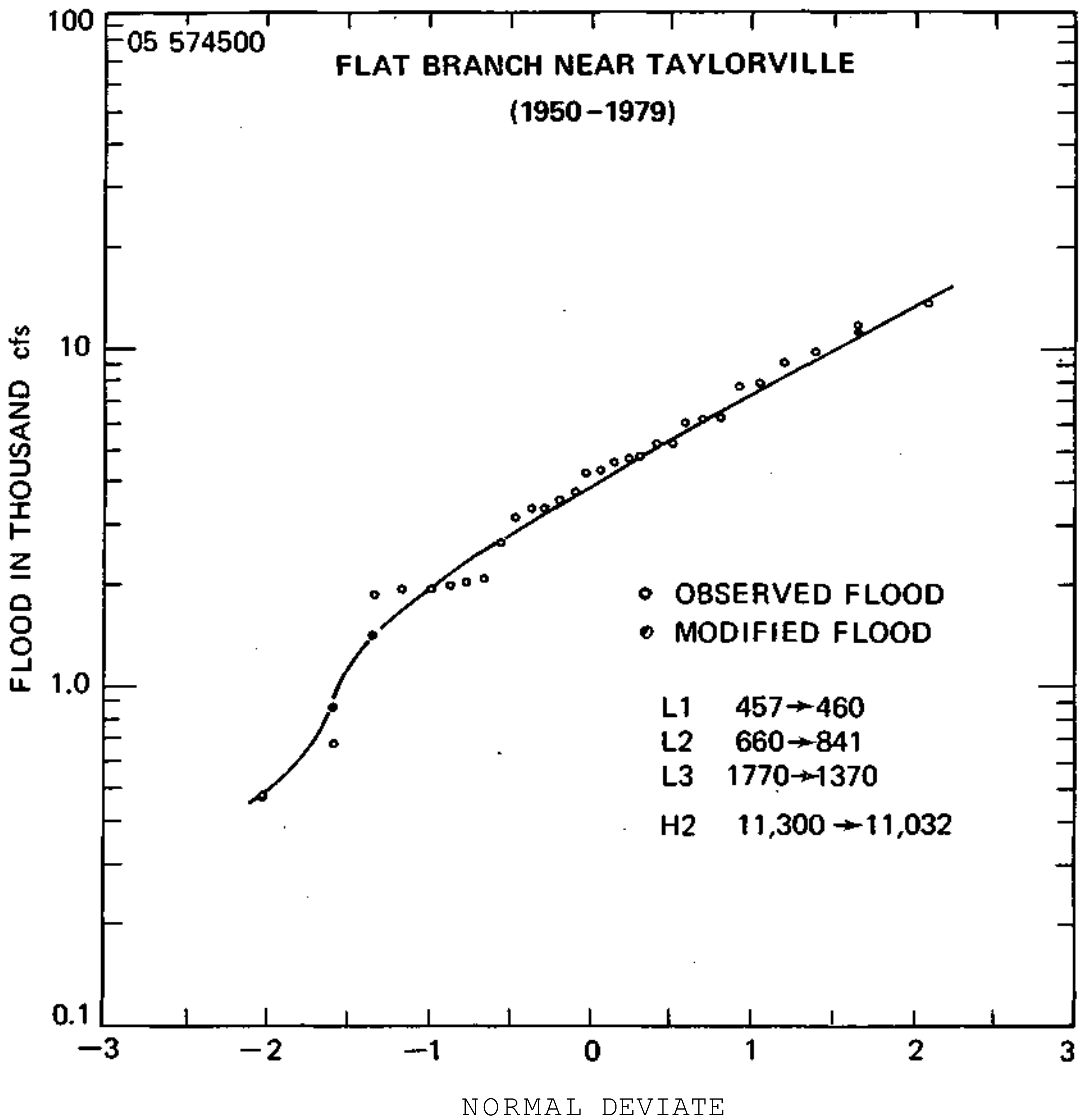


Figure 21. Observed and modified floods and the fitted mixed distribution curve (window 5) for the Flat Branch near Taylorville

$\mu_2 = 3.612$. The σ_1 is also smaller than σ_2. Because of the small value of *a* and the large difference in μ_1 and μ_2, the effect of the first component distribution is felt only in the beginning position of the fitted distribution as shown in figure 21. The LP3 cannot fit such a probability curve and the PT method may not be precise because the power transformed series is not exactly symmetrical (5th moment is not close to zero). Therefore, the flood estimates with the MD method are considered better than with the others.

4. Horse Creek near Keenes: High Outlier

The results obtained with the computer program are given in Table 20. The NO equals [19/10] or 1. Only one high outlier is indicated and it is modified from its value of 17,100 cfs to 9,170 cfs in the 4th window and 7,889 cfs in the 5th window. The LP3 statistics show that skew changes from 0.729 to -0.493 and standard deviation from 0.231 to 0.188 in going from window 0 to 5. The 100- and 1000-year floods are 17,873 and 35,635 cfs in window 0 and 8,858 and 10,756 cfs in window 5. Though the second highest observed flood in 19 years is 5,890 cfs, the 100- and 1000-year flood estimates are much lower than the observed flood of 17,100 cfs.

The PT statistics indicate a decrease in kurtosis from 4.004 in window 0 to 2.885 in window 5 and the corresponding 5th moment values are 2.264 and 0.877. The MD estimates of 100- and 1000-year floods are 19,893 and 36,193 cfs in window 0 and 8,476 and 11,823 cfs in window 5.

Because the flood estimates seem rather low and because a 19-year record is quite close to a 20-year record, analyses were made with NO = 2. The results are presented in Table 21. A summary of the flood estimates is:

Table 20. Flood Frequency Analyses: Horse Creek near Keenes

STATION NO. 3380475 HORSE CREEK NEAR KEENES
DRAINAGE AREA 97.2 Sq Mi Years of Record 19 (1960-1979)

LEVEL NO.		0	1	2	3	4	5	6
METHOD		100-Year Flood in cfs						
Power Transform, PT								
With kt = 3.0		17,246	17,246	17,246	17,246	9,704	8,565	8,099
With sample kt		20,839	20,839	20,839	20,839	9,931	8,488	7,900
Log Transform								
LP3, Sample skew		17,873	17,873	17,873	17,873	9,813	8,858	8,491
LP3, Weighted skew		11,191	11,191	11,191	11,191	9,464	9,150	9,009
Mixed Distrib., MD		19,893	19,893	19,893	19,893	10,308	8,476	8,159
Type	No.	Observed and Modified Floods in cfs						
Low	1*	1,550	1,550	1,550	1,550	1,550	1,550	1,550
	2	2,200						
	3	2,220						
	4	2,270						
	5	2,330						
High	5	5,260						
	4	5,420						
	3	5,840						
	2	5,890						
	1*	17,100	17,100	17,100	17,100	9,170	7,889	7,325
METHOD	STATISTICS	Values of Statistics						
PT	mean	2.729	2.729	2.729	2.729	30.871	171.591	506.877
	std dev	.030	.030	.030	.030	4.124	39.215	142.134
	skew	-.018	-.018	-.018	-.018	-.021	-.069	-.107
	kurtosis,kt	4.004	4.004	4.004	4.004	3.261	2.885	2.669
	5th moment	2.264	2.264	2.264	2.264	1.724	.877	.312
	lambda	-.345	-.345	-.345	-.345	.270	.550	.710
LP3	mean	3.596	3.596	3.596	3.596	3.582	3.579	3.577
	std dev	.231	.231	.231	.231	.195	.188	.185
	sample skew	.729	.729	.729	.729	-.300	-.493	-.572
	kurtosis,kt	5.446	5.446	5.446	5.446	3.072	2.830	2.763
	5th moment	12.918	12.918	12.918	12.918	-1.109	-2.914	-3.569
MD	weight 'a'	.489	.489	.489	.489	.162	.303	.723
	mu1	3.533	3.533	3.533	3.533	3.716	3.726	3.667
	mu2	3.657	3.657	3.657	3.657	3.556	3.514	3.342
	sigma1	.005	.005	.005	.005	.034	.069	.111
	sigma2	.311	.311	.311	.311	.202	.187	.124
	Test Stat	7.645	7.645	7.645	7.645	2.729	2.804	2.235

* High & low floods considered for outlier detection and modification

Table 20. Concluded

STATION NO. 3380475 HORSE CREEK NEAR KEENES
DRAINAGE AREA 97.2 Sq Mi Years of Record 19 (1960-1979)

VARIOUS RECURRENCE-INTERVAL FLOODS

METHOD	#	Flood in cfs for Recurrence Intervals (Years)						
		2	10	25	50	100	500	1000
PT, kt=3.0	0	3,775	7,895	10,928	13,771	17,246	28,927	36,270
PT, sample kt		3,775	7,613	11,190	15,146	20,839	48,257	73,728
LP3, sample skew		3,702	8,017	11,265	14,275	17,873	29,158	35,635
weighted skew		4,121	7,536	9,093	10,174	11,191	13,353	14,211
MD, mixed dist.		3,429	8,385	12,512	16,017	19,893	30,634	36,193
PT, kt=3.0	1							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	2							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	3							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	4	3,918	6,635	7,908	8,820	9,704	1.1,688	12,525
PT, sample kt		3,918	6,596	7,944	8,941	9,931	12,224	13,216
LP3, sample skew		3,906	6,676	7,981	8,913	9,813	11,813	12,645
weighted skew		3,937	6,630	7,837	8,674	9,464	11,156	11,835
MD, mixed dist.		4,034	6,256	7,823	9,051	10,308	13,414	14,846
PT, kt=3.0	5	3,972	6,309	7,284	7,947	8,565	9,881	10,409
PT, sample kt		3,972	6,323	7,269	7,903	8,488	9,719	10,208
LP3, sample skew		3,926	6,420	7,481	8,197	8,858	10,225	10,756
weighted skew		3,898	6,462	7,607	8,401	9,150	10,756	11,401
MD, mixed dist.		4,081	6,196	7,002	7,672	8,476	10,758	11,823
PT, kt=3.0	6	3,999	6,156	7,009	7,577	8,099	9,185	9,613
PT, sample kt		3,999	6,193	6,967	7,460	7,900	8,783	9,121
LP3, sample skew		3,930	6,310	7,277	7,915	8,491	9,649	10,085
weighted skew		3,879	6,386	7,503	8,278	9,009	10,576	11,207
MD, mixed dist.		4,102	6,135	6,985	7,583	8,159	9,445	9,987

= level number

Table 21. Flood Frequency Analyses: Horse Creek near Keenes

STATION NO. 3380475 HORSE CREEK NEAR KEENES
DRAINAGE AREA 97.2 Sq Mi Years of Record 19 (1960-1979)

LEVEL NO.		0	1	2	3	4	5	6
METHOD		100-Year Flood in cfs						
Power Transform, PT								
With kt = 3.0		17,246	17,452	17,770	18,006	13,332	10,975	9,172
With sample kt		20,839	20,922	21,108	21,273	14,194	11,105	9,012
Log Transform								
LP3, Sample skew		17,873	17,971	18,144	18,285	13,277	11,078	9,486
LP3, Weighted skew		11,191	11,308	11,458	11,558	10,640	10,084	9,531
Mixed Distrib., MD		19,893	19,866	19,906	19,932	13,897	11,388	9,258
Type	No.	Observed and Modified Floods in cfs						
Low	1*	1,550	1,550	1,550	1,550	1,550	1,550	1,550
	2*	2,200	2,200	2,200	2,200	2,171	2,107	2,040
	3	2,220						
	4	2,270						
	5	2,330						
High	5	5,260						
	4	5,420						
	3	5,840						
	2*	5,890	6,311	6,846	7,199	7,633	7,633	6,913
	1*	17,100	17,100	17,100	17,100	12,166	9,694	8,090
METHOD	STATISTICS	Values of Statistics						
PT	mean	2.729	2.756	2.769	2.770	5.524	15.012	85.919
	std dev	.030	.031	.032	.032	.209	1.410	16.962
	skew	-.018	-.015	-.011	-.008	.005	-.013	-.060
	kurtosis,kt	4.004	3.948	3.877	3.834	3.450	3.105	2.805
	5th moment	2.264	2.174	2.029	1.929	1.585	1.048	.550
	lambda	-.345	-.341	-.339	-.339	-.105	.132	.442
LP3	mean	3.596	3.598	3.600	3.601	3.594	3.588	3.581
	std dev	.231	.232	.234	.235	.217	.206	.196
	sample skew	.729	.710	.691	.683	.155	-.154	-.413
	kurtosis,kt	5.446	5.317	5.162	5.067	3.615	3.027	2.767
	5th moment	12.918	12.343	11.659	11.243	3.210	-.225	-2.307
MD	weight 'a'	.489	.476	.463	.454	.187	.090	.786
	mu1	3.533	3.533	3.532	3.531	3.544	3.729	3.655
	mu2	3.657	3.657	3.658	3.659	3.606	3.574	3.309
	sigma1	.005	.005	.005	.005	.005	.050	.139
	sigma2	.311	.309	.307	.306	.239	.211	.121
	Test Stat	7.645	7.403	7.122	6.937	4.440	2.883	2.309

* High & low floods considered for outlier detection and modification

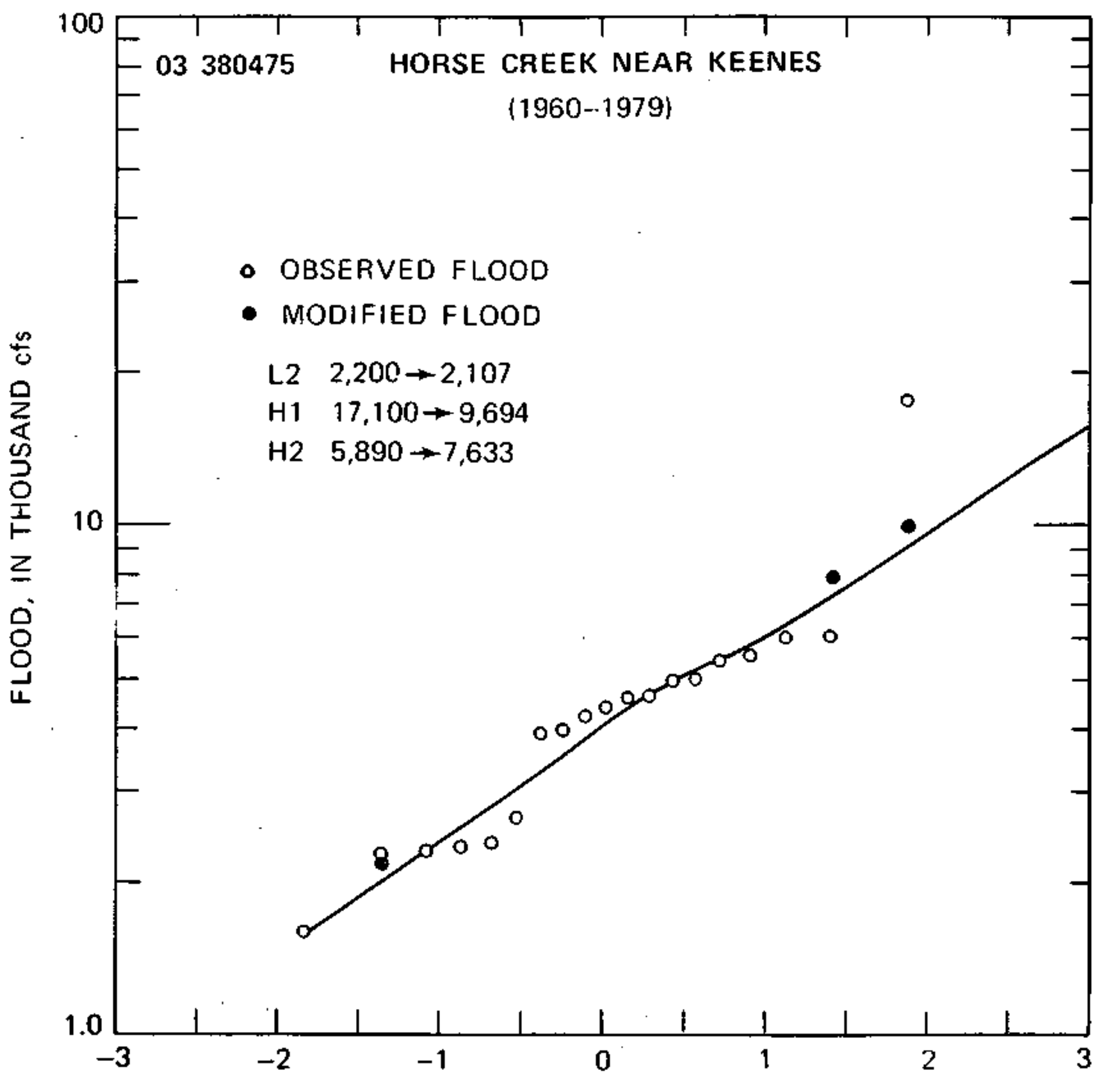
Table 21. Concluded

STATION NO. 3380475 HORSE CREEK NEAR KEENES
DRAINAGE AREA 97.2 Sq Mi Years of Record 19 (1960-1979)

VARIOUS RECURRENCE-INTERVAL FLOODS

METHOD	#	Flood in cfs for Recurrence Intervals (Years) 2	10	25	50	100	500	1000
PT, kt=3.0	0	3,775	7,895	10,928	13,771	17,246	28,927	36,270
PT, sample kt		3,775	7,613	11,190	15,146	20,839	48,257	73,728
LP3, sample skew		3,702	8,017	11,265	14,275	17,873	29,158	35,635
weighted skew		4,121	7,536	9,093	10,174	11,191	13,353	14,211
MD, mixed dist.		3,429	8,385	12,512	16,017	19,893	30,634	36,193
PT, kt=3.0	1	3,789	7,960	11,037	13,923	17,452	29,317	36,773
PT, sample kt		3,789	7,688	11,296	15,258	20,922	47,697	72,034
LP3, sample skew		3,720	8,076	11,344	14,367	17,971	29,241	35,689
weighted skew		4,137	7,593	9,174	10,273	11,308	13,511	14,387
MD, mixed dist.		3,425	8,460	12,564	16,036	19,866	30,476	35,953
PT, kt=3.0	2	3,803	8,043	11,187	14,144	17,770	30,019	37,751
PT, sample kt		3,803	7,785	11,440	15,434	21,108	47,482	71,037
LP3, sample skew		3,740	8,154	11,458	14,509	18,144	29,483	35,956
weighted skew		4,155	7,666	9,278	10,400	11,458	13,714	14,613
MD, mixed dist.		3,421	8,560	12,645	16,094	19,906	30,422	35,824
PT, kt=3.0	3	3,811	8,098	11,290	14,303	18,006	30,577	38,553
PT, sample kt		3,811	7,848	11,539	15,567	21,273	47,611	70,982
LP3, sample skew		3,752	8,206	11,540	14,618	18,285	29,718	36,244
weighted skew		4,167	7,715	9,347	10,485	11,558	13,849	14,763
MD, mixed dist.		3,413	8,626	12,699	16,138	19,932	30,371	35,742
PT, kt=3.0	4	3,879	7,507	9,674	11,438	13,332	18,322	20,763
PT, sample kt		3,879	7,401	9,779	11,843	14,194	20,954	24,557
LP3, sample skew		3,877	7,504	9,660	11,408	13,277	18,160	20,529
weighted skew		4,075	7,238	8,681	9,688	10,640	12,682	13,501
MD, mixed dist.		3,548	7,636	10,013	11,906	13,897	18,986	21,407
PT, kt=3.0	5	3,929	7,039	8,620	9,799	10,975	13,738	14,950
PT, sample kt		3,929	7,018	8,640	9,867	11,105	14,057	15,367
LP3, sample skew		3,922	7,060	8,670	9,873	11,078	13,910	15,154
weighted skew		4,005	6,942	8,275	9,205	10,084	11,974	12,735
MD, mixed dist.		3,983	6,832	8,578	9,962	11,388	14,925	16,554
PT, kt=3.0	6	3,973	6,540	7,661	8,437	9,172	10,765	11,415
PT, sample kt		3,973	6,567	7,631	8,347	9,012	10,423	10,987
LP3, sample skew		3,931	6,639	7,852	8,693	9,486	11,184	11,865
weighted skew		3,927	6,645	7,871	8,724	9,531	11,267	11,967
MD, mixed dist.		4,045	6,512'	7,640	8,456	9,258	11,107	11,907

= level number



NORMAL DEVIATE

Figure 22. Observed and modified floods and the fitted mixed distribution curve (window 5) for the Horse Creek near Keenes

	100-year flood window 0	window 5	1000-year flood window 0	window 5
PT, kt = 3.0	17,246	10,975	36,270	14,950
Sample kt	20,839	11,105	73,728	15,367
LP3, sample g	17,873	11,078	35,635	15,154
weighted g	11,191	10,084	14,211	12,735
MD	19,893	11,388	36,193	16,554

With NO = 2, the changes in distribution statistics are less than with NO = 1. Also, the flood estimates are more in line with those indicated by storm frequency and runoff conditions (Singh, 1980). The observed floods as well as the modified floods in the 5th window and the probability curve fitted by the MD method are shown in figure 22.

5. *Sangamon River near Oakley: High Inlier*

The results obtained with the computer program are given in Table 22. The NO equals [27/10] or 2. Only one high inlier is indicated and it is modified from its value of 16,000 cfs in window 0 to 20,085 cfs in window 5. An insignificant low inlier is also indicated. The value changes from 2,390 cfs to 2,321 cfs in window 5. The LP3 statistics show that skew increases from 0.398 to 0.486 and standard deviation from 0.250 to 0.258 in going from window 0 to 5. The 100- and 1000-year floods, with LP3 and sample skew, increase from 25,630 and 46,891 cfs in window 0 to 28,136 and 54,588 cfs in window 5.

The PT statistics indicate an increase in kurtosis from 2.202 in window 0 to 2.311 in window 5 and the corresponding 5th moment values are 0.322 and 0.403. The 100- and 1000-year flood estimates with PT and sample kurtosis are 23,346 and 35,558 cfs for window 0 and 27,145 and 48,531 cfs for window 5. The flood estimates with kurtosis = 3.0 are much higher. The MD

Table 22. Flood Frequency Analyses: Sangamon River near Oakley

STATION NO. 5572500 SANGAMON RIVER NEAR OAKLEY
DRAINAGE AREA 774.0 Sq Mi Years of Record 27 (1951-1977)

LEVEL NO.		0	1	2	3	4	5	6 .
METHOD		100-Year Flood in cfs						
Power Transform, PT								
With kt = 3.0		32,191	32,191	32,191	32,191	34,458	36,775	38,599
With sample kt		2.3,346	23,346	23,346	23,346	24,824	27,145	30,550
Log Transform								
LP3, Sample skew		25,630	25,630	25,630	25,630	26,685	28,136	30,069
LP3, Weighted skew		18,472	18,472	18,472	18,472	18,840	19,357	20,109
Mixed Distrib., MD		21,888	21,888	21,888	21,888	22,929	24,557	26,840
Type	No.	Observed and Modified Floods in cfs						
Low	1*	2,390	2,390	2,390	2,390	2,390	2,321	2,191
	2*	2,660	2,660	2,660	2,660	2,660	2,660	2,565
	3	3,020						
	4	3,020						
	5	3,120						
High	5	11,800						
	4	13,200						
	3	13,700						
	2*	15,300	15,300	15,300	15,300	15,300	15,300	15,619
	1*	16,000	16,000	16,000	16,000	17,766	20,085	22,841
METHOD	STATISTICS	Values of Statistics						
PT	mean	2.409	2.409	2.409	2.409	2.282	2.225	2.287
	std dev	.017	.017	.017	.017	.014	.013	.015
	skew	.093	.093	.093	.093	.098	.097	.086
	kurtosis,kt	2.202	2.202	2.202	2.202	2.232	2.311	2.455
	5th moment	.322	.322	.322	.322	.388	.403	.325
	lambda	-.402	-.402	-.402	-.402	-.427	-.439	-.426
LP3	mean	3.755	3.755	3.755	3.755	3.757	3.758	3.759
	std dev	.250	.250	.250	.250	.253	.258	.266
	sample skew	.398	.398	.398	.398	.439	.486	.528
	kurtosis,kt	2.309	2.309	2.309	2.309	2.395	2.551	2.773
	5th moment.	1.997	1.997	1.997	1.997	2.373	2.934	3.612
MD	weight 'a'	.378	.378	.378	.378	.385	.362	.391
	mu1	3.523	3.523	3.523	3.523	3.531	3.536	3.558
	mu2	3.896	3.896	3.896	3.896	3.898	3.885	3.889
	sigma1	.089	.089	.089	.089	.095	.097	.120
	sigma2	.207	.207	.207	.207	.216	.235	.253
	Test Stat	2.955	2.955	2.955	2.955	2.654	2.331	2.132

* High & low floods considered for outlier detection and modification

Table 22. Concluded

STATION NO. 5572500 SANGAMON RIVER NEAR OAKLEY
DRAINAGE AREA 774.0 Sq Mi Years of Record 27 (1951-1977)

VARIOUS RECURRENCE-INTERVAL FLOODS

METHOD	#	Flood in cfs for Recurrence Intervals (Years)						
		2	10	25	50	100	500	1000
PT, kt=3.0	0	5,349	12,244	18,056	24,073	32,191	66,028	93,010
PT, sample kt		5,349	12,750	16,874	20,058	23,346	31,621	35,558
LP3, sample skew		5,479	12,144	16,795	20,909	25,630	39,454	46,891
weighted skew		5,902	11,581	14,409	16,466	18,472	22,979	24,866
MD, mixed dist.		5,305	12,636	16,261	19,039	21,888	28,984	32,290
PT, kt=3.0	1							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	2							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	3							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	4	5,342	12,414	18,605	25,220	34,458	76,489	113,754
PT, sample kt		5,342	12,921	17,403	20,995	24,824	34,974	40,045
LP3, sample skew		5,476	12,332	17,223	21,604	26,685	41,819	50,090
weighted skew		5,927	1.1,736	14,646	16,767	18,840	23,511	25,471
MD, mixed dist.		5,231	12,912	16,808	19,820	22,929	30,746	34,428
PT, kt=3.0	5	5,338	12,629	19,209	26,421	36,775	87,423	136,530
PT, sample kt		5,338	13,106	18,104	22,363	27,145	41,021	48,531
LP3, sample skew		5,465	12,582	17,803	22,554	28,136	45,122	54,588
weighted skew		5,951	11,943	14,971	17,186	19,357	24,268	26,336
MD, mixed dist.		5,186	13,220	17,560	20,978	24,557	33,731	38,116
PT, kt=3.0	6	5,339	12,920	19,858	27,522	38,599	93,404	147,001
PT, sample kt		5,339	13,302	19,019	24,253	30,550	51,083	63,687
LP3, sample skew		5,445	12,920	18,578	23,820	30,069	49,547	60,637
weighted skew		5,969	12,228	15,432	17,790	20,109	25,385	27,618
MD, mixed dist.		5,153	13,677	18,643	22,622	26,840	37,733	42,944

= level number

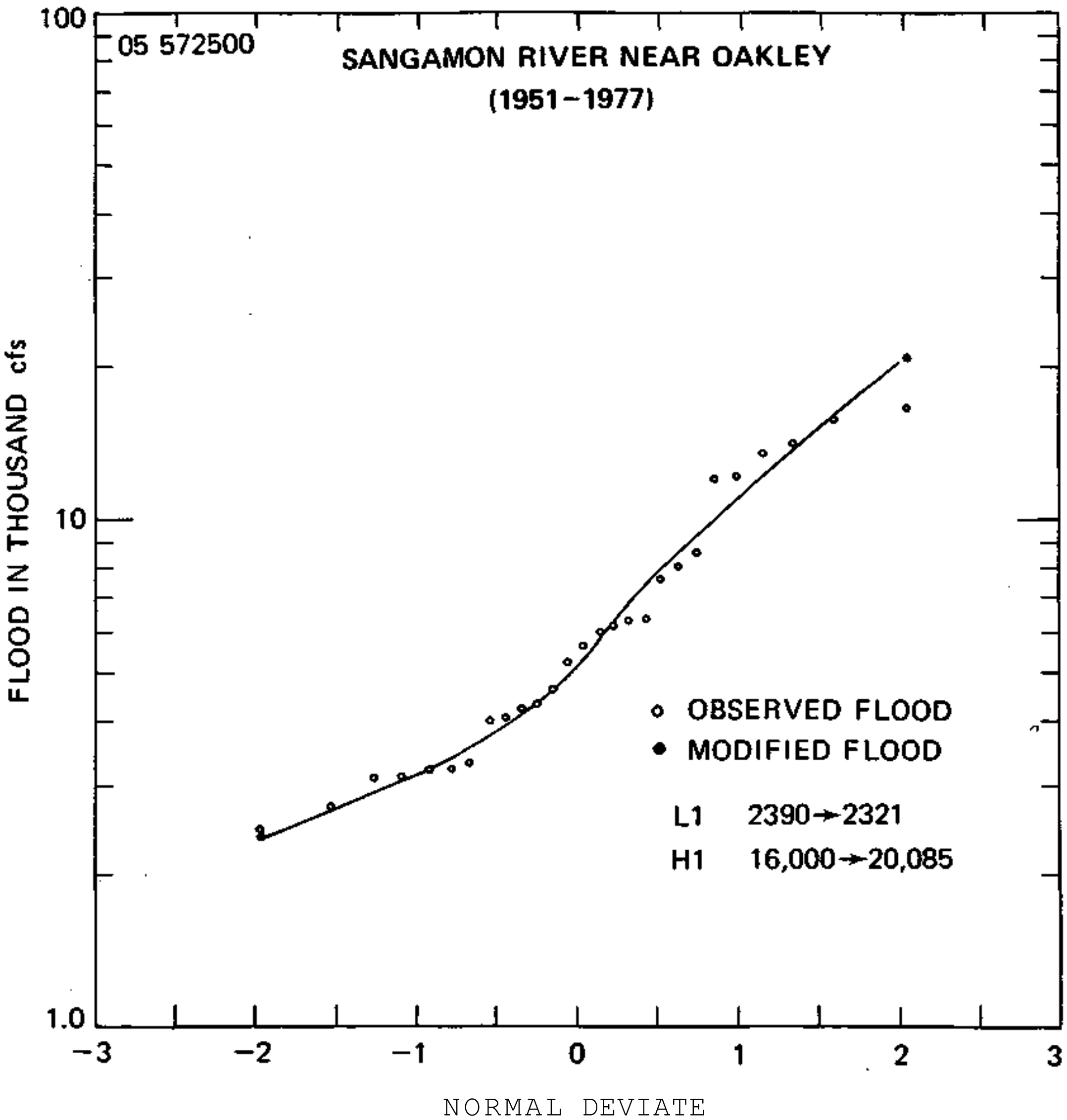


Figure 23. Observed and modified floods and the fitted mixed distribution curve (window 5) for the Sangamon River near Oakley

estimates are 21,888 and 32,290 and 24,557 and 38,116 cfs, respectively. The Q_{100}/Q_2 and Q_{1000}/Q_2 curves in figures 15 and 16 show that the regional estimate lies somewhere in between the MD values for windows 0 and 5, and that estimates by LP3 with sample skew and PT are much higher. The observed floods as well as the modified floods in the 5th window and the probability curve fitted by the MD method are shown in figure 23.

6. Skillet Fork near Wayne City: High Outlier and High Inliers

The results obtained with the computer program are given in Table 23. The NO equals [51/5] or 5. Only one high outlier is indicated but there are 4 high inliers as shown in figure 24. Two rather insignificant low inliers are also detected. The high outlier and inliers are modified as shown below:

	H1	H2	H3	H4	H5
Window 0	51,000	22,800	20,000	18,500	18,000
Window 5'	37,862	26,139	23,188	21,157	19,593

The LP3 statistics show a minor change in skew, from -0.394 to -0.467 and in standard deviation from 0.334 to 0.336. The 100- and 1000-year floods change from 36,719 and 54,194 cfs in window 0 to 35,593 and 50,946 cfs in window 5. The change is rather small.

The PT statistics show that kurtosis decreases from 3.495 in window 0 to 2.884 in window 5 and the 5th moment decreases from 1.317 to -0.056. The corresponding 100- and 1000-year flood estimates with PT and sample kurtosis change from 40,644 and 69,165 cfs to 35,230 and 51,250 cfs. The MD estimates change from 39,943 and 66,909 cfs to 37,723 and 71,302 cfs.

Table 23. Flood Frequency Analyses: Skillet Fork near Wayne City

STATION NO. 3380500 SKILLET FORK NEAR WAYNE CITY
DRAINAGE AREA 464.0 Sq Mi Years of Record 51 (1929-1979)

	LEVEL NO.	0	1	2	3	4	5	6
METHOD				100-Year Flood in cfs				
Power Transform, PT								
	With kt = 3.0	37,886	37,886	38,215	38,322	36,859	35,883	35,682
	With sample kt	40,644	40,644	40,896	40,468	37,167	35,230	34,286
Log Transform								
	LP3, Sample skew	36,719	36,719	37,068	37,322	36,275	35,593	35,852
	LP3, Weighted skew	36,642	36,642	36,897	37,165	36,782	36,498	36,454
	Mixed Distrib., MD	39,943	39,943	40,276	40,421	40,856	37,723.	36,132
Type	No.			Observed and Modified Floods in cfs				
Low	1*	858	858	858	858	858	858	927
	2*	1,450	1,450	1,450	1,450	1,450	1,450	1,471
	3*	2,110	2,110	2,110	2,110	2,110	2,110	2,071
	4*	2,860	2,860	2,860	2,842	2,654	2,518	2,415
	5*	3,040	3,040	3,040	3,040	2,955	2,828	2,727
High	5*	18,000	18,000	18,000	18,482	19,240	19,593	19,895
	4*	18,500	18,500	19,103	19,874	20,761	21,157	21,487
	3*	20,000	20,000	20,778	21,673	22,760	23,188	23,579
	2*	22,800	22,800	23,022	24,162	25,597	26,139	26,641
	1*	51,000	51,000	51,000	49,290	41,984	37,862	35,540
METHOD	STATISTICS			Values of Statistics				
PT	mean	19.883	19.883	19.560	20.012	24.431	28.125	28.459
	std dev	3.123	3.123	3.050	3.172	4.321	5.340	5.421
	skew	.007	.007	.005	.002	-.011	-.020	-.027
	kurtosis,kt	3.495	3.495	3.471	3.370	3.055	2.884	2.758
	5th moment	1.317	1.317	1.197	.906	.243	-.056	-.173
	lambda	.158	.158	.155	.159	.194	.218	.220
LP3	mean	3.886	3.886	3.887	3.888	3.888	3.887	3.887
	std dev	.334	.334	.334	.336	.336	.336	.335
	sample skew	-.394	-.394	-.388	-.389	-.437	-.467	-.444
	kurtosis,kt	3.599	3.599	3.583	3.526	3.375	3.301	3.135
	5th moment	-3.744	-3.744	-3.705	-3.814	-4.412	-4.679	-4.173
MD	weight 'a'	.108	.108	.098	.135	.206	.612	.598
	mu1	3.454	3.454	3.425	3.472	3.618	3.781	3.763
	mu2	3.939	3.939	3.937	3.953	3.958	4.053	4.070
	sigma1	.357	.357	.349	.337	.367	.364	.350
	sigma2	.290	.290	.292	.287	.288	.190	.202
	Test Stat	6.254	6.254	6.121	5.830	5.243	4.708	4.249

* High & low floods considered for outlier detection and modification

Table 23. Concluded

STATION NO. 3380500 SKILLET FORK NEAR WAYNE CITY
DRAINAGE AREA 464.0 Sq Mi Years of Record 51 (1929-1979)

VARIOUS RECURRENCE-INTERVAL FLOODS

METHOD	#	Flood in cfs for Recurrence Intervals (Years) 2	10	25	50	100	500	1000
PT, kt=3.0	0	8,054	19,797	26,715	32,187	37,886	52,054	58,572
PT, sample kt		8,054	19,422	27,065	33,527	40,644	59,739	69,165
LP3, sample skew		8,098	19,837	26,478	31,568	36,719	48,884	54,194
weighted skew		8,101	19,829	26,450	31,519	36,642	48,723	53,987
MD, mixed dist.		8,007	19,623	27,031	33,219	39,943	57,990	66,909
PT, kt=3.0	1							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	2	8,062	19,892	26,888	32,432	38,215	52,622	59,262
PT, sample kt		8,062	19,531	27,232	33,734	40,896	60,092	69,563
LP3, sample skew		8,104	19,931	26,656	31,825	37,068	49,496	54,939
weighted skew		8,113	19,914	26,594	31,715	36,897	49,136	54,477
MD, mixed dist.		8,025	19,716	27,202	33,450	40,276	58,578	67,630
PT, kt=3.0	3	8,092	19,978	26,991	32,540	38,322	52,701	59,318
PT, sample kt		8,092	19,695	27,272	33,594	40,468	58,604	67,396
LP3, sample skew		8,125	20,037	26,818	32,032	37,322	49,864	55,358
weighted skew		8,133	20,021	26,761	31,931	37,165	49,534	54,934
MD, mixed dist.		8,080	19,898	27,401	33,628	40,421	58,549	67,476
PT, kt=3.0	4	8,161	19,763	26,391	31,551	36,859	49,810	55,669
PT, sample kt		8,161	19,721	26,439	31,714	37,167	50,611	56,733
LP3, sample skew		8,165	19,908	26,417	31,344	36,275	47,720	52,630
weighted skew		8,140	19,963	26,606	31,674	36,782	48,763	53,957
MD, mixed dist.		8,028	19,939	27,555	33,929	40,856	59,518	68,773
PT, kt=3.0	5	8,196	19,590	25,963	30,873	35,883	47,960	53,365
PT, sample kt		8,196	19,679	25,861	30,529	35,230	46,354	51,250
LP3, sample skew		8,176	19,799	26,135	30,881	35,593	46,379	50,946
weighted skew		8,131	19,899	26,474	31,474	36,498	48,231	53,296
MD, mixed dist.		8,457	18,873	25,197	30,881	37,723	59,361	71,302
PT, kt=3.0	6	8,197	19,531	25,856	30,721	35,682	47,625	52,964
PT, sample kt		8,197	19,710	25,623	29,977	34,286	44,245	48,531
LP3, sample skew		8,151	19,782	26,191	31,025	35,852	47,008	51,776
weighted skew		8,121	19,848	26,41.6	31,418	36,454	48,244	53,346
MD, mixed dist.		8,407	19,150	25,281	30,392	36,132	53,146	62,528

= level number

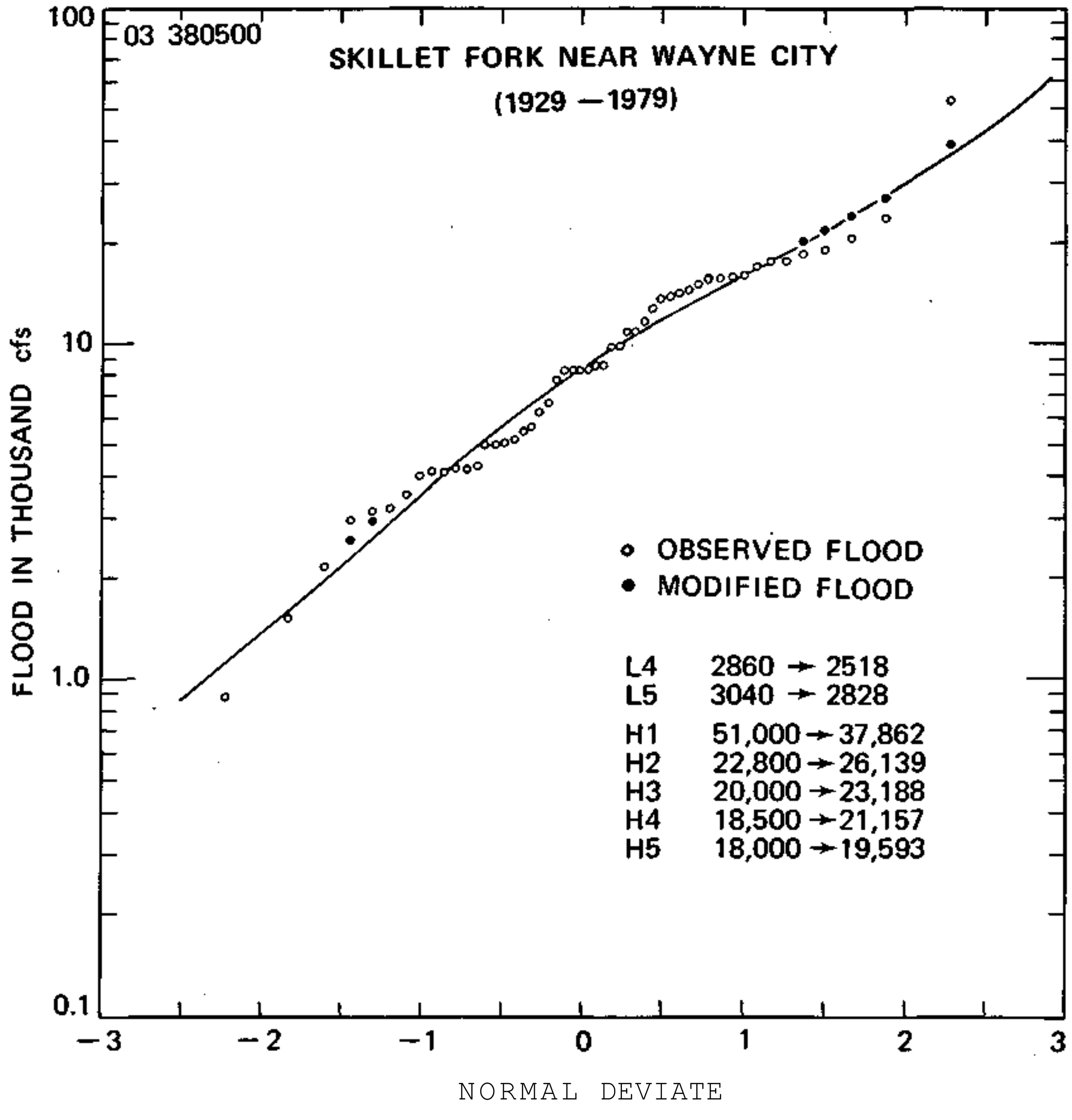


Figure 24. Observed and modified floods and the fitted mixed distribution curve (window 5) for the Skillet Fork near Wayne City

For this basin, the effect of a rather high outlier is largely balanced by 4 high inliers. The 100-year flood estimates with different methods are very close but the 1000-year flood with the MD is about 1.3 to 1.4 times that from the others. The top flood of 51,000 cfs was caused by a 2-3 day storm producing about 10 inches of catchment rainfall; the estimated recurrence interval is 300 to 500 years. The MD gives a 500-year flood of 59,361 cfs and the observed top flood of 51,000 cfs would correspond to somewhat higher than a 300-year flood.

7. Kishwaukee River near Perryville: Low Inlier and High Inlier

The results obtained with the computer program are given in Table 24. The NO equals [40/4] = 4. One significant and one insignificant low inliers and one insignificant low outlier, and one significant and three less significant high inliers are shown in figure 25.

The LP3 statistics show a minor change in skew, from -0.541 to -0.601, and in standard deviation, from 0.282 to 0.301. The 100- and 1000-year floods are 24,980 and 32,832 cfs in window 0 and 26,412 and 34,545 cfs in window 5, with the sample skew. Modification of low inliers generally reduces the skew and of high inliers increases the skew. When both low and high inliers are present, the opposite effects are cancelled to some extent.

The PT statistics show that kurtosis increases from 1.912 in window 0 to 2.254 in window 5 and the absolute value of the 5th moment decreases from 0.685 to 0.373. A summary of 100- and 1000-year floods with different methods is given on the next page.

Table 24. Flood Frequency Analyses: Kishwaukee River near Perryville

STATION NO. 5440000 KISHWAUKEE RIVER NEAR PERRYVILLE
DRAINAGE AREA 1099.0 Sq Mi Years of Record 40 (1940-1979)

LEVEL NO.		0	1	2	3	4	5	6
METHOD		100-Year Flood in cfs						
Power Transform, PT								
With kt = 3.0		22,164	22,164	22,409	22,593	23,519	24,682	25,953
With sample kt		18,562	18,562	18,864	19,261	20,577	22,140	23,939
Log Transform								
LP3, Sample skew		24,980	24,980	25,179	25,232	25,741	26,412	27,034
LP3, Weighted skew		26,385	26,385	26,515	26,822	27,682	28,698	29,960
Mixed Distrib., MD		18,665	18,665	18,663	18,930	20,605	22,608	25,516
Type	No.	Observed and Modified Floods in cfs						
Low	1*	2,020	2,020	2,020	1,789	1,483	1,281	1,096
	2*	2,080	2,080	2,080	2,080	2,080	1,980	1,787
	3*	2,340	2,340	2,340	2,340	2,340	2,340	2,299
	4*	2,360	2,360	2,360	2,360	2,403	2,505	2,582
	5	2,620						
High	5	14,800						
	4*	14,800	14,800	14,800	14,800	14,958	15,702	16,492
	3*	15,200	15,200	15,200	15,200	16,000	16,884	17,842
	2*	16,400	16,400	16,400	16,400	17,393	18,514	19,732
	1*	16,700	16,700	17,449	18,199	19,690	21,285	23,078
METHOD	STATISTICS	Values of Statistics						
PT	mean	224.414	224.414	201.455	198.579	154.926	118.079	95.519
	std dev	71.137	71.137	62.554	62.003	46.616	33.925	26.616
	skew	-.158	-.158	-.152	-.148	-.130	-.108	-.088
	kurtosis,kt	1.912	1.912	1.936	1.983	2.110	2.254	2.426
	5th moment	-.685	-.685	-.633	-.603	-.491	-.373	-.268
	lambda	.534	.534	.519	.517	.482	.443	.412
LP3	mean	3.855	3.855	3.855	3.854	3.855	3.856	3.856
	std dev	.282	.282	.283	.286	.293	.301	.311
	sample skew	-.541	-.541	-.533	-.556	-.580	-.601	-.641
	kurtosis,kt	2.244	2.244	2.246	2.330	2.517	2.726	2.999
	5th moment	-2.744	-2.744	-2.699	-2.959	-3.496	-4.093	-4.940
MD	weight 'a'	.524	.524	.543	.563	.578	.609	.619
	mu1	3.638	3.638	3.649	3.661	3.676	3.702	3.720
	mu2	4.094	4.094	4.100	4.103	4.099	4.094	4.077
	sigma1	.215	.215	.220	.234	.258	.285	.316
	sigma2	.086	.086	.082	.082	.098	.108	.115
	Test Stat	3.651	3.651	3.393	2.847	2.102	2.027	2.381

* High & low floods considered for outlier detection and modification

Table 24. Concluded

STATION NO. 5440000 KISHWAUKEE RIVER NEAR PERRYVILLE
DRAINAGE AREA 1099.0 Sq Mi Years of Record 40 (1940-1979)

VARIOUS RECURRENCE-INTERVAL FLOODS

METHOD	#	Flood in cfs for Recurrence Intervals (Years) 2	10	25	50	100	500	1000
PT, kt=3.0	0	7,930	14,948	18,025	20,154	22,164	26,504	28,268
PT, sample kt		7,930	15,454	17,108	17,934	18,562	19,616	19,962
LP3, sample skew		7,587	15,705	19,608	22,364	24,980	30,593	32,832
weighted skew		7,496	15,882	20,171	23,312	26,385	33,301	36,194
MD, mixed dist.		8,723	14,623	16,370	17,551	18,665	21,149	22,230
PT, kt=3.0	1							
PT, sample kt								
LP3, sample skew					SAME AS ABOVE			
weighted skew								
MD, mixed dist.								
PT, kt=3.0	2	7,921	15,027	18,167	20,347	22,409	26,877	28,697
PT, sample kt		7,921	15,523	17,271	18,168	18,864	20,049	20,446
LP3, sample skew		7,590	15,764	19,716	22,515	25,179	30,916	33,214
weighted skew		7,503	15,931	20,250	23,416	26,515	33,498	36,421
MD, mixed dist.		8,582	14,700	16,411	17,566	18,663	21,159	22,296
PT, kt=3.0	3	7,920	15,105	18,288	20,499	22,593	27,130	28,980
PT, sample kt		7,920	15,574	17,458	18,464	19,261	20,658	21,140
LP3, sample skew		7,598	15,845	19,802	22,591	25,232	30,874	33,115
weighted skew		7,496	16,046	20,440	23,664	26,822	33,939	36,918
MD, mixed dist.		8,520	14,781	16,545	17,755	18,930	21,825	23,359
PT, kt=3.0	4	7,911	15,445	18,853	21,243	23,519	28,494	30,538
PT, sample kt		7,911	15,861	18,152	19,473	20,577	22,617	23,351
LP3, sample skew		7,637	16,135	20,199	23,051	25,741	31,447	33,695
weighted skew		7,515	16,380	20,977	24,361	27,682	35,186	38,334
MD, mixed dist.		8,468	15,175	17,408	18,997	20,605	24,966	27,499
PT, kt=3.0	5	7,906	15,862	19,554	22,171	24,682	30,228	32,527
PT, sample kt		7,906	16,223	18,987	20,676	22,140	24,987	26,051
LP3, sample skew		7,685	16,486	20,695	23,641	26,412	32,256	34,545
weighted skew		7,545	16,774	21,610	25,182	28,698	36,667	40,018
MD, mixed dist.		8,428	15,498	18,199	20,286	22,608	30,303	34,842
PT, kt=3.0	6	7,903	16,321	20,323	23,186	25,953	32,125	34,705
PT, sample kt		7,903	16,617	19,900	22,025	23,939	27,833	29,340
LP3, sample skew		7,739	16,880	21,214	24,224	27,034	32,886	35,146
weighted skew		7,566	17,253	22,390	26,200	29,960	38,509	42,111
MD, mixed dist.		8,488	15,644	18,914	21,803	25,516	38,081	44,734

= level number

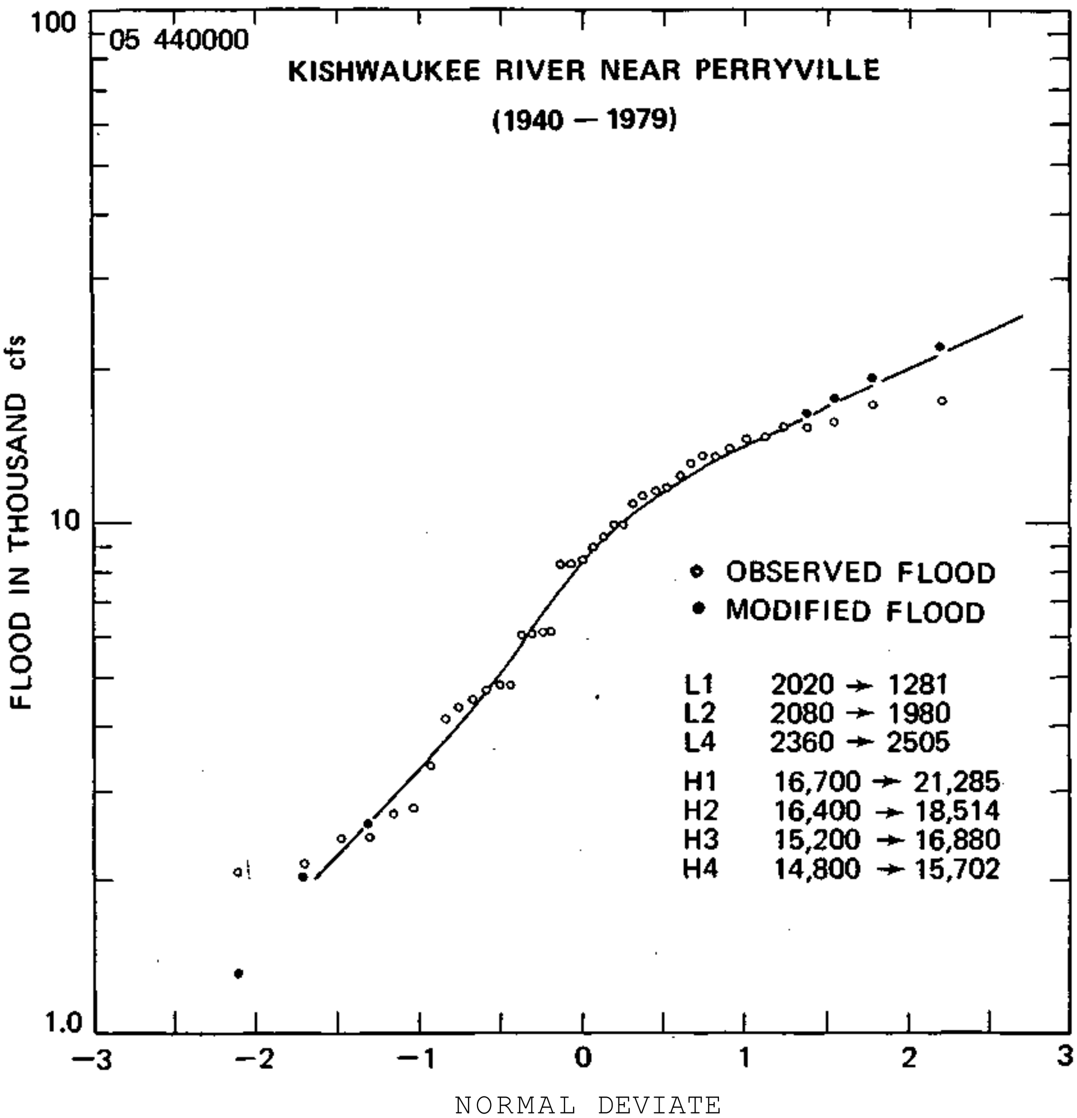


Figure 25. Observed and modified floods and the fitted mixed distribution curve (window 5) for the Kishwaukee River near Perryville

	100-year flood window 0	100-year flood window 5	1000-year flood window 0	1000-year flood window 5
PT, kt = 3.0	22,164	24,682	28,268	32,527
PT, sample kt	18,562	22,140	19,962	26,051
LP3, sample skew	24,980	26,412	32,832	34,545
LP3, weighted skew	26,385	28,698	36,194	40,018
MD	18,665	22,608	22,230	34,842

When both low and high inliers are present, the flood estimates are less sensitive to the modification of inliers for LP3 than with the MD.

8. Sangamon River at Riverton: Low Outliers and High Outliers

The results obtained with the computer program are given in Table 25. The NO equals [67/10] or a maximum of 5. Four low outliers and three high outliers (out of which H1 is a very significant high outlier) are shown in Figure 26. The modified values for these outliers in window 5 are also shown in the figure.

The LP3 statistics show that the skew decreases from -1.227 in window 0 to -1.386 in window 5 and the standard deviation decreases from 0.312 to 0.291. With sample skew, the 100- and 1000-year floods of 38,917 and 42,416 cfs in window 0 are replaced by 33,931 and 35,780 cfs, which are much lower than the observed flood of 68,700 cfs.

The PT statistics indicate that kurtosis and 5th moment decrease from 4.538 and 3.513 in window 0 to 3.231 and 0.141 in window 5. The 100- and 1000-year floods change from 54,386 and 84,445 cfs in window 0 to 41,641 and 53,410 cfs in window 5, with sample kurtosis. However, the corresponding MD estimates change from 53,173 and 123,725 cfs in window 0 to 44,041 and 63,018 cfs in window 5. Thus, only the 5th window 1000-year flood with

Table 25. Flood Frequency Analyses: Sangamon River at Riverton

STATION NO. 5576500 SANGAMON RIVER AT RIVERTON
DRAINAGE AREA 2618.0 Sq Mi Years of Record 67 (1908-1979)

LEVEL NO.		0	1	2	3	4	5	6
METHOD		100-Year Flood in cfs						
Power Transform, PT								
With kt = 3.0		48,573	48,573	46,552	45,177	42,341	40,858	39,667
With sample kt		54,386	54,386	50,740	48,344	43,952	41,641	39,735
Log Transform								
LP3, Sample skew		38,917	38,917	37,621	36,719	34,823	33,931	34,489
LP3, Weighted skew		46,750	46,750	45,577	44,685	42,686	41,370	40,746
Mixed Distrib., MD		53,173	53,173	49,872	47,912	45,146	44,041	42,643
Type	No.	Observed and Modified Floods in cfs						
Low	1*	1,040	1,040	1,040	1,040	1,040	1,040	1,280
	2*	1,830	1,830	1,830	1,830	1,860	2,177	2,668
	3*	2,540	2,540	2,540	2,540	2,778	3,151	3,640
	4*	2,840	2,840	2,902	3,141	3,523	3,931	4,407
	5*	4,260	4,260	4,260	4,260	4,260	4,599	5,060
High	5*	30,600	30,600	30,600	30,654	30,984	30,984	30,372
	4*	32,900	32,900	32,900	32,900	32,900	32,900	31,786
	3*	41,000	41,000	41,000	40,501	37,018	35,097	33,594
	2*	44,200	44,200	44,200	44,200	40,418	37,998	36,154
	1*	68,700	68,700	60,108	54,410	47,205	43,513	40,834
METHOD	STATISTICS	Values of Statistics						
PT	mean	122.449	122.449	171.779	217.877	373.890	475.065	518.865
	std dev	32.354	32.354	48.749	64.479	119.821	154.829	166.183
	skew	.088	.088	.063	.042	.007	-.016	-.034
	kurtosis,kt	4.538	4.538	4.137	3.873	3.468	3.231	3.022
	5th moment	3.513	3.513	2.265	1.474	.631	.141	-.123.
	lambda	.408	.408	.453	.484	.553	.583	.594
LP3	mean	4.144	4.144	4.143	4.143	4.142	4.144	4.148
	std dev	.312	.312	.310	.307	.300	.291	.276
	sample skew	-1.227	-1.227	-1.280	-1.313	-1.381	-1.386	-1.254
	kurtosis,kt	5.791	5.791	5.802	5.841	5.968	6.069	5.498
	5th moment	-15.139	-15.139	-15.893	-16.532	-17.963	-19.160	-16.481
MD	weight 'a'	.281	.281	.264	.242	.163	.159	.166
	mu1	3.904	3.904	3.877	3.847	3.704	3.719	3.753
	mu2	4.238	4.238	4.239	4.237	4.228	4.225	4.227
	sigma1	.442	.442	.433	.426	.384	.379	.338
	sigma2	.166	.166	.168	.173	.185	.182	.176
	Test Stat	4.054	4.054	4.062	3.972	3.569	3.674	4.127

* High & low floods considered for outlier detection and modification

Table 25. Concluded

STATION NO. 5576500 SANGAMON RIVER AT RIVERTON
DRAINAGE AREA 2618.0 Sq Mi Years of Record 67 (1908-1979)

VARIOUS RECURRENCE-INTERVAL FLOODS

METHOD	#	Flood in cfs for Recurrence Intervals (Years) 2	10	25	50	100	500	1000
PT, kt=3.0	0	15,289	30,867	38,233	43,494	48,573	59,888	64,613
PT, sample kt		15,289	29,673	38,860	46,378	54,386	74,834	84,445
LP3, sample skew		16,070	30,244	34,683	37,090	38,917	41,665	42,416
weighted skew		15,424	31,891	38,665	42,988	46,750	53,758	56,174
MD, mixed dist.		15,757	28,378	35,849	43,004	53,173	97,011	123,725
PT, kt=3.0	1							
PT, sample kt								
LP3, sample skew				SAME AS ABOVE				
weighted skew								
MD, mixed dist.								
PT, kt=3.0	2	15,374	30,264	37,105	41,932	46,552	56,718	60,918
PT, sample kt		15,374	29,405	37,632	44,081	50,740	67,003	74,340
LP3, sample skew		16,119	29,784	33,867	36,022	37,621	39,941	40,551
weighted skew		15,439	31,529	37,997	42,070	45,577	51,997	54,171
MD, mixed dist.		15,757	28,300	35,391	41,698	49,872	85,042	107,980
PT, kt=3.0	3	15,429	29,823	36,311	40,854	45,177	54,615	58,489
PT, sample kt		15,429	29,182	36,746	42,513	48,344	62,162	68,247
LP3, sample skew		16,150	29,429	33,273	35,264	36,719	38,781	39,309
weighted skew		15,452	31,224	37,462	41,356	44,685	50,713	52,731
MD, mixed dist.		15,715	28,304	35,194	41,009	47,912	74,905	93,932
PT, kt=3.0	4	15,520	28,842	34,617	38,597	42,341	50,390	53,649
PT, sample kt		15,520	28,516	34,876	39,467	43,952	54,017	58,260
LP3, sample skew		16,189	28,604	31,963	33,638	34,823	36,424	36,812
weighted skew		15,460	30,479	36,213	39,727	42,686	47,927	49,642
MD, mixed dist.		15,540	28,240	34,844	39,911	45,146	58,361	64,829
PT, kt=3.0	5	15,563	28,268	33,682	37,388	40,858	48,271	51,256
PT, sample kt		15,563	28,114	33,816	37,819	41,641	49,985	53,410
LP3, sample skew		16,187	28,077	31,249	32,822	33,931	35,421	35,780
weighted skew		15,477	29,872	35,292	38,596	41,370	46,260	47,854
MD, mixed dist.		15,515	27,798	34,155	39,022	44,041	56,752	63,018
PT, kt=3.0	6	15,572	27,739	32,881	36,389	39,667	46,648	49,452
PT, sample kt		15,572	27,728	32,897	36,432	39,735	46,800	49,641
LP3, sample skew		15,990	27,804	31,278	33,117	34,489	36,504	37,042
weighted skew		15,403	29,196	34,548	37,886	40,746	45,961	47,726
MD, mixed dist.		15,525	27,384	33,423	37,995	42,643	54,073	59,419

= level number

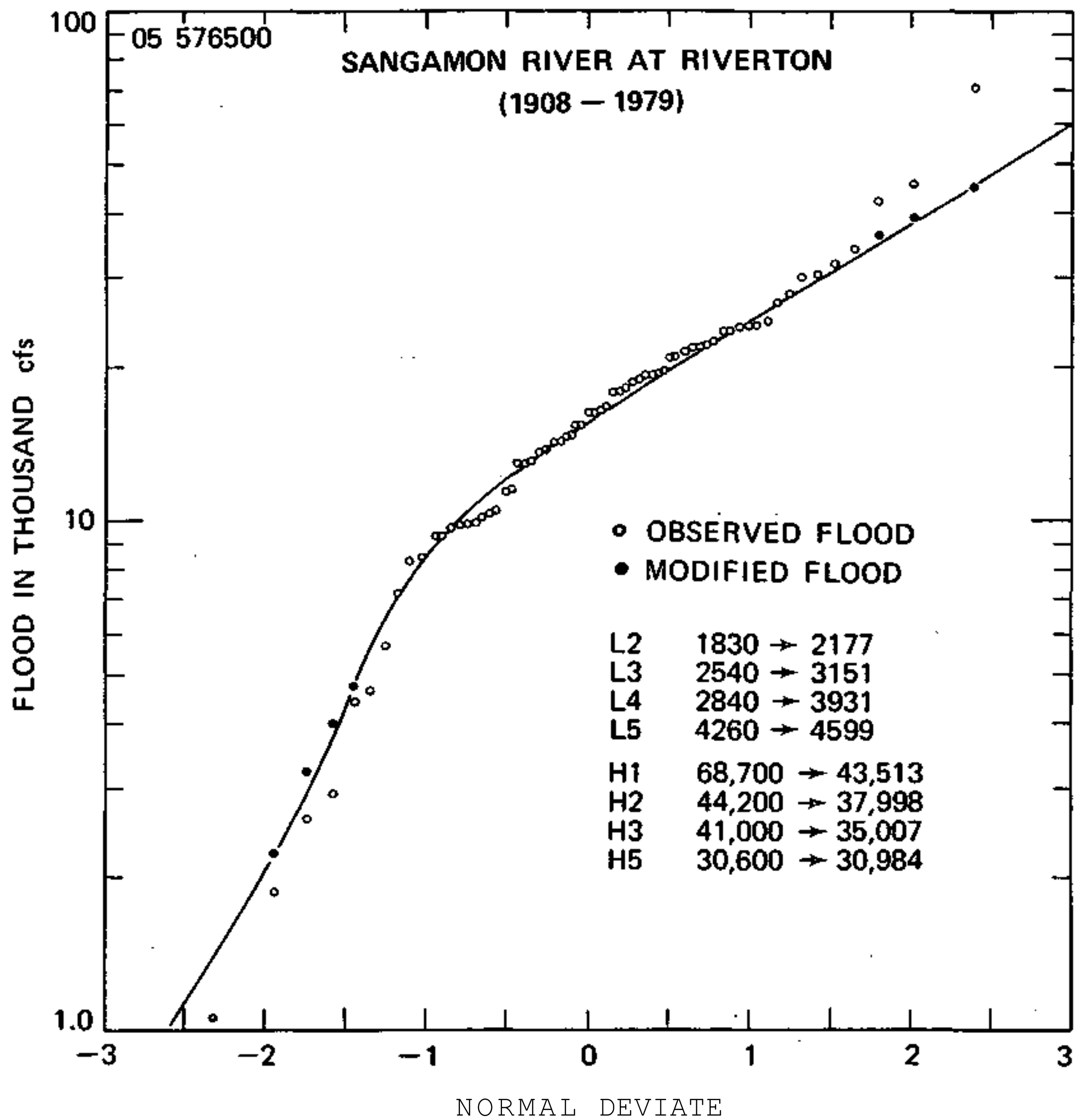


Figure 26. Observed and modified floods and the fitted mixed distribution curve (window 5) for the Sangamon River at Riverton

MD is close to the observed top flooded 68,700 cfs. The other methods yield estimates varying from 35 to 53 thousand cfs.

When the sample skew is very small in the algebraic sense, the LP3 flood estimates of 500, 1000, or higher recurrence-interval floods are not much higher than the 100-year flood. The Sangamon River at Riverton has flood data for 67 years. The MD estimates are considered better than those from the other four methods. However, the MD flood estimate in window 0 for high recurrence-interval floods can be very high. The MD does give good results after the outliers/inliers have been modified.

CONCLUSIONS

The main objectives of this study were: 1) the development of satisfactory tests for detecting outliers and inliers at various levels of significance in the two extreme tails of a suitably transformed flood series; 2) the extensive testing of available transformations in converting a number of observed flood series to series distributed approximately as $N(\mu, \sigma^2)$ and to determine the best transformation for general use; 3) the development and computerization of a flood-frequency methodology that not only detects and modifies outliers/inliers at different levels but also computes 2-year to 1000-year floods at those levels with the power transformation, log-Pearson type III, and mixed distribution methods; and 4) the overall conceptualization, theoretical basis, testing, and validation of a versatile and accurate new flood frequency method. These objectives have been met satisfactorily by the research, analyses, and comparative studies contained in this report. Some main conclusions, derived from this study, are given below.

1. An extensive testing of four methods or algorithms, for generating normally distributed random numbers, regarding their suitability, stability, and effectiveness in generating such numbers has indicated the Polar Method by Box, Muller, and Marsaglia to be the best.

2. Departure has been defined as the standard normal deviate corresponding to the plotting position of the high or low point of the series under consideration, minus the sample standard deviate for that point. The higher the absolute value of the departure, the more severe is the outlier/inlier. The distribution of the departures for up to 5 points on both the

high and the low end of various sample sizes has been determined from thousands of generated series. Both an extensive and a compact departure table have been developed for general testing of outliers at 0.01, 0.05, 0.10, 0.20, 0.30, and 0.40 levels of significance. Departures for only 0.01 and 0.05 levels are available in the literature for the top outlier and these are within 0.01 of the departures developed in this study. However, the statistical tests for inliers at any significance level and for outliers at 0.10 to 0.40 levels of significance and for up to five outliers/inliers are not available in the literature at the present time. The developed departures allow a step-by-step detection and modification of outliers/inliers at various levels.

3. Generally, the literature has dealt with outliers — a flood significantly higher than that indicated by the trend of the rest of the data at the high end, or a flood significantly lower than that indicated by the rest of the data at the low end. The introduction and designation of inliers — a flood lower than that indicated by the rest of the data at the high end or higher than that indicated at the low end — in this study is a welcome addition and fills the information gap. Statistically, both outliers and inliers can occur. However, the absolute value of departure for an inlier is generally less than that for an outlier because the inlier cannot be less than the next lower flood in ranked series at the high end or more than the next higher flood at the low end.

4. Transformation of an observed flood series to an approximately normally distributed series is necessary for checking any outliers/inliers with statistical tests developed in this study. Three transformations — power, Wilson-Hilferty, and 3-parameter lognormal — were tested on 28 flood

series. The results indicate that the power transformation is superior to the others in terms of yielding consistent and satisfactory statistical parameters for the transformed series. Values of g in Table 26 for the power transformed series are very close to zero and those for the 5th are considerably lower than the values with log transformation only (e.g., for LP3) .

5. Flood frequency methods have been put together in a computer program. These methods include power transformation method with kurtosis equal to 3.0 as for a normal distribution as well as with sample kurtosis, log-Pearson type III method with the sample skew as well as the weighted skew, and the mixed distribution. The kurtosis correction with the power transformation method is satisfactory if the transformed series approximates a symmetrical distribution. The relevant distribution statistics and measures of goodness of fit with the observed flood series and with the series after modification of outliers/inliers at various levels, as well as the 2-year to 1000-year floods at various levels, are presented in a tabular format. The output enables the analyst to follow the detection and modification of outliers/inliers at various levels and to choose the level he thinks is the best to use for a particular basin.

6. Results of flood frequency analyses of 37 observed flood series in Illinois indicate the following:

a) Absolute value of skew, g, with the power transformation (Table 26) is <0.05, 0.05 to 0.10, and 0.10 to 0.20 for 21, 8, and 8 basins in window 0 and for 19, 12, and 6 basins in window 5, respectively. The power transformation reduces the skew close to zero.

b) Kurtosis with the power transformation (Table 26) is < 3 for 26 basins and >3 for 11 basins in window 0, and <3 for 28 basins and >3 for

Table 26. Values of g, kt, and 5th with Power and Log Transformation

USGS No.	Trans	Window 0			Window 5		
		g	kt	5th	g	kt	5th
SANGAMON RIVER BASIN							
05 571000	Power	-0.008	2.601	-0.248	-0.007	2.597	-0.262
	Log	-0.073	2.621	-0.760	-0.070	2.617	-0.751
05 572000	Power	-0.016	2.726	-0.430	-0.019	2.716	-0.412
	Log	-0.312	2.991	-3.246	-0.336	3.003	-3.386
05 572500	Power	0.093	2.202	0.322	0.097	2.311	0.403
	Log	0.398	2.309	1.997	0.486	2.551	2.934
05 574000	Power	0.116	2.433	0.765	0.111	2.444	0.716
	Log	0.788	3.420	6.815	0.771	3.395	6.692
05 574500	Power	-0.016	3.256	-0.558	-0.022	3.154	-0.513
	Log	-0.803	4.302	-8.841	-0.717	4.046	-7.872
05 575500	Power	-0.002	2.939	-0.064	-0.004	2.936	-0.071
	Log	-0.051	2.944	-0.464	-0.054	2.942	-0.486
05 576000	Power	-0.053	2.633	-0.014	-0.056	2.620	0.027
	Log	-0.453	2.840	-3.205	-0.463	2.818	-3.192
05 576500	Power	0.088	4.538	3.513	-0.016	3.231	0.1*41
	Log	-1.227	5.791	-15.139	-1.386	6.069	-19.160
05 577500	Power	-0.032	2.605	-1.034	-0.023	2.575	-0.768
	Log	-0.325	2.959	-3.463	-0.203	2.736	-2.200
05 578500	Power	0.020	2.207	0.451	0.018	2.267	0.274
	Log	0.104	2.263	1.101	0.110	2.308	0.955
05 579500	Power	0.036	4.045	-1.494	0.005	3.162	-0.593
	Log	-0.730	5.438	-12.793	0.140	3.136	0.614
05 580000	Power	0.046	2.596	0.742	0.038	2.671	0.657
	Log	0.456	3.026	3.977	0.417	3.036	3.726
05 580500	Power	0.040	2.474	0.946	0.030	2.571	0.835
	Log	0.345	2.850	3.456	0.299	2.866	3.123
05 581500	Power	0.012	3.079	0.970	0.015	3.013	0.887
	Log	0.465	3.630	5.382	0.438	3.498	4.878

Table 26. Continued

USGS No.	Trans	Window 0 g	Window 0 kt	Window 0 5th	Window 5 g	Window 5 kt	Window 5 5th
05 582000	Power	-0.014	2.464	0.456	-0.016	2.484	0.391
	Log	-0.099	2.435	-0.146	-0.128	2.458	-0.411
05 582500	Power	-0.136	2.093	-0.629	-0.108	2.336	-0.425
	Log	-0.525	2.347	-2.687	-0.579	2.687	-3.634
05 583000	Power	0.018	3.820	3.900	-0.055	2.792	0.416
	Log	-0.562	3.421	3.027	-0.764	3.317	-5.653
ROCK RIVER BASIN							
05 435500	Power	-0.003	2.387	0.006	-0.005	2.508	-0.046
	Log	-0.024	2.387	-0.115	-0.047	2.512	-0.324
05 437000	Power	-0.077	2.175	-0.209	-0.071	2.354	-0.188
	Log	-0.345	2.241	-1.640	-0.410	2.504	-2.409
05 437500	Power	-0.051	2.190	-0.007	-0.054	2.310	-0.120
	Log	-0.258	2.211	-1.187	-0.329	2.414	-1.921
05 438250	Power	0.004	4.245	1.947	-0.064	3.676	0.252
	Log	-0.926	4.274	-7.214	-1.133	4.299	-9.109
05 438500	Power	-0.084	1.936	-0.436	-0.057	2.239	-0.234
	Log	-0.310	2.068	-1.593	-0.318	2.379	-2.007
05 439500	Power	-0.196	2.231	-0.558	-0.183	2.334	-0.591
	Log	-0.869	2.923	-5.292	-1.108	3.632	-8.388
05 440000	Power	-0.158	1.912	-0.685	-0.108	2.254	-0.373
	Log	-0.541	2.244	-2.744	-0.601	2.726	-4.093
05 440500	Power	-0.171	2.396	-0.660	-0.156	2.498	-0.527
	Log	-1.011	3.520	-7.537	-1.062	3.797	-8.691
05 441000	Power	-0.136	2.235	-0.686	-0.109	2.372	-0.414
	Log	-0.755	3.047	-5.650	-0.712	3.039	-5.334
05 443500	Power	-0.073	2.581	-0.039	-0.071	2.591	-0.071
	Log	-0.789	3.403	-6.455	-0.800	3.462	-6.709
05 444000	Power	-0.095	2.702	-0.567	-0.094	2.697	-0.538
	Log	-1.049	4.137	-9.658	-1.048	4.175	-9.924
05 445500	Power	0.002	3.572	0.525	-0.006	3.283	0.316
	Log	-0.210	3.573	-1.496	-0.348	3.380	-2.824
05 446500	Power	-0.116	1.945	-0.448	-0.087	2.205	-0.252
	Log	-0.421	2.167	-2.271	-0.453	2.472	-2.916

Table 26. Concluded

USGS No.	Trans	Window 0			Window 5		
		g	kt	5th	g	kt	5th
05 447000	Power	-0.115	2.783	-0.494	-0.092	2.752	-0.427
	Log	-1.147	3.975	-8.540	-1.106	3.904	-8.381
05 447500	Power	-0.031	3.117	0.036	-0.042	3.026	-0.065
	Log	-0.940	4.133	-8.700	-0.956	4.162	-9.123
05 448000	Power	-0.041	2.468	-0.048	-0.041	2.502	0.014
	Log	-0.317	2.544	-1.824	-0.336	2.580	-1.958
LITTLE WABASH RIVER BASIN							
03 379500	Power	-0.014	2.807	-0.647	-0.015	2.778	-0.472
	Log	-0.324	3.132	-3.715	-0.318	3.051	-3.329
03 380475	Power	-0.018	4.004	2.264	-0.013	3.105	1.048
	Log	0.729	5.446	12.918	-0.154	3.027	-0.225
03 380500	Power	0.007	3.495	1.317	-0.020	2.884	-0.056
	Log	-0.394	3.599	-3.744	-0.467	3.301	-4.679
03 381500	Power	0.007	3.975	-1.418	0.004	3.215	-0.377
	Log	-0.126	4.085	-2.896	0.306	3.272	2.286

9 basins in window 5. The values range from 1.912 to 2.939 and 3.079 to 4.538 in window 0 and from 2.205 to 2.936 and 3.013 to 3.676 in window 5. The kurtosis range decreases in window 5 becuase of the modification of any outliers and inliers.

c) Absolute value of the 5th with the power transformation (Table 26) are

	Number of basins with the |5th| in the range			
	<0.5	0.5-1.0	1.0-2.0	>2.0
Window 0	16	13	5	3
Window 5	26	11	0	0

The modification of outliers/inliers reduces significantly the absolute value of the 5th. The transformed series in window 5 are closer to normal distribution than are those in window 0.

d) The kurtosis correction with the power transformation method is reasonably valid if the 3rd and higher odd moments are close to zero. Though the values of g are close to zero for a majority of the transformed series, the 5th moment is not. Thus, the power transformed series are generally asymmetrical. The asymmetry is considered in the mixed distribution method.

e) The mixed distribution parameters a, μ_1, μ_2, σ_1, and σ_2 for the 37 study basins in Table 15 and window 5 show that $0.4 \leq a \leq 0.6$ for 12 basins, $|\sigma_1 - \sigma_2| \leq 0.05$ for 1 out of 12 basins with a varying from 0.4 to 0.6 and $|\mu_2 - \mu_1| \leq 0.25$ for none out of 12 basins. Thus, the conditions of a - 0.5 and $\mu_1 = \mu_2$ or $a = 0.5$ and $\sigma_1 = \sigma_2$ are not satisfied. The analysis of power transformed series with or without correction for kurtosis is not the best solution because of the apparent asymmetry exhibited by the transformed series. The mixed distribution is the better answer to the problem.

f) Plots of Q_{100}/Q_2 and Q_{1000}/Q_2 versus drainage area for the Sangamon and Rock River basins, with floods estimated from the mixed distribution and window 5, are well-defined and indicate a decrease in the ratio with increase in drainage area, except for areas less than 200 square miles. For smaller areas, the trend line steepens considerably. Corresponding data points with the LP3 and sample skew exhibit considerable scatter.

g) The flood estimates with the mixed distribution are generally found to be very satisfactory in window 5.

h) In the case of extreme high outliers, the storm statistics for the top 3 to 4 floods may be used in confirming the severity of the outlier with the methodology developed in a previous report (Singh, 1980).

i) The mixed distribution is highly versatile in simulating various observed distribution shapes. The method coupled with the detection and modification of outliers/inliers may perhaps be the best available at the present.

j) The regionalization of skew as recommended by the Water Resources Council and the use of LP3 may not be the best solution for the flood-frequency problem. The analyses presented in this report, together with the values of g in windows 0 and 5, do not suggest that regionalization of skew is worthwhile.

REFERENCES

Ang, A. H-S., and W. H. Tang. 1975. *Probability concepts in engineering, planning and design, Vol. I.* John Wiley and Sons, New York. 409 p.

Barnett, V.,and T. Lewis. 1978. *Outliers in statistical data.* John Wiley and Sons, New York. 365 p.

Blom, G. 1958. *Statistical estimates and trans formed beta variables.* John Wiley and Sons, New York. 176 p.

Box, G. E. P., and D. R. Cox. 1964. *An analysis of transformation.* Journal of Royal Statistical Society, Vol. B26. p. 211-252.

Box, G. E. P., and E. M. Muller. 1958. *A note on the generation of random normal deviates.* Annals of Mathematical Statistics, Vol. 29. p. 610-611.

Box, G. E. P., and G. C. Tiao. 1973. *Bayesian inference in statistical analysis.* Addison-Wesley, Reading, Massachusetts. 588 p. .

Cohen, A. C. 1967. *Estimation in mixtures of two normal distributions.* Technometrics, Vol. 9, No. 1. p. 15-28.

Cramer, H. 1946. *Mathematical methods of statistics.* Princeton University Press, Princeton, New Jersey. 575 p.

Cunnane, J. 1978. *Unbiased plotting position - a review.* Journal of Hydrology, Vol. 37. p. 205-222.

Harter, H. L. 1971. *Some optimization problems in parameter estimation* (in *Optimizing methods in statistics, edited by J. S. Rustagi).* Academic Press, New York. p. 33-62

Knuth, D. E. 1969. *The art of computer programming, Vol. 2.* Addison-Wesley, Reading, Massachusetts. p. 105-113.

Singh, K. P. 1968. *Hydrologic distributions resulting from mixed populations and their computer simulation.* International Association of Scientific Hydrology Tucson Symposium Publication. p. 336-344.

Singh, K. P., and R. A. Sinclair. 1972. *Two-distribution method for flood frequency analysis.* Journal of the Hydraulics Division, American Society of Civil Engineers, Vol. 98, No. HY 1. p. 29-44.

Singh, K. P. 1974.. *A two-distribution method for fitting mixed distributions in hydrology.* U.S. Department of Agriculture Miscellaneous Publication 1275. p. 371-382.

Singh, K. P. 1980. *Regional and sample skew values in flood-frequency analyses of streams in Illinois.* Illinois State Water Survey Contract Report No. 227. 169 p.

Singh, K. P. 1980. *Flood frequency analysis by power transformation — discussion.* Journal of the Hydraulics Division, American Society of Civil Engineers. Vol. 106, No. HY3. p. 462-465.

Water Resources Council, Hydrology Committee. 1967. *A uniform technique for determining flood flow frequencies, Bulletin 15.* Washington, D.C.

Water Resources Council, Hydrology Committee. 1976. *Guidelines for determining flood flow frequency, Bulletin 17.* Washington, D.C. r

Water Resources Council, Hydrology Committee. 1977. *Guidelines for determining flood flow frequency, Bulletin 17A.* Washington, D.C.

Wilson, E.B., and M. M. Hilferty. 1931. *Distribution of chi-squares.* Proceeding of Natural Academy of Sciences, Vol. 17, No. 12. p. 684-688.