ESTIMATION OF MAGNITUDE AND FREQUENCY OF FLOODS IN

PIMA COUNTY, ARIZONA, WITH COMPARISONS OF

ALTERNATIVE METHODS

By James H. Eychaner

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 84-4142

The opinions, findings, and conclusions expressed in this report are not necessarily those of Pima County or the City of Tucson.

Prepared in cooperation with

PIMA COUNTY AND CITY OF TUCSON

Tucson, Arizona
August 1984

CONTENTS

ILLUSTRATIONS

TABLES

CONVERSION FACTORS

For readers who prefer to use metric units rather than inch-pound units, the conversion factors for the terms used in this report are listed below:

Multiply	By	To obtain
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
square mile (mi^2)	2.590	square kilometer (km^2)
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

National Geodetic Vertical Datum of 1929 (NGVD of 1929): A geodetic datum derived from a general adjustment of the first-order level nets of both the United States and Canada, formerly called mean sea level.

ESTIMATION OF MAGNITUDE AND FREQUENCY OF FLOODS IN PIMA COUNTY, ARIZONA, WITH COMPARISONS OF ALTERNATIVE METHODS

By

James H. Eychaner

ABSTRACT

In Pima County, Arizona, a semiarid region of large relief, new regression equations estimate 5- to 100-year flood discharges with standard errors of 42 to 49 percent. Standard errors for 2- and 500-year discharges are about 60 percent. Predictor variables are drainage area (0.013 to 4,471 square miles), channel slope (0.3 to 13 percent), and shape factor. Second-order regression models represent the logarithmically nonlinear relations found across the wide range of basin characteristics. Flood estimates are reduced if channel conditions cause large attenuation of peaks. Estimates for gaged sites are a variance-weighted average of estimates from regressions and from gage data. Estimates for the Tucson urban area are based on equations developed in a nationwide study. Research on nonlinear logarithmic regressions and variables that index channel conditions might be useful.

Two methods for estimating flood discharges from gage records, two sets of new regressions, and two previously published regional methods are compared. Distribution-free tests against maximum observed floods show differences in accuracy between the methods, and comparisons with base methods show differences in variability. The tests and comparisons indicate that the new equations are more accurate and less variable than methods previously published.

INTRODUCTION

Reliable estimates of the magnitude and frequency of floods are needed for cost-effective design of bridges, culverts, drainage channels, and flood-protection structures and for effective management of flood plains. Estimating techniques are constantly being developed as needs change or as additional data become available. In Pima County, Arizona, continued urban development has caused increased interest in changes in flood runoff in urban areas. Several techniques of flood estimation are available for Pima County that focus on different aspects of flood estimation. In this study, more than 2,000 station-years of systematic flood data at 101 streamflow-gaging stations in and near Pima County were used; 6 years of recent flood data at 74 stations were used that had not been used in previous analyses.

1

Pima County is in the Basin and Range lowlands water province of southern Arizona and includes 9,240 mi^2 of the upper Sonoran Desert. Wide valley floors receive an average rainfall of 6 to 12 in./yr at elevations of less than about 3,500 ft above the National Geodetic Vertical Datum (NGVD) of 1929. Average precipitation in scattered areas above 7,500 ft is as much as 30 in./yr. Most floods are the result of intense summer thunderstorms, although tropical storms and widespread winter storms can produce serious flooding in the largest basins. The population of Pima County increased from 350,000 in 1970 to 530,000 in 1980. Tucson is the largest city and has about 60 percent of the population (U.S. Census Bureau, 1981).

In 1965 the U.S. Geological Survey, in cooperation with Pima County and the City of Tucson, began an investigation of the flood characteristics of streams in Pima County (Condes de la Torre, 1967). A network of crest-stage and recording streamflow stations was established to collect flood data, primarily for small streams. Major data collection in Pima County was completed at the end of the 1981 water year; a similar statewide program ended after the 1975 water year. Data collected in the Pima County program were used in a statewide flood-frequency study (Roeske, 1978) and in flood-insurance studies for Tucson and Pima County (Federal Emergency Management Agency, 1982a, b).

Purpose and Scope

The purpose of the study was to develop convenient and reliable techniques for estimating the magnitude and frequency of floods in Pima County, Arizona. A stepwise analysis explored many possibilities in order to determine the best available techniques. To clarify the advantages of several methods that might be used, the results of the techniques were compared with those of several other methods. New techniques are described that apply to rural and urban watersheds. Flood-frequency characteristics at gaging stations were correlated to basin characteristics such as drainage area and channel slope. Flood magnitudes (peak discharges) were estimated for recurrence intervals of 2, 5, 10, 25, 50, 100, and 500 years using guidelines of the U.S. Water Resources Council (1981b) and data through water year 1981. Recurrence interval is the reciprocal of the annual exceedance probability. Thus, a 50-year peak discharge has one chance in 50 (2 percent) of being exceeded at least once during any year. Generalized relations between flood-frequency and basin characteristics at gaging stations were derived by multiple-regression analysis. The relations were considered to apply to ungaged basins, and the applicability was tested statistically. Flood magnitudes for any basin in Pima County may be estimated using the equations presented in this report.

All results presented in this report are in inch-pound units. The coefficients of the estimating equations have implied units which depend on the measurement units of the original data. If the original data were transformed to equivalent values on a different scale, the

regression analyses would produce different values for the coefficients. Therefore, users who wish to use the metric system must convert input values to inch-pound units, compute the estimates as given, and reconvert the results to metric units. A table of conversion factors precedes the abstract.

Previous Studies

Techniques developed in this study were compared with those presented by Roeske (1978), Zeller (1979), and Reich and Renard (1981). Roeske (1978) presented sets of equations to estimate 2- through 500-year recurrence-interval flood discharges for any rural basin in Arizona on the basis of flood records through 1975. Separate sets of equations were applied to eastern Pima County, western Pima County, and areas above an elevation of 7,500 ft. Drainage area was the only predictor variable in the equations for Pima County. Roeske's equations superseded the method given by Patterson and Somers (1966) and the regional method of the Arizona Water Commission (1973).

Zeller (1979) presented a procedure to estimate 2- through 100-year peak discharges for basins as large as 10 mi^2 in Pima County from precipitation intensity, soil characteristics, vegetation, flow distance, slope, and roughness. Recommended ranges were given for several coefficients for rural and urban basins. Zeller dealt explicitly with factors that have been shown to affect peak discharge in various studies and calibrated his model using data from specific storms in Pima County. The Pima County Department of Transportation and Flood Control District recommended the use of Zeller's method for prediction of flood peaks from ungaged watersheds.

Reich and Renard (1981) recommended a graphical approach in frequency analysis of flood records. They plotted flood data on several types of graph paper that were designed to linearize possible distributions of annual floods. The graph for which a straight line best fit the data was selected as the correct distribution. Reich and Renard (1981, p. 69) stated that "Such linearity permits confident extrapolation to rare probabilities. The only line with predictive use on flood-frequency paper is a straight line."

In extensive comparisons of 10 methods for estimating flood magnitudes in ungaged basins, the U.S. Water Resources Council (1981a) showed that the regional-regression method was at least as good as other common approaches in terms of accuracy, reproducibility, and ease of application. Estimates by each method were compared to the "true" flood frequency, which was determined from gage records for each site. The rankings of the ungaged basin methods were the same regardless of the method used to determine "true" flood probabilities.

4

Acknowledgments

APPLICATIONS

Techniques for estimating the magnitude and frequency of floods in Pima County, Arizona, are described in this section. Three techniques allow the user to compute the peak discharge at specified recurrence intervals for any rural or urban basin in the county. The discussion of each technique includes an example of its use. Techniques are presented to adjust the estimated discharges for uncertainty, for extreme attenuation, or if gaging-station data are available. A summary of the maximum known floods in the area is included. The techniques are based on data collected at 55 gaging stations in Pima County and 46 in adjacent counties (fig. 1). Development of the techniques and comparison with several other methods are discussed later in this report.

Ungaged Rural Sites

Flood magnitudes at ungaged sites in Pima County that are unaffected by urbanization may be computed for recurrence intervals of 2, 5, 10, 25, 50, 100, and 500 years by using the equations in table 1. The equations define the relations of flood magnitudes to drainage area, channel slope, and basin shape. An alternate set of equations for estimating peak discharges requires drainage area only (table 2); however, the results are less accurate than results from equations in table 1. Basin characteristics required for the equations are defined in table 3. All basin characteristics are entered into the equations as base-10 logarithms, and the computed results are the logarithms of the desired peak discharges. Methods applicable to basins that include some urban development are discussed later in this report.
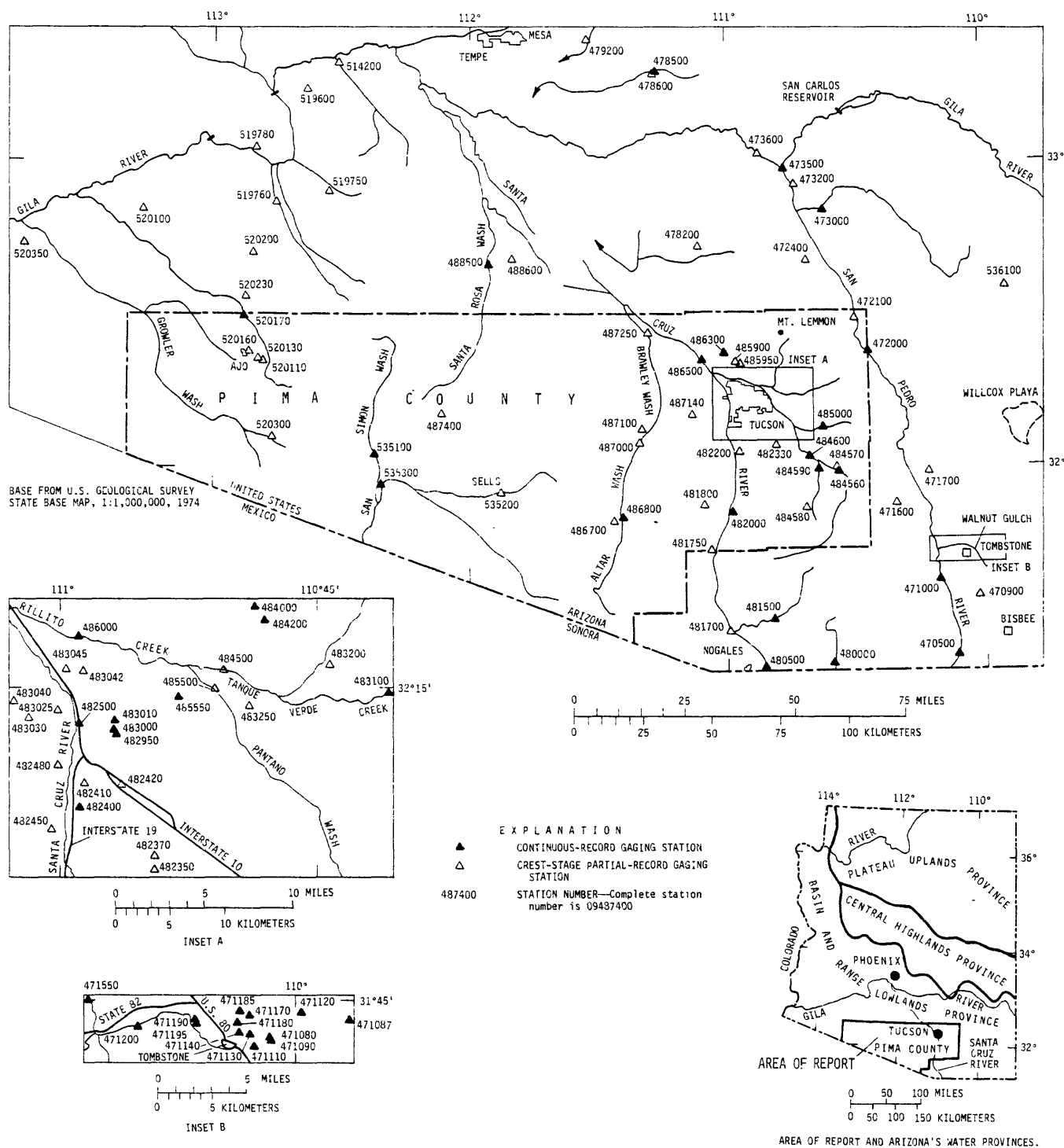
Figure 1.--Area of report and location of gaging stations.

Table 1.--Primary estimating equations for ungaged rural sites

| | | Applicable range | |
		Minimum	Maximum
A:	Drainage area, in square miles.	0.013	4,471
E:	Mean basin elevation, in feet.	600	6,300
L:	Main channel length, in miles.	.28	176
Log:	Base-10 logarithm.	--	--
S:	Main channel slope (10 to 85 percent of L), in percent.	.29	13.4
Sh:	Shape factor (L^2/A) dimensionless.	1.47	20.6
RQ:	Peak discharge, in cubic feet per second, at recurrence interval specified by the subscript. To compute peak discharge, raise 10 to the power of the computed logarithm.	9	52,000

| Equation | Coefficient of determination (R^2) | Standard error of regression | |
		Per-cent	Log units
$LogRQ_2 = 2.049 + 0.547LogA - 0.003(LogA)^2 + 0.299LogS - 0.194(LogS)^2 - 0.253(LogS)(LogSh)$	0.88	60	0.248
$LogRQ_5 = 2.430 + 0.591LogA - 0.023(LogA)^2 + 0.489LogS - 0.275(LogS)^2 - 0.408(LogS)(LogSh)$	.93	43	.181
$LogRQ_{10} = 2.621 + 0.609LogA - 0.031(LogA)^2 + 0.633LogS - 0.288(LogS)^2 - 0.578(LogS)(LogSh)$	.93	42	.176
$LogRQ_{25} = 2.814 + 0.625LogA - 0.039(LogA)^2 + 0.679LogS - 0.329(LogS)^2 - 0.590(LogS)(LogSh)$	.92	43	.180
$LogRQ_{50} = 2.936 + 0.636LogA - 0.044(LogA)^2 + 0.706LogS - 0.350(LogS)^2 - 0.601(LogS)(LogSh)$	.91	45	.191
$LogRQ_{100} = 3.044 + 0.646LogA - 0.049(LogA)^2 + 0.729LogS - 0.367(LogS)^2 - 0.614(LogS)(LogSh)$	.90	49	.205
$LogRQ_{500} = 3.260 + 0.665LogA - 0.058(LogA)^2 + 0.776LogS - 0.396(LogS)^2 - 0.651(LogS)(LogSh)$	.86	58	.241

Table 2.--Alternate estimating equations for ungaged rural sites

[These equations use the same notation and apply over the same ranges of variables as the primary estimating equations shown in table 1]

Equation	Coefficient of determination (R^2)	Standard error of regression	
		Per-cent	Log units
$\text{LogRQ}_2 = 2.051 + 0.551\text{LogA} - 0.011(\text{LogA})^2$	0.87	60	0.247
$\text{LogRQ}_5 = 2.447 + 0.592\text{LogA} - 0.035(\text{LogA})^2$	.91	45	.191
$\text{LogRQ}_{10} = 2.648 + 0.605\text{LogA} - 0.044(\text{LogA})^2$	.91	46	.193
$\text{LogRQ}_{25} = 2.846 + 0.621\text{LogA} - 0.054(\text{LogA})^2$	.90	48	.200
$\text{LogRQ}_{50} = 2.970 + 0.632\text{LogA} - 0.060(\text{LogA})^2$	.89	50	.211
$\text{LogRQ}_{100} = 3.080 + 0.643\text{LogA} - 0.066(\text{LogA})^2$	-.87	54	.224
$\text{LogRQ}_{500} = 3.297 + 0.662\text{LogA} - 0.077(\text{LogA})^2$	.84	63	.260

Table 3.--Evaluation of basin characteristics for rural estimating equations

Basin characteristics required for the equations in tables 1 and 2 should be determined on the basis of the best available topographic maps. Using the boundary or drainage divide of the basin tributary to the site in question:

Drainage area (A), in square miles, is measured by planimeter.

Mean basin elevation (E), in feet above the National Geodetic Vertical Datum (NGVD) of 1929, is obtained by placing a transparent grid over the maps, determining the elevation at each grid intersection within the boundary, and averaging the elevations. The grid size should be chosen so that at least 20 elevation points are sampled in the basin. For large basins, up to 100 elevation points may be needed. In most applications, inspection of the maps will be sufficient to determine that mean elevation lies between the 600- and 6,300-foot limits of applicability.

Channel length (L), in miles, is measured along the main channel from the basin outlet to the drainage divide. At each junction of tributary channels, the main channel is chosen by taking the fork that has the largest drainage area.

Channel slope (S), in percent, is the average slope of the main channel between points that are 10 and 85 percent of the channel length upstream from the basin outlet.

Basin shape (Sh) is the dimensionless ratio of the square of the channel length divided by the drainage area (L^2/A).

8

The equations in tables 1 and 2 were developed from gaging-station data within the ranges of basin characteristics given in table 1. Factors such as soils, geology, channel shape, roughness, and infiltration rate are not included in the equations; therefore, the equations represent an average condition for these factors among the watersheds of Pima County. Apply the equations with caution to basins whose characteristics approach or exceed the limits in table 1 or for which any excluded factors are atypical. For basins that have mean elevations greater than 7,500 ft, equations by Roeske (1978) should be used instead of the equations in tables 1 or 2. Limitations and accuracy of the equations are discussed later in this report, and a method is presented to compensate for uncertainty.

Example: Estimate the 100-year peak discharge in Amigo Wash at Arivaca Road 0.4 mi northwest of Arivaca using both the primary and alternate estimating equations.

1. The drainage basin is shown on the U.S. Geological Survey 7½-minute topographic map for Arivaca quadrangle, and the drainage area is 2.84 mi^2.

2. The channel length is 4.46 mi, and the 10- and 85-percent points are 0.45 and 3.79 mi above the basin outlet. Channel elevations at the 10- and 85-percent points are 3,640 and 3,920 ft, and the average slope between them is 1.59 percent.

3. The basin shape factor is (4.46 mi)^2/2.84 mi^2 = 7.00.

4. Each basin characteristic is within the applicable range given in table 1, including elevation, which ranges from 3,620 to 4,220 ft.

5. Substituting the basin-characteristic values in the primary estimating equation (table 1),

$$LogRQ_{100} = 3.044 + 0.646 Log2.84 - 0.049(Log2.84)^2$$
$$+ 0.729 Log1.59 - 0.367(Log1.59)^2$$
$$- 0.614(Log1.59)(Log7.00) = 3.354.$$

The primary estimate is therefore

$$RQ_{100} = 10^{3.354} = 2,260 \ ft^3/s.$$

6. Using the alternate estimating equation (table 2),

$$LogRQ_{100} = 3.080 + 0.643Log2.84$$

$$- 0.066(Log2.84)^2 = 3.358.$$

The alternate estimate is therefore

$$RQ_{100} = 10^{3.358} = 2,280 \text{ ft}^3/s.$$

Adjustment for Extreme Attenuation

In many basins in Pima County, the peak discharges of individual floods decrease or attenuate as the floods move downstream. This normal attenuation is reflected in the negative coefficients for $(LogA)^2$ in the estimating equations. In a few basins, however, the attenuation is so extreme that the equations may need to be adjusted. In such basins, a channel reach of extremely limited capacity causes flood-water to spread at shallow depths over a very wide flood plain. Large infiltration losses and an uncommonly large volume of water that is stored on the flood plain cause extreme attenuation of peaks. These extreme conditions, however, do not occur in most basins in Pima County, and do not represent a general predevelopment situation.

A reach that causes extreme attenuation is identified by its length, position within the basin, channel capacity, and flood-plain width. The identification of such a reach usually requires inspection in the field because published topographic maps are normally inadequate for reliable determinations. The following criteria for identifying basins subject to extreme attenuation are preliminary and should be used with discretion. The reach that causes extreme attenuation will be in the downstream third of the length of the main channel (L) and will extend for at least 10 percent of L. Within the reach, the bankfull capacity of the well-defined channel or channels will be less than 10 percent of the estimated peak discharge for a 2-year flood, and the width of the area inundated by a major flood will be more than 50 times the width of the well-defined channels.

For basins that meet all the criteria in the previous paragraph, flood magnitudes as computed from the equations in tables 1 or 2 may be reduced by 50 percent. This reduction applies to floods of all recurrence intervals. Flood magnitudes, however, could increase to their unadjusted levels if the channel capacity increases in the attenuating reach because of erosion or construction. The data supporting this reduction are discussed in the section on attenuation later in this report.

Example: Estimate the 50-year peak discharge in the unnamed wash which crosses Camino Verde Boulevard near the center of section 10, T. 15 S., R. 12 E., about 5 mi west of Tucson.

1. The drainage basin is shown on the U.S. Geological Survey 7½-minute topographic map for Cat Mountain quadrangle. The drainage area is 1.16 mi², channel length is 2.06 mi, slope is 0.92 percent, and shape is 3.66. On the basis of the primary estimating equations, the unadjusted 2-year peak is 120 ft³/s, and the 50-year peak is 918 ft³/s.

2. Field inspection showed that the basin has many small braided channels, particularly in the reach from 0.4 to 0.7 mi east of Camino Verde Boulevard. The channels are poorly defined, and branches separate and rejoin every 100 ft or less. In a major flood the inundated area might be 1,000 to 1,200 ft wide. A typical channel in the reach is triangular, measures 6 ft wide and 0.5 ft deep, and has a slope of 0.6 percent and a roughness coefficient of 0.040. According to Manning's formula, the channel capacity is less than 2 ft³/s.

3. The identifying criteria are satisfied so the initial estimate may be reduced. The reach is in the downstream third of the channel (0.7 of 2.06 mi) and exceeds 10 percent of the channel length (0.3 mi). Channel capacity is less than 10 percent of a 2-year flood (2 and 120 ft³/s), and flood-plain width is more than 50 times the channel width (1,000 and 6 ft).

4. The adjusted estimate is Q_{50} = 0.5(918) = 459 ft³/s.

Ungaged Urban Sites

Flood magnitudes at ungaged sites that are affected by urban development may be computed for recurrence intervals of 2, 5, 10, 25, 50, 100, and 500 years by using the equations in table 4. The equations were derived by Sauer and others (1983) from a nationwide sample of 199 urban basins with drainage areas between 0.2 and 100 mi². Comparison with data from gaged basins in Tucson showed that the equations are unbiased for Pima County. The equations account for climatic differences between various regions of the country by using an equivalent rural discharge as one of the predictor variables. In Pima County the equivalent rural discharge is computed from the equations in table 1. The effect of urbanization enters the equations as a basin development factor (BDF), which is an index of the efficiency of the drainage system (fig. 2). The BDF is determined by the prevalence of (1) channel improvements, (2) channel linings, (3) storm sewers, and (4) curb-and-gutter streets in each of the upper, middle, and lower thirds of the drainage basin. As an index variable, the BDF is generally easy to estimate. A fully developed system of improved drainageways would result in each of the four factors being counted in each of the three portions of the basin, which would yield a BDF of 12. Projected values of the BDF could be used to estimate flood frequency for future stages of

Table 4.--Estimating equations for ungaged urban sites

A: Drainage area, in square miles (limited to 0.2 to 100 mi^2).
BDF: Basin development factor, dimensionless.
RQ: Equivalent rural peak discharge, in cubic feet per second, at
     recurrence interval specified by the subscript (from table 1).
UQ: Peak discharge, in cubic feet per second, at recurrence interval
     specified by the subscript at sites affected by urban development.

Equation[1] [Do not use if BDF = 0]	Coefficient of determination (R^2)	Standard error of regression	
		Per-cent	Log units
$UQ_2 = 13.2A^{0.21}(13-BDF)^{-0.43}RQ_2{}^{0.73}$	0.91	43	0.180
$UQ_5 = 10.6A^{0.17}(13-BDF)^{-0.39}RQ_5{}^{0.78}$	.92	40	.170
$UQ_{10} = 9.51A^{0.16}(13-BDF)^{-0.36}RQ_{10}{}^{0.79}$	.92	41	.172
$UQ_{25} = 8.68A^{0.15}(13-BDF)^{-0.34}RQ_{25}{}^{0.80}$	.92	43	.180
$UQ_{50} = 8.04A^{0.15}(13-BDF)^{-0.32}RQ_{50}{}^{0.81}$	.91	44	.186
$UQ_{100} = 7.70A^{0.15}(13-BDF)^{-0.32}RQ_{100}{}^{0.82}$	.91	46	.195
$UQ_{500} = 7.47A^{0.16}(13-BDF)^{-0.30}RQ_{500}{}^{0.82}$	.89	52	.217

[1]From Sauer and others (1983).

urban development (Sauer and others, 1983, p. 8). A BDF of zero indicates the absence of significant channel improvements but does not indicate a total absence of urbanization. If the BDF is zero, the equations for rural basins should be used even though some urban development may be present.

Example: Estimate the 25-year peak discharge in Rose Hill Wash at Broadway Boulevard, 0.25 mi west of Kolb Road in Tucson.

1. The drainage basin is shown on the U.S. Geological Survey 7½-minute topographic map for Tucson East quadrangle. Street construction may change the previous drainage pattern; therefore, the drainage divide was checked in the field. The drainage area is 0.91 mi^2.

12

A. On a map showing the drainage divide, identify the lower, middle, and upper thirds of the basin. The division can generally be made without precise measurements, but each third should include approximately one-third of the drainage area and within each third the travel distances of different streams should be about equal. The subdivision of three typical basin shapes is shown below.

L   Lower third
M   Middle third
U   Upper third
───────── Drainage divide
─··─ Main channel and
           principal tributaries
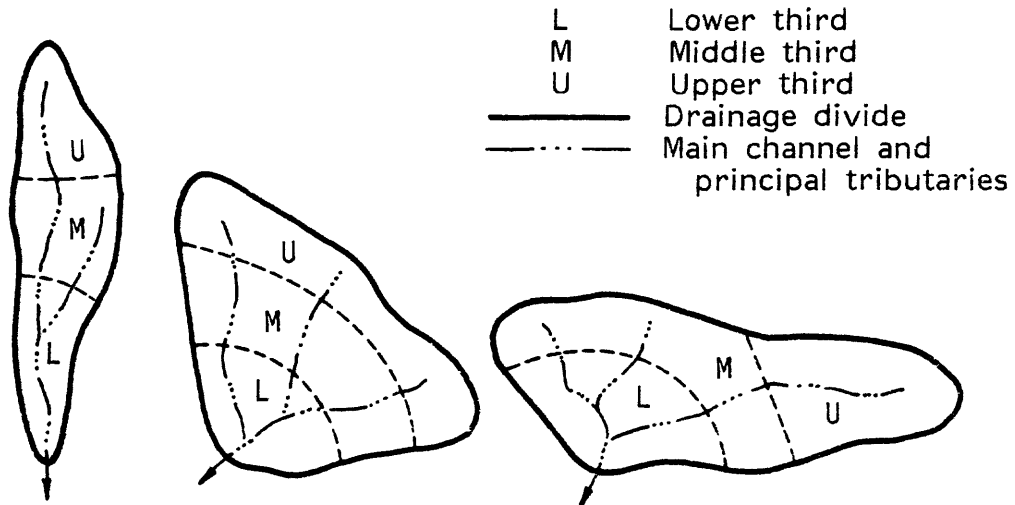
B. For each third, answer these questions:

1. Is at least 50 percent of the length of the main channel and principal tributaries improved to some degree over natural conditions? Improvements include straightening, enlarging, deepening, and clearing.

2. Has at least 50 percent of the length of the main channel and principal tributaries been lined with an impervious material, such as concrete?

3. Do at least 50 percent of the secondary tributaries consist of storm sewers? Storm sewers are enclosed drainage structures—generally pipes. They receive water directly from impervious areas and empty into the main channel or principal tributaries.

4. Is at least 50 percent of the area covered by urban development and are at least 50 percent of the streets and highways in the area constructed with curbs and gutters? Inverted streets, in which water flows at the center, are equivalent to curb-and-gutter streets.

C. The basin development factor is the number of "yes" answers (maximum 4 in each third of the basin for a maximum total of of 12). Field checking is recommended for accuracy.

Adapted from Sauer and others (1983).

Figure 2.--Evaluation of the basin development factor.

2. Channel length is 2.00 mi, and channel slope is 70 ft in 1.50 mi or 0.88 percent.

3. Basin shape factor is 4.40, and basin elevation is entirely between 600 and 6,300 ft.

4. The equivalent rural discharge, from the equation in table 1, is 590 ft³/s.

5. Field inspection identified substantial channel improvements and curb-and-gutter streets in each third of the basin, storm sewers in the upper two thirds, and channel linings in the upper third. The BDF is therefore 3+3+2+1=9. In the upper third of the basin, street gutters that empty directly into the principal tributaries were considered equivalent to storm sewers.

6. Substituting the basin-characteristic values in the equation from table 4,

$$UQ_{25} = 8.68(0.91)^{0.15}(13-9)^{-0.34}(590)^{0.80} = 880 \text{ ft}^3/\text{s}.$$

Alternatively, the effect of urbanization may be approximated by the average ratio of flood magnitude at an urban gaging station to the equivalent rural discharge computed from equations in table 1. The ratio shows the change in flood magnitude in the urban basin in comparison to a similar rural basin. For example, a ratio of 1.25 indicates that the flood peak for a given recurrence interval would average 1.25 times greater for an urban basin than for a similar rural basin. For urban basins in Tucson, the average ratio was:

Recurrence interval, in years	2	5	10	25	50	100	500
Urban/rural ratio	2.25	1.68	1.47	1.33	1.25	1.19	1.09

The gaged basins are all in areas of relatively low slope; none are in the surrounding foothills. Average BDF is 5.6 (range 3 to 9); drainage areas range from 0.95 to 8.2 mi²; slopes range from 0.6 to 1.1 percent; shape factors range from 2.6 to 10.5; and elevations range from 2,300 to 2,700 ft. Many basins in the Tucson metropolitan area have basin characteristics outside the stated ranges, where the ratios would not apply. A reliable accuracy statement cannot be made because of the small number of basins used to derive the average ratios. In addition, the ratios reflect only the average current state of urbanization in the these gaged basins. The ratios do not distinguish different levels of urban development in Tucson or potential changes in flood runoff that might result from continued development of partially developed basins.

Gaged Sites

At gaging stations, two estimates of flood magnitudes are available—one from analysis of the observed flood record and another from regional estimating equations. The user might choose either one, but a better estimate is available from an average of the two estimates, especially if the average is weighted by the accuracy of the original estimates. Accuracy of the estimates is measured by standard error; smaller values indicate greater accuracy. The standard error for the regression estimates is given in tables 1, 2, and 4. The standard error for the observed-flood estimates is a function of the number of years of record, recurrence interval, and standard deviation and skew of the logs of the annual floods; it may be estimated using the method shown in figure 3.

The weighted average is computed as:

$$LogQ_W = \frac{(LogQ_R)(SE_G)^2 + (LogQ_G)(SE_R)^2}{(SE_G)^2 + (SE_R)^2} \tag{1}$$

where

Q_W is the weighted estimate of peak discharge;

Q_R is the estimate from the regression equation (table 1, 2, or 4);

Q_G is the estimate from the gage record (table 8);

SE_R is the standard error of regression, in log units (last column of table 1, 2, or 4); and

SE_G is the standard error of the gage estimate, in log units (fig. 3).

Weighted estimates for all gaging stations in the study are listed in table 8 in the Data section. The standard error of the weighted estimate, in log units, is

$$SE_W = \sqrt{(SE_G)^2(SE_R)^2 / (SE_G^2 + SE_R^2)} \tag{2}$$

with the terms as previously defined. The standard error of the weighted estimate will always be less than either independent standard error, which reflects the improved accuracy of the weighted estimate.

The standard error, in log units, of a flood magnitude estimated from records of observed floods by using the log-Pearson III distribution is

$$SE_G = S_{LQ} R / \sqrt{N}, \text{ where}$$

S_{LQ} is the logarithmic standard deviation of annual floods,

R is a function of recurrence interval and skew whose value is shown in the graph below, and

N is the systematic record length, in years.

For Pima County, estimate S_{LQ} as the arithmetic average of 0.43 and the standard deviation used for the station frequency curve (after adjusting for outliers and historic peaks). The value 0.43 is the average standard deviation for gaging stations in the study. Select R on the basis of the skew coefficient used for the station frequency curve (after adjusting for outliers, historic peaks, and average regional skew). See text for example.
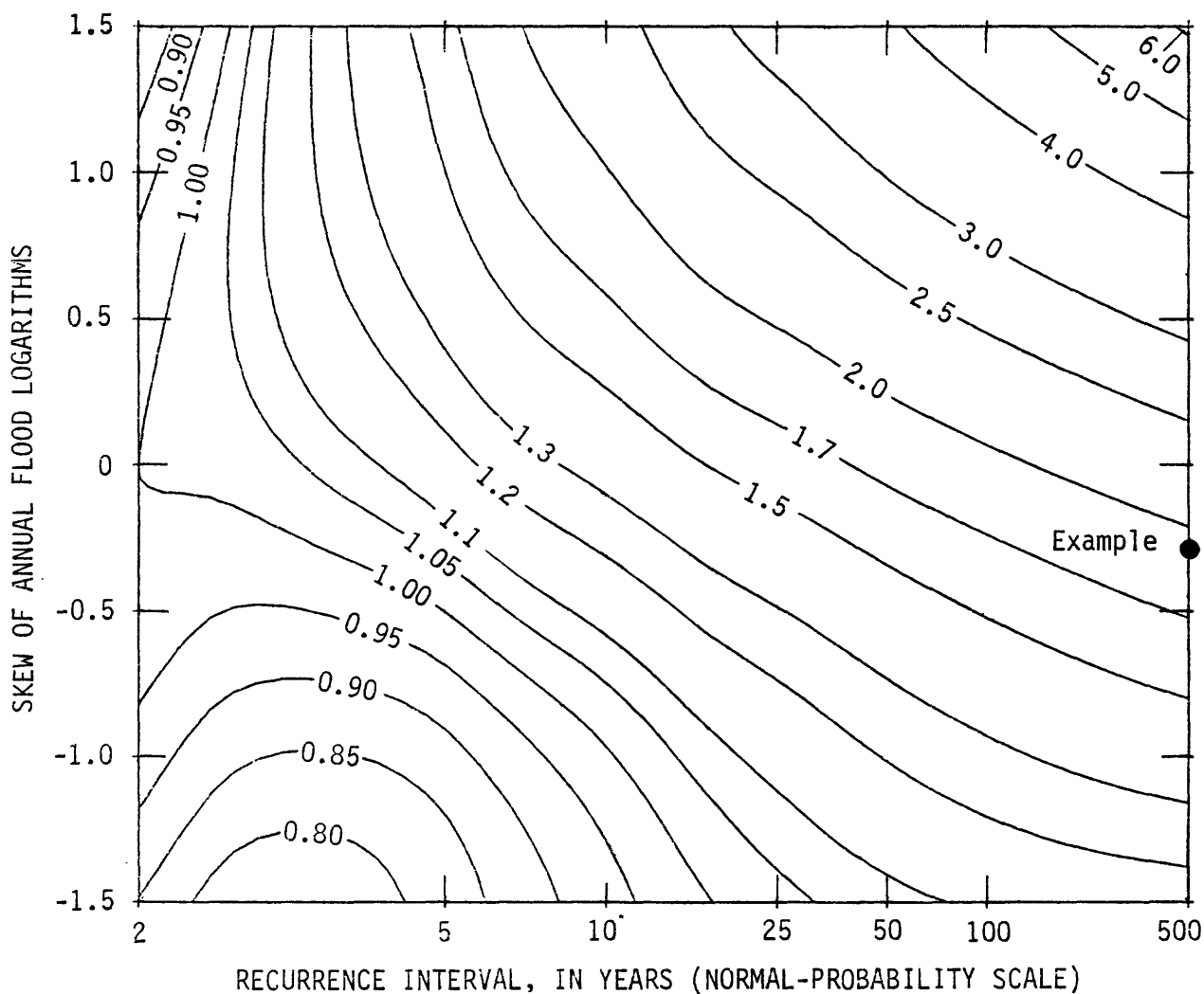

Figure 3.--Standard error of flood magnitudes estimated from records of observed floods.

Example: Compute the weighted estimate of the 500-year peak discharge and its standard deviation for Pima Wash at the intersection of First Avenue and Ina Road about 3 mi north of the Tucson city limits, which is at the Pima Wash near Tucson gaging station (09485900).

1. Data taken from table 8:

   Drainage area = 4.93 mi²   Slope = 10.1 percent
   Shape factor = 6.14   Elevation = 4,430 ft
   Q_G = 998 ft³/s $LogQ_G$ = 2.999
   Years of record = 18.
   Statistics of annual flood logarithms:
        Standard deviation = 0.48
        Skew = -0.27

2. From the last column of table 1:

   SE_R = 0.241 or about 58 percent.

   By applying the 500-year equation from table 1:
   $LogQ_R$ = 3.558 Q_R = 3,610 ft³/s.

3. The standard error of the gage estimate is based on figure 3:

   S_{LQ} = 0.5(0.48+0.43) = 0.455   R = 1.94

   therefore,

   $$SE_G = S_{LQ}R/\sqrt{N}$$
   $$= (0.445)(1.94)/\sqrt{18} = 0.208 \text{ or about 50 percent.}$$

4. The weighted discharge estimate (equation 1) is

   $$LogQ_W = \frac{[3.558(0.208)^2 + 2.999(0.241)^2]}{(0.208^2 + 0.241^2)} = 3.238$$

   $$Q_W = 10^{3.238} = 1,730 \text{ ft}^3/s,$$

   which is also given in table 8. If additional flood data became available for the site giving a new gage estimate, then a new weighted estimate could be computed by this method.

5. The standard error of the weighted estimate (equation 2) is

   $$SE_W = \sqrt{\frac{(0.208)^2(0.241)^2}{(0.208^2+0.241^2)}} = 0.157 \text{ log units}$$

   or about 37 percent.

Limitations and Accuracy

Estimating techniques presented in this report are generally applicable to rural and urban basins throughout Pima County except for basins that are significantly affected by regulation or diversion (fig. 1). The relationships, however, should be considered primarily a method for interpolating between a few known values because the flood-producing mechanisms are complex and incompletely understood. Statistical uncertainty of the techniques is small in comparison to most methods, but the uncertainty is large in absolute terms. A method to compensate for the uncertainty is discussed in the next section of this report.

On the average, peak discharges estimated by the techniques in this report are unbiased; that is, they are not consistently larger or smaller than values derived from observed floods (fig. 4). At a single site, however, there is a 50-percent chance that an estimate is too high and 50-percent chance that it is too low. The magnitude of the possible inaccuracy at a site is measured statistically by standard error.

The standard error of regression measures the scatter of the observed values used to develop the equations around the estimated values. About 68 percent of the observed values were within one standard error of regression, and about 95 percent were within two standard errors of regression (fig. 4). Similarly, the standard error of prediction measures the scatter of the unknown true flood magnitudes around the estimates for sites not used in developing the equations.

A single standard error of regression applies to each equation, but the standard error of prediction varies in a complex way for different sites. The standard error of prediction is smallest for a basin whose characteristics are near the average characteristics used to develop the equations. Standard error of prediction may become large for basins that approach or exceed the limits given in table 1.

Within the study area, however, the average standard error of prediction was approximately equal to the standard error of regression (see Prediction Error). The approximate predictive accuracy of the equations can be assessed using the standard errors of regression listed in tables 1, 2, and 4. The standard error of a peak-discharge estimate that was adjusted for extreme attenuation would be at least as large as for the unadjusted estimate. The standard error for a weighted estimate at a gaging station (equation 1) is given by equation 2. Standard errors were calculated in log units and converted to percentages for convenience.

For characteristics not represented in the equations, the results represent the average condition of the gaged basins. Thus an estimate is most accurate for a basin in which channel capacity, flood-plain width, and infiltration rates are near the average for gaged basins that have similar area, slope, and shape. A basin with most channels on smooth bare rock would probably produce larger peaks, and one with major infiltration losses would have smaller peaks.
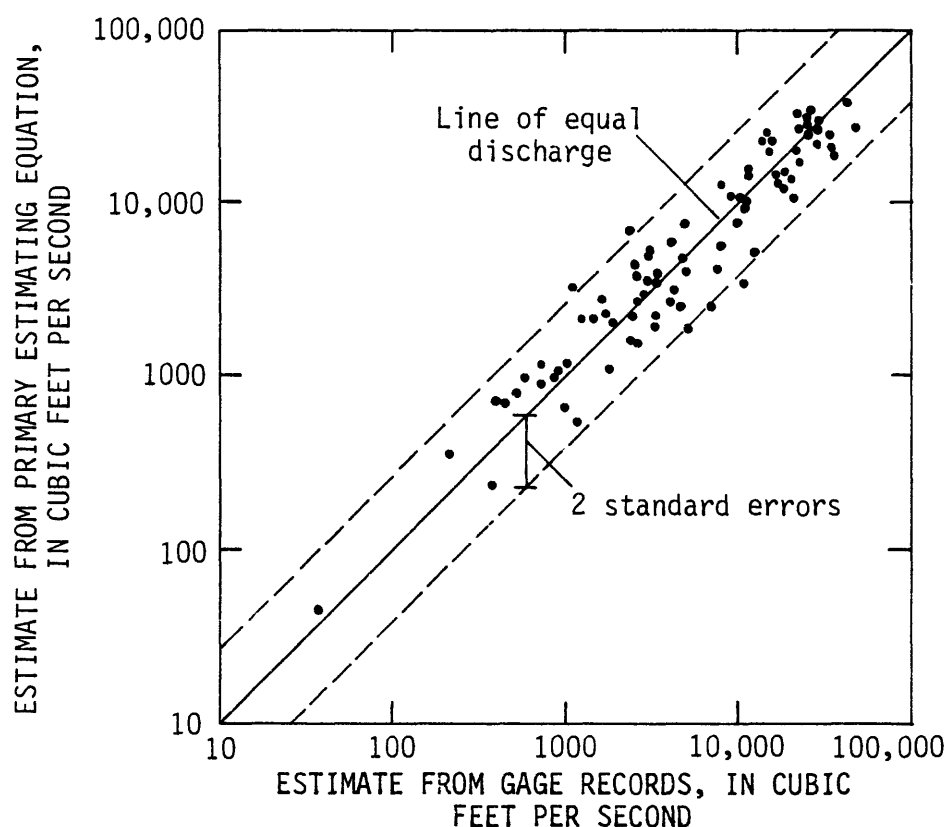
Figure 4.--Comparison of 100-year peak-discharge estimates from
gage records and by regression.

Although the equations presented in tables 1, 2, and 4 give
explicit answers for the peak discharge of any basin in Pima County, the
preceding discussion demonstrates that the user's judgment is essential to
achieving reliable results. The user must decide whether conditions in
the basin are sufficiently unusual to warrant adjustment of the computed
values. The ranges of characteristics given in table 1 and shown in
figure 9 (see Basin Characteristics) are a guide to whether a basin is
unusual. Although the applicable range for each characteristic is large,
a few basins in Pima County have more extreme values. The user might
decide to adjust for attenuation, or for uncertainty as described in the
next section, or subjectively by comparison with other methods that
address the unusual conditions.

Most of this section has considered the uncertainty of discharge
when the recurrence interval is known, but when the discharge is known
the recurrence interval is also uncertain. For example, a particular
discharge might have an estimated recurrence interval of 100 years, which
is an annual exceedance probability of 1 percent. If the uncertainty of
the probability were 0.5 percent, the limits of uncertainty would be 1.5
and 0.5 percent, which correspond to recurrence intervals of 67 and 200
years. Thus there is little point in defining the precise recurrence
interval of a particular large flood.

Adjustment for Uncertainty

In some circumstances, the possible losses due to underdesign of channels or hydraulic structures may be high, and the designer may want better than a 50-percent chance that an estimated peak discharge is at least as large as the true discharge for a specific recurrence interval. An adjusted estimate that makes use of the standard error can be computed for any desired confidence level:

$$Q^* = Q \ (10^{z(p)S_L}), \tag{3}$$

where

Q* is the adjusted estimate of peak discharge at confidence level p,

Q is an unadjusted estimate,

S_L is the standard error, in log units, and

z(p) is the standard normal deviate for cumulative probability p.

Equation 3 is based on the approximately log-normal error distribution associated with the regression analyses. Values of z(p) can be taken from the table below or from tables of the cumulative normal distribution given in most statistics textbooks. The uncertainty-adjustment factor $10^{z(p)S_L}$ is shown in figure 5 for several values of S_L.

p, in percent	50	60	70	80	90
z(p)	0.00	0.25	0.52	0.84	1.28

For a 50-percent confidence level—an equal chance that an estimate is too high or too low—figure 5 shows that the factor is 1.0. A 25-percent increase in a peak-discharge estimate, however, will improve the odds against the estimate being too low to about 2 to 1 (67 percent confidence level) for standard errors in the range of most values in tables 1, 2, and 4. This adjustment procedure can be used with any regression equation in this report.

Example: Estimate the 100-year peak discharge in Amigo Wash at Arivaca Road with 70-percent confidence that the true 100-year peak is no larger.

1. The unadjusted estimate is 2,260 ft³/s (see the previous example for ungaged rural sites).

2. The standard error is 0.205 log units (table 1).

3. From figure 5, $10^{z(p)S_L}$ = 1.28.

4. The adjusted estimate is Q* = 2,260(1.28) = 2,890 ft³/s.
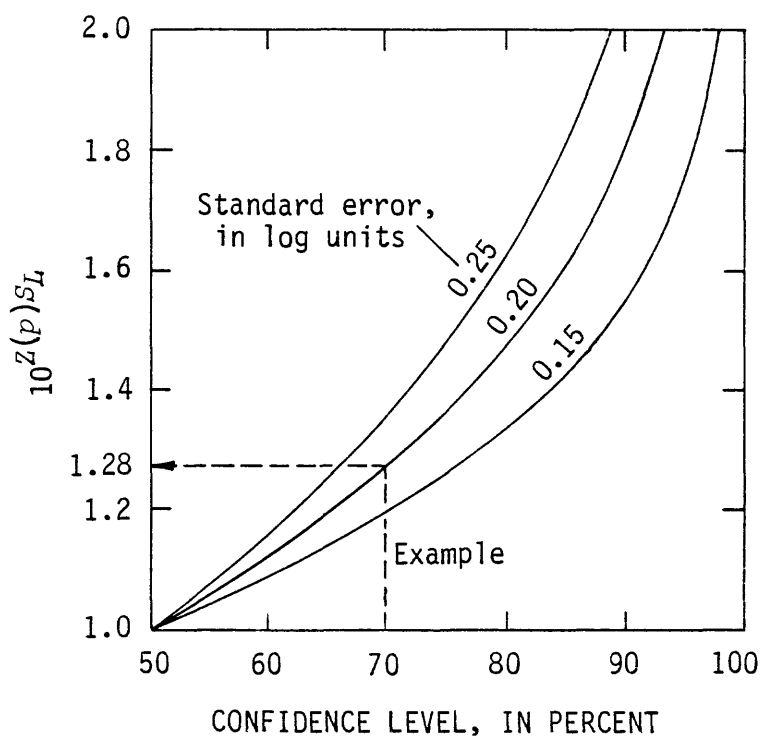
Figure 5.--Uncertainty-adjustment factor.

Maximum Known Floods

The maximum known peak discharges in southern Arizona through September 1981 are shown in figure 6. The discharge values were selected from table 9 in the Data section for gaging stations and from Roeske (1978, table 5) for miscellaneous sites. An enveloping curve for maximum known discharges in the United States (Crippen and Bue, 1977, fig. 2) and the alternate regression equation for the 100-year peak discharge (table 2) are shown for comparison. The enveloping curve is an estimate of the upper limit of potential flood discharges. Peak discharges on the enveloping curve are 8 to 35 times larger than those on the 100-year regression line. Floods more than three times larger than the value on the regression line have been recorded at four sites in southern Arizona. Comparison of a flood magnitude computed by any method with the values in figure 6 provides a simple check on the reasonability of the result.

For drainage areas less than 1 mi^2, most of the maximum floods plotted in figure 6 fall below the 100-year flood curve. The primary reason for these low values is that few floods are measured in such small basins. A 100-year flood that occurs in a 100-square-mile basin has a
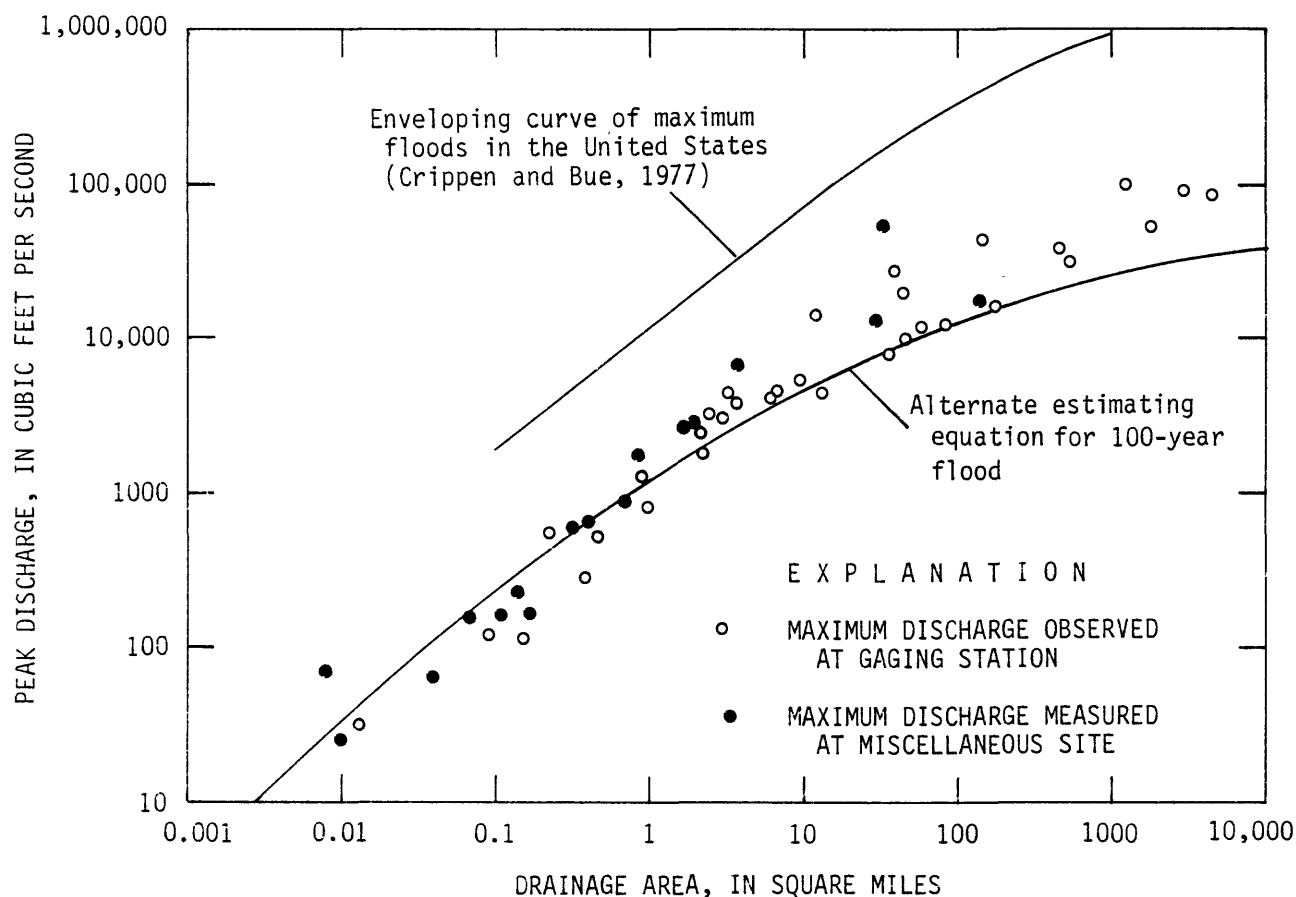
Figure 6.--Maximum floods measured in southern Arizona.

good chance of being measured. Roughly 10 percent of the 100-square-mile basins in Pima County have been gaged (see Basin Characteristics), and a measurement might very well be made for a flow exceeding 10,000 ft³/s even in an ungaged basin. The sampling intensity for small basins, however, is much lower. Only three basins among the 101 studied have drainage areas less than 0.2 mi², but Pima County contains more than 40,000 such basins. A 100-year flood in a basin of 0.1 mi² is only about 200 ft³/s, which is small in absolute magnitude and not easily recognized as outstanding.

DATA ANALYSIS

The development of the techniques presented in Applications is described in this section. Discussions of available data, computation of flood-frequency curves from records of observed floods, division of the complete data set into subsets for each analysis, and regression analyses and statistical tests used to produce the estimating equations are included.

Readers should have a basic familiarity with statistical methods in hydrology. Riggs (1968a, b; 1973) and Reich (1976) provided a good discussion of many of the basic principles. Most statistical analyses were done by using a group of computer programs described by SAS Institute (1982a, b).

A major objective of this study was to describe flood-frequency relationships for basins in Pima County that have drainage areas of less than about 100 mi^2. The data base, however, was not restricted by location or drainage area because the use of more data improved the reliability and applicability of the regression equations. Imposing hydrologically arbitrary boundaries could have resulted in discontinuities between the results of estimating equations for large and small basins or basins within and outside of Pima County.

Flood Records and Frequency Analyses

Records of flood peaks through water year 1981 at 101 gaging stations were used in this study and included 55 in Pima County and 46 in adjacent counties (fig. 1 and table 8). These stations were selected on the basis of a statewide analysis of flood-frequency regions (Roeske, 1978) and include essentially all gaging stations in Arizona that are in and near Pima County. The U.S. Geological Survey maintained continuous discharge records at 29 stations and crest-stage records at 56 stations. Continuous records were maintained by the U.S. Agricultural Research Service at 13 stations and by the Water Resources Research Center, University of Arizona, at 3 stations.

For each station at least 10 years of systematic annual flood records were available, and flow occurred in at least 75 percent of the years. Systematic data are the result of regular observations over a period of time; each peak is the largest in 1 year. From 10 to 67 years of systematic data were available. Seven stations had less than 13 years of record, and 22 stations had more than 20 years of record. Record length tends to increase with drainage area because most early data-collection efforts were concentrated in the larger basins (table 5). Historic flood information was available to extend the period of record at 41 stations. A historic flood peak is the largest in a known period beyond the systematic record. Combined record length ranged from 10 to 97 years, and 23 stations had more than 40 years.

Flood-frequency curves, which relate flood magnitudes to recurrence intervals, were developed from annual and historic peak-discharge data for each station. The analyses followed the guidelines of the U.S. Water Resources Council (1981b) in fitting a log-Pearson Type III distribution to the data. The distribution is defined by estimates of the mean, standard deviation, and coefficient of skewness of the logarithms of annual-flood discharges. Historic flood information, no-flow years, and outliers were treated according to the guidelines. Data from 4 of the 101 stations gave sufficient evidence under the guidelines to reject

Table 5.--Number of gaging stations and variation of record length

Drainage area (mi²)	All stations		Stations used in rural regressions	
	Number	Average years of record	Number	Average years of record
Less than 1	15	16.2	13	16.6
1-10	36	16.1	25	15.4
10-100	22	19.8	20	19.6
100-1000	18	24.7	18	24.7
More than 1000	10	32.7	8	38.0

the log-Pearson distribution; the frequency relations were defined by graphical analysis. All graphical analyses were compared with the results of standard methods and reviewed by the Arizona District Surface-Water Specialist before being accepted for this study.

Because of the high variability of skew estimates from statistically small samples, generalized skew values were weighted with estimates from station data. For rural stations, the generalized skew ranged from -0.2 to 0.1 (U.S. Water Resources Council, 1981b, pl. 1). The generalized skew for urban stations was zero, which is the average of the individual estimates for 10 stations with basin development factors greater than zero. Generalized skew and station skew were weighted in inverse proportion to the standard error of each estimate. The approximate standard error of station skew is a function of record length and the skew magnitude (U.S. Water Resources Council, 1981b); in this study the standard error was generally between 0.1 and 0.6. The standard error of rural generalized skew was 0.2 on the basis of a study of about 30 long-term records in Arizona. Standard error of urban generalized skew was 0.4. Skew estimates from station data ranged from -2.78 to 1.10, and weighted estimates for the 101 stations ranged from -0.63 to 0.34.

An essential assumption of flood-frequency analysis is that the annual flood data series is stationary; that is, the probability of a flood exceeding a given discharge is the same in the last year of the record as it was in the first. Factors such as construction of dams, urban development, and changes in channel systems or climate, however, could cause the probability to change. For an abrupt change, flood data after the change can be analyzed separately, assuming they form a stationary series. If the change is gradual or progressive, no generally accepted method is available to estimate future flood risk. The wide acceptance of

the assumption is reflected by the fact that the possibility of non-stationarity was not even mentioned in a general review of 147 recent flood-frequency research papers (Greis, 1983).

Flood records for the period 1915-81 for the Santa Cruz River at Tucson (09482500) show more large floods in the second half of the period than in the first. Reich (1976, p. 301-302) suggested that urban development or increases in upstream channel capacity could cause the apparent change, and he treated the second half of the record as a stationary series. The presumed causes of change definitely occurred, but they are progressive. Statistical evidence of change was also needed to justify excluding part of the data. In this study, Kendall's rank correlation test (Snedecor and Cochran, 1967, p. 193-195) was used to detect any time trend in the record without considering when the change occurred. The test showed a weak increasing trend that was not statistically significant, even at the 80-percent confidence level. In addition, the record was divided into two equal parts, and each was analyzed as a stationary series. The second half of the record had a larger mean but smaller standard deviation. For large recurrence intervals, the two frequency curves converged and, for recurrence intervals greater than 4 years, both short-record curves were within or below a 60-percent confidence band around a complete-record curve. Because no significant difference was found, the frequency curve that was based on the entire record was accepted for this study.

Basin Characteristics

The 101 basins cover a wide range of conditions, both rural and urban. Eleven basin characteristics, which were selected for their possible value in predicting flood magnitudes, were determined for each gaging station. These characteristics were drainage area, mean elevation, main-channel length and slope, average annual precipitation, 24-hour precipitation depth at 2- and 50-year recurrence intervals, forest cover, average annual snowfall, mean minimum January temperature, and basin development factor. In addition, a shape factor was computed from main-channel length and drainage area. The basin characteristics used in applying the estimating equations are described in table 3 and figure 2, and values for each station are listed in table 8.

Additional characteristics were determined for most of the stations. The portion of drainage area near mountain fronts was determined for 75 basins with drainage areas less than about 200 mi^2 because increased precipitation rates in this zone might influence flood magnitudes. For 87 stations with drainage areas of less than about 500 mi^2, the portions of drainage area in four hydrologic soil groups were determined. The soil groups, which describe the runoff potential of different soils (U.S. Soil Conservation Service, 1972), were determined from a generalized soils map of each county at scales of 1:250,000 or 1:500,000. Because the additional characteristics were not statistically significant in preliminary regression analyses, they were not determined
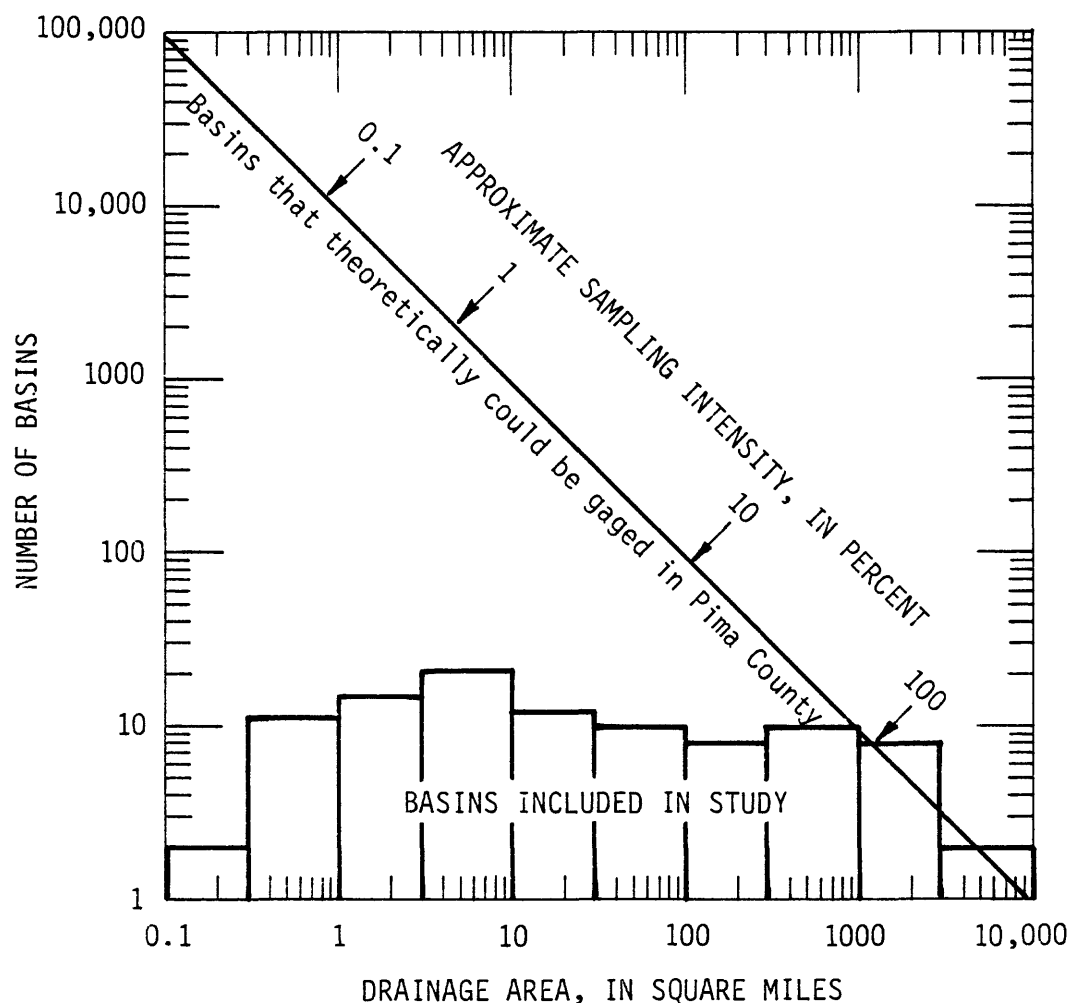
Figure 7.--Number of basins and sampling intensity.

for the larger basins and are not included in the estimating equations. An unusually high error rate in determining hydrologic soil groups from the generalized maps was an additional reason for excluding soil data.

The 101 gaging stations were considered representative of the wide range of drainage-basin characteristics in Pima County. Stations are scattered throughout the county, with a concentration in the more populated eastern part; 46 stations are outside Pima County (fig. 1). The gaged basins are evenly distributed over a wide range of drainage areas from 0.013 to 4,471 mi^2 (table 5 and fig. 7). In terms of the ratio of basins sampled to the number of available basins, however, the larger drainage areas were sampled much more heavily. In the 9,240 mi^2 of Pima County, there theoretically could be 92 basins of 100 mi^2, 924 basins of 10 mi^2, and so on, assuming no overlap between basins or across county

lines. The sample of 101 basins is not exactly comparable, because some basins overlap and some are outside the county. Approximately 10 percent of all possible 100-square-mile basins were sampled but only 0.1 percent of 1-square-mile basins (fig. 7).

Land elevation in Pima County ranges from 660 ft in Growler Wash west of Ajo to 9,160 ft on Mount Lemmon north of Tucson; 93 percent of the area of the county is between 1,000 and 5,000 ft (fig. 8). Mean elevations of the 101 basins range from 600 to 6,300 ft; 92 basins are between 1,000 and 5,000 ft. However, 52 percent of the basins are between 3,000 and 5,000 ft compared to only 26 percent of the area of the county. More gages were in the 3,000- to 5,000-foot elevation range because of data needs for the metropolitan area of Tucson and the Walnut Gulch research watershed near Tombstone (fig. 1). Elevation is not correlated with drainage area in this sample of 101 basins (fig. 9A).

Main-channel slope for the basins ranges from 0.3 to 13 percent and has a weak negative correlation with drainage area (fig. 9E). Slope tends to increase and become more variable as drainage area decreases. None of the gaged basins that have areas less than 1 mi^2 have slopes less than 1 percent. Slope also tends to increase with elevation (fig. 9B), but the relation actually is between slope and topographic position in mountainous areas. Channel slopes greater than 13 percent occur along mountain fronts, generally in basins of less than 5 mi^2, and probably in less than 5 percent of the county.

The shape factor (main-channel length squared, divided by drainage area) for the 101 basins ranges from 1.47 to 20.6 and is not correlated with drainage area, elevation, or slope (fig. 9C, D, and F). The factor is roughly equivalent to the ratio of basin length to width. Elongated basins in Pima County are generally found on alluvial slopes that extend away from mountain ranges. Parallel gullies develop on the slopes, and the drainage area of each gully is generally small. A few such basins in Pima County may have shape factors greater than 20. Main-channel length was not used directly in the equations because it was highly correlated with drainage area (r>0.95).

The basin development factor (BDF, fig. 2) has a maximum possible value of 12. For 10 urban basins in the study, BDF ranges from 1 to 9, and 8 different values of BDF are included. The BDF is zero for 91 rural basins.

Grouping of Basins for Analyses

Although the 101 basins are fairly uniformly distributed over the range of predictor variables used in the rural estimating equations (fig. 9), the basins are a disproportionate sample of some other characteristics. Initial regression analyses, which attempted to include all basins in a single equation, were unsuccessful because lightly sampled characteristics were statistically insignificant. The regression coefficients
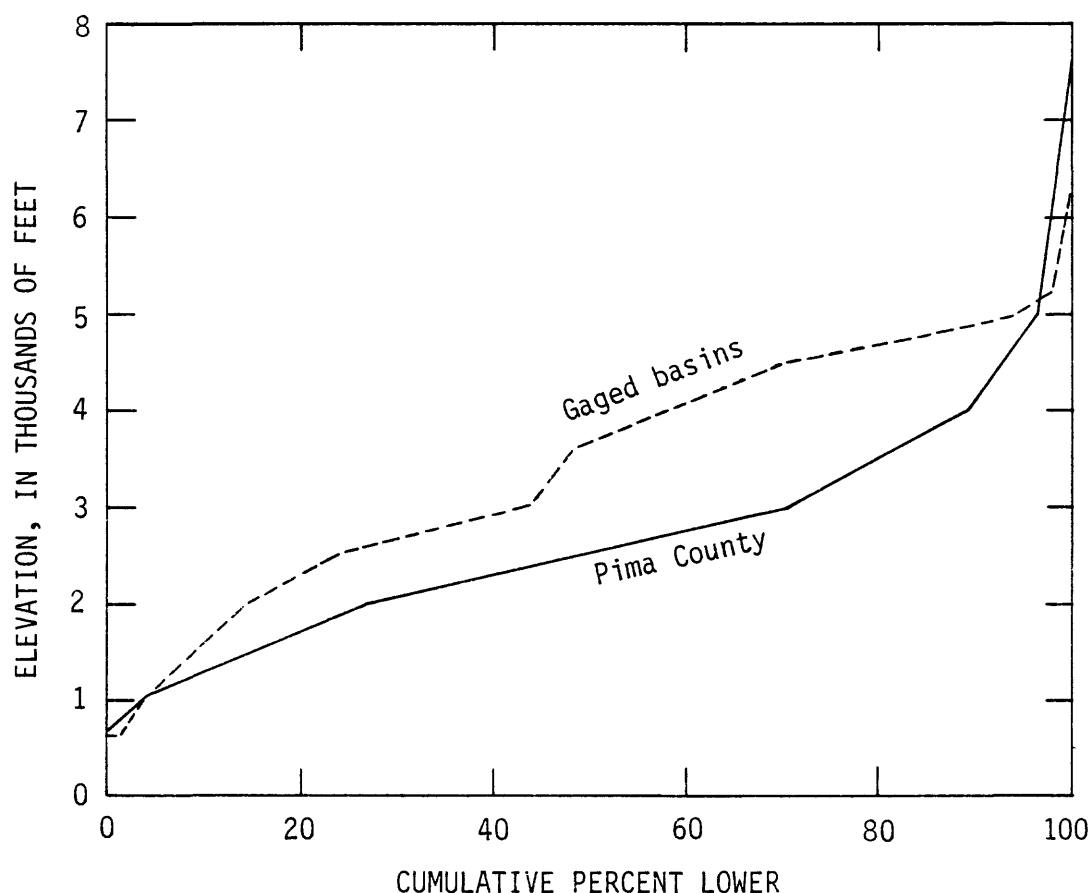
Figure 8.--Distribution of elevation in Pima County and gaged basins.

were biased toward the characteristics of a heavily sampled area. Subsets of the 101 basins therefore were selected for various analyses.

The Walnut Gulch research watershed of the U.S. Agricultural Research Service includes 13 gaging stations within one 57.7-square-mile basin near Tombstone. Because the stations measure flow from a limited range of climate, elevation, topographic position, and soil, their flood-frequency characteristics are strongly correlated. Four stations were excluded from the regression analysis because of the correlation (group C, table 8); however, they were used in estimating the standard error of prediction. In general, one of the subbasins was excluded if its drainage area differed by less than 40 percent from the area of the next larger or smaller basin. Half of the retained basins were more than three times larger than the next smaller basins. Although one pair of basins differed in area by only 23 percent, both were retained because they had the largest and one of the smallest shape factors (20.6 and 2.43) of the entire 101 basins. Drainage areas of the nine retained basins in the Walnut Gulch watershed range from 0.013 to 57.7 mi^2.
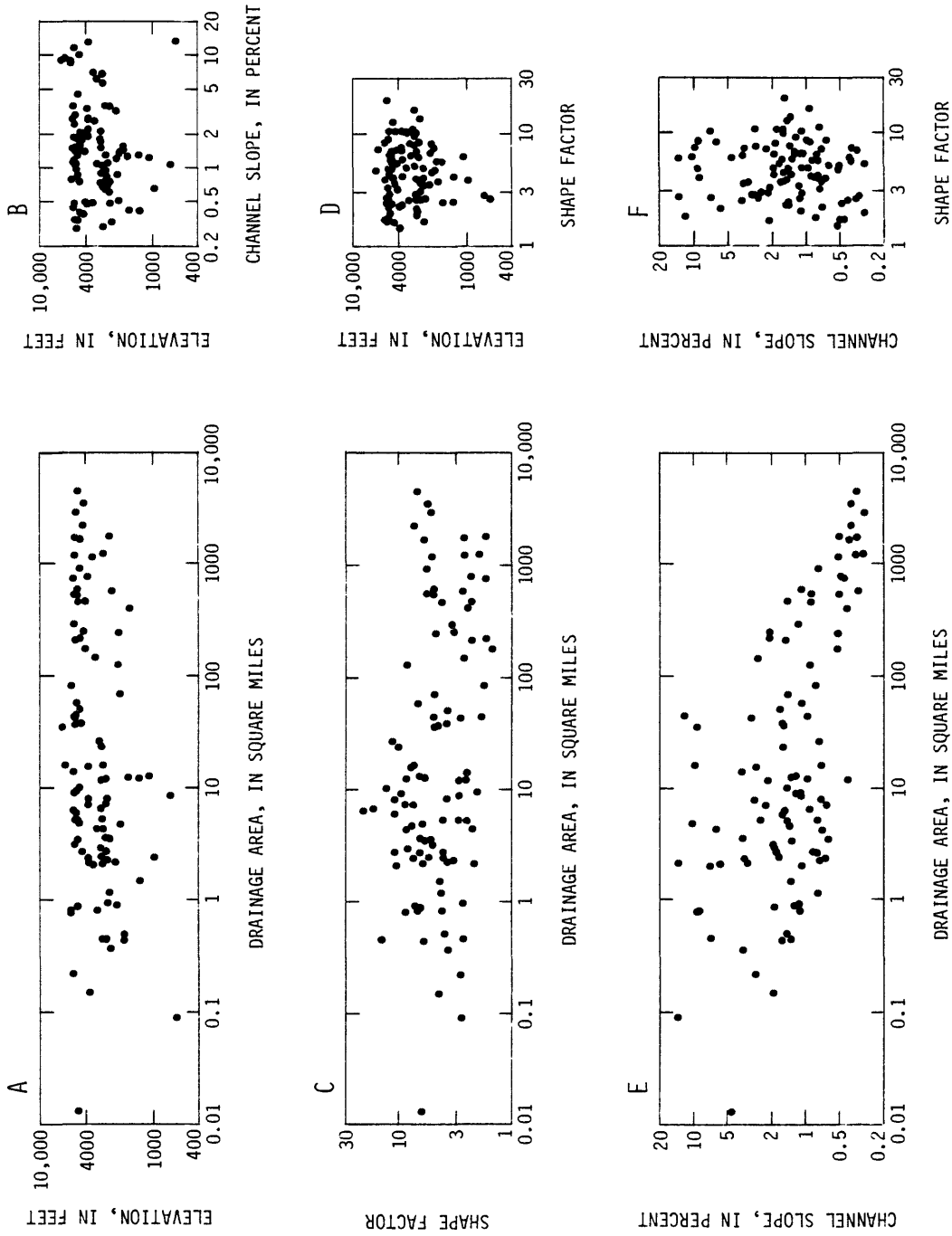
Figure 9.--Joint distributions of basin characteristics.

Ten stations for which the BDF reflects some degree of urban development were initially analyzed separately from the rural stations. Urbanization is generally assumed to increase flood peaks. For three of the stations, however, flood magnitudes at all recurrence intervals were much smaller than those estimated by preliminary rural regressions. The three basins were the least urbanized (BDF<3), and the area of urban development is in the lower part of each basin. In addition, the three basins are more elongated (shape>7.2) than six of the other seven basins. Because of the channel improvements, flood peaks from the upper and lower parts of the basin may be desynchronized. Therefore, the three least urbanized stations were analyzed with the rural stations, and only seven stations were used for urban analyses (group U, table 8).

Three rural stations were excluded from the regression analyses because channel characteristics in the basins can cause extreme attenuation of flood peaks (group A, table 8). In a reach that has an extremely small channel and wide flood plain, both infiltration losses and temporary storage can be large enough to reduce peak discharge substantially. Such a channel reach is present in each of the three rural basins.

Three other rural stations were excluded from the regression analyses because their flood magnitudes at all recurrence intervals were much smaller than magnitudes estimated by preliminary rural regressions (group L, table 8). They were used, however, in some comparisons of alternative methods. The largest peak at each station during 13 to 18 years of systematic record was less than the 5-year flood estimated by regression. The stations have no extraordinary combination of basin characteristics and are widely separated in the study area. The probability that all floods at a site during 15 years would be less than a 5-year flood is 0.035. In a sample of 101 stations, it would not be surprising to find 3.5 such sites (0.035 times 101). The conclusion, therefore, is that these three stations were low outliers only because of the chance absence of a large flood during the period of record and not because of any unique basin or climatic conditions.

The Cienega Creek near Pantano gaging station (09484560) provides an example of the variation of flood-frequency estimates from short records. The 100-year flood was estimated from 10 years of peak-flow record (1968-77) as 6,140 ft³/s, which is less than the regression estimate of the 5-year flood. The next 4 years produced the largest flood in the record and three of the four smallest floods. The revised estimate from 14 years of record is 14,100 ft³/s (table 8), which is still less than the regression estimate for a 100-year flood; however, the station is no longer an extremely low outlier. Thus at this site, 10 years was an inadequate period of time to sample the variability of floods.

In summary, 84 stations were used to develop estimating equations for rural sites (group R, table 8), and 4 more stations were used in estimating prediction error. Seven stations were analyzed for the effects of urbanization. Three stations were used to check the possible magnitude of peak attenuation, and three other stations were used only in comparing alternative methods.

Regression Analyses for Rural Basins

Statistical regression analyses were performed to develop a method of transferring flood information from the few gaged stations in and near Pima County to many ungaged sites (fig. 7). The analyses related measurable basin or climatic characteristics to flood magnitudes that were derived from observed floods. The analyses in this study were generally multiple regressions, which means that more than one predictor variable was used.

A fundamental assumption of regression analysis is that the variance of residual errors is equal throughout the range of each predictor variable (Draper and Smith, 1981, p. 22-24). Residual error is the difference between a flood magnitude estimated from a regression equation and the corresponding observed magnitude. Logarithms of the variables are usually used in regression analysis because the variance of flood magnitudes measured in cubic feet per second increases as the absolute magnitude increases (Riggs, 1968a, p. 10-11). All regression analyses in this study used base-10 logarithms of the data.

Draper and Smith (1981, p. 45) pointed out that regressions measure the correlation between variables but do not prove that the predictor variables cause the observed response. The appropriate form of the estimating equations cannot be specified precisely in advance because the mechanisms of runoff generation from watersheds are imprecisely known. Previous studies, however, provide guidance in the selection of predictor variables and the range of reasonable regression coefficients. The variables selected for analysis describe possible causal mechanisms and the range of conditions present in the study area. Many alternative models were compared to select the best model form and predictor variables.

Statistical techniques are available to distinguish lack of fit in a model from pure error, and therefore, to test the adequacy of the regression model to describe the process (Draper and Smith, 1981, p. 33-42). Pure error is the inescapable variation of the physical process. The techniques, however, require independent repeated observations of the response variable for which the values of all predictor variables are identical. Repeated observations were not available in this study and are rarely available in hydrology because of the large number of possible predictor variables. The estimating equations, therefore, should be considered as a method of interpolating between a few known values of an unknown function. The equations may produce unpredictably large errors if applied beyond the range of variables used in their development.

Alternative regression models can be compared quantitatively and qualitatively. Quantitative comparisons include the standard error and coefficient of determination of each model and the statistical significance of the coefficients. Standard error measures how well flood magnitudes estimated from a regression equation agree with the observed

magnitudes. Standard error measured in log units is equal above and below the regression line (fig. 4) but becomes unequal when transformed to percent. Percent standard errors given in this report are the average of the positive and negative errors. The coefficient of determination (R^2) is the proportion of the variance of the response variable that is explained by the predictor variables. Statistical significance is the probability that one or several coefficients would be smaller than the estimates from a regression analysis if no relation existed between the predictor and response variables. One qualitative test of a potential model form or predictor variable is whether it is reasonably related to the conceptual mechanisms which produce floods.

Many regression models were examined in this study generally using the 25-year flood as the response variable during exploratory analyses. Models that appeared promising on the basis of quantitative and qualitative tests were applied to 2-, 5-, 10-, 25-, 50-, 100-, and 500-year floods. For each model form, predictor variables were selected using both step-forward and step-backward techniques (Draper and Smith, 1981, p. 305-312). A step-forward analysis adds one predictor variable to the model at each step by selecting the most significant variable not previously selected. A step-backward analysis initially includes all potential predictor variables and deletes one at each step by selecting the least significant remaining variable. Both techniques generally identified the same final set of predictor variables.

Linear Model

The analyses initially considered models of the form

$$Q_T = 10^{B_0} C_1^{B_1} C_2^{B_2} \ . \ . \ . C_n^{B_n} 10^E$$

where

Q_T is T-year peak discharge, in cubic feet per second;

B is an estimated coefficient;

C is a measurable basin or climatic characteristic; and

E is residual error.

Taking the common logarithms of both sides and letting $C_0 = 10$ ($LogC_0=1$), the model assumes the linear form

$$LogQ_T = \sum_{i=0}^{n} B_i LogC_i + E, \tag{4}$$

for which the coefficients $B_0, ..., B_n$ can be estimated by standard multiple-regression techniques. This model has been widely used in regional flood-frequency studies (Thomas and Benson, 1970; Riggs, 1973; Roeske, 1978). Conceptually, it says that flood magnitudes respond linearly to changes in basin characteristics for all values of the characteristics and that the effect of changing one characteristic does not depend on the value of any other characteristic. For regions where most characteristics have small ranges, the model is a reasonable approximation. However, it tends to break down where the range of characteristics is large. Riggs (1973, p. 10) noted that the ". . .model is not adequate for semiarid regions of large relief."

The best linear regression equations included only drainage area and slope as predictor variables. All the available characteristics were tested, primarily by using step-backward methods. Standard errors for the regressions were 47 to 71 percent (fig. 10), and R^2 was 0.85 to 0.91 for recurrence intervals up to 100 years. The regression coefficients, however, are not reported here because the quadratic model, which used drainage area as the only predictor variable, had smaller standard errors and equal or larger R^2 for all recurrence intervals.

Quadratic Model

A quadratic model with drainage area (A) as the single predictor variable was considered next, because it had been useful in a flood-insurance study of Pima County (Federal Emergency Management Agency, 1982a, b). The model was

$$LogQ_T = B_0 + B_1 LogA + B_2(LogA)^2 + E \qquad (5)$$

with the terms as previously defined. For all recurrence intervals greater than 2 years, the B_2 coefficient was significantly less than zero at the 99.8-percent confidence level. The quadratic model gave standard errors of regression 2 to 7 percentage points less than those for the linear model (fig. 10). Coefficients for the quadratic model are given in table 2 as the alternate estimating equations for Pima County.

Conceptually, the quadratic model assumes that the rate at which flood magnitudes increase with basin area is less for large basins than for small ones. The high significance level of the B_2 coefficient is strong evidence that the relation of flood magnitudes to drainage area is curvilinear in log space. The quadratic model approximates the curvature by a transformation of LogA; some other transformation might represent the physical processes more accurately.

An attempt was made to improve the quadratic model by including additional linear terms for other characteristics. This represents the smallest conceptual departure from the linear model by
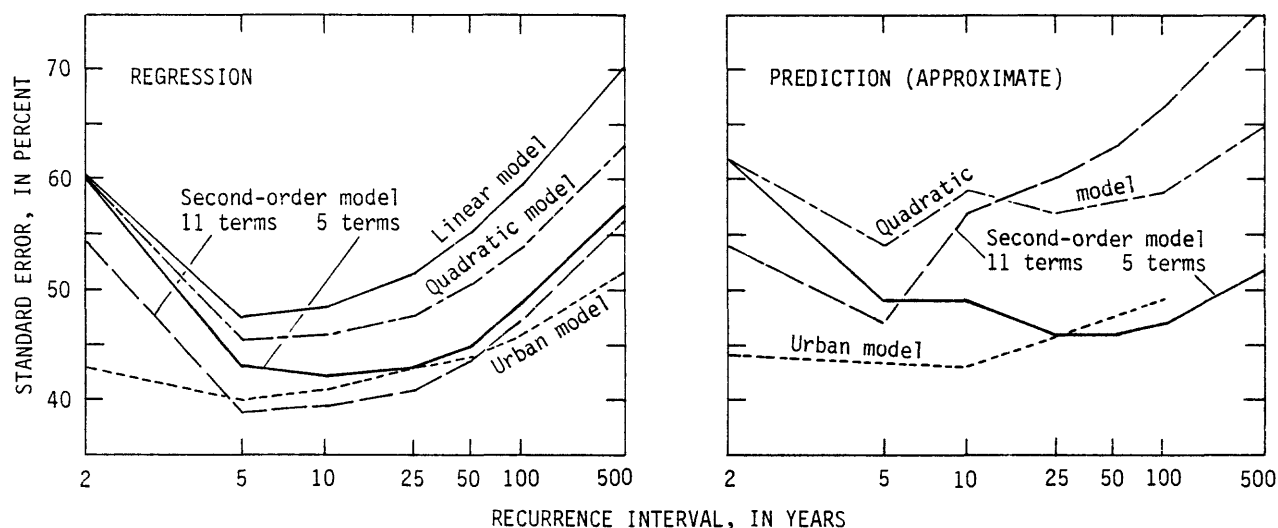
Figure 10.--Comparison of standard errors for selected equations.

allowing a nonlinear term only for drainage area. Using primarily a step-backward approach, only forest area and minimum January temperature were found to be statistically significant at the 95-percent confidence level. The variables were of little practical significance in reducing standard error, however, and their direct effect on flood-producing mechanisms was not obvious. Forest area is close to zero in most of the 101 basins and is weakly correlated with elevation, precipitation, and snowfall. Temperature is correlated with elevation and longitude in this data set and may enter the regressions as a surrogate for unmeasurable geographic variation in soils or vegetation. Because of the uncertainties, equations using these variables were not acceptable. Several variables describing soil type or mountain-front area were tested with the quadratic model in smaller data sets; none of the variables was statistically significant.

Generalized Second-Order Model

Although the quadratic model produced standard errors that were small relative to most previous studies, the errors were still large in absolute terms. Because one objective of the study was to determine the best possible estimating methods, further trials were made to discover whether some other model could transfer more information from gaged to ungaged sites and result in smaller errors.

A variety of evidence led to an expanded conceptual model. First, several variables that conceptually should affect flood magnitudes had not been used. Second, lines enveloping the maximum known floods in 17 regions of the United States are all curvilinear and concave

downward with respect to log drainage area (Crippen and Bue, 1977). The same shape might be expected for floods of a specific recurrence interval. Third, a curvilinear and concave downward effect for area had been proven statistically in this study. Fourth, it seemed likely that doubling the slope could have a different effect if the original slope were 1 percent or 10 percent. Finally, the effect of a given slope might be different in a long narrow basin and a short wide basin. Conceptually, therefore, a basin's flood response in log space might be curvilinear with respect to several basin characteristics, and the characteristics might interact.

A generalized second-order polynomial is a simple first approximation of the conceptual model in the presence of a large error term, although the physical processes may be log-log or involve some other transformation. The statistical model was

$$LogQ_T = \sum_{i=0}^{n} \sum_{j=i}^{n} B_{ij} LogC_i LogC_j + E \qquad (6)$$

using the terms defined previously for the linear model. Because $LogC_0$ was defined as 1, the model included a constant (when $j = i=0$), linear terms ($j > i=0$), quadratic terms ($j=i>0$), and cross-product terms ($j>i>0$). Equation 6 can produce complex models with many terms from only a few predictor variables; the analysis can easily degenerate into an exercise in computation. Although the following discussion is quite detailed, it provides the necessary hydrologic reasoning that supports the numerical manipulations and the final model.

The second-order model was tested initially by a stepwise process of adding or removing at one time all the terms that included a given basin characteristic. It was conceptually simpler to assume all effects were second order than to assume a different order or interaction for each predictor variable. Draper and Smith (1981, p. 220) noted that ". . .omission of terms implies possession of definite knowledge that certain types of [equations] (those which cannot be represented without the omitted terms) cannot possibly occur. Knowledge of this sort is not often available. When it is, it would usually enable a more theoretically based study to be made."

Because almost all cross-product terms which included drainage area were not significant, they were eliminated:

$$LogQ_T = B_1 LogA + B_2 (LogA)^2 + \sum_{i=0}^{n} \sum_{j=i}^{n} B_{ij} LogC_i LogC_j + E. \qquad (7)$$

The B_1, B_2, and B_{00} terms represent the large differences in flood magnitudes between basins of different sizes. The remaining terms of the summation describe smaller variations between basins of similar size.

Step-forward and step-backward analyses showed that minimum standard errors of regression were achieved by an 11-term model that had slope, shape, and elevation terms (equation 7). Standard errors of regression ranged from 39 to 57 percent and averaged 6.3 percentage points less than those for the quadratic model (fig. 10); R^2 was about 0.03 larger than for the quadratic model. Each of the four predictor variables in the resulting equations can be measured on topographic maps. The linear, quadratic, and cross-product classes of terms each were significant. Slope, shape, and elevation factors—all terms that included each variable—were significant for most recurrence intervals. The 11-term model seemed to describe the data best.

For each recurrence interval, however, at least four terms were not statistically significant in the regressions. Small changes in the sample of gaging stations produced large changes in the estimated coefficients. The unstable coefficients indicated the model might be overspecified. A regression model that includes too many terms may yield a small standard error of regression, but errors in predictions for sites that were not used to develop the regression may be large. The standard error of prediction for the 11-term model was substantially greater than the standard error of regression for recurrence intervals greater than 5 years (fig. 10). Standard error of prediction was greater than that of the quadratic model for recurrence intervals of 25 years or more. The 11-term model was rejected. The standard error of prediction is described in the section on prediction error.

After rejecting the 11-term model, the generalized second-order model was re-examined by treating each term individually in step-forward and step-backward analyses. Standard errors of both regression and prediction were lowest using area, slope, and shape as predictor variables:

$$LogQ_T = B_0 + B_1 LogA + B_2(LogA)^2 + B_3 LogS + B_4(LogS)^2$$

$$+ B_5(LogS)(LogSh) + E. \tag{8}$$

For most recurrence intervals, each of the five terms was significant at the 99-percent confidence level. The residuals had approximately a normal distribution and showed no consistent trends when plotted against predicted values or most excluded variables. Standard errors of regression were about 5 percent less than for the quadratic model, and standard errors of prediction were about 10 percent less (fig. 10). R^2 was about 0.02 greater than for the quadratic model and about 0.01 less than for the 11-term model. The five-term second-order model produced the best overall results, and the coefficients are used as the primary rural estimating equations for Pima County (table 1).

For all recurrence intervals, the coefficients of the two linear terms in equation 8 are positive, and the quadratic and cross-product coefficients are negative (table 1). The predicted values are curvilinear

with respect to log drainage area and concave downward (fig. 11A, curve P). The largest variation in flood magnitudes is related to the size of the basins; predicted values vary by about three orders of magnitude from the smallest to largest basin in the study. Combinations of slope and shape produce departures from the base curve (fig. 11B). For most recurrence intervals, the maximum departure is about four times the minimum. Predicted 100-year discharge generally increases with increasing slope for slopes up to about 2 percent. For larger slopes the estimates decrease, especially for elongated basins.

The final statistical model fits comfortably with the conceptual model. Geomorphology should integrate the runoff characteristics of basins so most information necessary for flood estimates should exist on topographic maps. Nonlinear effects should be found for variables with large ranges, and interactions among variables should be possible. The statistical model has these attractive conceptual characteristics. Because all the terms in the model were highly significant, area, slope, and shape seem to measure factors that are closely related to flood-producing mechanisms. The primary estimating equations are more complex than previous equations but are substantially more accurate. For users who are uncomfortable with this approach, the quadratic model is available with a fairly small decrease in accuracy.

For most recurrence intervals, one or two additional terms were statistically significant in predicting flood magnitudes. Each of the significant terms included elevation, which also can be measured from topographic maps. No individual term, however, was significant for more than three of the seven recurrence intervals examined. Inclusion of the two most commonly significant terms gave little practical improvement in standard error for recurrence intervals greater than 5 years. Even though a few equations would have smaller standard errors if a different regression model were used for some recurrence intervals, the five-term model was used for all recurrence intervals in order to maintain the correlation structure of the estimated flood magnitudes.

Residuals of the five-term model were examined further by plotting them on maps and against elevation and longitude. For most recurrence intervals, a cluster of positive residuals was present in the upper San Pedro River basin. The elevation and longitude plots showed the same cluster. Flood-frequency estimates from observed floods were generally greater than estimates from the five-term model for the Walnut Gulch experimental watersheds near Tombstone and for the main stem of the San Pedro River. None of the 16 topographic, climatological, and soils variables examined was adequate to explain the difference in flood production in the upper San Pedro River area. Basins with similar characteristics in other parts of the study area did not show similar differences in flooding. Basins in the San Pedro River area may differ from the remainder of the study area in some unmeasured basin characteristic, or the data in that area may be subject to a consistently different time-sampling error. The differences, however, were not large enough to justify adjusting the equations.
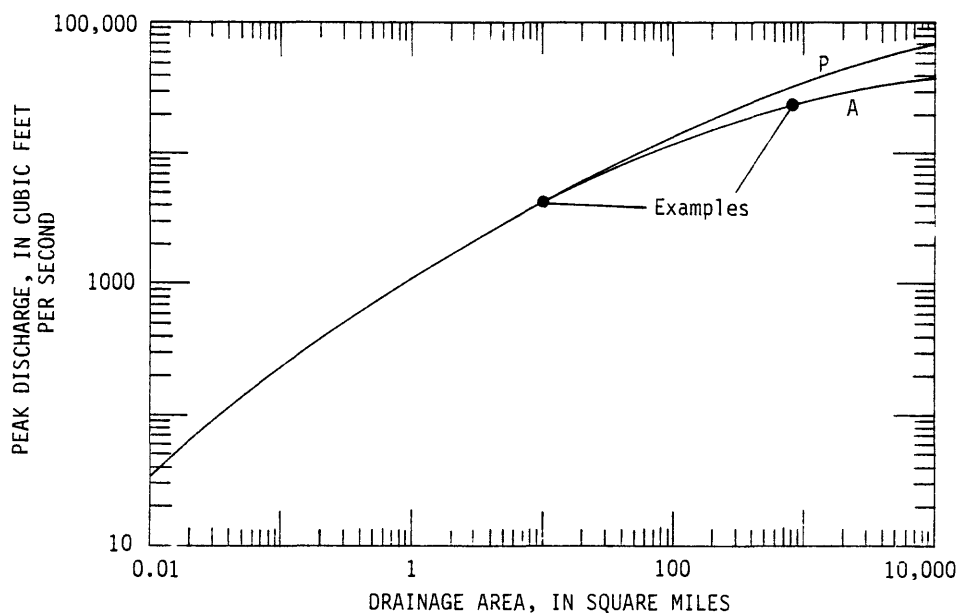
A. Base curves
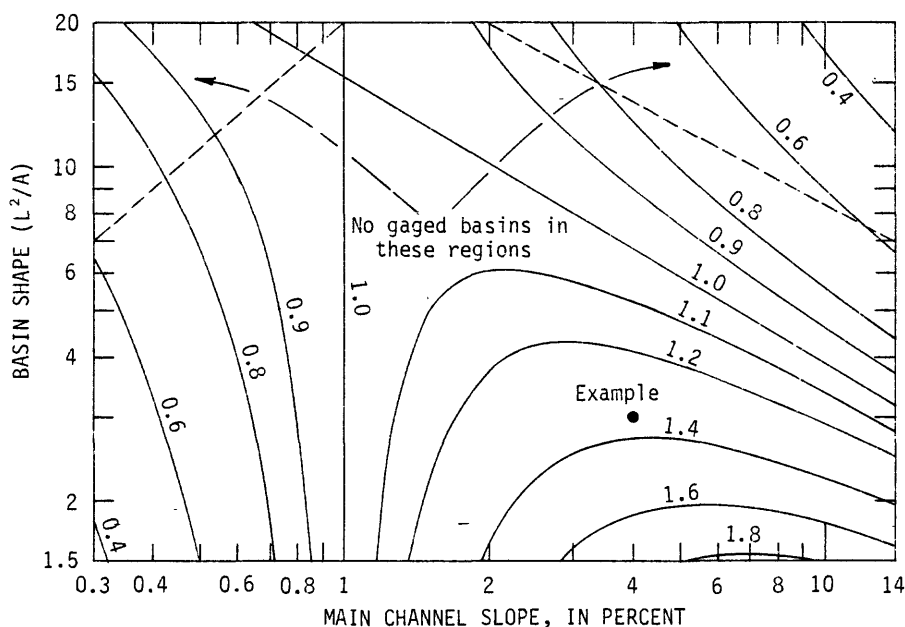
B. Multiplier for primary equation

EXPLANATION

1. To estimate a 100-year flood using the primary equation (table 1):
   a. Read a value from curve P in graph A for the required drainage area,
   b. Read a value from graph B for the required slope and shape, and
   c. Multiply the two values.
       Example: For area=10 mi^2, slope=4 percent, and shape=3:
           From graph A   4,400 ft^3/s
           From graph B   X1.35
           100-year flood estimate   5,900 ft^3/s
2. To estimate a 100-year flood using the alternate equation (table 2), read the value directly from curve A in graph A.
       Example: For area=800mi^2, the 100-year flood estimate is 25,000ft^3/s.

Figure 11.--100-year flood estimates.

Prediction Error

The standard error of regression has been used to compare the accuracy of several regression models, but it only measures how well each equation fits the original data. The standard error of prediction is a measure of an equation's accuracy for sites not used in developing the equation. The standard error of prediction, which is usually larger than the standard error of regression, is a function of the pure error, the lack of fit of the model, and the values of the new predictor variables relative to the original data (Draper and Smith, 1981, p. 83-85, 117-121). No simple equation will give the exact standard error of prediction for any site at which the equations might be applied.

Standard error of prediction can be approximated, however, by split-sample analysis (Snee, 1977). The complete data set was divided into groups A and B, and regression coefficients were computed using only group A. Predicted values and their differences from observed values were computed for group B. The ratio of the standard error of the differences to the standard error of regression from group A should be approximately equal to the ratio of the standard errors of prediction and regression for the complete data set. Mathematically,

$$SE_B/SE_A = SE_P/SE_R$$

or

$$SE_P = (SE_R)(SE_B)/SE_A \qquad\qquad (9)$$

where

SE_B is the standard error about zero of the differences between estimated and observed values for group B;

SE_A is the standard error of regression for group A;

SE_P is the standard error of prediction for the complete data set; and

SE_R is the standard error of regression for the complete data set.

Estimates of prediction error were made concurrently with the analysis of the various regression models. Because as many as 12 coefficients were estimated for equation 7, 12 more sites were assigned to group A than to group B. The computed standard errors for each group, therefore, had the same number of degrees of freedom. The basins with the largest and smallest values of each predictor variable were assigned to group A. The four basins near Tombstone that were

excluded from previous analyses because of correlation (table 8) were put in group B, and four nearby basins with similar drainage areas were put in group A. The remaining usable rural basins were sorted by drainage area and assigned to the two groups in the necessary proportion, with approximately every other one going to each group. Group A included 50 basins; group B included 38. Different groupings of the basins would yield different estimates of the standard error of prediction, but a new grouping was not made when equation 7 was rejected.

The standard errors of regression and prediction for the primary (5-term) and alternate (quadratic) estimating equations are shown in figure 10. The difference between the regression and prediction errors is reasonably small, so the standard error of regression is adequate for computing weighted estimates at gaging stations (equation 1) and for computing the adjustment for uncertainty (equation 3). The standard errors of regression and prediction for the 11-term, second-order model (equation 7) are also shown in figure 10. Although the 11-term model gave lower standard errors of regression and eliminated the cluster of positive residuals near Tombstone, the coefficients could not be estimated accurately. The standard error of prediction revealed the problem and led to rejection of the 11-term model.

Attenuation

The peak discharges of many floods in southern Arizona decrease as they move downstream. Certain combinations of channel characteristics, however, can cause extreme attenuation. For example, during the flood of October 1977, the peak discharge of the Santa Cruz River at Cortaro was 23,000 ft^3/s or about a 25-year flood (Aldridge and Eychaner, 1984). In the 90 river miles downstream to the next gaging station, the floodwaters spread to a width of as much as 4 mi in two reaches totaling 25 mi in length. The resulting shallow depths and low velocities reduced the peak to 2,010 ft^3/s at the downstream station, which is less than a 4-year flood. The qualitative understanding of such attenuation, however, has been difficult to apply in improving quantitative estimates of flood magnitudes.

Three of the 101 basins, which attracted attention as low outliers in the initial regressions, were found to contain channel reaches that cause extreme attenuation. The significant reaches are characterized by their length, channel capacity, flood-plain width, and position within the basins. The identifying characteristics given in "Adjustment for Attenuation" were based on the reach characteristics in the three basins.

The effect of attenuation was estimated by the ratio of a flood magnitude computed from gage records to the equivalent magnitude computed from the estimating equations (tables 1 and 2). The individual ratios ranged from 0.11 to 0.45 for the three sites and seven recurrence intervals; that is, at-site flood magnitudes were one-tenth to one-half as large as those computed from regional equations. The average ratio for a

single recurrence interval was 0.26 to 0.30. Because the sample ratios may be affected by the short records or particular basin characteristics of the three sites, the largest computed ratio was used to adjust for extreme attenuation. Rounding the largest computed ratio to one significant figure gives 0.5 as a multiplier to adjust the estimating equation values at all recurrence intervals.

Urban Basins

Sauer and others (1983) developed two sets of estimating equations in a national study of urban flood frequency. The simpler equations use drainage area, equivalent rural discharge, and basin development factor (BDF) to compute urban flood magnitudes. The more complex equations add slope, storage, impervious area, and rainfall intensity as predictors. The standard error of prediction, however, is as small for the simpler equations as for the more complex (Sauer and others, 1983, table 3). The predictors of slope, storage, impervious area, and rainfall intensity are either included in the primary estimating equations in table 1 or have a small range in the Tucson area.

Seven urban basins in Tucson were available for analysis, but that number was too few to develop reliable new estimating equations. Four of the seven basins were also used by Sauer and others (1983). The national equations were tested by computing the residual difference between flood magnitudes estimated from observed floods at the seven sites and those computed from the equations. The means of the residuals were nearly zero, and their standard deviations were about one-third less than the equations' standard errors of regression. The national equations, therefore, are unbiased for the Tucson area, and the reported estimating errors are reasonable. The three-parameter national equations were selected for use in this report (table 4).

For the same seven basins, ratios were computed of the flood magnitudes estimated from observed floods and from the primary estimating equations. Averages of these ratios, which could be used to approximate urban flood magnitudes, are shown in the section "Ungaged Urban Sites."

Weighting of Gage and Regression Estimates

At gaged sites, two nearly independent estimates of flood magnitudes are available—one from the regional estimating equations and one from analysis of the observed flood record. Weighting the two estimates inversely by their variances forms an estimate that has a smaller variance than either one. The following paragraphs discuss two methods of weighting the estimates.

The variance of the regression estimate is approximately the square of the standard error of regression. For a log-Pearson III distribution with known skew, the variance of the gage estimate is a function of the number of years of record, the recurrence interval of interest, and the standard deviation and skew of the annual floods (Hardison, 1969, p. D213):

$$SE_G = S_{LQ} \; R/\sqrt{N} \qquad\qquad (10)$$

where

SE_G is the standard error—square root of the variance—of the gage estimate of the T-year flood, in log units;

S_{LQ} is the standard deviation of the logs of the annual floods;

R is a given function of recurrence interval, T, and the skew of the logs of the annual floods; and

N is the record length.

The gage and regression variances can be indexed in terms of N (Hardison, 1971). The record length N_G for a single gage indexes $(SE_G)^2$. Manipulation of equation 10 gives:

$$N = R^2(S_{LQ}/SE_G)^2$$

from which an equivalent record length N_R is obtained by substituting the standard error of regression SE_R for SE_G and using average regional values of S_{LQ} and skew:

$$N_R = R^2(S_{LQ}/SE_R)^2 \qquad\qquad (11)$$

Hardison (1971, table 2) reported approximate values for R. Note that R^2 in equation 11 is not the regression coefficient of determination, for which the term R^2 is also used.

The equivalent-years method has been suggested for computing weighted estimates of flood magnitudes (U.S. Water Resources Council, 1981b, appendix 8):

$$LogQ_W = \frac{N_G LogQ_G + N_R LogQ_R}{N_G + N_R}, \qquad\qquad (12)$$

where Q_W, Q_G, and Q_R are the weighted, gage, and regression esti-mates. The method, however, causes some conceptual difficulties. The method assumes that all equal-length records will support equally accurate gage estimates, and it ignores differences among stations in the standard deviation and skew of the annual floods. The index of regression error is calculated from the average variance and skew of gage records, and several different forms of the averages might be used. Variance of the weighted estimate is not generally computed, but would be given by

$$V_W = C/(N_G + N_R),$$

where C is an unevaluated constant.

In this study, equation 10 was evaluated directly for each station and led to the following weighting equation, which was given previously in Applications:

$$LogQ_W = \frac{(LogQ_R)(SE_G)^2 + (LogQ_G)(SE_R)^2}{(SE_G)^2 + (SE_R)^2} \tag{1}$$

where SE_R is the standard error of regression, in log units.

Stedinger (1983, p. 21) stated that Hardison's analysis produced an asymptotic first-order approximation for SE_G. By combining Stedinger's equations 6 and 20, R can be stated as

$$R = \sqrt{1 + GK + 0.5K^2(1 + 0.75G^2)} \tag{13}$$

where

    G   is the skew of the logs of the annual floods; and
    K   is a percentage point of the Pearson Type III distribution, which depends on G and the recurrence interval (U.S. Water Resources Council, 1981b, appendix 3).

Values computed from equation 13 are very close to those reported by Hardison (1971, table 2). By using equation 13, R can be computed for points he did not tabulate. Values of R are shown in figure 3 by a contour plot.

The derivation of equation 13 assumes that G, the skew coefficient, is known, but in application G can only be estimated. Because some uncertainty is ignored, equations 10 and 13 understate SE_G. The weighted discharge estimate (equation 1), therefore, leans

somewhat too heavily toward the gage estimate. Hardison (1971, table 2) suggested increasing R by 10 and 20 percent for 50- and 100-year floods, respectively, to compensate for estimated skew. Resolution of this problem must await further research.

Equations 10 and 13 were applied by selecting appropriate values for G and S_{LQ}. Because short records may yield poor estimates of both parameters, the use of some regional information was attractive. Analysis of flood records under Water Resources Council guidelines produces an estimate of G that combines regional and station information; that estimate was selected for G in equation 13. For S_{LQ}, the average of the gage estimate and its regional mean was selected as the best estimate to use in equation 10 because neither value was acceptable by itself. At the 101 stations in southern Arizona, gage estimates of S_{LQ} ranged from 0.20 to 0.90 and had a mean of 0.43 and median of 0.41 (table 8). The lowest values were for short-record stations where no major floods were measured. Using a small S_{LQ} in equation 10 gives a small value for SE_G and, hence, an unreasonably high weight to the gage estimate of flood magnitude. The hypothesis that a single value of S_{LQ} exists for all sites studied was tested by using a statistic that measures homogeneity of variance among samples of unequal sizes (Snedecor and Cochran, 1967, p. 296-298) and was rejected at the 99.9-percent confidence level. Application of equations 10 and 13 is illustrated in figure 3.

Although the weighted estimates that result from equations 1 and 12 are not greatly different, the method used in this study recognizes more of the variability in gage records and makes fewer statistical assumptions. It also assigns a specific value to the standard error of the weighted estimate (equation 2), which can be used to compensate for the remaining uncertainty (see Adjustment for Uncertainty). The method described in that section should not be used to construct confidence intervals for gage estimates (unweighted) of flood magnitudes; Stedinger (1983, equation 21) provides a good method for that purpose.

COMPARISON OF ALTERNATIVE METHODS

In presenting a new analysis of the magnitude and frequency of floods in Pima County, it is worthwhile to compare the results with results of other methods. These comparisons fulfill one objective of the study and may help the practicing hydrologist select the method best suited for a particular application. The many possible comparisons range from simple graphs to complex statistical procedures. Tests that are more statistically sophisticated than those selected for this study are available but were not justified because 75 percent of the sites had less than 20 years of systematic record. Six flood-frequency estimating methods were compared in this study. Each method was applied to a number of sites

and the different estimates were compared. Estimates by all six methods were available for only 18 of the 101 sites, so the comparisons used all available sites for each method. More extensive comparisons of methods for ungaged sites were reported by the U.S. Water Resources Council (1981a).

Methods recommended by the U.S. Water Resources Council (1981b) and Reich and Renard (1981) were compared for the analysis of flood-peak data for gaged sites. The U.S. Water Resources Council (WRC) method is based on fitting a Pearson Type III distribution to the logarithms of the annual floods. The distribution is defined by three parameters—mean, variance, and skew; its application was described in the section on frequency analysis. Reich and Renard (1981) recommended a graphical method that assumes the true distribution of floods at a site can be selected among several candidate distributions by plotting the observed floods on graph papers that linearize each distribution. The plot that best matches a straight line is assumed to permit reliable extrapolation to large recurrence intervals. In this study the extreme-value distribution was used for the annual floods or their logarithms, and the normal distribution was used for annual flood logarithms. Riggs (1968b, p. 4-8) noted that graph paper on which a theoretical distribution will plot as a straight line can be prepared for any two-parameter distribution and that the extreme-value distribution has a fixed skew of 1.139. Estimates by the graphical method were provided by Dr. Brian M. Reich (City of Tucson Engineering Division, written commun., 1982) for 46 sites in the study area using the same flood data analyzed by the WRC method.

Four methods for estimating flood magnitudes at ungaged sites were compared. Two methods were based on the primary and alternate estimating equations given in tables 1 and 2. Estimates for all 101 sites included the adjustments for urbanization and attenuation described under Applications. Weighted estimates adjusted on the basis of observed flood data at each site were not used. The third method used estimating equations presented by Roeske (1978) for rural sites. His approach was similar to that used in this study, but he analyzed data for all of Arizona and used a linear regression model. Estimates were computed using Roeske's equations for 91 sites.

The fourth method (Zeller, 1979) applies to rural or urban basins of less than 10 mi^2 in Pima County. Peak discharge is computed as the product of drainage area and runoff supply rate, which is a function of rainfall rate, soil type, vegetative cover, flow distance, slope, channel roughness, and urban development. The procedure has been used primarily to estimate the 100-year flood, with floods of other recurrence intervals estimated in a fixed ratio to the 100-year discharge. Suzanne J. Shields (Pima County Department of Transportation and Flood Control District, written commun., 1982) provided 100-year flood estimates by Zeller's method for 49 sites.

Inspection of scatter plots revealed some differences among the methods (fig. 12). Each plot shows the values of two estimates for each
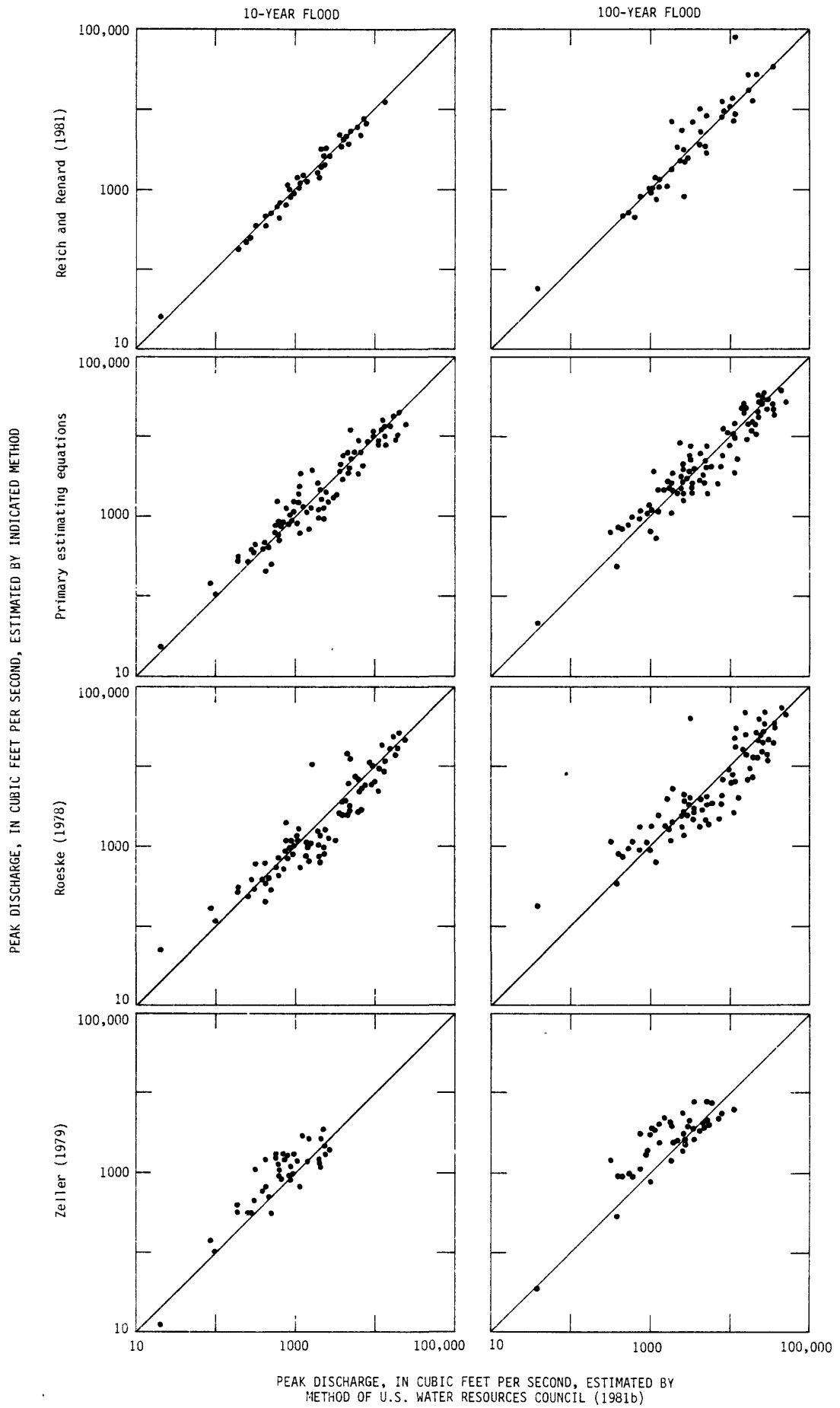
Figure 12.--Comparison of flood estimates by five methods.

site and an equal-value line. For gaged sites the WRC and graphical methods gave similar values for 10-year floods. For 100-year floods, the scatter was larger, but neither method gave distinctly larger estimates. Methods for ungaged sites generally showed larger differences from the WRC estimates; the scatter for the primary estimating equations was slightly smaller than for the other regional methods. For 10-year floods, the primary equations show a few more points above the line of equality than below it. Graphs for the alternate estimating equations were similar to those for the primary equations. Estimates from Roeske's equations showed a curved pattern of points at both recurrence intervals; most points were below the equal-value line in the middle of the range of discharges. Estimates from Zeller's method were greater than WRC estimates for most sites.

Comparison with Maximum Observed Floods

Flood estimates by the six methods were compared statistically with the maximum flood observed at each site. The result of the comparison is a measure of the statistical bias of each method at several recurrence intervals. This method of comparison does not depend on the true distribution of annual floods at each site.

Although a 100-year flood is a rare occurrence at any site, it may be exceeded in any year; the probability that it will be exceeded increases with the number of years considered. In this section, flood-frequency estimating methods are tested by comparing the number of sites at which 100-year flood estimates have been exceeded with the number that would be expected from the length of record. If the observed and expected numbers of exceedance sites are about the same, the estimating method probably is unbiased. If the numbers are extremely different, the method probably is statistically biased. This approach requires a fairly large number of sites, has no minimum record length for each site, and can be applied to floods of different recurrence intervals. By examining only the maximum flood at each site, some information about other large floods at the site was lost, but information was gained by using the historical record, which averaged 44 percent longer than the systematic record. In addition, the correlation is small between the single maximum flood and a flood magnitude estimated from all the observed floods. If several large floods were considered at each site, the correlation would increase and some results might be biased.

The derivation of the statistical test begins with the definition of the T-year flood, which is that particular discharge that has a probability of 1/T of being exceeded in any year. The probability or risk that the unknown, true T-year flood at a site would be exceeded at least once during n years is

$$p = 1 - (1 - 1/T)^n. \tag{14}$$

To count whether a site is an exceedance site, a variable is defined as

$$x = \begin{cases} 1, & \text{if the maximum observed flood was greater than} \\ & \text{or equal to the estimated T-year magnitude; or} \\ 0, & \text{otherwise.} \end{cases} \qquad (15)$$

Assuming the estimating method is statistically unbiased, x is a Bernoulli random variable whose distribution parameter is defined by equation 14. The mean or expected value of a Bernoulli random variable is its parameter, p, and the variance is

$$v = p(1-p). \qquad (16)$$

If k uncorrelated sites each had n years of record, then kp exceedance sites would be expected with variance kv. Real flood data, however, have different record lengths at each site.

Rather than discard much of the data in order to have a constant record length, the distribution of the number of exceedance sites was examined while allowing for unequal record lengths. For a single estimating method and recurrence interval, the observed number of exceedance sites is

$$X = \sum_{i=1}^{k} x_i, \qquad (17)$$

where i denotes each of the k sites. The mean, or expected value, of X is

$$P = \sum_{i=1}^{k} p_i, \qquad (18)$$

and the variance of X is

$$V = \sum_{i=1}^{k} v_i + 2 \sum_{i=1}^{k} \sum_{j=i+1}^{k} \text{Cov}(x_i, x_j) \qquad (19)$$

where $\text{Cov}(x_i, x_j)$ is the covariance of x at sites i and j. Because X is a sum of random variables, its distribution is approximately normal, provided that k is fairly large and the p_i are not too close to 0 or 1. Assuming the covariance is small, equation 19 can be approximated by

$$V = \sum_{i=1}^{k} v_i, \qquad (20)$$

and the number of exceedance sites can be standardized as

$$Z = (P - X)/\sqrt{V}. \qquad (21)$$

Z has a standard normal distribution if (1) the estimating method for the T-year flood is unbiased; (2) the sites and years for which floods were observed are a representative sample of all floods in the study area; (3) the number of sites, k, is large enough that the normal distribution is a good approximation of the distribution of X; and (4) the sum of covariance terms in equation 19 is a small portion of the total variance. The objective of the analysis is to test for bias (assumption 1). Assumption 2 seems reasonable because of the wide range of sites and years studied; estimates by the 6 methods were available for 46 to 101 sites. The assumption that time-sampling error is small is probably more accurate in this analysis than in one using only systematic record, which includes fewer years. The assumption of normality (3) appears reasonable for recurrence intervals of 25, 50, and 100 years because the range of individual site risk probabilities (equation 14) was greater than 0.5. For 2, 5, and 500 years, however, the range was less than 0.16, and the distribution of Z was too skewed to be considered normal. The covariance (4) is related to the correlation of large annual floods at different sites, which is generally small but positive in the study area. Equation 20 probably underestimates V, so using Z to compute probability levels could be misleading. As many as 16 sites were tentatively dropped from the analysis to reduce the covariance, but the results did not change substantially. Therefore, all available sites were used.

The expected number of exceedance sites (P) and the variance (V) depend only on the record lengths and are independent of the true distribution of annual floods at each site. The observed number of exceedance sites (X) depends on the estimating method and the available flood histories. The standardized difference (Z), therefore, tests the estimating methods independently of the distribution of annual floods. Z will generally be positive if a method overestimates flood magnitudes and negative if the method underestimates. Although probability levels associated with Z are uncertain, the relative bias of different methods can be evaluated. The bias test is sensitive to the number of sites that have had extreme floods.

This approach was applied to each estimating method and recurrence interval by computing (a) the expected number of exceedance sites using equations 14 and 18; (b) the variance using equations 16 and 20; (c) the observed number of exceedance sites using equations 15 and 17; and (d) the standardized difference using equation 21. Values of X,

P, $\sqrt{V}$, and Z are shown in table 6 for the 25-, 50-, and 100-year recurrence intervals for each estimating method.

The distribution-free test against maximum observed floods (table 6) produced two unexpected results. First, the WRC method tended to overestimate flood magnitudes at each recurrence interval even though an expected-probability adjustment was not used. Theory suggests that the WRC method as used here should underestimate flood magnitudes, because the average or expected probability that a flood would exceed a given discharge is greater than the probability used to compute the discharge estimate (U.S. Water Resources Council, 1981b, appendix 11). Second, the absolute value of the bias statistic Z was smaller for both sets of equations developed in this study than for either method for gaged sites. Thus, for the record lengths available in this study, the best ungaged-site methods have smaller bias than the methods based on gage data.

The results also show that both methods for gaged sites overestimated 25-, 50-, and 100-year floods (table 6). The positive bias was larger by the WRC method at 25 years and by the graphical method at 50 and 100 years. Among the methods for ungaged sites, estimates by Zeller's method were largest and those using Roeske's equations were smallest for 25-, 50-, and 100-year floods in comparison to the maximum observed floods (table 6). A trend from overestimating 25-year floods toward underestimating 100-year floods was noted for the primary, alternate, and Roeske's equations. However, standardized differences greater than two in absolute value were found only for Roeske's equations at the 50- and 100-year recurrence intervals and for Zeller's method at 25 and 100 years. Zeller's method showed a positive bias at all recurrence intervals for the sites that were available for this study.

Comparison by Differences Between Methods

Statistical comparisons can also be made between pairs of flood-frequency estimating methods. Differences from a consistent base can be used to measure the variability of each method and to determine whether one method generally yields larger values than another. Estimated flood magnitudes by several methods were available for each site and recurrence interval. Using the estimates, logarithmic differences from base methods were computed:

$$D_T = LQ_T - B_T,$$

where

LQ_T is the logarithm of the T-year flood estimated by one method for a single site, and

B_T is the logarithm of the base estimate for the same site.

Table 6.--Comparisons of maximum observed floods to discharges
estimated by selected methods

Estimating method	Total sites	Recurrence interval (years)	Number of exceedance sites			
			Observed (X)	Expected		Standardized difference (Z)
				Mean (P)	Standard deviation ($\sqrt{V}$)	
Methods for Gaged Sites						
U.S. Water Resources Council (1981b)	101	25	48	60.9	4.53	2.84
		50	35	40.1	4.52	1.13
		100	19	23.9	4.01	1.21
Reich and Renard (1981)	46	25	21	26.5	3.15	1.75
		50	7	16.9	3.08	3.21
		100	4	9.8	2.65	2.18
Methods for Ungaged Sites						
Primary estimating equations[1]	101	25	57	60.9	4.53	.85
		50	37	40.1	4.52	.69
		100	27	23.9	4.01	-.78
Alternate estimating equations[1]	101	25	56	60.9	4.53	1.07
		50	40	40.1	4.52	.02
		100	28	23.9	4.01	-1.03
Roeske (1978)	91	25	54	56.1	4.27	.49
		50	47	37.3	4.30	-2.26
		100	34	22.3	3.85	-3.03
Zeller (1979)	49	25	16	25.3	3.42	2.73
		50	11	15.1	3.17	1.29
		100	2	8.3	2.59	2.43

[1]Adjusted for attenuation and urbanization. Not adjusted for gage data or uncertainty.

The standard deviation and mean of the differences were computed for each estimating method and recurrence interval (table 7). The standard deviation of the differences measures the scatter of estimates by the method after adjusting for the average difference from the base method. This standard deviation is comparable to the standard error of prediction of the regression equations. The mean of the differences is positive if estimates by the method are generally larger than those by the base method. The mean difference, therefore, is similar to the previous test of bias but also depends on the method selected as base. The mean difference summarizes the difference of two estimates at all sites, but it is affected by the particular base estimates used. In contrast, the preceding test using maximum observed floods is sensitive to the number of sites that have had extreme floods and is independent of the distribution of annual floods.

For each group of methods in table 7, a method was selected from the opposite group as the base for comparisons. A base method was used as a standard to compare methods within a group without assuming one of the group to be superior. For methods applicable to gaged sites, the base estimate was computed from the primary estimating equations (table 1) with adjustments but without weighting. The base estimate for methods applicable to ungaged sites was computed from the gage record using the methods described in the section on frequency analysis. The base methods were available for all sites and showed the smallest overall bias when compared to maximum observed floods.

At three sites, estimates by the WRC method were known to be low compared to regional trends (see the section entitled "Grouping of Basins for Analyses"). The three sites were included in the tests of methods for gaged sites because they represent part of the possible variability of those estimates. They were excluded, however, from the tests of methods for ungaged sites because they would incorrectly indicate overestimation by the regional methods.

Because the primary estimating equations are based on WRC estimates at 84 sites, table 7 is somewhat redundant by including two comparisons of these methods. The estimates by the primary equations in table 7, however, apply to more sites and include adjustments for attenuation and urbanization. As noted above, the two comparisons used slightly different groups of sites. Thus, the comparisons are not duplicates nor simply a restatement of the regression residuals. The results shown in table 7 are considered to be a reasonable approximation of the variance and statistical bias of estimates by each method.

At the 10-year recurrence interval, the standard deviation—square root of variance—of the differences was similar for the WRC and graphical methods (table 7). The standard deviation for WRC estimates, however, increased 4 percentage points from the 10- to 100-year recurrence interval, while the standard deviation for graphical estimates increased 32 points. The mean differences indicate that, on the average, WRC estimates were slightly less and graphical estimates slightly more than the base estimates computed from the adjusted primary equations.

Table 7.--Comparisons between discharges estimated by selected methods

Estimating method	Total sites	Recurrence interval (years)	Difference from base estimate			
			Mean		Standard deviation	
			Log units	Percent	Log units	Percent
Methods for Gaged Sites[1]						
U.S. Water Resources Council (1981b)	101	10 100	-0.022 -.021	-5 -5	0.205 .221	49 53
Reich and Renard (1981)	46	10 100	.022 .031	+5 +7	.194 .311	46 78
Methods for Ungaged Sites[2]						
Primary estimating equations[3]	98	10 100	.003 .005	+1 +1	.176 .199	42 47
Alternate estimating equations[3]	98	10 100	.009 .012	+2 +3	.199 .226	47 54
Roeske (1978)	88	10 100	-.052 -.031	-11 -7	.228 .280	55 69
Zeller (1979)	46	10 100	.142 .145	+39 +40	.209 .230	50 55

[1]Base estimate is primary estimating equations with adjustments. See footnote 3.

[2]Base estimate is U.S. Water Resources Council (1981b). Three sites were excluded that are regional low outliers by that method.

[3]Adjusted for attenuation and urbanization. Not adjusted for gage data or uncertainty.

For recurrence intervals of 25 and 50 years, the results were intermediate to those shown in table 7. Results for this sample of sites indicate that the uncertainty of flood-frequency estimates, as measured by standard deviation, increases more rapidly with recurrence interval for the graphical method than for the WRC method, but that neither method has a large statistical bias.

Among the methods for ungaged sites, the standard deviation of flood estimates in comparison to a base method was smallest for the primary estimating equations and successively larger for the alternate equations, Zeller's method, and Roeske's equations. For each method, the standard deviations and means for 25- and 50-year floods were intermediate to those shown in table 7. For 100-year floods, the standard deviation was about 47 percent for the primary equations and about 69 percent for Roeske's method. The standard deviations for Zeller's method and the alternate equations were nearly equal.

The mean differences show that, on the average, estimates from the primary or alternate equations were within 3 percent of WRC estimates. Estimates by Roeske's equations, however, were about 10 percent smaller and those by Zeller's method were about 40 percent larger than the base estimates. Assuming a log-normal error distribution (see Adjustment for Uncertainty), the probability is about 63 percent that a 100-year flood estimate by Zeller's method would be larger than a WRC estimate on the basis of the values provided by the Pima County Department of Transportation and Flood Control District. Similarly, the probability is about 55 percent that a 10-year flood estimate by Roeske's equations would be smaller than a WRC estimate.

SUMMARY AND CONCLUSIONS

New techniques were developed to estimate the magnitude and frequency characteristics of floods in Pima County, Arizona, which includes 9,240 mi^2 in a semiarid region of large relief. Flood magnitudes at rural or urban sites with drainage areas between 0.013 and 4,471 mi^2 may be computed for recurrence intervals of 2, 5, 10, 25, 50, 100, and 500 years. The standard error of regression for the primary estimating equations developed in this study is 42 to 49 percent for 5- to 100-year floods, which is a substantial reduction from the standard errors of methods published previously. A technique was presented to compensate for the uncertainty of the estimates.

The new techniques were based on more than 2,000 station years of systematic flood data at 101 sites in and near Pima County. Flood-frequency estimates were computed from the data for each gaging station by following the guidelines of the U.S. Water Resources Council (1981b). Relationships applicable to ungaged sites were developed by multiple-regression analysis of the gage estimates with measureable characteristics of the tributary drainage basins. Weighted estimates of

flood magnitudes at gaging stations were computed by using the standard errors of gage estimates and regression estimates.

The analyses for rural basins showed that a second-order regression among the logarithms of the variables gave better results than a linear or first-order regression. The predictor variables in the primary estimating equations were drainage area, channel slope, and basin shape; they were used in two first-order and three second-order terms. Alternative quadratic estimating equations included only drainage area as a predictor variable. The nonlinear terms in both models apparently were needed because of the large range of many basin characteristics within the study area. The magnitudes and signs of the coefficients for non-linear terms were consistent with a qualitative understanding of the hydrologic processes involved. Analyses of urban basins showed that estimating equations developed in a national study (Sauer and others, 1983) were applicable in Pima County. A tentative adjustment factor was presented to reduce peak-discharge estimates for basins where channel conditions cause large attenuation of flood peaks.

Six methods of estimating flood magnitudes in Pima County were compared. A technique was developed to test flood estimates against the maximum observed floods independently of the distribution of annual floods. The test showed that both methods applicable to gaged sites (U.S. Water Resources Council, 1981b; Reich and Renard, 1981) over-estimated 25-, 50-, and 100-year floods for the period of record in the study area, even though expected-probability theory indicates the WRC method should have underestimated. Comparisons with a base method showed that the variability of the estimates was somewhat larger for Reich and Renard's method. For ungaged sites, the primary and alternate estimating equations were compared with methods presented by Roeske (1978) and Zeller (1979). When tested against maximum floods in the study area, Roeske's method tended to underestimate and Zeller's method tended to overestimate to a greater degree than the other two methods. The primary and alternate estimating equations showed smaller absolute bias than either method applicable to gaged sites. Estimates by Roeske's method averaged about 10 percent less than estimates from analysis of gaging-station records and showed the greatest variability. Estimates by Zeller's method averaged 40 percent larger than gage estimates and were about equal in variability to the alternate estimating equations. On the average, the alternate and primary estimates were equal to gage estimates; the primary equations showed the smallest variability.

The results of this study suggest potentially useful work in two areas of flood hydrology. The logarithmically nonlinear relationships between flood magnitudes and basin characteristics were represented in this study by second-order regression models. Other nonlinear models might have been used, but they were not investigated because of limited time. A comparative study might show whether the second-order model or some other nonlinear model produces lower standard errors.

The second area is the effect of channel characteristics on flood magnitudes. We know that some channels convey flood hydrographs with

little change in peak discharge, while others cause substantial attenuation. A consistent channel effect on individual floods will be reflected in a change in flood magnitudes at various recurrence intervals. Zeller's (1979) method includes a factor for channel conditions, and the adjustment for attenuation in this report recognizes an extreme channel effect. Definition of variables in applying channel information to flood prediction, however, is a problem. The characteristics of a channel reach offer more information than a single cross section; a useful additional predictor of flood magnitudes must summarize reach characteristics, perhaps as an index variable such as the urban basin development factor. Identification of the appropriate variables might be helped by simulation of flood movement through channels and field investigations of channel characteristics in basins whose observed flood experience differs from the prediction equations.

<h1 style="text-align:center">REFERENCES CITED</h1>

Aldridge, B. N., and Eychaner, J. H., 1984, Floods of October 1977 in southern Arizona and March 1978 in central Arizona: U.S. Geological Survey Water-Supply Paper 2223, 143 p.

Arizona Water Commission, 1973, Flood plain delineation criteria and procedures: Arizona Water Commission Report 4, 12 p.

Condes de la Torre, Alberto, 1967, Streamflow and flood characteristics, Pima County, Arizona (a progress report): U.S. Geological Survey open-file report, 24 p.

Crippen, J. R., and Bue, C. D., 1977, Maximum floodflows in the conterminous United States: U.S. Geological Survey Water-Supply Paper 1887, 52 p.

Draper, N. R., and Smith, Harry, 1981, Applied regression analysis, second edition: New York, John Wiley, 709 p.

Federal Emergency Management Agency, 1982a, Flood insurance study, City of Tucson, Arizona: Federal Emergency Management Agency report, Community number 040076, 45 p.

_____1982b, Flood insurance study, Pima County, Arizona: Federal Emergency Management Agency report, Community number 040073, 94 p.

Greis, N. P., 1983, Flood frequency analysis, a review of 1979-1982: Reviews of Geophysics and Space Physics, v. 21, no. 3, p. 699-706.

Hardison, C. H., 1969, Accuracy of streamflow characteristics, in Geological Survey Research 1969: U.S. Geological Survey Professional Paper 650-D, p. D210-D214.

56

_____1971, Prediction error of regression estimates of streamflow characteristics at ungaged sites, in Geological Survey Research 1971: U.S. Geological Survey Professional Paper 750-C, p. C228-C236.

Patterson, J. L., and Somers, W. P., 1966, Magnitude and frequency of floods in the United States—Part 9, Colorado River basin: U.S. Geological Survey Water-Supply Paper 1683, 475 p.

Reich, B. M., 1976, Magnitude and frequency of floods: CRC Critical Reviews in Environmental Control, v. 6, no. 4, p. 297-346.

Reich, B. M., and Renard, K. G., 1981, Application of advances in flood frequency analysis: Water Resources Bulletin, v. 17, no. 1, p. 67-74.

Riggs, H. C., 1968a, Some statistical tools in hydrology: U.S. Geological Survey Techniques of Water-Resources Investigations, Book 4, Chapter A1, 39 p.

_____1968b, Frequency curves: U.S. Geological Survey Techniques of Water-Resources Investigations, Book 4, Chapter A2, 15 p.

_____1973, Regional analyses of streamflow characteristics: U.S. Geological Survey Techniques of Water-Resources Investigations, Book 4, Chapter B3, 15 p.

Roeske, R. H., 1978, Methods for estimating the magnitude and frequency of floods in Arizona: Arizona Department of Transportation Report ADOT-RS-15(121), 82 p. [Also distributed as National Technical Information Service (NTIS) report PB 289 424/U.C.]

Sauer, V. B., Thomas, W. O., Stricker, V. A., and Wilson, K. V., 1983, Flood characteristics of urban watersheds in the United States: U.S. Geological Survey Water-Supply Paper 2207, 63 p.

SAS Institute, 1982a, SAS User's guide: Basics, 1982 edition: Cary, N.C., SAS Institute, Inc., 923 p.

_____1982b, SAS User's guide: Statistics, 1982 edition: Cary, N.C., SAS Institute, Inc., 584 p.

Snedecor, G. W., and Cochran, W. G., 1967, Statistical methods, sixth edition: Ames, Iowa State University Press, 593 p.

Snee, R. D., 1977, Validation of regression models—methods and examples: Technometrics, v. 19, no. 4, p. 415-428.

Stedinger, J. R., 1983, Confidence intervals for design events: Journal of Hydraulic Engineering, v. 109, no. 1, p. 13-27.

Thomas, D. M., and Benson, M. A., 1970, Generalization of streamflow characteristics from drainage-basin characteristics: U.S. Geological Survey Water-Supply Paper 1975, 55 p.

U.S. Census Bureau, 1981, Arizona final population and housing unit counts: 1980 Census of Population and Housing report PHC80-V-4, 8 p.

U.S. Soil Conservation Service, 1972, Hydrology: U.S. Soil Conservation Service, National Engineering Handbook, section 4, 547 p.

U.S. Water Resources Council, 1981a, Estimating peak flow frequencies for natural ungaged watersheds—A proposed nationwide test: Water Resources Council, Hydrology Committee report, 346 p.

_____1981b, Guidelines for determining flood flow frequency (Errata corrections included through March 1982): Water Resources Council Bulletin 17B, 183 p.

Zeller, M. E., 1979, Hydrology manual for engineering design and flood plain management within Pima County, Arizona: Tucson, Arizona, Pima County Department of Transportation and Flood Control District, 147 p.

Table 8.--Flood-frequency data and basin characteristics for gaging stations

Basin characteristics: AREA, drainage area in square miles; SLOPE, channel slope in percent; SHAPE, basin shape factor (length squared divided by area); ELEV, mean basin elevation, in feet above NGVD of 1929; BDF, basin development factor.
Group: A, site subject to attenuation of peaks; C, site correlated with nearby sites; L, site is a low outlier; R, site used to develop rural estimating equations; U, site used in urban analyses.
Annual floods: YRS, number of years of systematic record; STD, standard deviation of logarithms of annual floods; SKEW, coefficient of skew of logarithms of annual floods. STD and SKEW include all adjustments specified in U.S. Water Resources Council (1981b).
Flood estimates: Peak discharge in cubic feet per second. First line is weighted average of gage and regional estimates using equation 1; second line is estimate from gage record only.

STATION NUMBER AND NAME					GROUP	ANNUAL FLOODS			FLOOD ESTIMATES AT GIVEN RECURRENCE INTERVAL, IN YEARS: LINE 1, WEIGHTED; LINE 2, GAGE RECORD						
BASIN CHARACTERISTICS															
AREA	SLOPE	SHAPE	ELEV	BDF		YRS	STD	SKEW	2	5	10	25	50	100	500
09470500 SAN PEDRO RIVER AT PALOMINAS															
741	0.45	1.69	4950	0	R	41	0.24	-0.12	6260	9750	12200	15600	18400	21400	28900
									6460	10300	13000	16600	19400	22300	29500
09470900 SAN PEDRO RIVER TRIBUTARY NEAR BISBEE															
5.25	2.44	2.96	4770	0	R	16	.45	-.26	387	897	1390	2170	2870	3680	5950
									406	929	1400	2120	2750	3450	5370
09471000 SAN PEDRO RIVER AT CHARLESTON															
1219	.35	2.59	4840	0	R	64	.28	.05	7260	12300	16100	21700	26600	31900	46400
									7410	12900	17300	23700	29200	35100	51200
09471080 WALNUT GULCH 63.010 NEAR TOMBSTONE[1]															
6.42	1.51	20.6	4970	0	P	15	.55	-.24	297	795	1260	2050	2790	3660	6250
									296	821	1360	2290	3160	4200	7300
09471087 WALNUT GULCH 63.111 NEAR TOMBSTONE[1]															
.22	2.75	2.77	5020	0	P	20	.52	-.28	90	213	325	495	652	835	1370
									99	258	413	666	896	1160	1910
09471090 WALNUT GULCH 63.009 NEAR TOMBSTONE[1]															
9.11	1.20	9.31	4840	0	R	15	.53	-.23	538	1320	2040	3210	4310	5630	9600
									588	1570	2560	4220	5770	7600	13000
09471110 WALNUT GULCH 63.015 NEAR TOMBSTONE[1]															
9.24	1.07	1.99	4680	0	C	27	.49	-.21	496	1210	1890	2980	3980	5170	8660
									513	1290	2030	3260	4380	5690	9470
09471120 WALNUT GULCH 63.011 NEAR TOMBSTONE[1]															
3.18	1.89	4.86	4880	C	P	19	.42	-0.23	515	1050	1520	2240	2900	3670	5900
									588	1300	1930	2890	3720	4640	7170
09471130 WALNUT GULCH 63.008 NEAR TOMBSTONE[1]															
5.98	1.55	10.7	4760	C	C	19	.39	-.18	583	1160	1660	2440	3150	3970	6280
									642	1340	1930	2830	3600	4450	6760
09471140 WALNUT GULCH 63.006 NEAR TOMBSTONE[1]															
36.7	1.53	4.42	4790	0	R	20	.45	-.28	1210	2750	4130	6300	8270	10500	16900
									1290	2990	4500	6840	8870	11100	17300
09471170 WALNUT GULCH 63.004 NEAR TOMBSTONE[1]															
.88	1.84	6.55	4550	C	R	24	.65	-.19	77	250	432	747	1050	1420	2610
									72	244	448	842	1250	1780	3550
09471180 WALNUT GULCH 63.003 NEAR TOMBSTONE[1]															
3.47	1.29	5.84	4550	0	C	28	.61	-.15	168	522	904	1570	2210	2990	5500
									160	505	902	1650	2410	3380	6600
09471185 WALNUT GULCH 63.103 NEAR TOMBSTONE[1]															
.013	4.51	6.03	4500	0	P	19	.28	-.16	9	15	21	28	34	40	54
									9	15	20	27	32	37	50
09471190 WALNUT GULCH 63.002 NEAR TOMBSTONE[1]															
43.9	.95	4.86	4700	C	C	28	.51	-.30	1370	3300	5070	7830	10400	13300	21900
									1450	3710	5880	9370	12500	16100	26100
09471195 WALNUT GULCH 63.007 NEAR TOMBSTONE[1]															
5.22	1.42	2.43	4480	0	P	16	.75	-.21	276	939	1650	2830	4010	5480	10400
									270	1100	2220	4580	7230	10800	23700
09471200 WALNUT GULCH 63.001 NEAR TOMBSTONE[1]															
57.7	1.04	6.66	4700	0	P	25	.37	-.20	1580	3150	4510	6600	8420	10500	16000
									1660	3320	4700	6740	8440	10300	15200
09471550 SAN PEDRO RIVER NEAR TOMBSTONE															
1740	0.34	2.61	4820	0	R	13	.30	-.22	7280	11900	15200	19800	23800	28100	39300
									7870	13900	18400	24600	29500	34600	47400

Table 8.--Flood-frequency data and basin characteristics for gaging stations—Continued

STATION NUMBER AND NAME					GROUP	ANNUAL FLOODS			FLOOD ESTIMATES AT GIVEN RECURRENCE INTERVAL, IN YEARS: LINE 1, WEIGHTED; LINE 2, GAGE RECORD						
BASIN CHARACTERISTICS															
AREA	SLOPE	SHAPE	ELEV	BDF		YRS	STD	SKEW	2	5	10	25	50	100	500

09471600 CANARY WASH NEAR BENSON															
0.79	8.94	8.56	5240	0	L	13	0.90	-0.19	11	65	132	258	373	510	933
									4	20	46	111	192	313	817
09471700 FENNER WASH NEAR BENSON															
2.71	1.78	10.8	4180	0	R	14	.49	-.17	167	431	686	1120	1510	1970	3310
									161	405	646	1050	1420	1860	3160
09472000 SAN PEDRO RIVER NEAR REDINGTON															
2939	.29	5.25	4660	0	R	49	.35	-.17	8480	15700	21600	29500	36300	43600	63200
									8610	16500	23000	32300	40100	48500	70600
09472100 PECK CANYON TRIBUTARY NEAR REDINGTON															
8.02	2.77	10.8	3680	0	R	14	.61	-.11	162	559	986	1750	2480	3350	6050
									129	414	752	1400	2090	2980	6050
09472400 MAMMOTH WASH NEAR MAMMOTH															
2.40	3.43	7.35	3700	0	R	14	.83	-.63	177	594	980	1600	2190	2890	4960
									178	755	1440	2670	3810	5130	8710
09473000 ARAVAIPA CREEK NEAR MAMMOTH															
541	.87	5.67	4530	0	R	30	.33	-.21	4390	8160	11200	15700	19500	23600	34100
									4490	8330	11300	15500	19000	22600	31800
09473200 GREEN LANTERN WASH NEAR WINKELMAN															
3.63	3.58	6.35	2590	0	R	13	.52	-.17	430	977	1470	2270	3040	3970	6830
									517	1380	2260	3780	5220	6950	12200
09473500 SAN PEDRO RIVER AT WINKELMAN															
4471	.34	6.96	4520	0	R	18	.33	.02	7450	13900	19400	27000	33400	40300	59300
									7330	13800	19300	27500	34700	42700	65100
09473600 TAM O'SHANTER WASH NR HAYDEN															
4.37	6.14	8.49	3050	0	R	14	.23	-.09	353	569	748	1030	1270	1540	2190
									376	585	733	929	1080	1240	1620
09478200 DURHAM WASH NEAR FLORENCE															
15.6	2.71	7.76	3670	0	R	18	.58	-.21	465	1320	2180	3630	5010	6660	11700
									459	1350	2320	4040	5730	7780	14200
09478500 QUEEN CREEK AT WHITLOW DAMSITE NEAR SUPERIOR															
144	2.60	2.61	3180	0	R	16	.44	-.06	2850	6240	9510	14700	19500	25200	42000
									3100	7050	10700	16700	22200	28700	48000
09478600 QUEEN CREEK TRIBUTARY NO.3 AT WHITLOW DAM															
.37	3.54	3.51	2320	0	R	14	.33	-.13	72	145	214	322	413	512	780
									72	136	187	261	322	388	563
09479200 QUEEN CREEK TRIBUTARY AT APACHE JUNCTION															
.51	1.44	3.84	1760	0	R	19	.48	-.19	53	138	223	362	487	628	1030
									49	120	188	300	403	522	870
09480000 SANTA CRUZ RIVER NEAR LOCHIEL															
82.2	.80	1.75	5150	0	R	32	.37	-.18	1440	2920	4200	6170	7870	9770	14900
									1470	2960	4210	6060	7630	9350	13900
09480500 SANTA CRUZ RIVER NEAR NOGALES															
533	.49	4.92	4850	0	R	50	.31	-.04	4270	7620	10400	14400	17700	21500	31600
									4340	7820	10600	14700	18000	21700	31600
09481500 SONOITA CREEK NEAR PATAGONIA															
209	1.45	2.25	4800	0	R	43	.36	-.26	2680	5370	7630	11000	13900	16900	24800
									2720	5370	7500	10600	13100	15700	22600
09481700 CALABASAS CANYON NEAR NOGALES															
10.3	1.43	12.8	4360	0	R	13	.42	-.17	318	775	1220	1950	2610	3350	5460
									303	675	1010	1530	1990	2510	3980
09481750 SOPORI WASH AT AMADO															
176	.50	1.47	3840	0	R	19	.55	-.24	1470	3760	5770	8970	11900	15400	25800
									1470	4100	6820	11500	15900	21200	37000
09481800 DEMETRIE WASH TRIBUTARY NEAR CONTINENTAL															
.15	1.89	4.27	3620	0	R	14	.42	-.17	28	67	104	164	215	270	426
									26	57	85	129	168	211	334

Table 8.--Flood-frequency data and basin characteristics for gaging stations—Continued

STATION NUMBER AND NAME BASIN CHARACTERISTICS					GROUP	ANNUAL FLOODS			FLOOD ESTIMATES AT GIVEN RECURRENCE INTERVAL, IN YEARS: LINE 1, WEIGHTED; LINE 2, GAGE RECORD						
AREA	SLOPE	SHAPE	ELEV	BDF		YRS	STD	SKEW	2	5	10	25	50	100	500
09482000 SANTA CRUZ RIVER AT CONTINENTAL															
1662	0.40	5.92	4350	0	R	37	0.30	0.07	4000	7930	14700	19000	22500	26300	36000
									3940	7700	14900	19000	22200	25600	34100
09482200 FLATO WASH NEAR SAHUARITA															
6.65	0.90	16.6	2820	0	R	17	.33	.00	532	1000	1440	2120	2740	3450	5490
									572	1090	1530	2200	2770	3420	5230
09482330 PUMPING WASH NEAR VAIL															
.81	1.09	6.58	3010	0	R	16	.36	-.22	100	208	305	456	585	726	1100
									100	198	277	392	487	589	B55
09482350 SOUTH FORK AIRPORT WASH NEAP TUCSON															
4.38	.69	2.19	2740	0	R	15	.50	-.20	203	522	827	1320	1780	2300	3860
									198	511	822	1340	1830	2400	40B0
09482370 NORTH FORK AIRPORT WASH NEAR TUCSON															
5.28	.77	4.01	2780	0	R	16	.55	-.22	191	545	904	1490	2040	2680	4600
									176	495	826	1400	1950	2610	4610
09482400 AIRPORT WASH AT TUCSON															
16.2	.70	7.20	2700	1	R	16	.35	.34	430	977	1540	2420	3340	4410	7530
									367	740	1100	1700	22B0	3000	5330
09482410 RODEO WASH AT TUCSON															
7.24	.64	8.66	2560	2	R	12	.28	-.12	298	579	832	1210	1560	1950	2970
									256	440	579	772	927	1090	1510
09482420 JULIAN WASH AT TUCSON															
26.5	0.74	11.3	2900	2	R	12	.35	-.20	550	1170	1750	2620	3460	4400	6860
									427	816	1130	1580	1940	2340	3360
09482450 WEST BRANCH SANTA CRUZ RIVER AT TUCSON															
23.6	1.53	9.92	2800	0	A	16	.42	-.16	238	573	899	1440	1930	2470	4000
									226	498	742	1120	1460	1840	2900
0948248O BIG WASH AT TUCSON															
2.94	1.80	8.17	2850	0	R	16	.83	-.24	83	370	703	1300	1880	2600	4940
									58	275	593	1310	2140	3290	7640
09482500 SANTA CRUZ RIVER AT TUCSON															
2222	.38	7.34	4050	0	R	65	.29	-.17	5310	9210	12200	16400	19800	23200	31900
									5290	9080	11900	15800	18900	22100	30200
09482950 RAILROAD WASH AT TUCSON [2]															
2.30	.73	3.17	2488	5	U	11	.33	.22	339	655	914	1310	1680	2120	3300
									372	716	1020	1520	1970	2510	4130
09483000 TUCSON ARROYO AT VINE AVENUE, AT TUCSON															
8.20	.70	3.69	2510	3	U	26	.31	.20	789	1440	1990	2820	3590	4490	7030
									842	1540	2150	3090	3920	4890	7710
09483010 HIGH SCHOOL WASH AT TUCSON [2]															
.95	1.11	2.63	2462	3	U	14	.29	.02	230	417	565	784	984	1210	1800
									260	458	616	845	1040	1250	182C
09483025 SILVERCROFT WASH AT TUCSON															
2.74	.84	3.95	2540	0	R	14	.68	-.19	129	431	747	1270	1780	2390	4330
									113	408	776	1510	2290	3310	6830
09483030 ANKLAM WASH AT TUCSON															
2.11	5.64	2.09	2700	0	R	17	.29	.00	71	438	1610	2250	2820	3480	5390
									63	410	2020	2780	3400	4050	5900
09483040 WEST SPEEDWAY WASH NEAR TUCSON															
.46	6.89	2.63	2750	0	R	17	.44	-.25	87	201	315	485	635	803	1270
									89	201	300	452	585	732	1130
09483042 CEMETERY WASH AT TUCSON															
1.17	.76	4.13	2367	7	U	16	.23	-.07	266	437	568	754	916	1090	1540
									281	439	553	705	823	946	1250
09483045 FLOWING WELLS WASH AT TUCSON															
3.53	.61	5.09	2376	9	U	11	.30	-.03	488	913	1260	1760	2220	2750	4110
									498	888	1200	1650	2030	2440	3530
09483100 TANQUE VERDE CREEK NEAR TUCSON															
43.0	2.95	2.81	4780	0	R	22	.38	-.20	1200	2540	3810	5790	7520	9450	14700
									1230	2520	3610	5230	6600	81O0	12100

Table 8.--Flood-frequency data and basin characteristics for gaging stations—Continued

STATION NUMBER AND NAME BASIN CHARACTERISTICS					GROUP	ANNUAL FLOODS			FLOOD ESTIMATES AT GIVEN RECURRENCE INTERVAL, IN YEARS: LINE 1, WEIGHTED; LINE 2, GAGE RECORD						
AREA	SLOPE	SHAPE	ELEV	BDF		YRS	STD	SKEW	2	5	10	25	50	100	500
09483200 AGUA CALIENTE WASH TRIBUTARY NEAR TUCSON															
2.04	6.97	10.4	3300	0	R	16	0.39	-0.15	103	233	348	540	708	896	1410
									99	210	306	453	581	725	1120
09483250 ROB WASH AT TUCSON															
2.08	1.05	10.5	2713	4	U	11	.39	-.03	264	567	826	1220	1590	2020	3180
									270	571	844	1280	1670	2120	3430
09484000 SABINO CREEK NEAR TUCSON															
35.5	9.12	4.76	6300	0	R	50	.44	-.18	1060	2400	3650	5600	7380	9430	15300
									1090	2510	3830	5920	7800	9950	16100
09484200 BEAR CREEK NEAR TUCSON															
16.3	9.52	7.42	5860	0.	R	16	.43	-.27	331	780	1200	1860	2450	3100	4860
									320	718	1070	1600	2070	2580	3950
09484500 TANQUE VERDE CREEK AT TUCSON															
219	2.06	1.65	4340	0	R	22	.54	-.19	2000	5600	9390	15600	21300	28100	48400
									1950	5410	9020	15300	21300	28600	50800
09484560 CIENEGA CREEK NEAR PANTANO															
289	1.13	3.37	4890	0	R	14	.37	-.07	2060	4570	6940	10700	14100	17800	28200
									2000	4110	5940	8770	11300	14100	22000
09484570 MESCAL ARROYO NEAR PANTANO															
38.4	1.57	3.69	4260	0	R	17	.50	-.02	816	2150	3530	5840	8040	10700	18800
									808	2110	3480	5920	8330	11300	21000
09484580 BARREL CANYON NEAR SONOITA															
14.1	3.60	2.47	5000	0	R	15	.47	-.22	499	1280	2120	3450	4670	6060	10100
									493	1190	1840	2890	3830	4910	7970
09484590 DAVIDSON CANYON WASH NR VAIL															
50.5	1.63	3.66	4340	0	R	14	.53	-.27	1370	3390	5310	8360	11200	14600	24600
									1490	3970	6420	10500	14300	18600	31300
09484600 PANTANO WASH NEAR VAIL															
457	.88	4.14	4500	0	R	23	.46	-.20	3410	7840	11800	18000	23600	30000	48500
									3470	8250	12700	19900	26400	33800	55100
09485000 RINCON CREEK NEAR TUCSON															
44.8	11.6	1.81	4850	0	R	29	.56	-.13	938	2740	4910	8390	11800	15800	28600
									938	2740	4720	8340	12000	16500	31100
09485500 PANTANO WASH AT TUCSON															
602	1.07	4.92	4560	0	R	16	.55	-.23	2240	6310	10300	16800	22600	29200	48500
									2000	5570	9260	15600	21700	28900	50600
09485550 ARCADIA WASH AT TUCSON [2]															
2.72	.77	6.39	2559	8	U	14	.29	-.08	391	720	987	1380	1730	2130	3150
									396	690	918	1240	1500	1780	2510
09485900 PIMA WASH NEAR TUCSON															
4.93	10.1	6.14	4430	0	L	18	.48	-.27	77	220	370	623	841	1080	1730
									64	155	239	373	491	625	998
09485950 GERONIMO WASH NEAR TUCSON															
2.15	13.2	6.03	3600	0	R	18	.43	-.17	122	277	424	653	857	1090	1730
									121	271	407	619	807	1020	1620
09486000 RILLITO CREEK NEAR TUCSON															
918	.76	5.73	4400	0	R	67	.33	-.24	4940	9230	12700	17500	21500	25600	36300
									4970	9250	12600	17200	21000	24900	34900
09486300 CANADA DEL ORO NEAR TUCSON															
250	2.05	3.23	4000	0	R	16	.46	-.19	2580	6110	9480	14900	19700	25200	41000
									2620	6170	9470	14800	19500	25000	40500
09486500 SANTA CRUZ RIVER AT CORTARO															
3503	.38	5.59	4000	0	R	38	.23	-.20	8560	13200	16500	20900	24200	27600	35700
									8610	13200	16300	20300	23300	26300	33400
09486700 CHILTEPINES WASH NEAR SASABE															
7.13	2.20	7.27	3660	0	L	13	.26	-.06	159	334	508	802	1040	1300	1940
									142	236	306	404	482	565	776

Table 8.--Flood-frequency data and basin characteristics for gaging stations—Continued

STATION NUMBER AND NAME BASIN CHARACTERISTICS					GROUP	ANNUAL FLOODS			FLOOD ESTIMATES AT GIVEN RECURRENCE INTERVAL, IN YEARS: LINE 1, WEIGHTED; LINE 2, GAGE RECORD						
AREA	SLOPE	SHAPE	ELEV	BDF		YRS	STD	SKEW	2	5	10	25	50	100	500
09486800 ALTAR WASH NEAR THREE POINTS															
463	1.40	2.27	3920	0	R	15	0.33	-0.08	4280	8180	11700	17100	21800	27000	41200
									4470	8370	11600	16200	20200	24500	36100
09487000 BRAWLEY WASH NEAR THREE POINTS															
776	0.48	2.32	3710	0	R	16	.26	-.08	3800	6380	8450	11500	14100	16800	23800
									3860	6370	8250	10800	12900	15100	20600
09487100 LITTLE BRAWLEY WASH NEAR THREE POINTS															
11.9	2.12	2.93	2800	0	R	14	.33	-.14	733	1410	2050	3050	3920	4900	7590
									789	1470	2010	2790	3430	4120	5940
09487140 SAN JOAQUIN WASH NEAR TUCSON															
.45	1.31	13.9	2530	0	R	13	.33	-.18	165	285	382	531	666	816	1230
									192	356	486	670	822	983	1400
09487250 LOS ROBLES WASH NEAR MARANA															
1170	.49	5.09	3350	0	A	13	.47	-.27	1340	3320	5110	7800	10100	12600	19400
									1180	2860	4400	6850	9020	11500	18200
09487400 QUIJOTOA WASH TRIBUTARY NEAR QUIJOTOA															
2.44	1.68	5.31	2800	0	R	13	.42	-.09	168	405	641	1030	1380	1790	2970
									163	367	557	862	1140	1460	2400
09488500 SANTA ROSA WASH NR VAIVA VO															
1782	0.48	1.70	2340	0	R	20	.47	-.01	1450	4100	6720	11100	14800	19000	31200
									1180	2940	4740	7890	11000	14700	26800
09488600 SILVER REEF WASH NEAR CASA GRANDE															
12.8	1.28	5.78	1620	0	R	14	.47	-.08	301	819	1360	2280	3130	4100	7010
									273	666	1050	1710	2330	3060	5320
09514200 WATERMAN WASH NEAR BUCKEYE															
403	.42	2.48	1570	0	R	16	.40	-.08	1510	3390	5070	7690	9960	12500	19600
									1410	3020	4460	6730	8760	11100	17700
09519600 RAINBOW WASH TRIBUTARY NEAR BUCKEYE															
2.43	.65	3.95	950	0	R	17	.36	-.12	354	654	908	1300	1650	2050	3210
									395	779	1100	1580	1990	2440	3670
09519750 BENDER WASH NEAR GILA BEND															
68.8	1.40	4.81	1900	0	R	17	.72	-.13	548	2150	4000	7250	10400	14300	27100
									434	1720	3460	7180	11400	17200	39100
09519760 SAUCEDA WASH NEAR GILA BEND															
126	.88	8.38	1980	0	R	17	.71	-.11	649	2530	4680	8420	12000	16400	30500
									499	1930	3830	7880	12500	18800	42300
09519780 WINDMILL WASH NEAR GILA BEND															
12.9	1.22	6.42	1050	0	R	15	.24	.00	155	1630	2710	4560	6420	8760	16600
									136	1800	3190	5840	8650	12300	25200
09520100 MILITARY WASH NEAR SENTINEL															
8.70	1.06	2.87	674	0	R	17	.68	.03	160	633	1200	2200	3200	4440	8570
									126	470	939	1970	3200	4940	12000
09520110 HOT SHOT ARROYO NEAR AJO															
.44	1.58	5.82	1760	0	R	16	.25	-.06	116	190	254	352	434	523	754
									122	195	249	322	380	441	593
09520130 DARBY ARROYO NEAR AJO															
4.72	1.35	7.63	1920	0	R	16	.32	-.03	469	870	1230	1800	2320	2910	4580
									510	954	1320	1870	2340	2850	4270
09520160 GIBSON ARROYO AT AJO															
2.18	3.24	3.60	2100	0	R	15	.37	.00	222	478	734	1150	1530	1970	3240
									229	473	691	1040	1340	1700	2740
09520170 RIO CORNEZ NEAR AJO															
243	.51	4.65	1950	0	R	14	.24	.04	2870	4680	6260	8620	10600	12900	18800
									3030	4860	6250	8170	9720	11400	15700
09520200 BLACK GAP WASH NEAR AJO															
12.1	.41	2.50	1280	0	R	18	.27	-.07	387	676	920	1290	1610	1940	2830
									392	657	857	1130	1360	1590	2190
09520230 CRATER RANGE WASH NEAR AJO															
1.49	1.31	4.23	1280	0	R	17	.63	-.05	106	339	587	1010	1430	1930	3580
									98	326	607	1180	1800	2620	5630

Table 8.--Flood-frequency data and basin characteristics for gaging stations—Continued

STATION NUMBER AND NAME BASIN CHARACTERISTICS					GROUP	ANNUAL FLOODS			FLOOD ESTIMATES AT GIVEN RECURRENCE INTERVAL, IN YEARS: LINE 1, WEIGHTED; LINE 2, GAGE RECORD						
AREA	SLOPE	SHAPE	ELEV	BDF		YRS	STD	SKEW	2	5	10	25	50	100	500
09520300 ALAMO WASH TRIBUTARY NEAR AJO															
0.90	1.21	6.94	2040	0	R	16	0.32	-0.02	156	290	411	603	773	966	1510
									164	302	416	586	732	895	1340
09520350 MOHAWK PASS WASH AT MOHAWK															
.09	13.4	2.67	601	0	R	14	.55	.03	21	58	101	162	217	282	482
									19	55	97	176	260	369	754
09535100 SAN SIMON WASH NEAR PISINIMO															
569	.33	2.73	2250	0	R	10	.54	.00	1370	3830	6020	9360	12300	15700	25800
									1090	3150	5460	9830	14400	20200	40400
09535200 SELLS WASH TRIBUTARY AT SELLS															
12.2	.95	8.53	2560	0	R	15	.20	-.13	1470	2100	2600	3370	4040	4760	6600
									1680	2470	3000	3690	4210	4720	5940
09535300 VAMORI WASH AT KOM VO															
1250	.30	1.92	2699	0	A	10	.29	.00	792	1610	2290	3310	4140	4990	7170
									672	1170	1570	2130	2610	3120	4490
09536100 PITCHFORK CANYON TRIBUTARY NEAR FORT GRANT															
.81	8.69	4.00	5210	0	R	14	.41	-.26	114	249	378	571	741	930	1450
									119	258	378	557	710	877	1320

[1]Stations maintained by U.S. Agricultural Research Service.

[2]Stations maintained by the Water Resources Research Center, University of Arizona.

Table 9.--Maximum observed floods at gaging stations

Station number	Station name	Water years	Drainage area, in square miles	Maximum observed flood	
				Date	Discharge, in cubic feet per second
09470500	SAN PEDRO RIVER AT PALOMINAS	1930-33,1935-41, 1950-79	741	08-14-40	22,000
09470900	SAN PEDRO RIVER TRIBUTARY NEAR BISBEE	1963-76,1978-79	5.25	09-04-65	1,460
09471000	SAN PEDRO RIVER AT CHARLESTON	1916-79	1,219	09-28-26	98,000
09471080	WALNUT GULCH 63.010 NEAR TOMBSTONE [1]	1967-81	6.42	09-10-67	2,220
09471087	WALNUT GULCH 63.111 NEAR TOMBSTONE [1]	1962-81	.22	07-27-76	541
09471090	WALNUT GULCH 63.009 NEAR TOMBSTONE [1]	1967-81	9.11	07-24-72	2,640
09471110	WALNUT GULCH 63.015 NEAR TOMBSTONE [1]	1955-81	9.24	10-04-54	5,290
09471120	WALNUT GULCH 63.011 NEAR TOMBSTONE [1]	1963-81	3.18	07-22-64	4,390
09471130	WALNUT GULCH 63.008 NEAR TOMBSTONE [1]	1963-81	5.98	07-22-64	4,010
09471140	WALNUT GULCH 63.006 NEAR TOMBSTONE [1]	1962-81	36.70	07-22-64	6,590
09471170	WALNUT GULCH 63.004 NEAR TOMBSTONE [1]	1954-77	.88	07-19-55	1,270
09471180	WALNUT GULCH 63.003 NEAR TOMBSTONE [1]	1954-81	3.47	07-19-55	2,860
09471185	WALNUT GULCH 63.103 NEAR TOMBSTONE [1]	1963-81	.013	07-17-75	30.8
09471190	WALNUT GULCH 63.002 NEAR TOMBSTONE [1]	1954-81	43.90	08-17-57	19,200
09471195	WALNUT GULCH 63.007 NEAR TOMBSTONE [1]	1966-81	5.22	08-12-72	2,590
09471200	WALNUT GULCH 63.001 NEAR TOMBSTONE [1]	1957-81	57.70	08-17-57	11,500
09471550	SAN PEDRO RIVER NEAR TOMBSTONE	1967-79	1,740	10-09-77	24,200
09471600	CANARY WASH NEAR BENSON	1963-75,1978	.79	- -63	84
09471700	FENNER WASH NEAR BENSON	1962-76,1978	2.71	UNKNOWN	950
09472000	SAN PEDRO RIVER NEAR REDINGTON	1926,1931-41, 1943-80	2,939	09-28-26	90,000
09472100	PECK CANYON TRIBUTARY NEAR REDINGTON	1968-81	8.02	08-12-72	4,340
09472400	MAMMOTH WASH NEAR MAMMOTH	1956,1963-76	2.40	UNKNOWN	3,200
09473000	ARAVAIPA CREEK NEAR MAMMOTH	1919-21,1931-41, 1965-80	541	08-02-19	20,000
09473200	GREEN LANTERN WASH NEAR WINKELMAN	1964-76,1981	3.63	05-01-81	3,700
09473500	SAN PEDRO RIVER AT WINKELMAN	1919,1926,1930, 1935,1940,1963-80	4,471	09-28-26	85,000
09473600	TAM O'SHANTER WASH NEAR HAYDEN	1963-76,1981	4.37	08-02-74	1,570
09478200	DURHAM WASH NEAR FLORENCE	1954-57,1963-76, 1980	15.60	08-20-71	3,500
09478500	QUEEN CREEK AT WHITLOW DAMSITE NEAR SUPERIOR	1917-20,1948-59, 1961-81	144	08-19-54	42,900
09478600	QUEEN CREEK TRIBUTARY NO.3 AT WHITLOW DAM	1966-79	.37	09-13-66	280
09479200	QUEEN CREEK TRIBUTARY AT APACHE JUNCTION	1961-79	.51	09-30-71	262
09480000	SANTA CRUZ RIVER NEAR LOCHIEL	1949-80	82.20	10-09-77	12,000
09480500	SANTA CRUZ RIVER NEAR NOGALES	1930-79	533	10-08-77	19,300
09481500	SONOITA CREEK NEAR PATAGONIA	1930-72,1978	209	09-30-46	14,000
09481700	CALABASAS CANYON NEAR NOGALES	1963-65,1967-76, 1978	10.30	10-09-77	1,200
09481750	SOPORI WASH AT AMADO	1948,1954-58, 1964-76,1978	176	08-15-48	16,000
09481800	DEMETRIE WASH TRIBUTARY NEAR CONTINENTAL	1963-76	.15	09-07-75	110
09482000	SANTA CRUZ RIVER AT CONTINENTAL	1940-47,1952-80	1,662	10-09-77	26,500
09482200	FLATO WASH NEAR SAHUARITA	1955,1961,1965-80	6.65	- -55	4,500
09482330	PUMPING WASH NEAR VAIL	1966-81	.81	07- -71	337

Table 9.--Maximum observed floods at gaging stations__Continued

Station number	Station name	Water years	Drainage area, in square miles	Maximum observed flood	
				Date	Discharge, in cubic feet per second
09482350	SOUTH FORK AIRPORT WASH NEAR TUCSON	1966-80	4.38	07-08-74	1,890
09482370	NORTH FORK AIRPORT WASH NEAR TUCSON	1961,1965-80	5.28	08-22-61	1,350
09482400	AIRPORT WASH AT TUCSON	1961,1965-80	5.28	08-22-61	1,350
09482410	RODEO WASH AT TUCSON	1970-81	7.24	07-20-70	898
09482420	JULIAN WASH AT TUCSON	1970-81	26.50	07-19-70	1,270
09482450	WEST BRANCH SANTA CRUZ RIVER AT TUCSON	1966-81	23.60	09-25-76	910
09482480	BIG WASH AT TUCSON	1966-81 1966-81	2.94	08-17-71	3,000
09482500	SANTA CRUZ RIVER AT TUCSON	1915-79	2,222	10-10-77	23,700
09482950	RAILROAD WASH AT TUCSON[2]	1970-74,1976-81	2.30	07-19-71	1,590
09483000	TUCSON ARROYO AT VINE AVENUE, AT TUCSON	1956-81	8.20	08-22-61	5,000
09483010	HIGH SCHOOL WASH AT TUCSON[2]	1968-81	.95	08-12-72	800
09483025	SILVERCROFT WASH AT TUCSON	1965,1969-81	2.74	07-20-70	1,500
09483030	ANKLAM WASH AT TUCSON	1965-81	2.11	08-17-71	2,420
09483040	WEST SPEEDWAY WASH NEAR TUCSON	1965-81	.46	09-25-76	240
09483042	CEMETERY WASH AT TUCSON	1966-81	1.17	08-20-68	600
09483045	FLOWING WELLS WASH AT TUCSON	1971-81	3.53	08- -71	1,250
09483100	TANQUE VERDE CREEK NEAR TUCSON	1960-81	43	07-30-81	6,700
09483200	AGUA CALIENTE WASH TRIBUTARY AT TUCSON	1965-80	2.04	08-19-71	430
09483250	ROB WASH AT TUCSON	1971-81	2.08	- -71	1,400
09484000	SABINO CREEK NEAR TUCSON	1932-81	35.50	09-06-70	7,730
09484200	BEAR CREEK NEAR TUCSON	1960-74,1979	16.30	12-18-78	1,400
09484500	TANQUE VERDE CREEK AT TUCSON	1940-45,1966-81	219	12-18-78	12,700
09484560	CIENEGA CREEK NEAR PANTANO	1958,1968-81	289	08-11-58	20,000
09484570	MESCAL ARROYO NEAR PANTANO	1958,1965-81	38.40	08-11-58	27,000
09484580	BARREL CANYON NEAR SONOITA	1962-76	14.10	08- -71	1,900
09484590	DAVIDSON CANYON WASH NEAR VAIL	1968-81	50.50	07-20-70	6,860
09484600	PANTANO WASH NEAR VAIL	1958-81	457	08-11-58	38,000
09485000	RINCON CREEK NEAR TUCSON	1953-81	44.80	08-19-71	9,660
09485500	PANTANO WASH AT TUCSON	1940,1958,1965-76, 1979-81	602	08-12-58	20,000
09485550	ARCADIA WASH AT TUCSON[2]	1966,1968-81	2.72	08-17-71	1,210
09485900	PIMA WASH NEAR TUCSON	1964-81	4.93	07-26-78	300
09485950	GERONIMO WASH NEAR TUCSON	1964-81	2.15	08-12-71	705
09486000	RILLITO CREEK NEAR TUCSON	1915-81	918	09-23-29	24,000
09486300	CANADA DEL ORO NEAR TUCSON	1959,1961,1964,	250	12-20-67	13,900
09486500	SANTA CRUZ RIVER AT CORTARO	1940-47,1950-79	3,503	10-10-77	23,000
09486700	CHILTEPINES WASH NEAR SASABE	1963-75	7.13	09-10-64	560
09486800	ALTAR WASH NEAR THREE POINTS	1966-80	463	09-04-70	22,000
09487000	BRAWLEY WASH NEAR THREE POINTS	1962,1966-81	776	09-04-70	13,700
09487100	LITTLE BRAWLEY WASH NEAR THREE POINTS	1962,1968-81	11.90	09-26-62	13,800
09487140	SAN JOAQUIN WASH NEAR TUCSON	1969-81	.45	07-28-81	520
09487250	LOS ROBLES WASH NEAR MARANA	1962,1966-78	1,170	09-26-62	32,000
09487400	QUIJOTOA WASH TRIBUTARY NEAR QUIJOTOA	1963-75	2.44	07-24-64	715
09488500	SANTA ROSA WASH NEAR VAIVA VO	1955-74	1,782	09-27-62	53,100

Table 9.--Maximum observed floods at gaging stations__Continued

Station number	Station name	Water years	Drainage area, in square miles	Maximum observed flood	
				Date	Discharge, in cubic feet per second
09488600	SILVER REEF WASH NEAR CASA GRANDE	1963-75	12.80	08-03-71	1,400
09514200	WATERMAN WASH NEAR BUCKEYE	1964-78,1980	403	09-03-67	6,300
09519600	RAINBOW WASH TRIBUTARY NEAR BUCKEYE	1963-79	2.43	09-03-67	1,430
09519750	BENDER WASH NEAR GILA 8END	1963-79	68.80	08- -71	2,670
09519760	SAUCEDA WASH NEAR GILA BEND	1963-79	126	09-26-76	3,150
09519780	WINDMILL WASH NEAR GILA BEND	1964-78	12.90	12-19-67	4,430
09520100	MILITARY WASH NEAR SENTINEL	1963-79	8.70	08-02-74	1,530
09520110	HOT SHOT ARROYO NEAR AJO	1966-81	.44	09-05-76	240
09520130	DARBY ARROYO NEAR AJO	1966-81	4.72	09-06-67	1,670
09520160	GIBSON ARROYO AT AJO	1967-81	2.18	08-02-70	1,800
09520170	RIO CORNEZ NEAR AJO	1967-80	243	09-04-76	8,030
09520200	BLACK GAP WASH NEAR AJO	1962-79	12.10	07-20-79	940
09520230	CRATER RANGE WASH NEAR AJO	1963-79	1.49	09-04-69	590
09520300	ALAMO WASH TRIBUTARY NEAR AJO	1963-76,1978-79	.90	08-31-72	510
09520350	MOHAWK PASS WASH AT MOHAWK	1963-76	.09	08-01-70	117
09535100	SAN SIMON WASH NEAR PISINIMO	1972-81	569	09-24-76	12,500
09535200	SELLS WASH TRIBUTARY AT SELLS	1962-76	12.20	09-13-66	2,800
09535300	VAMORI WASH AT KOM VO	1972-81	1,250	10-20-72	1,880
09536100	PITCHFORK CANYON TRIBUTARY NEAR FORT GRANT	1963-76	.81	08-15-65	375

[1]Station maintained by U.S. Agricultural Research Service.

[2]Station maintained by the Water Resources Research Center, University of Arizona.