IBM

International Business Machines Corporation

Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, New York 10598
914/945-3000

May 18, 1979

Joseph M. Hendrie
Director
Nuclear Regulatory Commission
Malomic Building
1717 H Street NW
Washington, D.C. 20555

Dear Mr. Hendrie:

While watching the Three Mile Island, 3MI, debacle on television I was struck with the appearance of the island. It appears that 3MI is a comparatively flat, low-lying island situated in the middle of a large river, and probably comprised of alluvial sands and gravels. The geomorphology literature attests to the fact that this is the sort of topography that gets redistributed during the really extreme, say 1/1,000 or 1/10,000 year flood events. This raises two obvious questions, first what happens to a nuclear power plant when it is submerged in flood waters, and second what is the flood risk at the 3MI site?

To evaluate the second of the above questions I have investigated the Susquehanna river data base, and asked a few questions here and there. There are 87 years of streamflow measurements for the Susquehanna river at Harrisburg, a fortunate circumstance as this is only a little upstream of the 3MI site. The highest instantaneous flow for each year can be obtained from the published records, ranked, and plotted on cumulative probability paper (see Figure 1). A log. Pearson III curve following the official U.S. Water Resources Council guidelines (Bulletin 17a) has also been shown on the figure. One might feel happier with this flood risk estimate (the fitted curve) if the procedures were not known to be totally fallacious.

The original design for the 3MI site estimated that "the maximum possible flood", MPF, for the site was 1.1 million cfs (cubic feet per second). In 1972 hurricane Agnes arrived and subsequently a new MPF of 1.6 million cfs. Will the next big event raise the MPF another 45%? Do you know the mechanics of how such estimates are made? From the name it sounds like the procedure developed in the thirties that has little to do with science, nothing to do with

7908150463

696 076

probability theory, and with uncertainties at every step of the calculation. The cumulative error is unknowable.

The use of MPF does not appear to be the sort of decision theoretic approach to estimating geophysical hazards that one might have expected of an NRC. It is to be hoped that by now the NRC has a sufficiently large competent in-house staff that thev do not have to rely upon the log. Pearson III or similarly technically unsound and hap-hazard flood risk appraisals that may be generated by the more tradition-bound Federal agencies.

The risk of flooding does seem to be rather high for the 3MI site. One can't help but wonder why 3MI was selected, when the use of Hill Island would have removed this consideration. It is understood that a levee has been constructed that is meant to maintain the integrity of the 3MI site from an event the size of the latest estimated MPF. All well and good, but levees have been known to fail, and MPF's are not decreed by God, but estimated by fallible man. Are the geophysical hazards at other nuclear plants of similar magnitude, and are risk evaluations, as at 3MI, always based upon a single number?

Sincerely,

James R Wallis

J. R. Wallis
President-elect
Section of Hydrology
American Geophysical Union

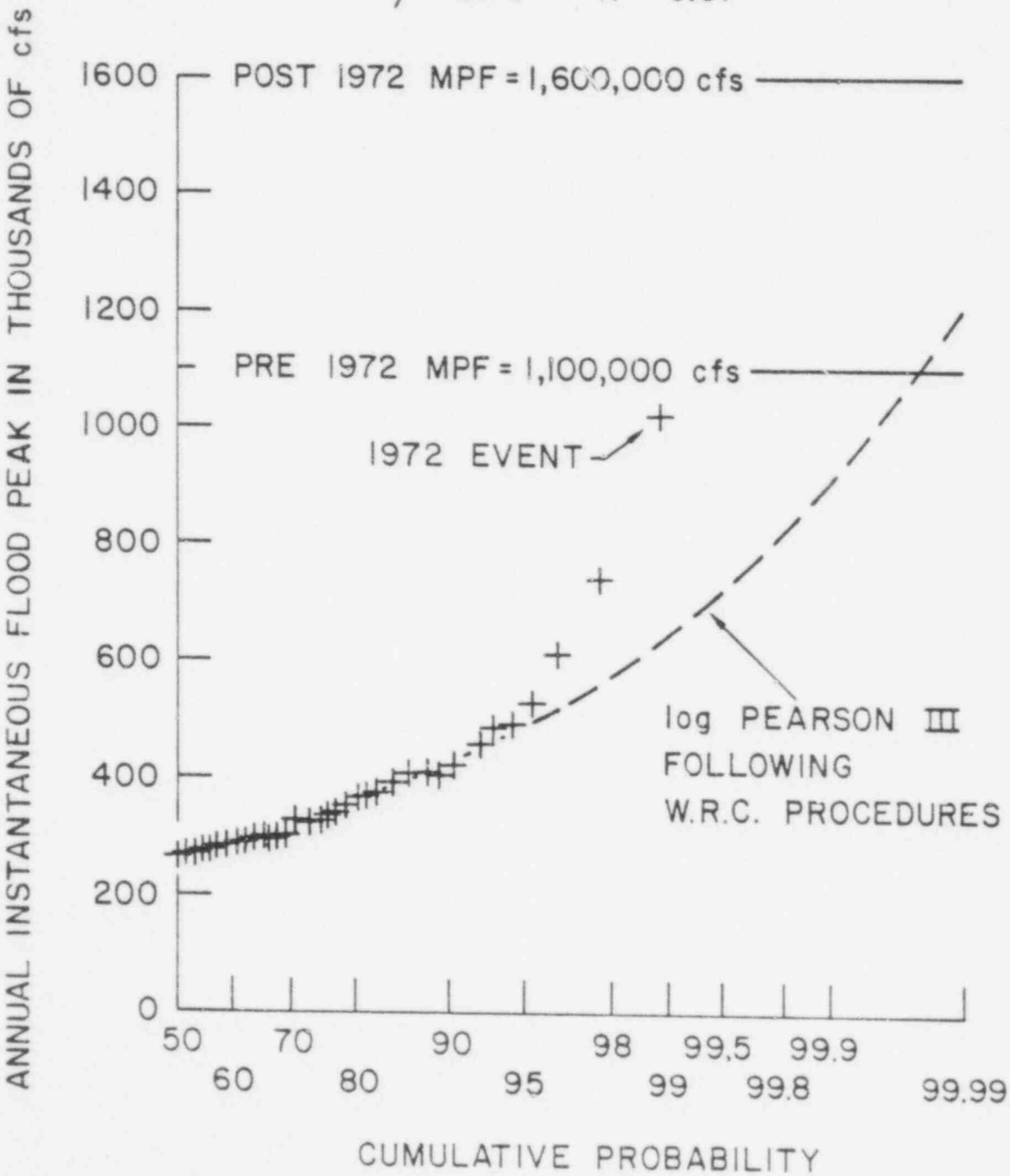
enc.

SUSQUEHANNA RIVER AT HARRISBURG, PA.

1891 - 1977

μ = 2341 σ = 1296

γ = 2.70 λ = 13.87



January, 1979

Dr Leo Eisel, Director
United States Water Resources Council
2120 L street, N.W.
Washington, D.C. 20037

Dear Leo,

It was great to see you and Vicki once again, and to be brought up to date on your respective careers and other doings. Shall we make it an annual event, or just continue to have it just once every hydrologic decade ?

Since meeting with you a couple of papers have been published (reprints enclosed) that have implications for flood frequency estimates made using the WRC's flood frequency guidelines (Bulletin 17-A). In the WRR paper it is pointed out (figure 7, and appendix C) that if flood sequences were distributed as Weibull and positively skewed in real space, γ_R, then infinite samples would always yield estimates of skew in log space, γ_L, of -1.14. Of course, small finite samples, n, would in expectation have other values of γ_L, but these estimates would not be primary functions of γ_R, but rather of n, and of the means, μ, and standard deviations, σ, of the individual records in real space. In view of this it would seem necessary that we be absolutely sure that none of the records used to estimate regional γ_L be Weibull distributed in real space. Further, if one proceeds without such an assurance then there appears to be no possible statistical justification for attempting to find a new regional γ_L, or for that matter to use any pre-existing γ_L.

696 079

In real space it is easy to see that flood hazard is, other things being equal, a function of γ_R. However, we have yet to identify a distribution for which $\gamma_L \neq f(n, \mu, \sigma)$. Hence, it now becomes difficult if not impossible to justify any flood estimating procedure that depends upon γ_L, and all past estimates made using such techniques must automatically be suspect. In particular, because log Pearson III with widely different γ_L's can lead to identical γ_R (see figure 9, and table A1 of the WRR paper) it would appear that log Pearson III flood freqency analyses would be the most suspect of all.

These are hardly new ideas. They were dicussed at the December, 1977 AGU symposium dedicated to the foibles of using the WRC flood freqency procedures. At that meeting there was an interesting split displayed between representatives of the Federal bureaucracy and those in the private sector. The latter group were mostly afraid that if they used WRC procedures for engineering consulting work, that they would be making themselves liable for civil suits charging professional incompetence. They have a point, the gulf between the scientific literature and the concepts of Bulletin 17-A does seem to be getting excessively large.

Recently, I have had several invitations to lecture on the statistics of flood frequency analysis, and how current knowledge relates to WRC procedures. Presumably your hydrology committee is concerned with such matters. Please keep me posted on all late breaking developments because it would be nice if my lectures could end on an up-beat note.

Hope to see you again one of these days.

Sincerely,

Jim Wallis.

Some Comparisons of Flood Statistics in Real and Log Space

J. MACIUNAS LANDWEHR AND N. C. MATALAS

U.S. Geological Survey, National Center, Reston, Virginia 22092

J. R. WALLIS

IBM Corporation, Watson Research Laboratory, Yorktown Heights, New York 10598

Some statistics of historical and simulated flood sequences were examined in real and log space. It was found that several statistical properties of floods in real space could not be inferred from those in log space without extensive knowledge of the distribution of floods in real space as well as information about their sampling characteristics. It is shown that the construction and use of regional skew maps in log space are most likely counterproductive.

What is reasonable is real; that which is real is reasonable.—George Hegel, 1821.

INTRODUCTION

The *U.S. Water Resources Council* [1976] recently issued guidelines for flood frequency analysis stating that 'the Pearson Type III distribution with log transformation of the flood data (log-Pearson Type III) is recommended as the basic distribution for defining the annual flood series.' The guidelines, however, are not strictly binding inasmuch as 'in those cases where the procedures of this guide are not followed, deviations must be supported by appropriate study and accompanied by a comparison of results using the recommended procedures.' In this work we show that such cases are more likely to be the rule than the exception.

Previously, *Matalas et al.* [1975] investigated several distribution functions, noting that samples derived from these distributions yielded estimates of skewness whose statistical properties did not accord with those derived from historical flood sequences. In this earlier work the statistical comparisons were conducted entirely in real space. This work has been extended into log space, and in addition to the distributions formerly treated, two further distributions have been added. One is the log Pearson type III distribution, recommended by the Water Resources Council, and the other is the Wakeby distribution, defined by H. A. Thomas (personal communication, 1976) and examined by *Houghton* [1977].

The extent to which skewness in log space is affected by the lower bound of the distribution of floods is examined. The implication of these results in the construction of regional skew maps in log space as advocated by the Water Resources Council is detailed.

More mathematical aspects of this work are included in appendices.

FLOOD STATISTICS

Floods are denoted as X in real space (RS) and $Z = \ln X$ in log space (LS). As random variables the distributions of X and Z may be characterized, though not necessarily uniquely, by the coefficients of variation, $C_v[\ ]$; skewness, $\gamma[\ ]$; and kurtosis, $\lambda[\ ]$. To assess some sampling properties of estimates of $C_v[\ ]$, $\gamma[\ ]$, and $\lambda[\ ]$ in RS and LS, sequences of length n = 10, 20, and 30 derived from 1351 historical flood records were considered. In Figure 1 the area of the United States is partitioned into 14 regions, and in Table 1 the distribution of the historical flood records over the regions is given, as well as the number $k(n)$ of nonoverlapping sequences of length n.

Let Y denote any of the parameters $C_v[\ ]$, $\gamma[\ ]$, and $\lambda[\ ]$, and let y denote a moment estimate of Y derived from a sequence of length n. Given the $k(n)$ values of y, the mean $\bar{y}$ and standard deviation $\hat{\sigma}(y)$ were calculated and are given in Tables 2 (RS) and 3 (LS).

Matalas et al. [1975] noted that the increase in $\hat{\sigma}(y)$ in RS with n suggests that floods are characterized by high values of γ (>2). The values of C_v and λ, as well as γ, presented in Tables 2 and 3, will serve as a guide for further analysis. Briefly, it is noted that the variations of $\bar{y}$ and $\hat{\sigma}(y)$ with n in LS are somewhat inconsistent with those in RS. For $Y \equiv C_v$, $\bar{y}$ and $\hat{\sigma}(y)$ are not very sensitive to n in either RS or LS. In the case $Y \equiv \gamma$, $\bar{y}$ decreases with n, and $\hat{\sigma}(y)$ tends to decrease with n in LS, though not in such a consistent manner; in RS, both $\bar{y}$ and $\hat{\sigma}(y)$ tend to increase with n. For $Y \equiv \lambda$, $\bar{y}$ and $\hat{\sigma}(y)$ tend to increase with n in both RS and LS.

It is of particular importance to note that in RS for $Y \equiv \gamma$, $\bar{y} > 0$ for $n \leq 30$ in all regions. However, in LS the tendency is for $\bar{y} < 0$, except in regions 1, 2, and 14. For a given region, $Q(\ )$, the proportion of sequences yielding estimates of γ in RS and LS of like or opposite signs, is given in Table 4. It is noted that $Q(++)$ tends to decrease with n for all regions except 1 and 4. For regions 3–11, $Q(++) < 0.50$ for all n. $Q(-+) = 0$ for all regions and n. That is, no sequence yielded a negative estimate of γ in RS and a positive estimate in LS. Except for region 1, $Q(+-)$ tends to increase with n, and in general for $n \geq 20$, $Q(+-) > 0.50$. For all regions, $[Q(--) + Q(+-)]$ tends to increase with n. For all regions, $[Q(++) + Q(+-)]$ tends to increase with n, where for $n = 30$, $[Q(++) + Q(+-)] \geq 0.90$. Thus as n becomes large, flood sequences yield dominantly positive estimates of γ in RS and negative estimates in LS. Rarely do flood sequences yield negative estimates of γ in both RS and LS, and even more rarely, seemingly never, do they yield negative estimates of γ in RS and positive estimates in LS.

CONDITION OF SEPARATION

Previously, *Matalas et al.* [1975] investigated some sampling properties of y, where $Y \equiv \gamma$, in RS and noted that the relation between $\bar{y}$ and $\hat{\sigma}(y)$ derived from the historical flood sequences was not in accord with that derived from Monte Carlo experiments for several well-known distribution functions. For the Monte Carlo results, let ω denote an estimate of γ. Based on

Paper number 8W0334.

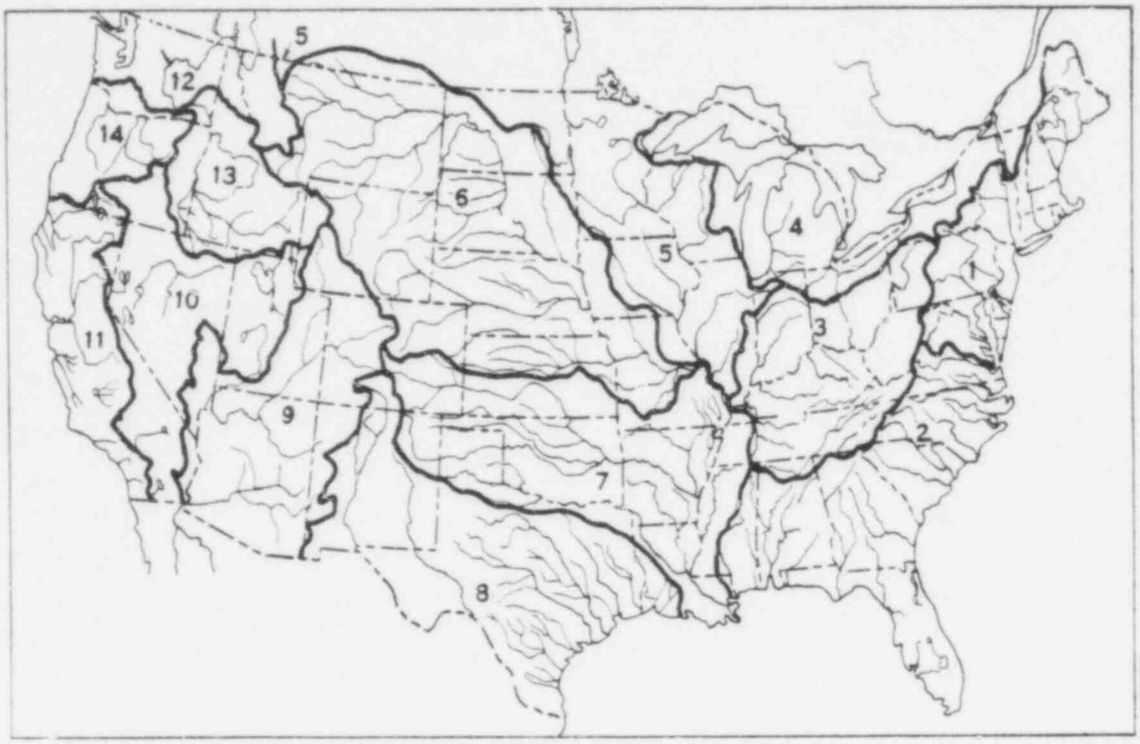


Fig. 1. Regional hydrologic division of the United States.

100,000 generated 'flood' sequences of length n, the mean $\bar{\mu}(\omega)$ and standard deviation $\bar{\sigma}(\omega)$ were determined for feasible values of γ in the range [0, 100]. The discordance between the historical floods and the generated floods was defined by what was called the condition of separation; that is, for $\bar{y} = \bar{\mu}(\omega)$, $\bar{\sigma}(y) > \bar{\sigma}(\omega)$ for each element of the distribution set Φ, where the elements of Φ were the uniform (U), normal (N), log normal (LN), Gumbel extreme value type I (G), Pearson type III (PIII), Weibull (W), and Pareto (P) distributions. The separation is a function of n and tends to become more pronounced as n increases.

The separation cannot be explained by the small number $k(n)$ of historical flood sequences relative to the 100,000 generated flood sequences, by autocorrelation [*Matalas et al.*, 1975], or by cross correlation [*Wallis et al.*, 1977]. However, the separation can be accounted for by spatial mixing of values of γ within a region and by nonstationarity in γ [*Wallis et al.*, 1977]. There is, of course, the possibility that a distribution which is not an element of Φ may explain the separation.

Some insight into the nature of such a distribution may be gained by reference to Figure 2, depicting the condition of separation for $n = 10$ for the distribution set Φ. From Figure 2 it is seen that the points $(\bar{y}, \bar{\sigma}(y))$ for all regions lie above the curves formed by the points $(\bar{\mu}(\omega), \bar{\sigma}(\omega))$ for each element of Φ. The upper bound on the curves is LN except for $\bar{y}$ greater than about 1.3 (corresponding to $\gamma \geq 15$), where the upper bound is seemingly formed by the curve for P. It has previously been observed [*Matalas et al.*, 1975] that the condition of separation becomes more pronounced as n increases.

In reference to Table 5, giving the relation between skewness γ and kurtosis λ for the distribution set Φ, it is noted that the relative positions of the curves for given values of γ can be explained as functions of λ. Given two elements of Φ, say, ϕ_1 and ϕ_2, the curve for ϕ_1 will be above that for ϕ_2 over that range of values for which λ of ϕ_1 is larger than λ of ϕ_2. Thus it seems that a necessary but perhaps not sufficient condition for minimizing the separation is that floods follow a distribution with λ larger than that for LN for $\gamma \leq 15$ and larger than that for P for $\gamma > 15$. Two distributions, log Pearson III and Wakeby, not previously in the set Φ, are considered below as possible explanations of the condition of separation.

TABLE 1. Distribution of Historical Records and Sequences of Length n

Region	No. of Records	No. of Sequences, $k(n)$		
		$n = 10$	$n = 20$	$n = 30$
1	193	691	286	184
2	88	291	116	80
3	164	582	237	154
4	50	182	72	46
5	130	418	160	104
6	114	368	147	84
7	100	314	121	78
8	79	267	105	70
9	80	274	111	70
10	35	133	55	30
11	81	321	135	74
12	98	311	128	70
13	41	148	61	34
14	98	356	145	91

Log Pearson Distribution

The random variable X (flood in RS) is said to be log Pearson type III (LP) distributed if $Z = \ln(X - \nu)$ (flood in LS) is distributed as Pearson type III (PIII), where the constant $\nu \leq x$. The density function of X is

$$f(x) = [1/(x - \nu)]f(z) \tag{1}$$

where $f(z)$, the density function of Z, is

696 082

TABLE 2. Flood Statistics in RS

	n = 10			n = 20			n = 30		
	C_v	γ	λ	C_v	γ	λ	C_v	γ	λ
Region 1									
$\bar{y}$	0.53	0.94	3.40	0.60	1.43	5.29	0.62	1.67	6.60
$\hat{\sigma}(y)$	0.24	0.73	1.61	0.27	0.88	3.36	0.27	0.96	4.63
Region 2									
$\bar{y}$	0.58	0.86	3.24	0.65	1.35	5.09	0.69	1.56	6.32
$\hat{\sigma}(y)$	0.23	0.74	1.60	0.22	0.90	3.45	0.22	1.00	4.69
Region 3									
$\bar{y}$	0.49	0.76	3.09	0.53	1.10	4.51	0.54	1.28	5.41
$\hat{\sigma}(y)$	0.21	0.71	1.46	0.22	0.87	3.09	0.21	0.91	4.06
Region 4									
$\bar{y}$	0.41	0.59	2.79	0.44	0.86	3.75	0.45	1.08	4.70
$\hat{\sigma}(y)$	0.16	0.66	1.24	0.16	0.72	2.26	0.15	0.85	3.70
Region 5									
$\bar{y}$	0.62	0.78	3.14	0.70	1.26	4.96	0.70	1.41	5.92
$\hat{\sigma}(y)$	0.27	0.75	1.55	0.31	0.93	3.44	0.29	1.04	4.95
Region 6									
$\bar{y}$	0.63	0.84	3.17	0.71	1.24	4.68	0.75	1.50	6.24
$\hat{\sigma}(y)$	0.34	0.76	1.54	0.40	0.89	3.26	0.43	1.14	5.51
Region 7									
$\bar{y}$	0.67	0.84	3.22	0.77	1.29	4.89	0.79	1.47	5.97
$\hat{\sigma}(y)$	0.32	0.77	1.59	0.39	0.94	3.42	0.36	1.07	5.15
Region 8									
$\bar{y}$	0.81	1.00	3.45	0.94	1.54	5.76	0.98	1.97	8.07
$\hat{\sigma}(y)$	0.38	0.78	1.79	0.50	1.04	4.11	0.40	1.12	5.39
Region 9									
$\bar{y}$	0.46	0.53	2.82	0.50	0.73	3.66	0.51	0.95	4.38
$\hat{\sigma}(y)$	0.24	0.76	1.28	0.25	0.86	2.34	0.26	0.86	2.99
Region 10									
$\bar{y}$	0.56	0.68	3.01	0.61	0.91	[illegible]	0.60	1.02	4.73
$\hat{\sigma}(y)$	0.34	0.75	1.38	0.42	0.99	3.21	0.43	1.01	4.19
Region 11									
$\bar{y}$	0.85	1.12	3.62	0.95	1.61	5.75	0.99	1.95	7.70
$\hat{\sigma}(y)$	0.35	0.76	1.78	0.39	0.94	3.64	0.43	1.06	5.40
Region 12									
$\bar{y}$	0.39	0.59	2.79	0.42	0.88	3.69	0.42	0.93	3.97
$\hat{\sigma}(y)$	0.15	0.64	1.14	0.14	0.68	3.69	0.12	0.64	2.10
Region 13									
$\bar{y}$	0.41	0.57	2.76	0.43	0.82	3.52	0.47	0.94	4.53
$\hat{\sigma}(y)$	0.21	0.71	1.30	0.17	0.70	2.13	0.19	0.93	3.51
Region 14									
$\bar{y}$	[illegible]	0.81	3.13	0.48	1.04	4.09	0.47	1.11	4.52
$\hat{\sigma}(y)$	0.20	0.69	1.42	0.16	0.72	2.41	0.15	0.69	2.56

$$f(z) = \frac{1}{|a|\Gamma(b)}\left(\frac{z-c}{a}\right)^{b-1}\exp\left[-\left(\frac{z-c}{a}\right)\right] \quad (2)$$

The density functions $f(x)$ and $f(z)$ are defined for $a \neq 0$ and $b > 0$. If $a > 0$, then $c \leq z \leq \infty$, and $m \leq x \leq \infty$, and if $a < 0$, then $-\infty \leq z \leq c$, and $0 \leq x \leq m$, where $m = \nu + e^c$. Thus if $a < 0$, there is a 'maximum certain flood' of value m. The moments and some properties of the distributions of X and Z are given in Appendix A. It is noted that if $a > 0$, then γ in both RS and LS is positive; if $a < 0$, then γ in LS is negative, but γ in RS may be positive or negative.

In Appendix A, tables of C_v, γ, and λ in RS and LS are given as functions of a and b conditioned on $c = 0$ and $m = 0$, for which $\nu = -1$. Values of γ and λ in RS and LS depend only upon a and b. In RS, as a becomes large, there are some values of γ with associated values of λ larger than those for LN. For example, with $a = 0.15$ and $b = 2.5$, $\gamma = 2.99$ (Table A5), and $\lambda = 25.03$ (Table A7). In the case of LN, $\lambda \simeq 22.4$ for $\gamma = 2.99$ (Table 5). This is further illustrated in Table 6. Thus LP may potentially minimize the condition of separation.

In Figure 3 a portion of the curve defined by the points $(\bar{\mu}(\omega), \bar{\sigma}(\omega))$ for PIII and N is shown. The values of $\bar{\mu}(\omega)$ and $\bar{\sigma}(\omega)$ were originally interpreted as being in RS [*Matalas et al.*, 1975]. However, they also can be interpreted as being in LS, where in LS the distribution is PIII. Superimposed in Figure 3 are the points $(|\bar{y}|, \hat{\sigma}(y))$ in LS for each of the regions for $n = 10$. All but one of the regional points (region 12) lie above the N-PIII curve, indicating that in LS, historical floods do not accord well with PIII or N.

Values of $\bar{\mu}(\omega)$ and $\bar{\sigma}(\omega)$ in RS derived from 50,000 LP sequences for each of several values of γ are given in Table 7. The algorithm for generating LP sequences, involving exponentiation of the algorithm for generating PIII sequences [*Johnk*, 1964; *Berman*, 1971], is described in Appendix A. The points $(\bar{\mu}(\omega), \bar{\sigma}(\omega))$ are shown in Figure 4 in relation to the LN curve. From Figure 4 it is seen that most of the LP points lie very near the LN curve. In the case where the LP λ is less than the LN λ for given values of γ the points lie below the LN curve. Of those points for which the LP λ is greater than the LN λ for given values of γ, 6 lie on or slightly above the LN curve, and 2 somewhat below it. Thus LP offers at best only marginal improvement over LN with respect to lessening the condition of separation. Large kurtosis may be necessary, but it is not sufficient for explaining the condition of separation.

TABLE 3. Flood Statistics in LS

	n = 10			n = 20			n = 30		
	C_v	γ	λ	C_v	γ	λ	C_v	γ	λ
Region 1									
$\bar{y}$	0.06	0.19	2.60	0.07	0.31	3.01	0.07	0.30	3.22
$\hat{\sigma}(y)$	0.02	0.63	0.84	0.02	0.57	1.10	0.02	0.53	1.06
Region 2									
$\bar{y}$	0.07	−0.01	2.54	0.07	0.07	2.85	0.08	0.01	2.92
$\hat{\sigma}(y)$	0.02	0.64	0.86	0.02	0.56	0.89	0.02	0.49	0.78
Region 3									
$\bar{y}$	0.06	−0.01	2.59	0.06	−0.09	3.13	0.06	−0.11	3.25
$\hat{\sigma}(y)$	0.04	0.66	0.90	0.04	0.66	1.13	0.04	0.58	1.08
Region 4									
$\bar{y}$	0.05	−0.08	2.55	0.06	−0.14	2.98	0.06	−0.03	3.16
$\hat{\sigma}(y)$	0.02	0.64	0.84	0.02	0.60	1.19	0.02	0.62	1.33
Region 5									
$\bar{y}$	0.09	−0.19	2.57	0.09	−0.23	3.16	0.09	−0.29	3.37
$\hat{\sigma}(y)$	0.05	0.65	0.93	0.05	0.68	1.44	0.05	0.69	1.84
Region 6									
$\bar{y}$	0.10	−0.06	2.57	0.10	−0.08	2.94	0.11	−0.09	3.06
$\hat{\sigma}(y)$	0.07	0.69	0.87	0.06	0.65	0.94	0.08	0.66	0.90
Region 7									
$\bar{y}$	0.09	−0.16	2.60	0.10	−0.17	2.97	0.10	−0.21	3.23
$\hat{\sigma}(y)$	0.10	0.69	0.94	0.07	0.63	1.08	0.10	0.66	1.38
Region 8									
$\bar{y}$	0.12	−0.17	2.49	0.13	−0.18	2.84	0.13	−0.18	3.02
$\hat{\sigma}(y)$	0.06	0.62	0.79	0.05	0.56	0.97	0.05	0.51	0.95
Region 9									
$\bar{y}$	0.07	−0.16	2.50	0.07	−0.29	2.98	0.07	−0.24	3.14
$\hat{\sigma}(y)$	0.03	0.63	0.85	0.03	0.62	1.16	0.03	0.59	1.23
Region 10									
$\bar{y}$	0.10	−0.14	2.47	0.10	−0.34	3.09	0.10	−0.40	3.38
$\hat{\sigma}(y)$	0.06	0.62	0.89	0.06	0.67	1.21	0.05	0.57	1.35
Region 11									
$\bar{y}$	0.13	0.01	2.53	0.13	−0.04	2.86	0.13	−0.04	3.07
$\hat{\sigma}(y)$	0.08	0.68	0.97	0.07	0.64	0.98	0.07	0.57	1.25
Region 12									
$\bar{y}$	0.05	−0.02	2.44	0.05	0.04	2.65	0.05	−0.04	2.81
$\hat{\sigma}(y)$	0.02	0.57	0.75	0.02	0.47	0.65	0.02	0.57	0.69
Region 13									
$\bar{y}$	−0.06	−0.01	2.40	0.07	−0.02	2.63	0.07	−0.14	2.77
$\hat{\sigma}(y)$	0.04	0.63	0.83	0.03	0.51	0.72	0.03	0.46	0.70
Region 14									
$\bar{y}$	0.06	0.12	2.52	0.06	0.06	2.78	0.06	0.01	2.83
$\hat{\sigma}(y)$	0.03	0.62	0.86	0.02	0.56	0.90	0.02	0.49	0.67

TABLE 4. Proportion $Q(\ )$ of Estimates of γ in RS and LS of Like and Opposite Signs

Value of n	Region													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
							$Q(+\ +)$							
10	0.60	0.51	0.48	0.45	0.37	0.45	0.41	0.38	0.36	0.44	0.48	0.50	0.52	0.59
20	0.71	0.47	0.41	0.43	0.31	0.43	0.34	0.34	0.36	0.26	0.46	0.52	0.48	0.52
30	0.74	0.50	0.38	0.49	0.28	0.43	0.28	0.37	0.33	0.27	0.40	0.43	0.38	0.53
							$Q(-\ -)$							
10	0.10	0.14	0.12	0.19	0.13	0.12	0.14	0.09	0.24	0.19	0.06	0.18	0.19	0.12
20	0.02	0.02	0.04	0.11	0.07	0.05	0.06	0.02	0.21	0.14	0.02	0.07	0.07	0.05
30	0.01	0.01	0.01	0.01	0.02	0.04	0.03	0.01	0.10	0.07	0.00	0.06	0.06	0.03
							$Q(+\ -)$							
10	0.30	0.35	0.40	0.36	0.50	0.43	0.45	0.53	0.40	0.37	0.46	[illegible]	0.29	0.29
20	0.27	0.52	0.55	0.46	0.62	0.52	0.60	0.64	0.43	0.60	0.52	0.41	0.45	0.43
30	0.25	0.49	0.61	0.50	0.70	0.55	0.69	0.62	0.57	0.66	0.60	0.51	0.56	0.44

For $Q(-\ +)$, no cases were observed.

696 084

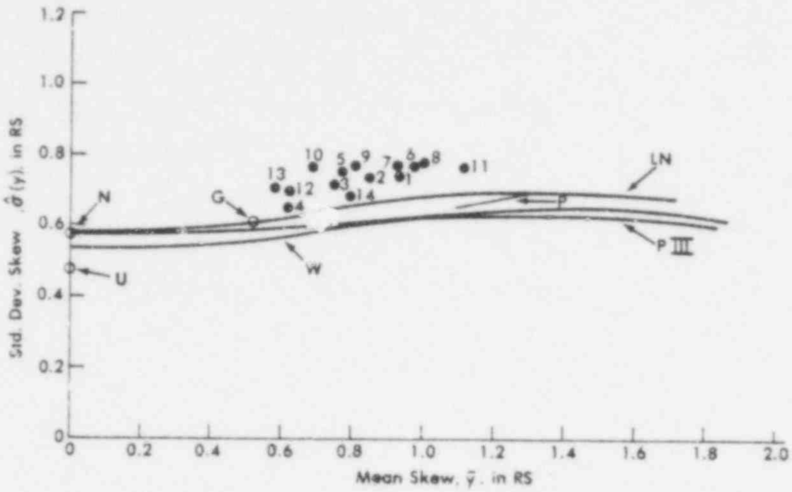


Fig. 2. Mean skew $\bar{y}$ versus standard deviation of skew, $\hat{\sigma}(y)$, for n = 10: historical and simulated data in RS [see *Matalas et al.*, 1975].

Although they are not shown, similar results in LS and RS were obtained for n = 20 and 30.

Wakeby Distribution

Following H. A. Thomas (personal communication, 1976), the random variable X is said to be distributed as Wakeby (WA) if

$$x = m + a[1 - (1 - F)^b] - c[1 - (1 - F)^{-d}] \quad (3)$$

where $F \equiv F(x) = P[X \leq x]$ and $x \geq m$. The moments and some properties of X are given in Appendix B. Also in Appendix B, tables are given for values of the coefficients of skewness and kurtosis for various values of (c/a), b, and d, as well as values of the coefficient of variation conditioned on m = 0. (Note that the ratio (c/a) is used rather than individual values of c and a, to allow for concise tabular presentation.)

Houghton [1977] has developed a technique for fitting WA to observed sequences and has shown that WA can account for the condition of separation. Thus WA offers an alternative to accounting for the condition of separation by the mixing of values of γ within a region as suggested by *Wallis et al.* [1977]. However, all WA's do not account for the condition of separation even if for a given value of γ, λ is larger than that for LN. Exploratory work has indicated that for given values of γ, where λ for WA is larger than that for LN, the condition of separation may be accounted for by WA with $b > 1$ but not with $b < 1$. In Appendix B it is shown that if a, c, and $d > 0$ and $0 < b < 1$, then the density function of WA monotonically decreases with x.

In Table 8 the values of $\bar{\mu}(\omega)$ and $\bar{\sigma}(\omega)$ for estimates of γ in RS based on 10,000 WA sequences of length n = 10 are given for selected values of b and d and for (c/a) = 5 with a = 1. The algorithm for generating WA sequences is given in Appendix B. Also in Table 8 the values of γ and λ corresponding to those of (c/a), b, and d are given. For each value of γ, λ is larger than that for LN. From Figure 5, showing the points ($\bar{\mu}(\omega)$, $\bar{\sigma}(\omega)$) relative to the LN curve, it is noted that those points conditioned on b = 8 lie above the LN curve, whereas those conditioned on b = 0.4 lie below it. Thus for a given value of γ, λ larger than that for LN is not sufficient to account for the condition of separation.

A WA with a = 1, b = 16, c = 5, and d = 0.19 can be transformed so that $\mu[X]$ = 10[illegible] and $\sigma[X]$ = 660, i.e., $C_v[X]$ = 0.66. Estimates of $C_v[X]$ derived from 10,000 sequences of length n = 10 have mean 0.55 and standard deviation 0.20, which agree closely with the mean and standard deviation of estimates of $C_v[X]$ for region 1 (see Table 2). Furthermore, in RS for n = 10 this WA yields estimates of γ with $\bar{\mu}(\omega)$ = 0.95 and $\bar{\sigma}(\omega)$ = 0.75, which are in close agreement with $\bar{y}$ = 0.94 and $\hat{\sigma}(y)$ = 0.73 for region 1. The corresponding WA values of

TABLE 5. Values of Skewness γ and Kurtosis λ for Elements of Φ

Value of γ	Value of λ						
	U	N	G	LN	PIII	W	P
0	1.8	3				2.72	
$\frac{1}{4}$				3.11	3.00	2.77	
$\frac{1}{2}$				3.45	3.38	3.03	
$\frac{1}{2}^{1/2}$				3.90	3.75	3.40	
1				4.83	4.50	4.16	
1.14			5.4	5.39	4.95	4.62	
$2^{1/2}$				6.75	6.00	5.73	
2				10.86	9.00	9.00	
3				22.40	16.50	17.6[illegible]	20.72
4				41.00	27.00	30.60	44.67
5				68.26	40.50	48.25	98.49
10				387.78	153.00	218.43	*
15				1139.73	340.50	546.73	*

* Here λ does not exist.

696 085

TABLE 6. Selected Values for γ and λ for LP in RS and LS and LN in RS

Parameter Values		LP				LN in RS	
		LS		RS			
a	b	γ	λ	γ	λ	γ	λ
−0.10	15	−0.52	3.40	0.55	3.21	0.50	3.45
−0.15	15	−0.52	3.40	1.04	4.41	1.00	4.83
0.01	4	1.00	4.50	1.08	4.82		
0.09	10	0.63	3.60	1.93	11.06	2.00	10.86
−0.08	4	−1.00	4.50	2.04	7.96		
0.01	1	2.00	9.00	2.06	9.50		
0.14	5	0.89	4.20	2.84	23.60	3.00	22.40
0.12	20	0.45	3.30	3.72	43.42	4.00	41.00
0.18	3	1.15	5.00	3.81	50.19		
0.16	10	0.63	3.60	4.55	82.41	5.00	68.26
0.20	1	2.00	9.00	4.65	73.80		
0.24	1	2.00	9.00	6.39	467.33	10.00	387.78
0.16	20	0.45	3.30	8.23	509.47		

Note that if X is distributed as LN, then log X is distributed as N with $\gamma = 0$ and $\lambda = 3$.

γ and λ are 3.95 and 51.94, respectively. From Table 5 it is noted that $\lambda = 51.94$ is larger than $\lambda = 41$ for LN with $\gamma = 4$ and hence may satisfy the condition of separation. For region 1 the γ mixing algorithm [*Wallis et al.*, 1977] provided a regional value of $\gamma \simeq 4.5$, which is similar to that of the aforementioned WA. Thus both the WA and the mixing algorithm suggest that for floods, γ is large, at least for region 1.

γ in RS and LS

It was noted above that historical flood sequences yielded estimates of γ which tended to be positive in RS and negative in LS. It was further noted that if X is distributed as LP, then γ may be positive in RS and negative in LS. But if γ in LS is negative, then X distributed as LP is bounded above. Other distributions, which are unbounded above, may be characterized by positive γ in RS and negative γ in LS. This point is illustrated below.

Consider the case where X is distributed as Weibull. The density function of X is

$$f(x) = \frac{b}{a}\left(\frac{x-m}{a}\right)^{b-1} \exp\left[-\left(\frac{x-m}{a}\right)^{b}\right] \quad (4)$$

where $b > 0$. If $a > 0$ (< 0), then in RS, $\gamma > 0$ (< 0), and m denotes the lower (upper) bound. From the moments of X, given in Appendix C, with $a > 0$, m is defined as

$$m = \mu[X] - \frac{\sigma[X]\Gamma(1 + 1/b)}{[\Gamma(1 + 2/b) - \Gamma^2(1 + 1/b)]^{1/2}}$$

$$= \mu[X] - \sigma[X]\theta(b) \quad (5)$$

where $\mu[X]$ and $\sigma[X]$ denote the mean and standard deviation of X. If $C_v[X] = \sigma[X]/\mu[X] = \theta^{-1}(b)$, then $m = 0$, and if $C_v[X] < \theta^{-1}(b)$, then $m > 0$. It is mathematically possible, but not physically reasonable, to have $m < 0$, in which case $C_v[X] > \theta^{-1}(b)$.

To examine the effect of sample size n on the relation between the skews in RS and LS as a function of the lower bound m in RS, Monte Carlo experiments were conducted. To generate W sequences with $a > 0$ ($\gamma > 0$ in RS), the variate x was expressed as

$$x = m + a[-\ln(1 - u)]^{1/b} \quad (6)$$

where $u \equiv F(x) = \int_m^x f(x)\,dx$ is distributed uniformly on the interval [0, 1]. For convenience, $\mu(x)$ was set equal to 1000. As previously defined, $\tilde{\mu}(\omega)$ represents the mean of estimates of skew derived from simulated sequences. Values of $\tilde{\mu}(\omega)$ for estimates of $\gamma[Z]$, where $Z = \ln(X)$, are given in Tables 9 and 10 for $\gamma[X] = \frac{1}{4}$ ($b \simeq 2.77$) and $\gamma[X] = 2$ ($b = 1$), respectively. The values of $\tilde{\mu}(\omega)$ were obtained from 10,000 sequences of length n = 10, 20, and 30 and from 10 sequences of length n =

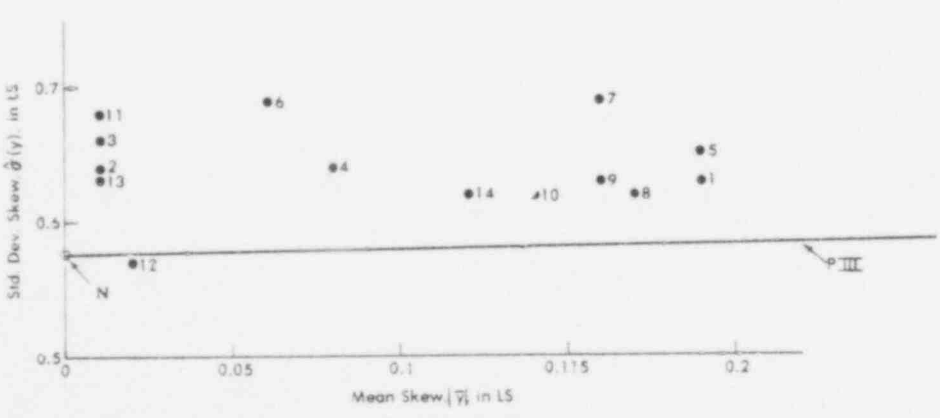


Fig. 3. Mean skew $\bar{y}$ versus standard deviation of skew, $\hat{\sigma}(y)$, for n = 10: historical and simulated PIII data in LS.

696 086

TABLE 7. Mean $\tilde{\mu}(\omega)$ and Standard Deviation $\tilde{\sigma}(\omega)$ of Estimates of γ in RS for 5000 Sequences of Length $n = 10$ Generated as LP With Parameters a and b

a	b	$\tilde{\mu}(\omega)$	$\tilde{\sigma}(\omega)$
*LP Characterized by Smaller λ Than LN**			
−0.80	4	1.15	0.61
−0.15	15	0.56	0.58
−0.10	15	0.30	0.56
0.01	4	0.54	0.60
0.01	1	0.98	0.67
*LP Characterized by Larger λ Than LN**			
0.09	10	0.73	0.66
0.14	5	0.88	0.68
0.12	20	0.97	0.69
0.18	3	0.98	0.68
0.20	1	1.01	0.69
0.16	10	1.18	0.67
0.16	20	1.16	0.70
0.24	1	1.22	0.67

*For given values of γ.

TABLE 8. Values of γ, λ, $\tilde{\mu}(\omega)$, and $\tilde{\sigma}(\omega)$, $\omega = \hat{\gamma}$, in RS for WA Sequences of Length $n = 10$ and $(c/a) = 5$

	Value of d		
	0.16	0.24	0.32
		$b = 0.40$	
γ	2.99	5.52	35.43
λ	23.78	346.25	*
$\tilde{\mu}(\omega)$	1.05	1.17	1.24
$\tilde{\sigma}(\omega)$	0.64	0.67	0.69
		$b = 8$	
γ	2.91	5.78	37.70
λ	24.49	404.46	*
$\tilde{\mu}(\omega)$	0.75	0.98	1.13
$\tilde{\sigma}(\omega)$	0.76	0.75	0.74

*Here λ is not defined; see Appendix B.

10,000. From Tables 9 and 10 it is noted that for $m = 0$, $\tilde{\mu}(\omega)$ is negative for all n, and as m increases, in which case $C_v[x]$ decreases, $\tilde{\mu}(\omega)$ increases and becomes positive. Moreover, it is noted that for $m = 0$, $\tilde{\mu}(\omega|n = 10{,}000) = -1.14$ for both values of $\gamma[X]$. In Appendix C it is shown that if X is distributed as W with $\gamma[X] > 0$ and $m = 0$, then $Z = \ln X$ is distributed as G with $\gamma[Z] = -1.139 \cdots$ for all $\gamma[X] > 0$, and thus $\tilde{\mu}(\omega|n) \to -1.139 \cdots$ as $n \to \infty$.

To assess the effect of $\lambda[X]$ on estimates of $\gamma[Z]$, where $Z = \ln X$, with X unbounded above and distributed with $\gamma[X] > 0$, X was assumed to be distributed as WA. Consider five random variables $X_1, \cdots, X_5$, each distributed as WA with $\gamma[X] \simeq 2$ such that $\lambda[X_1] < \cdots < \lambda[X_5]$. The values of the WA parameters a, b, c, and d for X_i, $i = 1, \cdots, 5$, and the corresponding values of $\gamma[X_i]$ and $\lambda[X_i]$ are given in Table 11 (see Appendix B). Note that X_1 has $\gamma[X_1] = 1.91$ and $\lambda[X_1] = 10.73$ and therefore X_1 is distributed approximately as LN (see Table 5).

Each X_i was transformed to $X_i' = \sigma[X_i']X_i + \mu[X_i']$, where $\mu[X_i'] = 1000$ for all i. Note that $\gamma[X_i] = \gamma[X_i']$ and $\lambda[X_i] = \lambda[X_i']$. For $\gamma[Z_i']$, where $Z_i' = \ln X_i'$, values of $\tilde{\mu}(\omega)$ were determined for each X_i' based on 10,000 WA sequences of lengths $n = 10$, 20, and 30 and 10 WA sequences of length $n = 10{,}000$. The values of $\tilde{\mu}(\omega)$ are given in Tables 12–16 for X_1', $\cdots$, X_5', respectively.

From Tables 12 and 16 it is noted that for X_i' with $m = 0$, $\tilde{\mu}(\omega)$ is negative for all i, and as m increases and $C_v[X']$ decreases, $\tilde{\mu}(\omega)$ increases and becomes positive. To this extent, X_i' distributed as WA behaves as though it were distributed as W. The effect of increasing $\lambda[X_i']$ is that for given values of $\mu[X_i']$, $C_v[X_i']$, and $\gamma[X_i']$, $\tilde{\mu}(\omega)$ decreases and becomes negative. The tables illustrate that although $\gamma[Z_i']$, approximated by $\tilde{\mu}(\omega)$ with $n = 10{,}000$, may be positive, $\tilde{\mu}(\omega)$ with n small may be negative but will become positive and approach $\gamma[Z_i']$ as $n \to \infty$.

If X is unbounded above and distributed with $\gamma[X] > 0$, then $Z = \ln X$ distributed with $\gamma[Z] < 0$ is realizable. It does not follow that this is so for all such X. If X is distributed as LN with $\gamma[X] > 0$, then $Z = \ln x$ is distributed with $\gamma[Z] \geq 0$ for $m \geq 0$. Thus if X distributed as WA is distributed approximately a[illegible], as in the case of X_1', the sampling properties of $Z = \ln X$ do not accord with those where X is indeed distributed as LN. In general, the value of $\gamma[Z]$ depends upon $C_v[X]$, $\gamma[X]$, and $\lambda[X]$. Estimates of $\gamma[Z]$ are in expectation dependent upon

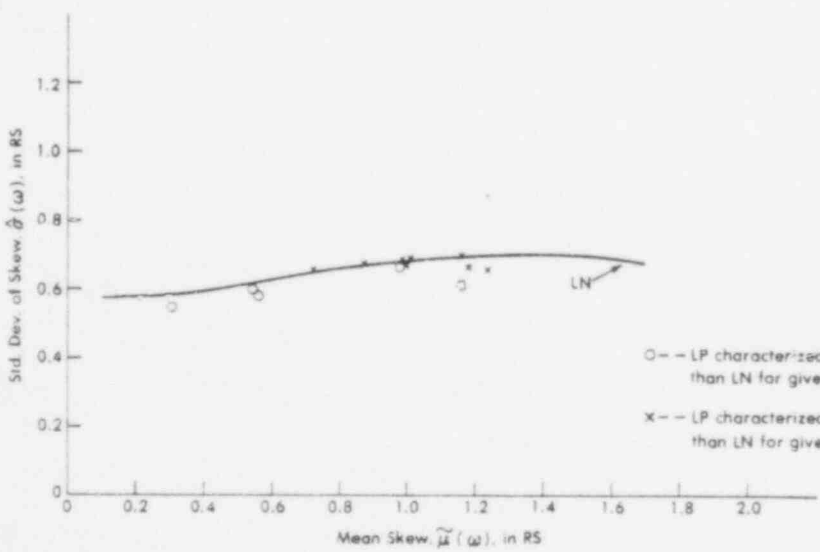


Fig. 4. Mean skew $\tilde{\mu}(\omega)$ versus standard deviation of skew, $\tilde{\sigma}(\omega)$, for $n = 10$: simulated LN and LP sequence in RS.

696 087

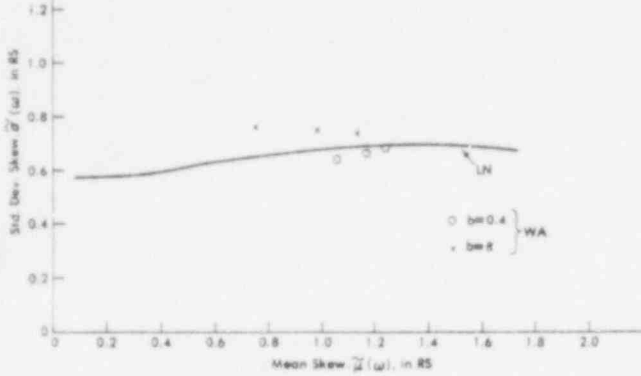


Fig. 5. Mean skew $\hat{\mu}(\omega)$ versus standard deviation of skew, $\hat{\sigma}(\omega)$, for n = 10: simulated LN and WA data in RS.

n as well as $C_v[X]$, $\gamma[X]$, and $\gamma[X]$. Thus X bounded above is not a necessary condition for $\gamma[Z] < 0$ to be realized.

Regional Skew Maps

The biases and large sampling errors inherent in estimates of $\gamma[X]$ from historical flood sequences have led the *U.S. Water Resources Council* [1976] to suggest that if $n < 100$, the use of regional skew maps is preferable to the estimate of $\gamma[X]$ obtainable from using only the specific record. In particular, the Council suggested the construction of regional skew maps in LS. However, since historical records in the United States are almost always less than 100 years, the Water Resources Council is effectively advocating the use of regional skew maps at all sites.

The above discussions indicate that the usefulness of regional skew maps in LS is questionable. It was noted that $\gamma[Z]$ is not simply related to $\gamma[X]$ but depends upon $C_v[X]$, $\lambda[X]$, and the distribution of X as well. Thus it is difficult to infer properties of X from those of Z without considerable knowledge about the distribution of X and its sampling characteristics.

A regional skew map consists of a set of contours of equiskew, smoothing out the variability in the estimates of skew at a large number of sites in the region. Quite similar regional maps in LS may be realized whether $\gamma[X]$ is a constant for all sites in a region or varies among the sites. But even if distinctly different LS maps were realized, it would be difficult to determine if a particular LS map was generated by constant or varying values of $\gamma[X]$. Given only the estimates of $\gamma[Z]$ at the various sites, little can be said with confidence about $\gamma[X]$ or $\gamma[Z]$, however smooth or rugged the contours of the estimates of $\gamma[Z]$ may be.

TABLE 9. Values of $\hat{\mu}(\omega)$ for Estimates of $\gamma[Z]$

$C_v[X]$	m	$\hat{\mu}(\omega)$ n = 10	n = 20	n = 30	n = 10,000
0.39	0	−0.51	−0.73	−0.83	−1.14
0.34	130	−0.41	−0.56	−0.61	−0.75
0.29	260	−0.31	−0.42	−0.45	−0.53
0.24	390	−0.23	−0.30	−0.33	−0.37
0.19	510	−0.15	−0.18	−0.20	−0.23
0.14	640	−0.07	−0.09	−0.10	−0.08
0.09	770	0.01	0.01	0.02	0.04

X is distributed as W with $\mu[X]$ = 1000 and $\gamma[X] = \frac{1}{4}$ (b = 2.77).

TABLE 10. Values of $\hat{\mu}(\omega)$ for Estimates of $\gamma[Z]$

$C_v[X]$	m	$\hat{\mu}(\omega)$ n = 10	n = 20	n = 30	n = 10,000
1	0	−0.52	−0.74	−0.83	−1.14
0.9	100	−0.10	−0.11	−0.11	−0.11
0.8	200	0.09	0.13	0.14	0.17
0.7	300	0.25	0.31	0.33	0.38
0.6	400	0.37	0.46	0.49	0.55
0.5	500	0.48	0.59	0.63	0.73
0.4	600	0.58	0.73	0.78	0.90
0.3	700	0.66	0.85	0.92	1.07
0.2	800	0.77	1.00	1.08	1.29

X is distributed as W with $\mu[X]$ = 1000 and $\gamma[X]$ = 2 (b = 1).

Consider the case where X is distributed as W with $\gamma[X] > 0$ and m = 0. For a hypothetical region, depicted in Figure 6, there is a north to south flowing stream fed by first-order tributaries flowing east to west. One of three hydrologies is assumed to exist, where the hydrologies are as follows. For hydrology 1 (H_1), $\gamma[X]$ = 2.5 at all sites. For hydrology 2 (H_2), $\gamma[X]$ decreases from 4 to 1 in a north to south direction but is a constant along any tributary. For hydrology 3 (H_3), $\gamma[X]$ does not vary from one tributary to another, but along any tributary, $\gamma[X]$ decreases from 4 to 1 in the downstream direction. Thus for each hydrology the mean regional skew in RS is 2.5.

The regional skew maps in RS and LS for the three hydrologies are depicted in Figure 7. It is noted that although the skew contours in RS are distinctly different for the three hydrologies, the contours in LS are identical. Thus given the LS regional skew map, the RS contours cannot be uniquely specified. Even if it is known that X is distributed as W and a perfect regional skew map in LS is available, the contours in RS still cannot be uniquely inferred.

For the hypothetical region, consider a fourth hydrology which is defined as follows. For hydrology 4 (H_4), X is distributed as WA with $\gamma[X]$ = 2 and $\sigma[X]$ = 61 at all sites, where $X \equiv X_3'$ in the above WA discussions. Along any tributary, m increases in the downstream direction from 0 to 100. Consequently, $\mu[X]$ increases in the downstream direction from 100 to 680.

In Figure 8 the regional skew map in RS is shown in relation to LS maps based on expected values of estimates of $\gamma[Z]$ for three cases: (1) n = 10 at all sites, (2) n increases along each tributary in the downstream direction from 10 to 30, and (3) $n = \infty$ at all sites. Expected values of estimates of $\gamma[Z]$ for $n = \infty$ are approximated by those for n = 10,000.

From Figure 8 it is noted that the skew contours are dis-

TABLE 11. Characteristics of X_i Distributed as WA

i	a	b	c	d	$\gamma[X_i]$	$\lambda[X_i]$
1	1	1	5	0.12	1.91	10.73
2	1	1.5	2.5	0.17	1.96	15.14
3	1	4	2.5	0.17	2.00	18.45
4	1	5.5	1.25	0.21	2.02	33.77
5	1	8	0.83	0.23	2.05	67.75

696 038

TABLE 12. Values of $\hat{\mu}(\omega)$ for Estimates of $\gamma[Z_1']$

$C_v[X_1']$	m	$\hat{\mu}(\omega)$			
		$n = 10$	$n = 20$	$n = 30$	$n = 10{,}000$
0.87	0	−0.61	−0.87	−0.98	−1.32
0.77	115	−0.23	−0.26	−0.26	−0.27
0.67	230	−0.03	−0.01	0.00	0.03
0.57	345	0.13	0.18	0.20	0.25
0.47	460	0.26	0.34	0.37	0.45
0.37	575	0.37	0.48	0.54	0.66
0.27	690	0.49	0.64	0.70	0.88
0.17	805	0.58	0.78	0.88	1.11
0.07	920	0.70	0.96	1.07	1.48

X_1' is distributed as WA with $\mu[X_1'] = 1000$ and $\gamma[X_1'] = 1.91$.

tinctly different in RS and LS. Moreover, in LS the contours vary considerably with n, though their general structures are the same. Thus n finite and variable from site to site adds confusion to the already confusing task of inferring the RS skews from the LS skews or for that matter inferring the LS contour that would apply with $n = \infty$.

For the same hypothetical region, consider a fifth and a sixth hydrology defined as follows. For hydrology 5 (H_5), X with lower bound equal to zero is distributed as LP with $a = -1$ and $b = 15$ such that at all sites, $\gamma[X] = 0.55$ and $\gamma[Z] = -0.52$. For hydrology 6 (H_6), X with lower bound equal to 1 is distributed as LP with $a = 0.01$ and $b = 25$ such that at all sites, $\gamma[X] = 0.56$ and $\gamma[Z] = 0.40$.

For H_5 and H_6 the regional skew maps in RS are very nearly identical, but in LS they are distinctly different (see Figure 9). Thus with X distributed as LP, inference problems are no less confusing than they are with X distributed differently.

SUMMARY AND CONCLUSIONS

Flood statistics in both real space, X, and log space, $Z = \ln X$, were assessed with respect to specific distribution functions. Based on a partitioning of the United States into 14 regions, the mean $\bar{y}$ and standard deviation $\hat{\sigma}(y)$ of estimates of $\gamma[X]$ derived from available flood sequences of length n within a region were compared with the mean $\hat{\mu}(\omega)$ and standard deviation $\hat{\sigma}(\omega)$ of estimates of $\gamma[X]$ derived from Monte Carlo experiments conditioned on specific distribution functions. For several well-known distribution functions defined by three or fewer parameters, *Matalas et al.* [1975] showed that in real space there exists, for each region, what was called a condition of separation; that is, for $\bar{y} = \hat{\mu}(\omega)$, $\hat{\sigma}(y) > \hat{\sigma}(\omega)$. Among the distribution functions the condition of separation was less pronounced for the log normal distribution LN and for the Pareto distribution P at the higher values of $\gamma[X]$.

TABLE 13. Values of $\hat{\mu}(\omega)$ for Estimates of $\gamma[Z_2']$

$C_v[X_2']$	m	$\hat{\mu}(\omega)$			
		$n = 10$	$n = 20$	$n = 30$	$n = 10{,}000$
0.78	0	−0.69	−0.99	−1.12	−1.52
0.68	128	−0.33	−0.39	−0.40	−0.41
0.58	257	−0.13	−0.12	−0.11	−0.09
0.48	385	0.04	0.08	0.10	0.17
0.38	513	0.17	0.26	0.29	0.38
0.28	642	0.29	0.41	0.47	0.65
0.18	770	0.48	0.58	0.67	0.95
0.08	899	0.52	0.75	0.86	1.33

X_2' is distributed as WA with $\mu[X_2'] = 1000$ and $\gamma[X_2'] = 1.96$.

TABLE 14. Values of $\hat{\mu}(\omega)$ for Estimates of $\gamma[Z_3']$

$C_v[X_3']$	m	$\hat{\mu}(\omega)$			
		$n = 10$	$n = 20$	$n = 30$	$n = 10{,}000$
0.61	0	−0.76	−1.16	−1.36	−2.00
0.51	164	−0.44	−0.57	−0.61	−0.66
0.41	328	−0.24	−0.26	−0.26	−0.21
0.31	493	−0.06	−0.01	0.02	0.16
0.21	657	0.10	0.22	0.28	0.50
0.11	821	0.25	0.44	0.55	1.03
0.09	985	0.40	0.68	0.86	1.99

X_3' is distributed as WA with $\mu[X_3'] = 1000$ and $\gamma[X_3'] = 2.06$.

By means of Monte Carlo experiments, *Matalas et al.* [1975] showed that the condition of separation can be explained neither by the small number of historical flood sequences relative to the very large number of generated sequences conditioned on the specific distribution functions nor by autocorrelation. *Wallis et al.* [1977] noted that cross correlation cannot explain the condition of separation but that separation can be accounted for both by spatial mixing of values of $\gamma[X]$ within a region and by nonstationarity in $\gamma[X]$.

Analysis of historical flood data in addition to further Monte Carlo experiments in real space (RS) and log space (LS) led to the following conclusions.

1. In RS, estimates of $\gamma[X]$ derived from flood sequences are dominantly positive and become more so as n increases, whereas in LS, estimates of $\gamma[Z]$ are dominantly negative and become more so as n increases for most of the U.S. regions.

2. The log Pearson type III distribution (LP), recommended by the *U.S. Water Resources Council* [1976] for use by all federal agencies in conducting flood frequency studies, can accommodate $\gamma[X] > 0$ and $\gamma[Z] < 0$. However, if $\gamma[Z] < 0$, X is bounded above regardless of the sign of $\gamma[X]$.

3. The LP distribution is unlikely to explain the condition of separation, since it fails to do so for representative parameter values of the distribution.

4. In LS, all but one of the 14 regional points ($\bar{y}$, $\hat{\sigma}(y)$) lie above the $\hat{\mu}(\omega)$ versus $\hat{\sigma}(\omega)$ curve for the Pearson type III distribution (PIII), indicating that the condition of separation cannot be accounted for by LP.

5. For a distribution to yield a less pronounced condition of separation relative to LN it is apparently necessary but not sufficient that for a given value of $\gamma[X]$ the distribution should have a value of $\lambda[X]$ larger than that for LN.

6. Although for certain LP, $\lambda[X]$ is larger than that for LN for a given value of $\gamma[X]$, the condition of separation is only marginally less pronounced than that for LN.

7. *Houghton* [1977] has shown that the Wakeby distribution (WA), defined as

$$x = m + a[1 - (1 - F)^b] - c[1 - (1 - F)^{-d}]$$

TABLE 15. Values of $\hat{\mu}(\omega)$ for Estimates of $\gamma[Z_4']$

$C_v[X_4']$	m	$\hat{\mu}(\omega)$			
		$n = 10$	$n = 20$	$n = 30$	$n = 10{,}000$
0.49	0	−0.83	−1.35	−1.61	−2.61
0.39	202	−0.55	−0.76	−0.82	−0.93
0.29	405	−0.35	−0.41	−0.42	−0.36
0.19	607	−0.16	−0.12	−0.08	0.14
0.09	810	0.01	0.15	0.24	0.16

X_4' is distributed as WA with $\mu[X_4'] = 1000$ and $\gamma[X_4'] = 2.02$.

696 089

TABLE 16. Values of $\tilde{\mu}(\omega)$ for Estimates of $\gamma[Z_s']$

$C_v[X_s']$	m	$\tilde{\mu}(\omega)$ $n = 10$	$n = 20$	$n = 30$	$n = 10{,}000$
0.40	0	−0.82	−1.44	−1.78	−3.25
0.30	248	−0.57	−0.85	−0.97	−1.16
0.20	495	−0.36	−0.48	−0.51	−0.41
0.10	743	−0.17	−0.15	−0.10	0.33
0.04	990	0.00	0.17	0.30	1.34

X_s' is distributed as WA with $\mu[X_s'] = 1000$ and $\gamma[X_s'] = 2.05$.

where $F \equiv F(x) = P[X \leq x]$, can explain the condition of separation. Among those WA's where $\lambda[X]$ is larger than that for LN for a given value of $\gamma[X]$ it seems that the condition of separation can be explained with $b > 1$ but not with $b \leq 1$ when a, b, c, and $d > 0$.

8. WA offers an alternative to spatial mixing of values of $\gamma[X]$ within a region as an explanation for the condition of separation. A particular WA yielding values of $\tilde{\mu}(\omega)$ and $\tilde{\sigma}(\omega)$ in close agreement with $\bar{y}$ and $\hat{\sigma}(y)$ for region 1 has $\gamma[X] = 3.95$. This is close to the value 4.5 obtained by *Wallis et al.* [1977] by spatial mixing of values of $\gamma[X]$ within region 1 conditioned on LN. Whether WA and spatial mixing of $\gamma[X]$ would yield results in close agreement for the other regions remains to be determined.

9. Monte Carlo experiments with WA indicate that for a given value of $\gamma[X] > 0$, $\gamma[Z] < 0$ is realizable. For given values of $\mu[X]$, $C_v[X]$, and $\gamma[X]$ the expected value $\tilde{\mu}(\omega)$ of estimates of $\gamma[Z]$ decreases as $\lambda[X]$ increases.

10. If X is unbounded above and distributed with $\gamma[X] > 0$, then $\gamma[Z]$ may be negative, particularly if m, the lower bound on X, is small. If $\gamma[Z] < 0$ with m small, then $\gamma[Z]$ will increase and become positive as m increases. Thus the property of $\gamma[X] > 0$ but $\gamma[Z] < 0$ is realizable with X unbounded above. However, X would be bounded above if it were distributed as LP.

11. From values of $\gamma[Z]$ it is difficult to infer $\gamma[X]$. Distinctly different $\gamma[X]$ contours may give rise to identical $\gamma[Z]$ contours and vice versa. Thus, in general, the skew map in RS cannot be uniquely inferred from that in LS.

12. With small samples it is difficult to infer $\gamma[X]$ from estimates of $\gamma[X]$, and these difficulties are only compounded by attempting inference from estimates of $\gamma[Z]$.

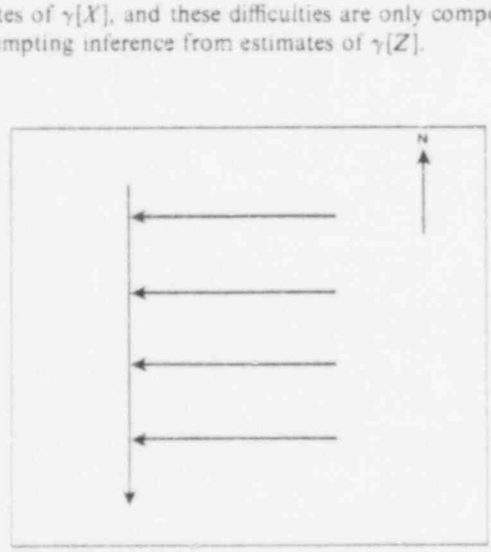


Fig. 6. Stream network for hypothetical region.

13. The construction and use of regional skew maps are most likely to be counterproductive.

Appendix A: Comments on the Log Pearson Type III Distribution

Definition

Let ν be a constant and $X > \nu$ a random variable distributed as log Pearson type III (LP), such that Z, defined as

$$Z = \ln(X - \nu) \tag{A1}$$

is a random variable distributed as Pearson type III (PIII). The probability density function of X can be written as

$$f(x) = [1/(x - \nu)]f(z) \tag{A2}$$

where

$$f(z) = \frac{1}{|a|\,\Gamma(b)} \left(\frac{z - c}{a}\right)^{b-1} \exp\left[-\left(\frac{z - c}{a}\right)\right] \qquad |a| > 0 \tag{A3}$$

The ranges of variation of X and Z are

$$\begin{aligned} a > 0 &\qquad e^c + \nu \leq x \leq \infty \\ a < 0 &\qquad \nu \leq x \leq e^c + \nu \\ a > 0 &\qquad c \leq z \leq \infty \\ a < 0 &\qquad -\infty \leq z \leq c \end{aligned} \tag{A4}$$

When $a > 0$, there exists a lower bound m on the values of x such that

$$m \equiv e^c + \nu \tag{A5}$$

and $x = m$ when $z = c$. Conversely, when $a < 0$, m is an upper bound on x; again, $x = m$ when $z = c$. If $c = 0$, then $m = 1 + \nu$, and if both $c = 0$ and $\nu = 0$, then $m = 1$.

Moments

Consider only the case where $a > 0$. By definition, the kth-order moment $E[X^k]$ is

$$E[X^k] = \int_m^\infty x^k f(x)\,dx \tag{A6}$$

so that

$$E[X^k] = \sum_{i=0}^{k} \binom{k}{i} \nu^{k-i} e^{ic}/(1 - ia)^b \tag{A7}$$

where $(1 - ia) > 0$ for $i = 0, 1, \cdots, k$. Thus the kth moment of X exists only if $a < 1/k$. For Z,

$$E[Z^k] = \int_m^\infty z^k f(z)\,dz \tag{A8}$$

so that

$$E[Z^k] = \frac{1}{\Gamma(b)} \sum_{i=0}^{k} \binom{k}{i} c^i a^{k-i} \Gamma(k + b - i) \tag{A9}$$

exists for all $a \neq 0$. The mean $\mu[\ ]$, the standard deviation $\sigma[\ ]$, and the coefficients of skewness, $\gamma[\ ]$, kurtosis, $\lambda[\ ]$, and variation, $C_v[\ ]$, for X and Z are defined in Table A1.

It is noted that if $a \geq 1/k_0$, then the moments of order $k \geq k_0$ are not defined for X, although the moments of order k exist for $Z \,\forall\, k$. Thus if $a \geq \frac{1}{4}$, $\lambda[X]$ is not defined; if $a \geq \frac{1}{3}$, $\gamma[X]$ is not defined; if $a \geq \frac{1}{2}$, $\sigma[X]$ is not defined; and if $a \geq 1$, $\mu[X]$ is not defined.

696 090

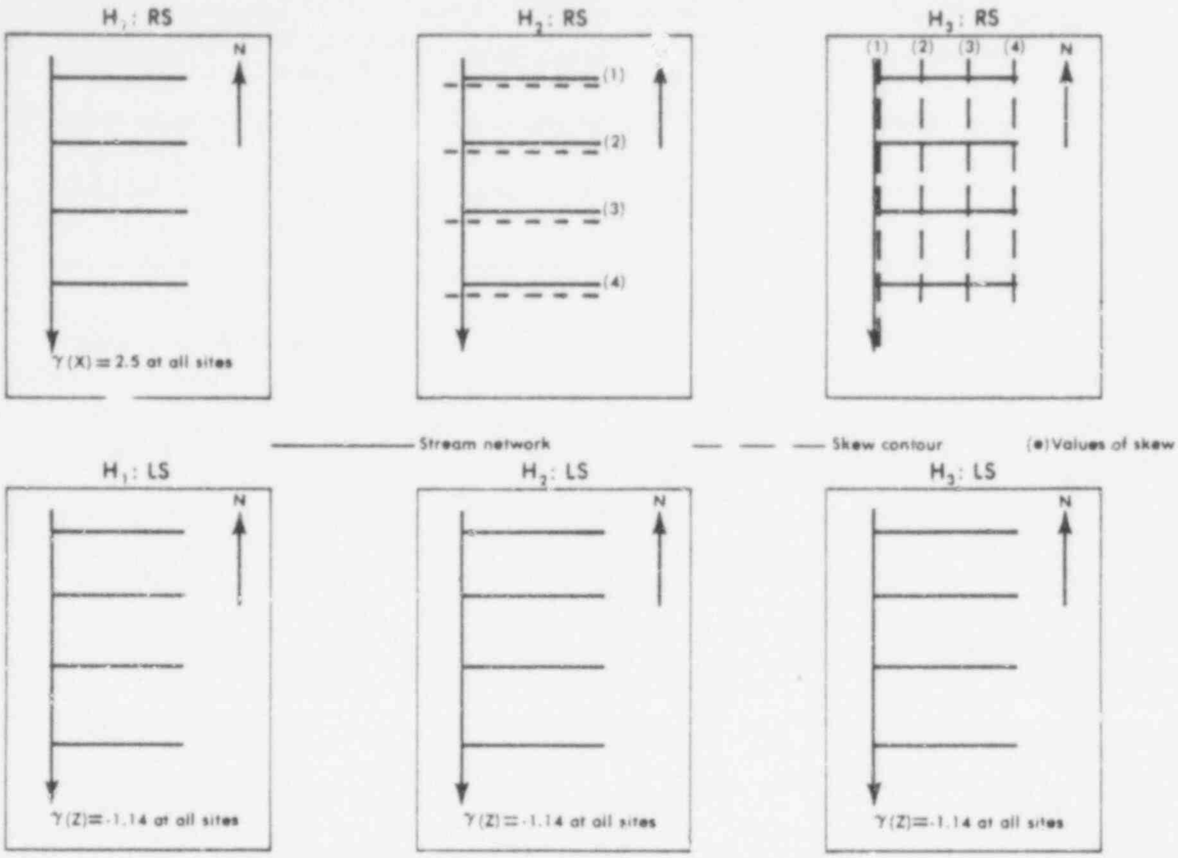


Fig. 7. Skew maps in RS and LS for hypothetical region: X distributed as W with $m = 0$.

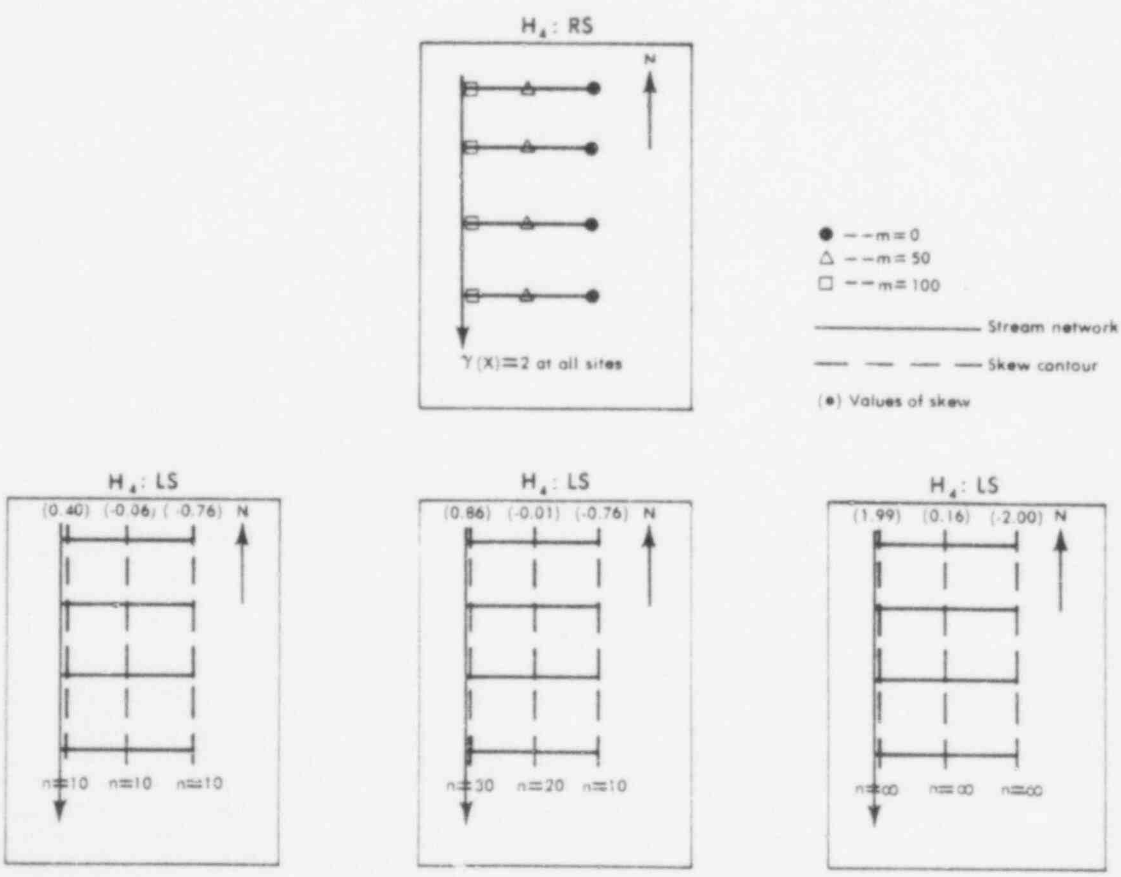


Fig. 8. Skew maps in RS and LS for hypothetical region: X distributed as WA.

696 091

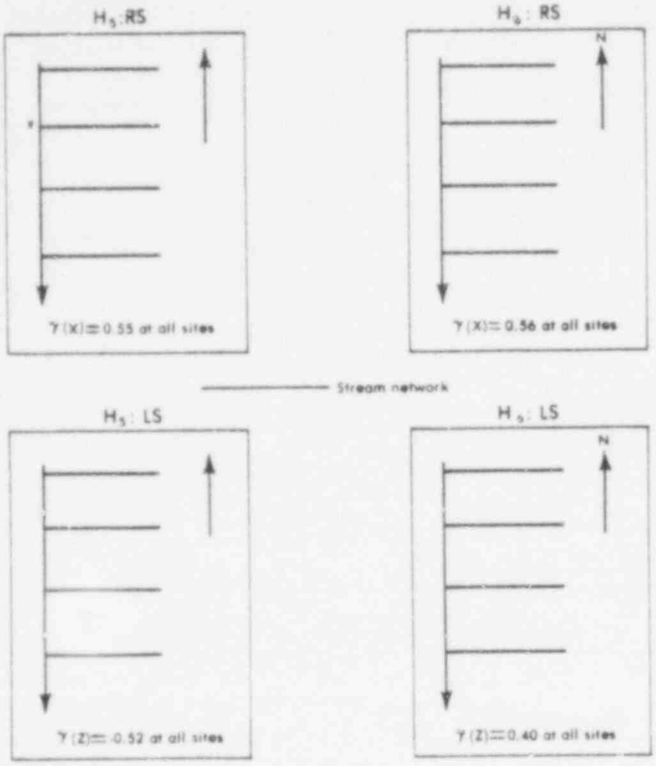


Fig. 9. Skew maps in RS and LS for hypothetical region: X distributed as LP.

Consider the case where $a < 0$. The equations for the moments of X and Z are the same as those for the case where $a > 0$. However, since $k \geq 0$, $(1 - ka) \geq 1$ for $a < 0$. Therefore if $a < 0$, then all moments, for both X and Z, are defined.

Tables: Coefficients of Variation, Skew, and Kurtosis

Tables A2–A7 present values for the coefficients of variation, skew, and kurtosis for a and b with $c = 0$ and $m = 0$. The following observations are made from Tables A2 and A3 with respect to $C_v[\ \ ]$. (1) As b increases, $C_v[Z]$ approaches 0. (2) For a given value of b, $C_v[X]$ increases as $|a|$ increases. (3) For a given value of a, $C_v[X]$ increases as $|\gamma[Z]|$ decreases.

The following observations are made from Tables A4 and A5 with respect to $\gamma[\ \ ]$. (1) As b increases, $\gamma[Z]$ approaches 0. (2) For a given value of b, that is, for a given $\gamma[Z]$, $\gamma[X]$ increases as $|a|$ increases. (3) If $a > 0$, then $\gamma[X] > 0$. Given $a = a^* > 0$, $\gamma[X]$ is a minimum for some $b = b^*$, i.e., for some $\gamma[Z] = \gamma^*$. As a increases in value, the associated γ^* also increases (b^* decreases). (4) If $a < 0$, then $\gamma[X]$ may be either positive or negative. However, given $b = b^*$, i.e., for $\gamma[Z] = \gamma^*$, there exists an $a = a^* < 0$ such that $\gamma[X] = 0$. If $a = a^*$ and $b < b^*$, i.e., for $\gamma[Z] < \gamma^*$, then $\gamma[X] < 0$; conversely, if $\gamma[Z] > \gamma^*$, then $\gamma[X] > 0$. Furthermore, as γ^* decreases in value, the associated $a^* < 0$ also decreases.

Thus suppose that floods X are distributed as LP. If X has no upper bound, then a must be positive, so that the skews of both X and Z are positive. Alternatively, if there is a 'maximum certain flood,' so that X is bounded above, then a and $\gamma[Z]$ are negative, although $\gamma[X]$ may be positive or negative. Conversely, if $\gamma[Z] > 0$, then $\gamma[X] > 0$, and floods are unbounded above in real space. However, if $\gamma[z] < 0$, then $\gamma[X]$ may be positive, zero, or negative, but there is a maximum certain flood in both log and real space.

The following observations are made from Tables A6 and A7 with respect to the $\gamma[\ \ ]$. (1) As b increases, $\lambda[Z]$ approaches 3. (2) For a given value of a, there exists some $b = b^*$, that is, some $\lambda[Z] = \lambda^*$, such that $\lambda[X]$ is a minimum. Furthermore, as $|a|$ increases, λ^* increases, and b^* decreases. (3) For $a < 0$ and any given b, that is, $\lambda[Z]$, $\lambda[X]$ decreases as a decreases. (4) For $a > 0$ and any given b, that is, $\lambda[Z]$, $\lambda[X]$ increases to infinity as a approaches 0.25 and decreases as a approaches 0.

From Tables A5 and A7, one can choose pairs of a and b values such that for X distributed as LP defined by these parameters, $\lambda[X] > \lambda[Y]$, even though $\gamma[X] \leq \gamma[Y]$ for Y distributed as log normal.

Log Pearson Random Variables

To generate random numbers which are distributed as LP, the following algorithm [*Johnk*, 1964; *Berman*, 1971] can be used. Let

$$x = \exp\left\{c + a\left[-\ln \prod_{k=1}^{[b]} u_k - B \ln u\right]\right\} + \nu \qquad \text{(A10)}$$

where $u \sim$ uniform (U) on [0, 1], $[b]$ denotes the greatest integer less than or equal to b, and B is defined as follows. (1) Set $r = b - [b]$ and $s = 1 - r = 1 - b + [b]$. (2) Generate u_1, $u_2 \sim U[0, 1]$. (3) Set $\zeta = u_1^{1/r}$ and $\xi = u_2^{1/s}$. (4) If $\zeta + \xi > 1$, return to step 2; otherwise, proceed as follows. (5) Set $B = \zeta/(\zeta + \xi)$. If b is an integer, $B = 0$, and therefore

$$x = \exp\left\{c + a\left[-\ln \prod_{k=1}^{b} u_k\right]\right\} + \nu \qquad \text{(A11)}$$

that is,

$$x = \exp\left\{c + a\left[-\sum_{k=1}^{b} \ln u_k\right]\right\} + \nu \qquad \text{(A12)}$$

TABLE A1. Statistical Properties of X Distributed as LP and Z Distributed as PIII

Property	Definition
$\mu[X]$	$\nu + e^c(1 - a)^{-b}$
$\sigma[X]$	$e^c[(1 - 2a)^{-b} - (1 - a)^{-2b}]^{1/2}$
$\gamma[X]$	$e^{3c}[(1 - 3a)^{-b} - 3(1 - a)^{-b}(1 - 2a)^{-b} + 2(1 - a)^{-3b}]\sigma^{-3}[X]$
$\lambda[X]$	$e^{4c}[(1 - 4a)^{-b} - 4(1 - a)^{-b}(1 - 3a)^{-b} + 6(1 - a)^{-2b}(1 - 2a)^{-b} - 3(1 - a)^{-4b}]\sigma^{-4}[X]$
$C_v[X]$	$e^c[(1 - 2a)^{-b} - (1 - a)^{-2b}]^{1/2}[\nu + e^c(1 - a)^{-b}]^{-1}$
$\mu[Z]$	$c + ab$
$\sigma[Z]$	$\lvert a\rvert b^{1/2}$
$\gamma[Z]$	$2a[\lvert a\rvert b^{1/2}]^{-1}$
$\lambda[Z]$	$3 + 6b^{-1}$
$C_v[Z]$	$\lvert a\rvert b^{1/2}[c + ab]^{-1}$

696 092

TABLE A2. Coefficient of Variation in RS (X) and LS (Z) for Log Pearson Distribution With $m = 0$ and $c = 0$

	$C_v[X]$										
b	$a =$ −2.00	$a =$ −1.50	$a =$ −1.00	$a =$ −0.80	$a =$ −0.50	$a =$ −0.30	$a =$ −0.15	$a =$ −0.10	$a =$ −0.05	$a =$ −0.01	$C_v[Z]$
0.05	0.17	0.15	0.12	0.11	0.08	0.05	0.03	0.02	0.01	0.00	−4.47
0.10	0.25	0.21	0.17	0.15	0.11	0.07	0.04	0.03	0.02	0.00	−3.16
0.25	0.40	0.34	0.27	0.24	0.17	0.12	0.07	0.05	0.02	0.00	−2.00
0.75	0.74	0.63	0.49	0.42	0.30	0.20	0.11	0.08	0.04	0.01	−1.15
1.00	0.89	0.75	0.58	0.50	0.35	[illegible]	0.13	0.09	0.05	0.01	−1.00
2.00	1.50	1.20	0.88	0.74	0.52	[illegible]	0.19	0.13	0.07	0.01	−0.71
2.50	1.83	1.43	1.03	0.86	0.59	[illegible]	0.21	0.14	0.08	0.02	−0.63
3.00	2.20	1.68	1.17	0.97	0.65	0.42	0.23	0.16	0.08	0.02	−0.58
3.50	2.61	1.94	1.32	1.08	0.71	0.46	0.25	0.17	0.09	0.02	−0.53
4.00	3.08	2.23	1.47	1.19	0.78	0.49	0.27	0.18	0.10	0.02	−0.50
5.00	4.23	2.88	1.79	1.42	0.90	0.56	0.30	0.21	0.11	0.02	−0.45
10.00	18.87	9.26	4.09	2.83	1.50	0.85	0.43	0.29	0.15	0.03	−0.32
15.00	82.13	28.40	8.59	5.11	2.20	1.13	0.54	0.36	0.19	0.04	−0.26
20.00	357.05	86.73	17.73	8.98	3.09	1.41	0.64	0.42	0.22	0.04	−0.22
25.00	*	264.70	36.44	15.62	4.24	1.71	0.73	0.48	0.24	0.05	−0.20

*Greater than 999.99.

If $b < 1$, then $[b] = 0$, so that

$$x = \exp\{c + a[-B \ln u]\} + v \qquad \text{(A13)}$$

Equation (A11) presents an algorithm which is faster computationally than (A12). However, for b large it becomes impossible to use (A11) on a computer, since the product of the u_k rapidly becomes too small to express on the machine. This problem is avoided by using the sum of the logarithms rather than the logarithm of the product of the u_k.

APPENDIX B: COMMENTS ON WAKEBY DISTRIBUTION

Definition

The Wakeby distribution (H. A. Thomas, personal communication, 1976) is defined in the following manner. Let X be a random variable such that

$$x = m + a[1 - (1 - F)^b] - c[1 - (1 - F)^{-d}] \qquad \text{(B1)}$$

where $F \equiv F(x) = P(X \leq x)$ and $x \geq m$. The density function $f \equiv f(x)$ is defined as

$$f = dF/dx = [ab(1 - F)^{b-1} + cd(1 - F)^{-d-1}]^{-1} \qquad \text{(B2)}$$

If $F = 0$, then $x = m$, and $f = 1/(ab + cd)$. Note that since $f \geq 0 \ \forall\ x$, $(ab + cd) \geq 0$. For $F = 1$ the values of x and f depend upon the values of the parameters of the distribution, the upper bound on x being either $+\infty$ or $(m + a - c)$. Furthermore, the definition of F precludes certain parameter combinations, as is shown in Table B1.

Moments

The moments are defined as $E[X^k] = \int_0^1 x^k\, dF$, where x is given by (B1). $E[X^k]$ cannot be computed if $d \geq 1/k$ or if $b \leq -1/k$, since the integral would not be defined properly. Thus if $d \geq 1/k$ or if $b \leq -1/k$, then the moments of order k and higher do not exist. That is, the mean, $\mu[x]$, exists only if $d < 1$ and $b > -1$; the standard deviation, $\sigma[x]$, exists if $d < \frac{1}{2}$ and $b > -\frac{1}{2}$; the coefficient of skew, $\gamma[x]$, exists if $d < \frac{1}{3}$ and $b > -\frac{1}{3}$; and the coefficient of kurtosis, $\lambda[x]$, exists if $d < \frac{1}{4}$ and $b > -\frac{1}{4}$.

To allow concise tabular presentation of these properties

TABLE A3. Coefficient of Variation in RS (X) and LS (Z) for Log Pearson Distribution With $m = 0$ and $c = 0$

	$C_v[X]$										
b	$a =$ 0.01	$a =$ 0.05	$a =$ 0.09	$a =$ 0.10	$a =$ 0.15	$a =$ 0.18	$a =$ 0.22	$a =$ 0.24	$a =$ 0.28	$a =$ 0.32	$C_v[Z]$
0.05	0.00	0.01	0.02	0.02	0.04	0.05	0.06	0.07	0.09	0.11	4.47
0.10	0.00	0.02	0.03	0.04	0.06	0.07	0.09	0.10	0.13	0.16	3.16
0.25	0.01	0.03	0.05	0.06	0.09	0.11	0.14	0.16	0.20	0.25	2.00
0.75	0.01	0.05	0.09	0.10	0.15	0.19	0.25	0.29	0.36	0.45	1.15
1.00	0.01	0.05	0.10	0.11	0.18	0.23	0.29	0.33	0.42	0.53	1.00
2.00	0.01	0.07	0.14	0.16	0.26	0.32	0.42	0.48	0.62	0.81	0.71
2.50	0.02	0.08	0.16	0.18	0.29	0.36	0.48	0.55	0.71	0.93	0.63
3.00	0.02	0.09	0.17	0.19	0.32	0.40	0.53	0.61	0.80	1.06	0.58
3.50	0.02	0.10	0.19	0.21	0.34	0.43	0.58	0.67	0.88	1.18	0.53
4.00	0.02	0.11	0.20	0.23	0.37	0.47	0.63	0.72	0.96	1.31	0.50
5.00	0.02	0.12	0.22	0.25	0.41	0.53	0.72	0.83	1.13	1.58	0.45
10.00	0.03	0.17	0.32	0.36	0.61	0.80	1.14	1.36	2.04	3.35	0.32
15.00	0.04	0.21	0.40	0.45	0.78	1.05	1.57	1.96	3.27	6.46	0.26
20.00	0.05	0.24	0.47	0.53	0.94	1.30	2.06	2.68	5.06	12.18	0.22
25.00	0.05	0.27	0.53	0.60	1.10	1.56	2.64	3.58	7.70	22.83	0.20

TABLE A4. Coefficient of Skew in RS (X) and LS (Z) for Log Pearson Distribution

	$\gamma(X)$										
b	$a =$ −2.00	$a =$ −1.50	$a =$ −1.00	$a =$ −0.80	$a =$ −0.50	$a =$ −0.30	$a =$ −0.15	$a =$ −0.10	$a =$ −0.05	$a =$ −0.01	$\gamma(Z)$
0.05	−3.81	−4.16	−4.70	−5.02	−5.73	−6.50	−7.40	−7.81	−8.32	−8.81	−8.94
0.10	−2.49	−2.75	−3.17	−3.41	−3.94	−4.51	−5.18	−5.49	−5.86	−6.22	−6.32
0.25	−1.17	−1.38	−1.69	−1.88	−2.27	−2.69	−3.18	−3.40	−3.67	−3.93	−4.00
0.75	0.22	−0.02	−0.34	−0.52	−0.88	−1.24	−1.64	−1.82	−2.04	−2.25	−2.31
1.00	0.64	0.36	−0.00	−0.19	−6.57	−0.93	−1.34	−1.52	−1.73	−1.94	−2.00
2.00	2.13	1.59	0.97	0.69	0.19	−0.25	−0.70	−0.90	−1.12	−1.35	−1.41
2.50	2.97	2.19	1.39	1.04	0.45	−0.04	−0.52	−0.73	−0.97	−1.20	−1.26
3.00	3.94	2.86	1.80	1.37	0.68	0.14	−0.38	−0.59	−0.85	−1.09	−1.15
3.50	5.12	3.60	2.23	1.70	0.90	0.30	−0.25	−0.48	−0.75	−1.00	−1.07
4.00	6.55	4.46	2.68	2.04	1.11	0.44	−0.15	−0.39	−0.66	−0.93	−1.00
5.00	10.55	6.64	3.71	2.76	1.50	0.70	0.03	−0.23	−0.53	−0.82	−0.89
10.00	108.33	42.81	14.18	8.59	3.67	1.75	0.62	0.24	−0.16	−0.53	−0.63
15.00	*	276.41	51.30	23.91	6.98	2.80	1.04	0.55	0.05	−0.40	−0.52
20.00	*	*	187.98	66.91	12.70	4.03	1.41	0.79	0.20	−0.31	−0.45
25.00	*	*	693.40	189.13	23.01	5.58	1.76	1.00	0.32	−0.25	−0.40

*Greater than 999.99.

(see Tables B1–B29 of the microfiche supplement[1]), the parameter c is expressed as a function of a; i.e., $c = ra$, where $r = c/a$. The definition of WA as given in (B1) is rewritten as

$$x = m + a\{[1 - (1 - F)^b] - r[1 - (1 - F)^{-d}]\} \quad \text{(B3)}$$

and the aforementioned statistical properties are given in Table B2.

From Table B2 it is seen that the statistical properties are functions of all four parameters a, b, $r(c)$, and d. Property $\mu[X]$ is also a function of m. However, if $m = 0$, then $C_v[X] \equiv \sigma[X]/\mu[X]$ is only a function of b, d, and r. Regardless of the value of m, $\gamma[X]$ and $\lambda[X]$ are only functions of b, d, and r.

If $\mu[X]$, $\sigma[X]$, and m are known, then a and b can be derived explicitly as functions of c and d. Similarly, c and d can be derived explicitly as functions of a and b plus $\mu[X]$, $\sigma[X]$, and m; a and c can be obtained as functions of b and d, given $\mu[X]$, $\sigma[X]$, and m. However, given a, c, $\mu[X]$, $\sigma[X]$, and m, explicit expressions for b and d cannot be readily obtained.

The value of r is a function, explicitly, of m, a, b, and d, as well as just c and a:

$$r = \frac{c}{a} = \frac{m/a - [1 - F(a)]^b}{1 - [1 - F(a)]^{-d}} \quad \text{(B4)}$$

so that if $m = 0$,

$$r = \frac{[1 - F(a)]^{b+d}}{1 - [1 - F(a)]^d} \quad \text{(B5)}$$

Special Cases

Three special cases of the WA distribution are highlighted, i.e., (1) $a = 0$, (2) $b = \infty$, and (3) $c = 0$, with not more than one of these conditions holding. $C_v[X]$, $\gamma[X]$, and $\lambda[X]$ for these cases are given in Table B3.

Note that if $b = 0$, the distribution is equivalent to that of case 1, and if $d = 0$, it is the same as case 3. If $d = \infty$, the

[1]These tables are available with the entire article on microfiche. Order from American Geophysical Union, 1909 K Street, N. W., Washington, D. C. 20006. Document W78-007; $1.00. Payment must accompany order.

TABLE A5. Coefficient of Skew in RS (X) and LS (Z) for Log Pearson Distribution

	$\gamma(X)$										
b	$a =$ 0.01	$a =$ 0.05	$a =$ 0.09	$a =$ 0.10	$a =$ 0.15	$a =$ 0.18	$a =$ 0.22	$a =$ 0.24	$a =$ 0.28	$a =$ 0.32	$\gamma(Z)$
0.05	9.09	9.75	10.59	10.83	12.38	13.69	16.25	18.15	24.86	50.10	8.94
0.10	6.43	6.92	7.54	7.72	8.87	9.85	11.76	13.19	18.25	37.90	6.32
0.25	4.08	4.43	4.88	5.01	5.84	6.55	7.96	9.02	12.89	29.29	4.00
0.75	2.37	2.66	3.02	3.13	3.81	4.41	5.64	6.61	10.46	32.76	2.31
1.00	2.06	2.34	2.70	2.81	3.50	4.10	5.37	6.39	10.61	39.60	2.00
2.00	1.48	1.78	2.17	2.28	3.04	3.75	5.32	6.69	13.56	112.34	1.41
2.50	1.33	1.65	2.05	2.17	2.99	3.75	5.53	7.15	16.06	209.13	1.26
3.00	1.23	1.55	1.98	2.11	2.97	3.81	5.81	7.73	19.32	406.63	1.15
3.50	1.14	1.48	1.93	2.07	2.99	[illegible]	6.16	8.41	23.54	817.57	1.07
4.00	1.08	1.43	1.90	2.04	3.02	4.01	6.55	9.21	28.95	*	1.00
5.00	0.98	1.36	1.86	2.02	3.12	4.28	7.49	11.14	44.90	*	0.89
10.00	0.74	1.24	1.93	2.16	3.95	6.28	15.81	32.79	562.20	*	0.63
15.00	0.64	1.24	2.12	2.42	5.09	9.44	36.18	111.49	*	*	0.52
20.00	0.59	1.28	2.33	2.71	6.54	14.36	88.33	417.56	*	*	0.45
25.00	0.56	1.33	2.57	3.03	8.41	22.18	226.70	*	*	*	0.40

*Greater than 999.99.

TABLE A6. Coefficient of Kurtosis in RS (X) and LS (Z) for Log Pearson Distribution

	λ(X)										
b	a = −2.00	a = −1.50	a = −1.00	a = −0.80	a = −0.50	a = −0.30	a = −0.15	a = −0.10	a = −0.05	a = −0.01	λ(Z)
0.05	17.73	21.18	27.40	31.58	42.23	56.20	76.52	87.41	102.12	118.16	123.00
0.10	8.42	10.17	13.35	15.49	20.98	28.19	38.74	44.40	52.08	60.47	63.00
0.25	3.04	3.74	5.05	5.95	8.29	11.43	16.09	18.61	22.06	25.85	27.00
0.75	1.66	1.66	1.90	2.15	2.94	4.14	6.09	7.19	8.73	10.46	11.00
1.00	2.14	1.87	1.80	1.90	2.40	3.31	4.87	5.78	7.07	8.54	9.00
2.00	7.56	5.06	3.15	2.60	2.15	2.33	3.14	3.72	4.59	5.66	6.00
2.50	13.32	8.19	4.52	3.46	2.42	2.28	2.85	3.33	4.11	5.08	5.40
3.00	22.76	12.87	6.38	4.62	2.82	2.33	2.68	3.09	3.79	4.70	5.00
3.50	38.22	19.83	8.88	6.10	3.34	2.46	2.59	2.93	3.56	4.42	4.71
4.00	63.58	30.15	12.17	7.96	3.96	2.64	2.55	2.83	3.40	4.22	4.50
5.00	174.06	68.36	22.25	13.23	5.53	3.10	2.54	2.71	3.18	3.93	4.20
10.00	*	*	386.71	134.66	24.42	7.30	3.23	2.79	2.83	3.36	3.60
15.00	*	*	*	*	94.90	15.72	4.41	3.21	2.82	3.18	3.40
20.00	*	*	*	*	367.78	32.10	5.98	3.77	2.90	3.09	3.30
25.00	*	*	*	*	*	64.34	7.98	4.44	3.01	3.04	3.24

*Greater than 999.99.

degenerate distribution, $x = \infty \;\forall\; F > 0$ and $x = m$ if $F = 0$, results. However, if $b = \infty$, the WA definition is almost the same as that in case 1, except that a is included in the definition of the lower bound. As $a \to 0$, $C_v[X]$ for case 2 approaches that of case 1; $\gamma[X]$ and $\lambda[X]$ are always the same for these two cases.

Since the $C_v[X]$, $\gamma[X]$, and $\lambda[X]$ for these three cases denote bounds for WA statistical properties, they are important for tabular presentation of information about WA (see Tables B4–B6).

Consider case 1 with $m = 0$ and $d > 0$, hence $c > 0$. The range of X is from 0 ($F = 0$) to ∞ ($F = 1$), and the probability density function (pdf), $f \equiv dF/dx$, is

$$f(x) = \frac{1}{cd}\left[\frac{c}{x + c}\right]^{(1+d)/d} \qquad \text{(B6)}$$

where

$$F(x) = 1 - \left[\frac{c}{x + c}\right]^{1/d} \qquad \text{(B7)}$$

for all d and $f(x)$ is a monotonically decreasing function with $f(0) = 1/cd$.

Consider case 3 with $m = 0$ and $b > 0$. The range of x is from 0 ($F = 0$) to a ($F = 1$), and the probability density function, $f \equiv dF/dx$, is

$$f(x) = \frac{1}{ab}\left(\frac{a - x}{a}\right)^{(1-b)/b} \qquad \text{(B8)}$$

where

$$F(x) = 1 - \left(\frac{a - x}{a}\right)^{1/b} \qquad \text{(B9)}$$

If $b = 1$, then X is uniformly distributed. If $0 < b < 1$, then $f(x)$ is a monotonically decreasing function with its maximal value $1/ab$ at the lower bound, $x = 0$. However, if $b > 1$, then $f(x)$ is J shaped, approaching infinity at the upper bound of x, a.

Note that the definition of WA can be considered a combination of cases 1 and 3. Thus the value of the parameter b

TABLE A7. Coefficient of Kurtosis in RS (X) and LS (Z) for Log Pearson Distribution

	λ(X)								
b	a = 0.01	a = 0.05	a = 0.09	a = 0.10	a = 0.15	a = 0.18	a = 0.22	a = 0.24	λ(Z)
0.05	128.26	154.67	194.44	207.67	311.79	438.70	905.09	*	123.00
0.10	65.75	79.62	100.58	107.57	162.85	230.78	484.29	*	63.00
0.25	28.25	34.60	44.30	47.56	73.71	106.63	235.30	541.09	27.00
0.75	11.59	14.62	19.41	21.06	34.95	53.83	139.98	415.98	11.00
1.00	9.50	12.13	16.36	17.83	30.50	48.37	136.57	467.33	9.00
2.00	6.38	8.44	11.97	13.25	25.28	44.85	178.49	*	6.00
2.50	5.76	7.72	11.18	12.47	25.03	46.90	224.15	*	5.40
3.00	5.34	7.25	10.72	12.03	25.36	50.19	290.61	*	5.00
3.50	5.04	6.92	10.43	11.79	26.06	54.51	385.82	*	4.71
4.00	4.82	6.68	10.26	11.66	27.04	59.79	522.10	*	4.50
5.00	4.51	6.47	10.13	11.65	29.64	73.44	*	*	4.20
10.00	3.89	5.94	11.06	13.45	54.23	252.17	*	*	3.60
15.00	3.69	6.03	12.97	16.62	108.34	*	*	*	3.40
20.00	3.60	6.25	15.46	20.86	227.47	*	*	*	3.30
25.00	3.55	6.54	18.52	26.33	498.37	*	*	*	3.24

*Greater than 999.99.

696 095

TABLE B1. Valid and Invalid Parameter Combinations for Wakeby Distribution

Sign of Parameter				Valid Distribution?		
a	b	c	d	Yes	Maybe	No
+	+	+	+	X		
−	+	+	+		*	
+	+	−	+			X
−	+	−	+			X
+	+	+	−		†	
−	+	+	−			X
+	+	−	−	X		
−	+	−	−		†	
+	−	+	+		‡	
−	−	+	+	X		
+	−	−	+			X
−	−	−	+		§	
+	−	+	−			X
−	−	+	−		*	
+	−	−	−			X
−	−	−	−	X		

*Valid if $ab + cd > 0$, i.e., valid pdf.

†Valid if asterisked footnote holds and $a > c$.

‡Valid if asterisked footnote holds and either $|b| < |d|$ or $c > a$ when $|b| = |d|$.

§Valid if asterisked footnote holds and either $|b| > |d|$ or $a > c$ when $|b| = |d|$.

strongly influences the shape of the distribution. If $0 < b < 1$, then the probability density function of X distributed as WA is a monotonically decreasing function; if $1 < b$, then $f(x)$ has a mode at some value other than the lower bound.

Tabular Presentation of $C_v[X]$, $\gamma[X]$, and $\lambda[X]$

Tables B1–B29 of the microfiche supplement present values of $C_v[X]$ (with $m = 0$), $\gamma[X]$, and $\lambda[X]$ for $0 \le d$, $0 \le b \le \infty$, and $0 \le r \le \infty$ ($c > 0$). As was noted in the previous section, special cases 1–3 form the bounds on the values in these tables. The following specific observations are also made.

1. Property $\mu[X]$ exists for $0 \le d < 1$ and $b > 0$.
2. Property $\sigma[X]$ exists for $0 \le d < 0.5$ and $b > 0$.
3. Property $\gamma[X]$ exists for $0 \le d < \frac{1}{3}$. (1) For $0 \le b \le 1.0$, $\gamma[X] > 0$. (2) For $1.0 < b \le \infty$, $\gamma[X] \lesseqgtr 0$, and as r decreases, the negative skew region increases.
4. Property $\lambda[X]$ exists for $0 < d < 0.25$ and $b > 0$.
5. If r and $0 < d$ are held constant and b is allowed to increase, in general, (1) there exists a minimum value for $\gamma[X]$, (2) there exist local minima for $\lambda[X]$, (3) the b value at which the minimum skew occurs need not be the b at which the local minimum kurtoses occur, (4) as d increases for a given r, the b value at which the minimum skew occurs decreases (the same is true if r decreases for a given d), (5) similarly, as d increases for a given r, the b values at which the local minimum kurtoses occur decrease (the same is true if r decreases for a given d), and (6) $C_v[X]$ decreases.
6. If r and b are fixed and d increases in value, in general, (1) as d approaches $\frac{1}{3}$, $\gamma[X]$ approaches ∞, (2) $\lambda[X]$ increases, and (3) $C_v[X]$ increases.
7. If b and d are fixed and r increases, (1) $C_v[X]$ increases, (2) $\gamma[X]$ increases, and (3) $\lambda[X]$ decreases to a minimum and then increases.

Extreme Points

To find the extreme points of f, that is, the values of x at which the derivative of f vanishes, df/dx is set equal to 0, so that

$$\frac{df}{dx} = \frac{ab(b-1)(1-F)^{b-2} + cd(-d-1)(1-F)^{-d-2}}{[ab(1-F)^{b-1} + cd(1-F)^{-d-1}]^3} = 0 \tag{B10}$$

This will occur if the numerator is zero, and the numerator will be zero if any one of the following three conditions holds: (1) $F(x) = 1$. (2) Either $ab = 0$ or $b = 1$, and either $cd = 0$ or $d = -1$. (3) $ab(b-1)(1-F)^{b-2} = cd(d+1)(1-F)^{-d-2}$; that is,

$$(1-F)^{b+d} = \frac{cd(d+1)}{ab(b-1)} \tag{B11}$$

Since the roots of this equation should be real and since $F(x) \le 1$, if an extreme point is to occur, then

$$1 \ge \frac{cd(d+1)}{ab(b-1)} > 0 \tag{B12}$$

TABLE B2. Statistical Properties of X Distributed as WA

Property	Definition
$\mu[X]$	$m + a\left(\frac{b}{1+b} + \frac{rd}{1-d}\right)$
$\sigma[X]$	$\lvert a\rvert\left\{\frac{b^2}{(1+b)^2(1+2b)} + \frac{r^2d^2}{(1-d)^2(1-2d)} + \frac{2rbd}{(1+b)(1-d)(1+b-d)}\right\}^{1/2}$
$\gamma[X]$	$\frac{E[V^3] - 3E[V]E[V^2] + 2E^3[V]}{\sigma^3[X]}$
$\lambda[X]$	$\frac{E[V^4] - 4E[V]E[V^3] + 6E^2[V]E[V^2] - 3E^4[V]}{\sigma^4[X]}$
V	$x - a(1-r) - m$
$E[V^k]$	$a^k \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i} \frac{r^i}{1 + (k-i)b - id}$

696 096

TABLE B3. $C_v[X]$, $\gamma[X]$, and $\lambda[X]$ for X Distributed as WA Special Cases

	Case 1, $a = 0$	Case 2, $b = \infty$	Case 3, $c = 0$
X	$m - c[1 - (1 - F)^{-d}]$	$m + a - c[1 - (1 - F)^{-d}]$	$m + a[1 - (1 - F)^{b}]$
$C_v[X]$ $(m = 0)$	$\left(\frac{1}{1 - 2d}\right)^{1/2}$	$\left(\frac{1}{1 - 2d}\right)^{1/2} \frac{d/(1 - d)}{(a/c) + d/(1 - d)}$	$\left(\frac{1}{1 + 2b}\right)^{1/2}$
$\gamma[X]$	$\frac{(1 - 2d)^{3/2}}{d^3}\left\{\frac{(1 - d)^3}{(1 - 3d)} - 3\frac{(1 - d)^3}{(1 - 2d)} + 2\right\}$		$\frac{(1 + 2b)^{3/2}}{b^3}\left\{\frac{-(1 + b)^3}{(1 + 3b)} + 3\frac{(1 + b)^3}{(1 + 2b)} - 2\right\}$
$\lambda[X]$	$\frac{(1 - 2d)^2}{d^4}\left\{\frac{(1 - d)^4}{(1 - 4d)} - 4\frac{(1 - d)^4}{(1 - 3d)} + 6\frac{(1 - d)^4}{(1 - 2d)} - 3\right\}$		$\frac{(1 + 2b)^2}{b^4}\left\{\frac{(1 + b)^4}{(1 + 4b)} - 4\frac{(1 + b)^4}{(1 + 3b)} + 6\frac{(1 + b)^4}{(1 + 2b)} - 3\right\}$

If condition (B12) is satisfied, then

$$F^* = 1 - \left[\frac{cd(d + 1)}{ab(b - 1)}\right]^{1/(b+d)} \quad \text{(B13)}$$

The first condition must be treated separately, since it results in an indeterminate form for df/dx. Note that $F = 1$ when x attains its upper bound. If the second condition occurs, either $f(x)$ is improperly defined, or X is uniformly distributed and has no extreme point. Therefore if condition (B12) is satisfied, there is generally only one finite value of x at which an extreme point can occur, and that point is

$$x^* = m + a\left[1 - \left(\frac{c}{a}\,\frac{d(1 + d)}{b(b - 1)}\right)^{b/(b+d)}\right] - c\left[1 - \left(\frac{c}{a}\,\frac{d(1 + d)}{b(b - 1)}\right)^{-d/(b+d)}\right] \quad \text{(B14)}$$

Note that if a, c, and $d > 0$, then it is necessary (although not sufficient) that $b > 1$ if (B12) is to be satisfied. If $0 < b \leq 1$ and a, c, and $d > 0$, it can be shown that $f(x)$ will be a monotonically decreasing function such that the maximum value of $f(x)$ occurs at $x = m$, where $f(m) = 1/(ab + cd)$. Let $m \leq x_1 < x_2$; then $0 = F(m) \leq F(x_1) \leq F(x_2) \leq 1$. Since $d > 0$, then $1 + d > 1$, and

$$[1 - F(m)]^{-(1+d)} \leq [1 - F(x_1)]^{-(1+d)} \leq [1 - F(x_2)]^{-(1+d)} \quad \text{(B15)}$$

Since $0 < b < 1$, then $-1 < b - 1 < 0$, and

$$[1 - F(m)]^{b-1} \leq [1 - F(x_1)]^{b-1} \leq [1 - F(x_2)]^{b-1} \quad \text{(B16)}$$

so that since $a, b, c, d > 0$,

$$ab[1 - F(m)]^{b-1} + cd[1 - F(m)]^{-(1+d)}$$
$$\leq ab[1 - F(x_1)]^{b-1} + cd[1 - F(x_1)]^{-(1+d)}$$
$$\leq ab[1 - F(x_2)]^{b-1} + cd[1 - F(x_2)]^{-(1+d)} \quad \text{(B17)}$$

but

$$f(X) = [ab(1 - F)^{b-1} + cd(1 - F)^{-d-1}]^{-1}$$

so that $f(m) \geq f(x_1) \geq f(x_2)$ when $m \leq x_1 < x_2$.

Inflection Points

If $f(x)$ has inflection points, the second derivative of $f(x) \equiv f$ is zero at these points. The second derivative is

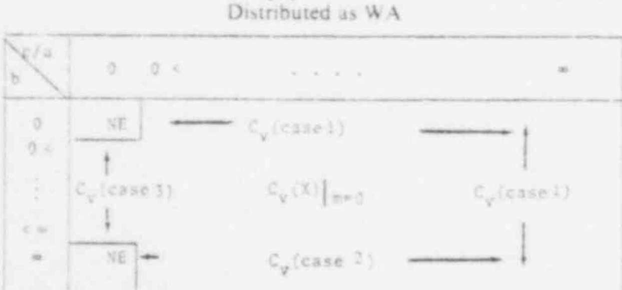
$$d^2f/(dx)^2 = [ab(1 - F)^{b-1} + cd(1 - F)^{-d-1}]^{-5}$$
$$\cdot[(ab)^2(b - 1)(2b - 1)(1 - F)^{2b-4}$$
$$- abcd(b^2 + d^2 + 6bd + 3b - 3d - 2)(1 - F)^{b-d-4}$$
$$+ (cd)^2(d + 1)(2d + 1)(1 - F)^{-2d-4}] \quad \text{(B18)}$$

For F to be defined properly, only F values between 0 and 1 are considered. Furthermore, unless the sign of the second derivative changes at F, the value will not denote an inflection point. The slope of f when $F = 0$, that is, $x = m$, is

$$\left.\frac{df}{dx}\right|_{x=m} \equiv f'(m) = \frac{ab(b - 1) - cd(d + 1)}{(ab + cd)^3} \quad \text{(B19)}$$

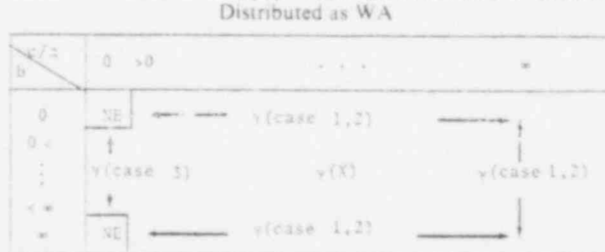
If a, b, c, and $d > 0$, so that $ab + cd > 0$, the following conjecture is proposed. (1) If $ab(b - 1) < cd(d + 1)$, then $f'(m) < 0$, and no inflection points exist. (2) If $ab(b - 1) = cd(d + 1)$, then $f'(m) = 0$, and the extreme point occurs at $x^* = m$. (3) If $ab(b - 1) \geq cd(d + 1)$, then $f'(m) \geq 0$, and there exists at least one inflection point. (4) If $b < 1$, then $f'(m) < 0$, f is 'J shaped', and there is no extreme point x^*, although the peak occurs at $x = m$.

TABLE B4. Bounds on $C_v[X]$ Values for Fixed d, $c > 0$, and X Distributed as WA



NE -- does not exist

TABLE B5. Bounds on $\gamma[X]$ Values for Fixed d, $c > 0$, and X Distributed as WA



NE -- does not exist

696 097

TABLE B6. Bounds on $\lambda[X]$ Values for Fixed d, $c > 0$, and X Distributed as WA

b \ c/a	0	0 <	. . .	∞
0	NE	←	λ(case 1,2) →	↑
0 <	↑			
⋮	λ(case 3)		λ(X)	λ(case 1,2)
< ∞	↓			
∞	NE	←	λ(case 1,2) →	↓

NE -- does not exist

Wakeby Random Numbers

To generate random variables distributed as WA, let $1 - F(x) = u$, where u is distributed uniformly on (0, 1). Then

$$x = m + a(1 - u^b) - c(1 - u^{-d}) \quad \text{(B20)}$$

Appendix C: Comments on the Weibull Distribution

Definition

Let X be a random variable distributed as Weibull (W), so that the probability density function of X is

$$f(x) = \frac{b}{a}\left(\frac{x-m}{a}\right)^{b-1} \exp\left[-\left(\frac{x-m}{a}\right)\right]^b \quad \text{(C1)}$$

where $|a| > 0$ and $b > 0$.

Moments

By definition, the moment of order k, $E[X^k]$, is

$$E[X^k] = \int_m^\infty x^k f(x)\,dx \quad \text{(C2)}$$

so that for X distributed as W,

$$E[X^k] = \sum_{i=0}^{k} \binom{k}{i} m^i a^{k-i} \Gamma(k + b - i) \quad \text{(C3)}$$

where $E[X^k] < \infty \ \forall\ k \geq 0$.

The mean, $\mu[X]$, standard deviation, $\sigma[X]$, and coefficients of skewness, $\gamma[X]$, kurtosis, $\lambda[X]$, and variation, $C_v[X]$, are given in Table C1.

Define $\theta(b) = C_v^{-1}[X]$ when $m = 0$. It is noted that m, the lower bound on X, can be defined as

$$m = \mu[X] - a\Gamma(1 + 1/b) = \mu[X] - \sigma[X]\theta(b) \quad \text{(C4)}$$

It is also noted that the sign of $\gamma[X]$ is the same as that of a.

Weibull Random Numbers

To generate Weibull-distributed random numbers, the following algorithm can be used:

$$x = m + a[-\ln(1 - u)]^{1/b} \quad \text{(C5)}$$

where u is distributed as uniform (U) on [0, 1]. For the special case where $m = 0$ and $a > 0$,

$$\ln x = \ln a + (1/b) \ln[-\ln(1 - u)] \quad \text{(C6)}$$

Special Case

If a random variable Z is distributed as Gumbel (G) with a probability density function

$$f(z) = \frac{1}{a'} \exp\left[-\left(\frac{z - m'}{a'}\right)\right] \cdot \exp\left\{-\exp\left[-\left(\frac{z - m'}{a'}\right)\right]\right\} \quad \text{(C7)}$$

then random numbers distributed accordingly can be generated using the algorithm

$$z = m' + a'(-\ln[-\ln u']) \quad \text{(C8)}$$

where $u \simeq U(0, 1)$. It is noted that if (1) $u' \equiv 1 - u$, (2) $m' \equiv \ln a$, and (3) $a' \equiv -1/b$, then Z, distributed as G, equals $\ln X$, for X distributed as W with $m = 0$ and $a > 0$.

The distribution G has a skew equal to ± 1.14, depending upon the sign of the parameter a'. Since the parameter a' is set opposite in sign to that of parameter $b > 0$, $\gamma(Z) = -1.14$. Furthermore, since the sign of $\gamma(X)$ for X distributed as W is equal to that of $a > 0$, then $\gamma(X) > 0$. Therefore if a random variable X is distributed as W with a lower bound of 0 and is positively skewed, then $Z = \ln X$ is distributed as G with a skew of $-1.14 \ \forall\ \gamma(X) > 0$.

It is further noted that if the Weibull parameter $b = 1$, X is also distributed as Pearson type III (PIII) with the PIII parameter $b = 1$ and $\gamma(X) = 2$. (Refer to Appendix A for the form of the PIII distribution.) Thus if a random variable X is distributed as PIII with $\gamma(X) = 2$ and lower bound 0, then $Z = \ln X$ is distributed as G with a skew of -1.14.

Notation

RS	real space.
LS	log space.
X	random variable in RS.
x	variate value of X.
ln	logarithm to base e.
Z	random variable in LS.
z	variate value of Z.
C_v	coefficient of variation.
γ	coefficient of skewness.
λ	coefficient of kurtosis.
n	length of historical or simulated flood sequence.
$k(n)$	number of historical flood sequences of length n.
Y	equivalent to C_v, γ, or λ.
y	moment estimate of Y derived from a sequence length n.
$\bar{y}$	mean value of y.
$\hat{\sigma}(y)$	standard deviation of y.

TABLE C1. Statistical Properties of X Distributed as W

Property	Definition
$\mu[X]$	$m + a\Gamma(1 + 1/b)$
$\sigma[X]$	$\{a^2[\Gamma(1 + 2/b) - \Gamma^2(1 + 1/b)]\}^{1/2}$
$\gamma[X]$	$a^3[\Gamma(1 + 3/b) - 3\Gamma(1 + 2/b)\Gamma(1 + 1/b) + 2\Gamma^3(1 + 1/b)]\sigma^{-3}[X]$
$\lambda[X]$	$a^4[\Gamma(1 + 4/b) - 4\Gamma(1 + 1/b)\Gamma(1 + 3/b) + 6\Gamma^2(1 + 1/b)\Gamma(1 + 2/b) - 3\Gamma^4(1 + 1/b)]\sigma^{-4}[X]$
$C_v[X]$	$\lvert a \rvert\,[\Gamma(1 + 2/b) - \Gamma^2(1 + 1/b)]^{1/2}[m + a\Gamma(1 + 1/b)]^{-1}$

696 098

$Q(\ )$ proportion of historical flood sequences yielding estimates of γ of like or opposite signs in RS and LS.
ω moment estimate of γ derived from simulated flood sequence of length n.
$\hat{\mu}(\omega)$ mean of ω.
$\hat{\sigma}(\omega)$ standard deviation of ω.
Φ set of distribution functions.
ϕ element of Φ.
U uniform distribution.
N normal distribution.
LN log normal distribution.
G Gumbel extreme value type I distribution.
PIII Pearson type III distribution.
W Weibull distribution.
P Pareto distribution.
LP log Pearson type III distribution.
WA Wakeby distribution.
u variate value of U.
e, a, b, c, d, m parameters of distribution functions.
$f \equiv f(x)$ probability density function.
$F \equiv F(x)$ cumulative distribution function.
$\Gamma[\ ]$ gamma function.
$\theta(b)$ reciprocal of C_v for W with lower bound $m = 0$.
X' transform of the random variable X.
Z' transform of the random variable Z.
$E[X^k]$ kth-order moment of X.
V random variable (Appendix B).

REFERENCES

Berman, M. B., Generating gamma distributed variates for computer simulation models, *Rep. R-641-PR*, 43 pp., Rand Corp., Santa Monica, Calif., Feb. 1971.

Houghton, J. C., Robust estimation of the frequency of extreme events in a flood frequency context, Ph.D. dissertation, Harvard Univ., Cambridge, Mass., 1977.

Johnk, M. D., Evzeugung von betaverteilten und gamma-verteilten Zufallszahlem, *Metrika*, 8(1), 5–15, 1964.

Matalas, N. C., J. R. Slack, and J. R. Wallis, Regional skew in search of a parent, *Water Resour. Res.*, 11(6), 815–826, 1975.

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

Wallis, J. R., N. C. Matalas, and J. R. Slack, Apparent regional skew, *Water Resour. Res.*, 13(1), 159–182, 1977.

(Received June 15, 1977;
revised February 27, 1977;
accepted March 22, 1978.)

696 099