

Estimating Magnitude and Frequency of Floods Using the PeakFQ 7.0 Program

Flood-Frequency Analysis

Flood-frequency analysis provides information about the magnitude and frequency of flood discharges based on records of annual maximum instantaneous peak discharges collected at streamgages. The information is essential for defining flood-hazard areas, for managing floodplains, and for designing bridges, culverts, dams, levees, and other flood-control structures.

Bulletin 17B (B17B) of the Interagency Advisory Committee on Water Data (IACWD; 1982) codifies the standard methodology for conducting flood-frequency studies in the United States. B17B specifies that annual peak-flow data are to be fit to a log-Pearson Type III distribution. Specific methods are also prescribed for improving skew estimates using regional skew information, tests for high and low outliers, adjustments for low outliers and zero flows, and procedures for incorporating historical flood information.

The authors of B17B identified various needs for methodological improvement and recommended additional study. In response to these needs, the Advisory Committee on Water Information (ACWI, successor to IACWD; <http://acwi.gov>), Subcommittee on Hydrology (SOH), Hydrologic Frequency Analysis Work Group (HFAWG), has recommended modest changes to B17B. These changes include adoption of a generalized method-of-moments estimator denoted the Expected Moments Algorithm (EMA) (Cohn and others, 1997) and a generalized version of the Grubbs-Beck test for low outliers (Cohn and others, 2013). The SOH requested that the USGS implement these changes in a user-friendly, publicly accessible program.

A Brief Introduction to the PeakFQ Program

The Peak flow FreQuency (PeakFQ 7.0) analysis program, which runs interactively under the Windows¹ Operating System, implements both the existing Bulletin 17B and the HFAWG proposed EMA procedures for flood-frequency analysis of streamflow records. Single and multiple Grubbs-Beck outlier screening is available for both procedures. [Users are cautioned that the ACWI has not yet approved EMA or the multiple Grubbs-Beck outlier screening for standardized use, pending a public comment period.]

¹Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

PeakFQ Input Options

Annual peak-flow data must be supplied to PeakFQ in a standard WATSTORE text formatted file (Flynn and others, 2006). The user specifies processing options interactively or by supplying a program specification file (.psf), which also can be used to identify the file containing the peaks.

PeakFQ Output Options

PeakFQ provides estimates of flood magnitudes and their corresponding variance for 15 annual exceedance probabilities. The output file also provides estimates of the parameters of the log-Pearson Type III frequency distribution, including the logarithmic mean, standard deviation, skew, and mean square error of the skew. PeakFQ can also provide a graph (fig. 1) displaying the fitted frequency curve, systematic peaks, low outliers, censored peaks, interval peaks, historic peaks, thresholds, and confidence limits.

Flood-Frequency Analysis Methodology

Following the approach recommended in B17B, PeakFQ fits the log-Pearson Type III distribution to the logarithms of annual peak discharges, using the method-of-moments to compute mean, standard deviation, and skew of the log-transformed data. PeakFQ provides the user the option to improve the station skew estimate by computing a weighted average with a generalized/regional skew estimate

obtained from B17B or other sources. The station and generalized skews are weighted to reflect the relative accuracy of each estimate.

EMA addresses several methodological concerns identified in B17B, while retaining the essential structure and moments-based approach of the existing B17B procedures for determining flood frequency. EMA can accommodate interval data, which simplifies analysis of datasets containing censored observations, historic and (or) paleo data, low outliers, and uncertain data points, while also providing enhanced confidence intervals on the estimated discharges.

Unlike B17B, which recognizes two categories of data—systematic peaks (annual peaks observed in the course of the systematic streamgaging at the station) and historic peaks (records of floods that occurred outside the period of regular streamgaging)—EMA employs a more general description of flood information from the historical period that includes both systematic and historic peaks. For every year Y during the historical period, it is assumed that there was a peak discharge Q_Y , regardless of whether this discharge was recorded. In the framework of EMA, the hydrologist's knowledge of the peak flow Q_Y is described by the flow interval ($Q_{Y,lower}$; $Q_{Y,upper}$). When running EMA, a flow interval must be specified for each year in the historical record, including any gaps for which no discharge observation is recorded, as well as for censored and interval peaks.

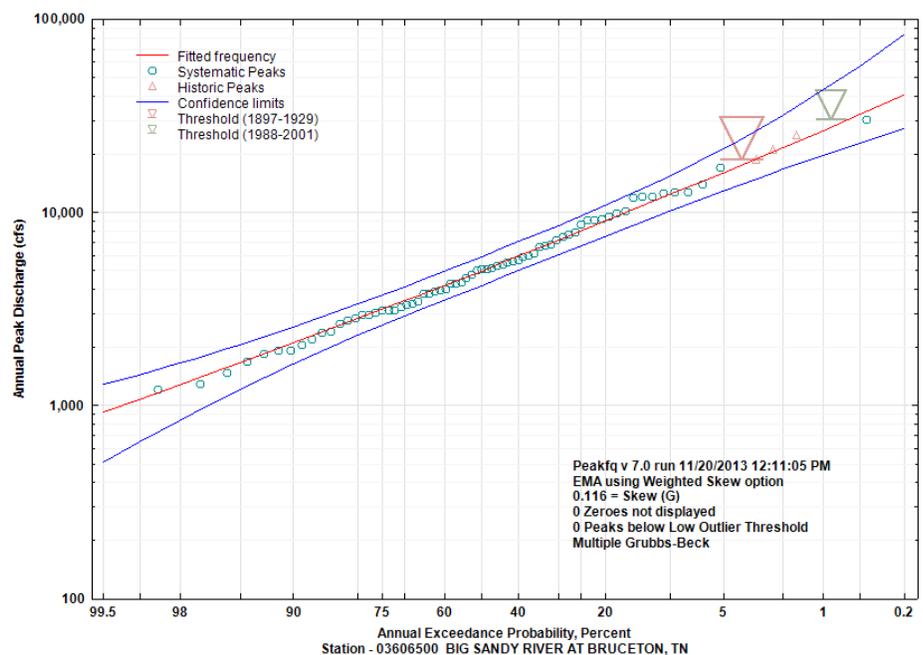


Figure 1. Example of PeakFQ probability graph of annual peaks, fitted frequency curve, confidence limits, and thresholds.

For both B17B and EMA procedures, the user enters into PeakFQ the sampling properties of each recorded peak flow. B17B differentiates between sampling properties of observations taken when the entire range of flows would have been recorded (that is, systematic data) and peak flows that have been recorded because their magnitudes led to the creation of a permanent record (that is, historic data; Stedinger and Cohn, 1986). The authors of B17B recognized that these two types of data need to be treated differently. EMA distinguishes among sampling properties by employing perception thresholds denoted ($T_{Y,lower}$; $T_{Y,upper}$), which reflect the range of flows that would have been measured/recorded had they occurred. Perception thresholds describe the range of measurable potential discharges and are independent of the actual peak discharges that have occurred. The lower bound, $T_{Y,lower}$, represents the smallest peak flow that would result in a recorded flow, while the upper bound, $T_{Y,upper}$, represents the largest peak flow that would result in a measured flow. Defining the thresholds is often complicated; for more information, see Cohn and others (1997) and Cohn and others (2001).

For every year in the historical period, EMA uses both a flow interval and a perception threshold. It is the user's responsibility to ensure that the flow interval and perception threshold for any year in the historical period are consistent. PeakFQ requires documentation for any addition or modification to the default flow intervals and (or) perception thresholds.

B17B recommends the use of the Grubbs-Beck test (Grubbs and Beck, 1972) to detect low outliers in flood-frequency analysis. As described by Cohn and others (2013), the multiple Grubbs-Beck test is a generalization of the Grubbs-Beck method that allows for a standard procedure for identifying multiple Potentially Influential Low Flows (PILFs). In flood-frequency analysis, PILFs are annual peaks that meet three criteria: their magnitude is much smaller than the flood quantile of interest; they occur below a statistically significant break in the flood-frequency data plot; and they have excessive influence on the estimated frequency of large floods. When an observation is identified as a PILF, all values smaller than that flood are also categorized as PILFs. Identifying PILFs and recording them as censored peaks can greatly improve estimator robustness with little or no loss of efficiency. Thus, the use of the multiple Grubbs-Beck test can improve the fit of the small annual exceedance probabilities, while minimizing lack-of-fit due to unimportant PILFs in an annual peak series (Cohn and others, 2013).

Using PeakFQ

PeakFQ is launched by double-clicking on the PeakFQ icon on the desktop or by selecting USGS and then PeakFQ under Programs in the Windows Start menu. The File pull-down menu at the top of the window

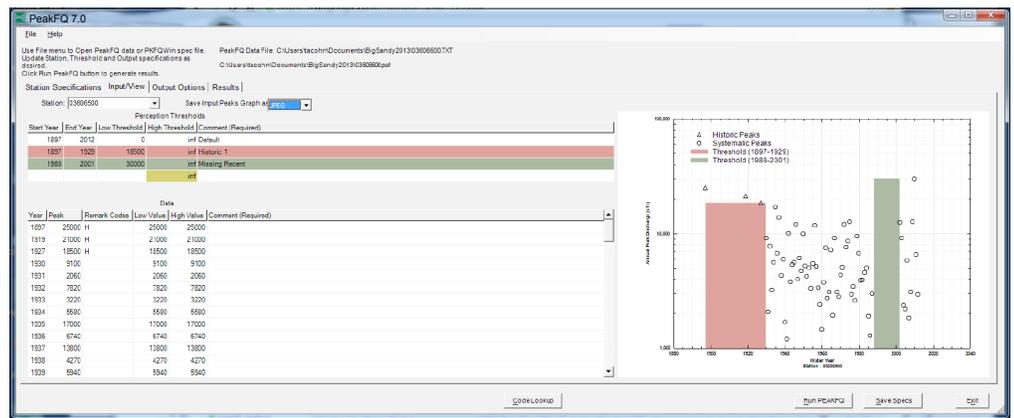


Figure 2. Input/View tab in main window of PeakFQ.

(fig. 2) contains options to *Open* an input file, *Save Specs* to a .psf file, or *Exit* the program. The main window contains four tabs. The *Station Specifications* tab, the *Input/View* tab, and the *Output Options* tab become active once an input file has been identified and read. The *Run PeakFQ* button at the bottom of the window runs the analysis. The *Results* tab becomes active after the Run PEAKFQ button has been selected. The list below highlights the main elements of the program.

- **Station Specifications:** The *Station Specifications* tab is populated with data read from the file of peak flows and, if used, the .psf file. Default values are used for parameters not specified in the input files. The gray shaded columns are informational and cannot be modified. Values in the white columns can be modified as appropriate for the analysis.
- **Input/View:** The *Input/View* tab (fig. 2) provides a representation of the data in both tabular and graphic form. The *Input/View* tab is populated with data read from the input file, as well as the *Station Specifications* tab. The *Perception Thresholds* table and the *Low Value* and *High Value* entries of the *Data* table apply only to an EMA analysis.
- **Output Options:** The *Output Options* tab is used to specify file and graphics output choices and to name the text output files.
- **Run PeakFQ:** The *Run PeakFQ* button runs the analysis.
- **Results:** The *Results* tab becomes active after the analysis has been completed and provides access to the text output files and the plots generated by the analysis.
- **Save Specs:** The *Save Specs* button at the bottom of the window is used to save the specifications and output options to a .psf file. A .psf file can be edited and used as an input file to rerun the analysis, facilitating evaluation of different parameters.

Software, Data Sources, and Additional Information

The current version of the PeakFQ program is available at <http://water.usgs.gov/software/PeakFQ/>. Future versions of the program as well as a complete user's manual,

which is under development, can be accessed at this location. USGS peak-flow data can be retrieved in the WATSTORE text format at the USGS National Water Information System at <http://nwis.waterdata.usgs.gov/usa/nwis/peak>.

For information on local USGS flood-frequency studies and research, contact the USGS Water Science Center Director for the area of interest. A contact list is available at http://water.usgs.gov/district_chief.html.

References

- Cohn, T.A., England, J.F., Berenbrock, C.E., Mason, R.R., Stedinger, J.R., and Lamontagne, J.R., 2013, A generalized Grubbs-Beck test statistic for detecting multiple potentially influential low outliers in flood series: *Water Resources Research*, v. 49, no. 8, p. 5047–5058.
- Cohn, T.A., Lane, W.L., and Baier, W.G., 1997, An algorithm for computing moments-based flood quantile estimates when historical flood information is available: *Water Resources Research*, v. 33, no. 9, p. 2089–2096.
- Cohn, T.A., Lane, W.L., and Stedinger, J.R., 2001, Confidence intervals for Expected Moments Algorithm flood quantile estimates: *Water Resources Research*, v. 37, no. 6, p. 1695–1706.
- 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, chap. B4, 42 p.
- Grubbs, F.E., and Beck, G., 1972, Extension of sample sizes and percentage points for significance test of outlying observations: *Technometrics*, v. 10, p. 211–219.
- Interagency Advisory Committee on Water Data (IACWD), 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.
- Stedinger, J.R., and Cohn, T.A., 1986, Flood frequency analysis with historical and paleoflood information: *Water Resources Research*, v. 22, no. 5, p. 785–793.

Prepared by Andrea G. Veilleux, Timothy A. Cohn, Kathleen M. Flynn, Robert R. Mason, Jr., and Paul R. Hummel