For information and correspondence:
WATER RESOURCES PUBLICATIONS
P.O. Box 2841
Littleton, Colorado 80161, U.S.A.

REPRINTED FROM THE BOOK:

STATISTICAL ANALYSIS OF RAINFALL AND RUNOFF

A PART OF THE
Proceedings of the International Symposium on Rainfall-Runoff Modeling held May 18-21, 1981 at Mississippi State University, Mississippi State, Mississippi, U.S.A.

WRP

IMPROVING POINT RAINFALL PREDICTION WITH EXPERIMENTAL WATERSHED DATA

Brian M. Reich
Assistant City Engineer
Tucson, Arizona

and

Herbert B. Osborn
U.S. Department of Agriculture
Tucson, Arizona

INTRODUCTION

When rainfall-runoff modeling is applied to real problems, it is often necessary for the rainfall input needs to be specified for events with probability as small as 0.01 (the so-called 100-yr event). Calculations of the 100-yr storm in desert regions are highly variable when predicted from relatively short raingage records, where raingages are widely spaced, and the areal extent of runoff-producing thunderstorm cells is far smaller than distances between gages. The 150 km^2 United States Department of Agriculture (USDA) Walnut Gulch experimental watershed (figure 1), in southeastern Arizona, is much larger than the high-intensity thunderstorm rainfall cells that dominate small watershed rainfall-runoff relationships in the region. The dense raingage network (95 recording gages with varying lengths of record) on the watershed provides a good measure of rainfall input. The relatively small size of the intense rainfall cells occurring within different parts of Walnut Gulch suggests that the maximum annual rainfall series can be generated using several well-spaced gages which may be assumed independent sampling points.

The first part of this paper examines the hypothesis of independent sampling points. The criteria established for selection of "independent" sampling points are then used to gain more information than would normally be yielded by only one relatively short raingage record. A combination of point records, the station-year concept, should provide more reliable estimates of, for example, the 100-yr storm. A further use of the station-year concept would be to determine what statistical distribution would be most appropriate for this type of rainfall input. Finally, such knowledge will benefit analysts in similar climatic regions who are limited to a short series from only one or a few widely scattered rain gages.

CORRELATION BETWEEN RAINGAGES

Osborn et al. (1979) investigated the correlation between pairs of gages with distance for storm rainfall using a 20-yr record at 26 recording raingages on the Walnut Gulch experimental

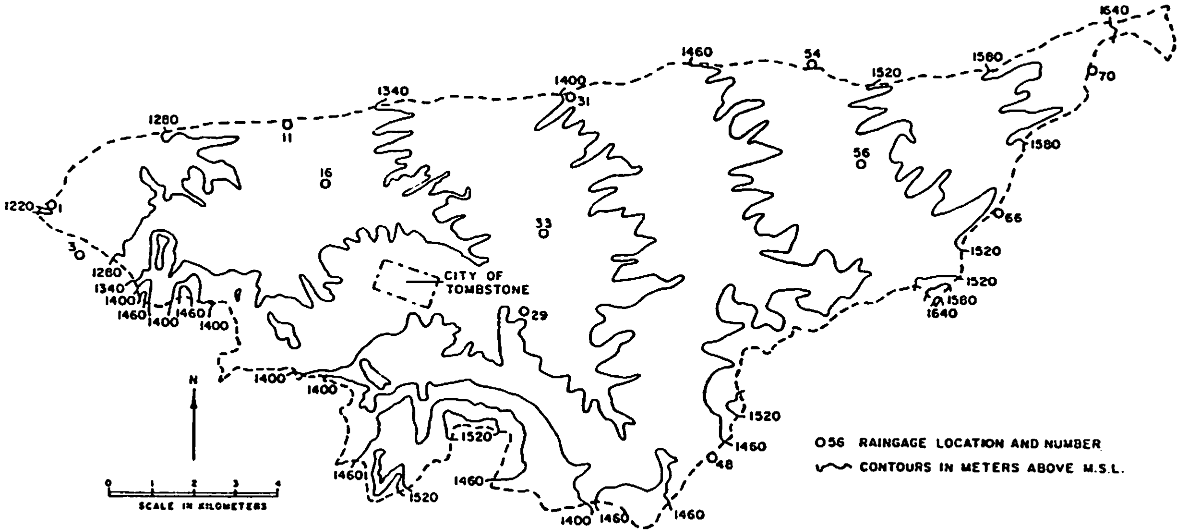

Figure 1. Topography and location of twelve raingages used in combined station-year sets for Walnut Gulch.

watershed. By eliminating time from the correlation, they established a good test for areal rainfall variability and found that the correlation coefficient, r, decreased rapidly with distances between pairs of gages. To verify and reinforce this test, we ran correlations between pairs of gages (13 total gages) for both storm rainfall and maximum 5-min rainfall within each storm (figure 2). The correlations between gage pairs and distance for storm rainfall were similar to that observed by Osborn et al. (1979). As might be expected, correlations for maximum 5-min depths decreased more rapidly with distance than did storm rainfall. For storm rainfall, gage pairs were included when at least one gage recorded 5 mm or more of rainfall; for maximum 5-min rainfall, one gage had to record 1 mm or more. To assume independent sampling points, however, other criteria must be proved or assumed. For one thing, rainfall cannot be significantly correlated with differences in elevation. On Walnut Gulch, the average annual rainfall is about 25 mm more on the upper than the lower end of the watershed (a vertical difference of about 700 m) (figure 1). However, the extreme events have seemed to occur randomly throughout the watershed. To determine the random occurrence of extreme events, the Gumbel (1958) extreme value distribution was used to calculate the 100-yr, 60-min rainfall depths for 40 gages on Walnut Gulch. If anything, there was a slight, although far from significant, decrease in 100-yr, 60-min depths with increased elevation (figure 3). For this reason, we assumed that the extreme events did occur randomly, and that gages at some determined distance were independent sampling points. We also assumed that 20 yr of record on Walnut Gulch is representative of a long-term record at a point in the region. Long-term daily rainfall records suggest that summer rainfall, during the period of record on Walnut Gulch, has been slightly below average (Knisel et al. 1979), but there is no evidence that the nature of extreme thunderstorm events has changed in 100 yr.

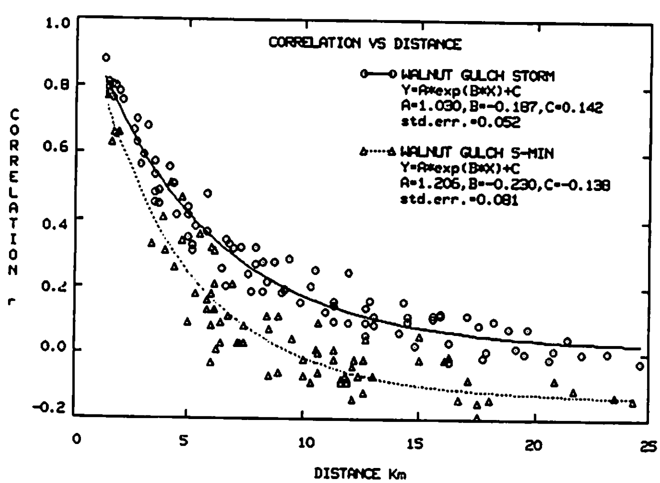

Figure 2. Correlation of maximum 5-min and storm rainfall with distance between pairs of raingages on Walnut Gulch.

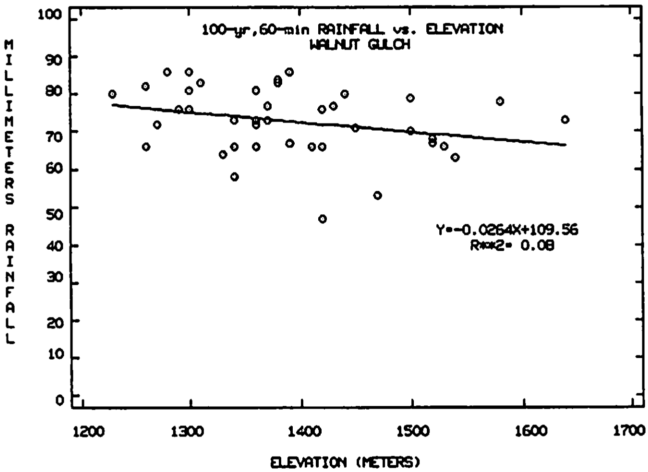

Figure 3. Correlation of estimated 100-yr, 60-min point rainfall depths with elevation on Walnut Gulch.

FIVE-MINUTE MAXIMA

The ensemble of 5-min periods of rain in one year from which each annual maximum amount is chosen is sufficiently large, even in a semiarid climate, that one would expect the annual events to behave according to extreme value theory (Gumbel 1958). On the other hand, there are so few independent 120-min rainy periods each year that the annual series of such a longer duration maxima may well deviate from extreme value theory. In contrast, extreme value theory is commonly applied to longer durations in the more humid East. We began our study with 5-min maxima, but also examined 30-, 60-, and 120-min maxima.

Individual Station Series

Based on figure 2 and the assumption on the randomness of thunderstorm rainfall, we assumed that gages spaced at 5 km or greater on Walnut Gulch could be considered independent sampling points. Twelve Walnut Gulch gages (figure 1) were grouped into three networks of 5 gages (each with a 20-yr record) whose locations were 5 km or greater apart. We were unable to select 15 gages with at least 20 yr of good record each which could be separated into sets of 5 gages with gages separated by at least 5 km. We wanted to look at an odd number of groups, so some gages were included in more than one set. This gave us the equivalent of three 100-yr records. Based on the maximum 5-min curve in figure 2, the correlation coefficient, r, is therefore expected to average less than 0.20 (gages $\geq$ 5 km apart), and the series should behave as independent observations.

Annual maxima for 20 yr of record at each site were plotted according to Cunnane's (1978) compromise formula:

$$p_e = \frac{m - .4}{N + .2} \tag{1}$$

where p_e = probability of occurrence,

where m = rank, with 1 for the largest, and
N = total number of items in the series.

In the following discussions, we consider extreme value (EV), log extreme value (LEV), normal (N), log normal (LN) frequency distributions, and the Log Pearson Type III (LP) method. Plots of the 5 min maxima for Walnut Gulch raingage 54 show the four types of probability paper (EV, LEV, N, and LN) used with each of the five data sets (figures 4a 4b, 4c, and 4d). Log Pearson III is used later in the paper in mathematical fitting only.

Criteria for Drawing a Line Through Data

Since engineering designs based upon rainfall data generally are for return periods of 2 through 100 yr, most attention was paid to rainfall probabilities within this range. Quite often, the smaller observations deviate progressively from the straight line which best fits the trend of the data for return periods longer than 2 yr. Linearity suggests that the correct paper, or probability function, has been used. At the same time, expected sampling error will cause some deviation about the straight line even if the correct distribution is used. Unfortunately, the effect is pronounced for the largest observations, the ones that, for practical reasons, we would like to give the greatest weight. The seriousness of this time-sampling problem worsens for individual stations with short records. The method of eye-fitting such a straight line has been outlined and illustrated elsewhere (Reich and Renard 1981). Also, the optical distortion produced by logarithmic scales must be considered when evaluating the vertical deviations. Our compromise judgment was that EV paper was the best of the four used in figures 4a thru 4d.

Selecting the Best Distribution from a Short Series

The same analyses were performed with 20 yr of record from raingages 3, 33, 48, and 66. The best probability distribution for each of the separate data sets is given along with the 100-yr estimate (P_{100}) in table 1. Predictions of the 100-yr depth vary by a factor of almost two.

Table 1.--Distributions producing best linearity and their 5-min, 100-yr rainfall estimates (mm) compared to LN predictions

Raingage number	3	33	48	54	66	Avg	CV
Best distribution	N	LEV	LEV	EV	LEV	---	---
$P_{5;100}$	13.1	19.4	23.9	22.7	22.0	20.2	.213
$P_{5;100}$ based on LN	14.0	17.3	20.6	23.1	18.3	18.7	.184

We must concede that attempts to analyze a group of 20-yr records lead to a dilemma--do certain points indicate the presence of a particular distribution, or are they merely a result of time sampling error? This does suggest an added justification for studying longer natural series or synthesized series (station-year concept).

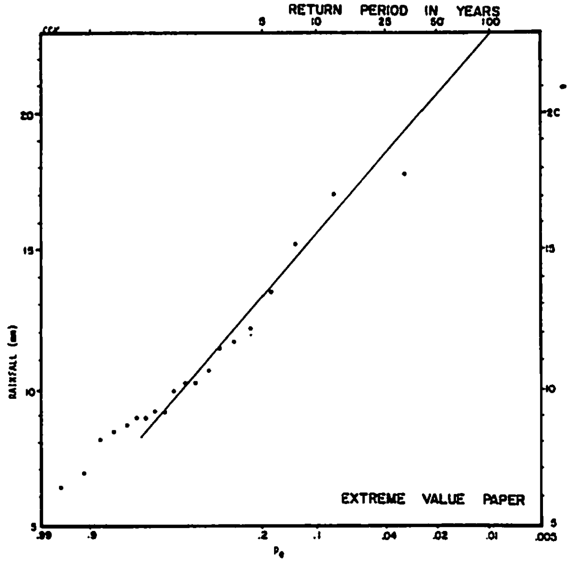

Figure 4a. Visual fit of twenty years of maximum annual 5-min rainfall depths for extreme value distribution for raingage 54 on Walnut Gulch.

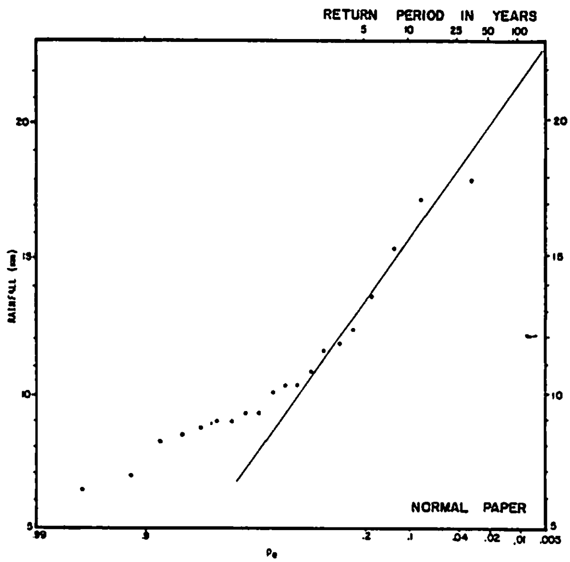

Figure 4b. Visual fit of twenty years of maximum annual 5-min rainfall depths for normal distribution for raingage 54 on Walnut Gulch.

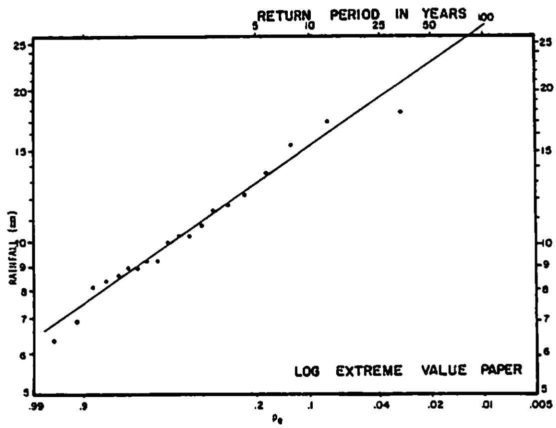

Figure 4c. Visual fit of twenty years of maximum annual 5-min rainfall depths for log extreme value distribution for raingage 54 on Walnut Gulch.

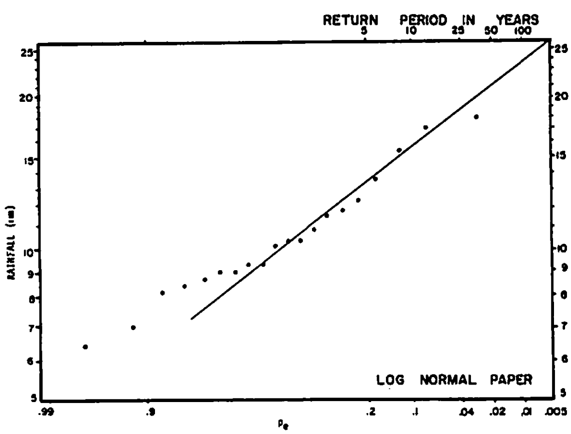

Figure 4d. Visual fit of twenty years of maximum annual 5-min rainfall depths for log normal distribution for raingage 54 on Walnut Gulch.

Station-Year Series

Combining annual series from five rain gages which were $\geq$ 5 km apart (100 yr of record) by assuming independence of the maximum 5-min rainfall depths gives us a much better chance of fitting a distribution to the plotted points. The combined station-year data for gages 3, 11, 33, 54, and 60 fitted a log normal (LN) distribution (figure 5a, 5b, 5c, and 5d). The other two 5-gage combinations also plotted extremely well on LN paper (figures 6a and 6b). Estimates for the maximum 5-min, 100-yr depths ($P_{5;100}$) averaged 19.2 mm.

Applying this station-year evidence that 5-min rainfall maxima are distributed log normally to the five separate 20-yr records requires reexamination of the last row in table 1. There is

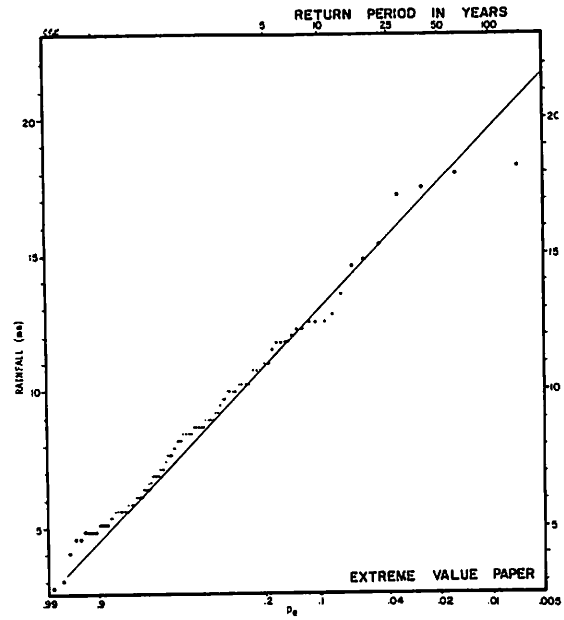

Figure 5a. Visual fit of a combined 100-yr record (raingages 3, 11, 33, 54, and 60) with extreme value distribution for maximum annual 5-min rainfall depths on Walnut Gulch.

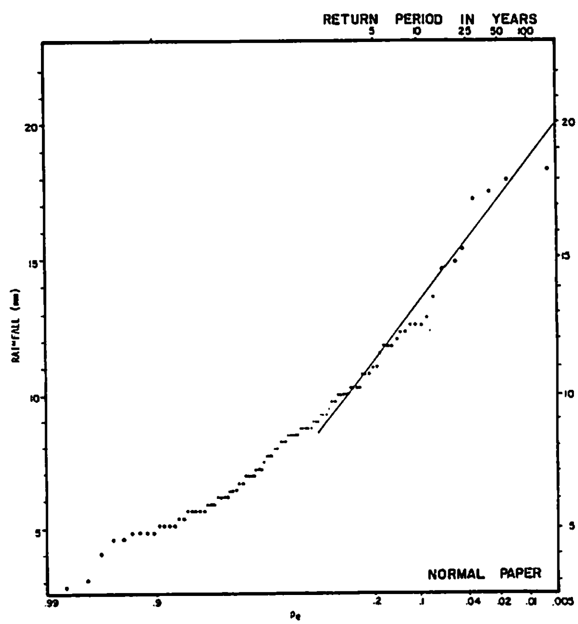

Figure 5b. Visual fit of a combined 100-yr record (raingages 3, 11, 33, 54, and 60) with normal distribution for maximum annual 5-min rainfall depths on Walnut Gulch.

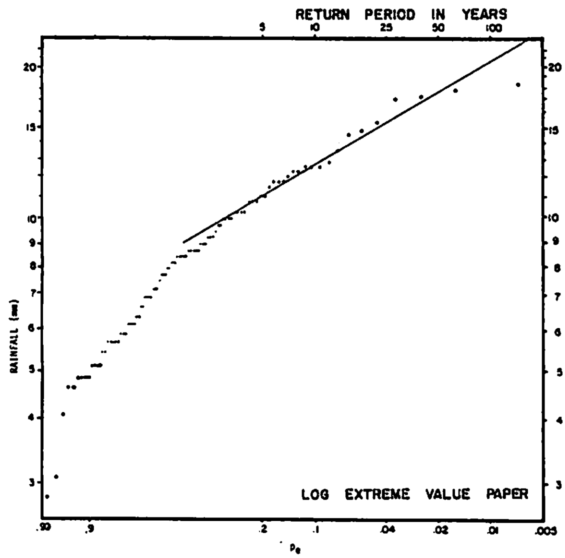

Figure 5c. Visual fit of a combined 100-yr record (raingages 3, 11, 33, 54, and 60) with log extreme value distribution for maximum 5-min rainfall depths on Walnut Gulch.

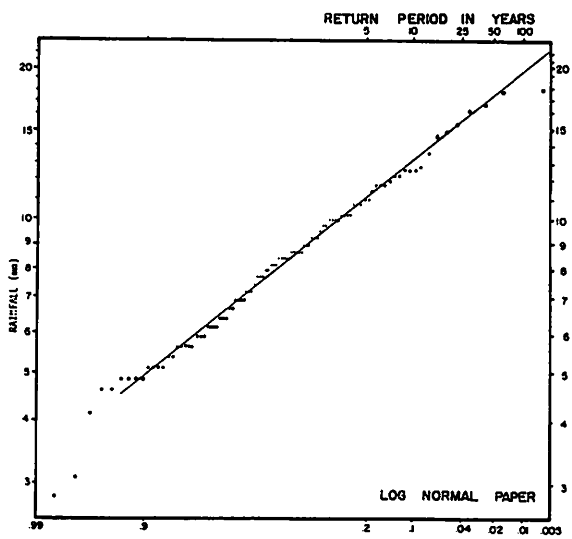

Figure 5d. Visual fit of a combined 100-yr record (raingages 3, 11, 33, 54, and 60) with log normal distribution for maximum annual 5-min rainfall depths on Walnut Gulch.

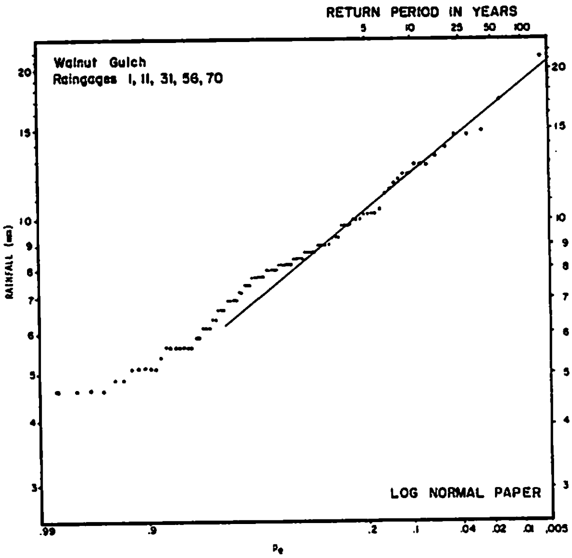

Figure 6a. Visual fit of a combined 100-yr data set (raingages 1, 11, 31, 56, and 70) with log normal distribution for 5-min maxima on Walnut Gulch.

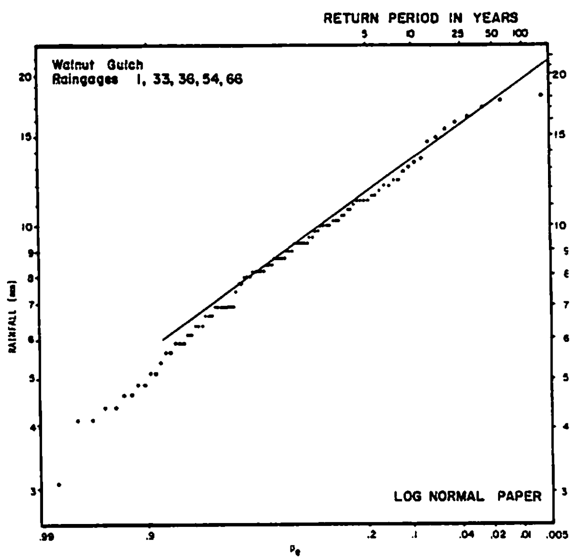

Figure 6b. Visual fit of a combined 100-yr data set (raingages 1, 33, 36, 54, and 66) with log normal distribution for 5-min maxima on Walnut Gulch.

still sampling error. Raingage 3 gave low values of $P_{5;100}$ for the LN. However, it did so for its best model (N), also. Likewise, raingage 54 gave higher estimates with both the LN and LEV models. Coefficients of variability (CV) were computed for the set of individually best distributions and for the set of LN distributions. The variability among the five gage LN estimates is 20% lower than the variability obtained using separately fitted distributions with the 20-yr data sets. The average $P_{5;100}$ obtained by separate model selection is 8% higher than the $P_{5;100}$ averaged from the five LN models. Another reason for confidence in the LN results is that the average of 18.7 mm is close to the 19.2 mm from the three best fits of the 100-yr combined series.

LONGER DURATIONS

Based on data from elsewhere in southern Arizona, annual series for the maximum 120-min depth (Reich 1978) appeared to fit a different probability distribution than did the 5-min maxima. The best linearity for durations of 5-, 30-, 60-, and 120-min, based on Walnut Gulch data and their station-year series, are shown in table 2. For 120-min duration, all three of the combined series indicate a best fit using normal distribution, but LN is second best in all three cases. When LN was used, $P_{120;100}$ was overestimated by only 7, 6, and 5%. These slightly conservative estimates are acceptable if we choose to assume one distribution, LN, for all durations up to 120-min. For 30-min and 60-min durations, all four distributions appear at least once to give the best fit to the data, but LN is best or second best in four of the six cases. Considering all 12 cases in table 2, using LN distribution would lead to an average overestimation of only 2.5%, with the range from -4 to +13%. It would appear that using the LN distribution for all durations from 5- to 120-min is reasonable.

MATHEMATICAL FITTING

It is very easy to fit curves mathematically through an entire set of annual maxima. This eliminates the need for visual fitting of lines through larger events in order to produce estimates appropriate to the return periods ranging from 10 through 100 yr. Unfortunately, an a priori choice must be made of the mathematical model to be used, and such mathematical procedures are oblivious to any systematic trend away from a straight line which may be seen in the entire data series. Thus, a progressive falling off of the data smaller than the 2-yr magnitude may increase the standard deviation and produce an unrealistically large P_{100} by theoretically fitting a steeper straight line through all of the data. Additionally, abnormally high or low outliers will greatly alter the mean and coefficient of variation of the data set. Sampling error will result in large changes of mathematical 100-yr estimates. These phenomena have been discussed elsewhere (Reich 1976). It is, however, interesting to compare the probabilities of rainfall depths for the 100-yr storm derived by five mathematical models (tables 3 and 4) with those derived by visually fitted lines (tables 1 and 2).

The variations covering the five gages with 20-yr records, using mathematical fitting, is summarized in table 3. The coefficients of variation (CV) for each of the five mathematical models

Table 2.--100-yr maximum rainfall estimates produced from the best line arity of four probability papers and from log normal for four short durations

Minutes	5	30			60			120		
Station-year series	Best	Best	LN	LN/Best	Best	LN	LN/Best	Best	LN	LN/Best
A 3,11,33,54,66	LN 19.4	LEV 56	53	0.96	LN 74	74	1.00	N 72	77 second best	1.07
B 1,11,31,56,70	LN 18.5	EV 51	53.5 2nd best	1.05	N 71	80 2nd best	1.13	N 77	82 second best	1.06
C 3,16,29,48,54	LN 19.7	EV 49	48	0.98	LN 58	58	1.00	N 63	66 second best	1.05
Average	19.2	52.0	51.5	1.00	67.7	70.7	1.03	70.7	75	1.06
CV	.033	.068	.059	----	.126	.161	----	.100	.109	----

Table 3.--The 5-min, 100-yr individual gage rainfall estimates (mm) from five mathematical models

Raingage number		3	33	48	54	66	Avg.	CV
MATH MODEL	EV	15.4	18.3	19.4	21.2	20.2	18.9	.1177
	LEV	14.4	16.5	18.1	21.3	17.3	17.5	.1442
	LN	15.3	18.3	19.0	19.6	21.6	18.7	.1235
	N	12.4	14.6	15.0	17.1	15.8	15.0	.1162
	LP	14.6	17.3	17.9	18.8	20.4	17.8	.1203
Average		14.4	17.0	17.9	19.6	19.1		.124
CV among models		.084	.091	.097	.090	.127	.097	

averaged 0.1244, with a range from 0.1162 for the normal distribution to 0.1442 for the log extreme value distribution. This variability was greater than the fluctuations caused by applying different mathematical models to a single gage (as shown in the last row of table 3), with average CV of only 0.0974. However, small variability, by itself, does not signal engineeringly acceptable answers. For example, the normal model has the least variability, but gives an average 100-yr estimate of only 15 mm. This is clearly too low based on the other mathematical and visual best fits.

The variability is reduced by using station-year analysis (table 4). One combination of five gages for the 5-min series varied so little that the 100-yr estimates from the five different mathematical models had a CV of 0.0746. The average of CV's from individual series was about 30% greater. The last row of table 4 also shows low variability among mathematical estimates from station-year series of 30-, 60-, and 120-min maxima. Again,

Table 4.--Precipitation depth estimates (mm) for 100-yr frequency using five models and 100 station-year records for gages 3, 11, 33, 54, and 66

Math models	Minutes	5	30	60	120
EV		19.2	56.7	69.0	73.4
LEV		17.9	56.8	72.0	76.1
LN		20.0	56.8	69.2	73.9
N		16.4	48.1	58.3	62.1
LP		18.8	53.6	65.1	69.6
	Average	18.4	54.4	66.7	71.0
	CV	.075	.070	.080	.078

however, some of these mathematical estimates were considerably lower than most P_{100}'s found by mathematical, as well as visual, best fit. For example, the recommended LN visual fit for 120 min was 77 mm (table 2) versus mathematical estimates ranging from 62 through 76 with an average of 71 mm. Similarly, the station-year 60-min visual estimate averaged 71 mm (with a range of 58 to 80) in table 2 compared to an average of 67 mm and a range of 58 to 72 mm from the mathematical models.

Some consistent contrasts between various mathematical models are worth noting (tables 3 and 4). First, the normal (N) model universally gives substantially smaller 100-yr estimates for both eye and mathematical fits. To a lesser extent, the Log Pearson Type III (LP) gives smaller P_{100}'s than the EV, LEV, and LN mathematical models. Mathematically computed LP curves in table 4 give P_{100}'s which are low compared to the EV, LEV, and LN values. The LP estimates in table 4 of 18.8, 53.6, 65.1, and 69.6 mm are also small compared to series values from table 2 of 19.4, 53, 74, and 77 mm. Thus, the fitting of a mathematical equation to an entire data series may produce consistent underprediction if the prior assumption of the probability model was inappropriate.

COMPARISON WITH PUBLISHED ESTIMATES

Designers usually do not have the advantage of extensive data from a raingage network such as the one on Walnut Gulch. They must rely on generalized maps, graphs, or tables. The National Weather Service has published regional rainfall maps (NOAA Atlas 2) for the Western States (Miller et al. 1973). The map for Arizona indicates values of 19, 52, 65, and 72 mm for the 5-, 30-, 60-, and 120-min, 100-yr storm for the center of gravity on Walnut Gulch. These compare well with average LN station-year estimates for Walnut Gulch of 19, 52, 71, and 75 mm. However, the maps in NOAA Atlas suggest that the eastern edge of Walnut Gulch will be struck by storms that are 20% more severe than on the western edge. Our own analysis, as indicated in figure 3, does not verify this. Perhaps, a future experiment can be designed to test the hypothesis more thoroughly.

CONCLUSIONS

1. Station-year combinations of annual maxima from raingages in southeastern Arizona that are more than 5 km apart can sta-

bilize estimates of 100-yr return periods.

2. The log normal distribution can be used successfully for visually fitting a straight line through the upper data points for durations from 5 through 120 min.

3. Both mathematical and eye fitting of thunderstorm rainfall data may lead to errors in estimates of 100-yr rainfall depths, but mathematical fitting more often leads to bias in estimates with a greater likelihood of overestimating.

4. Average point values for Walnut Gulch thus developed agree with values derived from NOAA Atlas 2; differences in extreme events with elevation, suggested in NOAA Atlas 2, are not verified by Walnut Gulch records.

REFERENCES

Cunnane, C.
1978. Unbiased Plotting Positions--A Review. J. of Hydrology 37(3):205-222.

Gumbel, E. J.
1958. Statistics of Extreme Values. Columbia University Press, NY, 375 p.

Knisel, W. G., Renard, K. G., and Lane, L. J.
1979. Hydrologic Data Collection: How Long is Long Enough? Proc. ASCE Specialty Conference on Irrigation and Drainage, 238-254.

Miller, J. F., Fredrick, R. H., and Tracey, R. J.
1973. PrecipitationFrequency Atlas of the Western United States, Vol. 8, Arizona, NOAA National Weather Service, Silver Spring, Md. 41 p.

Osborn, H. B., Renard, K. G., and Simanton, J. R.
1979. Dense Networks to Measure Convective Rainfall in the Southwestern United States. Water Resources Res. 15(6):1701-1711.

Reich, B. M.
1976. Magnitude and Frequency of Floods. Critical Reviews in Environmental Control 6:297-348.

Reich, B. M.
1978. Rainfall Intensity-Duration-Frequency Curves Developed from (not by) Computer Output. Transportation Research Record 685:35-43.

Reich, B. M., and Renard, K. G.
1981. Application of Advances in Flood Frequency Analysis. AWRA Water Resources Bulletin 17(1):67-74.