User's Manual for Program PeakFQ, Annual Flood-Frequency Analysis Using Bulletin 17B Guidelines

Chapter 4 of Book 4, Section B

Techniques and Methods 4-B4

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

***Cover*: The Sorlie Bridge between Grand Forks, North Dakota, and East Grand Forks, Minnesota, during the 1997 Red River of the North flood (photograph by Steven W. Norbeck, U.S. Geological Survey).**

User’s Manual for Program PeakFQ, Annual Flood-Frequency Analysis Using Bulletin 17B Guidelines

By Kathleen M. Flynn, William H. Kirby, and Paul R. Hummel

Chapter 4 of Book 4, Section B

Techniques and Methods 4-B4

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

U.S. Department of the Interior
DIRK KEMPTHORNE, Secretary

U.S. Geological Survey
Mark D. Myers, Director

U.S. Geological Survey, Reston, Virginia: 2006

For product and ordering information:
World Wide Web: http://www.usgs.gov/pubprod
Telephone: 1-888-ASK-USGS

For more information on the USGS--the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment:
World Wide Web: http://www.usgs.gov
Telephone: 1-888-ASK-USGS

Suggested citation:
Flynn, K.M., Kirby, W.H., and Hummel, P.R., 2006, User's Manual for Program PeakFQ Annual Flood-Frequency Analysis Using Bulletin 17B Guidelines: U.S. Geological Survey, Techniques and Methods Book 4, Chapter B4; 42 pgs.

Contents

Figures

Tables

SYMBOLS

Symbol, explanation:

a, a constant characteristic of a particular plotting position
$\overline{G}$, generalized skew coefficient.
$\tilde{G}$, historically-adjusted station skew coefficient.
$\tilde{G}'$, skew coefficient of frequency curve passing through $\tilde{Q}^{*}0.50$. $\tilde{Q}^{*}0.10$, and $\tilde{Q}^{*}0.01$.
$|G|$, absolute value of the station-skew coefficient.
G, station skew coefficient.
G_W, weighted-skew coefficient estimate used in final log-Pearson Type III frequency curve.
g, desired skew coefficient.
H, historical period length.
K, confidence coefficient.
K_N, 10-percent significance-level critical value for outlier test statistic for sample of size N from the normal distribution.
K_H, 10-percent significance-level critical value for the outlier test statistic for sample of size H, where H is the length of the historic record period.
$k_{\tilde{G}',p}$, ordinates for skew $\tilde{G}'$ and exceedance probability (p) for curve passing through $\tilde{Q}^{*}0.50$. $\tilde{Q}^{*}0.10$, and $\tilde{Q}^{*}0.01$.
$k_{g,p}$, Pearson Type III standardized ordinates for desired skew (g) and exceedance probability (p).
k_p, standard normal frequency factor for probability p.
k'_p, frequency factor for expected probability frequency curve.
$k_{(1-\alpha)}$, standard normal deviate with exceedance probability 1-α.
$k_{\gamma,p}$, the Pearson Type III frequency factor.
$\tilde{M}$, historically-weighted logarithmic mean.
$\hat{M}$, Bulletin 17B mean.
$\tilde{M}'$, mean of frequency curve passing through $\tilde{Q}^{*}0.50$. $\tilde{Q}^{*}0.10$, and $\tilde{Q}^{*}0.01$.
$\tilde{m}$, historically-weighted rank of the m^{th} largest observed peak.
m, rank of the m^{th} largest observed peak.
MSE, mean-square error (standard error of estimate squared).
$MSE_{\overline{G}}$, mean-square error of generalized skew coefficient.
MSE_G, mean-square error of station skew coefficient.
$\tilde{N}$, effective number of peaks above flood base, Q_O.
N_{BB}, number of peaks below the flood base, including any zeros and low outliers.
N_{HO}, number of high outliers.
N_{HP}, number of historic peaks.
N_S, number of systematic peaks.
N_X, number of peaks between Q_O and Q_H.
N, record length in years.
n, sample size from normal population of flood logarithms.
$\tilde{P}_m$, historically-weighted probability plotting position of the m^{th} ranked observed peak.
$\tilde{P}_O$, estimated probability of a flood exceeding the flood base.
p, exceedence probability.
p', normal exceedance probability corresponding to k'_p.
$\tilde{Q}$, conditional frequency curve describing only those peaks above the flood base.
$\tilde{Q}^{*}$, intermediate unconditional frequency curve.
$\tilde{Q}^{*}_p$, ordinates of the unconditional curve.
$\hat{Q}_p$, final Bulletin 17B-estimated frequency curve.
Q_H, historical threshold streamflow.
Q_O, flood base streamflow.

$\hat{Q}_{s,p}$, systematic frequency curve ordinate at exceedance probability p.
RMSE, root mean square error.
S, sample logarithmic standard deviation.
$\hat{S}$, Bulletin 17B standard deviation.
$\tilde{S}$, historically-weighted logarithmic standard deviation.
$\tilde{S}'$, standard deviation of frequency curve passing through $\tilde{Q}^{*}0.50$. $\tilde{Q}^{*}0.10$, and $\tilde{Q}^{*}0.01$.
t_{n-1}, Student's t random variate with n-1 degrees of freedom.
$t_{n-1,p}$, Student's t quantile with n-1 degrees of freedom and exceedance probability p.
W, weight given to systematic peaks below historical threshold.
X', logarithmic magnitudes of historic peaks and high outliers.
$\bar{X}$, sample logarithmic mean.
X, logarithmic magnitudes of systematic peaks between Q_O and Q_H.
X_H, logarithmic high-outlier test threshold.
X_L, logarithmic low-outlier test threshold.
α, confidence level.
γ, population skew coefficient.
μ, population mean.
σ, population standard deviation.
$>$, less than.
$<$, greater than.
$\leq$, greater than or equal to.

Note: Most symbols and explanations from Interagency Advisory Committee on Water Data (1982) and Lepkin and others (1979).

User's Manual for Program PeakFQ, Annual Flood-Frequency Analysis Using Bulletin 17B Guidelines

By Kathleen M. Flynn[1], William H. Kirby[1], and Paul R. Hummel[2]

Abstract

Estimates of flood flows having given recurrence intervals or probabilities of exceedance are needed for design of hydraulic structures and floodplain management. Program PeakFQ provides estimates of instantaneous annual-maximum peak flows having recurrence intervals of 2, 5, 10, 25, 50, 100, 200, and 500 years (annual-exceedance probabilities of 0.50, 0.20, 0.10, 0.04, 0.02, 0.01, 0.005, and 0.002, respectively). As implemented in program PeakFQ, the Pearson Type III frequency distribution is fit to the logarithms of instantaneous annual peak flows following Bulletin 17B guidelines of the Interagency Advisory Committee on Water Data. The parameters of the Pearson Type III frequency curve are estimated by the logarithmic sample moments (mean, standard deviation, and coefficient of skewness), with adjustments for low outliers, high outliers, historic peaks, and generalized skew. This documentation provides an overview of the computational procedures in program PeakFQ, provides a description of the program menus, and provides an example of the output from the program.

Introduction

Program PeakFQ performs statistical flood-frequency analyses of annual-maximum peak flows (annual peaks) following procedures recommended in Bulletin 17B of the Interagency Advisory Committee on Water Data (1982), referred to hereinafter as Bulletin 17B. The following sections document the implementation of the Bulletin 17B guidelines in program PeakFQ. This information is intended to assist the user with selection of program options and interpretation of the program output. Program users should refer to Bulletin 17B for the complete and definitive description of the recommended procedures.

The Bulletin 17B procedures treat the occurrence of flooding at a site as a sequence of annual random events or trials. The magnitudes of the annual events are assumed to be independent random variables following a log-Pearson Type III probability distribution; that is, the logarithms of the annual peak flows are assumed to follow a Pearson Type III distribution. This distribution defines the probability that any single annual peak will exceed a specified streamflow. Given this annual exceedance probability, other probabilities, such as the probability that a future design period will be free of exceedances, can be calculated by standard methods, as described in Appendix 10 of Bulletin 17B. Program PeakFQ estimates the parameters of the log-Pearson Type III frequency distribution from the logarithmic sample moments (mean, standard deviation, and coefficient of skewness) of the record of annual flows, with adjustments for low outliers, high outliers, historic peaks, and generalized peak skew. The parameter values are used to calculate the percentage points (or quantiles) of the log-Pearson Type III distribution for selected exceedance probabilities.

Peak Streamflow Records

The U.S. Geological Survey maintains a peak-flow file in the National Water Information System (NWIS) data base. The contents and format of data retrieved from the peak-flow file are described in Appendices B.2 (Station Header Records) and B.4 (Peak-Flow Records). Program PeakFQ uses the station identification number and name to label the printed output and may use the station latitude and longitude to look up the generalized skew. The Bulletin 17B statistical computations use only the annual-peak discharge and discharge-qualification codes from the peak-flow records; the gage-height information and all information about partial-duration or secondary peaks is ignored.

The annual peak-flow data fall into two classes: systematic and historic. The systematic record includes all annual peaks observed in the course of one or more systematic gaging programs at the site. In a systematic gaging program, the annual peak is observed (or estimated) for each year of the program. Several systematic records at one site can be combined, provided that the hydrologic conditions during the periods of record are comparable. The gaps between distinct, systematic-record periods can be ignored, provided that the lack of record in the interim was unrelated to the hydrologic conditions. Thus, if a flood record was interrupted for lack of funds for data collection, the interruption could be ignored and the available data could be used as if no interruption had occurred. On the other hand, if the record was interrupted because of prolonged drought or excessive flooding, the interruption should not be ignored but rather should be used, if possible, as evidence for adding one or more estimated peaks to the systematic record. Thus, the systematic record is intended

[1] U.S. Geological Survey

[2] Aqua Terra Consultants

to constitute an unbiased and representative sample of the population of all possible annual peaks at the site.

In addition to the systematic record, some stations have a historic record consisting of generally isolated high-magnitude peaks that occurred outside the period of systematic data collection. In contrast to the systematic record, the historic record consists of annual peaks that would not have been observed except for a recognition that an unusually large peak had occurred. Flood information acquired from old newspaper articles, letters, personal recollections, and other historical sources almost invariably refers to floods of noteworthy, and hence extraordinary, size. Similarly, paleoflood information, determined by analysis of geological or botanical evidence, is considered historic information and almost always refers to extraordinary floods. Thus, historic records, by the conditions of their collection, form a biased and unrepresentative sample of flood experience. Despite this bias, however, the historic record can be used to supplement the systematic record provided that all historic peaks above some historic threshold have been recorded.

The systematic record also may contain one or more large-magnitude peaks for which historic information is available or which exceed some historic peaks. Such peaks are called high outliers if they are greater than the high outlier threshold. They are not considered historic peaks because they are part of the systematic record. In particular, the peak of record is not considered a historic peak if it occurred during a period of systematic data collection. Although high outliers are part of the systematic record, they also are treated in the same way as historic peaks in the historic-record adjustment procedure.

Qualification codes are assigned to some peaks to identify (1) basin or environmental conditions that may have affected the magnitude of the streamflow, (2) measurement conditions that may have affected the accuracy of the recorded value and (3) historic peaks. These codes are described in Appendix B.4 and also in the NWIS-Web Help System at URL -- *http://nwis.waterdata.usgs.gov/nwis/help?codes_help#flow_qual_cd*. Note that an individual peak flow can have more than one qualification code associated with it. Program PeakFQ recognizes several of these codes and uses them to control the statistical computation. For example, discharge code 4 (discharge less than indicated value, reported as the mnemonic letter code L in the PeakFQ output file) automatically triggers the zero-flow and conditional-probability adjustments. Table 1 identifies the peak-flow file qualification codes used by program PeakFQ, explains how these codes are interpreted by the program, and briefly describes how the PeakFQ program handles the associated peaks.

Principles of Computations

The Bulletin 17B computational analysis is illustrated in figure 1. The following sections provide an overview of the major computational steps.

Systematic-Record Analysis

The systematic-record analysis involves the computation of the mean, standard deviation, and coefficient of skewness ($\overline{X}$, S, and G, respectively) of the common logarithms of the annual peak flows in the systematic record. At some sites, annual peaks of magnitude zero can occur; more generally, the annual peak may occasionally fall below or be equal to

Table 1. Peak-flow codes used by program PeakFQ

[NWIS--National Water Information System]

PeakFQ code	NWIS code	PeakFQ Interpretation	PeakFQ Action
D	3	Dam failure, non-recurrent flow anomaly	Peak always excluded.
G	8	Discharge greater than stated value	Peak always excluded.
X	3 and 8	Both D and G	Peak always excluded.
L	4	Discharge less then stated value	Conditional-probability adjustment
K	6 or C	Known effect of regulation, urbanization, or other watershed change	Peak excluded by default. Can be included by specifying "yes" in the "Urban/Reg Peaks" field of the PeakFQ station specifications.
H	7	Historic peak. (Note: Historic peaks are events that occur outside periods of systematic data collection. The peak of record is not a historic peak if it was observed as part of the systematic record collection. See text for additional details.)	Peak excluded by default. Can be included by specifying a value for historic period in the PeakFQ station specifications, in which case the historic adjustment will be applied.
-	1, 2, 5, 9, A, B, or E	Codes are not considered by PeakFQ	Peak always included.

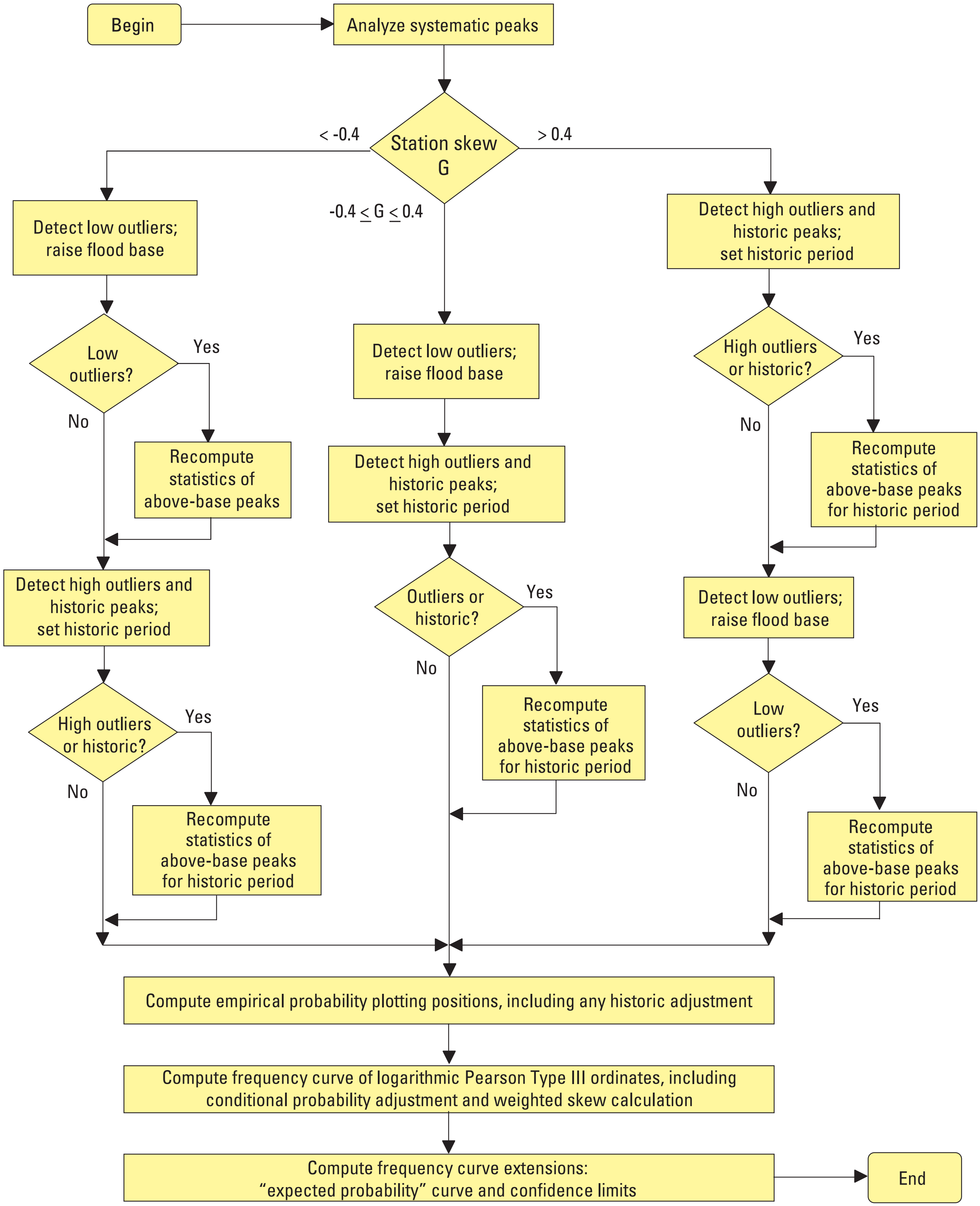


Figure 1. General flow chart for flood-frequency computations (modified from Interagency Advisory Committee on Water Data, 1982).

some lower limit of measurement called the gage base (which usually is zero but may be greater than zero). To account for this possibility, the number of peaks below the gage base is computed in addition to the mean, standard deviation, and skewness of the logarithms of the above-base systematic peaks. The statistics of the above-base systematic peaks and the number of peaks below the gage base are used to compute the systematic-record frequency curve. If there are no zeroes or below-base peaks, the systematic-record frequency curve is computed as follows:

$$\log \hat{Q}_{s,p} = \overline{X} + S k_{g,p}, \qquad (1)$$

where $\hat{Q}_{s,p}$ = systematic frequency-curve ordinate at exceedance probability p, and
$k_{g,p}$ = the Pearson Type III standardized ordinates for station skew g and exceedance probability p.

If there are zeroes or below-base peaks, the statistics of the above-base systematic peaks are used to define a conditional above-base systematic-record frequency curve, which is then adjusted by the conditional-probability adjustment, as described in a subsequent section and in Bulletin 17B, Appendix 5.

The systematic-record frequency curve is an initial estimate of the Bulletin 17B frequency curve. This initial estimate is adjusted to account for historic data, high and low outliers, and regional (generalized) skew information.

Outlier Tests

Peaks that depart substantially from the trend of the remaining peaks are outliers. The first adjustments of the initial frequency curve involve detecting and accounting for high and low outliers. The sequence of these tests and adjustments depends on the station skew coefficient, G, computed in the first step. Because a relatively large skew coefficient of either sign ($G > +0.4$ or $G < -0.4$) is likely to be caused by an outlier on the corresponding end of the sample, this possibility is checked first and any necessary adjustment is applied before checking for outliers on the other end. If the skew coefficient is of moderate size ($-0.4 \leq G \leq +0.4$), the existence of both high and low outliers can be checked before applying any adjustments.

Program PeakFQ tests for high outliers using the following equation:

$$X_H = \overline{X} + K_N S, \qquad (2)$$

where X_H = Logarithmic high-outlier test threshold, and
K_N = 10-percent significance-level critical value for outlier test statistic for samples of size N from the normal distribution. (See Bulletin 17 B Appendix 4.)

Program PeakFQ does not automatically use the high-outlier test threshold in the analysis. Flood peaks considered high outliers should be evaluated in the context of the flood record at the site and at nearby sites. If the records indicate a high outlier(s) is the maximum in an extended period of time, the outlier(s) should be treated as historic data. For this case, the program requires the user to specify the high-outlier threshold and length of historic period in order for the high-outlier and historic-peak adjustment to be applied. The computations for performing the adjustment are described in the next section. If the record does not contain sufficient information to adjust for high outliers, they should be retained as part of the systematic record. For this case, no values are specified for the high-outlier threshold and length of historic period.

Program PeakFQ tests for low outliers using the following equation:

$$X_L = \overline{X} - K_N S, \qquad (3)$$

where X_L = logarithmic low-outlier test threshold.

If an adjustment for historic data has previously been made, then the following equation is used to detect low outliers:

$$X_L = \tilde{M} - K_H \tilde{S}, \qquad (4)$$

where
$\tilde{M}$ = historically-weighted logarithmic mean,
K_H = 10-percent significance-level critical value for outlier test statistic for sample of size H, where H is the length of the historic record period, and
$\tilde{S}$ = historically-weighted logarithmic standard deviation.

The computation of $\tilde{M}$ and $\tilde{S}$ is described in the next section.

The frequency curve is automatically adjusted for the effect of low outliers using the conditional probability adjustment described later.

Historic-Record Adjustment

Recalculation of the statistics of the above-base peaks is required after the detection of outliers or historic information, as specified in Appendix 6 of Bulletin 17B. The logical basis for the calculation is the following:

Historic-adjustment criterion: It is assumed that every annual peak that exceeded some historic threshold streamflow (Q_H) during the historic period (H) has been recorded as either a historic peak or a systematic peak (high outlier). In other words, the record is complete for peaks above Q_H during the time period H.

The historic period H includes the systematic-record period plus one or more years that have no systematic record. This criterion implies that the unrecorded portion of the historic period contains only peaks below the threshold Q_H. Figure 2 presents a definition sketch showing the time periods and streamflows used in the historic-record adjustment.

The Bulletin 17B historic adjustment, in effect, fills in the ungaged portion of the historic period with an appropriate number of replications of the below-Q_H portion of the systematic record. The adjustment is accomplished by weighting the below-threshold systematic peaks in proportion to the number of the below-threshold years in the historic period, as illustrated in figure 2, as follows:

$$W = \frac{H - N_{HP} - N_{HO}}{N_S - N_{HO}}, \quad (5)$$

where W is the weight to be applied to the below-threshold systematic peaks and N_S, N_{HP}, and N_{HO} are the numbers of systematic peaks, historic peaks, and high outliers, respectively. Then the effective number of peaks, $\tilde{N}$, above the flood base (Q_O) is

$$\tilde{N} = N_{HP} + N_{HO} + W(N_S - N_{HO} - N_{BB}) = H - W(N_{BB}), \quad (6)$$

where N_{BB} is the number of peaks below the flood base, including any zeros and low outliers.

The corresponding estimated probability of a flood exceeding the flood base is

$$\tilde{P}_O = \frac{\tilde{N}}{H}. \quad (7)$$

Applying the historic weight W to those systematic peaks below the historic threshold Q_H (and above the flood base Q_O) yields the following formulas for the historically-weighted mean ($\tilde{M}$), standard deviation ($\tilde{S}$), and skewness ($\tilde{G}$):

$$\tilde{M} = \frac{(W\sum X + \sum X')}{\tilde{N}}, \quad (8)$$

$$\tilde{S} = \sqrt{\frac{\left[W\sum(X-\tilde{M})^2 + \sum(X'-\tilde{M})^2\right]}{(\tilde{N}-1)}}, \quad (9)$$

$$\tilde{G} = \sqrt{\frac{\left[W\sum(X-\tilde{M})^3 + \sum(X'-\tilde{M})^3\right]\tilde{N}}{(\tilde{N}-1)(\tilde{N}-2)\tilde{S}^3}}, \text{ and} \quad (10)$$

in which X′ denotes logarithmic magnitudes of historic peaks and high outliers and X denotes logarithmic magnitudes of systematic peaks between the flood base Q_O and the historic threshold Q_H. These formulas are equivalent to those given in Appendix 6 of Bulletin 17B.

These formulas remain correct even if there is no historic information (in which case $H = N_S$), no high or low outliers, and no below-gage base peaks. Thus, these formulas are used in PeakFQ to calculate the Bulletin 17B statistics in all cases, including the calculation of the unadjusted systematic record statistics.

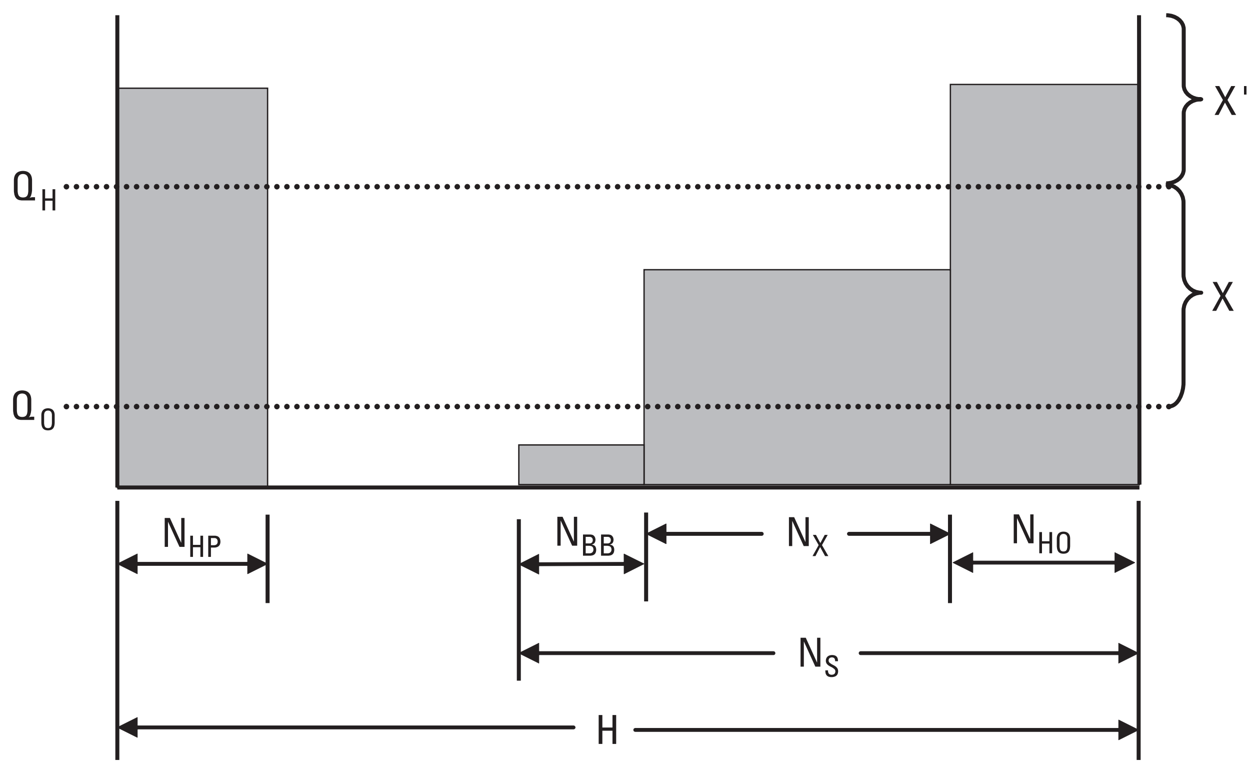


Figure 2. Definition sketch showing time periods and discharges used in historic record adjustment.

Conditional-Probability Adjustment

After the peak-streamflow frequency curve parameters have been determined, the historically weighted frequency curve can be tabulated. If no low outliers, zero flows, or below-gage base peaks are present, this process is simply a matter of looking up the Pearson Type III standardized ordinates, $k_{g,p}$, for the desired skew coefficient (g) and probability (p) and computing the logarithmic frequency curve ordinates by the formula:

$$\log\hat{Q}_p = \tilde{M} + \tilde{S}k_{g,p}. \quad (11)$$

When peaks below the flood base are present, however, the above calculation determines a conditional frequency curve $\tilde{Q}$ describing only those peaks above the base. To account for the fraction of the population below the flood base, the following argument is used: the probability that an annual peak will exceed a streamflow x (above the flood base) is the probability that the peak will exceed the base at all, multiplied by the conditional probability that the peak will exceed x, given that the peak exceeds the base. The first of these factors is just the probability, $\tilde{P}_o$, calculated in equation (7); the second factor is the probability on the conditional frequency curve at streamflow x. Thus the unconditional curve, $\tilde{Q}^*$, assigns a probability $[\tilde{P}_o]p$ to the streamflow having exceedance probability p on the above-base curve. Conversely, an exceedance probability p on the unconditional curve $\tilde{Q}^*$ corresponds to the probability $p/\tilde{P}_o$ on the original above-base curve $\tilde{Q}$. Thus, the ordinates of the unconditional curve can be computed directly by the formula:

$$\log\tilde{Q}^*_p = \tilde{M} + \tilde{S}k_{\tilde{G},(p/\tilde{P}_O)}, \quad (12)$$

in which $\tilde{M}$, $\tilde{S}$, and $\tilde{G}$ are the logarithmic mean, standard deviation, and skew coefficient, respectively, of the above-base distribution.

Because this distribution does not have the Pearson Type III shape, it is used only as an intermediate step in constructing an equivalent Pearson Type III curve. First, the three points $\tilde{Q}^*$ 0.50, $\tilde{Q}^*$0.10, and $\tilde{Q}^*$0.01 are computed using the above formula. Then a logarithmic-Pearson Type III curve is passed through these three points; the curve mean, standard deviation, and skew coefficient, $\tilde{M}'$, $\tilde{S}'$, and $\tilde{G}'$, respectively, are found by solving the three simultaneous equations:

$$\tilde{M}' + \tilde{S}'k_{\tilde{G},p} = \log\tilde{Q}^*_p \quad \text{(for p = 0.50, 0.10, and 0.01)}. \quad (13)$$

An exact solution requires a laborious interpolation in the Pearson Type III tables; the Bulletin 17B guidelines present a direct formula based on a linear approximation. Note that $\tilde{M}'$, $\tilde{S}'$, and $\tilde{G}'$ represent the contributions of all the observed peaks, those below the base as well as those above, whereas $\tilde{M}$, $\tilde{S}$, and $\tilde{G}$ do not. The resulting unconditional frequency curve, when floods below the base have been detected, then is:

$$\log\hat{Q}_p = \tilde{M}' + \tilde{S}'k_{\tilde{G}',p}. \quad (14)$$

This defines only the part of the distribution above the flood base; the part below the flood base is not defined, and is of no practical importance.

These conditional-probability adjustments are used not only to construct the final Bulletin 17B frequency curve, but also to construct a systematic-record frequency curve that takes into account any zero flows or below-gage base peaks (but not low outliers).

Computation of Weighted Skew Coefficient

The station- (or sample-) skew coefficient, which reflects the average of the cubed deviations from the sample mean, is highly sensitive to the observations in both the upper and lower tails of the sample. As a result, the estimated station-skew coefficient and extreme-flood quantiles may be strongly affected by idiosyncrasies of the sample, and may be unrepresentative of long-term flood characteristics. To help counter this problem, Bulletin 17B uses a generalized skew, which is a skew coefficient representative of neighboring stations, as explained in a subsequent section.

The station skew and generalized skew are combined in a weighted average that is expected to be more accurate than either of its constituents. Guidelines for estimating generalized skew are given in Bulletin 17B and are summarized in a subsequent section of this manual. Program PeakFQ does not perform generalized skew estimation. Instead, program PeakFQ either looks up the generalized skew from a digitized copy of the map in Bulletin 17B or reads it from user-supplied input (see preceding section). The following paragraphs explain the weighted skew computation.

The station skew, the generalized skew, and the weighted skew are quantities that are estimated from flood records at and near the station under study. As such, they are subject to estimation errors. The error in each of the skew statistics is characterized by two properties, the expected value (bias) and the standard deviation (standard error), representing systematic errors and random-sampling variability, respectively. The random and systematic errors are combined in a single composite property called mean-square error (MSE), which is the expected value of the difference between the estimated and true values of the statistic (station, generalized, or weighted skew). The MSE is the sum of the squares of the bias and standard error. The MSE often is reported in terms of its square root, the Root Mean Square Error (RMSE), which is directly comparable to the quantity being estimated (rather than to its square) and can be expressed as a percentage. If the bias is small relative to the standard error, the RMSE is approximately equal to the standard error. Because of its wide availability and usefulness, the RMSE is used as input to program PeakFQ; the program squares the input RMSE to obtain the MSE values used in equation 15.

The station- and generalized-skew coefficients are combined in a weighted average to form a better estimate of the skew coefficient for a given watershed. Under the assumption that the generalized-skew coefficient is unbiased and independent of the station-skew coefficient, the MSE of the weighted-skew estimate is minimized by weighting the station- and generalized- skew coefficients in inverse proportion to their individual MSE's. This concept is expressed in equation 15, adapted from Tasker (1978), which is used in computing the weighted-skew coefficient:

$$G_W = \frac{MSE_{\overline{G}}(G) + MSE_G(\overline{G})}{MSE_{\overline{G}} + MSE_G} \qquad (15)$$

where G_w = weighted-skew coefficient,
G = station-skew coefficient,
$\overline{G}$ = generalized-skew coefficient,
$MSE_{\overline{G}}$ = mean-square error of generalized-skew coefficient, and
MSE_G = mean-square error of station-skew coefficient.

The MSE (or RMSE) of the generalized skew is estimated in conjunction with the development of the generalized-skew value. In program PeakFQ, if the user does not specify a value for the generalized-skew coefficient, the value is obtained from a digitized version of Plate 1 of Bulletin 17B, and the corresponding value of $MSE_{\overline{G}}$= 0.302 is used in equation 15. (The corresponding RMSE value is 0.55.) Otherwise, the user must supply the RMSE of the generalized skew as input data along with the user-supplied generalized-skew value.

The MSE of the station skew for log-Pearson Type III random variables is obtained from the results of Monte Carlo experiments by Wallis and others (1974). Their results show that the MSE of the logarithmic station skew is a function of record length and population skew. This function (MSE_G) is approximated with sufficient accuracy for use in calculating the weighted skew by the equation:

$$MSE_G = 10^{[A - B[\log_{10}(N/10)]]} \qquad (16)$$

where A = $-0.33 + 0.08\,|G|$ if $|G| \le 0.90$,
$-0.52 + 0.30\,|G|$ if $|G| > 0.90$,
B = $0.94 - 0.26\,|G|$ if $|G| \le 1.50$, and
0.55 if $|G| > 1.50$,

in which $|G|$ is the absolute value of the station-skew coefficient (used as an estimate of population-skew coefficient) and N is the record length in years. If the historic adjustment (Bulletin 17B, Appendix 6) has been applied, the historically-adjusted skew coefficient, $\tilde{G}$, and historic period, H, are used for G and N, respectively, in equation 16.

Bulletin 17B indicates that equations 15 and 16 may underestimate the weight given to the station skew if the station skew is large and the record is long, or if the magnitude differs from the generalized skew by more than 0.5. In these cases, Bulletin 17B suggests that the peak-flow data and the flood-producing characteristics of the basin be examined to determine whether greater weight should be given to the station skew.

Expected-Probability Adjustment

The final steps in the Bulletin 17B analysis, as implemented in program PeakFQ, are to compute the expected-probability frequency curve and a set of upper and lower confidence limits. These computations are optional and are intended primarily as an aid to the interpretation of the principal Bulletin 17B-estimated frequency curve given by $\hat{Q}p$ above.

The expected probability concept deals with the following problem. A sample of size n will be drawn from a normal population (of flood logarithms), and the flood having a specified exceedance probability p will be estimated by the quantity $\overline{X} + k_pS$, in which $\overline{X}$ and S are the ordinary sample mean and standard deviation, respectively, and k_p is the standard normal frequency factor for probability p. Because it is computed from a random sample, the estimate $\overline{X} + k_pS$ is a random variable, that usually will differ from the true p-probability flood. Thus, one is led to ask how the probability of another flood exceeding the estimate $\overline{X} + k_pS$ compares with the specified (nominal) probability p. For a normal population, one has:

$$P\left\{X > \overline{X} + k_pS\right\} = P\left\{\frac{X - \overline{X}}{S} > k_p\right\} = P\left\{t_{n-1} > k_p\left[\frac{n}{n+1}\right]^{0.5}\right\}, \qquad (17)$$

where t_{n-1} is Student's t random variate with n-1 degrees of freedom. This probability has come to be known as the "expected probability" (Beard, 1960; Bulletin 17B, Appendix 11). For nominal exceedance probabilities less than 0.50 (floods above the median), the expected probability exceeds the nominal probability. The expected probability can be made equal to the nominal probability by replacing k_p by the frequency factor:

$$k'_p = t_{n-1,p}\left[(n+1)/n\right]^{0.5}, \qquad (18)$$

in which $t_{n-1,p}$ is the Student's t value with n-1 degrees of freedom and exceedance probability p. The visible effect of this adjustment is to increase the slope of the estimated frequency curve in proportion to the statistical variability of the sample statistics.

This normal-population result is applied to the Bulletin 17B-estimated Pearson Type III distribution with mean, standard deviation, and skew coefficient, $\hat{M}$, $\hat{S}$, and G_W, by first looking up the normal exceedance probability p′ corresponding to k'_p and, second, applying the Pearson Type III frequency factor, $k_{G,p'}$ having skew coefficient and probability, to the sample mean and standard deviation, as follows:

$\hat{M} + \hat{S}\,(k_{G_W,p})$. Even this estimate, however, when evaluated for any particular sample, normally will misrepresent the true p-probability flood. With respect to a large number of flood records, however, the fraction of floods actually exceeding the estimated p-probability floods will be correct. Nonetheless, the Bulletin 17B guidelines specify that the basic flood-frequency curve (without expected probability) is the curve to be used for estimating flood risk and forming weighted averages of independent flood-frequency estimates.

Confidence Limits

Finally, one-sided confidence limits for the p-probability flood are computed. A one-sided confidence limit is a sample statistic—hence a random variable—having a specified probability (confidence level) of exceeding (or not exceeding) a specified population characteristic. In the Bulletin 17B analysis, these statistics are of the form $\overline{X}$ + KS, where $\overline{X}$ and S are the sample mean and standard deviation, respectively, after all Bulletin 17B tests and adjustments and K is a confidence coefficient chosen to satisfy the following equation:

$$P\left\{\overline{X} + KS > \mu + k_{\gamma,p}(\sigma)\right\} = \alpha, \qquad (19)$$

in which α is the confidence limit, μ, σ, and γ are the population mean, standard deviation, and skew coefficient, respectively, $k_{\gamma,p}$ is the Pearson Type III frequency factor, and the right-hand side of the inequality is the population p-probability flood. The population parameters are unknown, but constant. The idea is to find a K-value such that $\overline{X}$ + KS, which can be computed from the sample, and is a random variable, will have a high probability of being an upper (or lower) bound on the unknown population p-probability flood. In any particular sample the computed value $\overline{X}$ + KS may fail to bound the population characteristic, but, over a number of samples, the specified fraction, -α (or 1-α), will yield correct bounds. A value of α close to unity yields upper confidence limits and a value close to zero yields lower limits. In particular, the upper 95-percent confidence limit has α = 0.95; the lower 95-percent limit has α = 0.05. The value of K is found by rearranging the probability statement as follows:

$$P\left\{\frac{\left((\mu-\overline{X})/(\sigma/\sqrt{n})+\sqrt{n}(k_{\gamma,p})\right)}{(S/\sigma)} < \sqrt{n}(K)\right\} = \sigma \qquad (20)$$

in which n is the sample size. If the underlying population were normally distributed (γ = 0), and if $\overline{X}$ and S were the ordinary sample mean and standard deviation, respectively, then the random variable on the left-hand side of the inequality would have the noncentral t distribution with n-1 degrees of freedom and noncentrality parameter $\sqrt{n}\,(k_{\gamma,p})$. If the underlying population had a small skew, if the sample were large, and if the population skew coefficient, γ, could be replaced by the Bulletin 17B estimated skew coefficient, G_W, then one could assume that the variate would have approximately the noncentral t distribution. Building upon this foundation, one obtains:

$$\sqrt{n}(K) = t_{n-1},\sqrt{n}\left(k_{G_W,p}\right),(1-\alpha), \qquad (21)$$

which is the noncentral t value with exceedance probability l-α. A standard large-sample approximation for the noncentral t distribution then yields the result:

$$K = \frac{\left(k_{G_W,p}\right) + \dfrac{k_{(1-\alpha)}}{\sqrt{n}}\left\{1 + \dfrac{nk^2_{G_W,p} - k^2_{(1-\alpha)}}{2(n-1)}\right\}^{0.5}}{\dfrac{1-k^2_{(1-\alpha)}}{2(n-1)}}, \qquad (22)$$

in which $k_{(1-\alpha)}$ is the standard normal deviate with exceedance probability 1-α and G_W is the Bulletin 17B weighted-skew coefficient. As stated above, an α-value near unity yields upper confidence limits whereas a value near zero yields lower limits. This result is equivalent to that in the Bulletin 17B guidelines.

Probability-Plotting Positions

Probability plotting positions are estimates of the exceedance probabilities of observed peak flows. They are computed by the formula p = (m-a)/(N-2a+1) (equation 10 in Bulletin 17B), in which m is the rank of an observed peak (m = 1 for highest peak), N is the sample size, and a is a constant characteristic of a particular plotting-position formula. Bulletin 17B and Program PeakFQ use the Weibull plotting-position formula (a = 0) by default, although other a-values can be specified. The probability-plotting positions are not used in the Bulletin 17B computations, but are used in graphic displays of the observed data in relation to the fitted frequency curve.

If there is historical information, the probability-plotting positions are adjusted using the same logic that underlies the calculation of the historically-weighted mean, standard deviation, and skew coefficient. The actual sample of size N is augmented by (W-1) "virtual" copies of the observed peaks below the historic threshold to fill out the entire historic period (H). In the ranked record, each below-threshold observed peak is preceded by (W-1)/2 of its virtual copies and followed by the remaining (W-1)/2 copies. In the augmented ranked series, if there are Z peaks above the historic threshold, then the rank of the first below-threshold observed peak is Z + (W-1)/2 + 1. The rank of the second below-threshold observed peak is Z + W + (W-1)/2 + 1. In general, the historically adjusted rank $\tilde{m}$ of the m[th] ranked observed peak is:

$$\begin{aligned}\tilde{m} &= m && \text{if } m <= Z \\ &= Z + W\,((m-Z) - 1/2) + 1/2 && \text{if } m > Z\end{aligned} \qquad (23)$$

where Z $= N_{HO} + N_{HP}$.

The historically-weighted plotting positions $\widetilde{P}_m$ then are:

$$\widetilde{P}_m = (\widetilde{m} - a) / (H - 2a + 1). \qquad (24)$$

These equations are equivalent to equations 6-6 through 6-8 in Appendix 6 of Bulletin 17B. As indicated above, program PeakFQ uses the Weibull plotting-position formula (a = 0) by default, although other a-values can be specified.

Estimation of Generalized Skew

The skew of a frequency distribution has a great effect on the shape and thus the values of the distribution, particularly in the extreme tail, which is of most concern in flood-risk estimation. The skew coefficient of the station record (station skew coefficient, G) is sensitive to extreme events; thus it is difficult to obtain an accurate estimate of the skew coefficient from a small sample. The accuracy of the estimated skew coefficient can be improved by weighting the station-skew coefficient with a generalized-skew coefficient estimated by pooling information from nearby sites.

Program PEAKFQ does not perform generalized-skew estimation. The program either looks up the generalized skew in a digitized version of the map (Plate I) in Bulletin 17B or reads the generalized skew and its associated RMSE from user-supplied input. The estimation of the generalized skew is performed by the flood-frequency analyst.

The discussion in this section concerns Bulletin-17B guidelines for development of appropriate generalized-skew coefficients for flood-frequency analysis. The following discussion is modified from Bulletin 17B (p. 10-14).

Bulletin 17B includes a map (Plate I) showing generalized-skew values that may be used in the absence of detailed generalized-skew studies. This map and its corresponding MSE of 0.302 (RMSE = 0.550) were developed when Bulletin 17 was first introduced in 1976 and have not been updated.

Additional peak-flow records have become available since that time. Also, the procedures used to develop the map do not conform in all respects to Bulletin 17B. Generalized-skew estimates developed in accordance with Bulletin 17B procedures should preferably be used if available. Nonetheless, Plate I is still considered an alternative for use with Bulletin 17B for those who prefer not to develop their own generalized-skew estimates. Program PeakFQ contains a digitized version of this map, which is used if the user does not specify a generalized skew and RMS error.

The Bulletin-17B recommended procedure for developing generalized-skew coefficients requires the use of at least 40 stations, or all stations within a 100-mile radius. The stations used should have 25 or more years of record. It is recognized that in some locations, a relaxation of these criteria may be necessary. The actual procedure includes analysis by three methods: (1) skew isolines drawn on a map; (2) skew prediction equation; and (3) the mean skew coefficient from selected stations. Each of the methods is discussed separately.

To develop the isoline map, each station-skew coefficient is plotted at the centroid of its drainage basin and the plotted data are examined for any geographic or topographic trends. If a pattern is evident, then isolines are drawn and the average of the squared differences between observed and isoline values, MSE, is computed. The square root of the MSE (RMSE or RMS error) should be computed to permit a better appraisal of the expected magnitude of the discrepancy between the generalized and station skews relative to the absolute magnitude of the skews. The MSE or RMSE will be used in appraising the accuracy of the isoline map. If no pattern is evident, then an isoline map cannot be drawn and is, not considered further.

A prediction equation should be developed that would relate either the station-skew coefficients or the differences from the isoline map to predictor variables that affect the skew coefficient of the station record. These would include watershed and climatologic variables such as drainage area, channel slope, and precipitation characteristics. The prediction equation should be used preferably for estimating the skew coefficient at stations with variables that are within the range of data used to calibrate the equation. The MSE (or RMSE) should be computed as the average (or square root of the average) of the residuals between the observed station skews and the fitted relation. If the relation is fitted by linear regression, then the standard error of regression can be taken as equivalent to the RMSE. The MSE (or RMSE) will be used to evaluate the accuracy of the prediction equation.

Determine the arithmetic mean and variance (or standard deviation) of the skew coefficients for all stations. In some cases, the variability of the runoff regime may be so large as to preclude obtaining 40 stations with reasonably homogeneous hydrology. In these situations, the arithmetic mean and variance of about 20 stations may be used to estimate the generalized-skew coefficient. The drainage areas and meteorologic, topographic, and geologic characteristics should be representative of the region around the station of interest. The variance (or standard deviation) is taken as comparable to the MSE (or RMSE) and is used to appraise the accuracy of the mean value as a prediction of the skew.

Select the method that provides the most accurate estimate of the skew coefficient. Compare the MSE from the isoline map to the MSE for the prediction equation. The smaller MSE should then be compared to the variance of the data. If the MSE is significantly smaller than the variance, the method with the smaller MSE should be used and that MSE used in equation 15 to predict the weighted skew coefficient. If the smaller MSE is not significantly smaller than the variance, neither the isoline map nor the prediction equation provides a significantly more accurate estimate of the skew coefficient than the mean value. In this case, the mean skew coefficient should be used because it provides as accurate an estimate as the more complicated alternatives; the variance should be used in equation 15 for the MSE of the generalized skew $MSE_{\bar{G}}$.

The accuracy of a regional generalized skew relations is generally not comparable to the accuracy of Plate I in Bulletin 17B. Whereas the average accuracy of Plate I is given, the accuracies of subregions within the United States are not given. A comparison should be made only between relations that cover approximately the same geographical area.

Computer Program

The following sections describe the computer program PeakFQ for performing the Bulletin 17B flood-frequency analysis. There are two methods for running PeakFQ: a Windows version (called PKFQWin) and a batch version (called PKFQBat).

Windows Version (PKFQWin)

The program PKFQWin provides a user interface to the PeakFQ batch model. The opening screen of the program is shown below in figure 3.

When first opened, most of the interface is disabled. The interface is designed to follow a logical procession toward running PeakFQ: Use the File:Open menu item to open a PeakFQ data file. View/edit the Station Specifications that are populated by the data. View/edit the Output Options. Click the **Run PEAKFQ** button and View Results. Click the Save Specs button to store a desired set of specifications for future use.

File Open

The **File:Open** menu item is used to open any of the file types that can be used by PeakFQ. These include:

- Standard WATSTORE format Peak Flow data, as described in Appendices B.2, B.3, and B.4. Data in this format can be found on the NWIS-Web at *http://nwis.waterdata.usgs.gov/usa/nwis/peak*.

- Watershed Data Management (WDM) files, as described in Appendix C.

- PeakFQ Specification (PSF) files, as described in Appendix B.1.

Selecting the **File:Open** menu item opens the Open PeakFQ File dialogue. As shown in figure 4, this dialogue can open any of the three file types discussed above. After opening a file, the Station Specifications tab will be populated based on the contents of the file. Initial station specification values are derived from different sources for the three file formats: WATSTORE - station header, station option, and peak-flows records; WDM - data and attributes from each station's data set; PSF - data file (WATSTORE or WDM) plus specifications for overriding initial values.

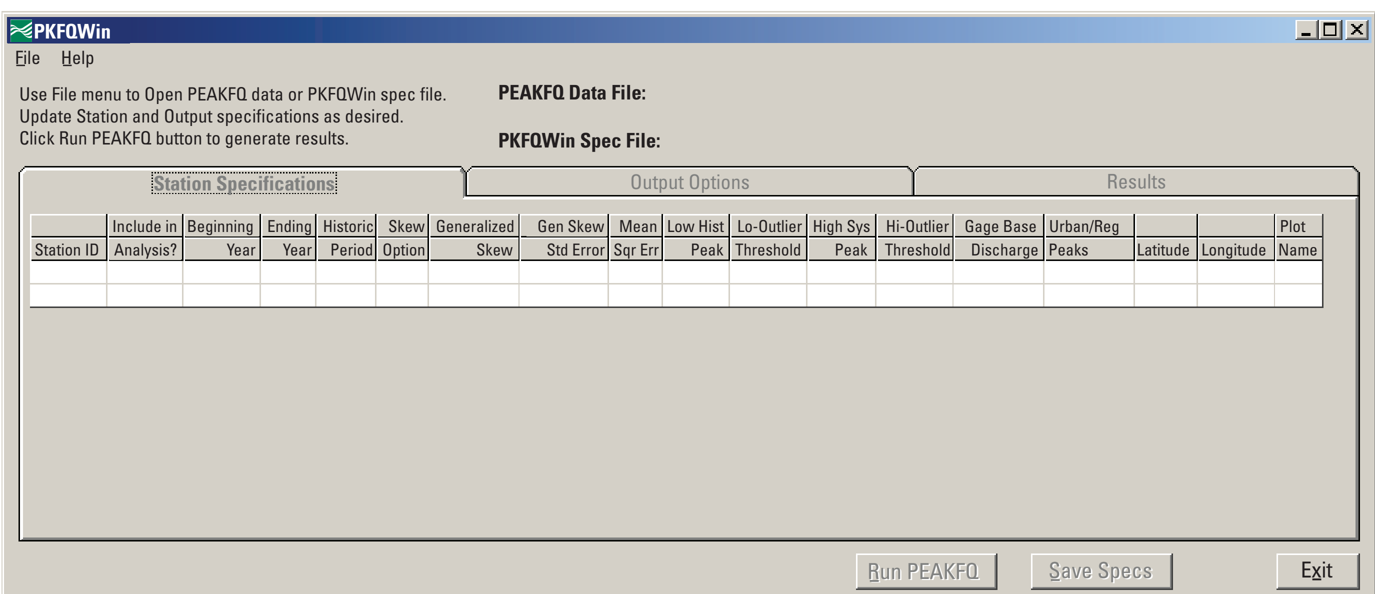


Figure 3. Example of opening screen of program PKFQWin showing the station specifications tab before an input file has been opened.

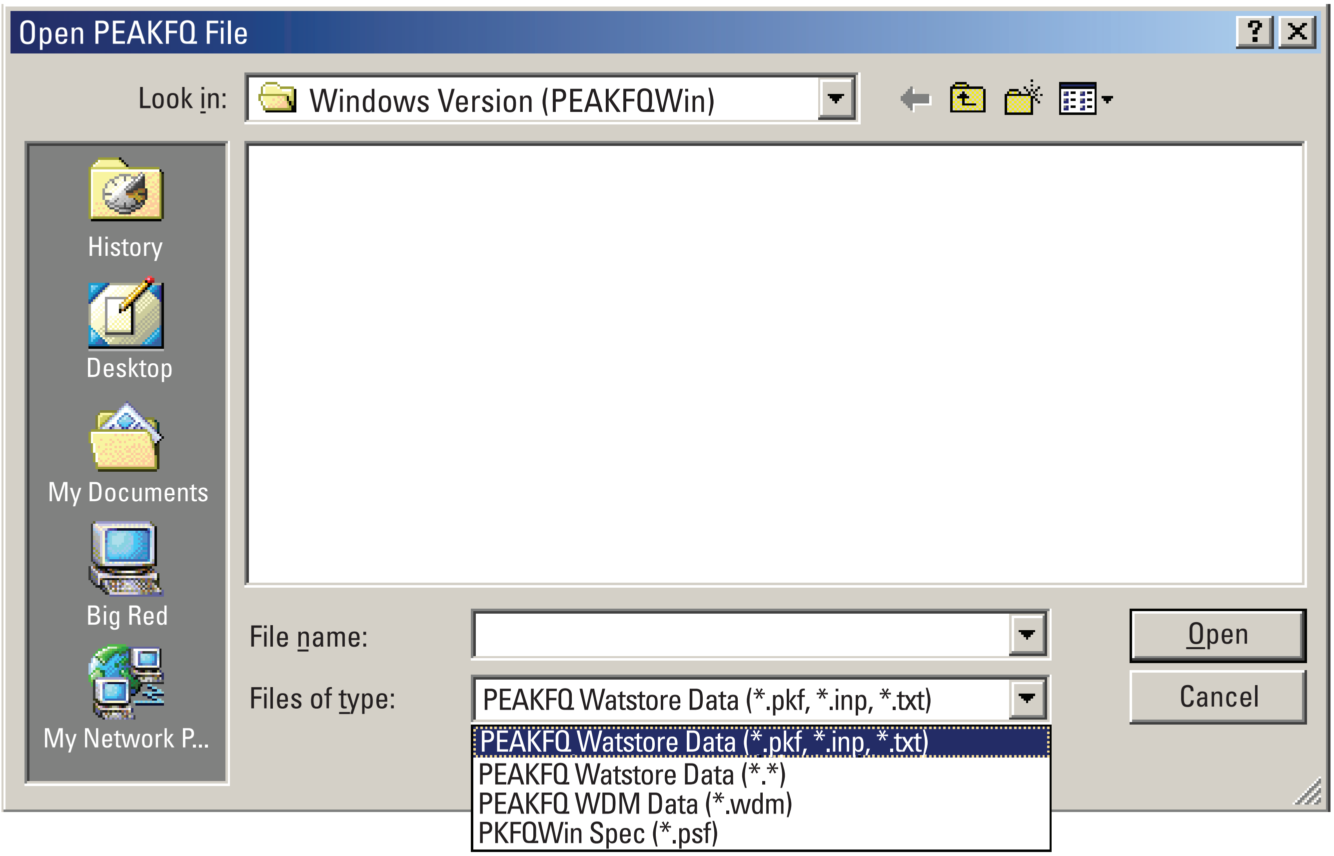


Figure 4. Example of the File Open window in program PKFQWin, obtained by selecting Open from the File menu.

Station Specifications

Once the selected data file has been opened and read, the **Station Specifications** tab is populated.

The default values on the **Station Specifications** tab are determined by the contents of the input file, including any WATSTORE "I" records or WDM attributes that may be present. The default values may be further updated if a PSF file was opened that contains station specifications for overriding defaults. The example in figure 5 shows many of the different options for the various fields. In particular, note how the same station may be run multiple times with different options between the runs (see Station IDs 03606500, 06600500, and 11274500 in figure 5).

By default, all stations found on the data file will be included in the analysis. If a station is not to be analyzed, the **Include in Analysis**? field may be changed to no by either typing "No" or double-clicking to activate the pull-down menu.

By default, the **Beginning Year** and **Ending Year** fields contain the water years of the first and last peaks, respectively, in the input file for the station. If a WATSTORE "I" record or WDM attribute is present, positive values in the **Beginning Year** and (or) **Ending Year** fields will be provided as the defaults. This will determine the period of record to be used in the calculations. These fields may be updated interactively by clicking in the desired station row and typing the new value.

The **Historic Period** field displays the value of any user-specified historic period that may have been present in a WATSTORE "I" record or WDM attribute. If no user-specified value has been given, then a value of zero (0) is displayed and no historic adjustment is applied during computation, the historic peaks are ignored, and any high outliers are treated as normal systematic peaks. If positive, the historic period contains the systematic record as a subset and historic adjustment will be applied during the computation. This field may be updated interactively by clicking in the desired station row and typing the new value.

The **Skew Option** is, by default, "Weighted" between "Station" and "Generalized" skews (WTD, Bulletin 17B weighted skew.) If a WATSTORE "I" record or WDM attribute is present with a station option code of S or G, the default skew option code will be "Station" or "Generalized." The three options are available for selection in a pull-down menu in the **Skew Option** field.

The **Generalized Skew** is, by default, based on the station latitude and longitude using the generalized-skew map accompanying Bulletin 17B. If a WATSTORE "I" record or WDM attribtute is present and the generalized skew field is non-blank, that value will be provided as the default. This field may be updated interactively by clicking in the desired station row and typing the new value.

The **Gen Skew Std Error** field is, by default, 0.55, corresponding to the standard error of the generalized-skew map accompanying Bulletin 17B. If a WATSTORE "I" record or WDM attribute is present and the standard error of generalized-skew field is non-blank, that value will be provided as the default. This field may be updated interactively by clicking in the desired station row and typing the new value.

The **Low-Outlier-Threshold** field displays the value of any user-specified low-outlier threshold that may have been present in a WATSTORE "I" record or WDM attribute. If no user-specified value has been given, then a value of 0.0 is displayed, and the low-outlier threshold is computed by PeakFQ using the Bulletin 17-B low outlier test. Any input peaks less than the low-outlier threshold are accounted for by the conditional-probability adjustment. Occasionally, it may be necessary or appropriate to override the Bulletin-17-B low-outlier test if, for example, the test criterion is very close to one of the input peaks or if there are several very low peaks. The Low-Outlier-Threshold field may be updated interactively by clicking in the desired station row and typing the new value.

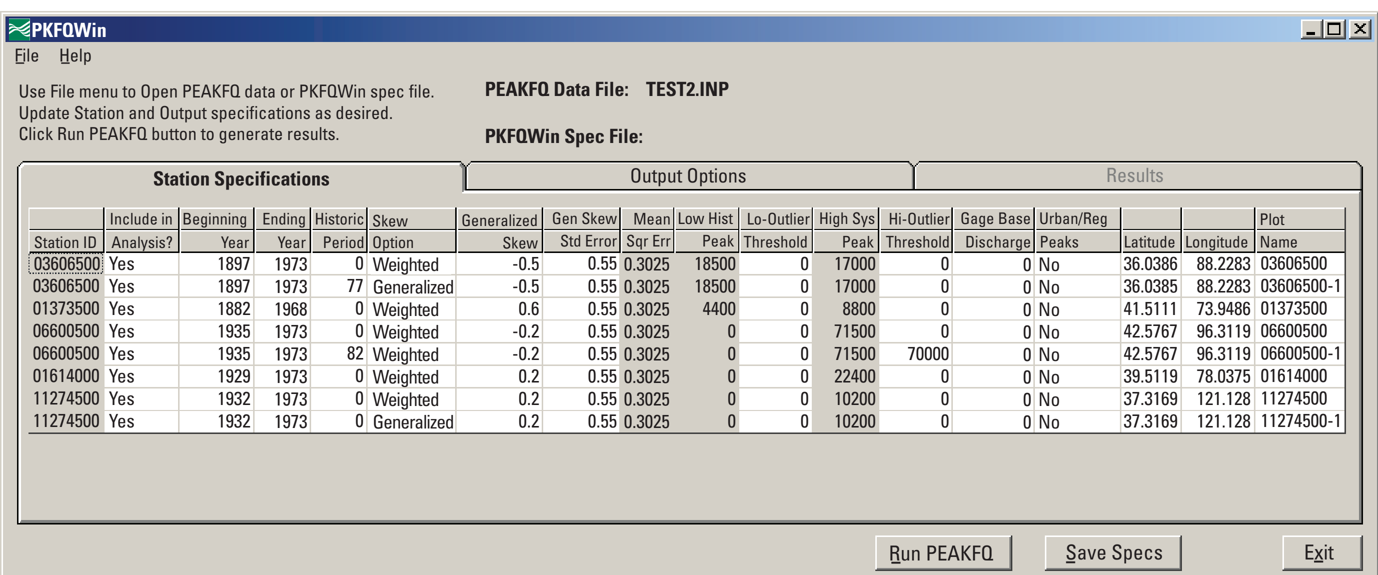
The **Hi-Outlier Threshold** field displays the value of any user-specified historic/high-outlier threshold that may have been present in a WATSTORE "I" record or WDM attribute. If no user-specified value has been given, then a value of 0.0 is displayed. This field is used only if the Bulletin-17-B historic-record adjustment has been specified by the user in the Historic-Period field. If a value greater than zero is displayed in the Hi-Outlier-Threshold field, that value will be used as the historic/high-outlier threshold in the Bulletin-17-B historic-record adjustment. Otherwise, the lowest historic-coded input peak will be used as the historic/high-outlier threshold. If there are no historic-coded peaks and a historic adjustment for a high outlier is needed, the user must specify the required high-outlier threshold. The Hi-Outlier-Threshold field may be updated interactively by clicking in the desired station row and typing the new value.

The **Gage Base Discharge** represents the lower limit of measureable flood peak at a station; this is zero (0) by default. If a WATSTORE "I" record or WDM attribute is present, a positive value in the field will be provided as the default. A negative or zero value will be ignored by the program. If positive, this gage-base discharge will supersede the gage base inferred from any "less than" NWIS qualification code (4) in the peak record. Note that this gage-base discharge is not the same as the partial-duration base discharge that may be in the station header "Y" record. This field may be updated interactively by clicking in the desired station row and typing the new value.

By default, **Urban** and (or) **Reg**ulated **Peaks** are not ("No") included in the computations. These peaks are indicated by a "6" or "C" in the NWIS qualification code field. If a WATSTORE "I" record or WDM attribute is present, this field will default to "Yes" if the Station option field contains a "K."

The **Latitude** and **Longitude** fields contain, by default, the values from the WATSTORE station header "H" record or the WDM attributes. They are used to compute the generalized skew if it is not entered. These fields may be updated interactively by clicking in the desired station row and typing the new value.

The **Mean Sq**uare **Err**or, **Low**est **Hist**oric **Peak**, and **High**est **Sys**tematic **Peak** fields are informational and cannot be modified. These values are determined from the peak record for the station.



Station ID	Include in Analysis?	Beginning Year	Ending Year	Historic Period	Skew Option	Generalized Skew	Gen Skew Std Error	Mean Sqr Err	Low Hist Peak	Lo-Outlier Threshold	High Sys Peak	Hi-Outlier Threshold	Gage Base Discharge	Urban/Reg Peaks	Latitude	Longitude	Plot Name
03606500	Yes	1897	1973	0	Weighted	-0.5	0.55	0.3025	18500	0	17000	0	0	No	36.0386	88.2283	03606500
03606500	Yes	1897	1973	77	Generalized	-0.5	0.55	0.3025	18500	0	17000	0	0	No	36.0385	88.2283	03606500-1
01373500	Yes	1882	1968	0	Weighted	0.6	0.55	0.3025	4400	0	8800	0	0	No	41.5111	73.9486	01373500
06600500	Yes	1935	1973	0	Weighted	-0.2	0.55	0.3025	0	0	71500	0	0	No	42.5767	96.3119	06600500
06600500	Yes	1935	1973	82	Weighted	-0.2	0.55	0.3025	0	0	71500	70000	0	No	42.5767	96.3119	06600500-1
01614000	Yes	1929	1973	0	Weighted	0.2	0.55	0.3025	0	0	22400	0	0	No	39.5119	78.0375	01614000
11274500	Yes	1932	1973	0	Weighted	0.2	0.55	0.3025	0	0	10200	0	0	No	37.3169	121.128	11274500
11274500	Yes	1932	1973	0	Generalized	0.2	0.55	0.3025	0	0	10200	0	0	No	37.3169	121.128	11274500-1

Figure 5. Example of the Station Specification Tab of program PKFQWin after a file containing 8 sets of data has been opened.

Output Options

The **Output Options** tab is used to modify options for output that will be used for all of the stations processed. These include the output file name, saving additional output to other files, including additional information in the output file, plot types and styles, and confidence limits.

The **Output File** panel in figure 6 contains the name of the file that will be used for all regular output from the program. This includes a summary of the input data, computed results in tabular format, and any warning or error messages that may be issued. By default, the output file name will use the prefix of the input file name and have the .prt suffix. If a .psf file is used and the O FILE record is included, that file name will be used. A different name may be specified for output by choosing the Select button and entering the name for the file. See appendix D.3 for an example output file. See appendix A for information on the error and warning messages that may be written to this file.

Three additional types of information may be included in the regular output file by selecting the appropriate check boxes. Selecting **Output Intermediate Results** will result in additional messages and tabulated information that may be useful for debugging purposes. Selecting this check box is equivalent to specifying O DEBUG YES in the .psf file. Selecting **Print Plotting Positions** will result in the Empirical Frequency Curves table being included in the regular output; this table contains the points used to generate the annual exceedance probability plot. Selecting this check box is equivalent to specifying O PLOT PRINTPOS YES in the .psf file. Selecting **Line Printer Plots** results in a plot rendered using keyboard characters; this option is included for consistency with older versions of the program and is equivalent to specifying PRINTER or BOTH for O PLOT STYLE in the .psf file.

The **Additional Output** panel contains check boxes for two other files. If the peaks are read from a Watershed Data Management (WDM) file, the computed statistics may be saved as attributes in the **WDM** file. These statistics are identified in appendix C, table C.2. The computed statistics may also be saved to a file in the **Watstore** standard Basin Characteristic format; see appendix B.5 for an example and a description of this file. By default, the Watstore output file name will use the prefix of the input file name and have the .bcd prefix. A different name may be specified for Watstore output by choosing the Select button and entering the name for the file. The Watstore output option is included for consistency with older versions of the program.

Within PKFQWin, a variety of **Graphic Plot Formats** are available for the annual exceedance probability plot. These include:

- Computer Graphics Metafile (**CGM**)
- PostScript (**PS**)
- Raster Bitmap (**BMP**)
- Windows Metafile (**WMF**)

By default, **NONE** are produced. Click on the appropriate radio button for the desired format. There will be one file for each station analysis. If the radio button for CGM, PS, or WMF is selected, temporary BMP files are generated to be used for viewing from within PKFQWin; these files

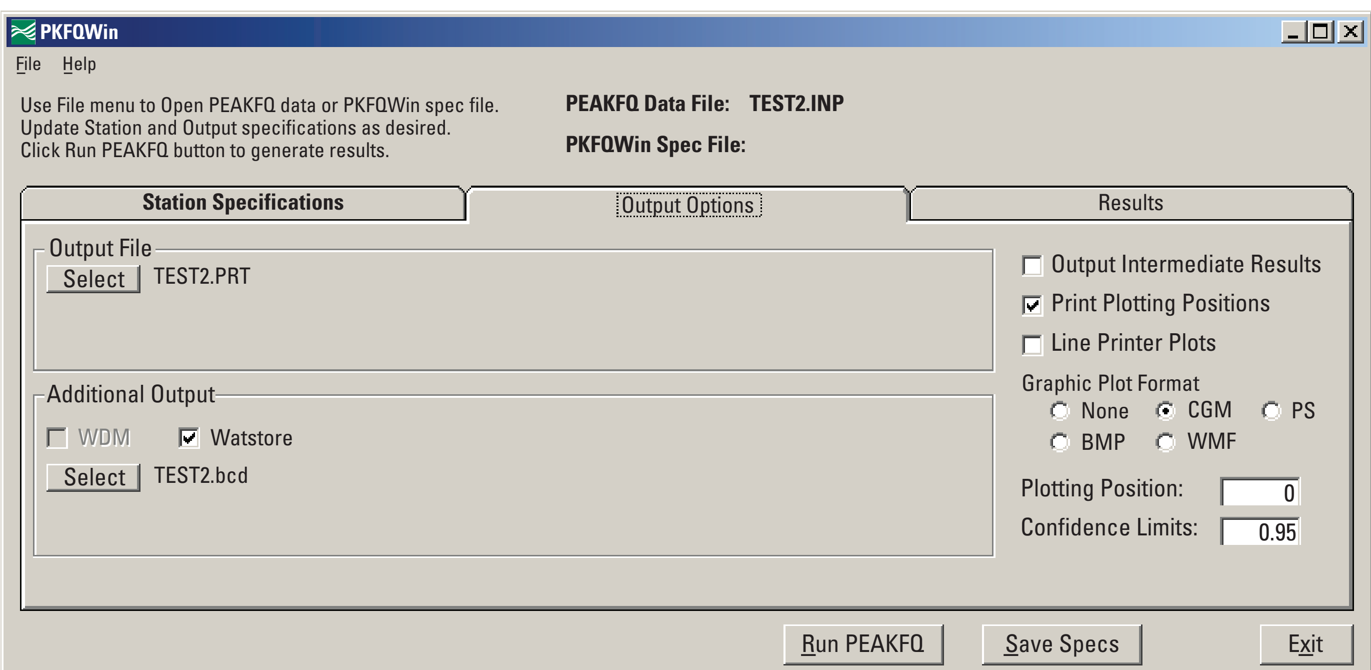


Figure 6. Example of the Output Options tab of program PKFQWin after an input file has been opened.

are deleted at the end of the session. If the BMP format is selected, the files are retained at the end of the session.

By default, the **Plotting Position** used is 0.0, this is the Weibull plotting position. Other named plotting positions include Median/Beard (0.3), Bolm (0.375), Cunnane (0.4), and Gringorten (0.44). The plotting position is entered as a numeric value and is not restricted to the named values. See the description of O PLOT POSITION in appendix B.1 for a description of how the plotting position is computed.

Upper and lower **Confidence Limits** for the Bulletin 17B estimates are drawn on the graph and also tabulated in the output file. By default, the 95-percent confidence limits are used (0.95).

Results

The **Results** tab shown in figure 7 allows viewing of the various forms of PeakFQ results. These include the main output file, the additional output file (if in use), and graphic plots.

The **View** buttons in the Output File and Additional Output frames are used to open those files for viewing. The files will be viewed with the system's default viewer of Text files.

The **Graphs** list displays the available plots from the stations that were analyzed. This list is populated only if the user selects something other than **None** for the **Graphic Plot Format** on the Output Options tab. The default base file names are the station IDs. If a station is analyzed more than once, an index is attached to the station ID to make its graph name unique. The **View** button under the list of graphs will cause the selected graphs to be displayed, each in a separate window.

The graphs viewed from the PKFQWin interface are in Bitmap (BMP) format. If, on the Output Options tab, the user selected another graphic format (for example, CGM, PS, WMF), the graphs will also be stored in that selected format for use outside of PeakFQWin. The Bitmap files will not be saved for later use unless that was the selected graphic format.

Save Specs

The Save Specs feature shown in figure 8 (menu option or command button) allows the user to save a set of specifications for future use. The specifications from the last PeakFQ run will be written to a PSF file. The PSF file will contain only specifications that are not the default values for the run.

Batch Version

The PeakFQ batch program is run from a command prompt by typing the executable file name followed by an input specification file. It may be desirable to pipe the output to a file to capture any messages. For example:

```
PEAKFQ TEST2.PSF>TEST2.RUN
```

Paths to any of these files may also be included.

The batch program is given instructions for the run through the PeakFQ specification file (*.psf). The .psf extension is not required, but is useful for file organization. The only required elements of the specification file are the input

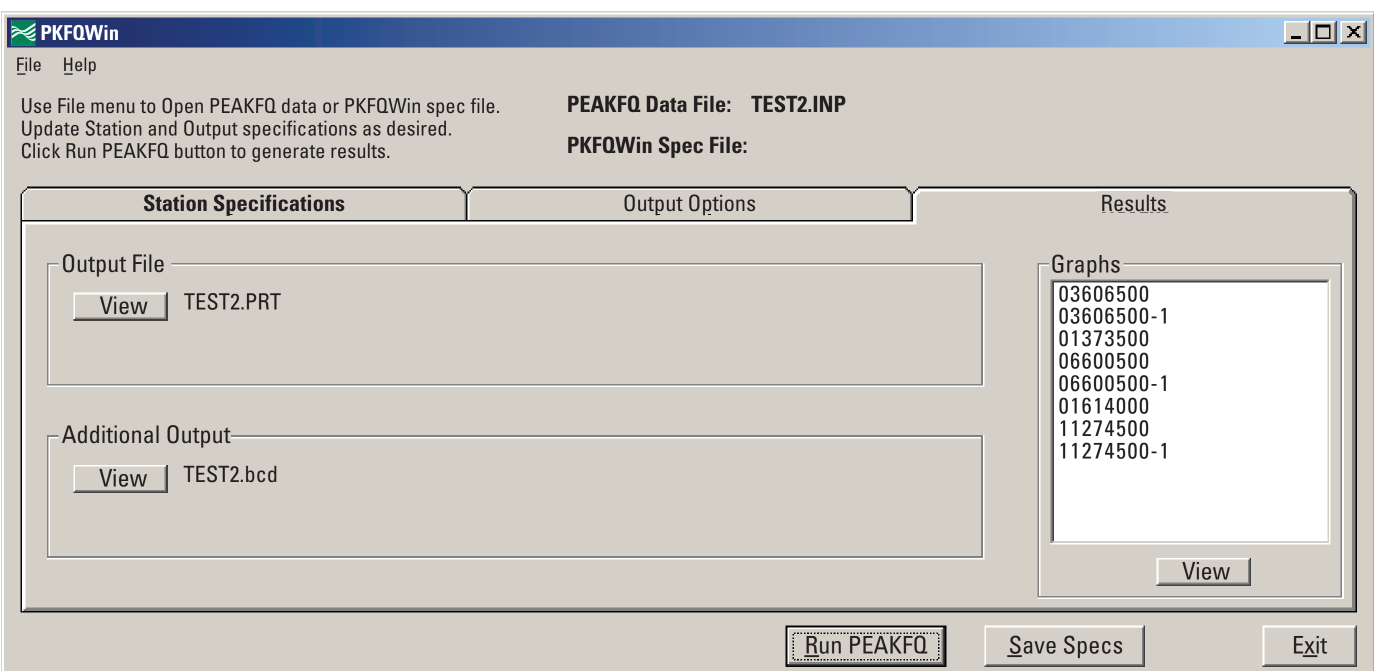


Figure 7. Example of the Results tab of program PKFQWin after the Run PEAKFQ button has been selected.

data file and the output file. Thus, the simplest example of a specification file might look like this:

I ASCI Test2.inp
O FILE Test2.OUT

Defaults for all output and station specifications are defined in the code. These specifications may then be updated by the input data file either through Watstore "I" cards or WDM attributes. Finally, specification updates may be made through the PeakFQ specification file.

Details of all specification file elements are found in Appendix B.1.

REFERENCES CITED

Beard, L.R., 1960, Probability estimates based on small normal-distribution samples: Journal of Geophysical Research, v. 65, no. 7, p. 2143-2148.

Flynn, K.M., Hummel, P.R., Lumb, A.M., and Kittle, J.L., Jr., 1995, Users manual for ANNIE, version 2, a computer program for hydrologic data management: U.S. Geological Survey Water-Resources Investigations Report 95-4085, 211 p., *http://water.usgs.gov/software/annie.html*.

Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood-flow frequency: Bulletin 17B of the Hydrology Subcommittee, Office of Water Data Coordination, U.S. Geological Survey, Reston, Va., 183 p., *http://water.usgs.gov/osw/bulletin17b/bulletin_17B.html*.

Kirby, W.H., 1981, Annual flood frequency analysis using U.S. Water Resources Council guidelines (program J407): U.S. Geological Survey Open-File Report 79-1336-I, WATSTORE User's Guide, v. 4, chap. I, sec. C, 56 p.

Lepkin, W.D., DeLapp, M.M., Kirby, W.H., and Wilson, T.A., 1979, National Water Data Storage and Retrieval System WATSTORE User's Guide: U.S. Geological Survey Open-File Report 79-1336-I, v. 4, ch. I, secs. A, B, and C.

Tasker, G.D., 1978, Flood Frequency Analysis with a Generalized Skew Coefficient: Water Resources Research, v. 14, no. 2, p. 373-376.

Wallis, J.R., Matalas, N.C., and Slack, J.R., 1974, Just a Moment: Water Resources Research, v. 10, no. 2, p. 211-219.

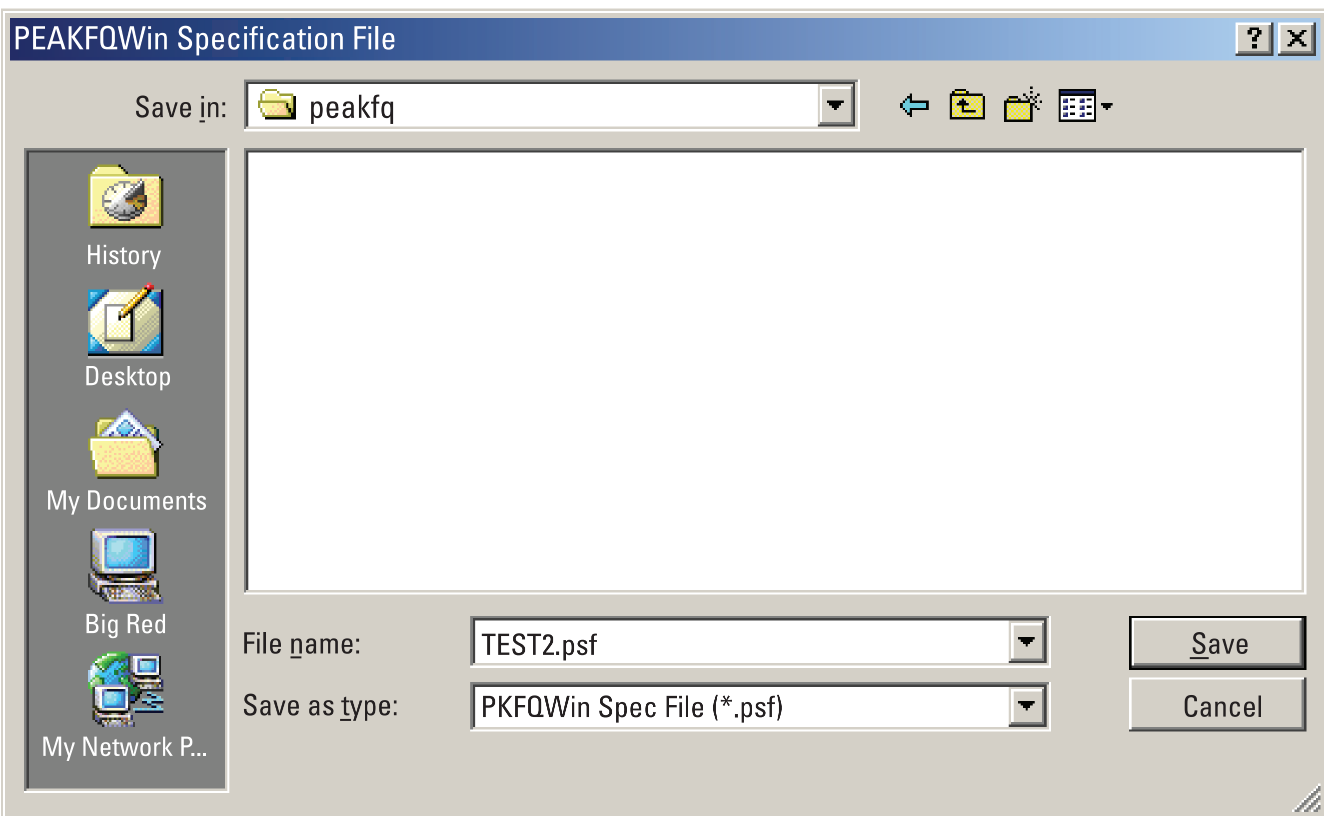


Figure 8. Example of the Save Specifications File window in program PKFQWin, obtained by selecting the Save Specs button at the bottom of the PKFQWin window.

APPENDIX A. PeakFQ Diagnostic Messages

Diagnostic messages are produced when real or potential errors are detected. The diagnostic messages included in the PeakFQ output file are substantially the same as those produced by the original J407 procedure in WATSTORE (Kirby, 1981). These messages are listed and briefly explained below.

Most of the messages have the following general format:

***iiinnns - text data

in which:

***	represents a variable number, possibly zero, of asterisks, to call attention to the message
iii	identifies the general part of the program producing the message as follows: INP - input processing PKF -reading the peak flow file retrieval records (WATSTORE card format) WCF - flood frequency calculations following Bulletin 17B guidelines
nnn	is a message number
s	is a severity indicator. E means error, W means warning, I and J mean routine information, and L means listing of data or results.
text	is the text of the message
data	is a list of numbers or words generally in the same order as items mentioned in the text

The messages are listed below approximately in alphabetic and numerical order by identifier and number.

FRQPLT WILL DROP POINTS BELOW PLOT BASE.

One or more points on the computed empirical frequency curves fall below the lower boundary of the plot. These points will not be plotted.

INPUT2 HISTORIC PEAKS OVERFLOWED. nhp i sta-id

The number of historic peaks (nhp) retrieved for station (sta-id) exceeds the capacity of program PeakFQ (20 historic peaks). Possible system error: check the input for validity; if there are more than 20 historic peaks (code 7), notify *h2osoft@usgs.gov*.

INPUT2 REQUESTED YEARS NOT IN RECORD. beg-yr end-yr first-yr last-yr sta-id

Probable user error. The years requested on the I-card (beg-yr, end-yr) do not overlap the years available in the record (first-yr, last-yr) at the station (sta-id).

INPUT2 STATION HAS NO PEAK FLOW DATA. STA-ID = xxxxx

Informative message. See preceding messages for explanation. Processing continues with next valid input record.

INPUT2 PEAK COUNT EXCEEDS STORAGE CAPACITY npks sta-id

The number of peaks (npks) retrieved for station (sta-id) exceeds the capacity of program PeakFQ (200 peaks). Possible system error: check the input for validity; if there are more than 200 peaks, notify *h2osoft@usgs.gov*.

PKFRD4 PEAK OVERFLOW. NPKS,MAX = n max

The number of peaks (n) exceeds the storage capacity (max) of program PeakFQ. Probable system error; notify WATSTORE User assistance.

PKFRD4 Insufficient data to process, only nnn peaks for station sta-id

Only nnn peaks were identified to be include in the analysis for station id sta-id. A minimum of 3 peaks is required.

PKFRD4 CARD types 4, 2, and * are ignored.

card-image

PKFRD4 Unrecognized CARD type. Must be Y, Z, N, H, I, 2, 3, 4, or *.

(2, 4, and * records are ignored.)
card-image

PKFRD4 Error reading input lat. or long. on H card.

card-image

PKFRD4 Error reading I-card

card-image

WCF001J FLOOD FREQUENCY, BULLETIN 17-B. VER n.n (dddddd).

Unedited machine computations. User is responsible for interpretation and use. n.n (dddddd) = version number and date of last revision. Normal beginning-of-job message, if requested.

WCF002J CALCS COMPLETED. RETURN CODE = n

Normal end-of-job message. Return codes: 0 = no error detected. 1 = non-standard data accepted, 2 = warning – calculations completed, but results may be incorrect.

WCF003E CALCS ABORTED. RETURN CODE = 3.

WCF … Routines were unable to complete the calculations for reasons explained in previous messages.

WCF004* INTERNAL PROGRAM LOGIC ERROR. Location-code data

This message should not occur. If it does, contact *h2osoft@usgs.gov*

WCF101L INPUT PARAMS

GEN SKU	OPT	GS.E RR	GAG EB	QLW OUT	QHI OUT	NHIS T	HIST PD	
xxx		xxx	xxx	xxx	xxx	xxx	xxx	xxx

Routine listing of input data, if requested.

WCF102E INVALID PEAK COUNTS. NPK, NHIST = nnn nnn

Either the number of historic peaks (HNIST) is negative or the total number of input peaks is less than NHIST. Probable error in counting input peaks.

WCF103L INPUT PEAKS, HISTORIC FIRST. TOTAL NO. = nnn

Routine listing of input data, if requested.

WCF104L INPUT LOG PKS, HIST FIRST. TOTAL NO. = nnn

Routine listing of input data, if requested.

WCF107I ACCEPTED GEN SKEW OUTSIDE MAP LIMITS GS m1 m2

Input generalized skew GS was outside range of values (m1, m2) set at program installation time. (Limits of Bulletin 17B skew map.)

WCF109W PEAKS WITH MINUS-FLAGGED DISCHARGE WERE BYPASSED. nnn

nnn negative input peaks were found. These are assumed to be codes for unknown discharges. These peaks are ignored in the computations, but large negative values are stored in corresponding locations in output logarithm vector. If the input has any unknown discharges coded as negative values, ensure that these peaks legitimately can be ignored. Otherwise, incorrect input peak counts may cause this message. Warning only—analysis continues.

WCF111E HISTORIC PEAK HAD MINUS-FLAGGED DISCHARGE .

One of the historic peaks was negative. Probable error in input data value or count.

WCF113W NUMBER OF SYSTEMATIC PEAKS HAS BEEN REDUCED TO NSYS - nnn

Missing-discharge peaks were noted and have been omitted from the sample (WCF109). The correct sample size for analysis is nnn.

WCF117E NO DATA IN SYST RECORD.

NSYS	NPK	NHIST	NMISS
xxx	xx	xxx	xxx

There is no systematic record at this station. Possible error in input data or input peak count. NPK = total number of input peaks. Other as in list of variables.

WCF118W SYSTEMATIC RECORD SHORTER THAN 17B SPEC. nnn

Systematic record length nnn is less than that specified in Bulletin 17B. Analysis proceeds, but sample size may be too small for reliable conclusions.

WCF133I SYSTEMATIC PEAKS BELOW GAGE BASE WERE NOTED.

nnn bbb
nnn = number of below-gage-base peaks. bbb = gage-base-discharge.

WCF134I NO SYSTEMATIC PEAKS WERE BELOW GAGE BASE. bbb

bbb = gage-base-discharge.

WCF141E SAMPLE SIZE TOO SMALL TO CALC STATS. lll nnn

NSYS,	NBG,	NLWOUT,	NHIOUT,	NHISTN,	HISTPN
xxx	xx	xx	xx	xx	xx

Either the systematic or the Bulletin 17B-adjusted record is too short for calculation of mean, variance, and skew (less than 3 items). Probable cause—insufficient record length or excessively high gage base or user low-outlier criterion. lll = either SYS (systematic) or 17B. nnn = effective sample size. Others as in list of variables.

WCF143E NEGATIVE VARIANCE OF LOGS. lll vvv

lll = either SYS (systematic) or 17B. vvv = the computed variance (should be near zero). Probable cause—round-off error in computing near-zero variance when all input peaks are (nearly) equal.

WCF151I 17B WEIGHTED SKEW REPLACED BY USER OPTION. www uuu igsopt

Bulletin 17B weighted skew calculation (www) has been superseded by user-specified skew uuu. An igsopt value of 1 means generalized skew; -1 means station skew.

WCF156I 17B HIGH-OUTLIER TEST SUPERSEDED BY MIN HIST PK. www

Routine information report of Bulletin 17B high-outlier test criterion (www). Historic peaks were present below this threshold, so the historic-high-outlier threshold was lowered to the level of the smallest historic peak.

WCF157W USER HIGH-OUTLIER CRIT LOWERED TO MIN HIST PK. uuu hhh

Probable user error—if historic peaks are given. The high-outlier base need not be set unless peaks smaller than the smallest historic peak are to be treated as high outliers. uuu = user high outlier criterion. hhh minimum historic peak.

WCF159E HIGH-OUT/HIST-PK BASE BELOW LOW-OUT/GAGE BASE. hhh lll

Probable user error—perhaps the high-outlier and low-outlier or gage-base data have been entered in the wrong order. hhh = high-outlier or historic base. lll = low-outlier or gage base.

WCF161I USER HIGH-OUTLIER CRITERION REPLACES 17B. uuu www

The user-specified historic-peak-high-outlier discharge threshold (uuu) has been noted. Its value supersedes the Bulletin 17B-recommended value (www).

WCF162I SYSTEMATIC PEAKS EXCEEDED BY HIGH-OUTLIER CRITERION. nho hhb

One or more (nho) systematic peaks exceeded the high-outlier discharge criterion (hhb). No historic adjustment was applied because the user did not specify the length of the historic period.

WCF163I NO HIGH OUTLIERS OR HISTORIC PEAKS EXCEEDED HHBASE. hhb

No high outliers or historic peaks were detected. The historic-peak-high-outlier discharge threshold is hhb.

WCF164W HISTORIC PERIOD IGNORED. histpd

A historic period length (histpd) was specified, but no high outliers or historic peaks were found. The historic period length is ignored and no Bulletin 17B historic adjustment is attempted. Probable user error—the historic period length should not be specified unless one or more historic peaks or high outliers are present.

WCF165I HIGH OUTLIERS AND HISTORIC PEAKS ABOVE HHBASE. nho nhp hhb

Historic adjustment was applied. nho = number high outliers noted, nhp = number historic peaks, hhb = high outlier/historic base flow.

WCF167E HIST PERIOD NO LONGER THAN SYS+HIST PEAKS. hhh nnn

Stated historic period hhh is no longer than actual count of observed peaks nnn. Probable user error - if both hhh and nnn are correct there is no point in doing the historic adjustment.

WCF169I ACCEPTED HISTORIC PERIOD GTR THAN T hhh ttt

The historic period hhh may be longer than can be justified under the Bulletin 17B criteria for historic information. T = 5 * systematic record, up to max of 300 yrs.

WCF171W NUMBER HI-OUT/HIST PKS EXCEEDS 10PCT OF SYS.PKS. nho nhp

Excessive number of historic peaks nhp and high outliers nho suggest that historic base may be set too low to ensure that every peak exceeding it has been recorded.

WCF191I USER LOW-OUTLIER CRITERION SUPERSEDES 17B. uuu www

uuu = user low-outlier criterion, www = Bulletin 17B low-outlier criterion

WCF193E LOW OUTLIER CRITERION EXCEEDS HIGH-HIST lll hhh

Probable user error—perhaps the high-outlier and low-outlier or gage-base data have been entered in the wrong order. hhh = high-outlier or historic base. lll - low-outlier or gage base.

WCF195I NO LOW OUTLIERS WERE DETECTED BELOW CRITERION xxxxx

No peaks above the gage base were below the low-outlier criterion. xxxxx = low outlier criterion adopted (user or 17B).

WCF198I LOW OUTLIERS BELOW FLOOD BASE WERE DROPPED. nnn bbb

Peaks above the gage base and below the low-outlier criterion were noted. The flood base of the Bulletin 17B frequency curve has been set at the low-outlier criterion. nnn = number of low outliers dropped. bbb = Bulletin 17B flood base.

WCF199W NUMBER OF PEAKS BELOW FLOOD BASE EXCEEDS 17B SPEC. nbb bbb maxnbb

Bulletin 17B specifies a maximum number of peaks that may fall below the flood base for this length of systematic record. The actual number nbb of below-base peaks exceeds this limit (maxnbb). The flood base = bbb. Warning—the calculation proceeds but the results may be unreliable.

WCF213E COND PROB ADJUST FAILED - EXCESSIVE lll PROB BELOW BASE. ppp

The conditional probability adjustment described in appendix 4 of Bulletin 17B cannot be performed when ppp fraction of the peaks are below the flood base. lll = (SYS for systematic rec freq curve, in which case flood base = gage base) or 17B.

WCF215E SKEW OUT OF TABLE RANGE. lll skew-a skew-u gensku

One or more of the skews to be used in constructing the Pearson Type III curve for either the systematic or Bulletin 17B record is out of the range of the Bulletin 17B Pearson Type III table (+ or - 9.0). lll = either SYS or 17B. Skew-a is the skew of the above base peaks. skew-u and gensku may not be present. skew-u is the unconditional skew after any conditional probability adjustment or weighted-skew calculation. gensku is the generalized skew, and is printed only if the error is detected after the Bulletin 17B weighted-skew calculations.

WCF217L	FREQUENCY CURVE PARAMS lll xxx	xxxxx	xxxxx	xxxxx
		xxxxx	xxxxx	xxxxx

Routine report of frequency curve parameters for curve lll (SYS or 17B). Parameters are—probability of exceeding base; mean, standard deviation, and skew of frequency curve (unconditional—after a conditional probability adjustment, if P (BASE) is less than 1.0); mean, standard deviation, and skew of above-base peaks (before conditional probability adjustment).

WCF219J	FREQ CURVE ORDINATES lll	2-YR(.50)	10-YR (.10)	100-YR (.01)
		xxxxxxx	xxxxxxx	xxxxxxx

Routine report of frequency curve ordinates at 2-yr, 10-yr, and 100-yr levels for lll (SYS or 17B) frequency curve.

WCF233W EXPECTED PROB OUT OF RANGE AT TAB PROB xxxxx yyyyy

Expected-probability calculation called for table lookup at expected probability xxxxx beyond the limits of the computed Bulletin 17B frequency curve. This message normally occurs several times when sample size is less than about 10 years and tabular probability yyyyy is less than about 0.10.

WCF238J FREQ CURVE 17B-EXPECT PROB xxxxxxx xxxxxxx xxxxxxx

Routine report of 'expected-probability'frequency curve ordinates at 2-, 10-, and 100-yr levels. 'Expected-probability' curve is based on Bulletin 17B frequency curve.

WCF239J	FREQ CURVE CONF LIMS 17B xx.x	xxxxxxx	xxxxxxx	xxxxxxx
		xxxxxxx	xxxxxxx	xxxxxxx

Upper and lower confidence limits for Bulletin 17B frequency curve ordinates reported in WCF219. xx.x = one-sided confidence level.

APPENDIX B. PeakFQ Data Formats

Appendix B contains detailed documentation of text files read by PeakFQ. Appendix B.1 describes the PeakFQ specification file used to run the batch program. Appendices B.2 - B.4 describe the WATSTORE standard formats used by PeakFQ.

- Appendix B.2 describes the station header records
- Appendix B.3 describes the station option records
- Appendix B.4 describes the peak-flows records

Appendix B.1. Batch Specifications

This appendix gives detailed descriptions of the PeakFQ specification (PSF) file. Running the batch version of PeakFQ, whether stand-alone or from the PKFQWin interface, requires a specification file as a command line argument. There are only two required records in PeakFQ specification files. These are the input data file and the main output file. The input data file record must start with "I", followed by either the "ASCI" (for Watstore text files) or "WDM" (for WDM files) keyword, followed by the name of the input data file. Here are examples of each:

```
I ASCI Test2.inp
I WDM Test.wdm
```

The main output file record must start with "O", followed by the "FILE" keyword, followed by the output file name. For example:

```
O FILE Test2.out
```

Other output specification records (also starting with "O"), are used to define output options that apply to the entire run. These specifications are described in table B.1.1.

Table B.1.1. Specification file output keywords that apply to the entire run.

Note: Each keyword is preceded by the letter O and a space.

Keyword	Valid Values	Default	Description
DEBUG	YES NO	NO	Yes provides additional printout of intermediate results in the analysis.
ADDITIONAL	WDM WAT BOTH NONE	NONE	WDM (or BOTH) puts computed statistics on each data set as attributes for further statistical analysis. WAT (or BOTH) puts computed statistics on file in the format described in Appendix B.5.
EMA	YES NO	NO	NO will run the traditional Bulletin 17B analysis. YES will run the new EMA analysis.
CONFIDENCE	0.nn	0.95	Where 0.nn is confidence limit percent as a fraction.
PLOT STYLE	GRAPHICS PRINTER BOTH NONE	NONE	GRAPHICS (or BOTH) will generate a graphic file for each station analyzed. PRINTER (or BOTH) uses ASCII character set on a line printer, 132 characters wide.
PLOT FORMAT	CGM PS BMP WMF	CGM	Format for GRAPHICS plots (Note: some file formats may not be available on all computer platforms).
PLOT PRINTPOS	YES NO	YES	YES provides additional table in the printout listing the observed peaks and assigned recurrence intervals.
PLOT POSITION	0.nn	0.00	Plotting position computed as (m-a)/(N+1-2a) where m is order number, N is total number of peaks, and a is a parameter where: a = 0.00 for Weibull a = 0.30 for Median/Beard a = 0.375 for Blom a = 0.4 for Cunnane a = 0.44 for Gringorten

The remaining specifications are made for each station being analyzed in the run. If specifications are to be made for a station, the first record must indicate the station to which the specifications apply:

STATION <staid>

where <staid> is either the alphanumeric Station ID from the WATSTORE file or the data-set number from the WDM file.

Table B.1.2 describes the available station specifications. This sequence of a STATION record followed by any desired specifications is then repeated for each station to be analyzed in the run.

Two additional keywords may be found in PSF files, particularly those generated and used by PKFQWin. These are VERBOSE and UPDATE. The VERBOSE keyword will only be found at the start of a PSF file and indicates that all possible specifications are written out in the file, even if they are the default value. The UPDATE keyword will only be found at the end of a PSF file and indicates that as PeakFQ performs the run, it should write out the PSF file in VERBOSE mode.

The following sample PSF File is written in VERBOSE mode and contains just the first station from the Test2.inp sample data file (included in program distribution.)

```
I ASCI TEST2.INP
O File TEST2.OUT
O Plot Style None
O Plot PrintPos Yes
O Plot Position     0.00000
O Additional None
O Debug No
O EMA No
O Confidence    0.950000
Station 03606500
   SkewOpt Weighted
   GenSkew   -0.500000
   SkewSE    0.550000
   BegYear         1897
   EndYear         1973
   HistPeriod     0.00000
   Urb/Reg No
   LoThresh     0.00000
   HiThresh     0.00000
   GageBase     0.00000
   Latitude     36.0386
   Longitude     88.2283
```

Table B.1.2. Specification file keywords that apply to a specific station.

Keyword	Valid Values	Default	Description
GENSKEW	n.nnn	From Generalized skew map using lat/lng	Where n.nnn defines the estimated skew based on experience at nearby stations or regional analysis.
SKEWSE	n.nnn	0.55	Where n.nnn defines the standard error of the generalized skew. If not specified, the standard error of the generalized skew map, 0.55, will be used.
BEGYEAR	nnnn	From data file	Where nnnn defines the first water year of data to be used in the analysis.
ENDYEAR	nnnn	From data file	Where nnnn defines the last water year of data to be used in the analysis.
HISTPERIOD	nn	0.0	Where nn defines the length of historic period in years (entering 0.0 will cause the historic peaks to be ignored). Must be greater than the systematic period.
SKEWOPT	GENERALIZED WEIGHTED STATION	WEIGHTED	STATION- station skew computed from recorded peaks GENERALIZED - generalized skew from map or regional analysis WEIGHTED - weighted between STA and GEN
URB/REG	YES NO	NO	Peaks affected by urban development or upstream regulation will be ignored unless this is YES.
LOHIST	nnnn	0.0	Where nnnn displays the lowest historic peak. This value is only informational for display in the PKFQWin interface.
LOTHRESH	nnnn	0.0	Where nnnn defines the low-outlier discharge criteria. If greater than 0.0, will override the Bulletin 17-B computed low-outlier criteria.
HISYS	nnnn	0.0	Where nnnn displays the highest systematic peak. This value is only informational for display in the PKFQWin interface.
HITHRESH	nnnn	0.0	Where nnnn defines the high outlier threshold.
GAGEBASE	nnnn	0.0	Where nnnn defines the lower limit of measurable flood peak discharge. If greater than 0.0, will supersede the gage base discharge inferred from any "less than" qualification codes.
LATITUDE	nn.nn	From data file	Where nn.nn defines latitude, in degrees, for computing generalized skew.
LONGITUDE	nnn.nn	From data file	Where nnn.nn defines longitude, in degrees, for computing generalized skew.

Appendix B.2. Station Header Records (WATSTORE standard format)

The optional station header records are described in table B.2. These records contain some fields not read by PeakFQ; for completeness, these fields are included in the description. If latitude and longitude are not provided on an H record, either latitude and longitude or generalized skew must be input elsewhere.

If included in the input file, the H, N, and Y records must contain the station identification number. The Record identifier is required for all records in the input file. Only fields described as required or optional are read by the PeakFQ program. Example:

```
columns           1         2         3         4         5         6         7         8
          ----+----0----+----0----+----0----+----0----+----0----+----0----+----0----+----0
          Z                              USGS
          H 03606500      3602190881342004747017SW 6040005 205.00         380.58
records   N 03606500      BIG SANDY RIVER AT BRUCETON, TENN
          Y 03606500      2000.00
```

Table B.2. WATSTORE station header record formats.

Record	Column	Format	WDM Attribute	Description
Z				**Agency Indentification Record - optional**
	1	“Z”	--	Record identifier. Required.
	33-37	A5	AGENCY	Agency code as assigned by WATSTORE.
H				**Station Header Record - optional** Note: If LATDEG and LNGDEG are not entered, latitude and longitude or generalized skew must be input elsewhere.
	1	“H”	--	Record identifier. Required.
	2-16	A15	STAID or ISTAID	Station identification number. Required.
	17-31	---		Station locator. Required.
		3I2	LATDEG	Latitude, DDMMSS. - Optional.
		I3,2I2	LNGDEG	Longitude, DDDMMSS. - Optional.
		I2		Sequence number.
	32-33	A2	STFIPS	Numeric state code where station is located.
	34-35	A2	DSCODE	For USGS sites only, the district (numeric state code) of the alpha project code of the office responsible for collecting and storing the data.
	36-38	A3	COCODE	FIPS county code where the station is located.
	39-40	A2	SITECO	Site code indicating the major class of data collected at the site: SW - stream GW - well SP - spring LK - lake or reservoir ES - estuary ME - meteorological
	41-48	I8	HUCODE	Hydrologic unit code from the USGS state hydrologic unit maps.
	49-55	F7.0	DAREA	Total drainage area, in square miles.
	56-62	F7.0	CONTDA	Contributing drainage area, in square miles.
	63-70	F8.0	DATUM	Datum, feet above mean sea level.
	71-79	F9.0	WELLDP	Well depth, in feet.
N				**Station Name Record - optional**
	1	“N”	--	Record identifier. Required.
	2-16	A15	STAID or ISTAID	Station identification number. Required.
	17-64	A64	STANAM	Station name. Required.
	65-72	A8	GUCODE	Major geologic unit codes as assigned by WATSTORE.

Table B.2. WATSTORE station header record formats—continued.

Record	Column	Format	WDM Attribute	Description
	73	A1	AQTYPE	Aquifer type code assigned by WATSTORE: U - unconfined single aquifer N - unconfined multiple aquifers C - confined single aquifers M - confined multiple aquifers X - mixed multiple aquifers
Y				**Base Discharge Record - ignored.**
	1	"Y"	--	Record identifier.
	2-16	A15	STAID or ISTAID	Station identification number. Required.
	7-23	F7.0	BASEQ	Base discharge. Note: this is not the Gage Base Discharge used in PeakFQ.

Appendix B.3. Station Option Record (WATSTORE standard format)

The station option record is described in table B.3; it is optional. If included in the input file, the I record must contain the station identification number, all other fields are optional and may be left blank. Program PeakFQ reads all of the fields on this record. The description column describes how blank fields are handled. Example:

```
columns            1         2         3         4         5         6         7         8
          ----+----0----+----0----+----0----+----0----+----0----+----0----+----0----+----0
records   I                 -.2    82.     70000.
```

Table B.3. WATSTORE station option record formats.

Record	Column	Format	WDM Attribute	Description
I				**Station option record - optional**
	1	"I"	--	Record identifier. Required.
	2-16	A15	STAID or ISTAID	Station identification number. Required.
	17-24	F8.0		Generalized skew. If not specified, the generalized skew will be determined based on gage latitude and longitude using the generalized skew map accompanying the Bulletin 17B guidelines.
	25-32	F8.0		Length of historic period in years. A positive value must be supplied in order for the historic adjustment to be applied. The historic period contains the systematic record period as a subset. If this field is left blank, any input historic peaks will be ignored and any high outliers will be treated as normal systematic peaks.
	33-40	F8.0		User-specified historic-high-outlier discharge threshold. Used only in conjunction with the historic period, this threshold is used to override the Bulletin 17B-computed high-outlier threshold. If this field is left blank, the Bulletin 17B threshold will be lowered automatically to equal the smallest historic peak(s) if one is known. If a positive value is specified in this field, all peaks that exceed this value will be used in the historic adjustment. Any historic peaks lying below this value will be ignored.
	41-48	F8.0		User-specified low-outlier discharge criterion. This criterion, if a postive number, will override the Bulletin 17B computed low-outlier criterion. A blank, negative value, or zero will be ignored.
	49-56	F8.0		Gage base discharge, representing a lower limit of measurable flood peak discharge at the site. This discharge, if a positive number, will supersede the gage base inferred from any "less than" qualification codes of the input peak flow records. A blank, negative value, or zero will be ignored. (The gage base discharge is not the same as the partial-duration base discharge that may be recorded in the Station Header record.)
	57-64	F8.0		Standard error for the generalized skew. If not specified, a value of 0.55, corresponding to the standard error of the generalized skew map accompanying the Bulletin 17B guidelines, will be used.
	65-69	5A1		Station-option codes selected from the following list. The codes may be in any order or combination and may be in any available column. In case of conflict, the rightmost code is used. The available options are: S - Station-skew option. Causes the station skew, adjusted for outliers and historic data, rather than the Bulletin 17B weighted skew, to be used for the final frequency curve. G - Generalized-skew options. Causes the generalized skew, rather than the Bulletin 17B weighted skew, to be used in the final frequency curve. K - Known regulation/urbanization input option. Allows peaks with the known regulation or urbanization codes (6 or C) to be included in the statistical analysis. H - Historic peak input option. Allows all historic peaks to be used, whether or not they exceed the user-specified historic-high-outlier discharge threshold. The program will print a warning message if it finds any below-threshold historic peaks and will lower the threshold to include them. Use of the option may cause the historic adjustment to include some historic and systematic peaks that are not representative of the historic period.
	71-74	F4.0		Begin year: first water year of retrieved records to be included in the statistical analysis; earlier years are ignored. This value must be either blank or a four-digit number. If blank or less than the first year of the input record, no years will be dropped from the beginning of the record.
	75-78	F4.0		End year: last water year of retrieved records to be included in the statistical analysis; later years will be ignored. This value must be either blank or a four-digit number. If blank or greater than the last year of the input record, no years will be dropped from the end of the record.

Appendix B.4. Peak-Flow Records (WATSTORE standard format)

The peak-flow data records are described in table B.4. The peak-flow record contains some fields that are not read by PeakFQ. The partial duration peak-flow data record is completely ignored by PeakFQ. For completeness, all fields for both record types are included in the description; the fields that are not read are shown with a light gray background. The peak-flow records may be preceded by station header records. Example:

```
columns           1         2         3         4         5         6         7         8
         ----+----0----+----0----+----0----+----0----+----0----+----0----+----0----+----0
records  3 03606500      192612    185007              16.50
         3 03606500      19300109   9100               13.98
```

Table B.4. WATSTORE peak-flow record formats.

Record	Column	Format	Description
3			**Peak-Flow Data Record - required.** **Fields in columns 44 - 75 are ignored.**
	1	"3"	Record identifier. Required
	2-16	A15	Station identification number. Required.
	17-20	I4	Year peak occurred - required: • Calendar year if columns 21-22 contain a valid moth. • Water year if columns 21-22 are blank or 99
	21-22	I2	Month the annual peak discharge occurred. Blank if month is not known.
	23-24	I2	Day of the month the annual peak discharge occurred. Blank if day is not known.
	25-31	F7.0	Annual peak discharge, right justified. Field may be blank.
	32-43	A12	Annual peak discharge qualification codes. More than one code may be associated with a peak, except as noted below. Field may be blank. 1 - discharge is a maximum daily average 2 - discharge is an estimate 3 - discharge is affected by dam failure 4 - discharge is less than indicated value, which is minimum recordable discharge at this site * 5 - discharge affected to unknown degree by regulation or diversion ** 6 - discharge affected by regulation or diversion ** 7 - discharge is an historic peak *** 8 - discharge actually greater than indicated value 9 - discharge due to snowmelt, hurricane, ice-jam or debris dam breakup A - year of occurrence is unknown or not exact B - month or day of occurrence is unknown or not exact C - all or part of the record affected by urbanization, mining, argicultural changes, channelization, or others D - base discharge changed during this year E - only annual maximum peak available for this year * Code 4 cannot occur simultaneously with codes 1, 2, 3, 7, or 8 ** Codes 5 and 6 cannot occur simultaneously. *** Code 7 should indicate that the value for the particular year is a historic peak and the particular year occurred before or after the systematic record, or during a break in systematic record.
	44-51	F8.0	Gage height associated with annual peak discharge, right justified in field. Ignored.
	52-55	A4	Gage height qualification codes. More than one code may be associated with a gage height. Field may be blank. Ignored. 1 - gage height affected by backwater 2 - gage height not the maximum for the year* 3 - gage height at different site and/or datum 4 - gage height below minimum recordable elevation (con't)

Table B.4. WATSTORE peak-flow record formats—continued.

Record	Column	Format	Description
	(con't)	(con't)	5 - gage height is an estimate 6 - gage datum changed during this year * If code 2 is given here, then a date and data entries should be made for the maximum annual gage height (cols 60-75)
	56-59	I4	"Highest since" year -- representing the calendar year after which the given peak discharge (cols 25-31) is known to be the highest. This year is determined from historic newspaper accounts, local information, or other sources.
	60-61	I2	Month in which the annual peak gage height occurred. While this month may not be in the same calendar year as the annual peak, it is in the same water year.
	62-63	I2	Day of the month of the annual peak gage height.
	64-71	F8.0	Annual peak gage height.
	72-75	A4	Annual peak gage height qualification codes. 1 - gage height affected by backwater 3 - gage height at different site and/or datum 5 - gage height is an estimate 6 - gage datum changed during this year
4			**Partial Duration Peak Flow Data Record - Ignored.**
	1	"4"	Record identifier. Required.
	2-16	A15	Station identification number. Required.
	17-20	I4	Year data on this record occurred. • Calendar year if columns 21-22 contain a valid month • Water year if columns 21-22 are blank or 99 This entry may or may not be the same as on the preceding type 3 record if one is present, but the month and year must have occurred in the same water year as the peak discharge.
	21-22	I2	Month the partial duration peak occurred. Blank if month is not known.
	23-24	I2	Day of the month the partial duration peak occurred. Blank if day is not known.
	25-31	F7.0	Partial duration peak discharge, right justified.
	32-43	A12	Partial duration peak discharge qualification codes. More than one code may be associated with a peak, 1 - discharge is a maximum daily average 2 - discharge is an estimate 3 - discharge affected by dam failure 4 - discharge less than indicated value, which is minimum recordable discharge at this site * 5 - discharge affected to unknown degree by regulation or diversion ** 6 - discharge affected by regulation or diversion ** 7 - discharge is an historic peak *** 8 - discharge actually greater than indicated value 9 - discharge due to snowmelt, hurricane, ice-jam or debris dam breakup A - year of occurrence is unknown or not exact B - month or day of occurrence is unknown or not exact C - all or part of the record affected by urbanization, mining, agricultural changes, channelization, or others D - base discharge changed during this year E - only annual maximum peak available for this year * Code 4 cannot occur simultaneously with codes 1, 2, 3, 7, or 8. ** Codes 5 and 6 cannot occur simultaneously. *** Code 7 should indicate that the value for the particular year is a historic peak and the particular year occurred before or after the systematic record, or during a break in the systematic record.
	44-51	F8.0	Partial duration peak gage height.

Table B.4. WATSTORE peak-flow record formats—continued.

Record	Column	Format	Description
	52-55	A4	Partial duration peak gage height qualification codes. More than one code may be associated with a gage height. 1 - gage height affected by backwater 3 - gage height at different site and/or datum 4 - gage height below minimum recordable elevation 5 - gage height is an estimate 6 - gage datum changed during this year

Appendix B.5. Basin Characteristics Records (WATSTORE standard format)

Table B.5 describes the format and contents of the basin characteristics file. PeakFQ outputs this file when the Watstore check box is selected for Additional Output on the Output Options tab. Example:

```
columns            1         2         3         4         5         6         7         8
           ----+----0----+----0----+----0----+----0----+----0----+----0----+----0----+----0
           103606500           BIG SANDY RIVER AT BRUCETON HIST B-17-B
           203606500        75   2940 76   5000 77   8280 78  10700 79  13800 80  16300
records    203606500        81  18900 82  21500 83  3.691 84  0.267 85 -0.187178  25100
           203606500       179 -0.188180  3.691181  0.267196     44197     44
```

Table B.5. WATSTORE basin characteristics record formats.

Record	Column	Format	WDM Attribute	Description
1				**Station identification record**
	1	"I"	--	Record identifier.
	2-16	A15	STAID or ISTAID	Station identification number. Required.
	17-18	I2	STFIPS	Numeric state code where station is locked. May be blank.
	19-20	I2	DSCODE	For USGS sites only, the district (numeric state code) or the alpha project code of the office responsible for collecting and storing the data. May be blank.
	21-68	A48	STANAM	Stations name and data identifer.
	69-75	F8.0		blank
	76-80	A5	AGENCY	Agency code. May be blank.
2				**Basin Characteristics, record 1**
	1	"2"	--	Record identifier.
	2-16	A15	STAID or ISTAID	Station identification number. Required.
	17-76			Pairs of basin characteristics index numbers and the value for the basin characteristic. In the order: index description
		I3,F7.0	P1.25	75 Annual flood peak, 2.5-year recurrence interval.
		I3,F7.0	P2.	76 Annual flood peak, 2-year recurrence interval.
		I3,F7.0	P5.	77 Annual flood peak, 5-year recurrence interval.
		I3,F7.0	P10.	78 Annual flood peak, 10-year recurrence interval.
		I3,F7.0	P25.	79 Annual flood peak, 25-year recurrence interval.
		I3,F7.0	P50.	80 Annual flood peak, 50-year recurrence interval.
2				**Basin Characteristics, record 2**
	1	"2"	--	Record identifier.
	2-16	A15	STAID or ISTAID	Station identification number. Required.
	17-76			Pairs of basin characteristics index numbers and the value for the basin characteristic. In the order: index description
		I3,F7.0	P100.	81 Annual flood peak, 100-year recurrence interval.
		I3,F7.0	P200.	82 Annual flood peak, 200-year recurrence interval.
		I3,F7.0	MEANPK	83 Mean of the logarithms, base 10, of systematic annual peak discharge.
		I3,F7.0	SDPK	84 Standard deviation of the logarithms, base 10, of systematic annual peak discharges.
		I3,F7.0	SKWPK	85 Skew of the logarithms, base 10, of systematic annual peak discharges.
		I3,F7.0	P500.	178 Annual flood peak, 500-year recurrence interval.
2				**Basin Characteristics, record 3**
	1	"2"	--	Record identifier.
	2-16	A15	STAID or ISTAID	Station identificatio number. Required.
	17-76			Pairs of basin characteristics index numbers and the value for the basin characteristics. In the order: index description
		I3,F7.0	WRCSKW	179 Skew of logarithms, base 10, of annual peak discharges after outlier and historic-peak adjustments.
		I3,F7.0	WRCMN	180 Mean of logarithms, base 10, of annual peak discharges after outlier and historic-peak adjustments.
		I3,F7.0	WRCSD	181 Standard deviation of logarithms, base 10, of annual peak discharges after outlier and historic-peak adjustments.
		I3,F7.0	YRSPK	196 Number of years of systematic peak flow record.
		I3,F7.0	YRSHPK	197 Number of consecutive years used for historic-peak adjustment of flood frequency data.

APPENDIX C. DATA-SET ATTRIBUTES

Data-set attributes in a Watershed Data Management (WDM) file are used to describe the data sets. Attributes may describe how the data are stored in the data set, where the data were gathered, physical features of the associated data, and statistics computed from the associated data. Over 300 attributes are available for describing data sets. Only a fraction of these attributes are used by PeakFQ, but any attribute may be present in the data set.

Table C.1 contains a list and description of the data-set attributes commonly found in annual peak-flow data sets. Table C.2 contains a list of the attributes that PeakFQ reads and (or) writes or that are commonly associated with annual peak-flow data sets and how these attributes are used by the PeakFQ program. The IOWDM program is used to write peak-flow data and most of these attributes to the data sets, but the attributes can be manually entered using the ANNIE program. The basin characteristic and station header formats are described in Appendices B.2 and B.3. See the IOWDM and ANNIE documentation (Flynn and others, 1995) for additional details. Some of these attributes may be modified each time the PeakFQ program is run. Some of these attributes are output from PeakFQ. Some of these attributes are ignored by PeakFQ.

Table C.1. Attributes associated with annual peak-flow data sets.

Name	Type	Length	Update	Data-set type		Description
				Time	Table	
AGENCY	Char	8	Yes	Opt	Opt	Agency code.
AQTYPE	Char	4	Yes	Opt	Opt	Aquifer type. U - unconfined single aquifer N - unconfined multiple aquifers C - confined single aquifer M - confined multiple aquifers X - mixed multiple aquifers
BASEQ	Real	1	Yes	Opt	Opt	Base discharge, in cubic feet per second.
COCODE	Int	1	Yes	Opt	Opt	County or parish code.
CONTDA	Real	1	Yes	Opt	Opt	Drainage area, in square miles, that contributes to surface runoff.
DAREA	Real	1	Yes	Opt	Opt	Total drainage area, in square miles, including noncontributing areas.
DATUM	Real	1	Yes	Opt	Opt	Reference elevation, to mean sea level.
DSCODE	Int	1	Yes	Opt	Opt	State code of the Geological Survey office that operates the station. Usually the same as the state code (STFIPS).
GUCODE	Char	12	Yes	Opt	Opt	Geologic unit code.
HUCODE	Int	1	Yes	Opt	Opt	Hydrologic unit code (8 digits). These codes are given in the U.S. Geological Survey map series "State Hydrologic Unit Maps," Open-File Report 84-708.
ISTAID	Int	1	Yes	Opt	Opt	Station identification number, as an integer.
J407BQ	Real	1	Yes	Opt	Opt	Base gage discharge.
J407BY	Int	1	Yes	Opt	Opt	Year to begin analysis, used to identify subset of available record.
J407EY	Int	1	Yes	Opt	Opt	Year to end analysis, used to identify subset of available record.
J407GS	Real	1	Yes	Opt	Opt	Generalized skew.
J407HO	Real	1	Yes	Opt	Opt	High outlier discharge criterion.
J407LO	Real	1	Yes	Opt	Opt	Low outlier discharge criterion.
J407NH	Int	1	Yes	Opt	Opt	Number of historic peaks.
J407SE	Real	1	Yes	Opt	Opt	Root mean square error of generalized skew.
J407SO	Int	1	Yes	Opt	Opt	Generalized skew option. -1 - station skew 0 - weighted skew 1 - generalized skew

Table C.1. Attributes associated with annual peak-flow data sets—continued.

Name	Type	Length	Update	Data-set type		Description
				Time	Table	
J407UR	Int	1	Yes	Opt	Opt	Include urban regulated peaks. 1 - no 2 - yes
LATDEG	Real	1	Yes	Opt	Opt	Latitude in decimal degrees.
LATDMS	Int	1	Yes	Opt	Opt	Latitude in degrees, minutes, seconds (dddmmss).
LNGDEG	Real	1	Yes	Opt	Opt	Longitude in decimal degrees.
LNGDMS	Int	1	Yes	Opt	Opt	Longitude in degrees, minutes, seconds (dddmmss).
P1.25	Real	1	Yes	Opt	Opt	Annual flood peak, in cubic feet per second, 1.25-year recurrence interval.
P10.	Real	1	Yes	Opt	Opt	Annual flood peak, in cubic feet per second, 10-year recurrence interval.
P100.	Real	1	Yes	Opt	Opt	Annual flood peak, in cubic feet per second, 100-year recurrence interval.
P2.	Real	1	Yes	Opt	Opt	Annual flood peak, in cubic feet per second, 2-year recurrence interval.
P200.	Real	1	Yes	Opt	Opt	Annual flood peak, in cubic feet per second, 200-year recurrence interval.
P25.	Real	1	Yes	Opt	Opt	Annual flood peak, in cubic feet per second, 25-year recurrence interval.
P5.	Real	1	Yes	Opt	Opt	Annual flood peak, in cubic feet per second, 5-year recurrence interval.
P50.	Real	1	Yes	Opt	Opt	Annual flood peak, in cubic feet per second, 50-year recurrence interval.
P500.	Real	1	Yes	Opt	Opt	Annual flood peak, in cubic feet per second, 500-year recurrence interval.
SITECO	Char	4	Yes	Opt	Opt	Site code. SW - stream SP - spring ES - estuary GW - well LK - lake or reservoir ME - meteorological
STAID	Char	16	Yes	Opt	Opt	Station identification, up to 16 alpha-numeric characters.
STANAM	Char	48	Yes	Opt	Opt	Station name or description of the data set.
STFIPS	Int	1	Yes	Opt	Opt	State FIPS code.
TSTYPE	Char	4	Yes	Opt	Opt	User-defined four-character descriptor. Used to describe the contents of the data set, for example: PRCP, RAIN, SNOW - precepitation FLOW, DISC, PEAK - discharge TEMP, TMIN, TMAX - temperature EVAP, PET - evapotranspiration Some models and application programs may require a specific TSTYPE for data sets they use.
WELLDP	Real	1	Yes	Opt	Opt	Depth of well, in feet. The greatest depth at which water can enter the well.
WRCMN	Real	1	Yes	Opt	Opt	Mean of logarithms, base 10, of annual peak discharges after outlier and historic-peak adjustments.
WRCSD	Real	1	Yes	Opt	Opt	Standard deviation of logarithms, base 10, of annual peak discharges after outlier and historic-peak adjustments.
WRCSKW	Real	1	Yes	Opt	Opt	Skew of logarithms, base 10, of annual peak discharge after outlier and historic-epak adjustments and generalized skew weighting.
YRSHPK	Int	1	Yes	Opt	Opt	Number of consecutive years used for historic-peak adjustment to flood-frequency data.

Table C.2. Sources of attributes associated with peak-flow data sets

WDM file		as processed by IOWDM				PeakFQ use *			
attribute		basin charac		statn header					
name	no.	name	no.	name	rec	i	r	w	p
istaid	51	sta id	--	sta id	N		x		
staid	2	sta id	--	sta id	N	x	x		
stanam	45	sta name	--	sta name	N	x	x		
latdms	54						x		
lngdms	55						x		
latdeg	8	lat gage	22	latitude	H	x	x		
lngdeg	9	lng gage	23	longitude	H	x	x		
yrshpk	81	yrshispk	197				x	x	
j407by	278						x	x	
j407ey	279						x	x	
j407lo	269						x	x	
j407ho	270						x	x	
j407so	271						x	x	
j407gs	272						x	x	
j407bq	273						x	x	
j407se	275						x	x	
j407ur	276						x	x	
j407nh	274						x	x	
wrcmn	78	wrc mn	180					x	x
wrcsd	79	wrc sd	181					x	x
wrcskw	77	wrc skew	179					x	x
p1.25	65	p1,25	75					x	x
p2.	66	p2	76					x	x
p5.	67	p5	77					x	x
p10.	68	p10	78					x	x
p25.	69	p25	79					x	x
p50.	70	p50	80					x	x
p100.	71	p100	81					x	x
p200.	72	p200	82					x	x
p500.	73	p500	178					x	x
meanpk	74	meanpk	83						x
sdpk	75	sdpk	84						x
skewpk	76	skewpk	85						x
tmtopk	98	timetopk	21						
yrspk	80	yrspk	196						x
darea	11	area	1	area	H				
contda	43	contda	2	cont area	H				
agency	40			agency	Z				
stfips	41	state c	--	state code	H				
dscode	42	dist c	--	dist code	H				
siteco	44			site code	H				
hucode	4			hyd unit c	H				
datum	264			datum	H				
welldp	47			well dpth	H				
gucode	46			geol unit	N				
aqtype	48			aquifer tp	N				
baseq	49			base q	Y				

* i - input by user
r - read from WDM
w - written to WDM
p - written to “punch”

APPENDIX D. SAMPLE FILES

Appendix D.1. Sample Specification File

```
I ASCI BigSandy.inp
O File BigSandy.out
O Plot Style Graphics
O Plot Format WMF
Station 03606500
     HistPeriod 77
     SkewOpt Weighted
     GenSkew -0.2
     PlotName 03606500
```

Appendix D.2. Sample Input File

```
H 03606500       3602190881342004747017SW 6040005 205.00         380.58
N 03606500       BIG SANDY RIVER AT BRUCETON HIST B-17-B
2 03606500
3 03606500       189703    250007              18.00
3 03606500       191903    210007              17.00
3 03606500       192612    185007              16.50
3 03606500       19300109   9100               13.98
3 03606500       19310327   2060               11.20
3 03606500       19320113   7820               13.60
3 03606500       19330321   3220               11.95
3 03606500       19331218   5580               12.94
3 03606500       19350121  17000               16.16
3 03606500       19360704   6740               13.28
3 03606500       19370121  13800               14.86
3 03606500       19380123   4270               12.67
3 03606500       19390204   5940               13.23
3 03606500       19400319   1680               10.91
3 03606500       19410802   1200               10.00
3 03606500       19420410  10100               14.52
3 03606500       19430320   3780               12.45
3 03606500       19440218   5340               13.07
3 03606500       19450102   5630               13.13
3 03606500       19460109  12000               14.92
3 03606500       19470104   3980               12.53
3 03606500       19480317   6130               13.31
3 03606500       19481120   4740               12.83
3 03606500       19491213   9880               14.37
3 03606500       19510104   5230               13.01
3 03606500       19511216   4260               12.70
3 03606500       19530519   5000               12.95
3 03606500       19540122   3320               12.32
3 03606500       19550322   5480               13.11
3 03606500       19560130  11800               14.85
3 03606500       19570130   5150               13.00
3 03606500       19571116   3350               12.33
3 03606500       19590216   2400               11.83
3 03606500       19591212   1460               10.94
3 03606500       19610609   3770               12.51
3 03606500       19620228   7480               13.71
3 03606500       19630305   2740               12.02
3 03606500       19640310   3100               12.21
3 03606500       19650212   7180               14.07
3 03606500       19660502   1920               11.64
3 03606500       19670515   9060               14.54
3 03606500       19680404   3080               12.64
3 03606500       19681130   2800               12.50
3 03606500       19700403   4330               13.11
3 03606500       19710824   5080               13.36
3 03606500       19720717  12000               15.14
3 03606500       19730421   7640               14.88
```

Appendix D.3. Sample Output File

```
1
  Program PeakFq            U. S. GEOLOGICAL SURVEY             Seq.000.000
  Ver. 5.0 Beta 8     Annual peak flow frequency analysis      Run Date / Time
  05/06/2005          following Bulletin 17-B Guidelines       04/28/2006 13:03

                         --- PROCESSING OPTIONS ---

                      Plot option         = Graphics device
                      Basin char output   = WATSTORE
                      Print option        = Yes
                      Debug print         = No
                      Input peaks listing = Long
                      Input peaks format  = WATSTORE peak file

                      Input files used:
                         peaks (ascii)  - D:\EX\BIGSANDY.INP
                         specifications - PKFQWPSF.TMP
                      Output file(s):
                         main - D:\EX\BIGSANDY.PRT
                         bcd  - BIGSANDY.BCD

1
  Program PeakFq            U. S. GEOLOGICAL SURVEY             Seq.001.001
  Ver. 5.0 Beta 8     Annual peak flow frequency analysis      Run Date / Time
  05/06/2005          following Bulletin 17-B Guidelines       04/28/2006 13:03

          Station - 03606500  BIG SANDY RIVER AT BRUCETON HIST B-17-B

                     I N P U T   D A T A   S U M M A R Y

                Number of peaks in record            =       47
                Peaks not used in analysis           =        3
                Systematic peaks in analysis         =       44
                Historic peaks in analysis           =        0
                Years of historic record             =        0
                Generalized skew                     =   -0.189
                     Standard error                  =    0.550
                     Mean Square error               =    0.303
                Skew option                          =   WEIGHTED
                Gage base discharge                  =      0.0
                User supplied high outlier threshold =   --
                User supplied low outlier criterion  =   --
                Plotting position parameter          =     0.00


  *********  NOTICE  --  Preliminary machine computations.        *********
  *********  User responsible for assessment and interpretation.  *********

  **WCF109W-PEAKS WITH MINUS-FLAGGED DISCHARGES WERE BYPASSED.       3
  **WCF113W-NUMBER OF SYSTEMATIC PEAKS HAS BEEN REDUCED TO NSYS =   44
    WCF134I-NO SYSTEMATIC PEAKS WERE BELOW GAGE BASE.                    0.0
    WCF195I-NO LOW OUTLIERS WERE DETECTED BELOW CRITERION.             921.3
    WCF163I-NO HIGH OUTLIERS OR HISTORIC PEAKS EXCEEDED HHBASE.     26151.7
    WCF002J-CALCS COMPLETED.  RETURN CODE =  2
```

1

Program PeakFq U. S. GEOLOGICAL SURVEY Seq.001.002
Ver. 5.0 Beta 8 Annual peak flow frequency analysis Run Date / Time
05/06/2005 following Bulletin 17-B Guidelines 04/28/2006 13:03

Station - 03606500 BIG SANDY RIVER AT BRUCETON HIST B-17-B

ANNUAL FREQUENCY CURVE PARAMETERS -- LOG-PEARSON TYPE III

	FLOOD BASE		LOGARITHMIC		
	DISCHARGE	EXCEEDANCE PROBABILITY	MEAN	STANDARD DEVIATION	SKEW
SYSTEMATIC RECORD	0.0	1.0000	3.6909	0.2672	-0.187
BULL.17B ESTIMATE	0.0	1.0000	3.6909	0.2672	-0.188

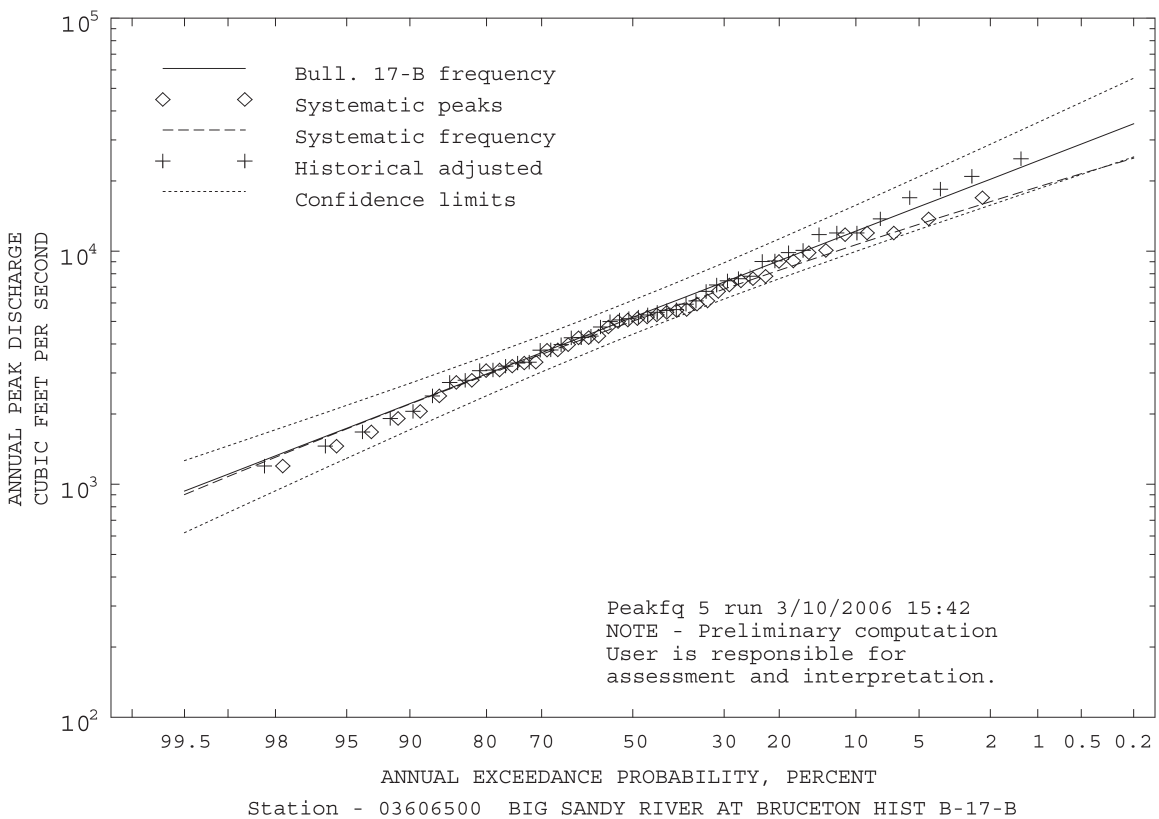
ANNUAL FREQUENCY CURVE -- DISCHARGES AT SELECTED EXCEEDANCE PROBABILITIES

ANNUAL EXCEEDANCE PROBABILITY	BULL.17B ESTIMATE	SYSTEMATIC RECORD	'EXPECTED PROBABILITY' ESTIMATE	95-PCT CONFIDENCE LIMITS FOR BULL. 17B ESTIMATES LOWER	UPPER
0.9950	902.7	903.1	810.3	604.0	1209.0
0.9900	1078.0	1078.0	991.8	746.7	1411.0
0.9500	1728.0	1728.0	1664.0	1306.0	2137.0
0.9000	2206.0	2206.0	2155.0	1736.0	2660.0
0.8000	2943.0	2943.0	2910.0	2415.0	3470.0
0.6667	3827.0	3827.0	3809.0	3229.0	4462.0
0.5000	5004.0	5004.0	5004.0	4288.0	5847.0
0.4292	5580.0	5580.0	5589.0	4790.0	6555.0
0.2000	8278.0	8278.0	8365.0	7017.0	10100.0
0.1000	10660.0	10660.0	10870.0	8855.0	13480.0
0.0400	13840.0	13840.0	14320.0	11200.0	18290.0
0.0200	16310.0	16310.0	17100.0	12960.0	22200.0
0.0100	18850.0	18860.0	20060.0	14720.0	26360.0
0.0050	21480.0	21490.0	23210.0	16500.0	30780.0
0.0020	25080.0	25090.0	27710.0	18890.0	37030.0

```
1


  Program PeakFq            U. S. GEOLOGICAL SURVEY            Seq.001.003
  Ver. 5.0 Beta 8     Annual peak flow frequency analysis      Run Date / Time
  05/06/2005          following Bulletin 17-B Guidelines       04/28/2006 13:03

          Station - 03606500  BIG SANDY RIVER AT BRUCETON HIST B-17-B

                        I N P U T   D A T A   L I S T I N G

     WATER YEAR    DISCHARGE   CODES      WATER YEAR    DISCHARGE   CODES

        1897       -25000.0        H          1951         5230.0
        1919       -21000.0        H          1952         4260.0
        1927       -18500.0        H          1953         5000.0
        1930         9100.0                   1954         3320.0
        1931         2060.0                   1955         5480.0
        1932         7820.0                   1956        11800.0
        1933         3220.0                   1957         5150.0
        1934         5580.0                   1958         3350.0
        1935        17000.0                   1959         2400.0
        1936         6740.0                   1960         1460.0
        1937        13800.0                   1961         3770.0
        1938         4270.0                   1962         7480.0
        1939         5940.0                   1963         2740.0
        1940         1680.0                   1964         3100.0
        1941         1200.0                   1965         7180.0
        1942        10100.0                   1966         1920.0
        1943         3780.0                   1967         9060.0
        1944         5340.0                   1968         3080.0
        1945         5630.0                   1969         2800.0
        1946        12000.0                   1970         4330.0
        1947         3980.0                   1971         5080.0
        1948         6130.0                   1972        12000.0
        1949         4740.0                   1973         7640.0
        1950         9880.0


        Explanation of peak discharge qualification codes

       PEAKFQ    NWIS
        CODE     CODE   DEFINITION

          D        3    Dam failure, non-recurrent flow anomaly
          G        8    Discharge greater than stated value
          X       3+8   Both of the above
          L        4    Discharge less than stated value
          K     6 OR C  Known effect of regulation or urbanization
          H        7    Historic peak

          -  Minus-flagged discharge -- Not used in computation
                -8888.0 -- No discharge value given
          -  Minus-flagged water year -- Historic peak used in computation
```

```
1
  Program PeakFq           U. S. GEOLOGICAL SURVEY           Seq.001.004
  Ver. 5.0 Beta 8     Annual peak flow frequency analysis     Run Date / Time
  05/06/2005          following Bulletin 17-B Guidelines      04/28/2006 13:03

        Station - 03606500  BIG SANDY RIVER AT BRUCETON HIST B-17-B


   EMPIRICAL FREQUENCY CURVES -- WEIBULL PLOTTING POSITIONS

      WATER         RANKED       SYSTEMATIC      BULL.17B
       YEAR       DISCHARGE        RECORD        ESTIMATE

       1935        17000.0         0.0222         0.0222
       1937        13800.0         0.0444         0.0444
       1946        12000.0         0.0667         0.0667
       1972        12000.0         0.0889         0.0889
       1956        11800.0         0.1111         0.1111
       1942        10100.0         0.1333         0.1333
       1950         9880.0         0.1556         0.1556
       1930         9100.0         0.1778         0.1778
       1967         9060.0         0.2000         0.2000
       1932         7820.0         0.2222         0.2222
       1973         7640.0         0.2444         0.2444
       1962         7480.0         0.2667         0.2667
       1965         7180.0         0.2889         0.2889
       1936         6740.0         0.3111         0.3111
       1948         6130.0         0.3333         0.3333
       1939         5940.0         0.3556         0.3556
       1945         5630.0         0.3778         0.3778
       1934         5580.0         0.4000         0.4000
       1955         5480.0         0.4222         0.4222
       1944         5340.0         0.4444         0.4444
       1951         5230.0         0.4667         0.4667
       1957         5150.0         0.4889         0.4889
       1971         5080.0         0.5111         0.5111
       1953         5000.0         0.5333         0.5333
       1949         4740.0         0.5556         0.5556
       1970         4330.0         0.5778         0.5778
       1938         4270.0         0.6000         0.6000
       1952         4260.0         0.6222         0.6222
       1947         3980.0         0.6444         0.6444
       1943         3780.0         0.6667         0.6667
       1961         3770.0         0.6889         0.6889
       1958         3350.0         0.7111         0.7111
       1954         3320.0         0.7333         0.7333
       1933         3220.0         0.7556         0.7556
       1964         3100.0         0.7778         0.7778
       1968         3080.0         0.8000         0.8000
       1969         2800.0         0.8222         0.8222
       1963         2740.0         0.8444         0.8444
       1959         2400.0         0.8667         0.8667
       1931         2060.0         0.8889         0.8889
       1966         1920.0         0.9111         0.9111
       1940         1680.0         0.9333         0.9333
       1960         1460.0         0.9556         0.9556
```

```
       1941         1200.0         0.9778         0.9778
       1927       -18500.0           --             --
       1919       -21000.0           --             --
       1897       -25000.0           --             --
1



 End PEAKFQ analysis.
   Stations processed :       1
   Number of errors   :       0
   Stations skipped   :       0
   Station years      :      47


Data records may have been ignored for the stations listed below.
(Card type must be Y, Z, N, H, I, 2, 3, 4,  or *.)
(2, 4, and * records are ignored.)

 For the station below, the following records were ignored:
2 03606500

 FINISHED PROCESSING STATION:  03606500            BIG SANDY RIVER AT BRUCETON H


 For the station below, the following records were ignored:

 FINISHED PROCESSING STATION:
```



Note: Graphic from a [wmf] file, curves enhanced for publication.

USGS

Flynn, Kathleen M., Kirby, William H., and Hummel, Paul R.—**User's Manual for Program PeakFQ**—Techniques and Methods Book 4, Chapter B4