USGS
science for a changing world

In cooperation with the
Lower Colorado River Authority

# Peak-Flow Frequency for Tributaries of the Colorado River Downstream of Austin, Texas

Water-Resources Investigations Report 98–4015

U.S. Department of the Interior
U.S. Geological Survey

# Peak-Flow Frequency for Tributaries of the Colorado River Downstream of Austin, Texas

By William H. Asquith

U.S. GEOLOGICAL SURVEY
Water-Resources Investigations Report 98–4015

In cooperation with the
Lower Colorado River Authority

Austin, Texas
1998

**U.S. DEPARTMENT OF THE INTERIOR**

Bruce Babbitt, Secretary

**U.S. GEOLOGICAL SURVEY**

Thomas J. Casadevall, Acting Director



**For additional information write to:**

**District Chief**
**U.S. Geological Survey**
**8011 Cameron Rd.**
**Austin, TX 78754–3898**

**Copies of this report can be purchased from:**
**U.S. Geological Survey**
**Branch of Information Services**
**Box 25286**
**Denver, CO 80225–0286**

# CONTENTS



## FIGURES



## TABLES

# Peak-Flow Frequency for Tributaries of the Colorado River Downstream of Austin, Texas

*By* William H. Asquith

## Abstract

A procedure to estimate the peak discharge associated with large floods is needed for tributaries of the Colorado River downstream of Austin, Texas, so that appropriate peak discharges can be used to estimate floodplain boundaries and used for the design of bridges and other structures. The U.S. Geological Survey, in cooperation with the Lower Colorado River Authority, studied flood peaks for streams in all or parts of 22 counties in that part of the Colorado River Basin extending downstream of Town Lake in Austin to the Gulf of Mexico. The study area was selected because the streams in this area either are tributaries to the Colorado River or have flood characteristics similar to those tributaries.

Peak-flow frequency for 38 stations with at least 8 years of data in natural (unregulated and nonurbanized) basins was estimated on the basis of annual peak-streamflow data through water year 1995. Peak-flow frequency represents the peak discharges for recurrence intervals of 2, 5, 10, 25, 50, 100, 250, and 500 years. The peak-flow frequency and drainage basin characteristics for the stations were used to develop two sets of regression equations to estimate peak-flow frequency for tributaries of the Colorado River in the study area. One set of equations was developed for contributing drainage areas less than 32 square miles, and another set was developed for contributing drainage areas greater than 32 square miles. A procedure is presented to estimate the peak discharge at sites where both sets of equations are considered applicable. Additionally, procedures are presented to compute the 50-, 67-, and 90-percent prediction interval for any estimation from the equations.

## INTRODUCTION

The tributaries of the Colorado River downstream of Austin, Texas, are in an area subject to continentally generated storms and hurricanes. These storms occasionally produce local or widespread flooding, sometimes with little warning. Many storms exceeding 10 inches (in.) of precipitation have been documented in the area, including some exceeding 36 in. (Ellsworth, 1923). These storms typically produce large peak discharges that can threaten lives and property. A procedure to estimate the peak discharge associated with large floods is needed for tributaries in this area so that appropriate peak discharges can be used to estimate floodplain boundaries and used for the design of bridges and other structures.

In 1997, the U.S. Geological Survey (USGS), in cooperation with the Lower Colorado River Authority, studied flood peaks for streams in all or parts of 22 counties in the Colorado River Basin (watershed) extending downstream of Town Lake in Austin to the Gulf of Mexico. The study area (fig. 1) was selected because the streams either are tributaries to the Colorado River or have flood characteristics similar to those tributaries. Another study, conducted simultaneously with this one, documents peak-flow frequency along the main stem of the Colorado River downstream of Austin (Asquith, 1997).

The mean annual precipitation in the study area for 1951–80 ranges from about 32 in. in the Austin area to about 44 in. near the mouth of the Colorado River in the Gulf of Mexico (Riggio and others, 1987, p. 23). Recently, storms producing severe flooding have occurred in or near the study area (Liscum and East, 1994; National Oceanic and Atmospheric Administration, 1995; Hejl and others, 1996).

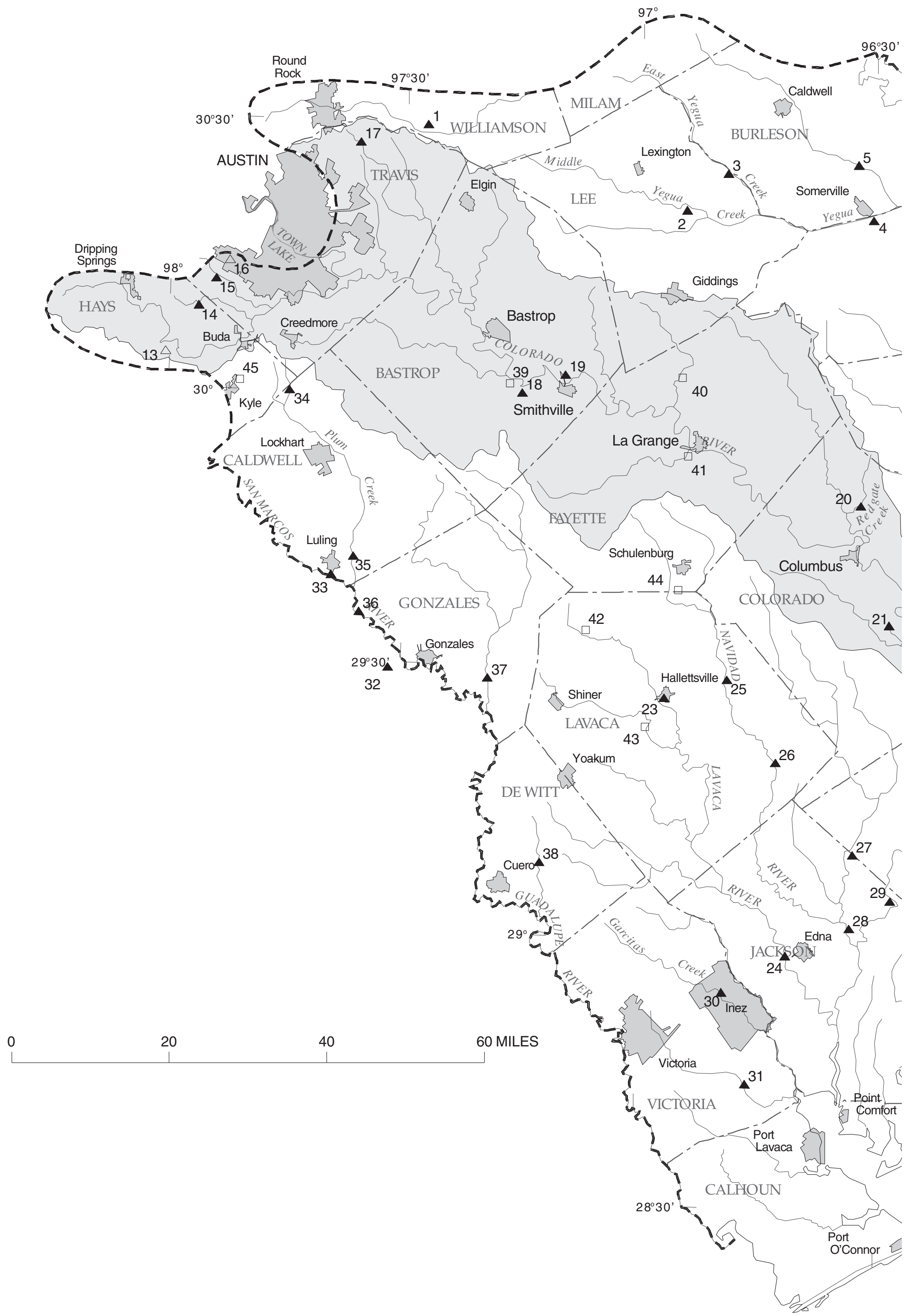
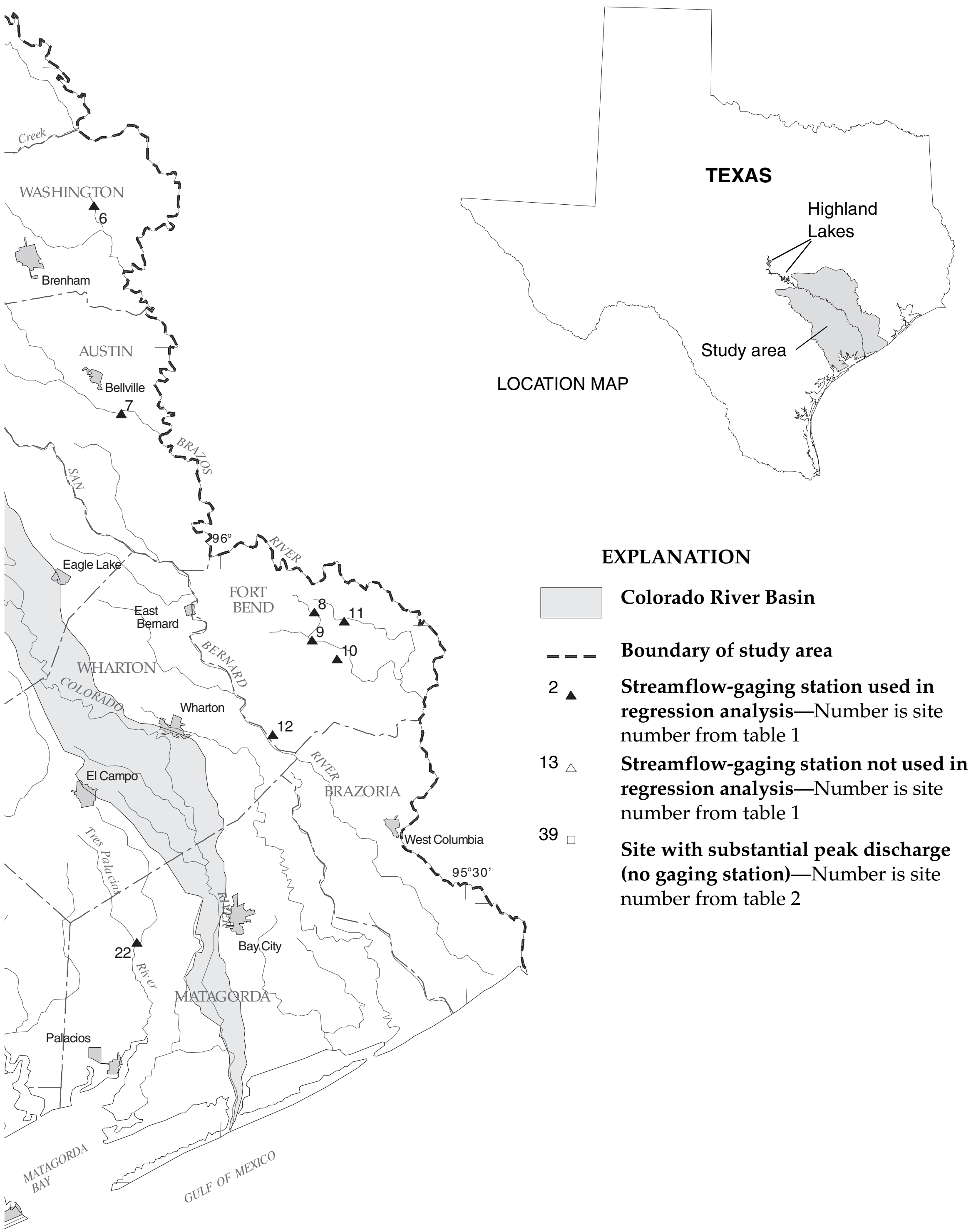



**Figure 1.** Location of streamflow-gaging stations and sites with substantial peak discharges in the vicinity of the




Colorado River Basin downstream of Austin, Texas.

## Purpose and Scope

The purpose of this report is to present equations to estimate peak-flow frequency for tributaries of the Colorado River in the study area. These tributaries have natural drainage basins; a natural drainage basin is defined as having less than 10-percent impervious cover and less than 10 percent of its drainage area controlled by reservoirs. Peak-flow frequency in this report refers to the peak discharges for recurrence intervals of 2, 5, 10, 25, 50, 100, 250, and 500 years.

Peak-flow frequency for 38 streamflow-gaging stations (stations) with at least 8 years of data in natural (unregulated and nonurbanized) basins in the study area (fig. 1, table 1 at end of report) was estimated on the basis of annual peak-streamflow data through water year 1995. The peak-flow frequency was estimated following guidelines established by the Interagency Advisory Committee on Water Data (IACWD, 1982). One station, East Pecan Branch near Gonzales (site 32), lies just outside the study area, but was included in the data base. This station represents a small basin (defined in the "Peak-Flow Frequency Estimation" section), and its addition to the data base was considered necessary to increase data availability.

Equations to estimate peak-flow frequency were developed for streams with natural drainage basins in the study area. These equations were developed from selected stations based on the relation between peak-flow frequency and basin characteristics for each station. The entire period of systematic record (through 1995) was used in the frequency analysis for each qualified station except for stations with regulated streamflow during part of the record. These stations are Yegua Creek near Somerville (site 4) and Plum Creek near Luling (site 35). One or more reservoirs were constructed in the basins upstream of each of these stations during the period of systematic record, resulting in regulated annual peak discharges. The regulated data were excluded from the analysis.

## Peak-Flow Frequency for Streamflow-Gaging Stations

Peak discharges are or have been monitored at each of the stations in the study area. These stations have various periods of systematic record (table 1). The peak discharges used in this study include the largest peak discharge for each year of systematic record (annual peak discharge) and all known historical peak discharges through 1995. A historical peak discharge—documented by newspaper articles, personal recollections, or other historical sources—represents the largest peak discharge since a known date preceding the beginning of the systematic record. Historical peak discharges can occur before or within the systematic record. The historical record is the number of years represented by the historical peaks. For example, 33 years of systematic record (1963–95) exist for the station Davidson Creek near Lyons (site 5). However, the October 1994 peak discharge is the highest since at least 1902; thus, the 1994 peak is the largest in at least 94 years (1902–95).

The annual and historical peak discharges for each station were used, together with the USGS computer program PEAKFQ (Slade and Asquith, 1996), to estimate peak discharges for the 2-, 5-, 10-, 25-, 50-, 100-, 250-, and 500-year recurrence intervals (table 1). The computer program for peak-flow frequency analysis follows the guidelines established by the Interagency Advisory Committee on Water Data (1982) and uses the log-Pearson Type III (LPIII) frequency distribution.

The skew in the distribution of annual peaks is characterized by a skew coefficient. A reliable skew coefficient is difficult to estimate for stations having short records. Therefore, the IACWD recommends using a weighted skew coefficient with the LPIII distribution. This weighted skew coefficient is computed by weighting the skew coefficient calculated for a station (station skew coefficient) with a generalized skew coefficient representative of the surrounding area. The weighted skew coefficient is based on the inverse of the respective mean square errors for the station and on generalized skew coefficients. Generalized skew coefficients were determined in a previous study for stations in the study area (Judd and others, 1996). A weighted skew coefficient then was used in the calculation of the peak-flow frequency for each station except Dry Creek near Rosenberg (site 11). For this station, only the station skew coefficient was used because, compared to the weighted skew coefficient, it produces a better fit of the LPIII frequency curve to the data.

Additionally, the IACWD provides a procedure for estimating low-outlier thresholds; annual peak discharges less than this threshold are excluded from the fitting of the LPIII frequency curve. The estimation of low-outlier thresholds is critical in the calculation of peak-flow frequencies. The IACWD procedure for estimating low-outlier thresholds is not always suitable for all stations in Texas because of the large variability and skewness of annual peaks at a single station. An equation to estimate low-outlier thresholds for Texas stations with natural basins was developed by Asquith and others (1995) and was used to estimate most low-outlier thresholds for the present investigation. The equation estimates low-outlier thresholds by using the mean, standard deviation, and skew of the logarithms for the systematic annual peak discharges. At one station—Mill Creek near Bellville (site 7)—the low-outlier threshold differed from the threshold from the equation. This alternative low-outlier threshold was identified by visually fitting the LPIII frequency curve to the peak-flow data.

### Basin Characteristics for Streamflow-Gaging Stations

Selected basin characteristics were aggregated for each station (table 1). The characteristics selected for estimation of peak-flow frequency are from previous investigations of peak-flow frequency in Texas (Schroeder and Massey, 1977; Slade and others, 1995; Asquith and others, 1996). The 2-year 24-hour precipitation comes from Hershfield (1962). The mean annual precipitation comes from U.S. Geological Survey (1986, p. 432) and is for the period 1951–80. The contributing drainage area (CDA) represents the areal extent of the drainage basin. The stream length represents the length, in miles, of the longest mapped channel from the station to the drainage divide at the headwaters, based on USGS quadrangle maps (scale, 1:100,000). The basin shape factor is the ratio of the square of the stream length to the CDA; the shape factor represents the ratio of the longest stream length to the mean width of the basin. The stream slope is the ratio of the change in elevation of the longest mapped channel (from the station upstream to the drainage divide at the headwaters) to the stream length in feet per mile.

### Large Recorded Floods

The USGS routinely has documented substantial peak discharges at stream sites without stations. Such sites in the Colorado River Basin downstream of Austin (from Asquith and Slade, 1995) are identified in figure 1 (sites 39–45), and ancillary information is listed in table 2 (at end of report). The relation to CDA of maximum peak discharge for the stations and substantial peak discharge at stream sites without stations is shown in figure 2. The relation provides an informative description of large recorded floods in the study area.

## PEAK-FLOW FREQUENCY ESTIMATION

Development of equations for estimation of peak-flow frequency is a multistep process requiring investigation of the relation between peak discharge and selected basin characteristics. The following sections discuss multiple regression, the relation between the 100-year peak discharge and selected basin characteristics, prediction interval calculation, and application of the equations. A detailed example of peak-discharge estimation is available in the appendix.

### Multiple-Regression Equations to Estimate Peak-Flow Frequency

Multiple-regression analysis was used to establish statistical relations between one dependent and one or more independent variables. The 2-, 5-, 10-, 25-, 50-, 100-, 250-, and 500-year peak discharges were used as dependent variables, and the selected basin characteristics were used as independent variables. Logarithmic transformations of the dependent and independent variables were used to increase linearity between variables. Logarithmic transformations traditionally have been used for hydrologic regression. In this report, logarithmic transformations are base 10 and are represented as "log."

A forward-stepwise weighted least-squares (WLS) regression procedure was used for the development of the equations to estimate peak-flow frequency. In WLS regression, each data point can be assigned a weighting factor. The weights generally are representative of the relative accuracy of each value for the dependent variable; greater weights are assigned to values having greater accuracy.

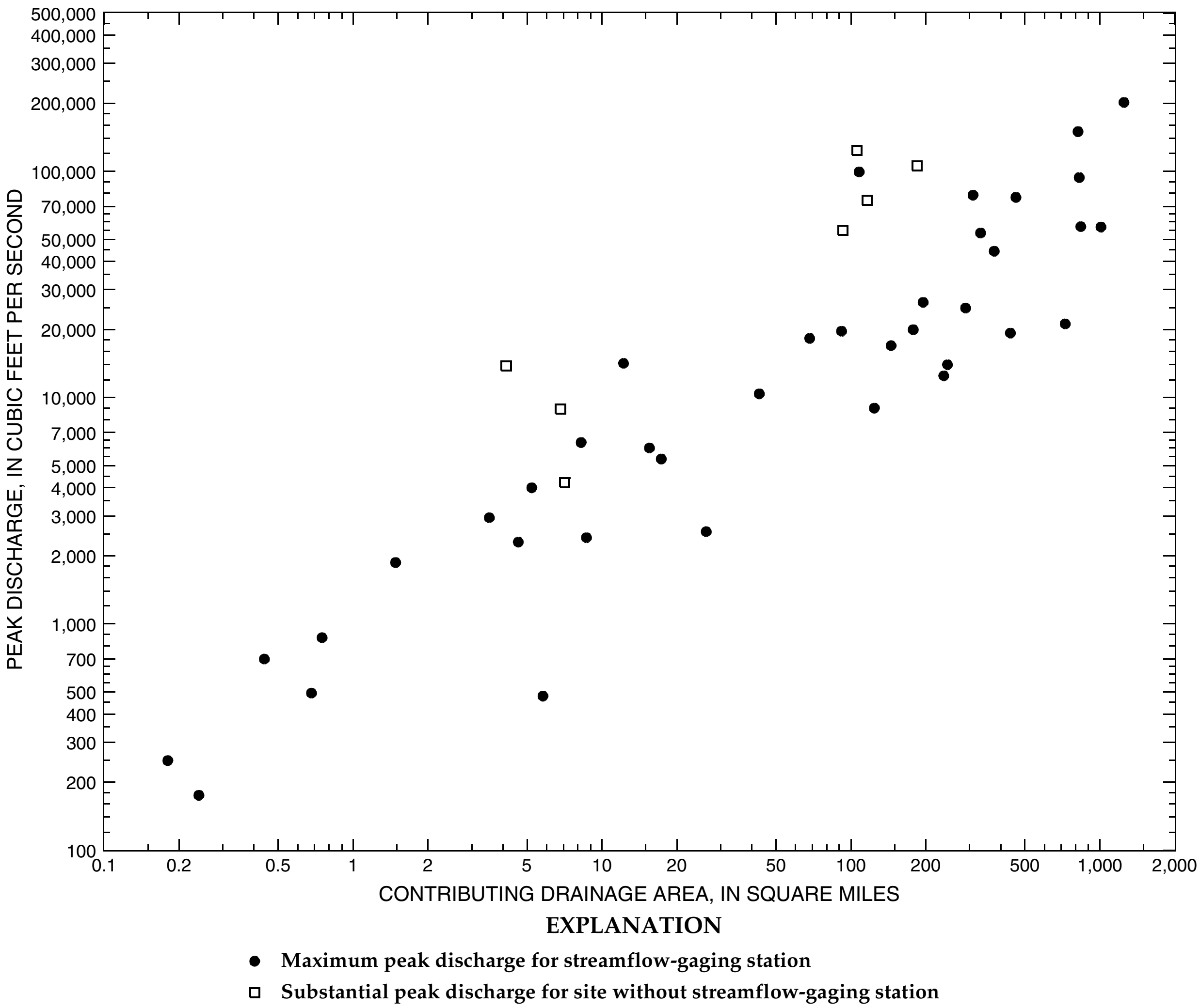



**Figure 2.** Large floods recorded in the Colorado River Basin downstream of Austin, Texas.

Empirical equations (G.D. Tasker, U.S. Geological Survey, written commun., 1994) based on Monte Carlo simulations (Tasker and Thomas, 1978; Stedinger and Cohn, 1986) were used to calculate a regression weight factor. The regression weight factor represents an equivalent "years of record" for each station with historical information. This factor is based on the length of systematic record, length of historical record, and number of high outliers; the regression weight factor for each station is listed in table 1. These factors were used as the weights for the WLS regression procedure in this report.

In forward-stepwise regression, the independent variable (basin characteristic) having the highest mathematical correlation to the dependent variable (peak-flow frequency) is entered into the equation, and successively, the remaining independent variables are tested for their statistical significance to the dependent variable. Each independent variable testing statistically significant (F ratio > 1.5) is entered into the equation. Thus, each independent variable in the final equation is considered statistically significant, and a basin characteristic's inclusion into the equation contributes to the explanation of the variance in the dependent variable.

Before the final equations were developed, an investigation of the relation between the 100-year peak discharge ($Q_{100}$) and the principal physical basin characteristics (CDA, stream slope, and basin shape factor) was conducted. The 100-year peak discharge was selected because this discharge information most often is needed by water managers and planners. Similar results, as described below, were recorded for the other recurrence intervals. The relation between $Q_{100}$ and CDA, stream slope, and basin shape factor is presented in figure 3. Additionally, 16 stations (including Fox Branch near Oak Hill, site 16) with a CDA less than 32 square miles ($mi^2$), and 22 stations (including Onion Creek near Driftwood, site 13) with a CDA greater than 32 $mi^2$ are shown in figure 3. Hereafter, in this report, stations with a CDA less than 32 $mi^2$ are referred to as "small basin" stations, and stations with a CDA greater than 32 $mi^2$ are referred to as "large basin" stations. A study of peak-flow frequency conducted by Asquith and Slade (1997) indicates that a change in flood characteristics for many parts of the State occurs at CDA of about 32 $mi^2$ or 1.5 log-$mi^2$.

The relation between $Q_{100}$ and the basin characteristics indicates that a slight curvilinear relation exists between $Q_{100}$ and CDA. By developing separate regression equations for the small and large basin stations, improved estimates of $Q_{100}$ are calculated because of increased linearity (Asquith and Slade, 1997). Also, classifying the stations as "small" and "large" produces a change in the slope of the relation between $Q_{100}$ and CDA. The slope (fig. 3b) is better correlated to $Q_{100}$ for the large basins than for the small basins. The apparent change in the relation of stream slope between the large and small basins indicates that separate regression equations would improve $Q_{100}$ estimation.

Also in figure 3b the relation between stream slope and $Q_{100}$ is proportional for each CDA class (small and large basin stations)—larger slopes are associated with larger values of $Q_{100}$. However, if all the stations are considered together, then the apparent relation between $Q_{100}$ and stream slope becomes inversely proportional—larger slopes are associated with smaller values of $Q_{100}$. Asquith and Slade (1997) found slope to be proportional to $Q_{100}$. It is thought that steep slopes produce more rapid runoff and therefore larger peak discharges. Thus, the classification of stations as small and large produces a stream slope and $Q_{100}$ relation with some conceptualized physical basis.

Two stations were excluded from the regression analysis; Onion Creek near Driftwood (site 13) and one of its tributaries, Fox Branch near Oak Hill (site 16). The watershed for each of these stations (fig. 1) lies entirely within the Texas Hill Country where topography and soil characteristics are significantly different than for the remaining stations in the study area. Both Onion Creek near Driftwood and Fox Branch near Oak Hill have stream slopes that are substantially different in comparison to other stations in their respective CDA classes (fig. 3b). Excluding these two stations reduces the number of stations available for the regression analysis to 15 small basin stations and 21 large basin stations.

Two sets of equations were developed using the 15 small basin and 21 large basin stations to estimate 2-, 5-, 10-, 25-, 50-, 100-, 250-, and 500-year peak discharges (table 3 at end of report). Of the six independent variables considered, stream length was excluded from the regression analysis because it correlates highly with the CDA. Many equations were developed for each recurrence interval using various combinations of the five remaining independent variables (2-year 24-hour precipitation, mean annual precipitation, CDA, basin shape factor, and stream slope) in the forward-stepwise regression procedure. CDA and stream slope consistently were the most significant independent variables. Additional basin characteristics were significant for some recurrence intervals but were not retained in the final equations. Testing of equations showed that if the independent variables were not consistent among the equations, some combinations of variable values can produce inconsistent peak discharges between recurrence intervals. For example, an inconsistent peak discharge would occur if the peak discharge for a given recurrence interval is less than the peak discharge for a smaller recurrence interval. Thus, only CDA and stream slope appear in the equations listed in table 3.

## Discussion of Regression Equations

CDA proved to be the most significant variable in all the regression equations (table 3); whereas, stream slope proved to be the second most significant variable. Also, the exponents on CDA and stream slope are positive, which indicates that peak discharge is directly proportional to each of these characteristics. Additionally, the exponents on CDA and stream slope increase with increasing recurrence interval, a characteristic

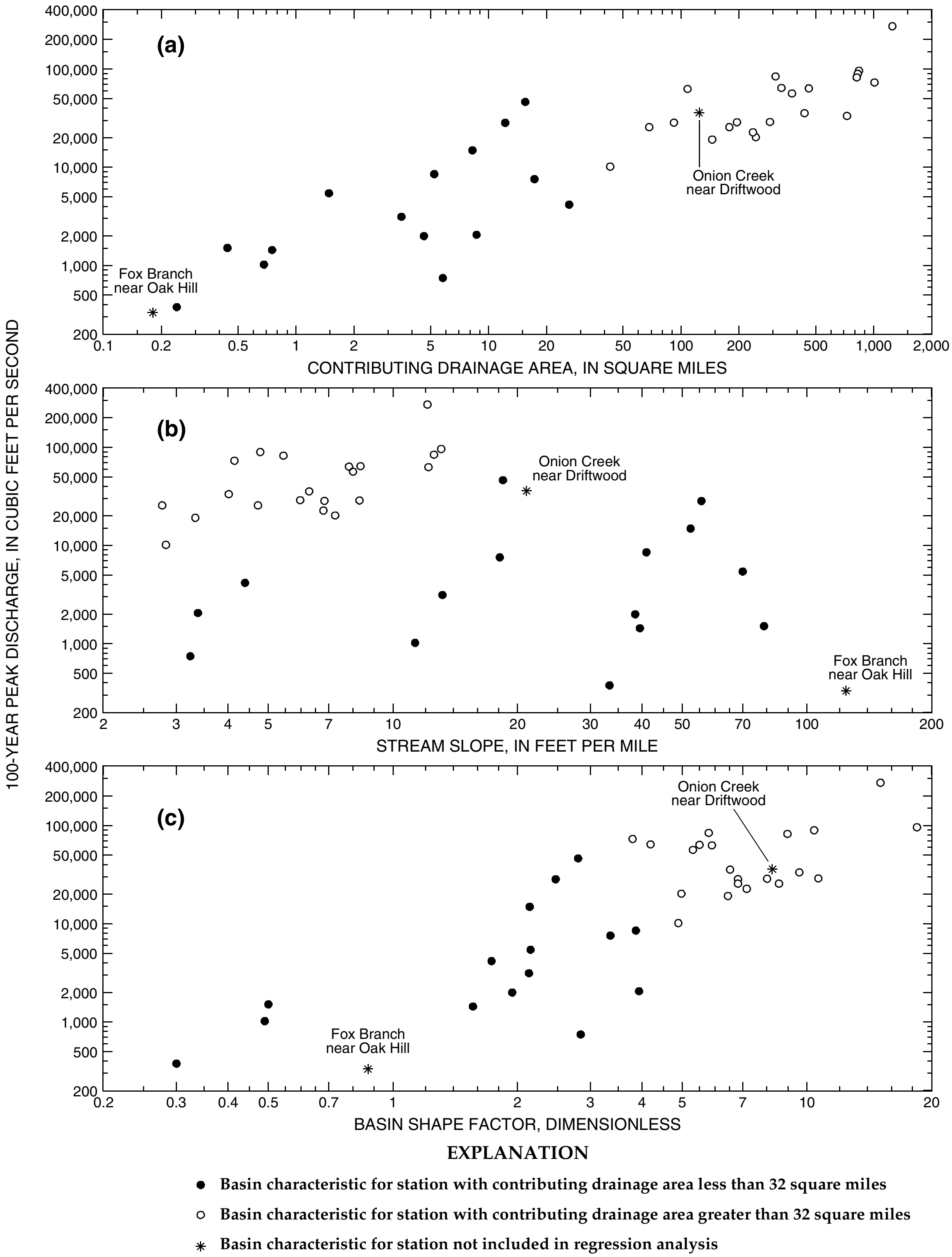



**Figure 3.** Relation between 100-year peak discharge and (a) contributing drainage area, (b) stream slope, and (c) basin shape factor for streamflow-gaging stations in the vicinity of the Colorado River Basin downstream of Austin, Texas.

frequently seen in peak-flow frequency equations (Schroeder and Massey, 1977; Slade and others, 1995; Asquith and others, 1996; and Asquith and Slade, 1997). The regression constants for each of the two sets of equations are similar as recurrence interval increases, which indicates that CDA and stream slope produce increasing peak discharge for increasing recurrence interval. The exponent on stream slope for each 2-year equation is considerably smaller than the analogous value for larger recurrence intervals (table 3). The influence of slope—as represented by large exponents—increases with increasing recurrence interval.

The (weighted) standard errors of estimate (standard deviation of residuals) for small basins generally are greater than the standard errors of estimate for large basins. The greater error is attributed to availability of short record (less than 20 years) for 13 of 15 small basin stations. Consequently, the peak-flow frequency for the short record stations shows greater variability. The mean record length for the small basin stations is about 14 years, whereas the mean record length for the large basin stations is about 33 years. Thus, the peak-flow frequency for the small basin stations has more potential error.

Finally, the standard error of estimate dramatically increases as the recurrence interval increases for the small basin stations (table 3), whereas the standard error remains relatively constant for the large basin stations. The increasing standard errors with recurrence interval for the small basins are also partially attributed to the short record availability, which results in increased extrapolation errors of the LPIII distribution for the larger recurrence intervals.

## Prediction Intervals

The regression equations of table 3 list individual estimates of peak discharge, and the adjusted R-squared and (weighted) standard errors of estimate provide overall measures of equation accuracy. However, frequently the accuracy (error) of individual estimates from the equation is desired. As estimates are made for basin characteristics (independent variables) further “away” from the basin characteristics of the stations on which the regression is based, the accuracy of the estimate decreases. The means for evaluating the accuracy of any estimate from the equations is the calculation of the prediction intervals (Helsel and Hirsch, 1992, p. 241–242, 300).

The prediction interval of any discharge estimate for a site from the regression equations in table 3 is computed from the relation (Helsel and Hirsch, 1992, p. 300, sec. 11.4.4):

$$\log Q_T - t_{(\alpha/2, n-p)}\sqrt{s^2(1+h_o)} \le \log Q_T \le \log Q_T + t_{(\alpha/2, n-p)}\sqrt{s^2(1+h_o)}\,, \qquad (1)$$

where

$\log Q_T$ = the logarithm of T-year recurrence interval peak discharge for a site from a regression equation, in log-cubic feet per second,

$t_{(\alpha/2, n-p)}$ = the critical value of t-distribution for a 100 • (1-$\alpha$)-percent confidence interval,

$n$ = the number of stations in a regression equation,

$p$ = the number of independent variables (basin characteristics) plus one (for regression constant),

$s^2$ = the square of standard error of estimate, and

$h_o$ = the leverage of the site.

The leverage of a site is an expression of the distance of the site’s basin characteristics from the center of the multidimensional space defined by the independent variables in the regression. The prediction interval is related directly to the square root of the magnitude of the leverage. The leverage of a site is computed by the following equation (Helsel and Hirsch, 1992, p. 300):

$$h_o = \mathbf{x_o}\{\mathbf{X}^{\boldsymbol{T}}\mathbf{W}^{-1}\mathbf{X}\}^{-1}\mathbf{x_o}^{\mathbf{T}}\,, \qquad (2)$$

where

$\mathbf{x_o}$ = a row vector of the log basin characteristics for the site,

$\{\mathbf{X}^{\boldsymbol{T}}\mathbf{W}^{-1}\mathbf{X}\}^{-1}$ = the covariance matrix of the regression equation, and

$\mathbf{x_o}^{\mathbf{T}}$ = a column vector of the log basin characteristics of the site.

The covariance matrices for the regression equations are listed in table 4 (at end of report); the covariance matrices are in log-space. Critical values for the t-distribution for selected prediction intervals pertinent to the regression equations are also listed in table 4. Application of the above equations for prediction interval calculation is presented later.

## Application of Regression Equations

The equations are designed to be used only for sites on the Colorado River tributaries in the study area. The equations are less applicable as the distance from the center of the study area increases. However, quantification of the diminishing applicability with increasing distance from the center of the study area is difficult. Also, the applicability of the equations for sites outside the study area is questionable.

In order to provide a smooth transition in determining peak discharge between small and large basins (separated at a CDA of 32 mi$^2$), a weighted peak-discharge estimate is used for sites within a CDA interval of 10 to 100 mi$^2$ (Asquith and Slade, 1997, p. 12). The logarithmically-weighted peak discharge (eq. 3) is based on estimates from both the equation for CDA less than 32 mi$^2$ and the equation for CDA greater than 32 mi$^2$.

$$Q_{TW} = (2 - \log A)Q_{T1} + (\log A - 1)Q_{T2} \quad , \qquad (3)$$

where

$Q_{TW}$ = the weighted peak discharge associated with the T-year recurrence interval, in cubic feet per second,

$Q_{T1}$ = the peak discharge associated with the equation for sites with CDAs less than 32 mi$^2$, in cubic feet per second,

$Q_{T2}$ = the peak discharge associated with the equation for sites with CDAs greater than 32 mi$^2$, in cubic feet per second, and

$A$ = the CDA of the site, in square miles.

Since a weighted peak discharge is used for sites within the CDA range of 10 to 100 mi$^2$, a weighted prediction interval is needed as well. The prediction interval (see "Prediction Intervals" section) of a weighted peak discharge for a site within the 10- to 100-mi$^2$ CDA range is estimated by the following equations, which are constructed similar to equation 3:

$$Q_{TW-UP} = (2 - \log A)Q_{T1-UP} + (\log A - 1)Q_{T2-UP}\text{, and} \qquad (4)$$

$$Q_{TW-LOW} = (2 - \log A)Q_{T1-LOW} + (\log A - 1)Q_{T2-LOW}\ , \qquad (5)$$

where

$Q_{TW-UP}$ = the upper prediction interval of the $Q_{TW}$ (eq. 3), in cubic feet per second,

$Q_{T1-UP}$ = the upper prediction interval of the $Q_{T1}$ (eq. 3), in cubic feet per second,

$Q_{T2-UP}$ = the upper prediction interval of the $Q_{T2}$ (eq. 3), in cubic feet per second,

$Q_{TW-LOW}$ = the lower prediction interval of the $Q_{TW}$ (eq. 3), in cubic feet per second,

$Q_{T1-LOW}$ = the lower prediction interval of the $Q_{T1}$ (eq. 3), in cubic feet per second, and

$Q_{T2-LOW}$ = the lower prediction interval of the $Q_{T2}$ (eq. 3), in cubic feet per second.

For sites at or near the stations included in the regression analyses, the peak discharge for such sites can be calculated from the regression equations $(Q_T^r)$. However, the regression equations might provide estimates that conflict with those from the LPIII frequency analysis of the station data $(Q_T^f)$. Thus, a weighted discharge $(Q_T^w)$ calculated from $Q_T^r$ and $Q_T^f$ can be used. The weighted discharge is calculated from equation 6 in which the relative weights assigned to $Q^r$ and $Q^f$ are inversely related for the error variances (squared standard errors of estimate in log units) for the values of $Q_T^r$ and $Q_T^f$.

$$Q_T^w = \frac{Q_T^r (SE_T^f)^2 + Q_T^f (SE_T^r)^2}{(SE_T^f)^2 + (SE_T^r)^2}\ , \qquad (6)$$

where:

$Q_T^w$ = the weighted peak-discharge estimate for recurrence interval $T$, in cubic feet per second,

$Q_T^r$ = the peak-discharge estimate from regression, in cubic feet per second,

$Q_T^f$ = the peak-discharge estimate from the LPIII frequency analysis, in cubic feet per second,

$SE_T^f$ = the square standard error of the LPIII distribution, in log-cubic feet per second, and

$SE_T^r$ = the square standard error from the regression, in log-cubic feet per second.

The standard error from the regression $(SE_T^f)$ in the above equation is estimated from one-half the difference of the logarithmic 67-percent prediction interval,

$$\frac{1}{2}[\log(Q_T^r - Q_{T-HIGH}^{67\text{ percent}}) - \log(Q_T^r - Q_{T-LOW}^{67\text{ percent}})]\ .$$

The standard error from the LPIII distribution is estimated as one-half the difference of the logarithmic 67-percent confidence limits, which can be determined from program PEAKFQ (confidence limits set at 83.5 percent[1] documented by Slade and Asquith (1996)).

Equation 6 provides a simple and convenient means of combining $Q^r$ and $Q^f$; however, the development of an always appropriate weighting procedure is difficult. The difficulties include: (1) the consideration of the number of years of record available at a station; peak discharges calculated from the regression equations are preferred in cases of very short record (about 15 years or less); (2) the consideration of the fit of the LPIII distribution to the station data; and (3) the knowledge that the presence of low and high outliers in the data can greatly affect the resulting LPIII distribution. Another difficulty in developing an always appropriate weighting procedure is how to quantify what is meant by "near" a station. The $Q_T^f$, derived from annual peak data, is increasingly less applicable for sites increasingly farther away (upstream or downstream) from the station.

A final consideration when using a specific equation for a specific site concerns the range in basin characteristics for which the equations are applicable. If a regression equation is used to estimate peak discharge for a site having a combination of basin characteristics substantially different from the stations on which the equation is based, erroneous estimates could result. Two techniques are available for evaluation of equation applicability. Each of these techniques involves analysis of the differences between the basin characteristics of the site and the basin characteristics involved in the regression. The first technique is to graphically compare the basin characteristics of the site with the characteristics involved in the regression. The second technique is to compare the leverage of the site with the maximum or mean leverage of the stations in the regression.

The graphical technique for evaluating the applicability of an equation is done by using the relation among the basin characteristics involved in the regression equations; the internal relation among the basin characteristics (CDA, stream slope, and basin shape factor) is shown in figure 4. If the basin characteristics for a site plot near or outside the approximate region defined by the basin characteristics (indicated by the shaded area), the resulting peak discharge could be erroneous. Consequently, the equations become less applicable when the site basin characteristics plot farther away from the shaded areas. Although, basin shape factor does not appear in the equations, the shape factor is intrinsically involved because it was considered in the regression analysis. For example, if stations with substantially different shape factors were available in the study area, then shape factor could have proven statistically significant. Therefore, the applicability of the equations is reduced for sites having basin shape factors substantially different from those in the data base.

The second technique for evaluation of equation applicability uses the leverage statistic, $h_o$ (eq. 2). Because leverage is a measure of the distance from the center of the multidimensional space defined by the basin characteristics involved in the regression analysis, the error of a peak-discharge estimate is directly proportional to the square root of leverage (eq. 1)—high leverage sites have large prediction intervals. The applicability of the equations diminishes as the site leverage increases. One guide in considering applicability of the equation is whether the site leverage is less than the maximum leverage of the regression. The maximum leverage for the small basin equations is 0.54, whereas the maximum leverage for the large basin equations is 0.38. Another guide in considering the equation applicability is whether the site leverage is less than 2 to 3 times the mean leverage. The mean leverage for the small basin equations is 0.24, whereas the mean leverage for the large basin equations is 0.14. More thorough discussion pertinent to interpretation of the leverage statistic is available in Helsel and Hirsch (1992, p. 300–305).

The most reliable peak-discharge estimates from the equations are for sites that plot within the shaded areas of figure 4 and have sufficiently small leverage

[1]Program PEAKFQ uses a different definition of confidence limits than used in the current report. The 67-percent confidence limits of this report are the {100 - (100 - 67)/2} = 83.5-percent confidence limits in PEAKFQ.

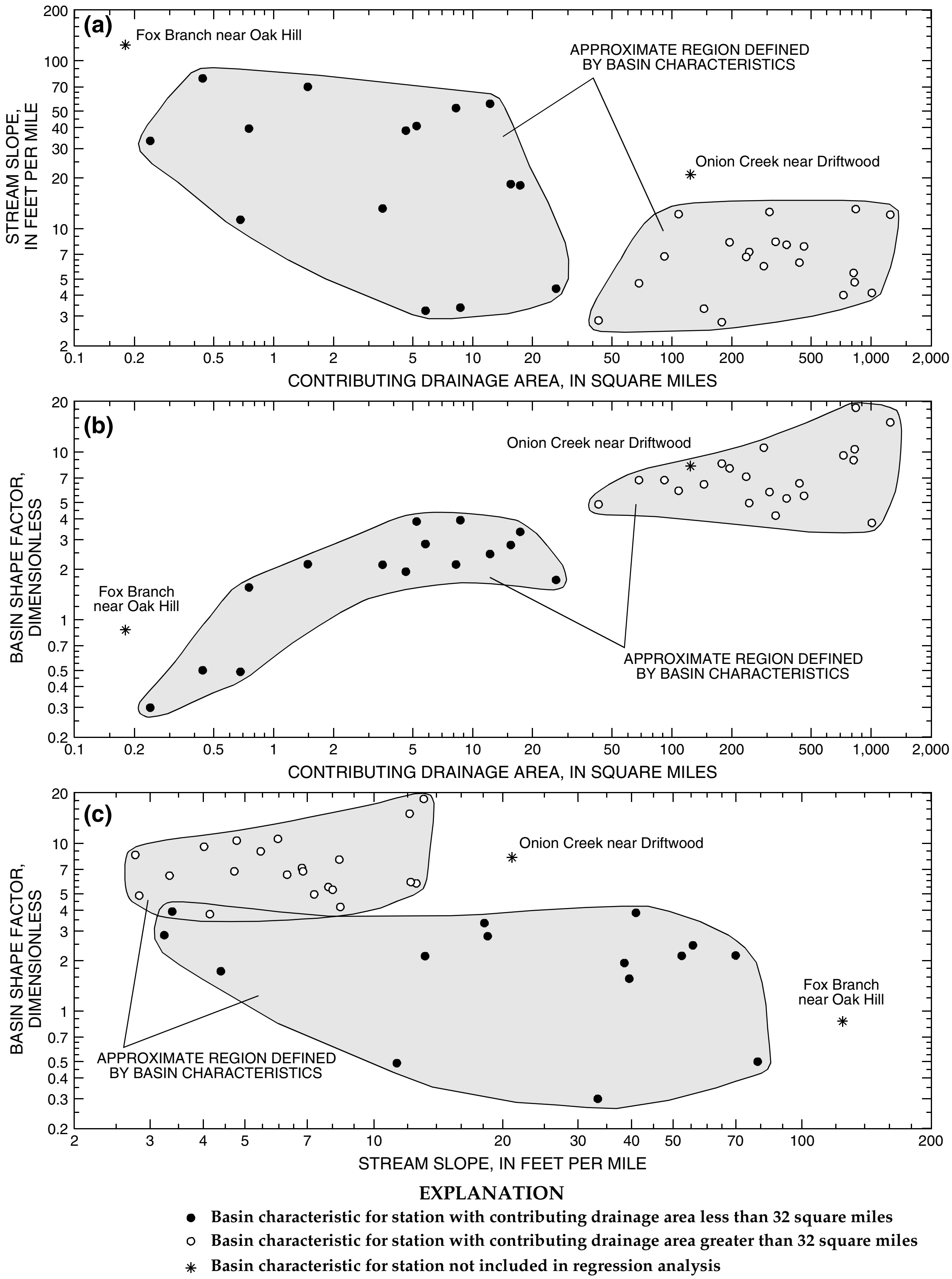



**Figure 4.** Relation between (a) stream slope and contributing drainage area, (b) basin shape factor and contributing drainage area, and (c) basin shape factor and stream slope for streamflow-gaging stations in the vicinity of the Colorado River Basin downstream of Austin, Texas.

statistics. The reliability of peak-discharge estimates is questionable for sites that plot near or outside the shaded areas in figure 4 or for sites that have large leverage. However, it is difficult to determine for which sites the equations become unapplicable.

## SUMMARY

Equations estimating peak-flow frequency are described for tributaries of the Colorado River in a study area encompassing all or parts of 22 counties downstream of Austin, Texas, to the Gulf of Mexico. The study area was selected because streams in the area are either tributaries to the Colorado River or have characteristics similar to the tributaries.

The peak-flow frequency was estimated for 38 qualified streamflow-gaging stations (stations) in the study area. Peak-flow frequency refers to the peak discharges for recurrence intervals of 2, 5, 10, 25, 100, 250, and 500 years. Qualified stations have at least 8 years of peak-flow data collected from natural basins. A natural drainage basin has less than 10-percent impervious cover and less than 10 percent of its drainage area controlled by reservoirs. The entire period of systematic record (through water year 1995) was used in the frequency analysis for each station, except at stations where streamflow was regulated during part of the record. Regulated data was excluded from the analyses.

Data from 36 of 38 stations were used to develop regressions equations to estimate peak-flow frequency. Multiple-regression analysis was performed using weighted least squares. Two sets of regression equations for peak-flow frequency were developed on the basis of the relation between peak-flow frequency and the basin characteristics of the stations. One set of equations was developed for contributing drainage area (CDA) less than 32 $mi^2$, and another set was developed for CDA greater than 32 $mi^2$. Procedures to compute the 50-, 67-, and 90-percent confidence limits for any estimation derived from the equations are presented. Additionally, a procedure is presented for determining a weighted discharge for sites where both sets of equations are applicable.

## REFERENCES CITED


Asquith, W.H., 1997, Peak-flow frequency along the Colorado River downstream of Austin, Texas: U.S. Geological Survey Water-Resources Investigations Report 97–4265, 19 p.

Asquith, W.H., and Slade, R.M., Jr., 1995, Documented and potential extreme peak discharges and relation between potential extreme peak discharges and probable maximum flood peak discharges in Texas: U.S. Geological Survey Water-Resources Investigations Report 95–4249, 58 p.

______1997, Regional equations for estimation of peak-streamflow frequency for natural basins in Texas: U.S. Geological Survey Water-Resources Investigations Report 96–4307, 27 p.

Asquith, W.H., Slade, R.M., Jr., and Judd, L.J., 1995, Analysis of low-outlier thresholds for log Pearson Type III peak-streamflow frequency analysis in Texas, *in* Texas Water '95, A Component Conference of the First International Conference on Water Resources Engineering, Water Resources Engineering Division, American Society of Civil Engineers, August 16–17, 1995, San Antonio, Proceedings: San Antonio, American Society of Civil Engineers, p. 379–384.

Asquith, W.H., Slade, R.M., Jr., and Lanning-Rush, Jennifer, 1996, Peak-flow frequency and extreme flood potential for streams in the vicinity of the Highland Lakes, central Texas: U.S. Geological Survey Water-Resources Investigations Report 96–4072, 1 sheet.

Ellsworth, C.E., 1923, The floods in central Texas in September, 1921: U.S. Geological Survey Water-Supply Paper 488, 56 p.

Hejl, H.R., Jr., Slade, R.M., Jr., and Jennings, M.E., 1996, Floods in central Texas, December 1991: U.S. Geological Survey Water-Resources Investigations Report 95–4289, 1 sheet.

Helsel, D.R., and Hirsch, R.M., 1992, Studies in environmental science 49—Statistical methods in water resources: Amsterdam, Elsevier, 522 p.

Hershfield, D.M., 1962, Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years: U.S. Weather Bureau Technical Paper 40, Washington D.C., 61 p.

Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood flow frequency: Reston, Va., U.S. Geological Survey, Office of Water Data Coordination, Hydrology Subcommittee Bulletin 17B [variously paged].

Judd, L.J., Asquith, W.H., and Slade, R.M., Jr., 1996, Techniques to estimate generalized skew coefficients of annual peak streamflow for natural basins in Texas: U.S. Geological Survey Water-Resources Investigations Report 96–4117, 28 p.

Liscum, Fred, and East, J.W., 1994, Floods in Southeast Texas, October 1994: U.S. Geological Survey Fact Sheet 94–073, 2 p.

National Oceanic and Atmospheric Administration, 1995, Southeast Texas tropical mid-latitude rainfall and flood event, October 1994: Natural Disaster Survey Report, 23 p.

Riggio, R.F., Bomar, G.W., and Larkin, T.J., 1987, Texas droughts—its recent history (1931–1985): Texas Water Commission Report LP 87–04, 74 p.

Schroeder, E.E., and Massey, B.C., 1977, Technique for estimating the magnitude and frequency of floods in Texas: U.S. Geological Survey Water-Resources Investigations/Open-File Report 77–110, 22 p.

Slade R.M., Jr., and Asquith, W.H., 1996, Peak data for U.S. Geological Survey gaging stations, Texas network; and computer program to estimate peak-streamflow frequency: U.S. Geological Survey Open-File Report 96–148, 57 p.

Slade R.M., Jr., Asquith, W.H., and Tasker, G.D., 1995, Multiple-regression equations to estimate peak-flow frequency for streams in Hays County, Texas: U.S. Geological Survey Water-Resources Investigations Report 95–4019, 1 sheet.

Stedinger, J.R., and Cohn, T.A., 1986, Flood frequency analysis with historical and paleoflood information: Water Resources Research, v. 22, no. 5, p. 785–793.

Tasker, G.D., and Thomas, W.O., 1978, Flood frequency analysis with pre-record information: Journal of the Hydraulics Division, American Society of Civil Engineers, v. 104, no. 2, p. 249–259.

U.S. Geological Survey, 1986, National water summary 1985: Water-Supply Paper 2300, 506 p.

**Table 1.** Selected basin characteristics and peak-flow frequency for streamflow-gaging stations in the vicinity of the Colorado River Basin downstream of Austin, Texas

[in., inches; FM, Farm Road; $mi^2$, square miles; mi, miles; ft/mi, feet per mile; yr, years; $ft^3/s$, cubic feet per second; $Q_{500}$, 500-year peak discharge; $Q_{100}$, 100-year peak discharge]

| Site no. (fig. 1) | Station no. | Station name | Latitude | Longitude | Available period of systematic record[1] | 2-year 24-hour precipitation (in.) | Mean annual precipitation 1951–80 (in.) |
|---|---|---|---|---|---|---|---|
| 1 | 08105900 | Avery Branch near Taylor | 30°29'11" | 97°27'27" | 1967–74 | 4.19 | 33.0 |
| 2 | 08109700 | Middle Yegua Creek near Dime Box | 30°20'21" | 96°54'16" | 1963–95 | 4.32 | 35.0 |
| 3 | 08109800 | East Yegua Creek near Dime Box | 30°24'26" | 96°49'02" | 1963–95 | 4.35 | 36.0 |
| 4 | 08110000 | Yegua Creek near Somerville | 30°19'18" | 96°30'26" | 1925–66[2] | 4.39 | 36.0 |
| 5 | 08110100 | Davidson Creek near Lyons | 30°25'10" | 96°32'24" | 1963–95 | 4.42 | 36.0 |
| 6 | 08111100 | Winkleman Creek near Brenham | 30°15'19" | 96°15'44" | 1966–74 | 4.62 | 39.0 |
| 7 | 08111700 | Mill Creek near Bellville | 29°52'51" | 96°12'18" | 1964–93 | 4.60 | 40.0 |
| 8 | 08114900 | Seabourne Creek near Rosenberg | 29°31'27" | 95°48'28" | 1967–74 | 4.90 | 43.0 |
| 9 | 08115000 | Big Creek near Needville | 29°28'35" | 95°48'45" | 1947–95 | 4.88 | 43.0 |
| 10 | 08115500 | Fairchild Creek near Needville | 29°26'45" | 95°45'41" | 1945–54 | 4.92 | 43.0 |
| 11 | 08116400 | Dry Creek near Rosenberg | 29°30'42" | 95°44'48" | 1959–79 | 4.93 | 43.0 |
| 12 | 08117500 | San Bernard River near Boling | 29°18'48" | 95°53'37" | 1955–95 | 4.75 | 42.0 |
| 13 | 08158700 | Onion Creek near Driftwood[3] | 30°04'59" | 98°00'29" | 1980–95 | 4.01 | 32.5 |
| 14 | 08158810 | Bear Creek below FM 1826 near Driftwood | 30°09'19" | 97°56'23" | 1979–95 | 4.06 | 33.0 |
| 15 | 08158840 | Slaughter Creek at FM 1826 near Austin | 30°12'32" | 97°54'11" | 1978–95 | 4.07 | 32.5 |
| 16 | 08158900 | Fox Branch near Oak Hill[3] | 30°14'01" | 97°52'29" | 1966–74 | 4.08 | 32.0 |
| 17 | 08159150 | Wilbarger Creek near Pflugerville | 30°27'16" | 97°36'02" | 1964–80 | 4.15 | 32.0 |
| 18 | 08159450 | Reeds Creek near Bastrop | 30°00'26" | 97°15'03" | 1965–74 | 4.32 | 36.0 |
| 19 | 08160000 | Dry Creek at Buescher Lake near Smithville | 30°02'32" | 97°09'34" | 1940–66 | 4.35 | 36.5 |
| 20 | 08160800 | Redgate Creek near Columbus | 29°47'56" | 96°31'55" | 1962–95 | 4.60 | 40.5 |
| 21 | 08161580 | Dry Branch tributary near Altair | 29°34'39" | 96°28'16" | 1967–74 | 4.65 | 41.0 |
| 22 | 08162600 | Tres Palacios River near Midfield | 28°55'40" | 96°10'15" | 1971–95 | 4.79 | 42.0 |
| 23 | 08163500 | Lavaca River at Hallettsville | 29°26'35" | 96°56'39" | 1940–95 | 4.43 | 38.0 |
| 24 | 08164000 | Lavaca River near Edna | 28°57'35" | 96°41'10" | 1939–95 | 4.47 | 37.5 |
| 25 | 08164300 | Navidad River near Hallettsville | 29°28'00" | 96°48'45" | 1962–95 | 4.47 | 38.5 |
| 26 | 08164350 | Navidad River near Speaks | 29°19'18" | 96°42'32" | 1982–89 | 4.48 | 39.0 |
| 27 | 08164450 | Sandy Creek near Louise | 29°09'34" | 96°32'47" | 1980–95 | 4.60 | 41.0 |
| 28 | 08164500 | Navidad River near Ganado | 29°01'32" | 96°33'08" | 1939–80 | 4.57 | 40.0 |
| 29 | 08164503 | West Mustang Creek near Ganado | 29°04'17" | 96°28'01" | 1980–95 | 4.70 | 41.0 |
| 30 | 08164600 | Garcitas Creek near Inez | 28°53'28" | 96°49'08" | 1971–95 | 4.51 | 37.0 |
| 31 | 08164800 | Placedo Creek near Placedo | 28°43'30" | 96°46'07" | 1971–95 | 4.53 | 38.0 |
| 32 | 08169850 | East Pecan Branch near Gonzales | 29°29'58" | 97°31'36" | 1966–74 | 4.28 | 34.5 |
| 33 | 08172000 | San Marcos River at Luling | 29°39'54" | 97°38'59" | 1940–95 | 4.03 | 33.0 |
| 34 | 08172100 | West Elm Creek near Niederwald | 29°59'04" | 97°44'39" | 1966–74 | 4.16 | 34.0 |
| 35 | 08173000 | Plum Creek near Luling | 29°41'58" | 97°36'12" | 1930–63[4] | 4.19 | 35.0 |
| 36 | 08173500 | San Marcos River at Ottine | 29°35'36" | 97°35'22" | 1916–43 | 4.10 | 33.5 |
| 37 | 08174600 | Peach Creek below Dilworth | 29°28'26" | 97°18'59" | 1960–79 | 4.32 | 36.5 |
| 38 | 08176200 | Irish Creek near Cuero | 29°08'02" | 97°12'10" | 1967–74 | 4.41 | 35.0 |

[1] Stations with available record ending 1995 were active as of 1996.
[2] Somerville Lake completed in 1967. Only peak discharges for 1925–66 used in analysis.
[3] Station not used in regression analysis.
[4] Many small reservoirs completed in 1964. Only peak discharges for 1930–63 used in analysis.

Table continued on next two pages.

**Table 1.** Selected basin characteristics and peak-flow frequency for streamflow-gaging stations in the vicinity of the Colorado

| Site no. (fig. 1) | Contributing drainage area ($mi^2$) | Stream length (mi) | Basin shape factor (dimensionless) | Stream slope (ft/mi) | Available unregulated systematic record (yr) | No. of historical peaks | Historical record (yr) | Equivalent years of record (regression weight factor) |
|---|---|---|---|---|---|---|---|---|
| 1 (Avery Branch-Taylor) | 3.52 | 2.74 | 2.13 | 13.17 | 8 | 0 | 0 | 8 |
| 2 (Middle Yegua Creek-Dime Box) | 236 | 41.08 | 7.15 | 6.79 | 33 | 1 | 83 | 51 |
| 3 (East Yegua Creek-Dime Box) | 244 | 34.84 | 4.97 | 7.25 | 33 | 0 | 0 | 33 |
| 4 (Yegua Creek-Somerville) | 1,009 | 61.83 | 3.79 | 4.14 | 42 | 0 | 0 | 42 |
| 5 (Davidson Creek-Lyons) | 195 | 39.52 | 8.01 | 8.30 | 33 | 2 | 94 | 56 |
| 6 (Winkleman Creek-Brenham) | .75 | 1.08 | 1.56 | 39.49 | 9 | 0 | 0 | 9 |
| 7 (Mill Creek-Bellville) | 376 | 44.62 | 5.30 | 8.01 | 30 | 0 | 0 | 30 |
| 8 (Seabourne Creek-Rosenberg) | 5.78 | 4.06 | 2.84 | 3.24 | 8 | 0 | 0 | 8 |
| 9 (Big Creek-Needville) | 42.8 | 14.47 | 4.89 | 2.83 | 48 | 0 | 0 | 48 |
| 10 (Fairchild Creek-Needville) | 26.2 | 6.73 | 1.73 | 4.39 | 8 | 0 | 0 | 8 |
| 11(Dry Creek-Rosenberg) | 8.65 | 5.83 | 3.93 | 3.38 | 21 | 1 | 48 | 31 |
| 12 (San Bernard River-Boling) | 727 | 83.45 | 9.58 | 4.01 | 41 | 0 | 0 | 41 |
| 13 (Onion Creek-Driftwood) | 124 | 31.99 | 8.25 | 21.02 | 16 | 0 | 0 | 16 |
| 14 (Bear Creek-Driftwood) | 12.2 | 5.49 | 2.47 | 55.55 | 17 | 1 | 81 | 40 |
| 15 (Slaughter Creek-Austin) | 8.24 | 4.20 | 2.14 | 52.32 | 18 | 0 | 0 | 18 |
| 16 (Fox Branch-Oak Hill) | .18 | .40 | .87 | 124.32 | 9 | 0 | 0 | 9 |
| 17 (Wilbarger Creek-Pflugerville) | 4.61 | 2.99 | 1.94 | 38.44 | 17 | 1 | 87 | 42 |
| 18 (Reeds Creek-Bastrop) | 5.22 | 4.49 | 3.86 | 40.91 | 10 | 0 | 0 | 10 |
| 19 (Dry Creek-Smithville) | 1.48 | 1.78 | 2.15 | 69.96 | 26 | 1 | 32 | 28 |
| 20 (Redgate Creek-Columbus) | 17.3 | 7.61 | 3.35 | 18.11 | 34 | 0 | 0 | 34 |
| 21 (Dry Branch tributary-Altair) | .68 | .58 | .49 | 11.31 | 8 | 0 | 0 | 8 |
| 22 (Tres Palacios River-Midfield) | 145 | 30.55 | 6.44 | 3.33 | 25 | 0 | 0 | 25 |
| 23 (Lavaca River-Hallettsville) | 108 | 25.23 | 5.89 | 12.18 | 56 | 2 | 156 | 92 |
| 24 (Lavaca River-Edna) | 817 | 85.59 | 8.97 | 5.44 | 57 | 3 | 116 | 80 |
| 25 (Navidad River-Hallettsville) | 332 | 37.31 | 4.19 | 8.35 | 34 | 0 | 0 | 34 |
| 26 (Navidad River-Speaks) | 437 | 53.39 | 6.52 | 6.27 | 8 | 0 | 0 | 8 |
| 27 (Sandy Creek-Louise) | 289 | 55.48 | 10.65 | 5.97 | 16 | 0 | 0 | 16 |
| 28 (Navidad River-Ganado) | 826 | 92.67 | 10.40 | 4.78 | 42 | 2 | 105 | 66 |
| 29 (West Mustang Creek-Ganado) | 178 | 39.05 | 8.56 | 2.77 | 16 | 0 | 0 | 16 |
| 30 (Garcitas Creek-Inez) | 91.7 | 24.98 | 6.81 | 6.83 | 25 | 2 | 93 | 51 |
| 31 (Placedo Creek-Placedo) | 68.3 | 21.57 | 6.81 | 4.72 | 25 | 1 | 29 | 27 |
| 32 (East Pecan Branch-Gonzales) | .24 | .27 | .30 | 33.33 | 9 | 0 | 0 | 9 |
| 33 (San Marcos River-Luling) | 838 | 124.26 | 18.43 | 13.07 | 56 | 0 | 0 | 56 |
| 34 (West Elm Creek-Niederwald) | .44 | .47 | .50 | 78.72 | 9 | 0 | 0 | 9 |
| 35 (Plum Creek-Luling) | 309 | 42.30 | 5.79 | 12.56 | 34 | 1 | 96 | 56 |
| 36 (San Marcos River-Ottine) | 1,249 | 137.12 | 15.05 | 12.11 | 28 | 0 | 0 | 28 |
| 37 (Peach Creek-Dilworth) | 460 | 50.28 | 5.50 | 7.83 | 20 | 1 | 40 | 28 |
| 38 (Irish Creek-Cuero) | 15.5 | 6.59 | 2.80 | 18.43 | 8 | 0 | 0 | 8 |

Continued from previous page.

River Basin downstream of Austin, Texas—Continued

| Date of maximum peak discharge of record | Maximum peak discharge of record ($ft^3/s$) | Peak discharge for indicated recurrence interval obtained from log-Pearson Type III frequency curve ($ft^3/s$) | | | | | | | | $\frac{Q_{500}}{Q_{100}}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 2 yr | 5 yr | 10 yr | 25 yr | 50 yr | 100 yr | 250 yr | 500 yr | |
| 05/01/1972 | 2,950 | 691 | 1,188 | 1,580 | 2,147 | 2,619 | 3,135 | 3,697 | 4,519 | 1.44 |
| 12/22/1991 | 12,500 | 1,490 | 4,650 | 7,860 | 13,060 | 17,640 | 22,270 | 28,220 | 36,080 | 1.62 |
| 05/24/1975 | 14,000 | 2,150 | 5,290 | 8,129 | 12,480 | 16,200 | 20,280 | 24,690 | 31,020 | 1.53 |
| 07/01/1940 | 56,800 | 6,790 | 16,960 | 26,680 | 42,460 | 56,760 | 73,200 | 91,890 | 120,200 | 1.64 |
| 10/17/1994 | 26,400 | 3,660 | 8,160 | 12,100 | 18,070 | 23,180 | 28,810 | 34,960 | 43,890 | 1.52 |
| 03/24/1973 | 870 | 302 | 555 | 748 | 1,010 | 1,220 | 1,440 | 1,660 | 1,970 | 1.37 |
| 06/13/1973 | 44,400 | 13,630 | 22,840 | 29,890 | 39,800 | 47,870 | 56,510 | 65,760 | 79,010 | 1.40 |
| 06/13/1973 | 480 | 269 | 395 | 479 | 586 | 667 | 747 | 827 | 935 | 1.25 |
| 06/26/1960 | 10,400 | 2,520 | 4,200 | 5,470 | 7,240 | 8,660 | 10,170 | 11,770 | 14,050 | 1.38 |
| 05/18/1953 | 2,560 | 997 | 1,680 | 2,210 | 2,940 | 3,530 | 4,160 | 4,840 | 5,800 | 1.39 |
| 10/31/1959 | 2,410 | 689 | 995 | 1,220 | 1,530 | 1,790 | 2,060 | 2,350 | 2,760 | 1.34 |
| 06/28/1960 | 21,200 | 7,080 | 12,350 | 16,550 | 22,660 | 27,780 | 33,400 | 39,550 | 48,570 | 1.45 |
| 06/06/1985 | 8,990 | 2,680 | 7,380 | 12,120 | 20,080 | 27,440 | 36,030 | 45,890 | 60,980 | 1.69 |
| 1939 | 14,200 | 793 | 3,100 | 6,120 | 12,370 | 19,240 | 28,380 | 40,250 | 60,930 | 2.15 |
| 12/20/1991 | 6,330 | 718 | 2,240 | 3,990 | 7,280 | 10,650 | 14,920 | 20,230 | 29,120 | 1.95 |
| 09/04/1967 | 249 | 49 | 99 | 142 | 208 | 267 | 332 | 407 | 518 | 1.56 |
| 1921 | 2,300 | 596 | 980 | 1,240 | 1,550 | 1,780 | 2,000 | 2,210 | 2,480 | 1.24 |
| 05/16/1965 | 4,000 | 230 | 840 | 1,660 | 3,450 | 5,550 | 8,520 | 12,640 | 20,410 | 2.40 |
| 06/30/1940 | 1,870 | 226 | 756 | 1,390 | 2,590 | 3,840 | 5,440 | 7,440 | 10,790 | 1.90 |
| 05/22/1979 | 5,360 | 1,790 | 3,120 | 4,110 | 5,460 | 6,510 | 7,590 | 8,710 | 10,250 | 1.35 |
| 06/13/1973 | 495 | 133 | 283 | 417 | 626 | 811 | 1,020 | 1,260 | 1,620 | 1.59 |
| 10/17/1983 | 17,000 | 5,130 | 8,250 | 10,590 | 13,820 | 16,420 | 19,180 | 22,110 | 26,280 | 1.37 |
| 08/31/1981 | 99,500 | 7,720 | 16,210 | 24,070 | 36,880 | 48,740 | 62,750 | 79,220 | 105,300 | 1.68 |
| 10/19/1994 | 150,000 | 11,590 | 22,910 | 33,120 | 49,520 | 64,540 | 82,190 | 102,800 | 135,400 | 1.65 |
| 09/13/1974 | 53,500 | 9,710 | 19,800 | 28,370 | 41,240 | 52,250 | 64,420 | 77,800 | 97,460 | 1.51 |
| 05/14/1982 | 19,300 | 6,480 | 12,110 | 16,730 | 23,540 | 29,290 | 35,620 | 42,560 | 52,740 | 1.48 |
| 10/19/1994 | 24,900 | 5,050 | 9,410 | 13,090 | 18,670 | 23,530 | 29,010 | 35,170 | 44,480 | 1.53 |
| 1936 | 94,000 | 13,620 | 25,760 | 36,660 | 54,240 | 70,450 | 89,650 | 112,300 | 148,500 | 1.66 |
| 10/19/1994 | 20,000 | 3,850 | 7,540 | 10,790 | 15,870 | 20,410 | 25,650 | 31,650 | 40,920 | 1.60 |
| 06/12/1981 | 19,700 | 4,550 | 8,860 | 12,540 | 18,140 | 23,020 | 28,510 | 34,670 | 43,930 | 1.54 |
| 10/31/1981 | 18,300 | 5,930 | 10,300 | 13,610 | 18,170 | 21,810 | 25,630 | 29,650 | 35,270 | 1.38 |
| 09/13/1974 | 175 | 81 | 141 | 188 | 257 | 314 | 377 | 445 | 545 | 1.45 |
| 09/12/1952 | 57,000 | 10,250 | 23,090 | 35,250 | 55,330 | 74,000 | 96,100 | 122,000 | 163,000 | 1.70 |
| 05/15/1970 | 700 | 185 | 402 | 599 | 910 | 1,190 | 1,510 | 1,880 | 2,440 | 1.62 |
| 07/01/1936 | 78,500 | 6,330 | 15,590 | 25,370 | 43,160 | 61,240 | 84,260 | 113,300 | 162,900 | 1.93 |
| 05/29/1929 | 202,000 | 14,800 | 39,880 | 68,870 | 126,100 | 188,600 | 273,200 | 386,000 | 592,300 | 2.17 |
| 04/20/1977 | 76,800 | 7,420 | 16,040 | 24,060 | 37,170 | 49,280 | 63,570 | 80,300 | 106,700 | 1.68 |
| 05/10/1972 | 6,000 | 1,550 | 5,380 | 10,250 | 20,260 | 31,380 | 46,410 | 66,280 | 101,900 | 2.20 |

**Table 2.** Substantial peak discharges for sites without streamflow-gaging stations in the vicinity of the Colorado River Basin downstream of Austin, Texas

[$mi^2$, square miles; $ft^3/s$, cubic feet per second]

| Site no. (fig. 1) | Stream site name and approximate location | Latitude | Longitude | Contributing drainage area ($mi^2$) | Date of substantial peak discharge | Substantial peak discharge ($ft^3/s$) |
|---|---|---|---|---|---|---|
| 39 | Little Piney Creek near Bastrop | 30°01'00" | 97°16'38" | 7.08 | 05/12/1969 | 4,220 |
| 40 | Rabbs Creek near Warda | 30°01'50" | 96°54'40" | 92.8 | 06/30/1940 | 55,000 |
| 41 | Buckners Creek near La Grange | 29°53'07" | 96°53'55" | 184 | 06/30/1940 | 106,000 |
| 42 | Youngs Branch near Moulton | 29°33'50" | 97°06'40" | 6.8 | 06/30/1940 | 8,900 |
| 43 | Rocky Creek near Hallettsville | 29°23'10" | 96°59'00" | 116 | 06/30/1940 | 74,700 |
| 44 | West Navidad River near Schulenburg | 29°38'20" | 96°55'00" | 106 | 06/30/1940 | 124,000 |
| 45 | Bunton Branch near Kyle | 30°01'00" | 97°51'00" | 4.12 | 06/30/1936 | 13,800 |

**Table 3.** Weighted least-squares regression equations for estimation of peak-flow frequency for the Colorado River Basin downstream of Austin, Texas

[$mi^2$, square miles; yr, year; CDA, contributing drainage area in square miles; SL, stream slope in feet per mile—ratio of change in elevation of (1) the length of the longest mapped channel to (2) the distance from the longest mapped channel of the site (or station) to the headwaters; SH, basin shape factor—ratio of length of longest mapped channel (stream length) squared to contributing drainage area (dimensionless)]

| Recurrence interval | Weighted least-squares regression equation for corresponding recurrence interval[1] (cubic feet per second) | Range of indicated independent variables (units as noted) | Adjusted R-squared[2] | Weighted standard error of estimate (percent) | Weighted standard error of estimate ($log_{10}$ units) | No. of stations in equation |
|---|---|---|---|---|---|---|
| **Sites[3] with contributing drainage area less than 32 $mi^2$ (small basin stations)** | | | | | | |
| **2 yr** | $Q_2 = 176.4\ CDA^{0.611}\ SL^{0.0460}$ | CDA: 0.24 to 26.2 square miles | 0.79 | 38 | 0.16 | 15 |
| **5 yr** | $Q_5 = 139.3\ CDA^{0.701}\ SL^{0.326}$ | | .87 | 31 | .13 | Do. |
| **10 yr** | $Q_{10} = 126.8\ CDA^{0.744}\ SL^{0.460}$ | | .85 | 38 | .16 | Do. |
| **25 yr** | $Q_{25} = 117.5\ CDA^{0.788}\ SL^{0.594}$ | SL: 3.24 to 78.72 feet per mile | .79 | 51 | .21 | Do. |
| **50 yr** | $Q_{50} = 113.3\ CDA^{0.814}\ SL^{0.676}$ | | .75 | 63 | .25 | Do. |
| **100 yr** | $Q_{100} = 110.6\ CDA^{0.836}\ SL^{0.746}$ | | .71 | 75 | .29 | Do. |
| **250 yr** | $Q_{250} = 108.9\ CDA^{0.856}\ SL^{0.807}$ | SH: 0.30 to 3.93 dimensionless | .67 | 88 | .33 | Do. |
| **500 yr** | $Q_{500} = 107.9\ CDA^{0.879}\ SL^{0.878}$ | | .63 | 103 | .37 | Do. |

Footnotes at end of table.

**Table 3.** Weighted least-squares regression equations for estimation of peak-flow frequency for the Colorado River Basin downstream of Austin, Texas—Continued

| Recurrence interval | Weighted least-squares regression equation for corresponding recurrence interval[1] (cubic feet per second) | Range of indicated independent variables (units as noted) | Adjusted R-squared[2] | Weighted standard error of estimate (percent) | Weighted standard error of estimate ($log_{10}$ units) | No. of stations in equation |
|---|---|---|---|---|---|---|
| **Sites[3] with contributing drainage area greater than 32 $mi^2$ (large basin stations)** | | | | | | |
| **2 yr** | $Q_2 = 460.6\ CDA^{0.379}\ SL^{0.239}$ | CDA: 42.8 to 1,249 square miles | 0.32 | 57 | 0.23 | 21 |
| **5 yr** | $Q_5 = 571.8\ CDA^{0.414}\ SL^{0.407}$ | | .54 | 43 | .18 | Do. |
| **10 yr** | $Q_{10} = 617.3\ CDA^{0.437}\ SL^{0.499}$ | | .63 | 38 | .16 | Do. |
| **25 yr** | $Q_{25} = 650.4\ CDA^{0.466}\ SL^{0.598}$ | SL: 2.77 to 13.07 feet per mile | .69 | 38 | .16 | Do. |
| **50 yr** | $Q_{50} = 662.6\ CDA^{0.486}\ SL^{0.664}$ | | .71 | 38 | .16 | Do. |
| **100 yr** | $Q_{100} = 667.0\ CDA^{0.506}\ SL^{0.725}$ | | .72 | 41 | .17 | Do. |
| **250 yr** | $Q_{250} = 665.4\ CDA^{0.525}\ SL^{0.781}$ | SH: 3.79 to 18.43 dimensionless | .73 | 41 | .17 | Do. |
| **500 yr** | $Q_{500} = 656.6\ CDA^{0.549}\ SL^{0.849}$ | | .72 | 43 | .18 | Do. |

[1] Order of independent variables (basin characteristics) in each equation from left to right indicates the relative statistical significance of each independent variable to the dependent variable (peak-flow frequency).

[2] Presented so that equations based on different numbers of stations can be compared.

[3] For stream sites with CDA within the range 10 to 100 $mi^2$, use equation 3 to calculate weighted discharge and equations 4 and 5 to calculate weighted prediction interval.

**Table 4.** Covariance matrices and critical values of the t-distribution for regression equations

[These matrices can be used to compute the prediction interval for any estimation from the regression equations (table 3). Qualifications as to which equations a matrix applies to are shown in this table. CDA, contributing drainage area; $mi^2$, square miles]

**Matrix $\{X^TW^{-1}X\}^{-1}$**

| | Sites with CDA less than 32 $mi^2$ | | | | Sites with CDA greater than 32 $mi^2$ | | |
|---|---|---|---|---|---|---|---|
| | **Area** | **Slope** | **Constant** | | **Area** | **Slope** | **Constant** |
| Area | 0.28411 | 0.10328 | -0.34482 | Area | 0.27877 | -0.052138 | -0.64721 |
| Slope | .10328 | .37361 | -.59347 | Slope | -.052138 | 1.1778 | -.85108 |
| Constant | -.34482 | -.59347 | 1.1372 | Constant | -.64721 | -.85108 | 2.3593 |

**Critical values of the t-distribution for selected confidence limits (two-tailed)**

| | Sites with CDA less than 32 $mi^2$ Degrees of freedom = 12 | Sites with CDA greater than 32 $mi^2$ Degrees of freedom = 18 |
|---|---|---|
| 50 percent ($\alpha = 0.25$) | 0.69548 | 0.68836 |
| 67 percent ($\alpha = 0.165$) | 1.0153 | 1.0012 |
| 90 percent ($\alpha = 0.05$) | 1.7823 | 1.7340 |

# Appendix—
# Example of Calculations

The following example illustrates the calculation of peak-flow frequency and prediction intervals for a site. The example involves the estimation of the 100-year peak discharge for a 50-$mi^2$ site with a stream slope of 10.5 feet per mile (ft/mi) and a basin shape factor of 5.0; additionally, the 67-percent prediction interval is desired. Because the CDA for the site is within the interval 10 to 100 $mi^2$, estimates are made from equations with CDA both less than and greater than 32 $mi^2$.

The estimate from the equation for CDA < 32 $mi^2$ (small basin estimate) is:

$$Q_{100<32} = 110.6\ (50)^{0.836}\ (10.5)^{0.746} = 16{,}800 \text{ cubic feet per second (ft}^3\text{/s)}.$$

The estimate from the equation for CDA > 32 $mi^2$ (large basin estimate) is:

$$Q_{100>32} = 667\ (50)^{0.506}\ (10.5)^{0.725} = 26{,}600 \text{ ft}^3\text{/s}.$$

Because the CDA of the site lies within the 10- to 100-$mi^2$ interval, the weighted estimate is:

$$Q_{100W} = [2 - \log(50)]\ (16{,}800) + [\log_{10}(50) - 1]\ (26{,}600) = 0.30\ (16{,}800) + 0.70\ (26{,}600)$$

$$Q_{100W} = 23{,}700 \text{ ft}^3\text{/s}.$$

A check on the applicability of the equations indicates that this site's basin characteristics plot in or near the shaded regions in figure 4. Notice that the site CDA and stream slope plot closer to the shaded region of the large CDA basins than to the shaded region of the small CDA basins; therefore, the large basin estimate has the larger weight factor (0.070) in the weighting equation than the small basin estimate (0.30). Thus, the large basin estimate is considered slightly more applicable than the small basin estimate.

The calculation of the 67-percent prediction interval for the large basin equation begins with calculation of the leverage, $h_o$, using the covariance matrix for sites with a CDA greater than 32 $mi^2$ from table 4. The $\mathbf{x_o}$ vector for this site is [log(50) log(10.5) 1] or [1.6990 1.0212 1]. The large basin leverage is:

$$h_o = [1.6990\ 1.0212\ 1] \begin{bmatrix} 0.27877 & -0.052138 & -0.64721 \\ -0.052138 & 1.1778 & -0.85108 \\ -0.64721 & -0.85108 & 2.3593 \end{bmatrix} \begin{bmatrix} 1.6990 \\ 1.0212 \\ 1 \end{bmatrix} = 0.27388\ .$$

Now, using equation 1 with $s^2$ equal to $(0.17)^2$ or 0.0289 from table 3 and a critical value for the t-distribution of 1.0012 from table 4, the prediction interval is calculated from equation 1:

$$\log Q_{100>32\text{-UP}} = \log\ (26{,}600) + 1.0012\ [0.0289\ (1 + 0.27388)]^{0.5} = 4.6170 \text{ log-ft}^3\text{/s}$$

$$Q_{100>32\text{-UP}} = 41{,}400 \text{ ft}^3\text{/s}.$$

$$\log Q_{100>32\text{-LOW}} = \log\ (26{,}600) - 1.0012\ [0.0289\ (1 + 0.27388)]^{0.5} = 4.2328 \text{ log-ft}^3\text{/s}$$

$$Q_{100>32\text{-LOW}} = 17{,}100 \text{ ft}^3\text{/s}.$$

There is a 67-percent chance that this prediction interval contains the true 100-year peak discharge for this site and equation, and the true discharge falls within the interval 17,100 to 41,400 $ft^3/s$. However, because this site lies within the CDA range of 10 to 100 $mi^2$, and two separate equations are considered applicable; the weighted prediction intervals are needed (eqs. 4 and 5). The 67-percent prediction interval for the 100-year peak discharge for a site with less than 32 $mi^2$ are (eq. 1):

$$\log Q_{100<32\text{-UP}} = \log\ (16{,}800) + 1.0153\ [0.0841\ (1 + 0.32152)]^{0.5} = 4.5638 \text{ log-ft}^3\text{/s}$$

$$Q_{100<32\text{-UP}} = 36{,}600 \text{ ft}^3\text{/s}.$$

$$\log Q_{100<32\text{-LOW}} = \log\ (16{,}800) - 1.0153\ [0.0841\ (1 + 0.32152)]^{0.5} = 3.8868 \text{ log-ft}^3\text{/s}$$

$$Q_{100<32\text{-LOW}} = 7{,}710 \text{ ft}^3\text{/s}.$$

The weighted prediction interval is calculated using the weighting equations 4 and 5:

$$Q_{100W\text{-}UP} = [2 - \log (50)]\,(36{,}600) + [\log (50) - 1]\,(41{,}400) = 0.30\,(36{,}600) + 0.70\,(41{,}400)$$

$$Q_{100W\text{-}UP} = 39{,}960 \text{ ft}^3/\text{s}.$$

$$Q_{100W\text{-}LOW} = [2 - \log (50)]\,(7{,}710) + [\log (50)-1]\,(17{,}100) = 0.30\,(7{,}710) + 0.70\,(17{,}100)$$

$$Q_{100W\text{-}LOW} = 14{,}300 \text{ ft}^3/\text{s}.$$

The prediction interval constructed for a prediction from any one regression equation can be statistically interpreted. However, the weighted prediction interval is only estimated, and therefore strict statistical interpretation of it requires relaxation.

A check on the leverage statistics on the large basin equation indicates that the site leverage (0.27) is less than the maximum leverage (0.38) and less than 2 to 3 times the mean leverage (0.28 to 0.42). A check on the leverage statistics on the small basin equation indicates that the site leverage (0.32) is less than the maximum leverage (0.54) and less than 2 to 3 times the mean leverage (0.48 to 0.72). The two site leverages (0.27 and 0.32) cannot be compared to each other to determine which equation is the most applicable for a site. Both sets of equations are considered applicable because both site leverages are sufficiently less than the maximum leverage and less than 2 to 3 times the mean leverage of their respective equations.