No. HY5, Paper No. 1808, Oct. 1958.
2. D. M. Thomas and M. A. Benson. Generalization of Streamflow Characteristics From Drainage-Basin Characteristics. U.S. Geological Survey, Water Supply Paper 1975, 1970.
3. B. O. Benn. Regional Planning Potential of Deterministic Hydrologic Simulation Models. Proc., Seminar on Hydrologic Aspects of Project Planning, HEC, Corps of Engineers, Davis, CA, March 7-9, 1972, pp. 13-26.
4. L. B. Leopold and T. Maddock, Jr. The Hydraulic Geometry of Stream Channels and Some Physiographic Implications. U.S. Geological Survey Professional Paper 252, 1953.
5. J. F. Orsborn. A Geomorphic Method for Estimating Low Flows. ASCE, Annual Meeting, Denver, Nov. 3-7, 1975.
6. J. F. Orsborn and M. N. Sood. Technical Supplement to the Hydrographic Atlas. Washington Department of Ecology, State Water Program, Lewis River Basin Study Area, 1973.
7. H. C. Riggs. The Relation of Discharge to Drainage Area in the Rappahannock River Basin, Virginia. U.S. Geological Survey, Professional Paper 501B, 1964, pp. B165-B168.
8. J. F. Orsborn and others. Relationships Between Low, Average and Flood Flows in the Pacific Northwest. Department of Civil and Environmental Engineering, Washington State Univ., Pullman, OWRT Project A-074-WASH, 1975.
9. J. F. Orsborn and others. Surface Water Resources of the Coeur d'Alene, St. Joe and St. Maries Rivers in Northern Idaho. In Preliminary Investigation of the Water Resources of the Northern Part of the Coeur d'Alene Indian Reservation, Department of Civil and Environmental Engineering, Washington State Univ., Pullman, 1975, pp. 160-197.
10. H. Conturk. Mean Discharge as an Index to Mean Maximum Discharge. Proc., Leningrad Symposium on Floods and Their Computation, IASH-UNESCO-WHO, Vol. 2, 1967, pp. 826-833.
11. H. P. Chrostowski. Stream Habitat Studies on the Uinta and Ashley National Forests. U.S. Forest Service, Intermountain Region, Ogden, UT, 1972.
12. E. R. Hedman and W. M. Kastner. Progress Report on Streamflow Characteristics as Related to Channel Geometry of Streams in the Missouri River Basin. U.S. Geological Survey, Open File Rept., Feb. 1974.
13. J. F. Orsborn and F. D. Deane. Investigation Into Methods for Developing a Physical Analysis for Evaluating Instream Flow Needs. Department of Civil and Environmental Engineering, Washington State Univ., Pullman, OWRT Project A-084-WASH Completion Rept., Sept. 15, 1976.

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

Rainfall Intensity-Duration-Frequency Curves Developed From (not by) Computer Output

Brian M. Reich, Pima County Flood Control District, Tucson

Thirty-two years of maxima observed at Tucson International Airport from the National Oceanic and Atmospheric Administration's recording raingage are used to prepare a sheet of intensity-duration-frequency curves commonly used in the design of storm drainage for small urban areas. The example is employed to stress the need for examining computer printouts of mathematical statistical analysis of the rains and their logarithms by plotting data on four types of probability paper. Stress is laid on dangers of blindly extrapolating a mathematical distribution that does not fit recorded amounts for the long return periods in which engineers are usually interested. Misapplication of scales involving a logarithmic transformation are discussed. The fact that longer durations may require a different type of frequency paper than do shorter durations is illustrated and rationalized on the basis of the physical process. Internal compatibility of results for 2-, 5-, 10-, 50-, and 100-year estimates of 5-, 10-, 20-, 30-, 45-, 60-, 120-, and 180-min rainfalls is preserved when examining a tabular array of as many as five frequency analyses on one of these 48 cells.

Intensity-duration-frequency (IDF) curves are a long-standing tool of the storm-drain designer (1, 2, 3). A U.S. Weather Bureau publication (4) gave depths of maximum rainfall for various durations and return periods on many separate maps. Since then, recording gages have provided additional data on rainstorms, often more than doubling record lengths at newer sites.

Local governments and consulting engineers may wish to prepare their own intensity-duration-frequency curves, like Figure 1, from their most up-to-date gage records. The purpose of this paper is to discuss topics that an engineer must consider while preparing such design curves.

There is an urgent need for engineers to gain at least a "feel" for statistical techniques. The availability of canned digital computer programs to fit preselected statistical distributions places the responsibility on the user for testing the validity of those automated analyses with respect to his or her particular data or engineering application. In outlining various means for exercising necessary discretion, this paper will refer to common statistical terms, concepts, and equations. They will be introduced in an informal, intuitive vein. Readers desiring additional pragmatic explanations of these extreme value statistics may wish to study Magnitude and Frequency of Floods (5). That 50-page review of terms and methods also contains complete tables needed in computation and various graph papers needed in plotting extreme rainfall data. Two excellent texts (6, 7) were recently published for engineers with deeper and wider interests in statistics.

Figure 1. Rainfall IDF curves from U.S. Weather Service recording gage, Tucson.

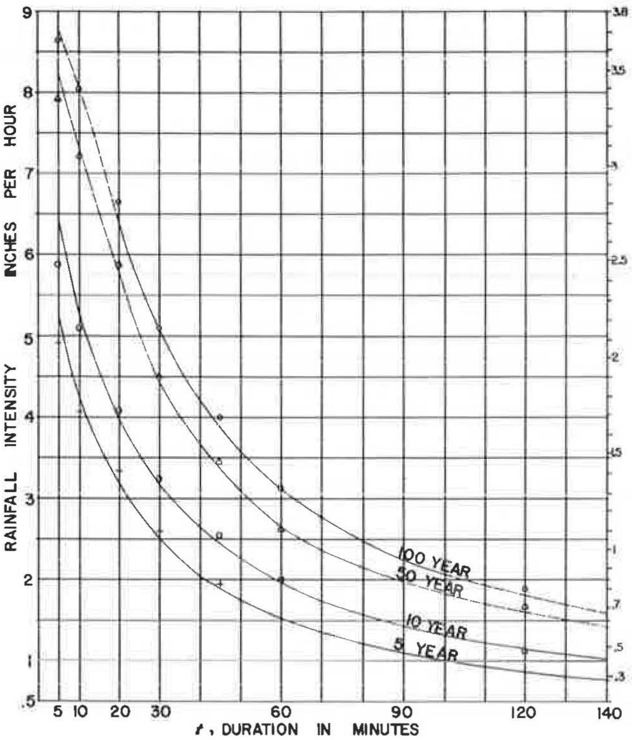


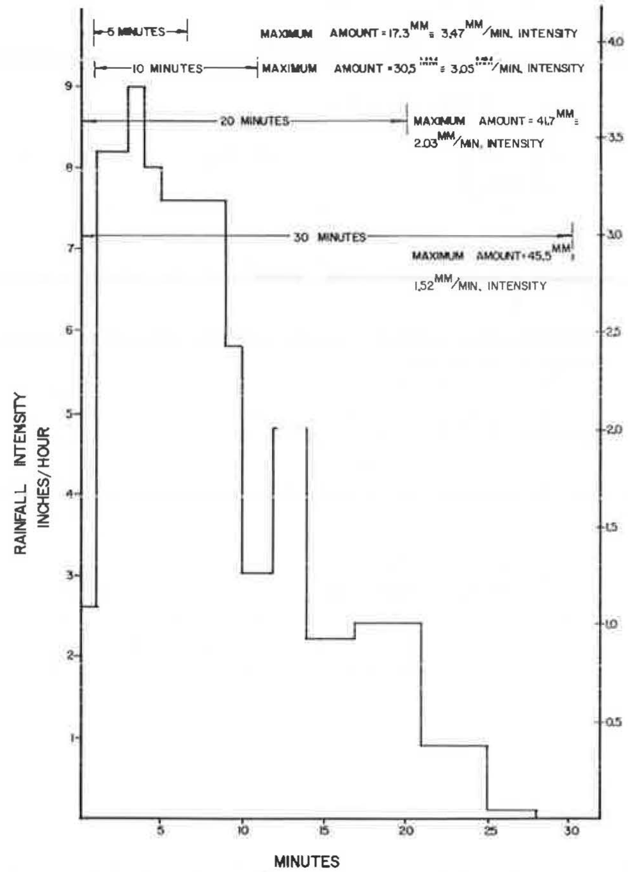
Table 1. Maximum rainfalls of various duration for each year at the Tucson Airport U.S. Weather Service recording gage.

	Maximum Rainfall (mm)							
	Duration of Maxima (min)							
Date[a]	5	10	20	30	45	60	120	180
9/24/1943	14.0	26.7	47.5	57.4	59.2	59.7	62.5	64.5
9/15/1944	10.2	15.2	22.9	26.4	26.9	27.2	30.0	31.0
7/27/1945	6.9	10.2	30.0	39.4	46.7	47.5	51.8	54.1
7/17/1946	9.4	13.5	19.6	21.1	21.3	24.1	26.9	26.9
8/8/1947	4.1	6.1	9.7	14.2	15.2	18.3	18.3	18.3
7/24/1948	14.0	23.6	34.0	40.1	45.0	46.2	47.0	47.5
8/8/1949	8.4	11.7	16.0	17.0	18.0	21.8	27.7	32.8
9/7/1950	8.1	12.2	17.3	19.3	21.1	21.1	24.4	26.7
8/2/1951	5.6	9.1	13.2	15.2	16.5	16.8	17.0	17.0
7/5/1952	9.4	12.2	22.1	23.9	25.1	25.4	25.7	25.7
7/29/1953	5.1	7.6	13.0	17.0	18.0	18.3	18.5	18.5
6/24/1954	9.9	15.7	26.7	30.2	30.5	30.5	31.5	32.0
8/3/1955	17.5	30.7	48.0	53.1	55.9	56.4	57.2	57.9
8/12/1956	7.4	10.2	11.7	11.7	17.0	18.0	18.5	18.5
8/3/1957	5.6	8.6	11.9	14.2	15.7	16.0	19.1	25.4
7/29/1958	10.7	18.8	29.0	34.5	38.9	42.2	68.8	79.2
7/3/1959	11.9	17.0	22.9	29.7	33.8	34.3	34.3	34.3
8/21/1960	4.8	7.6	9.4	10.7	11.7	12.2	12.4	12.4
8/22/1961	11.7	20.8	35.6	41.7	52.8	56.4	59.4	62.2
9/26/1962	3.6	5.1	10.2	11.9	14.2	15.7	17.5	17.5
9/4/1963	6.6	11.2	17.0	19.1	19.3	19.6	19.6	19.6
7/24/1964	11.4	17.5	19.8	22.9	25.9	28.7	45.0	45.2
7/25/1965	5.6	10.2	15.5	16.3	16.5	16.8	18.3	18.3
7/24/1966	7.6	13.0	17.5	22.1	26.2	28.2	34.5	35.1
7/29/1967	5.3	7.4	9.9	13.0	13.7	13.7	15.7	16.0
7/16/1968	7.1	8.4	11.7	11.9	12.2	14.0	14.7	15.0
8/1/1969	6.4	11.4	14.0	14.0	15.0	16.8	16.8	16.8
10/2/1970	8.1	15.2	29.0	29.7	29.7	30.5	30.5	40.9
8/12/1971	8.4	12.4	14.5	14.7	15.2	15.7	18.0	20.6
7/16/1972	10.2	14.0	25.1	34.0	36.8	37.6	39.4	39.4
8/23/1973	3.8	5.6	5.8	7.4	7.9	7.9	7.9	11.4
7/7/1974	9.4	11.4	21.3	25.7	26.9	27.9	30.7	31.8

Note: 1 mm = 0.039 in.

[a] Date refers to largest rainfall with shortest duration for that year. Maximum of longest duration often fell on a different day.

Figure 2. Pluviograph showing temporal variation of rainfall intensity within a convective storm.



With a view to side-stepping a dry discourse on probability theory and the many mathematical constraints on its application, I have adopted a "how-to" format here. An example of deriving IDF curves for Tucson is pursued as a setting for introducing discussions on the various decisions.

EXAMINING THE DATA

The speed of an electronic computer may stimulate the impulse to keypunch the data. After running the cards through one of the readily available statistical programs, such as the Water Resources Council's (8) log-Pearson Type III package, the engineer can look at the printout to see what, say, the 100-year value "is". However, this neglects one of our best resources, observed measurements. In addition to revealing erroneous entries, examination of raw data can provide useful clues to understanding the physical process of interest.

Origins of the Data

Table 1 presents the information on which the Tucson analysis was based. To eliminate any ambiguity, Figure 2 has been developed to illustrate where tabulated rainfall amounts typically come from.

In this sample storm, the largest amount of rain in any continuous 5-min period was the 17.3 mm (0.68 in) that fell in the second, third, fourth, fifth, and sixth minutes. If this had been the largest 5-min amount in one calendar year it would have been published by the National Weather Service and entered in Table 1. Furthermore, the maximum 10-min amount for the year can fall during the same rainstorm. The maximum rain for a longer duration, say, 20 min in the example of Figure 1, can also begin at a different time than that for 5 min.

Annual maximum amounts recorded in Table 1 for long durations such as 180 or 120 min can occur on different days or even in other seasons than that year's very short-duration extremes.

Unfortunately, digital recorders, which only punch their paper tape every 15 min, have recently been replacing many pen-and-chart recording raingages. The high intensities of very short duration will no longer be recorded, and underestimation will be accentuated by the random asynchrony between clock time and pulses of heavy rain.

Before 1935 the U.S. Weather Bureau analyzed tipping-bucket charts from first-order stations onto "excessive precipitation" forms. The latter quantity was defined as any portion of storm rainfall whose intensity exceeds 0.25 mm/min (0.01 in/min) with a threshold of 5.08 mm (0.20 in) of storm total. Even in the humid regions of the United States such storms have generally high-intensity rainfalls lasting 2 h or less. After 1935 the format was changed so that the most intense period of a storm was listed first, followed by its next most intense period, followed by its next, until the entire period of excessive precipitation was accumulated. An engineer fortunate enough to be working with records that go back so far should be aware of the difference.

Sampling Error

Returning to the real Tucson data in Table 1, we first observe that the highest values for short durations occurred in 1955. They were 50 percent greater than any extreme occurring in the subsequent 20 years. The second highest values had been encountered in the first year of recording, 12 years before 1955.

It is easy to realize how drastic the influence would be on the mean, or on other statistical computations, if one began the observations one year later. This should force us to see that even a 30-year record could, by chance, miss the high values that are of great significance to designs for 100-year or even 25-year floods.

These obvious comments derive from the statistical notion of sampling error, which simply states that any finite record length is merely a sample of a hypothetically infinite "population" of values. The mean of the population, μ, will never be known; all we can do is estimate it by the sample mean, $\overline{X}$. Greek symbols are reserved for population values, while Roman symbols are used for variables comprising samples, like X, and sample estimates derived from them.

The mathematical statistician will use population parameters when writing equations to describe, say, the distribution of 5-min annual maximum rainfalls throughout the possible size range of this variable. The applied statistician at best will only be able to substitute a sample estimate for each parameter needed in the theoretical equation. Some parameters or statistics are estimated fairly well from a sample; others are not. For example, if we only had the latest 16 years of this 5-min rainfall maximum, our estimate of the mean would be 7.62 mm (0.300 in). Had our father used only the first 16 years, his sample estimate would have been 9.12 mm (0.359 in), almost 20 percent greater than ours. The 32 years of data give 8.38 mm (0.330 in), which is a better estimate of the mean, but the population mean, μ, remains ellusive.

Plotting the Data

The human mind is limited in its ability to digest a column of numbers and can be greatly aided by a graphic display of the same information (see Figure 3). If points lie approximately on a straight line, then a straight line should be fitted through them by eye, and so labeled. In this way one may estimate a longer return period rain on the basis of many observations rather than simply on the largest rain recorded so far. Extrapolation is strictly justified only if the plotted set of data points displays no systematic deviation from a line. Implicit is the assumption that the distribution of these rainfall values follows the mathematical equation used in generating this particular type of probability paper. If the data exhibit a distinct curve away from the line or a marked s-shape, analysis with this type of paper or corresponding mathematical equation should be abandoned. Different types of probability paper should be tried until approximate linearity is achieved (5).

The horizontal placement of each point is achieved by assigning a rank m to each value in a list of the N observations rearranged from largest (m = 1) to smallest (m = N). These ranks enable us to assign a P_e value to each data point according to a plotting position formula. When using extreme value paper, sometimes called Gumbel for the man who introduced this statistical distribution in the United States, the best plotting position was shown by Gringorten (9) to be approximately

$$P_e = (m - 0.44)/(N + 0.12) \tag{1}$$

P_e is the probability that a rainfall equal to or larger than the specified number of millimeters will occur in one year. The return period in years for such an extreme is

$$T = 1/P_e \tag{2}$$

This cumulative probability, P_e, appears as the axis of Figure 3. It is seen to have a value of 0.01 toward the right. Commercially available paper may have high probability values, like 0.99, on the right that decrease toward the left. In that case, the numbers correspond to the probability of nonoccurrence,

$$P_n = 1 - P_e \tag{3}$$

Mathematically Fitting a Line

A generalized formula for hydrologic frequency analysis is

$$X = \overline{X} + K s_X \tag{4}$$

in which the mean is

$$\overline{X} = \Sigma X/N \tag{5}$$

and the standard deviation is

$$s_X = [(X - \overline{X})^2/(N - 1)]^{1/2} \tag{6}$$

Both $\overline{X}$ and s_x can be obtained from the series of annual maxima. K is the frequency factor that depends on the length of record used to estimate $\overline{X}$ and s_x, as well as on the probability paper selected, whose capital initial is subscripted. Table 2 gives K_E for use when the plot of data exhibits satisfactory linearity on extreme value

Figure 3. Plot of 1943 through 1974 annual maximum series of 5-min rainfall on EV probability paper.

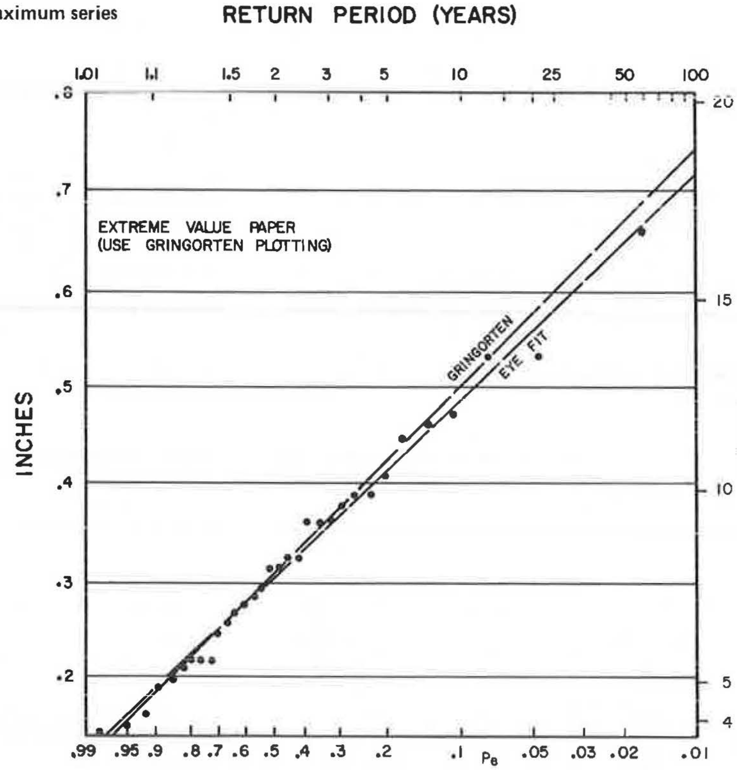


Table 2. Flood frequency factor (K_E) for extreme value line through Gringorten plotting.

Return Period	Frequency, Recording Length (years): 15	20	30	50	100	200
1.111	-1.167	-1.154	-1.140	-1.127	-1.116	-1.110
2	-0.155	-0.158	-0.160	-0.162	-0.163	-0.163
2.33	0.023	0.018	0.013	0.009	0.005	0.004
10	1.431	1.404	1.376	1.352	1.332	1.320
50	2.832	2.775	2.724	2.680	2.643	2.622
100	3.411	3.354	3.294	3.242	3.197	3.172
200	3.997	3.931	3.862	3.801	3.750	3.720
500	4.771	4.693	4.611	4.539	4.479	4.443

paper (see Figure 3). Earlier published tables had assumed Gumbel's mathematical fitting through data plotted according to the Weibull formula

$$P_e = m/(N + 1) \tag{7}$$

which has been in popular use for normal and log-normal paper as well. After this Tucson study a compromise formula

$$P_e = (m - 0.4)/(N + 0.2) \tag{8}$$

has been shown (10) mathematically to suit all four flood frequency papers considered here.

Application of Table 2 and Chow's Equation 4 to our 32 years of 5-min annual maximum rainfalls is simple. The 100-year estimate becomes

$$X_{100\,yr} = \bar{X} + 3.284 s_X = 8.374 + 3.284(3.251) = 19.051 \text{ mm} \tag{9}$$

The mean and standard deviation, s_X, can be obtained from pocket calculators. Similarly,

$$X_{2.33\,yr} = 8.374 + 0.012(3.251) = 8.413 \text{ mm} \tag{10}$$

This is virtually equal to the mean of the series of annual maxima, $\bar{X}$ = 8.382. The equality would be perfect for an infinitely long theoretical population. Computation of a third point with Equation 4 and Table 2 should verify the mathematical straight line (Gringorten) in Figure 3.

Attention should be drawn to the fact that the mathematically fitted extreme value line results from substituting $\bar{X}$ and s_X into Equation 4. It is simple to program a computer to print out estimated rain with 25-, 100-, or even 500-year return periods. The simple computer did not, however, examine a plotted data for linearity. If that criterion is violated, the computer output is misleading; that is, errors would result from using the wrong model.

Other Probability Papers

The extreme value distribution was introduced because it has long been used by the U.S. Weather Service in analyzing short-duration rainfall maxima. They have just produced maps (11) for 5- through 60-min rains for the 37 eastern states from mathematically fitting this distribution, which has the synonym Fisher-Tippett Type I. Their analysis comes close to the application of Equation 4, without the influence of Gringorten's theory. Vast amounts of hourly data were analyzed from about 1900 stations with 25 years of data, and minute-by-minute examinations were made for an additional 200 stations averaging 60 years' record length.

For such short durations there is also a theoretical justification for applying the EV, or Gumbel, distribution. This will be discussed later with respect to the Tucson data, where Figure 3 shows how close the 5-min annual series plots to a straight line. The Gringorten mathematical fit is also seen to closely approximate my own eye fit.

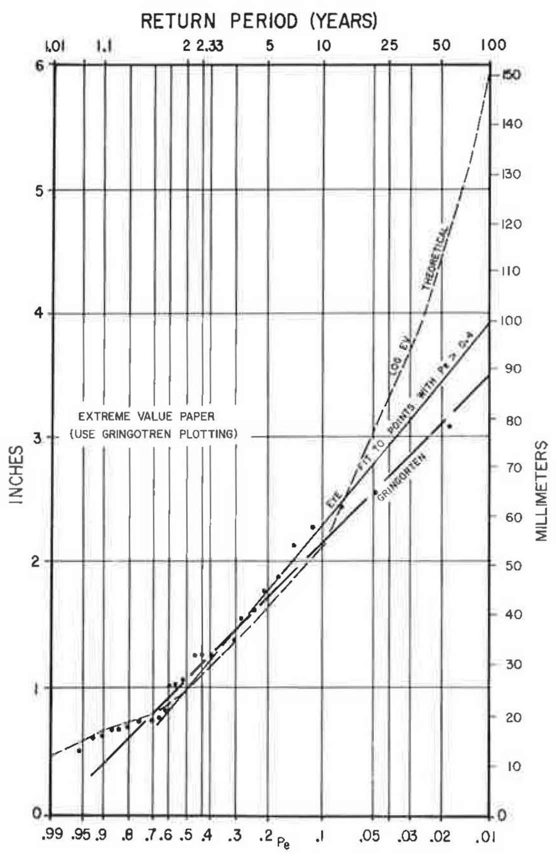
Local preparation of IDF curves will usually involve frequency analysis of only one raingage, so the engineer will be able to rapidly analyze the data graphically and with different types of probability distributions. Moreover, the search for IDF curves often necessitates the frequency analysis of rains for such long durations that the theoretical justification for the EV distribution is no longer valid. Figure 4 shows how systematic nonlinearity of plotted 180-min rainfall maxima shows the invalidity of extrapolating a Gringorten line, or use of Gumbel-type equations. It thus behooves an investigator to try other statistical or probability distributions that may suit those data better. The simple way to achieve this is to plot the annual series on different probability papers.

Log-extreme value paper is seen in Figure 5 to remove the curved toe exhibited by points with P_e greater than 0.2 on Figure 4. The horizontal grid of log-extreme value paper is basically the same as it is on EV paper. A theoretical line can be fitted by applying Equation 4 and Table 2 to the statistics computed from the logarithms of individual annual maxima. After replacing each annual maximum X by L = log X, one may proceed to obtain the mean and standard deviation of the transformed series $\bar{L}$ and s_L. The modified application of Equation 4 gives an estimate

$$\log X = \bar{L} + K_E s_L \tag{11}$$

This produces the theoretical straight line in Figure 5.

Figure 4. S-shape of data plotted by Gringorten's formula showing EV distribution invalid for 180-min rainfalls in Tucson.



Log Scales

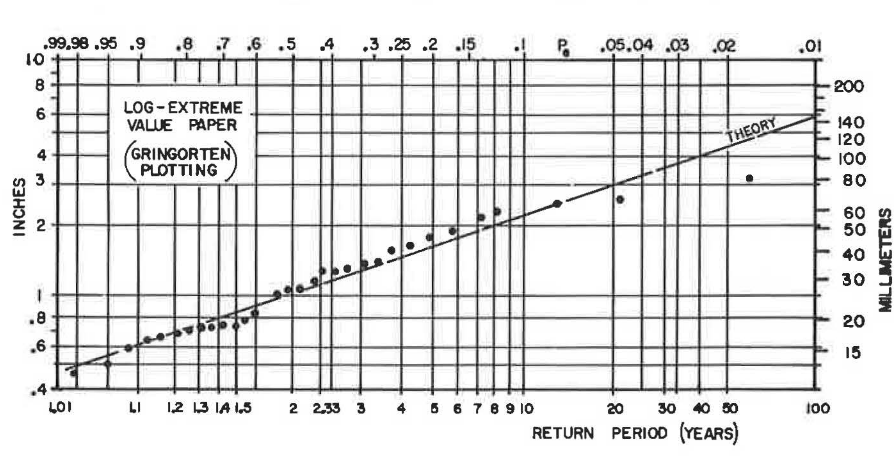
Data points may appear to be closer to the straight line in Figure 5 than to that in Figure 4. One should recognize the deceptive tendency for log-paper to apparently reduce the scatter of large values. The second-largest and largest appear closer to the logarithmic straight line in Figure 5 than to the eye-fit straight line in Figure 4. In fact, they are progressively 12.7 and 40.6 mm (0.5 and 1.6 in) of rain from the theoretical log-extreme value prediction. To emphasize this point the log-EV line, computed from Equation 11, was transformed back to linear units and added onto Figure 4 as a dotted curve.

A second problem of straight-line extrapolation on log paper is also emphasized by this dotted curve in Figure 4. The larger six or ten data points suggest a curve whose slope decreases with progressively longer return periods. The theoretical log-EV curve must, by definition, always have a constantly increasing slope. This is a violation of observations of large rains that particularly relate to our engineering interest in predicting for large return periods. Moreover, our understanding of the physical world suggests that the data should curve toward a horizontal asymptote representing a probable maximum precipitation (PMP) of all the moisture than can be drawn out from the finite overlying atmosphere. No one will deny that the log-EV curve gives the best representation of the 27 smaller data values in Figure 4, but that is not generally the domain of engineering interest.

Log-Normal and Normal Distributions

The classic log-normal (LN) paper, still favored by the Soil Conservation Service, is shown in Figure 6. This type of paper should always be tried in the search for the model that best fits a set of data. The attention that the Water Resources Council (WRC) has forced (8) on the log-Pearson Type III (LP III) makes it im-

Figure 5. Log-extreme value paper improving linearity of 180-min rainfalls from Tucson.



portant for engineers to understand the LN distribution, which is a special case of the LP III.

In Figure 6, the eye fit of this LN to 180-min rains was an acceptably straight line in the range around 5 to 10 years, for which it was used in Table 3. Graphically small deviations of points plotted at 20 and 40 years are actually more significant because the log transformation squeezed the vertical scale. True deviations were less when plotted on normal (N) paper. Extrapolation to the right side gave the best 50- and 100-year 180-min estimates, as signified by the arrows on the eye N line in Table 3.

One of the earliest probability distributions used was this so-called normal distribution. Its characteristic can be seen from Figure 7 to be a symmetrically changing spacing of the probability lines on either side of $P_e = 0.5$. This symmetry is due to its assumption that data will have a zero skewness coefficient, CSX. The latter statistical parameter can be evaluated, albeit with considerable trouble and risk of error, by hand computation as follows:

$$CSX = [N \Sigma (X - \bar{X})^3] / [(N - 1)(N - 2)(s_x)^3] \qquad (12)$$

Engineers soon found that much of their hydrologic data had positive skewness. For instance, the highest floods were often of an order of magnitude greater than floods that occurred rather frequently. This relative largeness in the numerator of Equation 12 was greatly amplified by cubing the terms before summing. An escape from this problem was sought by making a log transformation; this was accomplished through the nonlinear spacing along the vertical axis of Figure 6. The apparent scaling down of larger values was to have drawn the entire annual series of the logarithms into a straight line. If this were perfectly achieved, then the data set would have a zero value for the coefficient of skewness of the logs, where

$$CSL = [N \Sigma (L - \bar{L})^3] / [(N - 1)(N - 2)(s_L)^3] \qquad (13)$$

The advantage of the LN or N distributions is the extreme simplicity of fitting a mathematical curve. If a straight line effectively passes through a whole set of points on N paper, the theoretical line could be fitted as follows:

$$\bar{X} \text{ at } P_e = 0.5,$$

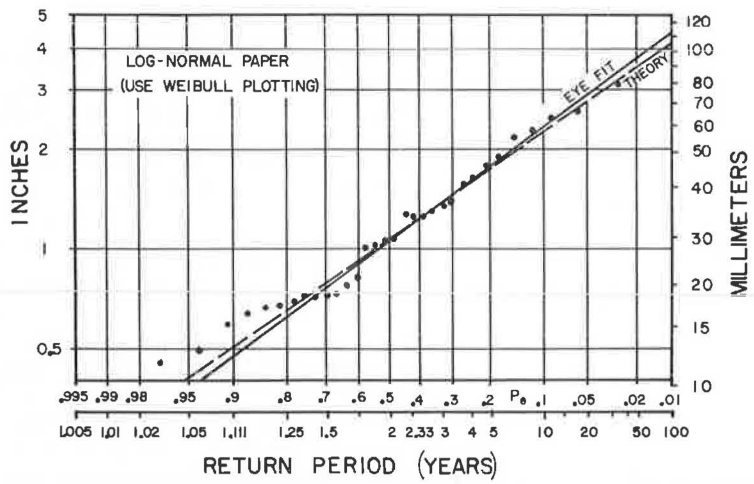


Figure 6. Example of classical log-normal analysis for 180-min rainfalls.

Table 3. Array of rainfalls expected for eight durations and four return periods from various curves.

Duration (min)	Curve	Preferred Order or Comment	Return Period (years) 5 mm	5 mm/min	10 mm	10 mm/min	50 mm	50 mm/min	100 mm	100 mm/min
5	Eye EV	Good	10.4	→ 3.08	13.4	→ 3.49	16.8	→ 3.35	18.3	→ 3.65
10		Good	17.3	→ 1.73	21.6	→ 2.16	30.5	→ 3.05	34.0	→ 3.40
20		Fair	28.2	→ 1.41	34.5	→ 1.73	49.8	→ 2.49	56.4	→ 2.82
30		S-shape	33.0	→ 1.10	41.1	→ 1.37	57.2	→ 1.90	64.8	→ 2.16
	Gringorten	Unacceptable	36.3		44.7		63.3		69.6	
	Eye EV		37.3	→ 0.83	47.5		71.4		80.8	
45	Eye LN		38.1		48.3	→ 1.07	74.4	→ 1.46	86.4	→ 1.69
	Theory LN		34.8		43.4		65.5		75.7	
	Eye N		39.1		48.8		66.0		71.6	
	Gringorten	Unacceptable	37.6		46.0		64.8		72.6	
	Eye EV		37.6		47.5		70.1		79.0	
60	Eye LN		38.1	→ 0.63	51.1	→ 0.85	83.6		99.3	→ 1.32
	Theory LN		35.8		45.2		66.0		76.2	
	Eye N		-		50.8		66.5	→ 1.11	79.3	
	Gringorten	Unacceptable	41.4		51.8		68.6		82.8	
	Eye EV	3	45.0	→ 0.36	55.1		73.2		92.7	
120	Eye LN	2	43.2		56.6	→ 0.47	94.0	→ 0.71	111.8	→ 0.80
	Eye N	1	45.5		56.9		76.7		83.6	
	Theory LN		40.6		51.1		76.7		88.7	
	Gringorten		43.4		55.1		78.2		88.9	
	Eye EV		45.2		57.7		87.4		99.1	
180	Eye LN		45.7	→ 0.25	60.2	→ 0.33	96.5		114.3	
	Eye N	1	46.2		59.5		81.8	→ 0.45	89.7	→ 0.50
	L-Gringorten	Unacceptable	41.4		55.1		112.0		149.9	

Notes: 1 mm = 0.039 in.
Arrows point to intensity corresponding to selected value.

$\overline{X} + s_X$ at $P_e = 0.159$, and
$\overline{X} - s_X$ at $P_e = 0.841$.

When fitting a theoretical LN line, antilogs must be taken, before plotting the line, according to

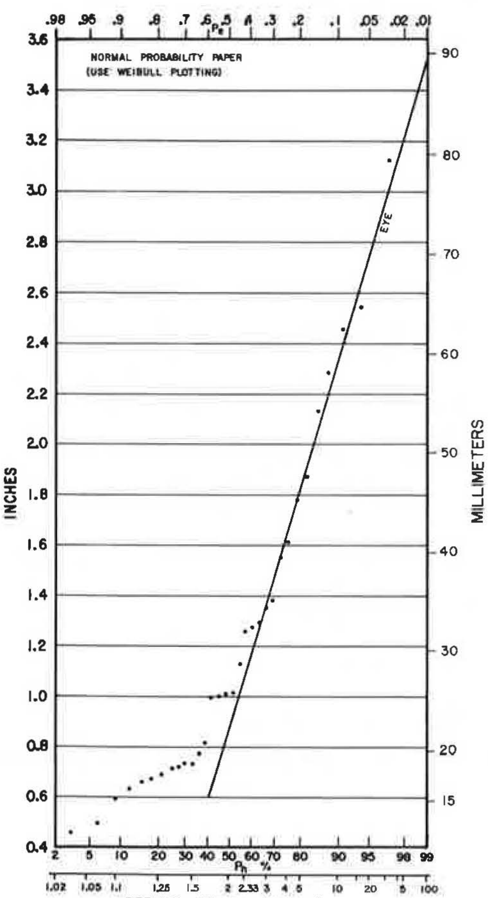
$\overline{L}$ at $P_e = 0.5$,
$\overline{L} + s_L$ at $P_e = 0.159$, and
$\overline{L} - s_L$ at $P_e = 0.841$.

Log-Pearson Type III

Nonlinearities of data, such as those seen in Figures 6 and 7, persisted to frustrate the mathematician. In 1923 Pearson developed a system of twelve types of curves to fit various degrees of upswing or flattening of data. The very next year this empirical system of curve fitting was applied to New York flood problems by H. A. Foster (12). Pearson's curve fitting was again discussed in E. E. Foster's excellent text (13) in 1948.

When the U.S. Army Corps of Engineers (14) proposed using Pearson's Type III (LP III) curve in 1962, they suggested that the log transformation be made to floods before proceeding with the computations. The procedure promulgated by the U.S. Water Resources Council first in 1967 and again in 1976 (8) requires the user to first take the log of each piece of data and then calculate their statistics: $\overline{L}$, s_L, and CSL. The last value is used to select K_P for various return periods, T, from Table 4. Substituting K_P values into Equation 14 and taking antilogs yield the LP III curve

Figure 7. Normal paper suggesting empirical prediction line for large return periods and long duration rainfalls.



$$\log X = \overline{L} + K_P\, s_L \tag{14}$$

The introduction of the third parameter, CSL, gives additional flexibility for closer fitting through observed points. Unfortunately, CSL is highly susceptible to sampling error caused perhaps by the presence of an abnormally large (outlier) maximum. Alternatively, another sample in time may by chance contain many values far smaller than its mean, which could cause CSL to become very negative. The great dangers of using LP III in predicting values for 100-year return periods (5, 15) lies in extrapolating curvature dictated by this error-prone CSL. The problem of determining CSL has led some authors to recommend the use of regional skewness, but that appears to be equally variable, and has led others

Table 4. K_P values for log-Pearson Type III analyses.

	P_e					
	0.90	0.429	0.10	0.02	0.01	0.002
	T					
CSL	1.111	2.33	10	50	100	500
3.0	-0.660	-0.284	1.180	3.152	4.051	6.205
2.9	-0.681	-0.274	1.195	3.134	4.013	6.117
2.8	-0.702	-0.263	1.210	3.114	3.973	6.017
2.7	-0.724	-0.251	1.224	3.093	3.932	5.922
2.6	-0.747	-0.238	1.238	3.071	3.889	5.825
2.5	-0.771	-0.226	1.250	3.048	3.845	5.728
2.4	-0.795	-0.213	1.262	3.023	3.800	5.628
2.3	-0.819	-0.200	1.274	2.997	3.753	5.527
2.2	-0.844	-0.185	1.284	2.970	3.705	5.425
2.1	-0.869	-0.171	1.294	2.942	3.656	5.321
2.0	-0.895	-0.155	1.302	2.912	3.605	5.215
1.9	-0.920	-0.139	1.310	2.881	3.553	5.108
1.8	-0.945	-0.125	1.318	2.848	3.499	5.000
1.7	-0.970	-0.108	1.324	2.815	3.444	4.890
1.6	-0.994	-0.092	1.329	2.780	3.388	4.779
1.5	-1.018	-0.075	1.333	2.743	3.330	4.667
1.4	-1.041	-0.058	1.337	2.706	3.271	4.553
1.3	-1.064	-0.041	1.339	2.666	3.211	4.439
1.2	-1.086	-0.025	1.340	2.626	3.149	4.323
1.1	-1.107	-0.008	1.341	2.585	3.087	4.206
1.0	-1.128	0.010	1.340	2.542	3.022	4.088
0.9	-1.147	0.026	1.339	2.498	2.957	3.969
0.8	-1.166	0.042	1.336	2.453	2.891	3.850
0.7	-1.183	0.058	1.333	2.407	2.824	3.730
0.6	-1.200	0.075	1.328	2.359	2.755	3.609
0.5	-1.216	0.095	1.323	2.311	2.686	3.487
0.4	-1.231	0.111	1.317	2.261	2.615	3.366
0.3	-1.245	0.126	1.309	2.211	2.544	3.244
0.2	-1.258	0.142	1.301	2.159	2.472	3.123
0.1	-1.270	0.158	1.292	2.107	2.400	3.001
0.0	-1.282	0.177	1.282	2.054	2.326	2.878
-0.1	-1.292	0.193	1.270	2.000	2.252	2.759
-0.2	-1.301	0.208	1.258	1.945	2.178	2.639
-0.3	-1.309	0.225	1.245	1.890	2.104	2.520
-0.4	-1.317	0.240	1.231	1.834	2.029	2.401
-0.5	-1.323	0.256	1.216	1.777	1.955	2.283
-0.6	-1.328	0.268	1.200	1.720	1.880	2.171
-0.7	-1.333	0.283	1.183	1.663	1.806	2.062
-0.8	-1.336	0.299	1.166	1.606	1.733	1.953
-0.9	-1.339	0.313	1.147	1.549	1.660	1.846
-1.0	-1.340	0.327	1.128	1.492	1.588	1.741
-1.1	-1.341	0.340	1.107	1.435	1.518	1.647
-1.2	-1.340	0.353	1.086	1.379	1.449	1.556
-1.3	-1.339	0.365	1.064	1.324	1.383	1.467
-1.4	-1.337	0.405	1.041	1.270	1.318	1.383
-1.5	-1.333	0.390	1.018	1.217	1.256	1.303
-1.6	-1.329	0.400	0.994	1.166	1.197	1.233
-1.7	-1.324	0.410	0.970	1.116	1.140	1.169
-1.8	-1.318	0.419	0.945	1.069	1.087	1.107
-1.9	-1.310	0.427	0.920	1.023	1.037	1.051
-2.0	-1.302	0.439	0.895	0.980	0.990	0.998
-2.1	-1.294	0.444	0.869	0.939	0.946	0.952
-2.2	-1.284	0.451	0.844	0.900	0.905	0.909
-2.3	-1.274	0.459	0.819	0.864	0.867	0.870
-2.4	-1.262	0.465	0.795	0.830	0.832	0.833
-2.5	-1.250	0.470	0.771	0.798	0.799	0.800
-2.6	-1.238	0.473	0.747	0.768	0.769	0.769
-2.7	-1.224	0.476	0.724	0.740	0.740	0.741
-2.8	-1.210	0.479	0.702	0.714	0.714	0.714
-2.9	-1.195	0.480	0.681	0.689	0.690	0.690
-3.0	-1.180	0.481	0.660	0.666	0.667	0.667

Table 5. Parameters from fitting IDF equation through observed intensities at eight durations.

Parameter	Return Period (years) 5	10	50	100	Approximate 2-Year Partial Duration ≤120 Min
a	50.1	66.0	91.1	102.1	33.7
b	17.29	18.97	18.64	17.91	14.49
r^2	0.999	0.999	0.995	0.993	0.999

Note: 1 in/h = 25 mm/h.

to suggest using a zero skewness of the logs. This causes the LP III curve to simplify back to the log normal.

Selecting the Frequency Curve

Mathematical statisticians are developing analytical tests for deciding which type of distribution, or model, best fits the data. Unfortunately, the sampling error in determining the parameters needed by the model's equation complicates the problem, which involves the interaction between the choice of model and uncertainty as to population parameters. Without knowing for certain the type of model, e.g., EV or LN, one cannot say whether the deviations of the data from the frequency curve are reasonable to expect from such a random process. Rather than discuss confidence bands within which a population estimate of, say, a 100-year rain can be expected to lie, this paper will simply consider the purpose for which each frequency analysis is performed.

Sometimes engineers require rainfall estimates for return periods from 20 to 100 years. In such a case, importance is ascribed to fitting a line through the 16 larger rains on Figure 7. On the other hand, designs may concern nuisance water with return periods below 2 years. In this case, attention would need to be paid to the lower part of the elbow in Figure 7.

With regard to those interested in larger rains, eye-fitted evaluation should be made to points with P_e greater than 0.4. For these, Figure 7 displays smaller variability of individual points than does Figure 4, and therefore takes precedence. Simultaneous consideration must be given to Figures 5 and 6, while observing the caution recommended with log-scales. The evaluation soon becomes very complex and beyond the capabilities of an electronic computer. If, on the other hand, computer output from fitting the EV, log-EV, LN, and LP III was simply read, the 100-year estimates would be 110.5, 294.4, 121.7, and 142.2 mm (4.35, 11.59, 4.79, and 5.60 in) respectively.

OBTAINING COMPATIBLE ESTIMATES FROM VARIOUS DURATIONS

An impression of the complexity can be had by studying the pros and cons of various frequency curves in Figures 4, 5, 6, and 7 simultaneously. Studies were also needed for 120-, 60-, 45-, 30-, 20-, 10-, and 5-min durations. For each of these an estimate had to be settled upon for the return periods of 5, 10, 50, and 100 years. Each cell could have involved choices among these six frequency curves: Gringorten, eye-fitted extreme value, eye-fitted log-normal, theoretically fitted log-normal, eye-fitted normal, or log-Gringorten. The array of results is presented in Table 3. The ultimate choice in each cell is marked with an arrow pointing to the equivalent intensity in millimeters per minute. In the 5-, 10-, 20-, and 30-min cases data were plotted so linearly on EV paper that estimates for all return periods should be read from that model rather than from any other paper. As durations increase, different probability papers serve better at fitting the observed maxima. With 45-min rain the eye EV and eye LN yield the best 5- and 10-year estimates. Longer durations' estimates sometimes are an average of two or more models from which arrows emanate.

Overall consistency within the IDF displayed by four curves in Figure 1 is a logical requirement. This constraint, as well as the frequency desiderata discussed above, must also be borne in mind. Thus for each duration in Table 3 intensities must increase toward the right. Likewise there must be a systematic decrease in the selected intensities down each column.

Another advantage of consolidating information across many frequency analyses is that the final curves smoothed through various durations in Figure 1 offset some sampling error. This is exemplified by the circles around the 100-year intensity-duration curve and crosses around the 5-year one. They represent the values settled on in Table 3. Eye fitting of curves through all four such sets of points to obtain generally concentric shapes provides further reinforcement across various return periods.

Intensity Duration Equations

Generalized relations along such curves and between various return periods have been found according to the classical equation

$$i = a/(b + t) \tag{15}$$

where a and b are different constants for each curve in Figure 1. They can be evaluated by the linear regression program wired into many pocket calculators by transforming Equation 15 to

$$it + ib = a \tag{16}$$

whence

$$a/i = b + t \tag{17}$$

and

$$1/i = (b/a) + (1/a)\,t \tag{18}$$

Reciprocals of finalized intensities from Table 3 are regressed as the dependent variable against t, the specified durations, that are considered free of error. The classical intercept and slope in parentheses can be manipulated to obtain a and b for Equation 15. Results of such an analysis are shown in Table 5. The coefficient of determination, r^2 (giving the fraction of the variation in 1/i that was explained by the equation), is highly satisfactory.

Once more the computer output need not be followed slavishly. The dashed portions of two curves in Figure 1 were sketched by eye where the mathematical model seemed to deviate too far from frequency estimates selected in Table 3.

Approximating the 2-Year Estimates

The smaller rainfall intensities that are exceeded every 2 years or more could cause damage or be a nuisance with economic impact, regardless of their occurring twice in one year or not at all in others. So strictly speaking a so-called "partial duration series" containing all events above a certain threshold value should be collected for an exact analysis of their statistics. Since

Table 1 of maximum annual rains does not contain that type of observation, we will have to content ourselves with approximations that seem rational.

A 2-year partial series (16) corresponds to an annual series return period of 2.33 years. The 2.33-year estimate made by extreme value (Gringorten) analysis was shown through Equation 4 and Table 2 to almost equal the mean of the annual series. Thus it is suggested that the mean, $\overline{X}$, be used as an approximation for the 2-year partial duration rainfall. The curve was not added to Figure 1, but a and b as determined from Equation 15, in this case for i millimeters per hour and t minutes are listed on the right of Table 5.

CONCLUSION

Only through the human integration of process and probability understanding can the mass of computer output be transformed into design curves for practitioners.

REFERENCES

1. J. W. Clark and W. Viessman, Jr. Water Supply and Pollution Control. International Textbook, Scranton, PA, 1965.
2. Subcommittee on Small Water Storage Projects of National Resources Committee. Low Dams—A Manual of Design for Small Water Storage Projects. U.S. Government Printing Office, 1938.
3. Rainfall Intensity-Duration-Frequency Curves. U.S. Weather Bureau, Technical Paper No. 25, 1955.
4. D. M. Hershfield. Rainfall Frequency Atlas of the United States for Durations From 30 Minutes to 24 Hours and Return Periods From 1 to 100 Years. U.S. Weather Bureau, Technical Paper No. 40, 1961.
5. B. M. Reich. Magnitude and Frequency of Floods. CRC Critical Reviews in Environmental Control, Vol. 6, No. 4, 1976, pp. 297-348.
6. C. T. Haan. Statistical Methods in Hydrology. Iowa State Univ. Press, Ames, 1977.
7. J. R. Benjamin and C. A. Cornell. Probability, Statistics, and Decisions for Civil Engineers. McGraw-Hill, New York, 1970.
8. Guidelines for Determining Flood Flows Frequency. U.S. Water Resources Council, Bulletin No. 17 of the Hydrology Committee, March 1976.
9. I. I. Gringorten. A Plotting Rule for Extreme Probability Paper. Journal of Geophysical Research, Vol. 68, No. 3, 1963, p. 813.
10. C. Cunnane. Unbiased Plotting Positions—A Review. Journal of Hydrology, Vol. 37, 1978, pp. 205-222.
11. R. H. Frederick, V. A. Meyer, and E. P. Auciello. 5- to 60-Minute Precipitation Frequency Maps for the Eastern and Central U.S. National Oceanographic and Atmospheric Administration, Technical Memo NWF HYDRO 35, 1977.
12. H. A. Foster. Theoretical Frequency Curves. Trans., ASCE, Vol. 87, 1924, p. 142.
13. E. E. Foster. Rainfall and Runoff. Macmillan, New York, 1948.
14. L. R. Beard. Statistical Methods in Hydrology. U.S. Army Corps of Engineers, Sacramento, CA, 1962, pp. 74.
15. B. M. Reich. "Lysenkoism" in U.S. Flood Determinations. Paper presented to the American Geophysical Union, Surface Runoff Committee, San Francisco, Dec. 1977.
16. R. K. Linsley, Jr., M. A. Kohler, and J. L. Paulhus. Hydrology for Engineers. McGraw-Hill, New York, 1958.

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