Guidelines for Determining Flood Flow Frequency
Bulletin 17C

Techniques and Methods 4–BXX

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

DRAFT: December 29, 2015

Cover. International bridge over the St. John River in Fort Kent, Maine at Station 01014000, during flood of April 30, 2008. Photo by M. Huard, USGS.

Guidelines for Determining Flood Flow Frequency Bulletin 17C

By John F. England Jr., Timothy A. Cohn, Beth A. Faber, Jery R. Stedinger, Wilbert O. Thomas Jr., Andrea G. Veilleux, Julie E. Kiang, and Robert R. Mason

Techniques and Methods 4–BXX

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

U.S. Department of the Interior
SALLY JEWELL, Secretary

U.S. Geological Survey
Suzette Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2015

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment, visit http://www.usgs.gov or call 1-888-ASK-USGS

For an overview of USGS information products, including maps, imagery, and publications, visit http://www.usgs.gov/pubprod

To order this and other USGS information products, visit http://store.usgs.gov

Suggested citation:
England Jr., J.F., Cohn, T.A., Faber, B.A., Stedinger, J.R., Thomas Jr., W.O., Veilleux, A.G., Kiang, J.E., and Mason, R.R., 2015, Guidelines for Determining Flood Flow Frequency – Bulletin 17C: U.S. Geological Survey Techniques and Methods 4–BXX, XXX p., at http://dx.doi.org/10.3133/tm4–BXX/.

Contents

Figures

Tables

Conversion Factors

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

Datum

Horizontal coordinate information is referenced to the North American Datum of 1983 (NAD 83).
Vertical coordinate information is referenced to the National Geodetic Vertical Datum of 1929 (NGVD 29).

Guidelines for Determining Flood Flow Frequency Bulletin 17C

By John F. England Jr., Timothy A. Cohn, Beth A. Faber, Jery R. Stedinger, Wilbert O. Thomas Jr., Andrea G. Veilleux, Julie E. Kiang, and Robert R. Mason

Abstract

Accurate estimates of flood frequency and magnitude are a key component of any effective nationwide flood risk management and flood damage abatement program. In addition to accuracy, methods for estimating flood risk must be uniformly and consistently applied because management of the Nation's water and related land resources is a collaborative effort involving multiple actors including most levels of government and the private sector.

Flood frequency guidelines have been published in the United States since 1967, and have undergone periodic revisions. In 1967, the U.S. Water Resources Council presented a coherent approach to flood frequency with Bulletin 15 (USWRC, 1967), "A Uniform Technique for Determining Flood Flow Frequencies." The method it recommended involved fitting the Log-Pearson Type III distribution to annual peak flow data by the method-of-moments. The first extension and update of Bulletin 15 was published in 1976 as Bulletin 17 (USWRC, 1976), "Guidelines for Determining Flood Flow Frequency" (*Guidelines*). It extended the Bulletin 15 procedures by introducing methods for dealing with outliers, historical flood information, and regional skew. Bulletin 17A was published the following year to clarify the computation of weighted skew. The next revision of the Bulletin, 17B (IACWD, 1982), provided a host of improvements and new techniques designed to address situations that often arise in practice, including better methods for estimating and using regional skew, weighting station and regional skew, detection of outliers, and use of the conditional probability adjustment (CPA) (Thomas, 1985; Griffis and Stedinger, 2007a).

The current version of the *Guidelines* are presented in this document, denoted Bulletin 17C. It incorporates changes motivated by four of the items listed as "Future Work" in Bulletin 17B and 30 years of post-17B research on flood processes and statistical methods. The updates include: adoption of a generalized representation of flood data that allows for interval and censored data types; a new method, called the Expected Moments Algorithm (Cohn et al., 1997, 2001), that extends the method-of-moments so that it can accommodate interval data; a generalized approach to identification of low outliers in flood data (Cohn et al., 2013); and an improved method for computing confidence intervals.

Federal agencies are requested to use these guidelines in all planning activities involving water and related land resources. State, local and private organizations are encouraged to use these guidelines to assure uniformity in the flood-frequency estimates that all agencies concerned with flood risk should use for Federal planning decisions.

This revision is adopted with the knowledge and understanding that review of these procedures will be ongoing. Updated methods will be adopted when warranted by experience and by examination and testing of new techniques.

Introduction

These *Guidelines* describe the data and procedures for computing flood flow frequency where systematic stream gaging records of sufficient length (at least 10 years) to warrant statistical analysis are available. The procedures do not cover watersheds where flood flows are appreciably altered by reservoir regulation or where the possibility of unusual events, such as dam failures, must be considered.

Background

In December 1967, Bulletin No. 15, "A Uniform Technique for Determining Flood Flow Frequencies", was issued by the Hydrology Committee of the Water Resources Council (USWRC, 1967). The report recommended use of the Pearson Type III distribution with log transformation of the data (log Pearson Type III distribution) as a base method for flood flow frequency studies. As pointed out in that report, further studies were needed covering various aspects of flow frequency determinations.

In March 1976, Bulletin 17, "Guidelines for Determining Flood Flow Frequency" was issued by the Water Resources Council (USWRC, 1976). The guide was an extension and update of Bulletin No. 15. It provided a more complete guide for flood flow frequency analysis incorporating currently accepted technical methods with sufficient detail to promote uniform application. It was limited to defining flood potentials in terms of peak discharge and exceedance probability at locations where a systematic record of peak flood flows is available. The recommended set of procedures was selected from those used or described in the literature prior to 1976, based on studies conducted for this purpose at the Center for Research in Water Resources of the University of Texas at Austin (Beard, 1974) that are summarized in IACWD (1982, Appendix 14) and other studies by the Work Group on Flood Flow Frequency.

The Guidelines were revised and reissued in June 1977 as Bulletin l7A, which clarified the procedure for computing weighted skew. Bulletin 17B is the next effort to improve and expand upon the earlier publications. Bulletin 17B was issued in 1981, and re-issued with minor corrections in 1982 (IACWD, 1982). Bulletin 17B provided revised procedures for weighting station skew values with results from a generalized skew study, detecting and treating outliers, making two station comparisons, and computing confidence limits about a frequency curve. Thomas (1985) and Griffis and Stedinger (2007a) present additional details on the history of Bulletins 17, 17A, and 17B.

In 2005, the Hydrologic Frequency Analysis Work Group (HFAWG), under the Subcommittee on Hydrology (SOH), began discussing recent research on flood frequency and potential significant revisions to Bulletin 17B. The HFAWG submitted a plan to SOH in 2006 (Hydrologic Frequency Analysis Work Group, 2006) to conduct studies on flood frequency improvements. The focus was on evaluating a generalized method of moments approach (Cohn et al., 1997), with tests on gaging station peak-flow data and with Monte-Carlo simulation (Cohn et al., 2014). New procedures were developed to deal with troublesome data sets, and new methods were extensively tested with selected data sets and in Monte Carlo studies (Cohn et al., 2014). In 2013, the HFAWG made recommendations to SOH to revise Bulletin 17B (Hydrologic Frequency Analysis Work Group, 2013). Additional background on revision efforts is available on the HFAWG webpage at http://acwi.gov/hydrology/Frequency/minutes/index.html. Appendix 1 listes HFAWG and SOH members involved in the study and revision effort.

This document is an update to the guidelines published earlier in Bulletins 17, 17A and 17B. Revisions incorporated in this document address major limitations of Bulletin 17B. Most of these limitations were well known, and are listed in Bulletin 17B (IACWD, 1982) on pp. 27-28 as topics needing future study.

A particularly important innovation in these new guidelines is elimination of the need, implicit in application of Bulletin 17B, that all annual peaks be either point-value flow estimates, or upper bounds on historical flows, or on low-flows and zeros. With new statistical and computational procedures, these *Guidelines* employ a new comprehensive data framework; flood data are now generalized as "interval estimates" that incorporate both standard point-value flood observations, as well as upper bound, lower bounds, or simple interval estimates describing the value of the peak flood in each year.

These *Guidelines* take advantage of the new data framework by utilizing the Expected Moments Algorithm (*EMA*) to analyze available flood data in a single, uniform and consistent framework that does not require the introduction of additional algorithms to adjust the flood-frequency curve to incorporate or account for the presence in the dataset of historic information, zero flows, or low outliers as is the case with Bulletin 17B. Thus it avoids the need for arbitrary selection of a sequence of such adjustments described on pages 12-2 through 12-4 of Bulletin 17B.

These *Guidelines* improve on Bulletin 17B by introducing a standardized Multiple Grubbs-Beck test to identify potentially influential low flood observations (PILFs) which can be given special treatment to prevent their exerting excessive influence on the fitting of the flood-frequency curve. This is a very important addition because the new procedure provides clear, reasonable and an objective steps for the identification of such PILFs.

In addition, these *Guidelines* improve on procedures for estimating regional skewness estimators and their precision, thus replacing the map provided in Plate 1 of Bulletin 17B. The recommended procedure employs Bayesian GLS regionalization concepts to develop improved estimates of regional flood skew reflecting the precision of available estimates, their cross-correlation, and the precision of the regional model.

Finally, taken together the use of the interval-data framework, *EMA*, and Bayesian skew coefficient regionalization permits development of a more accurate estimation of confidence intervals about the flood-frequency-curve than do procedures described in Bulletin 17B. Large differences in confidence intervals may be observed between intervals computed with Bulletin 17B and procedures in these *Guidelines* (Bulletin 17C) because the Bulletin 17B confidence intervals ignored uncertainty in the estimated skewness coefficient, and had no provision for recognizing the value of historical information.

Purpose and Scope

The present *Guidelines* incorporate updated flood frequency methods based on recent research summarized by Stedinger and Griffis (2008), concepts described by England Jr and Cohn (2007, 2008), testing by Cohn et al. (2014), and a substantial body of literature over the past 30 years cited thoughout this document (see References). These updated methods address some of the recommended research and limitations in Bulletin 17B. The following important improvements include:

1. the ability to accommodate a generalized form of peak-flow data, specifically interval estimates of peak discharge magnitudes;

2. a generalization of the method-of-moments that can accommodate interval, censored, and binomial-censored data called the Expected Moments Algorithm (*EMA*) (Cohn et al., 1997);

3. accurate confidence interval formulas that can account for historical and paleoflood information as well as regional skew information (Cohn et al., 2001; Cohn, 2015); and

4. a generalized low-outlier procedure, based on the existing Grubbs-Beck test, called the Multiple Grubbs-Beck Test (*MGBT*), that can identify multiple potentially-influential low floods in the peak flow dataset (Cohn et al., 2013).

These *Guidelines* are divided into nine sections which are summarized below.

Flood Flow Frequency Information – The following categories of flood data are recognized: systematic records, historical data, paleoflood and botanical data, regional information, comparison with similar watersheds, and flood estimates from precipitation. Common data issues and representation of data using intervals and thresholds are presented. How each can be used to define the flood potential is briefly described.

Data Assumptions and Specific Concerns – A brief discussion of basic data assumptions is presented as a reminder to those developing flood flow frequency curves to be aware of potential data issues and concerns. Flow measurement error, randomness of events, trends, long-term persistence, mixed populations, watershed changes, and climate variability are briefly discussed.

Determination of the Flood Flow Frequency Curve – This section provides guidance for determination of a frequency curve. The Pearson Type

III distribution with log transformation of the flood data (log-Pearson Type III) is recommended as the basic distribution for defining the annual flood series (USWRC, 1967; IACWD, 1982; Griffis and Stedinger, 2007b). The method of moments with the Expected Moments Algorithm is used to estimate the parameters of the distribution from station data, including historical and paleoflood data, when available. Adjustments are made for potentially-influential low floods. Regional information is used to estimate the skew coefficient. Optional record extension methods using nearby stations is presented. Statistical uncertainty in flood-quantile estimates, including the construction of confidence interval, is described.

Estimating Regional Skew – The general procedure that is recommended to estimate a regional skew is described.

Comparisons of Frequency Curves – Some concepts are described for making comparsions of frequency curves estimated using the procedures in this guide to those from similar watersheds and flood estimates from precipitation. In some situations, a weighted combination of frequency curves may provide an improved estimate.

Software and Examples – Software to estimate frequency curves and examples demonstrating the use of these procedures are described.

Future Studies – Recommended future studies are listed, including methods for ungaged sites and for regulated frequency and urbanization situations.

Applicability of These Guidelines – The applicability of these *Guidelines* and some limitations are discussed in this section.

Appendix – The appendixes provide information on data sources, procedures for initial data analysis, the methods and some computational details for the recommended procedures, flood frequency examples that implement the recommended procedures, and a Glossary.

It is possible to standardize many elements of flood frequency analysis. These *Guidelines* describe each major element of the process of defining the flood potential at a specific location in terms of peak discharge and annual exceedance probability (*AEP*). Flood quantiles with *AEP* ranges from 0.10 to about 0.002 are estimated using annual maximum flood series and methods described here. These estimates depend on the data used in the analysis. When longer historical and paleoflood records are used ($>$ 1,000 years), floods with *AEPs* < 0.002 can be estimated. Use is confined to stations where available records are adequate to allow reliable statistical analysis of the data. Special situations may require other approaches. In those cases where the procedures of this guide are not followed, deviations must be supported by appropriate study and accompanied by a comparison of results using the recommended procedures.

Flood records are limited. As more years of record become available at each location, the determination of flood potential may change. Thus, an estimate may be outdated a few years after it is made. Additional flood data alone may be sufficient reason for a fresh assessment of the flood potential. When making a new assessment, the analyst should incorporate in their study a review of earlier estimates. Where differences appear, they should be acknowledged and explained.

Risk Accumulates

It is important to realize that the probabilities computed here correspond to the annual exceedance probability, or the probability in any year that a flood threshold is exceeded. However, when considering the chance that homes, stores, factories and other public and private facilities are flooded, owners and occupants should consider the likelihood of flooding not just in a single year, but the chance over 10, 25 or even 100 years. Such permanent facilities are generally built with design lives, corresponding to a planning horizon, of 25 or more years.

As used in this guide, risk is defined as the probability that one or more events will exceed a given flood magnitude within a specified period of years n. Assuming the flow frequency curve is accurate and that events from year-to-year are independent, the probability p_n that a damage threshold is exceeded at least once in an n-year period is (Yen, 1970; Kite, 1988):

$$p_n = 1 - (1-p)^n \qquad (1)$$

where p is the annual exceedance probability (*AEP*) for each year.

Thus, given the probability that a threshold has an *AEP* of 0.01 (or 1%), over a 25 year period there is a 22% chance of the threshold being exceeded, over a 50 year period there is a 40% chance of the threshold being exceeded, and over a 100 year period there is a 63% chance of the threshold being exceeded. Or viewed another way, a new home or business that is protected to have only a 1% chance of being flooded in a single year has a 26% probability of being flooded over the life of a 30-year mortgage. Thus, there is a 1 in 4 chance the property will be flooded in that time period. While the probability of flooding in a single year may seem small when the *AEP* is just 1% or less, the chance of flooding accumulates over time so that the probability of flooding over 25 or 50 years is substantial. A full risk analysis that includes uncertainty (National Research Council, 2000) is an addition that could be considered, but is beyond the scope of these *Guidelines*.

Acknowledgments

These revised *Guidelines* were developed under the auspices of the Hydrologic Frequency Analysis Work Group (HFAWG), under the Subcommittee on Hydrology (SOH), of the Advisory Committee on Water Information (ACWI). HFAWG and SOH Work Group Members and participants in this revision are listed in Appendix 1.

Flood Flow Frequency Information

When developing a flood flow frequency curve, the analyst should consider all available information. The general types of data and information which can be included in the flood flow frequency analysis are described in the following sections, as well as how to best characterize available data. Flood frequency analysis relies primarily on systematic records, which typically can be represented as point observations. Other types of data, such as historical and paleoflood data, may be represented with intervals or thresholds, because the magnitudes of flood peaks might be known with less precision. The analyst also needs to consider the use of regional information and flood estimates from precipitation. Specific applications are discussed in subsequent sections of this guide.

Use of Annual Maximum Series

Flood events can be analyzed using either annual maximum series (AMS) or partial-duration series (PDS). The annual maximum flood series is based on the instantaneous maximum flood peak for each year. Annual maximum mean daily discharge or annual maximum *n-day* flood volumes (U.S. Army Corps of Engineers, 1993; Lamontagne et al., 2012) may also be considered, depending on the intended use of the flood-frequency relationship. A partial-duration series is obtained by taking all flood peaks equal to or greater than a predefined base flood. Thus an n-year record can produce m peaks with $m > n$.

Flood frequency estimates using these *Guidelines* are appropriate for the 0.10 *AEP* or less flood ($Q_p > Q_{0.10}$). The annual maximum flood series provides a satisfactory sample for this type of analysis. There is little difference in *AEP* estimates using AMS or PDS for these quantiles (Langbein, 1949). The AMS is also used due to widespread availability and extended length of AMS data. There are limited PDS records and challenges in defining PDS threshold(s) (Madsen et al., 1997).

If minor floods are of interest, with $Q_p \leq Q_{0.10}$ *AEP*, a partial-duration series may be appropriate. The PDS base is selected to assure that all events of interest are evaluated. A major problem encountered when using a partial-duration series is to define flood events to ensure that all events are independent. It is common practice to establish an empirical basis for separating flood events (Lang et al., 1999). The basis for separation will depend upon the investigator and the intended use. No specific guidelines are recommended for defining flood events to be included in a partial series.

Beard (1974) sought to determine if a consistent relationship existed between the annual and partial series which could be used to convert from the annual to the partial-duration series. Based on that work, it is recommended that the partial-duration series be developed from observed data. An alternative but less desirable solution is to convert from the annual to the partial-duration series using a factor.

The procedures described in this guide apply to the annual maximum flood series. If minor flood estimates are needed ($Q_p > Q_{0.10}$), a frequency analysis such as peaks-over-threshold (Stedinger et al., 1993; Coles, 2001) using partial-duration data may be appropriate. No specific guidelines are recommended for conducting a partial-duration frequency analysis.

Data Sources for a Site

The main data sources that are recommended for use in flood frequency include systematic records, historical flood information, and paleoflood and botanical information. These at-site flood data are briefly described; additional information on data sources is in Appendix 2. Refer to the Glossary for data-related definitions and notation.

Systematic Records

Systematic flood data consist of annual peak discharge data collected at regular, prescribed intervals at a gaging station (Salas et al., 1994; Wahl et al., 1995). Systematic flood data involve the continuous monitoring of flood properties by hydrologists (Rantz and Others, 1982; Baker, 1987). In the United States, the U.S. Geological Survey operates and maintains a nationwide gaging station network (Wahl et al., 1995), and is the primary source for systematic flood data. Stream gages are also operated by federal agencies (e.g., Bureau of Reclamation, U.S. Army Corps of Engineers), state agencies (e.g., California, Colorado), local agencies and private enterprises.

The data typically used for flood frequency analysis consist of annual peak discharge values or peak discharges above a base value (partial-duration series). Most annual peak records are obtained either from a continuous trace of river stages or from periodic observations provided by a crest-stage gage (Figure 1). Crest-stage records may provide information only on peaks above some pre-selected base. The records are usually continuous, although missing data or zero flow years may be present. A statistical analysis of these data is the primary basis for the determination of the flow-frequency curve for each station. A major portion of these data are available in the U.S. Geological Survey National Water Information System (NWIS) and other electronic files; additional information in published or unpublished form is available from many sources (Appendix 2).

Figure 1. Photograph of a streamflow-gaging station showing a water-stage recorder, sharp-crested weir and crest-stage gage at U.S. Geological Survey station 01589238, Gwynns Falls Tributary at McDonogh, Maryland.

Historical Flood Information

At many locations, particularly where people have occupied the flood plain for an extended period, or where civil works projects have been constructed by Federal agencies, there is information about major floods which occurred either before or after the period of systematic data collection. Similar information may be available at sites where the gage has been discon-

Figure 2. Historic flood high-water marks and flood of March 13-15, 2010, Potomac River at Great Falls Park, Virginia, upstream of U.S. Geological Survey streamflow-gaging station 01646580, Potomac River at Chain Bridge, at Washington, DC.

tinued, or where records are broken or incomplete. Data for recent floods that occurred outside the systematic data collection period are also treated as historical floods. This historical flood information can often be used to make estimates of peak discharge. It also may define an extended period during which the largest floods, either recorded or historic, are known. In many cases, people make a physical mark, that represents the approximate high-water mark of a flood, on a relatively permanent surface (Figure 2). The high-water mark elevation must be tied to a known datum in order to determine the peak discharge from a stage-discharge relation established after the flood.

Historical data are valuable information that are used in frequency analysis as follows. Let n_s denote the number of years in the systematic (gage) record, n_h be the number of years in the historical period and n be the total period of record, where $n_s + n_h = n$. Let T_h represent a discharge perception threshold that describes the knowledge that flood magnitudes exceeded this level, or were less than this level, during the historical period (Figure 3). The historical flood data are represented by the historic (e_h) peaks and the systematic (e_s) peaks that exceed the threshold T_h during the total flood period n. There is also knowledge that, during the historical period n_h, there are many years that no flood exceeded T_h (indicated with grey shading in Figure 3). The total number of floods that exceed the perception threshold is k, where $k = e_s + e_h$. The section Data Representation using Flow Intervals and Perception Thresholds discusses the determination of the historical period n_h and estimation of perception threshold(s) T_h.

Historical data for flood frequency typically consists of three types, that can extend the temporal information on flood magnitudes:

- large flood estimates prior to (outside of) the gaging station record (Figure 4);
- an extraordinary large flood and knowledge that one (or more) floods within the gaging record are actually the largest in a longer time period n than that of the gaging station record n_s (Figure 5);
- knowledge that floods did not exceed some value T_h (non-exceedance information) over a longer time n_h (Figure 6).

An example is used to illustrate each situation. In the first case, there are three historical floods that occurred prior to the establishment of the gaging station record. It is known that these floods exceeded a perception threshold of 18,000 ft^3/s. These three floods are the largest on record, extend the observational record by 35 years (1895-1929), and are the most important for estimating flood frequency (Figure 4). In the second case, there is one extraordinary flood that occurred in June, 1965 (Matthai, 1969, p. B39). This extraordinary flood is the largest in the 48-year gaging record (1948-1989), and there is historical flood and paleoflood information that indicates this flood might be the largest in over 900 years (Osterkamp and Costa, 1987) (Figure 5), rather than the largest in 48 years. Additional discussion for this extraordinary flood situation is in the Section Extraordinary Floods. In the third case, one has information from a physical feature, such as a bridge or river terrace, that no floods have exceeded a perception threshold. From detailed investigation of river terraces along the North Platte River near Seminoe dam, floods have not exceeded

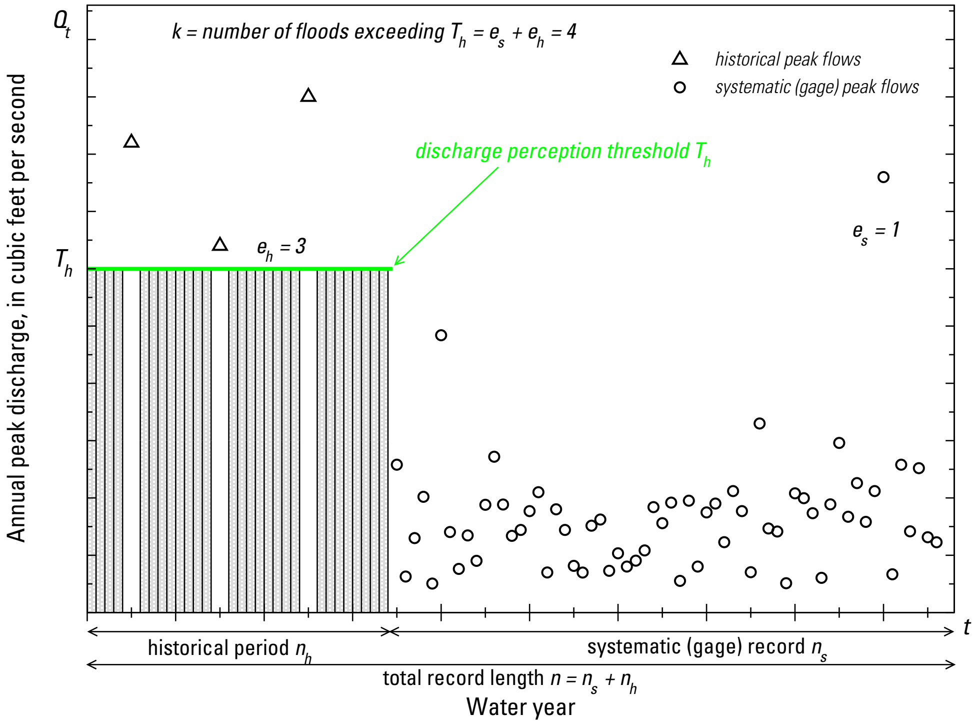

Figure 3. Example peak discharge time series with a historical period and a discharge perception threshold T_h. The grey shaded area represents floods of unknown magnitude less than T_h during the historical period n_h. Black vertical bars during the historical period represent flow intervals for each year for the unrecorded observations. Perception threshold T_h is shown as a green line. Historical floods that exceed the perception threshold (three years) are shown as black triangles. Systematic (gage) peak flows are shown as black circles.

45,000 ft^3/s in the past 7,000 years (Levish, 2002; Levish et al., 2003) (Figure 6). Additional discussion for this situation is in the Section Paleoflood and Botanical Information.

The USGS includes some historical flood information in its published reports and online. Additional information may be obtained from the files of other agencies, extracted from newspaper files, or obtained by intensive inquiry and investigation near the site for which the flood frequency information is needed (Thomson et al., 1964). Reports prepared by Federal agencies (such as the U.S. Army Corps of Engineers and Bureau of Reclamation) to Congress requesting funding for civil works projects often contain historical flood information that supports the need for the project. These reports are available at many university and public libraries around the country. Data sources that could be used to identify the historical period n_h, perception threshold(s) T_h, and the largest floods outside the gaging record are described in Appendix 2.

Over the past several decades, historical data and information have been shown to be extremely valuable in flood frequency analysis (Leese, 1973; Condie and Lee, 1982; Stedinger and Cohn, 1986, 1987; Cohn et al., 1997; England et al., 2003a). Dalrymple (1960) notes: "historical floods provide probably the most effective data available on which to base flood-frequency determinations, and where the data are reliable this information should be given the greatest weight in constructing the flood-frequency graph". Historical flood information should be obtained and documented whenever possible. Use of historical data assures that estimates fit community experience and improves the frequency determinations. This information is valuable in flood frequency analysis because it directly contributes extreme flood data on low annual exceedance probability floods.

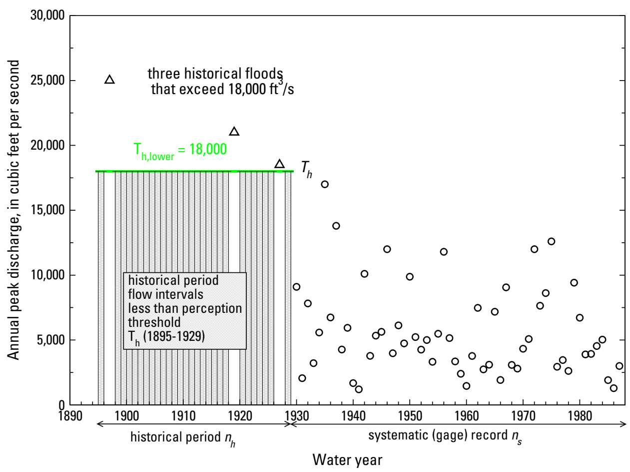

Figure 4. Example site with three large historical floods outside the gaging record, Big Sandy River at Bruceton, Tennessee, U.S. Geological Survey streamflow-gaging station 03606500. The historical floods are known to exceed the perception threshold T_h.

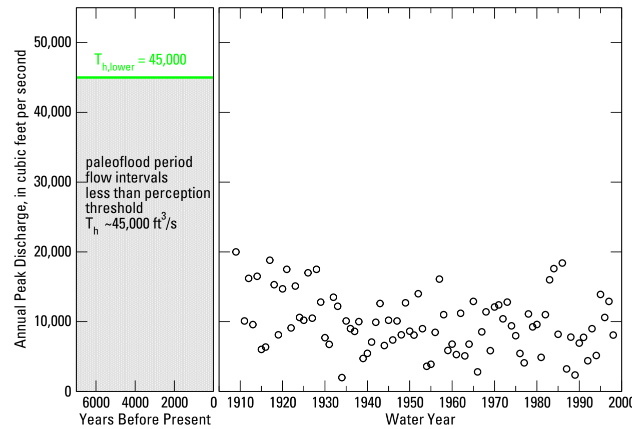

Figure 6. Example site with historical/paleoflood non-exceedance information, North Platte River near Seminoe Reservoir, Wyoming (Levish et al., 2003). A scale break is used to separate the gaging station data from the longer historical/paleoflood period. Floods have not exceeded a perception threshold T_h of 45,000 ft^3/s in the past 7,000 years along the river; the largest floods in the gage record are 20,000 ft^3/s.

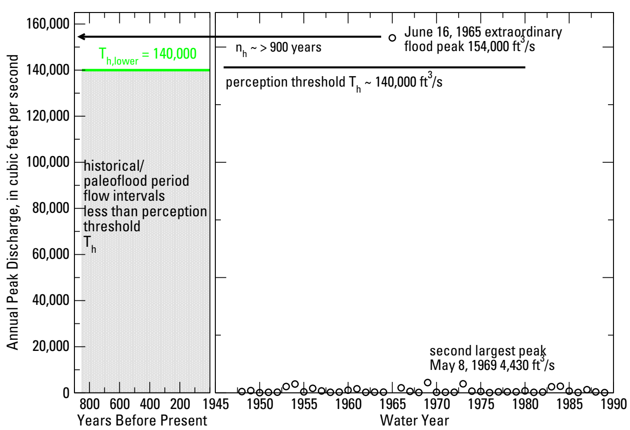

Figure 5. Example site with an extraordinary flood peak that represents a longer time frame, Plum Creek near Louviers, Colorado, U.S. Geological Survey streamflow-gaging station 06709500. A scale break is used to separate the gaging station data from the longer historical/paleoflood period. Horizontal lines indicate the approximate historical period and the perception threshold T_h.

Paleoflood and Botanical Information

Over the past 40 years, there have been significant developments and advances in paleoflood hydrology (Costa, 1978, 1987; Baker et al., 1988; Jarrett, 1991; House et al., 2002b), and increased use of paleoflood data in flood frequency studies by Federal agencies and many others (Jarrett and Tomlinson, 2000; Levish et al., 2003; Sutley et al., 2008; Harden et al., 2011). Paleoflood hydrology primarily involves the study of floods that occurred before human record. Paleofloods are different from historical floods in that they are determined by geologic and physical evidence of past floods rather than records based on community memory or referenced by built infrastructure. Paleoflood hydrology focuses on direct evidence of large, rare floods or the absence of such records. This is critical information for estimating the frequency of such floods (Baker, 1987; Baker et al., 2002).

Extraordinarily large floods often create geomorphologically significant changes to floodplains and terraces, and leave evidence of flood stages in the geologic record that are long-lasting in time. Paleoflood data that are relevant for flood frequency typically consist of: paleostage indicators (PSIs) – discrete evidence of maximum flood stages; and non-exceedance bound information – time intervals during which particular discharges have not been exceeded (Levish, 2002) (see Glossary for complete definitions). Paleoflood features that are typically used as PSIs for flood frequency are shown in Figure 7 (Jarrett and England, 2002), and consist of slackwater deposits (SWDs) (Kite et al., 2002; House et al., 2002a), cobble and gravel flood bars (FBs) (Jarrett and England, 2002), tree scars (Yanosky and Jarrett, 2002), erosional scars and scour lines (Jarrett and Malde, 1987), and silt lines (O'Connor et al., 1986). Geomorphic surfaces (primarily terraces) adjacent to rivers are used to place limits on flood discharges to estimate non-exceedance bounds (Levish, 2002).

Paleoflood data collection methods and applications, including comprehensive overviews and current state of knowledge, are described in Baker et al. (1988) and House et al. (2002b). In many cases, paleoflood evidence persists for hundreds to thousands of years. This allows flood hydrologists and geologists to obtain a great deal of relevant data and information about the largest floods that have occurred during an extended time period. Applied paleoflood and flood frequency studies (England et al., 2006; Harden et al., 2011) have shown that such evidence can greatly increase the precision of flood-frequency estimates at a relatively low cost. In addition, these data are available *now*; one does not have to wait decades to obtain a substantially-longer record.

Paleoflood data are treated in the same way as historical flood data for flood flow frequency analysis using these guidelines. Discharge perception thresholds for individual paleoflood magnitudes and non-exceedance bounds are used with age ranges for various paleoflood periods. In some cases, a single perception threshold, shown in Figure 3, is generalized to multiple thresholds for more complex paleoflood datasets (see Data Representation using Flow Intervals and Perception Thresholds Section). Paleoflood information should be obtained and documented whenever possible, particularly where the systematic record is relatively short, and/or the AEPs of interest are small (≤ 0.01). Some sources for paleoflood data, including regional approaches, are listed in Appendix 2.

Botanical information consists of vegetation that records evidence of a flood (or several floods) and/or indicates stability of a geomorphic surface for some time period. The types of botanical evidence utilized in paleohydrology studies consist primarily of age investigations, placement, distribution, and damage to trees. The four major types of botanical evidence of floods are (Hupp, 1987): corrasion scars, adventitious sprouts, tree age, and ring anomalies. Scars are the most easily observed damage feature, although outward evidence may disappear after a few years.

Sprouts generally occur from broken or inclined tree stems, sometimes called "clipper ships" (Figure 8). Tree age may be utilized to date a particular flood or a geomorphic surface that has been inundated by a flood or may indicate the relative stability of a surface. Vegetation ages in both cases represent a minimum age since the surface was created. In some cases, trees trunks may be partially buried by flood-transported sediments; tree ages in this case are older than the geomorphic surface. Different tree ring patterns (eccentric, shifts, vessel changes, etc.) occur due to floods. Currently, the most reliable and accurate method of tree-ring-determined dates of flooding is the analysis of increment cores or cross sections through scars (Hupp, 1988). Annual formation of rings permits flood dating to within a year, and sometimes to within several weeks Yanosky and Jarrett (2002). Detailed descriptions of each type of evidence are presented in Sigafoos (1964), Yanosky (1983), Hupp (1987, 1988), and Yanosky and Jarrett (2002). Hupp and Osterkamp (1996) review the role of vegetation in fluvial geomorphic processes, including extreme floods. In flood frequency analysis, it is common to describe botanical information as binomial-censored observations corresponding to exceedances of a perception threshold. Some sources for botanical data are listed in Appendix 2.

Common Issues with At-Site Data Records

There are several common issues associated with streamflow data records from gaging stations that may require investigation and treatment by the analyst. These issues include handling of incomplete records, extraordinary floods, and potentially-influential low floods (PILFs).

Broken, Incomplete and Discontinued Records

Annual peaks for certain years may be missing because of conditions not related to flood magnitude,

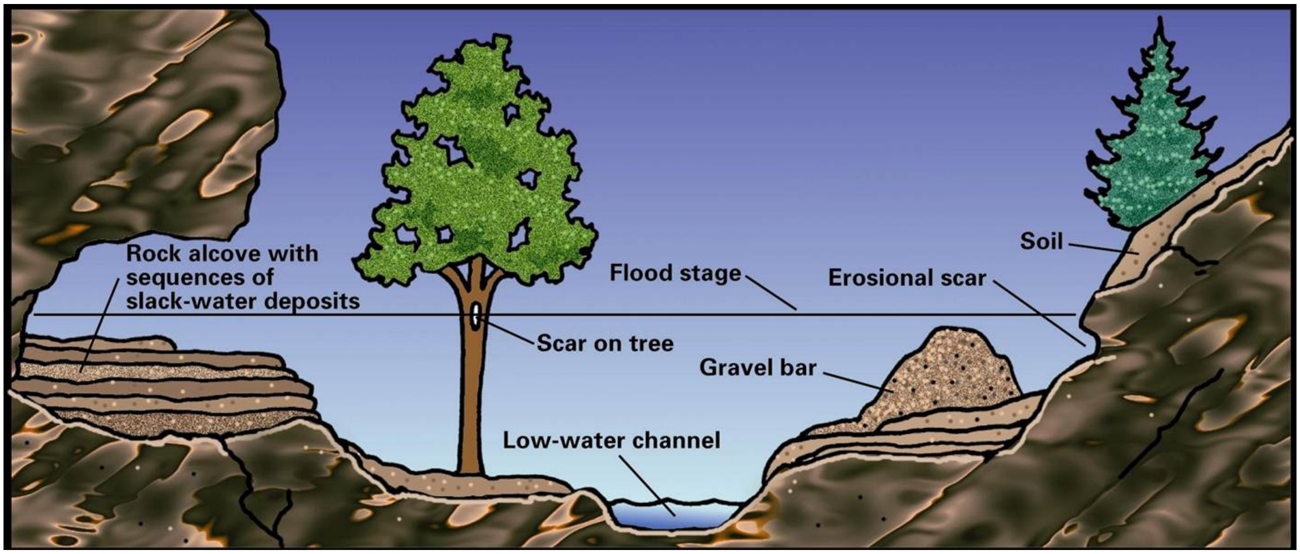

Figure 7. Diagrammatic section showing typical paleoflood features used as paleostage indicators (from Jarrett and England, 2002).

Figure 8. Inclined western juniper trees with upright branches from 1861-1862 flood, Crooked River near Prineville, Oregon.

such as gage removal. These records are considered "broken"; a typical example is shown in Figure 9. In this case, the analyst needs to determine if the records are equivalent, and if there is additional information such as historical or paleoflood information that can place the largest floods in a longer time context (Paretti et al., 2014a, Figure 4). The different record segments can be analyzed as a continuous record with length equal to the sum of both records if the gage is reestablished in a nearby location, unless there is some physical change in the watershed between segments which may make the total record non-homogeneous. Data from an upstream or downstream gage may provide additional information to estimate a perception threshold on the magnitude of floods that occurred during the missing or broken period.

An "incomplete" record refers to a streamflow record in which some peak flows are missing because they were too low or too high to record, or the gage was out of operation for a short period because of flooding. Missing high and low data require different treatment. When one or more high annual peaks during the period of systematic record have not been recorded, there is usually information available from which the peak discharge can be estimated, or a flow interval estimate can be made. A perception threshold is used to describe the knowledge that floods are not measured above a certain stage. For example, the USGS National Water Information System (NWIS) provides a code "8" that a discharge was greater than an indicated value. At some crest gage sites, the bottom of the gage is not reached in some years. the USGS NWIS provides a code "4" that a discharge was less than an indicated value. For this situation, a perception threshold is set to properly represent the

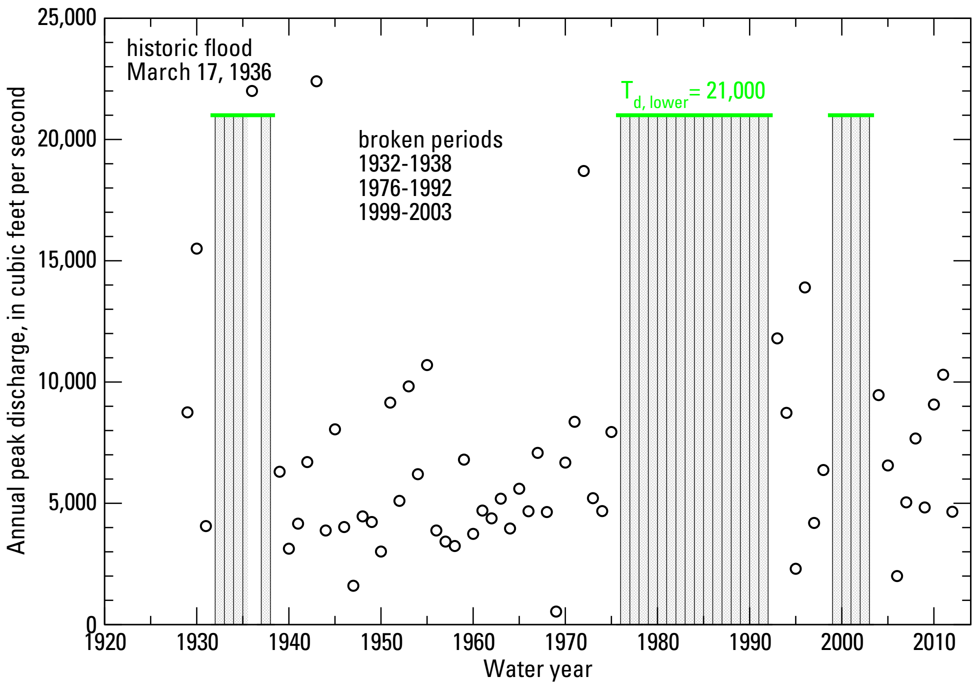

Figure 9. Example streamflow-gaging station with a broken record, U.S. Geological Survey station 01614000, Back Creek near Jones Springs, West Virginia. The grey shaded area represents floods of unknown magnitude less than a perception threshold T_h (shown as a green line) during the systematic record n_s. Black vertical bars during the systematic record represent flow intervals for each year for the unrecorded observations, with the perception threshold T_h based on the March 1936 flood.

incomplete observations less than some value. In most instances, the data collecting agency provides information to estimate peak discharges, flow intervals, and/or perception thresholds. Estimates that are made as part of the flood frequency analysis should be documented.

Streamflow-gaging data are available at many locations where records are no longer being collected. These stations and records are considered “discontinued”, are extremely valuable, and should be utilized for frequency analysis. Streamflow records in many watersheds have been discontinued due to watershed development, including construction of dams and reservoirs. These discontinued records can be extended with the use of reservoir records (Appendix 2) and a perception threshold (Figure 10).

Extraordinary Floods

Extraordinary floods are those floods that are the largest magnitude at a gaging station or miscellaneous site and that substantially exceed the other flood observations (Costa and Jarrett, 2008). Extraordinary floods may be from gaging station records, indirect measurements at miscellaneous sites or from historical flood, paleoflood, or botanical information as described in the Sections Historical Flood Information and Paleoflood and Botanical Information. These floods typically exceed the second largest observation at a gaging station by a factor of two or greater, and in some cases, can be 35 times larger (Figure 5). There are many examples of extraordinary floods throughout the United States, such as the June 1921 flood on the Arkansas River in Colorado (Hazen, 1930) (Figure 10), the record 1954 flood on the Pecos River in Texas (Kochel et al., 1982; Lane, 1987), the 1976 Big Thompson River flash flood in Colorado (Costa, 1978; Jarrett and Costa, 1988), and the June 2008 Cedar River, Iowa flood (Eash, 2010). Costa and Baker (1981) describe some extraordinary floods that represent substantially longer time frames than the gaging record length n_s at each site. Costa and Jarrett (2008) discuss the physical process recognition and indirect discharge issues in estimating extraordinary

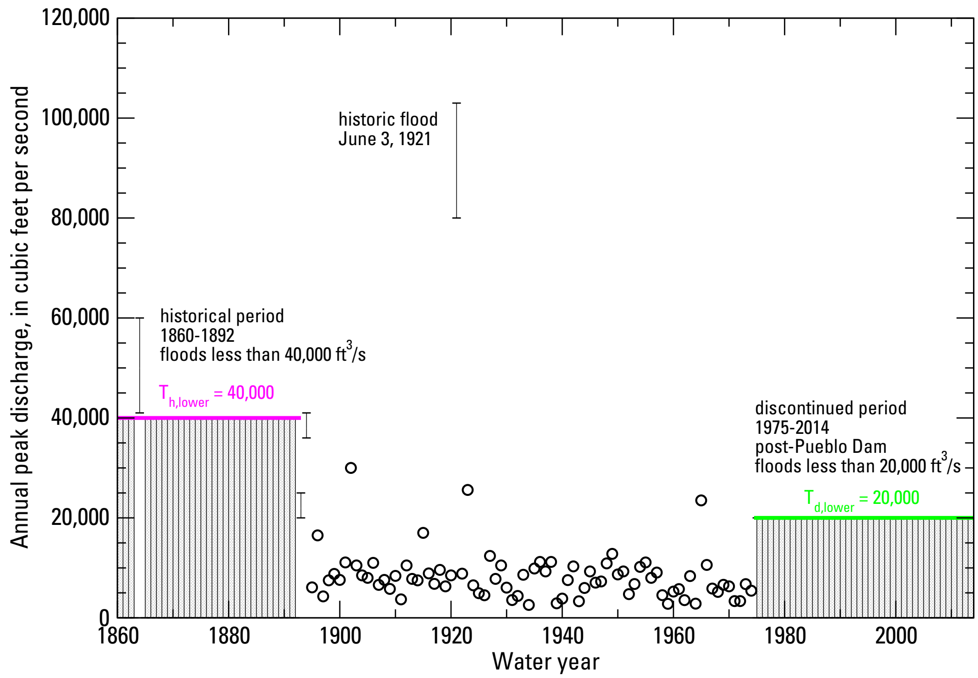

Figure 10. Example streamflow-gaging station with a discontinued record that is extended with a perception threshold from reservoir records, U.S. Geological Survey station 07099500, Arkansas River near Pueblo, Colorado. The gage was discontinued in 1975; floods since then have not exceeded a 20,000 ft^3/s perception threshold T_d shown as a green line. Large floods that occurred in 1864, 1893, 1894 and the June, 1921 extraordinary flood are described as interval observations and are shown as vertical bars with caps that represent lower and upper flow estimates. A perception threshold T_h for the historical period is shown as a magenta line.

flood magnitudes, and note the uncertainty of these estimates is large.

These extraordinary floods are of critical importance because these estimates have a direct and large influence on the fitting of the flood frequency distribution, and are the events of interest to estimate flood magnitude and frequency. Extraordinary floods should be identified by using flood peak ratios, time series plots, and regional flood peak envelope curves (Crippen and Bue, 1977; Asquith and Slade, 1995; O'Connor and Costa, 2004). The method used to estimate the extraordinary flood magnitude and relevant documentation should be reviewed to examine for potential errors, gather additional information about the flood, and to estimate uncertainty (Costa and Jarrett, 2008).

All extraordinary flood observations are to be retained and used in frequency analysis. These record floods represent a longer time frame than that of the gaging record length n_s. Historical flood, paleoflood, and botanical information should be collected within the watershed and region of interest, in order to estimate perception thresholds T_h and expand the record length n_h for the extraordinary flood(s). The recommended procedures, described in the section Determination of the Flood Flow Frequency Curve, are appropriate for analyzing extraordinary floods at gaging stations. The use of other frequency distributions, estimation procedures, or more complex models for extraordinary floods is not warranted. It is recommended to closely examine the fit of the flood frequency curve to the largest observations, and understand the influence of the any extraordinary observations on the fitted frequency curve. Confidence intervals should be used to estimate the range of *AEPs* for the flood. Examination of and comparison with regional flood information is also warranted. Regional flood peak envelope probabilities (Vogel et al., 2007) can be considered in order to assess frequency estimates.

Zero Flows and Potentially-Influential Low Floods

Many rivers and streams in arid and semi-arid regions within the western United States, such as in California (Lamontagne et al., 2012) and Arizona (Paretti et al., 2014a), have zero or very small flows for the entire year. The annual flood series for these streams will have one or more low-magnitude or zero flood values (Figure 11). Such observations merit special attention. In particular, the logarithm of zero is negative infinity, and the logarithm of unusually small values can also be anomalous. Moreover, small flood values can have a large influence on the fitting of the flood frequency distribution and the estimation of the magnitude of rare flood flows. These small observations are called Potentially-Influential Low Floods (PILFs) (Cohn et al., 2013).

In these watersheds, the processes that create very large floods – i.e. the floods of interest – may be different from the processes that cause the low (or zero) value annual peaks. Many low values can occur due to the influence of basin characteristics, such as channel-infiltration losses or evapotranspiration exceeding annual rainfall (Paretti et al., 2014a). The result is that the series of annual peaks appears to be generated from a mixed distribution. For example, peak flows in the range of 5,000 to 15,000 ft^3/s are of interest on Orestimba Creek (Figure 11), rather than the numerous zero values and small flows less than about 1,000 ft^3/s at this site. Consequently, the magnitudes of small annual peaks typically do not reveal much about the upper right-hand tail of the frequency distribution, and thus should not have a highly influential role when estimating the probabilities of large floods (Cohn et al., 2013). These low (or zero) flows are thus not relevant to estimating the probabilities of the largest flood events (Klemeš, 1986, 2000).

These *Guidelines* recommend the use of *robust* estimation procedures (Kuczera, 1982; Lamontagne et al., 2013) and a focus on the largest floods – the upper tail of the flood frequency distribution (National Research Council, 1988) – to eliminate PILFs. Robust estimation procedures are reasonably efficient when the assumed characteristics of the flood distribution are true, while not doing poorly when those assumptions are violated (Stedinger et al., 1993; Cohn et al., 2013). A focus on the most extreme events (upper tail) is based on the observation that hydrometeorological and watershed processes during extreme events are likely to be quite different from those same processes during more common events (National Research Council, 1988, p. 7). The statistical procedure presented in the Section Zeros and Identifying Potentially-Influential Low Floods is used to detect PILFs.

Data Representation using Flow Intervals and Perception Thresholds

Traditionally, flood flow frequency determination focused on the analysis of flood observations Q recorded in every year Y at continuous-record stream gages, which could be represented as point data Q_Y. The description of flood and streamflow data for frequency analysis, and knowledge of the statistical characteristics of the data, have changed over the past 30 years. Valuable flood data, that cannot usually be represented as point values, includes that from crest-stage gages, historical information, and paleoflood and botanical information. A generalized representation is needed to capture what is known about annual peak flows in a given year Y, or over a range of years n. This includes information about specific annual floods that are known to be within a range of values, or above or below an estimated perception threshold. Also, there may be information over a range of years in which it is known that no flood occurred above a known perception threshold. There may be sites where multiple perception thresholds are needed to represent different segments of the sample data across the historical period.

Representations of peak-flow observations are now generalized to include concepts such as: flow intervals, exceedances, nonexceedances, and multiple perception thresholds. These concepts are described in this section to provide a generalized data representation for flood frequency. Selected definitions for these concepts are presented in the Glossary. The recommended procedures in these *Guidelines*, described in the section Determination of the Flood Flow Frequency Curve, can readily incorporate these new types

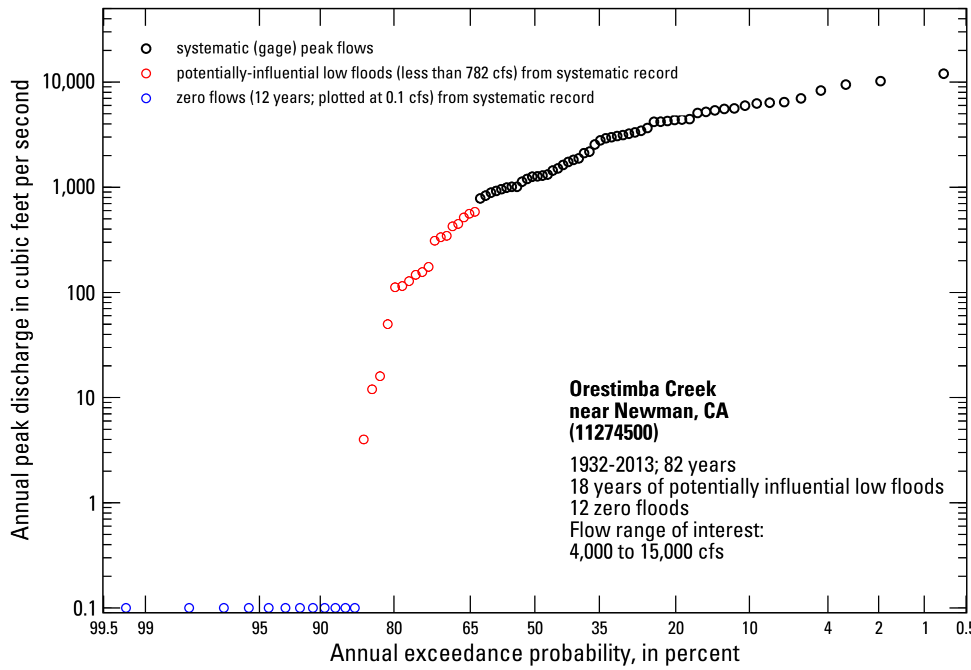

Figure 11. Example empirical frequency distribution and potentially-influential low floods (PILFs) at U.S. Geological Survey station 11274500, Orestimba Creek near Newman, California.

of information, and can use the data properly in frequency analysis of large floods. This allows use of all types of information in multiple combinations as necessary to best utilize the flood data available at a site. In these *Guidelines*, all flood data are represented by flow intervals and perception thresholds (Figure 12).

For each year Y, the magnitude of Q_Y is characterized as a flow interval $(Q_{Y,lower}, Q_{Y,upper})$. A lower estimate $Q_{Y,lower}$ and upper estimate $Q_{Y,upper}$ (interval) is made based on observations, written records, or physical evidence. For the majority of floods, such as those from a gaging station, the discharge is nearly “exactly” known (for all practical purposes), and $Q_{Y,lower} = Q_{Y,upper} = Q_Y$. Floods that are described by intervals or ranges currently address two situations (Figure 12): (1) a flood that is known to exceed some level, with no upper estimate (binomial-censored data); and (2) floods that are known to fall within a large range (interval data). In the binomial case, one only knows the lower estimate $Q_{Y,lower}$; the upper estimate $Q_{Y,upper} \cong +\infty$ and is represented by a dashed line (Figure 12). Flow intervals are used to describe, in some cases, the largest flood magnitudes that are estimated from historical and paleoflood records, and sometimes indirect measurements or field estimates at a gage that have large uncertainty ($> 25\%$). Flow intervals $Q_{Y,lower}$, $Q_{Y,upper}$ are not used to provide ranges on gaged flows and reflect measurement uncertainties that are within 5-25%. The interval observations are shown in Figure 12 with bars for the lower and upper estimates. For unobserved historical floods whose magnitudes are only known to be less than some perception threshold (T_h), the lower estimate $Q_{Y,lower} = 0$, and the upper estimate $Q_{Y,upper}$ corresponds to the perception threshold for that year, such as T_{h1L} or T_{h2L} (Figure 12). For crest-stage gages, flow intervals are determined with consideration of equipment recording limits of stage. There is usually a base (minimum) discharge Q_b established; this may vary each year.

Perception thresholds $(T_{Y,lower}, T_{Y,upper})$ are used to describe the knowledge in each year Y within the flood record, for which the value of Q_Y would have been observed or recorded. The lower bound $(T_{Y,lower})$ represents the smallest peak flow that would result in a recorded flow; the upper bound $(T_{Y,upper})$ represents the largest peak flow that could be observed or

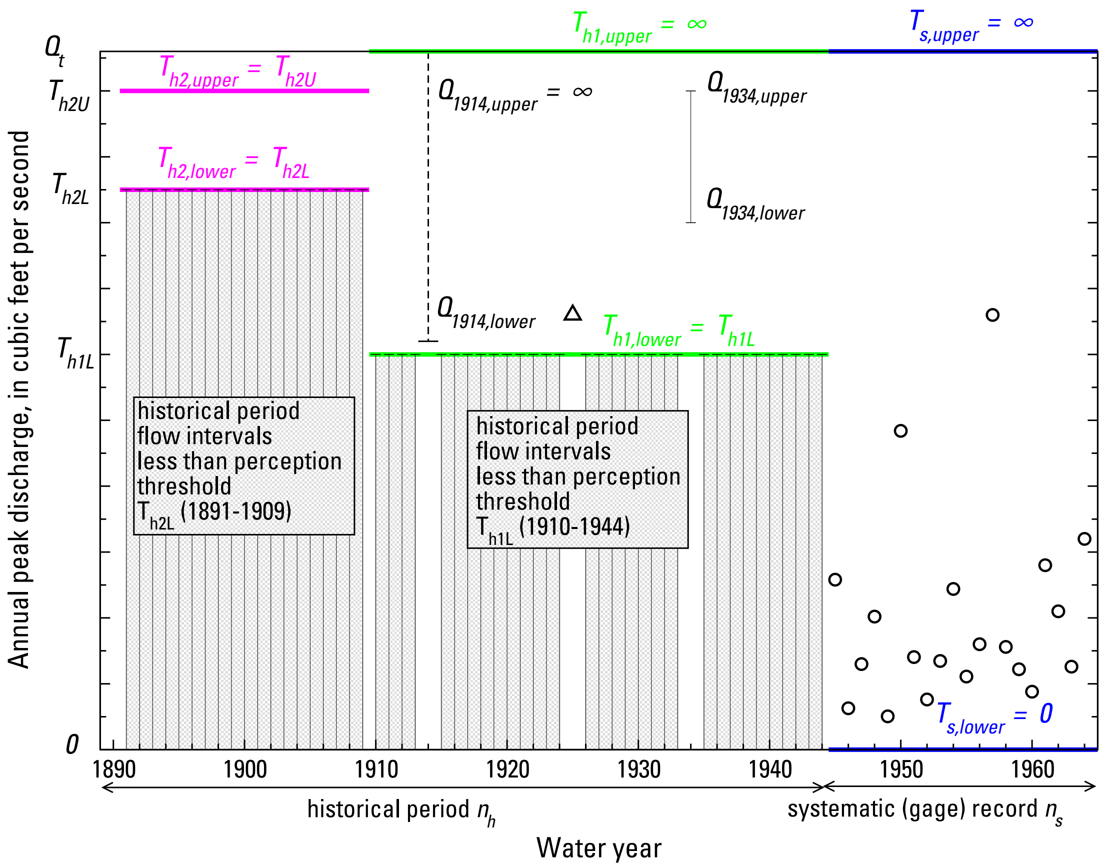

Figure 12. Example peak discharge time series showing peak flows, interval and binomial-censored flood observations, flow intervals, and perception thresholds. Systematic (gage) peak flows are shown as black circles, where $Q_{Y,lower} = Q_{Y,upper}$. During the historical period, there are three floods: a "binomial" observation in 1914; one flood with known magnitude in 1925; and an interval observation in 1934. The 1925 peak flow is shown as a black triangle, where the magnitude is exactly known ($Q_{1925,lower} = Q_{1925,upper}$). The 1914 flood is described as a binomial observation and shown with a dashed line; it is known that this flood exceeded $Q_{1914,lower}$ but the upper estimate is unknown. The flood in 1934 is known to fall within a certain range described with an interval ($Q_{1934,lower} < Q_{1934,upper}$). Flood intervals are shown as black vertical bars with caps that represent lower and upper flow estimates. The grey shaded areas represents floods of unknown magnitude less than the perception thresholds T_{h1L} and T_{h2L} during the historical period. The green lines represent the range in which floods would have been measured or recorded for the period 1910-1945, with lower and upper perception thresholds $T_{h1,lower}$ and $T_{h1,upper}$. The magenta lines are the perception thresholds $T_{h2,lower}$ and $T_{h2,upper}$ for the period 1891-1909. The perceptible range for the systematic (gage) period (1945-1965) $T_{s,lower}, T_{s,upper}$ $(0, \infty)$ is shown as blue lines.

recorded. The interval $(T_{Y,lower}, T_{Y,upper})$ defines the range of "perceptible values" – the range of potentially measurable flood discharges. These perception thresholds reflect the range of flows whose magnitude would have been recorded had they occurred, and are a function of the type of data collected at or near a gaging station and the physical characteristics of the river. In other words, the perception thresholds represent the "observable range" of floods. It is important to note that the perception thresholds T_Y do not depend on the actual peak discharges Q_Y that have occurred.

Lower and upper perception thresholds T_Y need to be estimated for each and every year of the record. The lower bound $T_{Y,lower}$ represents the smallest annual peak flow that would result in a permanent record. For systematic (gaging) records, this is typically represented by the "gage-base discharge," which is typically 0. At crest-stage gages, $Q_b > 0$, and may vary.

For historical floods, $T_{Y,lower}$ is typically estimated to be equal to a historical flood discharge threshold T_h (Figure 12). For most sites with a systematic, continuous gaging record, $T_{Y,upper}$ is assumed to be infinite; larger floods typically get recorded. At crest-stage gages, and for historical and paleoflood periods, $T_{Y,upper}$ needs to be estimated based on the CSG recording range, historical information (such as markers, bridges or buildings), or from geologic or botanical evidence. For periods where the gage has been discontinued (broken record) or ceased operation, the observation thresholds are both set to infinity, if there is no other information such as a gage base or historical information. By setting $T_{Y,lower} = T_{Y,upper} = \infty$, this means that there is no information about that particular year. If there is historical information that is used for record extension of the largest floods during broken record periods, $T_{Y,lower}$ can be set to a historical flood discharge threshold T_h.

In some situations, flood data sets need to be represented by multiple perception thresholds. This means more than one perception threshold is required to describe the data at hand. It is appropriate to utilize multiple perception thresholds, particularly with longer historical records and paleoflood data, to properly represent the data and information at hand. In this situation, the two perception thresholds shown in Figure 12 would be extended with additional perception thresholds that are larger in magnitude than T_{h2L} and represent longer time frames.

It is critical to collect historical data and determine the historical period n_h for flood frequency. The beginning of the historical period may be based on, for example, the earliest known historical settlement dates (such as 1860) along a river (Figure 10), from archaeological information, or from paleoflood information and dating of river terraces and non-exceedance bounds (Figures 5 and 6). The historical period does not begin at the earliest (first) observed flood, which is a biased estimate of n_h as it is a lower bound on the true historical period (Hirsch and Stedinger, 1987).

The lower perception threshold $T_{Y,lower}$ is particularly important to estimate. It represents our best judgment, for any given year, of the smallest size flood that would have left evidence that the investigator would know about today. The historical or paleoflood information needs to persist so that hydrologists and geologists can obtain the data from written records, historical investigations, or paleoflood studies. For example, for every year during the period 1891-1909, no evidence was found to indicate peak flows Q_Y had exceeded T_{h2L} (Figure 12). The investigator should recognize that the lower limit of the perception threshold may be a rough approximation, and that it usually changes (increases in magnitude) as one moves backwards in time. In some cases, only the most catastrophic events would have been recorded and the threshold is high (Figure 5); these are the events that are of interest.

Regional Information and Nearby Sites

Flood information from within a region surrounding the gage site or watershed of interest is useful to improve flood-frequency estimates, particularly when streamflow-gaging records are short (less than 30 years) (Stedinger et al., 1993). For these and other modest-length records, it is known that the station skew coefficient is sensitive to extreme events (Griffis and Stedinger, 2007a, 2009). Since Bulletin 17 (Beard, 1974), regional skew information G has been used to stabilize the station skew coefficient ($\hat{\gamma}$), which defines the shape of the fitted frequency distribution, through the use of a "weighted" skew coefficient $\tilde{G}$. The techniques for estimating regional skew have evolved over the past 30 years (Tasker and Stedinger, 1986; Griffis and Stedinger, 2007d; Parrett et al., 2011), with the result that estimates are now much more accurate and their statistical properties are better understood, than at the time Bulletin 17B was written. It is recommended that regional skew information G is considered and weighted appropriately when estimating flood-frequency curves. Some sources of regional skew information are listed in Appendix 2. Additional guidance is provided in the Sections Estimating Regional Skew and Weighted Skew Coefficient Estimator.

Other types of regional information that may be valuable for flood frequency can be considered, in addition to regional skew information. Griffis and Stedinger (2007a) describe several flood frequency estimators and show that regional estimates of the mean and standard deviation can be valuable. In arid and semi-arid regions, regional mean and standard deviation estimates from peak flows can be used to improve at-site flood frequency estimates, such as the desert region in California (Gotvald et al., 2012). Physiographic characteristics within a watershed or region, such mean basin elevation, drainage area, mean annual precipitation, and other physical factors, are useful in estimating regional parameters and in conducting regional flood frequency studies. Such studies usually employ generalized least-squares regression techniques (Tasker and Stedinger, 1989) to provide regional flood quantile estimates and quantile variances. These estimates are available for many states (Gotvald et al., 2012; Eash et al., 2013), and may be valuable for record extension and weighting of independent estimates. Additional guidance is provided in the Sections Record Extension with Nearby Sites and Weighting of Independent Frequency Estimates.

Flood Estimates from Precipitation

Flood discharges estimated from climatic data (rainfall and/or snowmelt) can be a useful adjunct to direct streamflow measurements. Estimates may be available from several cases, such as: (1) flood estimates from individual extreme events that are based on observed rainfall; (2) synthetic flood events and frequency curves from rainfall frequency estimates; and (3) continuous streamflow estimates and frequency curves from precipitation and climate information.

Such estimates require at least adequate climatic data and a valid watershed model for converting precipitation to discharge. In some situations, existing watershed models may be available that are already calibrated to the watershed of interest. For example, the National Weather Service (NWS) has calibrated watershed models for flood forecasting on major river basins through their River Forecast Centers (RFCs). Other Federal agencies (USACE, Reclamation, NRCS) may have calibrated flood watershed models for flood control, levee design, and other projects within their jurisdiction. As part of floodplain management studies for the Federal Emergency Management Agency (FEMA), state agencies, counties, and local watershed protection districts may have calibrated watershed models for large floods that may be used to supplement streamflow-gaging station records.

Individual extreme floods or flood frequency curves can be estimated from event-based or continuous rainfall-runoff models (National Research Council, 1988; Singh, 1995; U.S. Bureau of Reclamation and Utah State University, 1999; Beven, 2001; FEMA, 2009), using observed watershed precipitation, precipitation observed at nearby stations in a meteorologically homogeneous region, or from stochastically-generated precipitation. The rainfall-runoff model needs to be calibrated to extreme flood observations, using procedures such as those presented in Duan et al. (2003), in order to be useful for flood frequency estimation and prediction. It is recommended that an uncertainty analysis be conducted (Kjeldsen et al., 2014), including prediction uncertainty (Beven, 2001, Chapter 7), to reflect the range of variability associated with the estimated flood frequency curve from the rainfall-runoff model. The variance of flood quantile estimates from rainfall-runoff models is also needed for potential weighting the estimate, as described in the Section Weighting of Independent Frequency Estimates.

Flood frequency estimates from rainfall-runoff models can be be biased low (Thomas, 1982) or high and exhibit a loss of variance (Lichty and Liscum, 1978; Thomas, 1987) when model and other errors are not properly accounted for in uncertainty analysis. Including variability in precipitation and temperature inputs (Clark et al., 2004) helps in this situation. In some cases, rainfall-runoff models are calibrated to or parameters are adjusted to better match flood-frequency curves based on peak-flow statistics (Reed, 1999; Swain et al., 2006; MGS Engineering Consultants, 2009). Frequency curves from rainfall-runoff models need to be independent of the frequency curve estimated using the recommended procedures in these *Guidelines*, if curves are to be weighted and combined.

Analysts making use of such procedures should clearly document the rainfall-runoff method used for computing the floods and evaluate its performance based upon flood and storm experience in a hydrologically and meteorologically homogeneous region, including calibration and uncertainty analysis. Whether or not such studies are useful will depend upon the availability of the information, the adequacy of the existing flood records, and the purpose for which the watershed model was developed and calibrated. The magnitude and *AEP* of the precipitation or flood event are the most important factors to consider when including these estimates. The largest or most extreme flood events, with *AEPs* < 0.02 are very useful, especially for ungaged sites or in situations where gaging stations have been destroyed.

In addition to flood estimates from precipitation, hydroclimatological information (Maddox et al., 1980; Hirschboeck, 1991) is very useful and provides a broad perspective on data and flood processes for frequency analysis. Atmospheric circulation patterns (Hirschboeck, 1987a) and climate indices such as ENSO (Webb and Betancourt, 1992) can be coupled with streamflow records to gain insight to the types of flood-causing mechanisms and flood variability (National Research Council, 1999). Redmond et al. (2002) describe important connections between climate mechanisms, paleofloods, and flood variability. Some sources for precipitation and climate information are listed in Appendix 2.

Data Assumptions and Specific Concerns

The conventional assumptions for a statistical analysis are that the array of flood information is a reliable and representative time sample of random homogeneous events. Assessment of the adequacy and applicability of flood records is therefore a necessary first step in flood frequency analysis. This section discusses flow measurement error, randomness of events, mixed populations, watershed changes, and climate variability and change considerations for flood frequency analysis.

Flow Measurement Error

Peak-flow measurement errors exist in streamflow records, as in all other measured values. Sauer and Meyer (1992) describes sources of error in streamflow measurement. Errors in flow estimates are generally greatest during maximum flood flows. Peak flow estimates of the largest floods from systematic (gage) records, historical floods, paleofloods, or from other sources, can be substantially in error because of the uncertainty in both stage and stage-discharge relationships, and because the flows may be estimated from rating curve extensions or indirect methods, rather than by direct measurement. Many improvements have been made in direct measurements of streamflow by the USGS over the past several decades (Turnipseed and Sauer, 2010), with "good" (5%) accuracy of most discharge measurements. However, the largest flows are generally not directly measured because of problems with debris, inaccessibility issues, and safety considerations (Costa and Jarrett, 2008). The largest floods are usually estimated by rating-curve extensions or indirect methods, with estimation errors that can exceed 25% in many cases, to over 100% in high-gradient streams (Jarrett, 1987). Measurement errors can seriously degrade flood quantile estimates in some situations (Potter and Walker, 1985); therefore estimation errors in the largest floods should be investigated.

In many instances, annual peak discharges are estimated from rating-curve extensions. Significant errors in discharge estimation may occur from rating curve extensions (Cook, 1987; Kuczera, 1996), especially if the discharge value is more than twice the greatest measurement by current meter. Unfortunately, high outliers or significant flood peaks are usually never measured directly and are many times greater than twice the measured value (Klemeš, 1987). Kuczera (1996) indicates that rating curve extensions, in the presence of correlated errors, can significantly affect quantile estimates from such extrapolations.

Indirect methods are utilized to measure peak discharges after flood periods (Benson and Dalrymple, 1967; Rantz and Others, 1982), using high-water marks (HWMs) or PSIs. The slope-area method (Dalrymple and Benson, 1967) is most commonly used by the U.S. Geological Survey; other indirect methods are presented by Cook (1987) and Webb and Jarrett (2002). Slope-area methods have documented sources of uncertainty (Bathurst, 1986; Jarrett, 1987; Kirby, 1987; McCuen and Knight, 2006). Significant errors in indirect discharge estimates have been noted in mountain areas; the measurements typically are overestimated (Jarrett, 1987). Neglecting channel scour or fill is the most significant factor that may introduce large errors in indirect discharge estimates (Kirby, 1987). Quick (1991) presents sources of errors in the slope-area method and indicates that the method has a strong upward bias. Webb and Jarrett (2002) describe assumptions in estimating historical and paleoflood peak discharges; they also outline information needed to support discharge estimates.

At times errors will be apparent or suspected. Substantial efforts should be made to understand sources of flow measurement errors, and to quantify the uncertainty associated with such errors. If substantial, the errors should be brought to the attention of the data collecting agency with supporting evidence and a request for a corrected value.

Randomness of Events

In general, a time series of annual peak flow estimates may be considered to be a random sample of independent, identically-distributed random variables. The peak-flow time series is assumed to be a representative sample of the population of future floods. This assumption is contingent upon conducting exploratory data analysis (Appendix 3) and further physical knowledge of the system. In essence, the stochastic process that generates floods is assumed to be stationary, or invariant in time. Stationarity is a property of an underlying stochastic process, and not of observed data. Realizations from stationary processes can exhibit excursions and trends that persist for decades or centuries (Cohn and Lins, 2005). Nonstationary processes are difficult to detect in peak-flow series (Villarini et al., 2009) and may be challenging to determine (Koutsoyiannis, 2011). In some situations, long-term persistence concepts (Lins and Cohn, 2011) or shifting-mean models (Salas and Boes, 1980; Sveinsson et al., 2003) could be considered.

Before conducting flood frequency analysis, these *Guidelines* recommend that analysts perform an initial analysis of the data. Helsel and Hirsch (1992) and Hirsch et al. (1993) provide overviews and details on conducting exploratory data analysis. The recommended procedures for initial data analysis include plotting the series, estimating serial correlation, and examining for trends and abrupt shifts (changepoints), and are presented in Appendix 3.

In certain locations, flood records may indicate apparent nonrandomness and exhibit strong multidecadal trends or wet and dry cycles that are not explained by land use change, water management, or climate change. Such records are particularly challenging and this is one of the most vexing problems in flood frequency analysis. The Work Group did not evaluate methods to account for nonrandomness and/or multidecadal trends in flood frequency. Additional work in this area is warranted, as it is a seriously unresolved problem. If multidecadal trends of this sort are identified though appropriate statistical tests and data analysis, it is recommended that the underlying physical mechanisms be investigated to gain hydrological understanding (Lins and Cohn, 2011). How to adjust such a record for flood frequency is an unresolved problem.

Even when statistical tests of the serial correlation coefficients indicate a significant deviation from the independence assumption, the annual peak data may define an unbiased estimation of future flood activity if other assumptions are attained. The nonrandomness of the peak series will, however, result in error in the estimated uncertainty associated with the fitted frequency curve. Effective record length concepts (Tasker, 1983; Vogel and Kroll, 1991) should be used to correct uncertainty estimates in the presence of serial correlation.

Mixed Populations

Flooding in some watersheds is caused by different types of meteorological events associated with distinct physical processes. For example, flooding at some locations may arise from snowmelt, rainstorms, or by combinations of both snowmelt and rainstorms (Jarrett and Costa, 1988). Such a record may not be homogeneous and may require special treatment. This mixed population results in flood frequency curves with abnormally large skew coefficients reflected by abnormal slope changes when plotted on logarithmic normal probability paper. In some situations, the frequency curve of annual events can best be described by computing separate curves for each type of event, and combining the results.

One example of mixed population is rainfall-runoff mixed with snowmelt. In the Sierra Nevada region of California, hydrologic factors and relationships operating during general winter rain floods are usually quite different from those operating during spring snowmelt floods or during local summer cloudburst floods. In this region, peak flows are primarily caused by winter rainfall at lower elevations, while at higher elevations, peak flows are generally caused by spring snowmelt or rain-on-snow events (Parrett et al., 2011). Frequency studies in the Sierra Nevada have been made separately for rain floods which occur principally during the months of November through March, and for snowmelt floods, which occur during the months of April through July. Peak flows were segregated by cause – those predominately caused by snowmelt and those predominately caused by rain (Crippen, 1978). Likewise, in the Colorado Front Range, peak-flows are caused by both rainfall and snowmelt during the spring and summer (Elliott et al., 1982), especially in the lower elevation foothills zone (Jarrett and Costa, 1988).

Flooding in the eastern United States is caused by a mixture of flood-generating mechanisms, with tropical cyclones and extratropical systems playing a central role (Smith et al., 2011). Along the Atlantic and Gulf Coasts, in some instances floods from hurricane and non-hurricane events have been separated, thereby improving frequency estimates (Murphy, 2001). Ice-jam floods that occur in northern regions (Murphy, 2001) are another mixed-population example.

Hydroclimatological data, including the use of synoptic weather patterns (Hirschboeck, 1987b), is particularly useful to provide independent, physically-based information on climate-induced flood processes and to separate flood series by type. Additional data, such as paleohydrologic and paleoclimate data, may also be considered (Redmond et al., 2002). The flood types and particular causative mechanisms may also be explored using a watershed perspective and considering variables such as storm rainfall and duration, flood seasonality, timing, and runoff response (Merz and Blöschl, 2003).

When it can be shown that there are two or more distinct and generally independent causes of floods, it may be more reliable to segregate the flood data by cause, analyze and compute separate curves for each type of event, and then to combine the curves into an overall analysis of the flood frequency at the site. Procedures such as those described in Crippen (1978), U.S. Army Corps of Engineers (1982), Jarrett and Costa (1988), and Murphy (2001) may be considered. For ice jam flow situations, one may consider using the same mixed-population approach (Murphy, 2001), or a method that focuses on maximum elevation (Vogel and Stedinger, 1984). An example of combining frequency curves was performed for the Black Hills region as part of the peak flow frequency estimates for South Dakota (Sando et al., 2008); see also Alila and Mtiraoui (2002) and others. In some situations, there may not be sufficient data to perform a mixed-population analysis, or the results may not be as reliable (Gotvald et al., 2012). The Work Group did not conduct an evaluation of these procedures. Additional efforts are needed to provide guidance on the identification and treatment of mixed distributions.

Separation by calendar periods in lieu of separation by events is not considered hydrologically reasonable unless the events in the separate periods are clearly caused by different hydrometeorological conditions. The fitting procedures in these *Guidelines* can be used to fit each flood series separately, with the exception that regional skew coefficients cannot be used unless developed for the specific types of events being examined. If the flood events that are believed to comprise two or more populations cannot be identified and separated by an objective and hydrologically meaningful criterion, the record shall be treated as coming from one population.

Watershed Changes

It is becoming increasingly difficult to find watersheds in which the flow regime has not been altered by modifications to the river channel, to the river floodplain, creation or destruction of reservoirs and levees, or modifications to the characteristics of the watershed at large (e.g., urbanization, wildfires, change of cropping practices, erosion control, land drainage, or deforestation). Developments which can change flow conditions include urbanization, channelization, levees, and the construction of reservoirs, diversions, and alteration of land cover conditions (Sauer et al., 1983). Impervious areas within the watershed and their effects on runoff are important considerations (Moglen, 2009).

Watershed history and flood records should be carefully examined to assure that no major watershed changes have occurred during the period of record. Documents which accompany flood records often list such changes that occurred at discrete times. However, the effects of urbanization or the construction of numerous small reservoirs over a period of several years, will likely not be documented. Such incremental changes may not noticeably alter the flow regime from year to year but the cumulative effect can be significant.

Special effort should be made to identify those records which are not homogeneous. The data analysis tools described in Appendix 3 may be used to assess records for potential gradual trends or shifts that might be associated with watershed changes. Spatial and temporal estimates of land use data within watersheds should be obtained, where available. These data are particularly useful in quantifying urbanization impacts on flood frequency (Moglen and Beighley, 2002; McCuen, 2003).

Only records which represent relatively constant watershed conditions should be used for frequency analysis (Konrad, 2003; Moglen and Shivers, 2006). In some situations, flow records may be adjusted to account for watershed change so that they represent current watershed conditions, where physical evidence of watershed change exist in a significant portion of the watershed (McCuen, 2003). The Work Group did not evaluate methods to account for watershed changes and makes no particular recommendations, as additional work is needed in this area.

Climate Variability and Change

There is much concern about changes in flood risk associated with climate variability and long-term climate change. Time invariance was assumed in the development of this guide. In those situations where there is sufficient scientific evidence to facilitate quantification of the impact of climate variability or change in flood risk, this knowledge should be incorporated in flood frequency analysis by employing time-varying parameters or other appropriate techniques. All such methods employed need to be thoroughly documented and justified.

The Work Group did not evaluate methods to account for climate variability in flood frequency. Additional work in this area is warranted. Some information and background on nonstationarity is presented in Olsen et al. (2010) and Kiang et al. (2011). In the interim, analysts might consider the following:

- data on synoptic weather patterns (Hirschboeck, 1987b);
- paleoclimate information (Redmond et al., 2002);
- climate variability and climate projection information (Brekke et al., 2009);
- interannual and interdecadal variations in climate (Jain and Lall, 2001); and
- time-varying distribution parameters (Stedinger and Griffis, 2011; Salas and Obeysekera, 2014).

Determination of the Flood Flow Frequency Curve

This section presents the recommended procedures for determining a flood flow frequency curve. The procedures include: approaches for plotting positions; the flood distribution; parameter estimation; methods to handle zeros and identifying PILFs; the Expected Moments Algorithm; record extension; and confidence intervals for quantiles. Computer programs are required in order to make these calculations; see the Section Software and Examples for available ones.

Plotting Positions

Empirical frequency distributions are a “nonparametric” or distribution free method to infer the probability distribution function (mathmatical model) that describes flood risk. They are used to assess distribution function (e.g. LP-III) fits the data. Probability estimates are made using plotting positions. A basic plotting position formula for symmetrical distributions is (Stedinger et al., 1993, p. 18.24):

$$p_i = \frac{i-a}{n+1-2a} \quad (2)$$

where p_i is the exceedance probability of flood observations Q_i ranked from largest ($i=1$) to smallest ($i=n$), and a is a plotting position parameter ($0 \leq a \leq 0.5$) (see Table 4.1 in Appendix 4).

Historical flood peaks reflect the frequency of large floods and thus should be incorporated into flood frequency analysis. They can also be used to judge the adequacy of estimated flood frequency relationships. For this latter purpose, appropriate plotting positions or estimates of the average exceedance probabilities associated with the historical peaks and the remainder of the data are desired. Hirsch and Stedinger (1987) and Hirsch (1987) provide an algorithm for assigning plotting positions to censored data, such as historical floods. They emphasized the correct interpretation of the information conveyed by historical flood data, the recognition of the limited precision of estimates of the exceedance probabilities of historical floods, and showed that all estimators were relatively imprecise (Hirsch and Stedinger, 1987). The threshold exceedance plotting position formula is given in Appendix 4. It is applicable for potentially-influential low flood cases, in addition to historical data, as the censored-data principles are the same.

Flood Distribution

Flood records describe a succession of natural events which do not fit any one specific known statistical distribution. To make the problem of defining flood probabilities tractable, it is convenient to select a reasonable mathematical distribution. These *Guidelines* recommend the use of the log-Pearson Type III (LP-III) distribution. This distribution has been in use by Federal agencies since 1967 (USWRC, 1967).

Several studies have been conducted over the years to investigate which of many possible distributions and alternative parameter estimation procedures would best meet the purposes of these *Guidelines*. Beard (1974), summarized in IACWD (1982), found that the LP-III distribution with a regional skew coefficient performed well. Griffis and Stedinger (2007b) explored the the characteristics of the LP-III distribution, and showed that it is very flexible and encompasses a wide range of reasonable models for log-space skews $|\gamma| \leq 1.414$. The method of moments parameter estimation procedure works well with reasonable constraints on parameters (Griffis and Stedinger, 2007c), and an informative regional skew (Griffis and Stedinger, 2009). The Work Group concluded from these studies, many applications over the past 40 years, and recent testing (Cohn et al., 2014), that the Pearson Type III distribution with log transformation of the data (log-Pearson Type III distribution) with a regional skew coefficient is the base method for analysis of annual peak-flow data. The LP-III distribution also performs well and is appropriate for applications with historical and paleflood data (England, 1998; Bureau of Reclamation, 2002; Blainey et al., 2002; England et al., 2003a, 2010; Harden et al., 2011).

The base 10 logarithms $X_i \ldots X_n$ of peak flows $Q_i \ldots Q_n$ are assumed to follow a Pearson Type III (P-III) distribution; this probability density function $f(x)$ is:

$$f(x|\tau,\alpha,\beta) = \frac{\left(\frac{x-\tau}{\beta}\right)^{\alpha-1} \exp\left(-\frac{x-\tau}{\beta}\right)}{|\beta|\Gamma(\alpha)} \quad (3)$$

with $\left(\frac{x-\tau}{\beta}\right) \geq 0$ and distribution parameters τ, α and β, where τ is the location parameter, α is the shape parameter, β is the scale parameter, and $\Gamma(\alpha)$ is the gamma function, defined as:

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} \exp(-t)dt. \quad (4)$$

The shape parameter α is limited to positive values, and the scale parameter β may be positive or negative. When $\beta > 0$, the P-III distribution has a lower bound τ and is positive-skewed; the distribution is negative skewed when $\beta < 0$ (τ is an upper bound). This behavior may also be understood described using the skewness coefficient, rather than parameters. When the skewness coefficient is greater than zero ($\gamma > 0$) ($\beta > 0$), the distribution has a positive skew and floods are unbounded. When $\gamma < 0$, ($\beta < 0$), the distribution on the logarithm of floods has a negative skew and an upper bound. Griffis and Stedinger (2007b) present additional properties of the P-III distribution, including plots of the P-III probability density function.

Parameter Estimation: Simple Case

These *Guidelines* recommend the method of moments using the logarithms of flood flows to estimate the distribution parameters. The first three sample moments are used to estimate the P-III parameters. These include the mean ($\hat{\mu}$), standard deviation ($\hat{\sigma}$), and skewness coefficient ($\hat{\gamma}$).

Moments and Parameters

In the case where only systematic data are available, with no historical information or PILFs, the mean, standard deviation and skewness coefficient of station data may be computed using the following equations:

$$\hat{\mu} = \left(\frac{1}{n}\right)\sum_{i=1}^{n} X_i \qquad (5)$$

$$\hat{\sigma} = \sqrt{\left(\frac{1}{n-1}\right)\sum_{i=1}^{n}(X_i-\hat{\mu})^2} \qquad (6)$$

$$\hat{\gamma} = \left(\frac{n}{\hat{\sigma}^3(n-1)(n-2)}\right)\sum_{i=1}^{n}(X_i-\hat{\mu})^3 \qquad (7)$$

where n is the number of flood observations and ($\hat{}$) represents a sample estimate. The standard deviation ($\hat{\sigma}$) and skewness coefficient ($\hat{\gamma}$) include bias correction factors $(n-1)$ and $(n-1)(n-2)$ for small samples.

The parameters are estimated from the sample moments as:

$$\hat{\alpha} = \frac{4}{\hat{\gamma}^2} \qquad (8)$$

$$\hat{\beta} = sign(\hat{\gamma})\left(\frac{\hat{\sigma}^2}{\hat{\alpha}}\right)^{1/2} \qquad (9)$$

and

$$\hat{\tau} = \hat{\mu} - \hat{\alpha}\hat{\beta}. \qquad (10)$$

Flood quantiles $\hat{Q}_q$ for the P-III distribution can be estimated by

$$\hat{X}_q = \hat{\tau} + \hat{\beta}P^{-1}(\hat{\alpha},q) \qquad (11)$$

where $P^{-1}(\hat{\alpha},q)$ is the inverse of the incomplete Gamma function (Abramowitz and Stegun, 1964) and

$$\hat{Q}_q = 10^{\hat{X}_q} \qquad (12)$$

where q is the cumulative probability of interest (e.g. $q = 0.99$, and $q = 1-p$).

Weighted Skew Coefficient Estimator

There is relatively large uncertainty in the at-site sample skewness coefficient (third moment) $\hat{\gamma}$, because it is sensitive to extreme events in modest length records (Griffis and Stedinger, 2007a). The station skew coefficient $\hat{\gamma}$ and regional skew coefficient G can be combined to form a better estimate of skew $\tilde{G}$ for a given watershed, as illustrated by the concepts in Tasker (1978). Under the assumption that the regional skew coefficient G is unbiased and independent of the station skew $\hat{\gamma}$, the mean-square errors (MSEs) of the the station skew $MSE_{\hat{\gamma}}$ and the regional skew MSE_G can be used to estimate a weighted skew coefficient, as described in Appendix 6.

If the regional and station skews differ by more than 0.5, a careful examination of the data and the flood-producing characteristics of the watershed should be made. The MSE of the station skew is computed directly by *EMA*. The MSE of the regional skew is usually estimated through the procedures described in the Section Estimating Regional Skew.

Zeros and Identifying Potentially-Influential Low Floods

Potentially influential points ("outliers") are data points which depart significantly from the trend of the remaining data. In the case of annual peak flows, low outliers may be floods caused by different processes than the larger floods in the annual peak series, as defined in the Section Zero Flows and Potentially-Influential Low Floods. Because inclusion of these zero flow values and "outliers" can significantly affect the statistical parameters computed from the data, especially for small samples, the presence of PILFs in the dataset will bias parameter estimates.

The purpose of flood flow frequency estimation is to describe the relationship between discharge and exceedance probability at the high end of the frequency distribution where *AEPs* are values such as 0.05, 0.02, 0.01, and smaller. There are cases where observed values of some of the smaller annual floods can have a very strong effect on the shape of the estimated frequency distribution at the high discharge end. The purpose of the procedures described here is to eliminate the influence of low floods so that these small floods have little or no impact on the frequency estimates at high discharges. The ultimate goal is to obtain a good agreement between the high end of the observed frequency distribution and the high end of the estimated frequency distribution. This may result in a poor fit at the low end of the frequency distribution but there is generally no negative practical consequences to a lack of fit at the low end.

The smallest observations in the data set do not convey meaningful or valid information about the magnitude of significant flooding (Appendix 5), although they do convey valid information about the frequency of significant flooding. Therefore, if the upper tail of the frequency curve is sensitive to the numerical values of the smallest observations, then that sensitivity is a spurious artifact based on the mathematical form of the assumed but in fact unknown flood distribution, and has no hydrologic validity. Any procedure for treating outliers ultimately requires judgment involving both mathematical and hydrologic considerations. The analyst must use hydrological knowledge while applying a consistent and mathematically appropriate procedure.

These *Guidelines* recommend the use of a Multiple Grubbs Beck Test (*MGBT*) for the detection of PILFs. Statistical procedures for identifying outliers have been extensively studied, including methods for addressing the case of multiple low outliers considered here, as described in Cohn et al. (2013), Lamontagne et al. (2013), and Lamontagne et al. (2015), and citations therein. The new Multiple Grubbs-Beck test was developed as an improvement to the Grubbs-Beck (GB) test (Grubbs and Beck, 1972) used in Bulletin 17B. The GB test is easily defeated by the occurrence of multiple low outliers, which exert a large distorting influence on the fitted frequency curve, but also increase the standard deviation, thereby making the standardized distances between observations too small to trigger the Grubbs-Beck test.

The *MGBT* is a statistically appropriate generalization of the GB test, and is sensitive to the possibility that *several* of the smallest observations are "unusual," or are potentially very influential. The *MGBT* also correctly evaluates cases where one or more observations are zero, or are below a recording threshold (partial record sites). Thus it provides a consistent, objective and statistically defensible algorithm that considers whether a range of the smallest observations should be classified as outliers (or PILFs) for a much wider range of situations.

The *MGBT* follows the same reasoning as Rosner's R-statistic (RST) procedure (Rosner, 1983). Population mean and variance are computed from sample points that cannot be outliers under either the null or alternative hypotheses. The *MGBT* is a one-sided application of this procedure where only low outliers are believed to exist. In flood flow frequency analysis, high values are not treated as outliers. Low outliers are of concern because by using the logarithms of the flood peaks to fit a distribution, one or more unusual low-flow values can substantially distort the entire fitted frequency distribution. Therefore, the detection of such values is important. In addition, fitted distributions should be compared graphically with the data to check for problems.

The Multiple Grubbs-Beck test is applied to the systematic data of annual peaks from the station record. Let $\{X_1, \ldots, X_n\}$ be a series of logarithms of the annual peak floods. Consider the sorted dataset, $\{X_{[1:n]}, X_{[2:n]}, \ldots, X_{[n:n]}\}$, where $X_{[1:n]}$ is the smallest observation in the sample of size n. The null hypothesis (H_0) is that all observations $\{X_1, \ldots, X_n\}$ are drawn from the same population of independent and identically distributed normal variates. The alternative hypothesis is that the k-th smallest observation in the dataset, $X_{[k:n]}$, is unusually small compared to that population. If $X_{[k:n]}$ is declared a PILF, then all observations less than $X_{[k:n]}$ are also PILFs.

Annual peaks in the data set that are detected as potentially influential are then re-coded as less than a threshold discharge T_{PILF} and treated as interval data in the Expected Moments Algorithm, discussed below. Zero flow values, if observed in the peak-flow data set, are defined as PILFs. Computational details of the *MGBT* algorithm and $p-$values used for determining PILFs are described in Appendix 5. For the case of a single low outlier, the Multiple Grubbs-Beck test is identical to the Grubbs-Beck test (Grubbs and Beck, 1972) that was used in IACWD (1982). Where appropriate, if the *MGBT* does not adequately identify PILFs, the analyst may define a low outlier threshold based on hydrological considerations, knowledge of the watershed, and site characteristics. The justification for a PILF threshold T_{PILF} should be thoroughly documented.

Expected Moments Algorithm

The Expected Moments Algorithm (*EMA*) is a generalized method of moments procedure to estimate the P-III distribution parameters. *EMA* provides a direct fit of the P-III distribution using the entire data set, simultaneously employing regional skew information and a wide range of historical flood and threshold-exceedance information, while adjusting for any potentially influential low floods, missing values from an incomplete record, or zero flood years (Stedinger and Griffis, 2008). *EMA* utilizes multiple types of at-site flood information including Systematic Records, Historical Flood Information, and Paleoflood and Botanical Information. It also includes information about the magnitudes of historical floods and paleofloods, flow intervals, changing base discharges from crest-stage gages, and knowledge of the number of years in the historical period when no large flood occurred, as described in the Section Data Representation using Flow Intervals and Perception Thresholds. *EMA* also directly uses regional flood information (Section Regional Information and Nearby Sites) in the form of a regional skew coefficient G.

EMA is the reasonable extension of the Bulletin 17B LP-III method of moments approach to deal in a consistent statistical framework with all of the sources of information that are likely to be available. There have been numerous studies that document some weaknesses and potential improvements to the moments estimation methods in Bulletin 17B, including historical data, handling of low outliers, use of regional skew, and confidence intervals. Stedinger and Cohn (1986) and Lane (1987) recognized that there are historical and paleoflood data that are not efficiently used by Bulletin 17B. *EMA* was first developed as an alternative to Bulletin 17B (Lane, 1995; Lane and Cohn, 1996; Cohn et al., 1997) in order to fully use historical and paleoflood information (England et al., 2003b,a). *EMA* was then extended to consistently handle low outlier adjustments and regional skew information (Griffis et al., 2004; Griffis, 2008), in addition to historical information. Confidence intervals with *EMA* have been developed (Cohn et al., 2001; Cohn, 2015), as described in the Section Confidence Intervals for Quantiles; thus there is a consistent statistical framework for flood frequency. For simple cases with only a systematic record and a regional skew (see Section Parameter Estimation: Simple Case), the *EMA* algorithm reverts to the method of moments as recommended in IACWD (1982). Additional history, background, and perspectives are presented in Griffis and Stedinger (2007a) and Stedinger and Griffis (2008).

EMA employs the peak-flow intervals $Q_{Y,lower}$ and $Q_{Y,upper}$ to estimate the moments of the LP-III distribution. *EMA* requires the corresponding perception thresholds $T_{Y,lower}$ and $T_{Y,upper}$ to estimate the confidence intervals and other measures of uncertainty in frequency estimates. It is therefore important to estimate the flow intervals and thresholds accurately, based on all the data and information available that is presented in the Section Flood Flow Frequency Information. As described in the Section Data Representation using Flow Intervals and Perception Thresholds, peak-flow intervals and perception thresholds are defined for each data type and each year.

For the general case of a historical perception threshold T_h and a PILF threshold T_{PILF}, the inputs to *EMA* are determined by counting the floods greater than ($>$) (exceedances) and floods less than ($<$) (censored) for each year, relative to each perception threshold. Recall that $X = log_{10}(Q)$, and that X_h and X_{PILF} are the base 10 logarithms of T_h and T_{PILF}, respectively (see also the Glossary for notation and definitions). The logarithms of flood magnitudes are expressed as a union of four sets (Cohn et al., 1997):

$$\{X\} = \{X_s^>\} \cup \{X_h^>\}\{X_s^<\} \cup \{X_h^<\} \tag{13}$$

and where PILFs are identified, the systematic period is divided into floods above and below a PILF threshold X_l (Griffis, 2008), as:

$$\{X_s^<\} = \{X_l^>\} \cup \{X_l^<\} \tag{14}$$

with terms defined in Table 1.

The Expected Moments Algorithm for the general situation with a historical flood perception threshold X_h and a PILF threshold X_l includes the following steps.

1. Perception thresholds for the historical period X_h and PILFs X_l within the systematic period are defined.

2. Using the values that exceeded the thresholds $\{X_h^>\}$ and $\{X_l^>\}$, initial estimates of the sample moments $\{\hat{\mu}_1, \hat{\sigma}_1, \hat{\gamma}_1\}$ are computed as if one had a complete sample.

3. For iteration $i = 1, 2, \ldots$, the parameters of the P-III distribution $\{\hat{\alpha}_{i+1}, \hat{\beta}_{i+1}, \hat{\tau}_{i+1}\}$ are estimated using the previously computed sample moments:

$$\hat{\alpha}_{i+1} = 4/\hat{\gamma}_i \tag{15}$$

$$\hat{\beta}_{i+1} = \left(\frac{1}{2}\right)\hat{\sigma}_i\hat{\gamma}_i \tag{16}$$

$$\hat{\tau}_{i+1} = \hat{\mu}_i - \hat{\alpha}_{i+1}\hat{\beta}_{i+1} \tag{17}$$

4. New sample moments $\{\hat{\mu}_{i+1}, \hat{\sigma}_{i+1}, \hat{\gamma}_{i+1}\}$ are estimated using expected moments.

5. Convergence test – iterate *EMA* steps 3 and 4 until parameter estimates converge.

For example, using the mean, as shown in equation 5, the iteration $i+1$ is:

$$\hat{\mu}_{i+1} = \left(\frac{1}{n}\right)\sum_{i=1}^{n}\tilde{X}_i \tag{18}$$

Table 1. Flow and Year terms used in EMA moments.

Flow or Year	Definition
$\{X_s^>\}$	logarithms of floods that occurred in the systematic record with magnitudes that are greater than the historical threshold X_h
$\{X_h^>\}$	logarithms of historical floods or paleofloods with magnitudes greater than X_h that occurred during the historical period
$\{X_l^>\}$	logarithms of floods that occurred in the systematic record with magnitudes that are greater than the PILF threshold X_l and less than X_h
$\{X_s^<\}$	logarithms of the floods that occurred in the systematic record with magnitudes that are less than X_h and exceed X_l
$\{X_h^<\}$	logarithms of unmeasured historical floods or paleofloods less than X_h, because their magnitudes did not exceed X_h
$\{X_l^<\}$	logarithms of floods in the systematic record that are less than the PILF threshold X_l
$\{n_s^<\}$	number of floods in the systematic record with magnitudes that are less than X_h
$\{n_h^<\}$	number of unmeasured floods in the historical period with magnitudes that are less than X_h
$\{n_l^<\}$	number of floods in the systematic record with magnitudes that are less than X_l

where

$$\tilde{X}_i = \begin{cases} X_i & \text{if } X_i \text{ is measured or "exact"} \\ E[X|X_l < X_i < X_u & \text{if } X_l < X_i < X_u \end{cases} \quad (19)$$

and $E[X|X_l < X_i < X_u]$ is the expected value of an observation known to lie within a range. The equations and computation details for *EMA* are presented in Appendix 6. The *EMA* confidence intervals are described in the Section Confidence Intervals for Quantiles.

Record Extension with Nearby Sites

The minimum record length recommended for frequency analysis in Bulletin 17C is 10 years of annual maximum peak flows. Even with the use of regional skew, historical data and adjustment for low floods, 10 years of record may not be an adequate sample for estimating the more extreme floods like the 0.01 annual exceedance probability flood. Extending records in time is a way of achieving a more representative sample. There are a number of reasons why a short record station may not be representative of long term conditions:

- the short record may represent a wet period where one or more major floods occurred in a short period of time;
- the short record may represent a drought period where no major floods occurred; and
- it may be known that large historical floods occurred prior to or after systematic data collection at the short record station and estimates of these floods need to be incorporated into the frequency analysis.

Record extension involves estimating additional years of record at the short term station utilizing data at a nearby long term station. The estimated annual peak flows are then analyzed along with the observed data in a Bulletin 17C frequency analysis. The recommended approach for record extension is based on the Maintenance of Variance Extension (MOVE) techniques (Hirsch, 1982) with subsequent improvements (Vogel and Stedinger, 1985). The MOVE equations, with an example application, are presented in Appendix 7. A reasonable approach to implement MOVE is to use concurrent data at a nearby long term station that has similar watershed characteristics as the site of interest. There should be at least 10 years of overlapping data for the short record and long record stations and the correlation coefficient needs to exceed a critical value as defined in Appendix 7.

Confidence Intervals for Quantiles

The user of frequency curves should be aware that the curve is only an estimate of the population curve; it is not an exact representation. A streamflow record is only a sample. How well this sample will predict the flood experience (population) depends upon the sample size, its accuracy, and whether or not the underlying distribution is known or chosen wisely.

The record of annual peak flows at a site is a random sample of the underlying population of annual peaks and can be used to estimate the frequency curve of that population. If the same size random sample could be selected from a different period of time, a different estimate of the underlying population frequency curve probably would result. Thus, an estimated flood frequency curve can be only an approximation to the true frequency curve of the underlying population of annual flood peaks. To gauge the accuracy of this approximation, one may construct an interval or range of hypothetical frequency curves that, with a high degree of confidence, contains the population frequency curve. Such intervals are called confidence intervals and their end points are called confidence limits.

Confidence intervals provide either a measure of the uncertainty of the estimated exceedance probability of a selected discharge or a measure of the uncertainty of the discharge at a selected exceedance probability. Confidence intervals on the discharge for the P-III distribution can be estimated using the method described in Appendix 6. The *EMA* with all available data, including historical floods, PILFs, interval data, and regional skew, is used. Uncertainty in the at-site and regional estimates of the skewness coefficients is also included.

Application of confidence intervals in reaching water resource planning decision depends upon the needs of the user. This discussion is presented to emphasize that the frequency curve developed using this guide is only today's best estimate of the flood frequency distribution. As more data become available, the estimate will normally be improved and the confidence intervals narrowed.

Estimating Regional Skew

As described in the Section Weighted Skew Coefficient Estimator, it is recommended that the skew coefficient used be a weighted average of the station skew and a regional skew (Griffis and Stedinger, 2007a). A recommended procedure for estimating regional skew is using the Bayesian Weighted Least Squares/Bayesian Generalized Least Squares (B-WLS/B-GLS) method (Veilleux et al., 2011).

Tasker and Stedinger (1986) developed a weighted least squares (WLS) procedure for estimating regional skewness coefficients based on sample skewness coefficients for the logarithms of annual peak-discharge data. Their method of regional analysis of skewness estimators accounts for the precision of the skewness estimator for each station, which depends on the length of record for each station and the accuracy of an Ordinary Least Squares (OLS) regional mean skewness. More recently, Reis et al. (2005), Gruber et al. (2007), and Gruber and Stedinger (2008) developed a Bayesian generalized least squares (B-GLS) regression model for regional skewness analyses. Use of a GLS model allows the incorporation of the cross-correlation of skewness estimators. Cross-correlation arises as skewness estimators are dependent upon concurrent cross-correlation flood records. The Bayesian method allows for the computation of a posterior distribution of both the regression parameters and the model error variance. As shown in Reis et al. (2005), for cases in which the model error variance is small compared to the sampling error of the at-site estimates, the Bayesian posterior distribution provides a more reasonable description of the model error variance than both the generalized least squares (GLS) method-of-moments and maximum likelihood point estimates (Veilleux, 2011). While WLS regression accounts for the precision of the regional model and the effect of the record length on the variance of skewness coefficient estimators, GLS regression also considers the cross-correlations among the skewness coefficient estimators. The B-GLS regression procedures extend the GLS regression framework by also providing a description of the precision of the estimated model error variance, a pseudo analysis of variance and enhanced diagnostic statistics; see also Griffis and Stedinger (2009).

Due to complexities introduced by the use of the Expected Moments Algorithm (EMA) (Cohn et al., 1997) and large cross-correlations between annual peak discharges at some pairs of gages sites (Parrett et al., 2011), the B-WLS/B-GLS regression procedure was developed to provide both stable and defensible results for regional skewness coefficient models (Veilleux, 2011; Veilleux et al., 2011). It uses an OLS analysis to fit an initial regional skewness model; that OLS model is then used to generate a stable regional skewness coefficient estimate for each site. That stable regional estimate is the basis for computing the variance of each at-site skewness coefficient estimator employed in the WLS analysis. Then, Bayesian WLS is used to generate estimators of the regional skewness coefficient model parameters. Finally, B-GLS is used to estimate the precision of those B-WLS parameter estimators, to estimate the model error variance and the precision of that variance estimator, and to compute various diagnostic statistics including Bayesian plausibility values, pseudo adjusted R-squared, pseudo-Analysis of Variance table, two diagnostic error variance ratios, as well as leverage and influence metrics. This method has been successfully used to generate regional skew estimates around the nation.

It is recommended that regional skew coefficient G estimates and mean-square error MSE_G estimates be obtained from recent studies that use the B-WLS/B-GLS regression procedure completed by the USGS. Appendix 2 contains information regarding recent regional skew studies. In lieu of current published estimates, it is recommended that users consult with the U.S. Geological Survey to determine the availability of regional skew estimates that have been prepared using current methods, described in the Section Estimating Regional Skew. The regional skew estimates published in IACWD (1982, Plate 1) are not recommended for use in flood-frequency studies.

Comparisons of Frequency Curves

Major problems in flood frequency analysis at gaged locations are encountered when making flood estimates for probabilities more rare than defined by the available record. The accuracy of flood probability estimates based upon statistical analysis of flood data deteriorates for probabilities more rare than those directly defined by the at-site flood period of record that may include systematic, historical, and paleoflood data. This is due to several major factors, including the sampling error of the statistics from the station data, because the basic underlying distribution of flood data is not exactly known, and the physical flood processes may change at larger magnitudes.

Although other procedures for estimating floods on a watershed and flood data from adjoining watersheds can sometimes be used for evaluating flood levels at high flows and rare exceedance probabilities, procedures for doing so cannot be standardized to the same extent as the procedures discussed thus far. For these situations the *Guidelines* describe the information to incorporate in the analysis but allow considerable latitude in application.

Frequency curves that are estimated using the recommended procedures in the Section Determination of the Flood Flow Frequency Curve can be compared with frequency curves from similar watersheds using regional frequency methods, or frequency curves from precipitation using rainfall-runoff models. Independent estimates can in some cases be weighted and combined for an improved estimate as described in the Section Weighting of Independent Frequency Estimates. Prior to making comparisons, analysts should ensure that all data and information at the location of interest and within the region, as described in the Section Flood Flow Frequency Information and Appendix 2, have been adequately considered and incorporated into the frequency analysis. In this way, the flood frequency curve may reflect (as appropriate): temporal information such as historical and paleoflood data; spatial information such as regional skew and watershed characteristics; and causal information such as hydroclimate information and mixed-population data. Merz and Blöschl (2008a) and Merz and Blöschl (2008b) describe ways of including and combining various sources of flood frequency information.

The purpose for which the flood frequency information is needed will determine the amount of time and effort that can justifiably be spent to obtain additional data, make comparisons with other watersheds, utilize flood estimates from precipitation, and weight the independent estimates. All types of analyses should be incorporated when estimating flood magnitudes for exceedance probabilities less than 0.01 *AEP*.

The following sections describe the use of additional information to compare and potentially refine the flood frequency analysis using quantile weighting. Recommendations of specific procedures for regional comparisons or for appraising the accuracy of such estimates are beyond the scope of these *Guidelines*.

Comparisons with Similar Watersheds

Comparisons and potentially adjustment of a frequency curve based upon flood experience and flood statistics in nearby hydrologically similar watersheds can improve most flood frequency determinations. Use of the weighted skew coefficient recommended by these *Guidelines* is one form of transferring regional information to the site at hand. Additional comparisons may be helpful and are described in the following paragraphs.

A comparison between flood and extreme storm records such as those in U.S. Army Corps of Engineers (1973) (and others) and flood flow frequency analyses at nearby hydrologically similar watersheds will often aid in evaluating and interpreting both unusual flood experience and the flood frequency analysis of a given watershed. The shorter the flood record and the more unusual a given flood event, the greater will be the need for such comparisons.

When flood frequency curves are available for similar watersheds within a region, comparisons can be made with flood quantiles for selected exceedance probabilities or with the moments of the distribution. Flood quantile estimates from regional quantile regression models that use basin characteristics and physiographic factors, such as Paretti et al. (2014b), are usually available and are recommended for use in comparisons with at-site frequency curves. Regional flood quantile methods have a long history of use (Benson, 1962b,a; Feaster et al., 2009), and have been shown to perform well against alternatives (Griffis and Stedinger, 2007a). Comparisons of quantiles and frequency curve shapes can be made using the index flood method (Dalrymple, 1960; Hosking and Wallis, 1997), which may illustrate similarities or differences in flood runoff mechanisms (Bureau of Reclamation, 2002; England et al., 2010). Comparing regional moment estimates of the mean and standard deviation with at-site estimates is also informative (Griffis and Stedinger, 2007a); regional models of these moments can also be constructed (Gotvald et al., 2012). Simple drainage-area plots and peak-flow envelope curve comparisons can be useful, with an appropriate examination of flood processes and moments within a region (Blöschl and Sivapalan, 1997). If these estimates are independent of the station analysis, a weighted average of the two estimates will be more accurate than either alone. In many situations, the at-site estimate is used in a regional estimate; thus the two estimates are correlated (Moss and Thomas, 1982).

Comparisons with Flood Estimates from Precipitation

Floods and frequency curves developed from precipitation estimates can be used for comparison and to potentially adjust flood frequency curves, including extrapolation beyond experienced values. As described in the Section Flood Estimates from Precipitation, flood estimates from precipitation may be available based on reconstruction of specific flood events, synthetic flood events, or continuous streamflow estimates.

When a flood frequency curve is available from a calibrated rainfall-runoff model for the watershed of interest, comparisons can be made to estimates from the recommended procedures in the Section Determination of the Flood Flow Frequency Curve. Plotting of the flood estimates for a range of exceedance probabilities provides a guide for potentially combining and extrapolating the frequency curve. Quantile variance estimates from the rainfall-runoff model are needed in order to potentially combine estimates. Any potential weighting or combination of frequency curves must recognize the relative accuracy of the flood estimates and the other flood data used in the rainfall-runoff model. Whether or not such effort is warranted depends upon the procedures and data available and on the use to be made of the flood frequency estimates.

Because of the wide variety of rainfall-runoff models, parameters, and inputs, no specific procedures are recommended. Appraisal of the techniques to use flood estimates from rainfall-runoff models is currently outside the scope of these *Guidelines*, and future work is warranted in this area. Consequently, alternative procedures for making such studies or criteria for deciding when available flood records should be combined or extended by such procedures have not been evaluated.

Weighting of Independent Frequency Estimates

When flood frequency estimates are available from similar watersheds or from rainfall-runoff models and they are independent of the at-site estimates made using the procedures described in the Section Determination of the Flood Flow Frequency Curve, these flood quantile estimates $\hat{Q}_q$ may be weighted and combined. The weights are based on quantile variance and are assumed to be unbiased and independent. The weight given to each estimate is inversely proportional to its variance; Appendix 8 describes the recommended weighting method and provides an example.

It is recommended that weighting be done when dependable estimates of flood quantiles and the variances of quantiles are available. Prior to weighting and combining estimates, the quantiles and variances of the estimates need to be evaluated. Flood quantile estimates may be substantially different for a variety of reasons (Rogger et al., 2012). In some situations, highly-variable estimates (for example, from rainfall-runoff models) may be unreliable and should not be weighted, as they would degrade the at-site estimate.

Griffis and Stedinger (2007a) recently evaluated several weighting methods, including quantile weighting and moment weighting with 2 and 3 parameters, among other alternatives. As described in the Section Regional Information and Nearby Sites, regional mean and standard deviation estimates may be available. These moments could be considered in weighting frequency curves. The computational study by Griffis and Stedinger (2007a) demonstrates that the simple weighting of at-site and regional regression quantile estimates performs nearly as well as more complex alternatives, and for short records provides a substantial improvement in quantile accuracy. Quantile weighting, described in Appendix 8, is the recommended approach.

Analysts are encouraged to include flood frequency information from all sources, as appropriate. In some cases, information from numerous sources can be combined (Viglione et al., 2013). Other than the procedure recommended in Appendix 8, these methods have not been fully evaluated.

Software and Examples

Specialized software has been developed by various agencies that implements the recommended flood-frequency procedures in these *Guidelines*. This includes estimating the log-Pearson Type III distribution parameters using *EMA* with available historical and paleoflood data, PILFs, and regional skew information. Confidence intervals and plotting positions are also estimated. The software includes the methods and computations presented in the Section Determination of the Flood Flow Frequency Curve, PILFs described in Appendix 5 and *EMA* described in Appendix 6. A list of recommended software packages is provided on the HFAWG web page at http://acwi.gov/hydrology/Frequency/b17c/.

The initial data analysis (Appendix 3) and record extension techniques (Appendix 7) can be performed without the need for specialized software. Available ancillary materials and examples are provided on the HFAWG web page.

Some representative flood frequency examples that illustrate the recommended methods described in the Section Determination of the Flood Flow Frequency Curve are presented in Appendix 9. The main emphasis is on the data, flow intervals, and threshold inputs to *EMA*. The seven examples include: a systematic record; potentially-influential low floods record; a broken record; a historical record; a crest-stage record; a historical and PILF record; and a paleoflood record. Each example includes a detailed description of the data, a time series plot, and a flood frequency curve. Input and output files from software used to create the examples are also available on the HFAWG web page at http://acwi.gov/hydrology/Frequency/b17c/. These examples are meant to illustrate the main concepts presented in these *Guidelines*, and are not meant to be all-inclusive.

Future Studies

These *Guidelines* are designed to meet a current, ever-pressing demand that the Federal Government develop a coherent set of procedures for accurately defining flood potentials as needed in programs of flood damage abatement. Much additional study and data are required before the twin goals of accuracy and consistency will be obtained. It is hoped that this guide contributes to this effort by defining the essential elements of a coherent set of procedures for flood frequency determination. Although selection of the analytical procedures to be used in each step or element of the analysis has been carefully made based upon a review of the literature, the considerable practical experience of Work Group members, and special studies conducted to aid in the selection process, the need for additional studies is recognized.

The following is a list of some additional needed topics of study identified by the Work Group:

1. the identification and treatment of mixed distributions, including those based on hydrometeorological or hydrological conditions;

2. guides for defining flood potentials for ungaged watersheds and watersheds with limited gaging records as described below;

3. methods to include watershed hydrological processes and physical considerations into the analysis that can influence the frequency curve;

4. procedures for improving flood frequency analysis using precipitation data, rainfall-runoff models, and associated uncertainty analysis;

5. guides for defining flood potentials for watersheds altered by urbanization, wildfires, deforestation, and by reservoirs as described below;

6. guides for estimating dynamic flood frequency curves that vary with time, incorporating climate indices, changing basin characteristics, and addressing potential nonstationary climate conditions;

7. frequency estimation in cases where long-term trends are evident in the data but are not readily explainable by the history of land use, land use practices, or engineering modifications of the river or floodplain; and

8. an examination and redefinition of risk, reliability, and return periods under nonstationary conditions.

There is a need to develop guidance in three important areas: ungaged sites, regulated flow frequency, and urbanization. Some existing practices are listed below for each area; however no specific recommendations and guidance is made. While significant work has been done on these topics by researchers around the world, those efforts have not yet been evaluated for broad and systematic application as contemplated in these *Guidelines*.

Ungaged Sites

Many of the stream sites of interest do not have gages with sufficient records or are ungaged. One area of future work needed is to develop national guidance on methods for estimating flood flow frequency curves at ungaged sites. Two common methods that are used to estimate frequency curves for ungaged watersheds are (Thomas et al., 2001): (1) regional flood quantile regression equations based on generalized least squares (Tasker and Stedinger, 1989); and (2) rainfall-runoff models (Pilgrim and Cordery, 1993; McCuen, 2004). Regional regression equations are available through the USGS StreamStats software (Ries et al., 2008). A limited comparison of these two methods is in Thomas et al. (2001).

Regulated Flow Frequency

A large portion of the stream sites of interest have flows that are altered to some degree by regulating structures such as dams, reservoirs, and diversions, or flows are affected by levees. One area of future work needed is to develop national guidance on methods for estimating flood flow frequency curves at stream locations affected by varying degrees of regulation. Some common regulated flood frequency methods include estimating unregulated flows using empirical relationships or synthetic floods (U.S. Army Corps of Engineers, 1993), or by applying total probability concepts (Kubik, 1990; Sanders et al., 1990). Durrans (2002) summarizes these approaches and describes other methods that could be considered.

Urbanization and Watershed Change

At many stream sites of interest, flood-frequency relationships may be changing due to alterations within watershed and the stream corridor over time. This may be due to urbanization (Konrad, 2003), land development, and other factors described in the section Watershed Changes. National guidance for estimating flood flow frequency curves in watersheds experiencing urbanization and/or watershed change is an area needing further work. One option is to develop flood-frequency regression equations that include urbanization factors (Sauer et al., 1983). Other approaches for estimating flood frequency for watersheds undergoing landuse change are in McCuen (2003).

Applicability of These Guidelines

Bulletin 17C goes a long way towards addressing known concerns with Bulletin 17B. However, many concerns remain, such as the best methods of addressing regulated flows and mixed distributions, methods for addressing urbanizing areas and other land-use changes, better ways to use information provided by rainfall records and rainfall-frequency analyses, and better use of physiographic watershed characteristics to define the flood-flow frequency relationship. How to handle climate change and climate variability will be continuing concerns as science comes to better understand the likely impact of such atmospheric phenomena on hydrologic processes. Development of flood-flow frequency relationships between gaged and ungaged sites is an important topic not addressed here.

While many improvements have been made, there are significant limitations that apply to use of procedures recommended in these *Guidelines*. First and foremost, these *Guidelines* are predicated on the availability of flood data that constitute a reliable, representative, and homogeneous sample of expected future floods. Flood data that represent unique occurrences such as dam failures, ice jams, or importation or diversion of flood waters should not be used to characterize flood potential unless they are properly adjusted to represent prevailing (natural) watershed conditions. There are currently many concerns about potential changes in the distribution of floods due to watershed changes and anthropogenic climate change; such concerns may require special procedures as discussed in the Section Data Assumptions and Specific Concerns.

These *Guidelines* assume the use of the annual-maximum flood series and generally apply only to portions of the flood-frequency curve for *AEPs* less than 0.10. Flood-frequencies for larger, more common *AEPs* may be more appropriately determined from use of the partial-duration series data, that allow for more than one large flood per year rather than the annual-maximum flood series. Some procedures for these analyses are mentioned in the Section Flood Flow Frequency Information.

These *Guidelines* apply only to those situations for which there are sufficient data for carrying out the necessary computations. In general, flood-frequency computations are not reliable with records comprised of less than 10 annual flood observations. Accurate determination of floods for small *AEPs* ($<$0.01) generally requires more data; estimations of floods for *AEPs* smaller than 0.005 generally require augmentation of the systematically observed flood records with general regional information, insight from precipitation records, or paleoflood information, as available (Section Flood Flow Frequency Information).

These *Guidelines* permit augmentation of flood records by incorporation of community experience such as the documentation of floods in news reports, community accounts, or paleoflood indicators (see the Sections Historical Flood Information and Paleoflood and Botanical Information and Appendix 2). However, these conditions must be properly described by specification of accurate observation intervals and thresholds based upon consideration of the physical flood indicators and hydraulic conditions. These considerations must be well documented by a qualified analyst together with the necessary computations.

The *Guidelines* may be used to estimate flood-frequencies for urban conditions where there are flood observation datasets of sufficient length that represent stable development or that can be adjusted to account for changes in urban infrastructure and routing parameters (Sauer et al., 1983; McCuen, 2003). Similarly, any regional skewness estimator should be derived from flood records representing urban conditions.

These *Guidelines* describe the set of procedures recommended for defining flood potential as expressed by a flood flow frequency curve. Special situations may require other approaches, perhaps defining the frequency relationship for flood volumes or river stages. In those cases where the procedures of this guide are not followed, deviations must be supported by appropriate study, including a comparison of the results obtained with those obtained using the recommended procedures.

There is much concern about changes in flood risk associated with climate variability and long-term climate change. Time invariance was assumed in the development of this guide. In those situations where there is sufficient scientific evidence to facilitate quantification of the impact of climate variability or change in flood risk, this knowledge should be incorporated in flood frequency analysis by employing time-varying parameters or other appropriate techniques. All such methods employed need to be thoroughly documented and justified.

It is not anticipated that many special situations warranting other approaches will occur at sites which have reasonable flood flow records. These procedures should be followed unless there are compelling technical reasons for departing from the *Guidelines*. These deviations are to be documented and supported by appropriate study, including comparison of results. The Subcommittee on Hydrology requests that these situations be called to its attention for consideration in future modifications of these *Guidelines*.

References

Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions*. National Bureau of Standards, Washington.

Aldridge, B. N. and Eychaner, J. H. (1984). *Floods of October 1977 in southern Arizona and March 1978 in central Arizona*. U.S. Geological Survey Water-Supply Paper 2243, Reston, VA.

Alila, Y. and Mtiraoui, A. (2002). Implications of heterogeneous flood-frequency distributions on traditional stream-discharge prediction techniques. *Hydrological Processes*, 16(5):1065–1084.

Asquith, W. H. and Slade, R. M. (1995). Documented and potential extreme peak discharges and relation between potential extreme peak discharges and probable maximum flood peak discharges in Texas. Water-Resources Investigations Report 95-4249, U.S. Geological Survey, Reston, VA.

Baker, V. (2013). Global Late Quaternary Fluvial Paleohydrology: With Special Emphasis on Paleofloods and Megafloods. In Shroder, J. F. and Wohl, E., editors, *Treatise on Geomorphology*, volume 9.26, pages 511 – 527. Academic Press, San Diego.

Baker, V. R. (1987). Paleoflood hydrology and extraordinary flood events. *Journal of Hydrology*, 96(1-4):79–99.

Baker, V. R. (2008). Paleoflood hydrology: Origin, progress, prospects. *Geomorphology*, 101(1-2):1–13.

Baker, V. R., Kochel, R. C., and Patton, P. C. (1988). *Flood Geomorphology*. John Wiley and Sons, New York, 1st edition.

Baker, V. R., Webb, R. H., and House, P. K. (2002). The scientific and societal value of paleoflood hydrology. In House, P. K., Webb, R. H., Baker, V. R., and Levish, D. R., editors, *Ancient Floods, Modern Hazards*, volume 5 of *Water Science and Application Series*, pages 1–19. American Geophysical Union.

Bathurst, J. (1986). Slope-area discharge gaging in mountain rivers. *Journal of Hydraulic Engineering*, 112(5):376–391.

Beard, L. R. (1974). Flood flow frequency techniques. Technical Report CRWR-119, Center for Research in Water Resources.

Benito, G. and O'Connor, J. (2013). Quantitative Paleoflood Hydrology. In Shroder, J. F. and Wohl, E., editors, *Treatise on Geomorphology*, volume 9.24, pages 459 – 474. Academic Press, San Diego.

Benson, M. A. (1962a). Factors Affecting the Occurrence of Floods in the Southwest. Water-Supply Paper 1580-D, U.S. Geological Survey, Reston, VA.

Benson, M. A. (1962b). Factors Influencing the Occurrence of Floods in a Humid Region of Diverse Terrain. Water-Supply Paper 1580-B, U.S. Geological Survey, Reston, VA.

Benson, M. A. and Dalrymple, T. (1967). General field and office procedures for indirect discharge measurements. Techniques and Methods Book 3, Chapter A1, U.S. Geological Survey.

Beven, K. J. (2001). *Rainfall-Runoff Modeling, The Primer*. John Wiley and Son, Chichester, 1st edition.

Blainey, J., Webb, R., Moss, M., and Baker, V. (2002). Bias and information content of paleoflood data in flood-frequency analysis. *Water Science and Application*, 5:161–174.

Blöschl, G. and Sivapalan, M. (1997). Process controls on regional flood frequency: Coefficient of variation and basin scale. *Water Resources Research*, 33(12):2967–2980.

Brekke, L. D., Kiang, J. E., Olsen, J. R., Pulwarty, R. S., Raff, D. A., Turnipseed, D. P., Webb, R. S., and White, K. D. (2009). *Climate change and water resources management – A Federal perspective*. U.S. Geological Survey Circular 1331, Reston, VA.

Bureau of Reclamation (2002). Flood Hazard Analysis - Folsom Dam, Central Valley Project, California. Dam safety technical report, U.S. Department of Interior, Bureau of Reclamation, Denver, CO.

Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B., and Wilby, R. (2004). The Schaake shuffle: a method for reconstructing space-time variability in forecasted precipitation and temperature fields. *Journal of Hydrometeorology*, 5(1):243–262.

Cohen, A. C. (1991). *Truncated and Censored Samples - Theory and Application*. Marcel-Dekker, New York.

Cohn, T. A. (2015). A fast numerical algorithm for computing accurate confidence intervals for EMA-based flood quantile estimates. *(to be submitted to Water Resources Research)*.

Cohn, T. A., Barth, N. A., England, J. F., Faber, B., Mason, R. R., and Stedinger, J. R. (2014). Evaluation of Recommended Revision to Bulletin 17B. Scientific Investigations Report Draft 2.5 2014-XXXX, U.S. Geological Survey, Reston, VA.

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*, 49(8):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*, 33(9):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*, 37(6):1695–1706.

Cohn, T. A. and Lins, H. F. (2005). Nature's style: Naturally trendy. *Geophysical Research Letters*, 32(23):n/a–n/a.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer.

Condie, R. and Lee, K. A. (1982). Flood frequency analysis with historic information. *Journal of Hydrology*, 58:47–61.

Cook, J. L. (1987). Quantifying peak discharges for historical floods. *Journal of Hydrology*, 96(14):29 – 40.

Costa, J. E. (1978). Holocene stratigraphy in flood frequency analysis. *Water Resources Research*, 14(4):626–632.

Costa, J. E. (1986). A history of paleoflood hydrology in the United States, 1800-1970. *EOS, Transactions of the American Geophysical Union*, 67(17):425, 428–430.

Costa, J. E. (1987). A history of paleoflood hydrology in the United States, 18001970. In Gillmor, C. S., Landa, E. R., Ince, S., and Back, W., editors, *History of Geophysics: Volume 3*, volume 3, pages 49–53. American Geophysical Union.

Costa, J. E. and Baker, V. R. (1981). *Surficial Geology: Building With the Earth*. Wiley, 1st edition.

Costa, J. E. and Jarrett, R. D. (2008). An evaluation of selected extraordinary floods in the United States reported by the U.S. Geological Survey and implications for future advancement of flood science. Scientific Investigations Report 2008-5164, U.S. Geological Survey, Reston, VA.

Crippen, J. R. (1978). Composite Log-Type III Frequency-Magnitude Curve of Annual Floods. Open-File Report 78-352, U.S. Geological Survey.

Crippen, J. R. and Bue, C. D. (1977). *Maximum floodflows in the conterminous United States*. U.S. Geological Survey Water-Supply Paper 1887, Reston, VA.

Crowfoot, R. M., Paillett, A. V., Ritz, G. F., Smith, M. E., Jenkins, R. A., and O'Neill, G. B. (1997). Water resources data, Colorado, Water Year 1996, volume 2. Colorado River Basin. Water-Data Report WDR-CO-96-2, U.S. Geological Survey, Reston, VA.

Dalrymple, T. (1960). Flood-frequency analyses. manual of hydrology: Part 3. flood flow techniques. Water-Supply Paper 1543-A, U.S. Geological Survey, Reston, VA.

Dalrymple, T. and Benson, M. A. (1967). Measurement of peak discharge by the slope-area method. Techniques and Methods Book 3, Chapter A2, U.S. Geological Survey.

David, H. A. (1981). *Order Statistics*. Wiley, 2nd edition.

Duan, Q., Gupta, H. V., Sorooshian, S., Rousseau, A. N., and Turcotte, R. (2003). *Calibration of Watershed Models*, volume 6 of *Water Science and Application Series*. American Geophysical Union.

Durlin, R. R. and Schaffstall, W. P. (2002). Water resources data, Pennsylvania, Water Year 2002, volume 2. Susquehanna River Basin. Water-Data Report WDR-PA-02-2, U.S. Geological Survey, Reston, VA.

Durrans, S. R. (2002). Regulated flood frequency methods. Informal handouts, Hydrologic Frequency Analysis Work Group, Hydrology Subcommittee, Interagency Advisory Committee on Water Data.

Eash, D. A. (2010). Estimating flood frequency. In Mutel, C. F., editor, *A Watershed Year: Anatomy of the Iowa Floods of 2008*, chapter 7, pages 61–70. University of Iowa Press, Iowa City.

Eash, D. A., Barnes, K. K., and Veilleux, A. G. (2013). Methods for Estimating Annual Exceedance-Probability Discharges for Streams in Iowa, Based on Data through Water Year 2010. Scientific Investigations Report 2013-5086, U.S. Geological Survey, Reston, VA.

Elliott, J. G., Jarrett, R. D., and Ebling, J. L. (1982). Annual snowmelt and rainfall peak-flow data on selected foothills region streams, South Platte River, Arkansas River, and Colorado River Basins, Colorado. Open-File Report 82-426, U.S. Geological Survey.

England, J., Godaire, J., Klinger, R., Bauer, T., and Julien, P. (2010). Paleohydrologic bounds and extreme flood frequency of the upper arkansas river, colorado, usa. *Geomorphology*, 124(1):1–16.

England, J. F. (1998). Assessment of historical and paleohydrologic information in flood frequency analysis. Master's thesis, Colorado State University.

England, J. F., Jarrett, R. D., and Salas, J. D. (2003a). Data-based comparisons of moments estimators using historical and paleoflood data. *Journal of Hydrology*, 278(1):172–196.

England, J. F., Klawon, J. E., Klinger, R. E., and Bauer, T. R. (2006). Flood hazard analysis for Pueblo dam, Colorado, final report. Dam safety technical report, U.S. Department of Interior, Bureau of Reclamation, Denver, CO.

England, J. F., Salas, J. D., and Jarrett, R. D. (2003b). Comparisons of two moments-based estimators that utilize historical and paleoflood data for the log Pearson type III distribution. *Water Resources Research*, 39(9):1243.

England Jr, J. F. and Cohn, T. A. (2007). Scientific and practical considerations related to revising Bulletin 17B: The case for improved treatment of historical information and low outliers. In *American Society of Civil Engineers, EWRI World Water & Environmental Resources Congress*.

England Jr, J. F. and Cohn, T. A. (2008). Bulletin 17B Flood Frequency Revisions: Practical Software and Test Comparison Results. In *World Environmental and Water Resources Congress 2008 Ahupua*, pages 1–11. ASCE.

Feaster, T. D., Gotvald, A. J., and Weaver, J. C. (2009). Magnitude and Frequency of Rural Floods in the Southeastern United States, 2006: Volume 3, South Carolina. Scientific Investigations Report 2009-5156, U.S. Geological Survey, Reston, VA.

FEMA (2009). Guidelines and Specifications for Flood Hazard Mapping Partners. Appendix C: Guidance for Riverine Flooding Analyses and Mapping. Technical report, Federal Emergency Management Agency.

Follansbee, R. and Jones, E. E. (1922). The Arkansas River floods of June 3-5, 1921. Water-Supply Paper 487, U.S. Geological Survey, Reston, VA.

Follansbee, R. and Sawyer, L. R. (1948). Floods in Colorado. Water-Supply Paper 997, U.S. Geological Survey, Reston, VA.

Gerard, R. and Karpuk, E. W. (1979). Probability analysis of historical flood data. *Journal of Hydraulic Engineering*, 105(HY9):1153–1165.

Godaire, J. E. and Bauer, T. R. (2012). Paleoflood Study, North Fork Red River Basin near Altus Dam, Oklahoma. Technical Memorandum No.

86-68330-2012-14, U.S. Department of Interior, Bureau of Reclamation, Denver, CO.

Godaire, J. E. and Bauer, T. R. (2013). Paleoflood Study on the Rio Chama near El Vado Dam, New Mexico. Technical Memorandum No. 86-68330-2013-21, U.S. Department of Interior, Bureau of Reclamation, Denver, CO.

Godaire, J. E., Bauer, T. R., and Klinger, R. E. (2012). Paleoflood Study, San Joaquin River near Friant Dam, California. Technical Memorandum No. 86-68330-2012-24, U.S. Department of Interior, Bureau of Reclamation, Denver, CO.

Gotvald, A. J., Barth, N. A., Veilleux, A. G., and Parrett, C. (2012). Methods for Determining Magnitude and Frequency of Floods in California, Based on Data through Water Year 2006. Scientific Investigations Report 2012-5113, U.S. Geological Survey, Reston, VA.

Gotvald, A. J., Feaster, T. D., and Weaver, J. C. (2009). Magnitude and Frequency of Rural Floods in the Southeastern United States, 2006: Volume 1, Georgia. Scientific Investigations Report 2009-5043, U.S. Geological Survey, Reston, VA.

Griffis, V. and Stedinger, J. (2007a). Evolution of Flood Frequency Analysis with Bulletin 17. *Journal of Hydrologic Engineering*, 12(3):283–297.

Griffis, V. and Stedinger, J. (2007b). Log-Pearson Type 3 Distribution and Its Application in Flood Frequency Analysis. I: Distribution Characteristics. *Journal of Hydrologic Engineering*, 12(5):482–491.

Griffis, V. and Stedinger, J. (2007c). Log-Pearson Type 3 Distribution and Its Application in Flood Frequency Analysis. II: Parameter Estimation Methods. *Journal of Hydrologic Engineering*, 12(5):492–500.

Griffis, V. and Stedinger, J. (2009). Log-Pearson Type 3 Distribution and Its Application in Flood Frequency Analysis. III: Sample Skew and Weighted Skew Estimators. *Journal of Hydrologic Engineering*, 14(2):121–130.

Griffis, V. W. (2008). EMA with Historical Information, Low Outliers, and Regional Skew. In *World Environmental and Water Resources Congress 2008 Ahupua*, pages 1–10. ASCE.

Griffis, V. W. and Stedinger, J. R. (2007d). The use of GLS regression in regional hydrologic analyses. *Journal of Hydrology*, 344(1-2):82–95.

Griffis, V. W., Stedinger, J. R., and Cohn, T. A. (2004). Log Pearson type 3 quantile estimators with regional skew information and low outlier adjustments. *Water Resources Research*, 40(7):W07503.

Grover, N. C. (1937). *The floods of March 1936, Part 3, Potomac, James, and upper Ohio Rivers*. U.S. Geological Survey Water-Supply Paper 800, Reston, VA.

Grover, N. C. and Harrington, A. W. (1943). *Stream Flow*. John Wiley and Sons, New York.

Grubbs, F. and Beck, G. (1972). Extension of sample sizes and percentage points for significance tests of outlying observations. *Technometrics*, 14(4):847–854.

Gruber, A. M., Reis, D. S., and Stedinger, J. R. (2007). Models of regional skew based on Bayesian GLS regression. In Kabbes, K. C., editor, *World Environmental & Water Resources Conference, Tampa, FL*. ASCE EWRI.

Gruber, A. M. and Stedinger, J. R. (2008). Models of LP3 regional skew, data selection, and Bayesian GLS regression. In *World Environmental & Water Resources Conference 2008 AHUPUA'A*.

Harden, T. M., O'Connor, J. E., Driscoll, D. G., and Stamm, J. F. (2011). Flood-frequency analyses from paleoflood investigations for Spring, Rapid, Boxelder, and Elk Creeks, Black Hills, western South Dakota. Scientific Investigations Report 2011-5131, U.S. Geological Survey, Reston, VA.

Hazen, A. (1930). *Flood Flows*. John Wiley and Sons, New York, 1st edition.

Helsel, D. R. and Hirsch, R. (1992). *Statistical methods in water resources*. Elsevier.

Hirsch, R. (1982). A comparison of four streamflow record extension techniques. *Water Resources Research*, 18(4):1081–1088.

Hirsch, R. M. (1987). Probability plotting position formulas for flood records with historical information. *Journal of Hydrology*, 96(1-4):185–199.

Hirsch, R. M. and DeCicco, L. A. (2015). User Guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R Packages for Hydrologic Data. Techniques and Methods Book 4, Chapter A10, U.S. Geological Survey.

Hirsch, R. M., Helsel, D. R., Cohn, T. A., and Gilroy, E. J. (1993). Statistical analysis of hydrologic data. In Maidment, D. R., editor, *Handbook of Hydrology*, chapter 17, pages 17.1–17.55. McGraw Hill, Inc.

Hirsch, R. M. and Stedinger, J. R. (1987). Plotting positions for historical floods and their precision. *Water Resources Research*, 23(4):715–727.

Hirschboeck, K. K. (1987a). Catastrophic flooding and atmospheric circulation anomalies. In Mayer, L. and Nash, D., editors, *Catastrophic Flooding*, chapter 2, pages 23–56. Allen and Unwin, Boston.

Hirschboeck, K. K. (1987b). Hydroclimatically-defined mixed distributions in partial duration flood series. In Singh, V. P., editor, *Hydrologic Frequency Modeling*, pages 199–212, Louisiana State University, Baton Rouge. D. Reidel Publishing Company.

Hirschboeck, K. K. (1991). Climate and floods. In Paulson, R., Chase, E., Roberts, R., and Moody, D., editors, *National Water Summary 1988-89-Hydrologic Events and Floods and Droughts, U.S. Geological Survey Water-Supply Paper 2375*, pages 67–88. U.S. Geological Survey, Reston, VA.

Hosking, J. R. M. and Wallis, J. R. (1997). *Regional Frequency Analysis – An Approach Based on L-Moments*. Cambridge University Press.

House, P. K. and Baker, V. R. (2001). Paleohydrology of flash floods in small desert watersheds in western Arizona. *Water Resources Research*, 37(6):1825–1839.

House, P. K., Pearthree, P. A., and Klawon, J. E. (2002a). Historical flood and paleoflood chronology of the lower Verde River, Arizona: Stratigraphic evidence and related uncertainties. In House, P. K., Webb, R. H., Baker, V. R., and Levish, D. R., editors, *Ancient Floods, Modern Hazards*, volume 5 of *Water Science and Application Series*, pages 267–293. American Geophysical Union.

House, P. K., Webb, R. H., Baker, V. R., and Levish, D. R. (2002b). *Ancient Floods, Modern Hazards*, volume 5 of *Water Science and Application Series*. American Geophysical Union.

Hupp, C. R. (1987). Botanical Evidence of Floods and Paleoflood History. In Singh, V. P., editor, *Regional Flood Frequency Analysis*, pages 355–369. D. Reidel.

Hupp, C. R. (1988). Plant ecological aspects of flood geomorphology and paleoflood history. In Baker, V. R., Kochel, R. C., and Patton, P. C., editors, *Flood Geomorphology*, pages 335–356. John Wiley and Sons.

Hupp, C. R. and Osterkamp, W. R. (1996). Riparian vegetation and fluvial geomorphic processes. *Geomorphology*, 14(4):277–295.

Hydrologic Frequency Analysis Work Group (2006). Flood Frequency Research Needs. Memorandum to Subcommittee on Hydrology.

Hydrologic Frequency Analysis Work Group (2013). Recommendations Memorandum to the Subcommittee on Hydrology, Hydrologic Frequency Analysis Work Group Meeting Minutes.

IACWD (1982). Guidelines for determining flood flow frequency, Bulletin 17-B. Technical report, Interagency Committee on Water Data, Hydrology Subcommittee.

Jain, S. and Lall, U. (2001). Floods in a changing climate: Does the past represent the future? *Water Resources Research*, 37(12):3193–3205.

Jarrett, R. D. (1987). Errors in slope-area computations of peak discharges in mountain streams. *Journal of Hydrology*, 96(14):53 – 67.

Jarrett, R. D. (1991). Paleohydrology and its value in analyzing floods and droughts. In Paulson, R., Chase, E., Roberts, R., and Moody, D., editors, *National Water Summary 1988-89-Hydrologic Events and Floods and Droughts, U.S. Geological*

Survey Water-Supply Paper 2375, pages 105–116. U.S. Geological Survey, Reston, VA.

Jarrett, R. D. and Costa, J. E. (1988). Evaluation of the flood hydrology in the Colorado Front Range using precipitation, streamflow, and paleoflood data for the Big Thompson River Basin. Water-Resources Investigations Report 87-4117, U.S. Geological Survey, Reston, VA.

Jarrett, R. D. and England, J. F. (2002). Reliability of paleostage indicators for paleoflood studies. In House, P. K., Webb, R. H., Baker, V. R., and Levish, D. R., editors, *Ancient Floods, Modern Hazards*, volume 5 of *Water Science and Application Series*, pages 91–109. American Geophysical Union.

Jarrett, R. D. and Malde, H. E. (1987). Paleodischarge of the late pleistocene Bonneville flood, Snake River, Idaho, computed from new evidence. *Geological Society of America Bulletin*, 99:127–134.

Jarrett, R. D. and Tomlinson, E. M. (2000). Regional interdisciplinary paleoflood approach to assess extreme flood potential. *Water Resources Research*, 36(10):2957–2984.

Kiang, J. E., Olsen, J. R., and Waskom, R. M. (2011). Introduction to the featured collection on "nonstationarity, hydrologic frequency analysis, and water management". *JAWRA Journal of the American Water Resources Association*, 47(3):433–435.

Kirby, W. (1987). Linear error analysis of slope-area discharge determinations. *Journal of Hydrology*, 96(14):125 – 138.

Kite, G. W. (1988). *Frequency and Risk Analyses in Hydrology*. Water Resources Publications, Littleton, Colorado.

Kite, J. S., Gebhardt, T. W., and Springer, G. S. (2002). Slackwater deposits as paleostage indicators in canyon reaches of the Central Appalachians: Reevaluation after the 1996 Cheat River flood. In House, P. K., Webb, R. H., Baker, V. R., and Levish, D. R., editors, *Ancient Floods, Modern Hazards*, volume 5 of *Water Science and Application Series*, pages 257–266. American Geophysical Union.

Kjeldsen, T. R., Lamb, R., and Blazkova, S. D. (2014). Uncertainty in flood frequency analysis. In Beven, K. and Hall, J., editors, *Applied Uncertainty Analysis for Flood Risk Management*, chapter 8, pages 153–197. Imperial College Press.

Klemeš, V. (1986). Dilettantism in hydrology: Transition or destiny? *Water Resources Research*, 22(9):177S–188S.

Klemeš, V. (1987). Hydrological and engineering relevance of flood frequency analysis. In Singh, V. P., editor, *Hydrologic Frequency Modeling*, pages 1–18, Louisiana State University, Baton Rouge. D. Reidel Publishing Company.

Klemeš, V. (2000). Tall Tales about Tails of Hydrological Distributions. I. *Journal of Hydrologic Engineering*, 5(3):227–231.

Klinger, R. E. and Bauer, T. R. (2010). Paleoflood Study on the South Fork of the Boise River for Anderson Ranch Dam, Idaho. Technical Report SRH-2010-6, U.S. Department of Interior, Bureau of Reclamation, Denver, CO.

Klinger, R. E. and Godaire, J. E. (2002). Development of a Paleoflood Database for Rivers in the Western U.S. Dam safety technical report, U.S. Department of Interior, Bureau of Reclamation, Denver, CO.

Kochel, R. C., Baker, V. R., and Patton, P. C. (1982). Paleohydrology of southwestern Texas. *Water Resources Research*, 18(4):1165–1183.

Kohn, M. S., Jarrett, R. D., Krammes, G. S., and Mommandi, A. (2013). Web-based flood database for Colorado, water years 1867 through 2011. Open-File Report 2012-1225, U.S. Geological Survey.

Konrad, C. P. (2003). Effects of urban development on floods. Fact Sheet 076-03, U.S. Geological Survey.

Koutsoyiannis, D. (2011). Hurst-kolmogorov dynamics and uncertainty. *JAWRA Journal of the American Water Resources Association*, 47(3):481–495.

Kubik, H. E. (1990). Computation of Regulated Frequency Curves by Application of the Total Probability Theorem. Technical Report Research Memo No. 46, U.S. Army Corps of Engineers, Hydrologic Engineering Center.

Kuczera, G. (1982). Robust flood frequency models. *Water Resources Research*, 18(2):315–324.

Kuczera, G. (1996). Correlated rating curve error in flood frequency inference. *Water Resources Research*, 32(7):2119–2127.

Kuichling, E. (1917). Discussion of flood flows by W.E. Fuller. *Transactions, ASCE*, 77(1293):643–663.

Lamontagne, J., Stedinger, J., Yu, X., Whealton, C., and Xu, Z. (2015). Robust Flood Frequency Analysis: Performance of EMA with Multiple Grubbs-Beck Outlier Tests. *submitted to Water Resources Research*, page 39.

Lamontagne, J. R., Stedinger, J. R., Berenbrock, C., Veilleux, A. G., Ferris, J. C., and Knifong, D. L. (2012). Development of Regional Skews for Selected Flood Durations for the Central Valley Region, California, Based on Data Through Water Year 2008. Scientific Investigations Report 2012-5130, U.S. Geological Survey, Reston, VA.

Lamontagne, J. R., Stedinger, J. R., Cohn, T. A., and Barth, N. A. (2013). Robust national flood frequency guidelines: What is an outlier? In *Showcasing the Future*, pages 2454–2466. ASCE.

Lane, W. L. (1987). Paleohydrologic data and flood frequency estimation. In Singh, V. P., editor, *Application of Frequency and Risk in Water Resources*, pages 287–298. D. Reidel.

Lane, W. L. (1995). Method of moments approach to historical data. Informal handout, Bulletin 17B Working Group, Hydrology Subcommittee, Interagency Advisory Committee on Water Data.

Lane, W. L. and Cohn, T. A. (1996). Expected moments alogrithms for flood frequency analysis. In *North American Water and Environment Congress & Destructive Water*, pages 2185–2190. ASCE.

Lang, M., Ouarda, T., and Bobée, B. (1999). Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, 225(3-4):103–117.

Langbein, W. B. (1949). Annual floods and the partial-duration flood series. *Transactions of the American Geophysical Union*, 30(6):879–881.

Langbein, W. B. and Iseri, K. T. (1960). General introduction and hydrologic definitions. manual of hydrology: Part 1. general surface water techniques. Water-Supply Paper 1541-A, U.S. Geological Survey, Reston, VA.

Leese, M. N. (1973). Use of censored data in the estimation of gumbel distribution parameters for annual maximum flood series. *Water Resources Research*, 9(6):1534–1542.

Levish, D. R. (2002). Non-exceedance information for flood hazard assessment. In House, P. K., Webb, R. H., Baker, V. R., and Levish, D. R., editors, *Ancient Floods, Modern Hazards*, volume 5 of *Water Science and Application Series*, pages 91–109. American Geophysical Union.

Levish, D. R., England, J. F., Klawon, J. E., and O'Connell, D. R. H. (2003). Flood hazard analysis for Seminoe and Glendo dams, Kendrick and North Platte projects, Wyoming, final report. Dam safety technical report, U.S. Department of Interior, Bureau of Reclamation, Denver, CO.

Lichty, R. W. and Liscum, F. (1978). A rainfall-runoff modeling procedure for improving estimates of T-year (annual) floods for small drainage basins. Water-Resources Investigations Report 78-7, U.S. Geological Survey, Reston, VA.

Lins, H. F. and Cohn, T. A. (2011). Stationarity: Wanted dead or alive? *JAWRA Journal of the American Water Resources Association*, 47(3):475–480.

Maddox, R. A., Canova, F., and Hoxit, L. R. (1980). Meteorological characteristics of flash flood events over the Western United States. *Monthly Weather Review*, 108(11):1866–1877.

Madsen, H., Rasmussen, P. F., and Rosbjerg, D. (1997). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling. *Water Resources Research*, 33(4):747–757.

Mastin, M. C. (2007). Re-evaluation of the 1921 peak discharge at Skagit River near Concrete, Washington. Scientific Investigations Report 2007-5159, U.S. Geological Survey, Reston, VA.

Matalas, N. C. and Jacobs, B. (1964). A Correlation Procedure for Augmenting Hydrologic Data. Professional Paper 434-E, U.S. Geological Survey.

Matthai, H. F. (1969). *Floods of June 1965 in South Platte River Basin, Colorado*. U.S. Geological Survey Water-Supply Paper 1850-B, Washington, DC.

May, J., Gorman, J. G., Goodrich, R. D., Bobier, M. W., and Miller, V. E. (1996). Water resources data, Iowa, Water Year 1995. Water-Data Report WDR-IA-95-1, U.S. Geological Survey, Reston, VA.

McCabe, G. J. and Wolock, D. M. (2002). A step increase in streamflow in the conterminous United States. *Geophysical Research Letters*, 29(24):38–1–38–4.

McCord, V. A. S. (1990). *Augmenting flood frequency estimates using flood-scarred trees*. Ph.D. dissertation, University of Arizona.

McCuen, R. and Knight, Z. (2006). Fuzzy analysis of slope-area discharge estimates. *Journal of Irrigation and Drainage Engineering*, 132(1):64–69.

McCuen, R. H. (2003). *Modeling Hydrologic Change: Statistical Methods*. Lewis Publishers, Boca Raton.

McCuen, R. H. (2004). *Hydrologic Analysis and Design*. Prentice Hall, Upper Saddle River, NJ, 3 edition.

McGlashan, H. D. and Briggs, R. C. (1939). Floods of December 1937 in Northern California. Water-Supply Paper 843, U.S. Geological Survey, Reston, VA.

Merz, R. and Blöschl, G. (2003). A process typology of regional floods. *Water Resources Research*, 39(12):SWC 5–1 – SWC 5–20.

Merz, R. and Blöschl, G. (2008a). Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information. *Water Resources Research*, 44(8).

Merz, R. and Blöschl, G. (2008b). Flood frequency hydrology: 2. Combining data evidence. *Water Resources Research*, 44(8).

MGS Engineering Consultants (2009). General Storm Stochastic Event Flood Model (SEFM) - Technical Support Manual. Technical report, Bureau of Reclamation, Denver, CO.

Moglen, G. E. (2009). Hydrology and impervious areas. *Journal of Hydrologic Engineering*, 14(4):303–304.

Moglen, G. E. and Beighley, R. E. (2002). Spatially explicit hydrologic modeling of land use change. *JAWRA Journal of the American Water Resources Association*, 38(1):241–252.

Moglen, G. E. and Shivers, D. E. (2006). Methods for adjusting U.S. Geological Survey rural regression peak discharges in an urban setting. Scientific Investigations Report 2006-5270, U.S. Geological Survey, Reston, VA.

Moss, M. E. and Thomas, W. O. (1982). Discussion of "WRC standard flood frequency guidelines" by D.R. Jackson. *J. Water Resources Planning and Management*, 108(1):166–168.

Munn, J. and Savage, J. L. (1922). The Flood of June 1921 in the Arkansas River, at Pueblo, Colorado. *Transactions, ASCE*, 85(1480):1–65.

Murphy, P. (2001). Evaluation of mixed-population flood-frequency analysis. *Journal of Hydrologic Engineering*, 6(1):62–70.

National Research Council (1988). Estimating probabilities of extreme floods. Technical report, National Research Council.

National Research Council (1995). Flood risk managment and the American River basin: an evaluation. Technical report, National Academy Press.

National Research Council (1999). Improving American River flood frequency analyses. Technical report, National Research Council.

National Research Council (2000). Risk analysis and uncertainty in flood damage reduction studies. Technical report, National Academy Press.

O'Connor, J. E. and Costa, J. E. (2004). Spatial distribution of the largest rainfall-runoff floods from basins between 2.6 and 26,000 km2 in the united states and puerto rico. *Water Resources Research*, 40(1):n/a–n/a. W01107.

O'Connor, J. E., Webb, R. H., and Baker, V. R. (1986). Paleohydrology of pool-and-riffle pattern development: Boulder Creek, Utah. *Geological Society of America Bulletin*, 97(4):410–420.

Olsen, J. R., Kiang, J., and Waskom, R. (2010). Workshop on Nonstationarity, Hydrologic Frequency Analysis, and Water Management. Colorado Water Institute Information Series 109, Colorado Water Institute, Colorado State University, Fort Collins, CO.

Olson, S. A. (2014). Estimation of flood discharges at selected annual exceedance probabilities for unregulated, rural streams in Vermont, with a section on Vermont regional skew regression, by Veilleux, A.G. Scientific Investigations Report 2014-5078, U.S. Geological Survey, Reston, VA.

Osterkamp, W. R. and Costa, J. E. (1987). Changes accompanying an extraordinary flood on a sand-bed stream. In Mayer, L. and Nash, D., editors, *Catastrophic Flooding*, chapter 10, pages 201–224. Allen and Unwin, Boston.

Over, T. M. and Soong, D. T. (2015). Flood-Peak Magnitudes at Selected Stations in Northeastern Illinois Adjusted for Changes in Land-Use Conditions. Administrative Report 2012-XXXX, U.S. Geological Survey, Reston, VA.

Paretti, N. V., Kennedy, J. R., and Cohn, T. A. (2014a). Evaluation of the expected moments algorithm and a multiple low-outlier test for flood frequency analysis at streamgaging stations in Arizona. Scientific Investigations Report 2014-5026, U.S. Geological Survey, Reston, VA.

Paretti, N. V., Kennedy, J. R., Turney, L. A., and Veilleux, A. G. (2014b). Methods for estimating magnitude and frequency of floods in Arizona, developed with unregulated and rural peak-flow data through water year 2010. Scientific Investigations Report 2014-5211, U.S. Geological Survey, Reston, VA.

Parrett, C., Veilleux, A., Stedinger, J. R., Barth, N. A., Knifong, D. L., and Ferris, J. C. (2011). Regional skew for California, and flood frequency for selected sites in the Sacramento–San Joaquin River Basin, based on data through Water Year 2006. Scientific Investigations Report 2010-5260, U.S. Geological Survey, Reston, VA.

Pilgrim, D. H. and Cordery, I. (1993). Flood runoff. In Maidment, D. R., editor, *Handbook of Hydrology*, chapter 9, pages 9.1–9.72. McGraw Hill, Inc.

Potter, K. W. and Walker, J. F. (1985). An empirical study of flood measurement error. *Water Resources Research*, 21(3):403–406.

Prior, J. C. (1991). *Landforms of Iowa*. University of Iowa Press.

Quessy, J.-F., Favre, A.-C., Said, M., and Champagne, M. (2011). Statistical inference in Lombard's smooth-change model. *Environmetrics*, 22(7):882–893.

Quick, M. C. (1991). Reliability of flood discharge estimates. *Canadian Journal of Civil Engineering*, 18:624–630.

Rantz, S. E. and Others (1982). *Measurement and computation of streamflow: Volume 1, Measurement of stage and discharge*. U.S. Geological Survey Water-Supply Paper 2175, Reston, VA.

Redmond, K. T., Enzel, Y., House, P. K., and Biondi, F. (2002). Climate variability and flood frequency at decadal to millennial time scales. In House, P. K., Webb, R. H., Baker, V. R., and Levish, D. R., editors, *Ancient Floods, Modern Hazards*, volume 5 of *Water Science and Application Series*, pages 21–45. American Geophysical Union.

Reed, D. (1999). Flood Estimation Handbook, Volume 1: Overview. Technical report, Institute of Hydrology, Wallingford, Oxfordshire, UK.

Reis, D., Stedinger, J. R., and Martins, E. S. (2005). Bayesian GLS regression with application to LP3 regional skew estimation. *Water Resources Research*, 41:W10419.

Ries, K. G., Guthrie, J. G., Rea, A. H., Steeves, P. A., and Stewart, D. W. (2008). StreamStats: A Water Resources Web Application. Fact Sheet 2008-3067, U.S. Geological Survey.

Rogger, M., Kohl, B., Pirkl, H., Viglione, A., Komma, J., Kirnbauer, R., Merz, R., and Blöschl, G. (2012). Runoff models and flood frequency statistics for design flood estimation in Austria – Do they tell a consistent story? *Journal of Hydrology*, 456–457:30–43.

Rosner, B. (1983). Percentage points for a generalized ESD many outlier procedure. *Technometrics*, 25(2):165–172.

Russell, S. O. (1982). Flood probability estimation. *Journal of Hydraulic Engineering*, 108(HY1):63–73.

Salas, J. and Obeysekera, J. (2014). Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. *Journal of Hydrologic Engineering*, 19(3):554–568.

Salas, J. D. (1993). Analysis and Modeling of Hydrologic Time Series. In Maidment, D. R., editor, *Handbook of Hydrology*, chapter 19, pages 19.1–19.72. McGraw Hill, Inc.

Salas, J. D. and Boes, D. C. (1980). Shifting level modelling of hydrologic series. *Advances in Water Resources*, 3(2):59 – 63.

Salas, J. D., Wohl, E. E., and Jarrett, R. D. (1994). Determination of flood characteristics using systematic, historical and paleoflood data. In Rossi, G., Harmancioglu, N., and Yevjevich, V., editors, *Coping With Floods*, pages 111–134. Kluwer Academic Publishers, Netherlands.

Sanders, C. L., Kubik, H. E., Hoke, J. T., and Kirby, W. (1990). Flood frequency of the Savannah River at Augusta, Georgia. Water-Resources Investigations Report 90-4024, U.S. Geological Survey, Reston, VA.

Sando, S. K., Driscoll, D. G., and Parrett, C. (2008). Peak-flow frequency estimates based on data through water year 2001 for selected streamflow-gaging stations in South Dakota. Scientific Investigations Report 2008-5104, U.S. Geological Survey, Reston, VA.

Sauer, V. B. and Meyer, R. W. (1992). Determination of error in individual discharge measurements. Open-File Report 92-144, U.S. Geological Survey.

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

Sauer, V. B. and Turnipseed, D. P. (2010). Stage measurement at gaging stations. Techniques and Methods Book 3, Chapter A7, U.S. Geological Survey.

Sigafoos, R. S. (1964). Botanical evidence of floods and flood-plain deposition. Professional Paper 485-A, U.S. Geological Survey.

Singh, V. P. (1995). *Computer Models of Watershed Hydrology*. Water Resources Publications, Highlands Ranch, CO.

Smith, J. A., Villarini, G., and Baeck, M. L. (2011). Mixture distributions and the hydroclimatology of extreme rainfall and flooding in the Eastern United States. *Journal of Hydrometeorology*, 12:294–309.

Southard, R. E. and Veilleux, A. G. (2014). Methods for estimating annual exceedance-probability discharges and largest recorded floods for unregulated streams in rural Missouri. Scientific Investigations Report 2014-5165, U.S. Geological Survey, Reston, VA.

Stedinger, J. R. and Cohn, T. A. (1986). Flood frequency analysis with historical and paleoflood information. *Water Resources Research*, 22(5):785–793.

Stedinger, J. R. and Cohn, T. A. (1987). Historical flood-frequency data: Its value and use. In Singh, V. P., editor, *Regional Flood Frequency Analysis*, pages 273–286. D. Reidel.

Stedinger, J. R. and Griffis, V. W. (2008). Flood frequency analysis in the United States: Time to update. *Journal of Hydrologic Engineering*, 13(4):199–204.

Stedinger, J. R. and Griffis, V. W. (2011). Getting from here to where? Flood frequency analysis and climate. *JAWRA Journal of the American Water Resources Association*, 47(3):506–513.

Stedinger, J. R., Surani, R., and Therivel, R. (1988). Max user's guide: A program for flood frequency analysis using systematic-record, historical, botanical, physical paleohydrologic and regional hydrologic information using maximum likelihood techniques. Technical report, Cornell University.

Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E. (1993). Frequency analysis of extreme events. In Maidment, D. R., editor, *Handbook of Hydrology*, chapter 18, pages 18.1–18.66. McGraw Hill, Inc.

Stewart, J. E. and Bodhaine, G. L. (1961). Floods in the Skagit River basin, Washington. Water-Supply Paper 1527, U.S. Geological Survey, Reston, VA.

Sutley, D. E., Klinger, R. E., Bauer, T. R., and Godaire, J. E. (2008). Trinity Dam detailed hydrologic hazard analysis using the Stochastic Event Flood Model. Dam safety technical report, U.S. Department of Interior, Bureau of Reclamation, Denver, CO.

Suttie, R. H. (1928). Report on the Water Resources of Connecticut. Bulletin No. 44, Connecticut Geol. and Nat. Hist. Survey, Hartford, CT.

Sveinsson, O. G. B., Salas, J. D., Boes, D. C., and Sr, R. A. P. (2003). Modeling the dynamics of long-term variability of hydroclimatic processes. *Journal of Hydrometeorology*, 4:489–505.

Swain, R. E., England Jr, J. F., Bullard, K. L., and Raff, D. (2006). Guidelines for Evaluating Hydrologic Hazards. Technical report, U.S. Bureau of Reclamation, Denver, CO.

Tasker, G. D. (1978). Flood frequency analysis with a generalized skew coefficient. *Water Resources Research*, 14(2):373–376.

Tasker, G. D. (1983). Effective record length for the T-year event. *Journal of Hydrology*, 64:39–47.

Tasker, G. D. and Stedinger, J. R. (1986). Regional skew with weighted LS regression. *Journal of Water Resources Planning and Management*, 112(2):225–237.

Tasker, G. D. and Stedinger, J. R. (1989). An operational GLS model for hydrologic regression. *Journal of Hydrology*, 111(4):361–375.

Thomas, W. O. (1982). An evaluation of flood frequency estimates based on rainfall/runoff modeling. *JAWRA Journal of the American Water Resources Association*, 18(2):221–229.

Thomas, W. O. (1985). A uniform technique for flood frequency analysis. *Journal of Water Resources Planning and Management*, 111(3):321–337.

Thomas, W. O. (1987). Comparison of flood-frequency estimates based on observed and model-generated peak flows. In Singh, V. P., editor, *Hydrologic Frequency Modeling*, pages 149–161, Louisiana State University, Baton Rouge. D. Reidel Publishing Company.

Thomas, W. O., Grimm, M. M., and McCuen, R. H. (2001). Evaluation of Flood Frequency Estimates for Ungaged Watersheds. Technical report, Hydrologic Frequency Work Group, Subcommittee on Hydrology.

Thomson, M. T., Gannon, W. B., Thomas, M. P., and Hayes, G. S. (1964). Historical floods in New England. Water-Supply Paper 1779-M, U.S. Geological Survey, Reston, VA.

Turnipseed, D. P. and Sauer, V. B. (2010). Discharge measurement at gaging stations. Techniques and Methods Book 3, Chapter A8, U.S. Geological Survey.

U.S. Army Corps of Engineers (1973). Storm Rainfall in the United States, 1945 - 1973. Technical report, U.S. Army Corps of Engineers, Washington, DC.

U.S. Army Corps of Engineers (1982). Mixed-Population Frequency Analysis. Technical Report TD-17, U.S. Army Corps of Engineers, Hydrologic Engineering Center.

U.S. Army Corps of Engineers (1993). Hydrologic Frequency Analysis. Technical Report EM 1110-2-1415, U.S. Army Corps of Engineers.

U.S. Army Corps of Engineers (2008). Orestimba Creek Hydrology: A Reevalutaion Based on

Updated Peak and Volume Frequency Curves. Technical report, U.S. Army Corps of Engineers, Water Management Section, Sacramento District.

U.S. Bureau of Reclamation and Utah State University (1999). A Framework for Characterizing Extreme Floods for Dam Safety Risk Assessment. Technical report, U.S. Bureau of Reclamation.

USWRC (1967). *A uniform technique for determining flood flow frequencies, Bulletin No. 15*. U.S. Water Resources Council, Subcommittee on Hydrology, Washington, D.C.

USWRC (1976). *Guidelines for Determining Flood Flow Frequency, Bulletin No. 17*. U.S. Water Resources Council, Subcommittee on Hydrology, Washington, D.C.

Veilleux, A. (2011). *Bayesian GLS Regression For Regionalization Of Hydrologic Statistics, Floods And Bulletin 17 Skew*. Ph.D. dissertation, Cornell University.

Veilleux, A. G., Stedinger, J. R., and Lamontagne, J. R. (2011). Bayesian WLS/GLS Regression for Regional Skewness Analysis for Regions with Large Cross-Correlations among Flood Flows. In *World Environmental and Water Resources Congress 2011 Bearing Knowledge for Sustainability*. ASCE.

Viglione, A., Merz, R., Salinas, J. L., and Blöschl, G. (2013). Flood frequency hydrology: 3. A Bayesian analysis. *Water Resources Research*, 49(2):675–692.

Villarini, G., Serinaldi, F., Smith, J. A., and Krajewski, W. F. (2009). On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resources Research*, 45(8).

Vogel, R. M. and Kroll, C. N. (1991). The value of streamflow record augmentation procedures in low-flow and flood-flow frequency analysis. *Journal of Hydrology*, 125(34):259 – 276.

Vogel, R. M., Matalas, N. C., England, J. F., and Castellarin, A. (2007). An assessment of exceedance probabilities of envelope curves. *Water Resources Research*, 43(7):n/a–n/a. W07403.

Vogel, R. M. and Stedinger, J. R. (1984). Flood-Plain Delineation in Ice Jam Prone Regions. *Journal of Water Resources Planning and Management*, 110(2):206–219.

Vogel, R. M. and Stedinger, J. R. (1985). Minimum variance streamflow record augmentation procedures. *Water Resources Research*, 21(5):715–723.

Wahl, K. L., Thomas, W. O., and Hirsch, R. M. (1995). *The stream-gaging program of the U.S. Geological Survey*. U.S. Geological Survey Circular 1123, Reston, VA.

Weaver, J. C., Feaster, T. D., and Gotvald, A. J. (2009). Magnitude and Frequency of Rural Floods in the Southeastern United States, 2006: Volume 2, North Carolina. Scientific Investigations Report 2009-5158, U.S. Geological Survey, Reston, VA.

Webb, R. H. and Betancourt, J. L. (1992). *Climatic variability and flood frequency of the Santa Cruz River, Pima County, Arizona*. U.S. Geological Survey Water-Supply Paper 2379, Reston, VA.

Webb, R. H. and Jarrett, R. D. (2002). One-dimensional estimation techniques for discharges of paleofloods and historical floods. In House, P. K., Webb, R. H., Baker, V. R., and Levish, D. R., editors, *Ancient Floods, Modern Hazards*, volume 5 of *Water Science and Application Series*, pages 111–125. American Geophysical Union.

Wiley, J. B. and Atkins, J. T. J. (2010). Estimation of Flood-Frequency Discharges for Rural, Unregulated Streams in West Virginia. Scientific Investigations Report 2010-5033, U.S. Geological Survey, Reston, VA.

Wohl, E. E. and Enzel, Y. (1995). Data for palaeohydrology. In Gregory, K. J., Starkel, L., and Baker, V., editors, *Global Continental Palaeohydrology*, pages 24–59. John Wiley & Sons, Ltd.

Yanosky, T. M. (1983). Evidence of floods on the potomac river from anatomical abnormalities in the wood of flood-plain trees. Professional Paper 1296, U.S. Geological Survey.

Yanosky, T. M. and Jarrett, R. D. (2002). Dendrochronologic evidence for the frequency and magnitude of paleofloods. In House, P. K., Webb, R. H., Baker, V. R., and Levish, D. R., editors, *Ancient Floods, Modern Hazards*, volume 5 of *Water Science and Application Series*, pages 77–89. American Geophysical Union.

Yen, B. C. (1970). Risks in hydrologic design of engineering projects. *Journal of the Hydraulics Division, ASCE*, 96(HY4):959–966.

APPENDIXES

1

Appendix 1—Subcommittee and Work Group Members

Subcommittee and Work Group Members

The Subcommittee on Hydrology (SOH) is a sub-group under the Advisory Committee on Water Information (ACWI). The purpose of the SOH is to improve the availability and reliability of surface-water quantity information needed for hazard mitigation, water supply and demand management, and environmental protection. The SOH coordinates and oversees technical working groups, including the Hydrologic Frequency Analysis Work Group (HFAWG). The SOH sponsored this HFAWG work effort to prepare the update to these *Guidelines*. Current SOH membership is listed in Table 1.1. Further details about SOH and its activities are available at http://acwi.gov/hydrology/index.html.

The overall goal of the Hydrologic Frequency Analysis Work Group is to recommend procedures to increase the usefulness of the current guidelines for Hydrologic Frequency Analysis computations (e.g. Bulletin 17B) and to evaluate other procedures for frequency analysis of hydrologic phenomena. The work group forwards draft papers and recommendations to the Subcommittee on Hydrology of ACWI for appropriate action. As part of these activities, the HFAWG oversaw the revision to these *Guidelines*. Current HFAWG membership is listed in Table 1.2. Further details about HFAWG and its activities are available at http://acwi.gov/hydrology/Frequency/index.html.

Table 1.1. Subcommittee on Hydrology Members.

Member Organization	Representative
Association of State Floodplain Managers	Wilbert O. Thomas Jr.
BECKER	Martin Becker
DOI/Bureau of Land Management	Robert Boyd
DOI/Bureau of Reclamation	Dr. Ian Ferguson
DOI/Office of Surface Mining	TBD
DOI/US Geological Survey	Robert Mason *(Vice Chair)*
Federal Energy Regulatory Commission	Dr. S. Samuel Lin
Federal Highway Administration	Brian Beucler
Global Ecosystems Center	Don Woodward
NASA/Goddard Space Flight Center	David Toll
National Hydrologic Warning Council	Ben Pratt
National Science Foundation	Dr. Thomas Torgersen
NOAA/National Weather Service	Victor Hom *(Chair)*
US Army Corps of Engineers	Dr. Chandra Pathak
USDA/Agricultural Research Service (ARS)	Dr. David C. Goodrich
USDA/Natural Resources Conservation Service (NRCS)	Claudia Hoeft
USDA/U.S. Forest Service	Michael Eberle
U.S. Environmental Protection Agency	David Wells
USDHS/Federal Emergency Management Agency (FEMA)	Dr. Siamak Esfandiary
U.S. Nuclear Regulatory Commission (NRC)	Thomas J. Nicholson

Table 1.2. Hydrologic Frequency Analysis Work Group Members.

Member Name	Organization	Location
Wilbert O. Thomas Jr. *(Chair)*	Michael Baker International	Manassas, VA
Dr. Siamak Esfandiary	Federal Emergency Management Agency	Crystal City, VA
Don Woodward	Global Ecosystems Center	Derwood, MD
Martin Becker	BECKER	Atlanta, GA
Dr. Timothy Cohn	U.S. Geological Survey	Reston, VA
Dr. Beth Faber	U.S. Army Corps of Engineers	Davis, CA
Dr. John England	U.S. Army Corps of Engineers	Lakewood, CO
Prof. Jery Stedinger	Cornell University	Ithaca, NY
Dr. Zhida Song-James	Consulting Hydrologist	Fairfax, VA
Dr. Jerry Coffey	Mathematical Statistician	Middletown, VA
Joe Krolak	Federal Highway Administration	Washington, DC
William Merkel	Natural Resources Conservation Service	Beltsville, MD
Dr. Sanja Perica	National Weather Service	Silver Spring, MD
Thomas Nicholson	Nuclear Regulatory Commission	Rockville, MD
Dr. S. Samuel Lin	Federal Energy Regulatory Commission	Washington, DC
Mike Eiffe (through Sept. 2014)	Tennessee Valley Authority	Knoxville, TN
Curt Jawdy	Tennessee Valley Authority	Knoxville, TN

Appendix 2—Data Sources

Data Sources

This appendix provides some representative data sources for flood frequency. Systematic records, historical data, paleoflood data and botanical information, regional information, and precipitation and climate information are briefly described. These sources are intended to be used as references and starting points for data collection, and are not all-inclusive. Available sources and websites for these data can be found through Internet searches.

Systematic Records

Systematic records that may be useful for estimating flood frequency include: peak flows, daily flows, reservoir inflows and elevations, hydrograph data, and streamflow measurements. Annual maximum instantaneous peak streamflow and gage height data can be obtained from the USGS National Water Information System (NWIS). Daily streamflow data can be obtained from various sources. The main data source is the USGS NWIS. These data can be easily retrieved with software packages such as (Hirsch and DeCicco, 2015).

Other Federal agencies provide daily streamflow and extensive reservoir-related data, including elevations, inflows and outflows. These data can be of direct use for extending discontinued streamflow gages and estimating unregulated flows.

The Bureau of Reclamation http://www.usbr.gov/ provides data through its five regions in the 17 western states for numerous river locations and over 350 reservoirs. The Reclamation Hydromet data bases provide data for the Great Plains Region and Pacific Northwest Region. Data within the Upper Colorado Region is obtained through reservoir operations at http://www.usbr.gov/uc/. Data within the Lower Colorado Region is obtained through river operations at http://www.usbr.gov/lc/. Reclamation's Mid-Pacific region provides data for many locations, including the Central Valley, through the California Data Exchange Center.

The U.S. Army Corps of Engineers provides streamflow and reservoir information, within the conterminous United States, through seven divisions. A map with links to each division is at http://www.usace.army.mil. Streamflow and reservoir data can be provided for specific projects or river basins, within each division. For example, the Northwestern Division provides data for the Missouri River basin through their reservoir control center.

Individual state agencies provide streamflow information, typically through their Division of Water Resources or Division of Natural Resources. Some examples of streamflow data bases by states are: California, Colorado, Oregon, and Minnesota. Local flood-control districts and organizations may also have relevant streamflow data.

Instantaneous data (15-minute data, unit values, complete hydrographs), from 2007 to present for active streamgages, can be obtained from the USGS NWIS. Hydrograph data from about the mid-1980s to 2007 can be obtained from the instantaneous data archive.

Data on manual measurements of streamflow and gage height, including indirect measurement, can be obtained from the USGS. These measurements are used to supplement and (or) verify the accuracy of the automatically recorded observations, as well as to compute streamflow based on gage height. They are valuable for flood frequency studies to aid hydrologists in understanding how the largest flood estimates are made (such as an indirect), and in estimating uncertainty.

Historical Data

Historical flood data sources can be obtained from a variety of locations. This section describes some of those data sources useful for flood frequency, and is an excerpt from England (1998, Chapter 4), updated with additional recent studies.

One typically obtains U.S. Geological Survey records as a first step in the search for historical information. Information on observed floods, occurring after about 1900, that typically cause flooding of populated areas, damage and sometimes deaths are described in various USGS publications, such as Water-Supply Papers, Professional Papers and Scientific Investigations Reports. The information generally consists of basin rainfall estimates, types of discharge or indirect measurements made, damage estimates, pictures of damaged structures, and erosion and deposition in channels and floodplains. In some cases, past historical flood dates, stages, and peak discharge estimates in the region are described in each report. For example, the report on the Arkansas River flood of June 3-5, 1921 (Follansbee and Jones, 1922) lists previous floods back to about 1844, based primarily on Denver and Rio Grande railroad records. Stewart and Bodhaine (1961) describe recent floods and present a historical flood chronology back to 1815 for the Skagit River basin in Washington. In some cases, historic flood estimates are revised, such as the Skagit River near Concrete (Mastin, 2007). Many other Water-Supply Papers present historical flood information when documenting large regional floods, although in many cases the river stages and discharges are unknown (McGlashan and Briggs, 1939). In some cases, electronic databases of historical flood estimates are available, such as in Colorado by Kohn et al. (2013).

The U.S. Geological Survey Water Resources Data Reports, that have been published for each state (1962-2005), contain some limited historical flood descriptions and information that can be extremely valuable for frequency analysis. The information is provided on the site information sheet for individual gaging stations. Since 2006, this same information can be obtained for each individual gage, if the gaging station is currently in operation. Three types of data are typically presented in the reports and site information summaries: (1) dates, stages and sometimes discharges of observed floods prior to the gaging station period of record, e.g. Durlin and Schaffstall (2002, p. 210); (2) a large flood during the period of record that is known to be the *maximum stage and discharge since at least* some historic date, e.g. Crowfoot et al. (1997, p. 413); and (3) a large flood during the period of record that is known to be the *maximum stage and discharge since* some historic date, e.g. May et al. (1996, p. 193). The information provided in (2) and (3) sometimes only refers to either stage or discharge, depending on the observation or estimate made. In addition, there is a very subtle difference between the information provided in (2) and (3). Data provided as (2) indicate one *does not* have information on any flood discharges or stages prior to the date stated. One does have knowledge of a flood in the historical year stated in (3). The information for cases (1) and (3) is typically stored in electronic format in the U.S. Geological Survey NWIS data base. The data are generally summarized in two columns: discharge codes, where a "7" indicates that the discharge is a historic peak, and a "highest since" column, where the historic year is listed. These data need to be evaluated on an individual basis to properly estimate n_h and T_h.

State reports and publications are another major source of historical flood information. These publications can contain information on record floods, stages, historical periods, and inpacts to infrastructure. For example, Suttie (1928) states "there are three great storms affecting Connecticut that are worthy of particular mention: 1869, 1897, and 1927"; this information suggests that rainfall amounts and flood discharges are less than values estimated in the intervening time between these three events. For the 1869 flood, Suttie (1928, p. 120) states "the Connecticut River gage at Hartford registered 26.3 feet. This is the highest stage in over a century caused by rain alone"; this information can be utilized to estimate the historical period h one may use for the 1869 stage. Many other state reports contain relevant examples such as this one.

Journals and other Federal Agency reports are invaluable sources for historical flood information. The primary historical journal references are the Journal of the Boston Society of Civil Engineers and Transactions of the American Society of Civil Engineers. For example, Kuichling (1917, pp. 650-663) provides a table of maximum unit discharges for large floods in the United States to at least 1786; he also includes a reference list that includes many journals, Geological Survey and State reports. The U.S. Army Corps of Engineers retains flood files at District offices. Community flood information and experiences are usually included in Federal Emergency Management Agency (FEMA) Flood Insurance Studies. A detailed example of historical flood data collection is provided in Thomson et al. (1964); they present a flood chronology in New England from 1620 through 1955.

Paleoflood Data and Botanical Information

Paleoflood data and botanical information for river basins and specific locations can be obtained from existing, previously-published sources and institutions that have obtained the data, or by field data collection at the site of interest. The main sources of existing, previously-published paleoflood and botanical data are various institutions that have collected the data, such as Federal agencies, state agencies, and academic institutions. These data are routinely documented in journal articles, technical reports and data bases, books, and some electronic data bases. Over the past 20 years, the University of Arizona, Bureau of Reclamation, U.S. Geological Survey, and other agencies and institutions have embarked on numerous field campaigns to obtain paleoflood data relevant for flood frequency. Similar to historical information, paleoflood and botanical data are obtained by searching for relevant documents and contacting institutions that have interests and facilities within the watershed of interest.

Several journal articles and books are key references in obtaining previously-published paleoflood and botanical information, and are indispensible guides for data collection efforts. Wohl and Enzel (1995) provide a useful introduction and overview of available paleoflood data. Baker et al. (1988) and House et al. (2002b) are key references that describe data for numerous case studies and locations, present methods for paleoflood data collection, and contains numerous citations to other relevant works and data. Baker (2008, Table 2) summarizes many paleoflood studies and data collection completed in the United States. Benito and O'Connor (2013) and Baker (2013) provide current summaries on paleoflood and paleohydrology data and methods. Paleoflood data are readily used with *EMA*; for example England et al. (2003a) summarized paleoflood data and demonstrated its use in flood frequency with *EMA* for a number of sites in the United States.

Paleoflood data for many locations within the western United States have been collected by the Bureau of Reclamation for dam safety analyses. These data are typically available for many rivers and locations adjacent to Reclamation dams and other Department of Interior facilities, in order to document the most extreme floods and non-exceedance information in the Holocene. Reclamation staff typically collect paleoflood data at one of three levels: reconnaissance, intermediate, or detailed. As the level of study increases, more stratigraphic and soil-age data are obtained, and hydraulic models used to estimate discharge increase in complexity. These data are available in numerous Reclamation reports for specific projects and/or watersheds, and in databases (Klinger and Godaire, 2002). Some representative studies include: the American River and adjacent basins near Sacramento, California (Bureau of Reclamation, 2002); the North Platte River near Rawlins and Glendo, Wyoming (Levish et al., 2003); the Arkansas River near Pueblo, Colorado (England et al., 2006); the South Fork Boise River, Idaho (Klinger and Bauer, 2010); the North Fork Red River near Altus, Oklahoma (Godaire and Bauer, 2012); the San Joaquin River near Fresno, California (Godaire et al., 2012), and the Rio Chama near El Vado Dam, New Mexico (Godaire and Bauer, 2013). Peak-flow frequency estimates have been made at these sites using *EMA*. The U.S. Geological Survey has also conducted numerous paleoflood studies using reconnaissance or regional approaches (Jarrett and Tomlinson, 2000) and detailed methods for flood hazard assessments at specific locations (Harden et al., 2011). Some paleoflood data are available in electronic databases, such as Kohn et al. (2013).

Botanical information and data, such as tree scars and tree rings, are available in publications and some electronic data bases. Some essential publications on methods and data are Hupp (1987, 1988), and Yanosky and Jarrett (2002); these contain numerous citations to other relevant works and data. McCord (1990) provides tree-scar data at select sites in Arizona, Utah, New Mexico, and Colorado. Additional resources include the Laboratory of Tree-Ring Research at the University of Arizona and the International Tree-Ring Data Bank.

Regional Information

Regional information that can be considered for flood frequency typically consists of regional estimates of flow statistics. Regional skew coefficient G estimates and mean-square error MSE_G estimates can be obtained for some locations in current U.S. Geological Survey flood-frequency reports for regions or individual states. For example, regional skew estimates are available for the Southeastern US (Gotvald et al., 2009; Feaster et al., 2009; Weaver et al., 2009), California (Parrett et al., 2011; Gotvald et al., 2012), Iowa (Eash et al., 2013), Arizona (Paretti et al., 2014b), and Missouri (Southard and Veilleux, 2014). Regional flood quantile estimates Q_i and their variances $V_{regi,i}$ are also available in these reports, and are useful in record extension (Appendix 7) and in weighting of independent estimates (Appendix 8). These flood frequency reports and additional information on regional skew and regional quantile estimates for many locations are available from the U.S. Geological Survey and the HFAWG at http://acwi.gov/hydrology/Frequency/.

In lieu of current published estimates, it is recommended that users consult with the U.S. Geological Survey to determine the availability of regional skew estimates that have been prepared using current methods, described in the Section Estimating Regional Skew. The regional skew estimates published in IACWD (1982, Plate 1) are not recommended for use in flood-frequency studies. When no other regional skew information is available, it is recommended that users consider a regional skew equal to zero, or develop new estimates for the region of interest.

Precipitation and Climate Information

Precipitation and climate information that is potentially useful for flood rainfall-runoff modeling and flood-frequency analysis is generally available from various Federal and state agencies. Point precipitation data and radar rainfall products are available from the National Oceanic and Atmospheric Administration (NOAA), National Climatic Data Center (NCDC). National Weather Service River Forecast Centers also provide multi-sensor precipitation (combined radar and precipitation gage) estimates across the United States. Precipitation frequency estimates and time series are available from the National Weather Service Hydrometeorological Design Studies Center. Precipitation data for many of the largest historical rainfall events and floods can be obtained from extreme storm catalogs at the U.S. Army Corps of Engineers, Bureau of Reclamation, and through the Extreme Storm Events Working Group at http://acwi.gov/hydrology/extreme-storm/index.html. Precipitation and temperature data important for rainfall-runoff modeling of extreme floods can be obtained from the National Resources Conservation Service (NRCS) Snow Telemetry (SNOTEL) and snow course data. The National Weather Service's National Operational Hydrologic Remote Sensing Center (NOHRSC) SNOw Data Assimilation System (SNODAS), available through the National Snow and Ice Data Center, is another valuable data set for snow cover and associated variables. Climate information that is useful for a hydroclimatological perspective on floods is available from the NOAA Earth System Research Laboratory (ESRL). Information on climate models, downscaling information, and climate change, that is potentially relevant for floods, is presented in Brekke et al. (2009). Downscaled climate information for climate change assessment studies is available at http://www.usbr.gov/.

Appendix 3—Initial Data Analysis

Initial Data Analysis

When conducting a flood frequency analysis, a first step is to undertake basic analysis of the peak flow time series to check for obvious errors and to check that the data conform to the assumptions used in the frequency analysis. One of the main assumptions in flood frequency analysis is that the data in the peak flow time series are independent and identically distributed. Some tests that check that these assumptions are reasonable for a particular time series include tests for autocorrelation and non-stationarity. Visual inspection of the time series can also reveal issues which need to be addressed.

Visual Inspection: Plot the data

Before any formal statistical tests are employed, a visual inspection of a plot of the peak flow time series can be used to help identify any potential errors with the data. For example, any peaks that are orders of magnitude different from the others should be verified. Visual inspection of the time series plot may also reveal obvious changes in the mean or standard deviation of the peak flow data over time. For example, construction of a dam and reservoir may drastically alter peak flow time series and the entire pre-and post-dam time series should not be used together for a peak flow frequency analysis.

Autocorrelation

It is recommended that an annual flood series be examined for autocorrelation through the use of a correlogram (Salas, 1993). In an autocorrelated time series, the value in one time step is correlated with the value in a previous (and future) time step. Autocorrelated time series can also be said to exhibit persistence. Hydrologic time series will often exhibit long term persistence. Note that this can affect trend testing, as discussed below.

Trends and shifts

The peak flow frequency analysis methods described in this document are only applicable when the peak flow data are believed to be part of the same underlying population. Changes in peak flow generation processes can lead to gradual trends or abrupt shifts in the peak flow time series. Statistical tests for trends and shifts can be useful for detecting such changes in the peak flow time series. Depending on the likely causes and the magnitude of any detected changes, different treatments may be needed before Bulletin 17C methods can be applied. These will be discussed in a future update to this document. A particularly difficult case is when it is unknown whether the apparent trend will continue, level off, or reverse in the future. Possible approaches for dealing with such changes have been discussed in the research literature, but a consensus on best practices has not yet emerged. Consequently, substantial judgment must be exercised when trends are found.

Changes may occur gradually or abruptly and different tests are commonly used to test for the presence of either type of change. A visual inspection of a plot of the annual peak flow time series should be the first step in assessing the time series. It is recommended that this be followed by trend tests to help assess whether changes over time may be important for the flood frequency analysis. This can be followed by a change point test for an abrupt change if desired. The specific tests described below have been used frequently, but other tests may also be considered.

Trend tests and change point analysis are most commonly done on the mean values of a time series, but tests for change in the variance of a time series can also be considered. Note that these tests can be sensitive to the start and end points used in the analysis. For example, if the period of record happens to either start or end with a large peak, there may be an apparent trend in the data. However, this apparent trend may simply be the result of the particular sample that was used. A slightly longer or shorter record would not show the same apparent trend. In other cases, the period of record may include only the drying or wetting phase of an oscillation with long periodicity. The apparent trend results from a finite record length, but in this case a much longer period of record is needed to fully understand the data.

Statistical tests

A common test for trends in a time series is the Mann-Kendall test. This test uses Kendall's τ as the test statistic to measure the strength of the monotonic relationship between annual peak streamflow and the year in which it occurred. The Mann-Kendall test is nonparametric and does not require that the data conform to any specific statistical distribution. The statistic is calculated using the ranks of the observed streamflow peaks and not the actual data values. Positive values for τ indicate that occurrences of annual peak streamflows are increasing with time for the period of record while negative values of τ indicate that annual peak streamflows are decreasing with time for the period of record.

As with other statistical tests, a p-value can be calculated for the test. Note that the p-values will be correct only when there is no serial correlation in the annual time series. This requirement can be problematic for hydrologic time series which can exhibit short term and long term persistence (Cohn and Lins, 2005).

In addition to the statistical significance of a trend, the actual magnitude of the trend should be considered. The Theil-Sen slope (Helsel and Hirsch, 1992) can be calculated in conjunction with Kendall's τ for this purpose. It is calculated as the median of all the slopes calculated by using all the possible pairs of peak flow values and years.

In some situations, there may be an abrupt shift (McCabe and Wolock, 2002) or change in the time series, rather than a gradual trend. For example, there may be distinct periods, exhibiting different flood characteristics, before and after installation of flood control structures. In other cases, the reason for the step change may not be as evident, but abrupt changes may still be found. Villarini et al. (2009), for example, found step changes that appeared to coincide with changes in the streamgage location. To refine the analysis, the test for a monotonic trend could be followed by a test for a step change. The Wilcoxon rank-sum test (also known as the Mann-Whitney test) or the Kolmogorov-Smirnov test are both nonparametric tests can be used to test for differences between two samples, when there is a suitable hypothesis for separating the time series into two or more sections. The potential step change should not be identified solely on the basis of visual inspection of the data, as this biases the test. The Pettitt test and Lombard's Smooth Change Model have both been suggested as alternative tests for abrupt changes which do not require an analyst to predetermine where a likely change occurs (Villarini et al., 2009; Quessy et al., 2011). More information on changepoint tests is available in accompanying material on the B17C website. Additional information on these tests can be found in Helsel and Hirsch (1992) and other statistical textbooks.

Example: Skokie River near Highland Park, IL

This example uses data from U.S. Geological Survey gaging station 05535070, Skokie River near Highland Park, Illinois. Figure 3.1 shows a time series plot of the Skokie River. It is a 54.6 square kilometer watershed that has become more and more urbanized over time. The urbanized fraction was about 0.60 at the beginning of the period of record in 1967, and increased to about 0.90 by the 2014 (Over and Soong, 2015). Visual inspection of the time series reveals an increasing trend over time.

The visual trend is confirmed with the Mann-Kendall test. The results from the test are as follows:

$\tau = 0.321$,

$p-$value = 0.00156, and

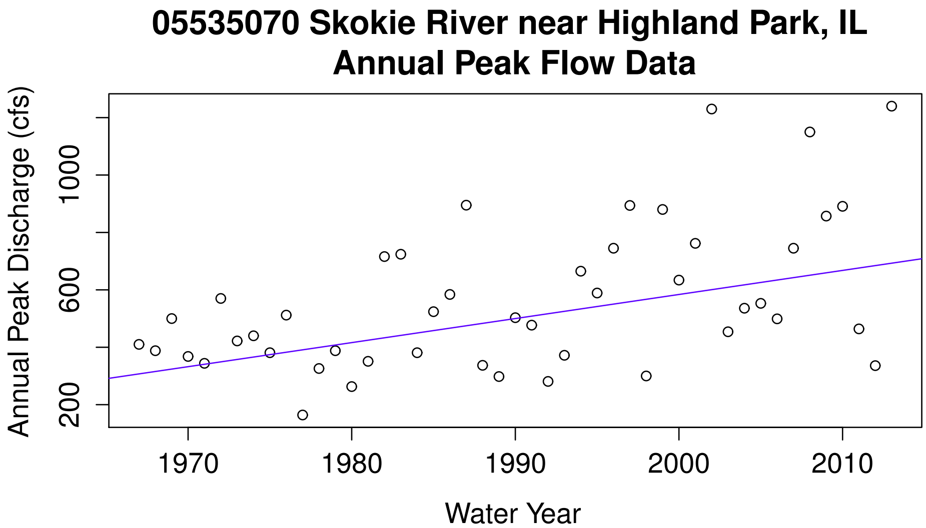

Figure 3.1. USGS 05535070 Skokie River near Highland Park, IL time series plot. Annual peak discharges have increased at this streamgage due to urbanization. The line is the fitted Theil-Sen line with slope 8.4 cfs/year.

Theil-Sen slope = 8.4 cfs/year.

The annual peak flows at this station are not significantly autocorrelated, as shown in Figure 3.2. This indicates that the estimated $p-$value is appropriate and is unaffected by autocorrelation.

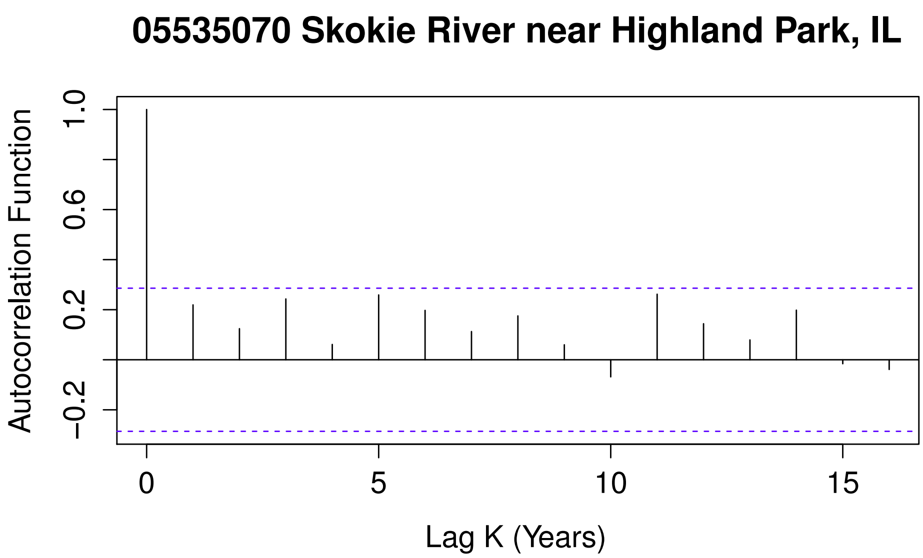

Figure 3.2. USGS 05535070 Skokie River near Highland Park, IL autocorrelation plot. The annual peaks do not exhibit any statistically significant autocorrelation for lag times between 1 and 15 years. The dashed lines are the thresholds for significant autocorrelation.

Appendix 4—Threshold-Exceedance Plotting Positions

Threshold-Exceedance Plotting Positions

This appendix provides an overview and equations for threshold-exceedance based plotting positions. Table 4.1 provides plotting position parameters. Some examples are shown in Appendix 9.

Consider a historical flood record with an n_h-year historical period in addition to a complete n_s-year gaged flood record. Assume that during the total $n = (n_s + n_h)$ years of record, a total of k floods exceeded a perception threshold for historical floods (Figure 3). If the k values which exceeded the threshold are indexed by $i = 1, \ldots,$ k, reasonable plotting positions approximating the exceedance probabilities with the interval $(0, p_e)$ are

$$p_i = p_e \left(\frac{1-a}{k+1-2a} \right) = \frac{k}{n} \left(\frac{i-a}{k+1-2a} \right) \tag{4.1}$$

where a is a value from Table 4.1 and $p_e = k/n$ is the probability of exceeding a threshold. For $k >> (1-2a)$, p_i is indistinguishable from $\dfrac{i-a}{n+1-2a}$ for a single threshold. Hirsch (1987) notes that for the first k floods, equation (4.1) is identical to the Hazen formula with $a = 0.5$, and is very close to the Gringorten formula with $a = 0.44$. Reasonable choices for a generally make little difference to the resulting plotting positions.

The plotting positions for systematic record floods below the threshold must be adjusted to reflect the additional information provided by the historical flood record, if the historical flood data and the systematic record are to be analyzed jointly in a consistent and statistically efficient manner (Hirsch and Stedinger, 1987). In this case, let e_s be the number of gaged-record floods that exceeded the threshold and hence are counted among the k exceedances of that threshold. Plotting positions within $(p_e, 1)$ for the remaining $(n_s - e_s)$ below-threshold gaged-record floods are

$$p_r = p_e + (1-p_e) \left(\frac{r-a}{n_s - e_s + 1 - 2a} \right) \tag{4.2}$$

for $r = 1$ through $n_s - e_s$, where again a is again a value from Table 4.1.

This approach directly generalizes to several thresholds. For the multiple exceedance threshold cases shown in Figure 12, equation (4.1) can be generalized (Hirsch and Stedinger, 1987; Stedinger et al., 1988, 1993). The number of thresholds is defined as j $(j = 1, \ldots, m)$, where the thresholds Q_j $(j = 1, \ldots, m)$ are ordered (sorted) from largest to smallest such that $Q_1 > Q_2 > \ldots Q_m$. The probability of exceedance p_{e_j} for each threshold j is defined as:

$$p_{e_j} = p_{e_{j-1}} + (1 - p_{e_{j-1}})\, q_{e_j} \tag{4.3}$$

where q_{ej} is the conditional probability that a flood falls between the j^{th} and $(j-1)^{th}$ threshold. It is defined by:

$$q_{e_j} = \frac{k_j}{n_j - \sum\limits_{l=1}^{j-1} k_l} \tag{4.4}$$

where k_j is the number of floods that exceed threshold j but not any higher thresholds $(j-1)$, and the denominator in equation (4.4) is the number of years (n_j) that threshold Q_j applies minus the sum of all floods k_l that exceed any higher $(j-1, j-2, \ldots)$ thresholds during period n_j. The above-threshold floods may be plotted by:

$$p_i = p_{e_{j-1}} + (1 - p_{e_{j-1}})\, q_{e_j} \left(\frac{i-a}{k_j + 1 - 2a} \right) \tag{4.5}$$

Table 4.1. Typical plotting position parameter a values.

Method	a
Weibull	0.00
Cunnane	0.40
Gringorten	0.44
Hazen	0.50

and the floods below all thresholds $(k_j + 1, \ldots, g)$ can be plotted using equation (4.2) with p_e equal to p_{e_j}.

Appendix 5—Potentially-Influential Low Floods (PILFs)

Potentially-Influential Low Floods (PILFs)

This appendix provides a general introduction to Potentially-Influential Low Floods (PILFs), and describes computation details for identifying PILFs – the Multiple Grubbs-Beck test (*MGBT*). Some details on *MGBT* and its performance are documented in Cohn et al. (2013), Lamontagne et al. (2013), and Lamontagne et al. (2015). Some examples of detecting PILFs with the *MGBT* are provided in Appendix 9.

PILF Background and Philosophy

There are recognized problems with the Grubbs-Beck (GB) test (Grubbs and Beck, 1972) used in Bulletin 17B (IACWD, 1982) when there are multiple low outliers. This issue was discussed in the Bulletin 17B FAQ, under the section "Low Outliers" written by Bill Kirby. Relevant portions of that FAQ are reproduced here, with *emphasis* on the important issues.

Bulletin-17-B detects low outliers by means of a statistical criterion (the Grubbs-Beck test) rather than by consideration of the influence of low-lying data points on the fit of the frequency curve. The test is based on the standardized distances, $(x_i - \hat{\mu})/\hat{\sigma}$, between the lowest observations and the mean of the data set. *The test is easily defeated by occurrence of multiple low outliers, which exert a large distorting influence on the fitted frequency curve, but also increase the standard deviation, $\hat{\sigma}$, thereby making the standardized distance too small to trigger the Grubbs-Beck test.*

The FAQ also provides further background, and a hydrological basis to deviate from the GB test as follows, with *emphasis* on the relevant text.

Obviously, the intention is to allow as many low outliers to be designated as necessary to achieve a good fit to the part of the data set that contains the significant flood and near-flood events. Equally obviously, the intention is that the Grubbs-Beck result be used unless the resulting poor fit gives compelling justification for not doing so. There is no universal method that can be followed blindly to achieve a good fit. The sensitivity analysis alluded to in Bulletin 17-B is *based on the engineering-hydrologic-common-sense proposition that the smallest observations in the data set do not convey meaningful or valid information about the magnitude of significant flooding, although they do convey valid information about the frequency of significant flooding. Therefore, if the upper tail of the frequency curve is sensitive to the numerical values of the smallest observations, then that sensitivity is a spurious artifact based on the mathematical form of the assumed but in fact unknown flood distribution, and has no hydrologic validity.*

Others have noted this hydrologic phenomenon. A key observation is from Klemeš (1986, p. 183S), reproduced as follows. "For it is by no means hydrologically obvious why the regime of the highest floods should be affected by the regime of flows in years when no floods occur, why the probability of a severe storm hitting this basin should depend on the accumulation of snow in the few driest winters, why the return period of a given heavy rain should be by an order of magnitude different depending, say, on slight temperature fluctuations during the melting seasons of a couple of years (p. 183S)."

Klemeš (2000, p. 229) also described this hydrological problem in the context of frequency distrbutions, as follows, with *emphasis* on the relevant text.

"…It is ironic that the only clue the FA (Frequency Analysis) theory inadvertently takes from hydrology is the wrong one. It derives the "distributional assumptions" [i.e., the general shape of F(X)] from a "probability plot" such as Fig. 1(b) whose shape is dominated by the small and medium observations. This shape is generally convex on the Gaussian plot, because hydrological phenomena like precipitation, runoff, snow cover, etc., have a zero lower bound, which "bends" the lower tail of the plot towards a horizontal asymptote. As a result, all the "standard" distribution models are convex on Gaussian frequency scale; they all are models with positive skewness. *Hence, it is the physical regime prevailing in the formation of the lower tail that determines the shape* *of the extrapolated upper tail; observations that are hydrologically least relevant to the high extremes and to the safety of facilities affected by them — have the greatest influence on their estimated "probabilities"! ...*"

These observations, as well as the data issues described in the Section Zero Flows and Potentially-Influential Low Floods, are handled with the *MGBT*.

Computational Details for Identifying PILFs with *MGBT*

The purpose of using the *MGBT* is to identify PILFs. PILFs are small observations (or zero flows) that potentially have a large influence on the fitted frequency curves. When data sets are negatively skewed, the smallest observations can be very influential in determining the estimated skewness coefficient and the estimated 1% *AEP* flood. The new *MGBT* is a statistically appropriate generalization of the GB test that is sensitive to the possibility that several of the smallest observations are "unusual," or are potentially very influential. The *MGBT* also correctly evaluates cases where one or more observations are zero, or are below a recording threshold (partial record sites). Thus it provides a consistent, objective and statistically defensible algorithm that considers whether a range of the smallest observations should be classified as PILFs for a wide range of situations that are observed in practice. See, for example, cases in Lamontagne et al. (2012), Paretti et al. (2014a), and examples in Appendix 9.

To provide an objective criteria for multiple low outlier identification, *MGBT* employs the actual distribution of the k^{th} largest observation in a sample of n independent normal variates, where the probability $p_{[k:n]}$ that the k^{th} largest observation in a normal sample of size n might have appeared to be smaller than the value observed (Cohn et al., 2013). If $p_{[k:n]}$ is small, then the k^{th} observation is unusually small.

To test H_0, we consider whether $\{X_{[1:n]}, X_{[2:n]}, \ldots, X_{[n:n]}\}$ are consistent with a normal distribution and the other observations in the sample by examining the statistic

$$\tilde{\omega} \equiv \frac{X_{[k:n]} - \hat{\mu}_k}{\hat{\sigma}_k} \tag{5.1}$$

where $X_{[k:n]}$ denotes the k-th smallest order statistic in the sample, and

$$\hat{\mu}_k = \frac{1}{n-k} \sum_{j=i+1}^{n} X_{[j:n]} \tag{5.2}$$

$$\hat{\sigma}_k^2 = \frac{1}{n-k-1} \sum_{j=i+1}^{n} (X_{[j:n]} - \hat{\mu}_k)^2. \tag{5.3}$$

The partial mean ($\hat{\mu}_k$) and partial variance ($\hat{\sigma}_k^2$) are computed based on all observations larger than $X_{[k:n]}$ to avoid swamping. These larger observations are not suspected of being low outliers, thus $\hat{\mu}_k$ and $\hat{\sigma}_k^2$ are assumed to correspond to the population of interest. From $\tilde{\omega}$, we calculate the $p-$value: the probability given H_0 of obtaining a value of $\tilde{\omega}_{[k:n]}$ as small or smaller than that observed in the sample. The $p-$value of interest is given by

$$p_k[\eta] \equiv P\left[\tilde{\omega}_{[k:n]} < \eta\right]. \tag{5.4}$$

Substituting the definition of $\tilde{\omega}_{[k:n]}$ from equation 5.1 and rearranging the terms yields (Cohn et al., 2013)

$$p_k[\eta] = P\left[\left(\frac{Z_{[k:n]} - \hat{\mu}_{Z,k}}{\hat{\sigma}_{Z,k}}\right) < \eta\right] \tag{5.5}$$

where $Z_{[k:n]}$ is the k-th order statistic in a standard normal sample of size n, and $\hat{\mu}_{Z,k}, \hat{\sigma}_{Z,k}$ are the partial mean and standard deviation of the normal sample. If that $p-$value is small (for example, less than α = 10%), then the k smallest observations are declared PILFs, such as those shown in Figure 11. The PILF threshold X_l that is used in *EMA* is set to the $(k+1)$-th value.

The *MGBT* for identifying Potentially Influential Low Floods has two steps. The input data are base-10 logarithms X_j of annual peak flows from the systematic (gaging) record (n_s), with flow values exactly known as point observations ($Q_{Y,lower} = Q_{Y,upper} = Q_Y$). Flows are ranked from smallest to largest, as noted in the Section Zeros and Identifying Potentially-Influential Low Floods.

1. Starting at the median and sweeping *outward* towards the smallest observation, each observation $X_{[k:n]}$ is tested and is identified as an outlier if $p(k;n) \leq \alpha_{out}$. If the k^{th} largest observation is identified as a low outlier, the outward sweep stops and the k^{th} and all smaller observations (i.e. for all $j \leq k$) are also identified as low outliers.

2. An *inward* sweep starts at the smallest observation $X_{[1:n]}$ and moves towards the median, where the j^{th} observation is identified as an outlier if $p(k;n) \leq \alpha_{in}$. If an observation $m = 1, 2, \ldots, n/2$ fails to be identified as an outlier by the inward sweep, the inward sweep stops.

The number of PILFs identified by the procedure is then the larger of k and $m-1$.

The algorithm has two parameters: an *outward sweep* significance level α_{out}, and an *inward sweep* significance level α_{in}. The recommended values used in *MGBT* are $\alpha_{out} = 0.005$ (0.5%) and $\alpha_{in} = 0.10$ (10%). These values were determined through extensive testing and evaluation by the HFAWG through careful examination of 82 sites (Cohn et al., 2014), testing and performance of alternatives (Lamontagne et al., 2013), and further investigations (Lamontagne et al., 2015).

The *outward* sweep seeks to determine if there is some break in the lower half of the data that would suggest the sample is best treated as if it had a number of low outliers. The *inward* sweep using a less severe significance level, $p(k;n) \leq 10\%$, mimics Bulletin 17B's willingness to identify one or more of the smallest observations as low outliers so that the analysis is more robust. Bulletin 17B also used a 10% significance test with its single GB threshold. However, a critical difference is that the *MGBT inward* sweep uses the $p(k;n)$ function which correctly describes whether the k^{th} largest observation in a normal sample of n variates is unusual.

For example, if a record has 5 zero flows, then the smallest non-zero flow is considered to be the 6^{th} smallest observation in the record. This correctly reflects the fact that the flood record included 5 smaller values. The GB test in Bulletin 17B includes no mechanism for correcting its threshold when testing the smallest non-zero flood value in a record containing zeros, or below-threshold discharges at sites with crest-stage gages. This is particularly problematic because sites with zero flows are very likely to include one or more very small or near-zero flood values which should legitimately be identified as low outliers were a statistically appropriate threshold employed. The *MGBT* solves this problem.

Computer programs (see the Section Software and Examples) are used to perform the *MGBT* and report critical values and PILFs.

Appendix 6—Expected Moments Algorithm (EMA)

Expected Moments Algorithm (EMA)

This appendix describes features of *EMA*, including some computation details, a generalized expected moments algorithm, and uncertainty of *EMA* moments, and confidence intervals with *EMA*.

EMA Computational Details

The *EMA* moments for the general situation with a historical flood perception threshold X_h and a PILF threshold X_l are as follows:

$$\hat{\mu}_{i+1} = \frac{\sum X_s^> + \sum X_l^> + \sum X_h^> + n_l^< \mathrm{E}[X_l^<] + n_h^< \mathrm{E}[X_h^<]}{n_s + n_h} \tag{6.1}$$

$$\hat{\sigma}^2_{i+1} = \frac{c_2}{n}\Bigg[\sum(X_s^> - \hat{\mu}_i)^2 + \sum(X_l^> - \hat{\mu}_i)^2 + \sum(X_h^> - \hat{\mu}_i)^2 + n_l^< \mathrm{E}[(X_l^< - \hat{\mu}_i)^2] + n_h^< \mathrm{E}[(X_h^< - \hat{\mu}_i)^2]\Bigg] \tag{6.2}$$

$$\hat{\gamma}_{i+1} = \frac{c_3}{n\hat{\sigma}^3_{i+1}}\Bigg[\sum(X_s^> - \hat{\mu}_i)^3 + \sum(X_l^> - \hat{\mu}_i)^3 + \sum(X_h^> - \hat{\mu}_i)^3 + n_l^< \mathrm{E}[(X_l^< - \hat{\mu}_i)^3] + n_h^< \mathrm{E}[(X_h^< - \hat{\mu}_i)^3]\Bigg] \tag{6.3}$$

where c_2 and c_3 are bias correction factors, defined as

$$c_2 = \frac{n_s + n_h^>}{n_s + n_h^> - 1} \tag{6.4}$$

$$c_3 = \frac{(n_s + n_h^>)^2}{(n_s + n_h^> - 1)(n_s + n_h^> - 2)} \tag{6.5}$$

and recalling $n_s + n_h = n$.

The expression $E[X_h^<]$ is the expected value of an observation known to have a value less than the historical threshold X_h, and is a conditional expectation given that $X < X_h$, and is evaluated with

$$E[X|X \le X_h; \hat{\tau}, \hat{\alpha}, \hat{\beta}] = \hat{\tau} + \hat{\beta}\,\frac{\Gamma\left(\frac{X_h - \hat{\tau}}{\beta}, \hat{\alpha} + 1\right)}{\Gamma\left(\frac{X_h - \hat{\tau}}{\beta}, \hat{\alpha}\right)} \tag{6.6}$$

where $\Gamma(y, \alpha)$ is the incomplete gamma function:

$$\Gamma(y, \alpha) = \int_0^y t^{\alpha - 1} \exp(-t)dt. \tag{6.7}$$

The expectation for higher-order moments is:

$$E[(X-\hat{\mu})^p|X \le X_h;\hat{\tau},\hat{\alpha},\hat{\beta}] = \sum_{j=0}^{p}\binom{p}{j}\hat{\beta}^j(\hat{\tau}-\hat{\mu})^{p-j}\left[\frac{\Gamma\left(\frac{X_h-\hat{\tau}}{\beta},\hat{\alpha}+j\right)}{\Gamma\left(\frac{X_h-\hat{\tau}}{\beta},\hat{\alpha}\right)}\right] \tag{6.8}$$

where p is the central moment index ($p = 2,3$). The conditional expectation for PILFs with $X < X_l$ and threshold X_l are similar to equations (6.6) and (6.8).

The *EMA* moments shown in equations (6.1)-(6.3), and expected values shown in equations (6.6) and (6.8), utilize observations whose magnitudes are exactly known, where $X_l = X_u$. In the cases where flow magnitudes are described by intervals or binomial observations, these equations are modified to account for logarithms of flow intervals $X_{Y,lower}$, $X_{Y,upper}$ and are presented in equation 6.9. Information from broken, incomplete, and discontinued records, crest-stage gages, and multiple thresholds (e.g., Figure 12) is easily represented by including additional expected value terms in the moments for each year Y or period where the flow interval or perception threshold varies.

The *EMA* employs the peak-flow intervals (Q_l, Q_u) to estimate the moments of the LP-III distribution. Using base 10 logarithms of flows, where $X_l = log_{10}(Q_l)$ and $X_u = log_{10}(Q_u)$, interval and binomial censored data are employed by replacing equation (6.8) with (Cohn et al., 1997):

$$E[(X-\hat{\mu})^p|X_l \le X \le X_u;\hat{\tau},\hat{\alpha},\hat{\beta}] = \sum_{j=0}^{p}\binom{p}{j}\hat{\beta}^j(\hat{\tau}-\hat{\mu})^{p-j} \cdot \left[\frac{\Gamma\left(\frac{X_u-\hat{\tau}}{\beta},\hat{\alpha}+j\right)-\Gamma\left(\frac{X_l-\hat{\tau}}{\beta},\hat{\alpha}+j\right)}{\Gamma\left(\frac{X_u-\hat{\tau}}{\beta},\hat{\alpha}\right)-\Gamma\left(\frac{X_l-\hat{\tau}}{\beta},\hat{\alpha}\right)}\right]. \tag{6.9}$$

When information from a regional skew coefficient G is available, it is included directly in the *EMA*, ensuring that the adjusted mean and standard deviation fit the data. Equation (6.3) for the skew coefficient $\hat{\gamma}_{i+1}$ is modified to include G, as:

$$\hat{\gamma}_{i+1} = \frac{1}{(n+n_G)\hat{\sigma}^3_{i+1}}\left[c_3\left\{\sum(X_s^> - \hat{\mu}_i)^3 + \sum(X_l^> - \hat{\mu}_i)^3 + \sum(X_h^> - \hat{\mu}_i)^3 + n_l^< \mathrm{E}[(X_l^< - \hat{\mu}_i)^3] + n_h^< \mathrm{E}[(X_h^< - \hat{\mu}_i)^3]\right\} + n_G G\hat{\sigma}^3_{i+1}\right] \tag{6.10}$$

where n_G is the additional years of record assigned to the regional skew G (Griffis et al., 2004). A skew constraint is imposed on each *EMA* iteration so that $\hat{\gamma}_{i+1} > -1.4$, as it is unlikely that the population skew would be less than -1.4.

A general listing of computations for flood flow frequency using *EMA*, that are implemented in software (see the Section Software and Examples), are as follows.

1. Check for low outliers with *MGBT*. If low outliers are detected, recode flows as censored data with an interval $Q_{Y,lower} = 0$; $Q_{Y,upper} = Q_l$. Adjust perception thresholds accordingly, $T_{Y,lower} = Q_l$; $T_{Y,upper} = \infty$.

2. Organize all flow intervals and perception thresholds for estimating parameters and confidence intervals.

3. Begin iterative fitting of the LP-III distribution using *EMA* with all data, including regional skew information. For each iteration, ensure that the weighted skew coefficient $\tilde{G} \geq -1.41$ and the largest observation is within the fitted support of the distribution (for skews < 0).

   (a) Fit the LP-III with *EMA* using at-site data, to estimate the at-site skew.

   (b) Estimate the at-site skew coefficient MSE with *EMA*.

   (c) Estimate a weighted skew coefficient.

   (d) Fit the LP-III with *EMA* using a weighted skew coefficient.

   (e) Test for convergence of *EMA* moments. If not converged, return to 3a.

4. Estimate quantile variances and compute confidence intervals based on the fitted LP-III model, including at-site and regional skew uncertainty.

The Generalized Expected Moments Algorithm (EMA)

This section presents parameterizations of the P-III distribution and a generalized Expected Moments Algorithm. The notation and terms are utilized to explain uncertainty of *EMA* moments and confidence intervals. Bold terms, such as $\mathbf{M}$ and $\boldsymbol{\theta}$ are used to indicate vectors or matrices. Carets (ˆ) represent a sample estimate, and tildes (˜) indicate non-central moments (on scalars) or estimators (on vectors).

The P-III distribution is typically characterized by three parameters that correspond to location $\{\tau\}$, scale $\{\beta\}$ and shape $\{\alpha\}$, where the vector $\boldsymbol{\theta} = \{\tau, \alpha, \beta\}$. The P-III distribution is also characterized by non-central moments $\boldsymbol{\mu} = \{\tilde{\mu}_1, \tilde{\mu}_2, \tilde{\mu}_3\}$ (about zero) for algebraic tractability, and central moments $\mathbf{M}$=$\{M, S, G\} = \{\mu, \sigma, \gamma\}$ for simplicity of explanation.

Central moments are defined as:

$$\mathbf{M} = \begin{bmatrix} M \\ S \\ G \end{bmatrix} \equiv \begin{bmatrix} E[X] \\ \sqrt{E[(X-M)^2]} \\ E[(X-M)^3/S^3] \end{bmatrix} \equiv \begin{bmatrix} \tilde{\mu}_1 \\ \sqrt{\tilde{\mu}_2 - \tilde{\mu}_1^2} \\ \frac{\tilde{\mu}_3 - 3\tilde{\mu}_2\tilde{\mu}_1 + 2\tilde{\mu}_1^3}{\sqrt{\tilde{\mu}_2 - \tilde{\mu}_1^2}^3} \end{bmatrix}$$

$$\equiv \begin{bmatrix} \tau + \alpha\beta \\ \sqrt{\alpha\beta^2} \\ \operatorname{sign}(\beta) 2/\sqrt{\alpha} \end{bmatrix}. \tag{6.11}$$

Non-central moments are:

$$\tilde{\boldsymbol{\mu}} \equiv \begin{bmatrix} \tilde{\mu}_1 \\ \tilde{\mu}_2 \\ \tilde{\mu}_3 \end{bmatrix} \equiv \begin{bmatrix} E_{\boldsymbol{\theta}}[X] \\ E_{\boldsymbol{\theta}}[X^2] \\ E_{\boldsymbol{\theta}}[X^3] \end{bmatrix} \equiv \begin{bmatrix} M \\ S^2 + M^2 \\ S^3 G + 3S^2 M + M^3 \end{bmatrix}$$

$$= \begin{bmatrix} \alpha\beta + \tau \\ \alpha(1+\alpha)\beta^2 + 2\alpha\beta\tau + \tau^2 \\ \alpha(1+\alpha)(2+\alpha)\beta^3 + 3\alpha(1+\alpha)\beta^2\tau + 3\alpha\beta\tau^2 + \tau^3 \end{bmatrix}. \tag{6.12}$$

And the P-III distribution parameters are:

$$\boldsymbol{\theta} = \begin{bmatrix} \alpha \\ \beta \\ \tau \end{bmatrix} = \begin{bmatrix} 4/G^2 \\ SG/2 \\ M-2S/G \end{bmatrix} = \begin{bmatrix} \frac{4(\tilde{\mu}_2-\tilde{\mu}_1^2)^3}{(\tilde{\mu}_3-3\tilde{\mu}_2\tilde{\mu}_1+2\tilde{\mu}_1^3)^2} \\ \frac{\tilde{\mu}_3-3\tilde{\mu}_2\tilde{\mu}_1+2\tilde{\mu}_1^3}{2(\tilde{\mu}_2-\tilde{\mu}_1^2)} \\ \frac{\tilde{\mu}_3\tilde{\mu}_1-2\tilde{\mu}_2^2+\tilde{\mu}_2\tilde{\mu}_1^2}{\tilde{\mu}_3-3\tilde{\mu}_2\tilde{\mu}_1+2\tilde{\mu}_1^3} \end{bmatrix}. \tag{6.13}$$

Here $E[]$ denotes the expectation operator, and the $\tilde{}$ is used to identify non-central moments. The formulas in equations 6.11, 6.13 and 6.12 facilitate converting one parametrization to another. When using sample estimates, the conversion from non-central moments $\hat{\boldsymbol{\mu}}$ to central moments $\mathbf{M}$ needs to include bias-correction factors with

$$\hat{\mathbf{M}} = \begin{bmatrix} 1 & \frac{N}{N-1} & \frac{\sqrt{N(N-1)}}{(N-2)} \end{bmatrix} * (\hat{\boldsymbol{\mu}}). \tag{6.14}$$

A generalized Expected Moments Algorithm, employing central moments $\mathbf{M}$, is as follows, where N is the total record length.

1. Initialize

   (a) Set $\hat{\mathbf{M}}_0 = \{0,1,0\}$

   (b) Define $\varepsilon > 0$ as a satisfactory level of convergence.

   A typical value for ε is 10^{-10}.

2. Iterate: for $j = 1,2...$

   (a) Update expected moments

$$\hat{\mathbf{M}}_j = \begin{Bmatrix} M_j \\ S_j^2 \\ G_j \end{Bmatrix} = \begin{Bmatrix} \frac{1}{N}\sum_{i=1}^{N} E_{\tilde{\mathbf{M}}_{j-1}}[X_i | X_{i,\mathrm{l}} \le X_i < X_{i,\mathrm{u}}] \\ \frac{1}{N-1}\sum_{i=1}^{N} E_{\tilde{\mathbf{M}}_{j-1}}[(X_i - M_j)^2 | X_{i,\mathrm{l}} \le X_i < X_{i,\mathrm{u}}] \\ \frac{N}{S^3(N-1)(N-2)}\sum_{i=1}^{N} E_{\tilde{\mathbf{M}}_{j-1}}[(X_i - M_j)^3 | X_{i,\mathrm{l}} \le X_i < X_{i,\mathrm{u}}] \end{Bmatrix} \tag{6.15}$$

   where

$$E_{\tilde{\mathbf{M}}_{j-1}}[(X_i - M_j)^k | X_{i,\mathrm{l}} \le X_i < X_{i,\mathrm{u}}] = \sum_{l=0}^{k} \binom{k}{l} E_{\tilde{\mathbf{M}}_{j-1}}[X_i^l | X_{i,\mathrm{l}} \le X_i < X_{i,\mathrm{u}}](-M_j)^{k-l} \tag{6.16}$$

   and, if the upper and lower bounds on X_i are equal (*i.e.* $X_{i,\mathrm{l}} = X_{i,\mathrm{u}}$, which means we know the exact value of X_i), then

$$E_{\tilde{\mathbf{M}}_{j-1}}[X_i^k | X_{i,\mathrm{l}} \le X_i < X_{i,\mathrm{u}}] = X_{i,\mathrm{l}}^k = X_{i,\mathrm{u}}^k \tag{6.17}$$

   If $X_{i,\mathrm{l}} < X_{i,\mathrm{u}}$, then we have to evaluate the expectation:

$$E_{\theta}[X^k | X_{i,\mathrm{l}} \le X < X_{i,\mathrm{u}}] = \begin{cases} \sum_{j=0}^{k} \binom{k}{j} \beta^j \tau^{k-j} \left(\frac{\Gamma(\alpha+j, \frac{X_{i,\mathrm{u}}-\tau}{\beta}, \frac{X_{i,\mathrm{l}}-\tau}{\beta})}{\Gamma(\alpha, \frac{X_{i,\mathrm{u}}-\tau}{\beta}, \frac{X_{i,\mathrm{l}}-\tau}{\beta})} \right) & \beta < 0 \\ \sum_{j=0}^{k} \binom{k}{j} \beta^j \tau^{k-j} \left(\frac{\Gamma(\alpha+j, \frac{X_{i,\mathrm{l}}-\tau}{\beta}, \frac{X_{i,\mathrm{u}}-\tau}{\beta})}{\Gamma(\alpha, \frac{X_{i,\mathrm{l}}-\tau}{\beta}, \frac{X_{i,\mathrm{u}}-\tau}{\beta})} \right) & \beta > 0 \end{cases} \tag{6.18}$$

where θ is the P-III parameters corresponding to $\tilde{\mathbf{M}}_{j-1}$, and

$$\Gamma(\alpha, X_{i,\mathrm{l}}, X_{i,\mathrm{u}}) = \int_{\max(0,X_{i,\mathrm{l}})}^{\max(0,X_{i,\mathrm{u}})} t^{\alpha-1} \exp(-t)\,dt \tag{6.19}$$

(b) If available, weight with regional skew. This can be done conceptually via:

$$\tilde{G}_j = \frac{MSE_G \hat{\gamma}_j + MSE_{\hat{\gamma}_j} G}{MSE_G + MSE_{\hat{\gamma}_j}} \tag{6.20}$$

and the number of years are used as weights, as in equation 6.10.

(c) Test for convergence. If $||\hat{\mathbf{M}}_j - \hat{\mathbf{M}}_{j-1}|| < \varepsilon$, return $\mathbf{M} = \hat{\mathbf{M}}_j$ as the EMA estimate. Otherwise, increment j and return to 2a.

Uncertainty of EMA Moments

Uncertainty of moments, specifically the at-site skew coefficient $(\hat{\gamma})$, are estimated with *EMA*. Details and equations are presented in Appendix A1 of Cohn et al. (2001) and in Cohn (2015).

For cases where there is historical information, PILFs, a gage base discharge, or some type of censored or interval data, *EMA* utilizes an approach to estimate $MSE_{\hat{\gamma}}$ that is based on all the data. This includes censored data, intervals, historical information and PILFs, including the P-III distribution parameters, as they are used in estimating the moments with *EMA*. Conceptually, this is done as follows:

$$MSE_{\hat{\gamma}} \approx VAR(\hat{\gamma}) \approx VAR(\hat{m}_3) \approx \frac{1}{n} f(X_{i,l}, X_{i,u}, T_{i,l}, T_{i,u}, \hat{\boldsymbol{\theta}}) \tag{6.21}$$

where $MSE_{\hat{\gamma}}$ is proportional to the variance (VAR) of the skew $(\hat{\gamma})$, and is proportional to the variance of the third non-central moment $VAR(\hat{m}_3)$, and is a function (f) of the observations (including censored data), and P-III parameters $\hat{\boldsymbol{\theta}}$. In this case n is the total record length (e.g. $n = n_s + n_h = N$), including any historical period, PILFs, censored and interval data.

As presented in Appendix A1 of Cohn et al. (2001), *EMA* estimates the variance of the each of the non-central moments $\{\hat{\boldsymbol{\mu}} = [\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3]\}$, where $\{\hat{\boldsymbol{\mu}} = \hat{\mathbf{M}} = [m_1, m_2, m_3]\}$ as outlined in Cohn et al. (2001). The non-central moment $\hat{m}_3$ (moment computed around zero) is used to estimate $MSE_{\hat{\gamma}}$. Key equations from Cohn et al. (2001, Appendix A1) are presented below.

The *EMA* estimates non-central moments $\hat{\mathbf{M}} = [\hat{m}_1, \hat{m}_2, \hat{m}_3]$, that directly take into account censored data, via:

$$\hat{\mathbf{M}} = (1/n) \sum_{i=1}^{n} \boldsymbol{\chi}(\psi(X_i), \hat{\mathbf{M}}) \mathscr{I}[\psi(X_i)] \tag{6.22}$$

where

$$\mathscr{I}[X] \equiv \begin{bmatrix} \mathscr{I}(X < a) \\ \mathscr{I}(a \le X \le b) \\ \mathscr{I}(X > b) \end{bmatrix} \tag{6.23}$$

$$\mathscr{I}(condition) \equiv \begin{cases} 1 & condition = true \\ 0 & otherwise \end{cases} \tag{6.24}$$

Table 6.1. *EMA* censored-data threshold categories.

Value of X_i	Category	$(T_{i,l}, T_{i,u})$
$x < a$	l	$(-\infty, a)$
$a \leq X \leq b$	b	(X, X)
$x > b$	g	(b, ∞)

and

$$\boldsymbol{\chi}(\psi(X),\mathbf{M}) = \begin{bmatrix} E_{\boldsymbol{\theta}[\mathbf{M}]}[X|X<a] & X_i & E_{\boldsymbol{\theta}[\mathbf{M}]}[X|X>b] \\ E_{\boldsymbol{\theta}[\mathbf{M}]}[X^2|X<a] & X_i^2 & E_{\boldsymbol{\theta}[\mathbf{M}]}[X^2|X>b] \\ E_{\boldsymbol{\theta}[\mathbf{M}]}[X^3|X<a] & X_i^3 & E_{\boldsymbol{\theta}[\mathbf{M}]}[X^3|X>b] \end{bmatrix}. \tag{6.25}$$

The function $\mathscr{I}[X]$ defines the censored data category for the flow logarithms X. Three categories are used: "less", where X is less than the "perception threshold" a; "between", where X is within the closed interval $[a,b]$; or "greater" if X is known to exceed some "perception threshold" b (Table 6.1). These threshold categories $[a,b]$ correspond to those described in the Section Data Representation using Flow Intervals and Perception Thresholds, where "between" is the "interval" category, "greater" is the "binomial" category. The "less than" category covers unobserved historical floods, flows less than a gage base, or low outliers. The magnitude of X is known if X is within $[a,b]$. Only a threshold on X can be identified if $X < a$ or $X > b$. The number of observations in each of these categories is a random variable, denoted n_l, n_b, and n_g, respectively. Because each X must fall into one of the three categories, the total sample size n is constant, where $n = n_l + n_b + n_g$.

The $MSE_{\hat{\gamma}}$ can be estimated by taking the variance of equation 6.22, as in equation 6.26. The formula for the asymptotic variance of the *EMA* moments estimator, denoted $\tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\mu}}}$, is derived in Cohn et al. (2001, Appendix A1). It is obtained by linearizing the expectations in equation 6.22 and solving for $\mathbf{M}$ in terms of the sample X_i values. The estimator $\tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\mu}}}$ is then expressed as a function of the population parameters, the record lengths, and the censoring thresholds. It can be used as an *estimator* of the variance-covariance matrix given estimated parameters ($\hat{\boldsymbol{\Sigma}}_{\hat{\mathbf{M}}}$).

$$VAR\,\hat{\mathbf{M}} = VAR \begin{bmatrix} \hat{m}_1 \\ \hat{m}_2 \\ \hat{m}_3 \end{bmatrix} \approx \tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\mu}}} \approx \hat{\boldsymbol{\Sigma}}_{\hat{\mathbf{M}}} \tag{6.26}$$

The variance of $\hat{\mathbf{M}}$ is (Cohn et al., 2001, equation 55):

$$\tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\mu}}} = \frac{1}{n^2}\mathbf{A}(Var[\mathbf{B}] + Var[\mathbf{C}])\mathbf{A}' \tag{6.27}$$

where

$$\begin{aligned}
\mathbf{B} &= \boldsymbol{\mu}_{\mathbf{X}}\mathbf{n} \\
\mathbf{C} &= \sum_{i=1}^{\mu_{n_b}} (\mathbf{X}_i - \boldsymbol{\mu}_{\mathbf{X}_b}) \\
\mathbf{D} &= \frac{\mu_{n_l}\mathbf{J}_l + \mu_{N_g}\mathbf{J}_g}{n} \\
\mathbf{A} &= (\mathbf{I} - \mathbf{D})^{-1}
\end{aligned} \tag{6.28}$$

and $\boldsymbol{\mu}_{\mathbf{X}}$ is the vector of expected values for non-central moments given parameters and value of X (Cohn et al., 2001, equations 50-51). The variance of $\mathbf{B}$ is given by:

$$Var[\mathbf{B}] = \boldsymbol{\mu}_{\mathbf{X}} Var[\mathbf{n}] \boldsymbol{\mu}_{\mathbf{X}}' \tag{6.29}$$

The large-sample variance of $\mathbf{C}$ is the expected value of the number of terms multiplied by the variance of each term:

$$Var[\mathbf{C}] = \mu_{n_B} \begin{bmatrix} V_{1,1} & V_{1,2} & V_{1,3} \\ V_{2,1} & V_{2,2} & V_{2,3} \\ V_{3,1} & V_{3,2} & V_{3,3} \end{bmatrix} \tag{6.30}$$

The MSE of the *EMA* at-site skewness coefficient is estimated using a first-order approximation (Cohn et al., 2001, equation 55), reproduced above as equation 6.27, with $\hat{m}_3$ as the non-central moment of interest.

Confidence Intervals with EMA

A simple formula for a confidence interval on a flood quantile $\hat{X}_q$ is (Stedinger et al., 1993; Cohn et al., 2001):

$$\hat{X}_q \pm z_{1-\alpha/2}\sqrt{\mathrm{var}(\hat{X}_q)} \tag{6.31}$$

where q is the quantile of interest (e.g. $q = 0.99$), $z_{1-\alpha/2}$ is the $(1-\alpha)/2$ quantile of the standard Normal distribution, α is the confidence level and

$$\sqrt{\mathrm{var}(\hat{X}_q)} = \hat{\sigma}_{\hat{X}_q} \tag{6.32}$$

is the estimated standard error of the flood quantile. Typically the confidence level $\alpha = 0.05$, resulting in a 90% confidence interval (5- and 95-% confidence limits).

Confidence intervals for flood quantiles $(\hat{X}_p)$ are estimated with *EMA*. Cohn et al. (2001) derive *EMA* confidence intervals in detail and provide key equations. Cohn (2015) improve the *EMA* confidence intervals for skews $|\hat{\gamma}| > 0.5$.

Confidence intervals are estimated using:

$$\left(\hat{X}_p + \frac{\hat{\sigma}_{\hat{X}_p} T_{\nu,(1-\varepsilon)/2}}{1-\kappa T_{\nu,(1-\varepsilon)/2}}, \hat{X}_p + \frac{\hat{\sigma}_{\hat{X}_p} T_{\nu,(1+\varepsilon)/2}}{1-\kappa T_{\nu,(1+\varepsilon)/2}}\right) \tag{6.33}$$

where T_ν is a Student's T variate (Abramowitz and Stegun, 1964), ε is the confidence level, $\hat{\sigma}_{\hat{X}_p}$ is the standard deviation of the quantile $\hat{X}_p$ and

$$\kappa \equiv \frac{\hat{Cov}[\hat{X}_p, \hat{\sigma}_{\hat{X}_p}]}{\hat{\sigma}^2_{\hat{X}_p}} \tag{6.34}$$

is a function of the sample size and the censoring threshold (and, to some extent, of α). Estimators for $Cov[\hat{X}_p, \hat{\sigma}_{\hat{X}_p}]$ and $\sigma^2_{\hat{X}_p}$ are available from Cohn et al. (2001, equation 70).

The asymptotic variance of $\hat{X}_p$ can be obtained from a first-order expansion of $\hat{X}_p$ as a function of $\mathbf{M}$:

$$\hat{X}_p \approx X_p + \mathbf{J}_{\hat{X}_p}(\mathbf{M} - \boldsymbol{\mu}_{\mathbf{M}}) \tag{6.35}$$

where

$$\mathbf{J}_{\hat{X}_p} = \begin{bmatrix} \frac{\partial \hat{X}_p}{\partial \hat{m}_1} & \frac{\partial \hat{X}_p}{\partial \hat{m}_2} & \frac{\partial \hat{X}_p}{\partial \hat{m}_3} \end{bmatrix} \tag{6.36}$$

The Jacobian can be evaluated by first computing derivatives with respect to $\{\alpha, \beta, \tau\}$ and then applying the chain rule.

The variance of $\hat{X}_p$ can be approximated by:

$$\tilde{\sigma}^2_{\hat{X}_p} \approx \mathbf{J}_{\hat{X}_p} \cdot \tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\mu}}} \cdot \mathbf{J}_{\hat{X}_p}{}' \tag{6.37}$$

where the linearized standard deviation, $\tilde{\sigma}_{\hat{X}_p}$, is defined as $\sqrt{\tilde{\sigma}^2_{\hat{X}_p}}$.

Cohn (2015) provides improved estimates of $Var[\hat{X}_p]$ using inverse quadrature.

Appendix 7—Record Extension with Nearby Sites

Record Extension with Nearby Sites

This appendix describes the background, guidance, computational details, and an example for record extension. Record extension uses information from a nearby longer flood record site to extend the time series at a short record site, when the cross-correlation between the two sites is high.

Background

Matalas and Jacobs (1964) developed an approach for obtaining unbiased estimates of the mean and variance of the lengthened time series (observed plus extended record) using OLS regression without adding a random noise term. This approach is the basis of the "Two Station Comparison" method that is described in Bulletin 17B (IACWD, 1982, Appendix 7). In that method, improved estimates of the mean and variance (standard deviation) are obtained but no additional years of record are estimated. In order to be useful in flood frequency with *EMA*, annual peak-flow estimates are needed, rather than improved estimates of the flow statistics (mean and variance).

The Maintenance of Variance Extension (MOVE) techniques, as described by Hirsch (1982), are useful approaches for extending flood data in time, based on records from a nearby site. As implied by the name, these techniques maintain the variance of the estimated data unlike ordinary least squares (OLS) regression. If an ensemble of points is estimated from OLS regression, the estimated values will have lesser variability than the true or population values unless a random noise term is added. However, adding a random noise term to the regression estimates does not achieve a unique extended record. Therefore, the loss of variance from regression analysis is a problem if the estimated values are used in subsequent statistical analyses such as frequency analyses.

Hirsch (1982) described two MOVE techniques, MOVE.1 and MOVE.2. The MOVE procedures are based on only one independent variable and the assumption is that there is a linear relation between the dependent and independent variables. If the annual peak flows are not linearly related, then it is common to transform the data using a logarithmic transformation because the logarithms of the flows tend to be linearly related. The MOVE techniques are another way of fitting a linear relation to data similar to OLS regression.

The MOVE.1 technique described by Hirsch (1982) is the simplest and uses the n_1 years of concurrent record at the two sites. Only the means and standard deviations for y_i and x_i for the concurrent record are used to define the MOVE.1 relation. The MOVE.2 technique described by Hirsch (1982) utilizes the Matalas-Jacobs estimators Matalas and Jacobs (1964) for the mean and variance of y_i at the short-term station plus the additional years of record x_{n_1+1} to $x_{n_1+n_2}$ at the long-term station that were not observed at the short-term station. The MOVE.1 and MOVE.2 techniques ensure that the moments of the historical sequence (y_1 to y_{n_1}) are preserved.

Vogel and Stedinger (1985) suggested several variations on the MOVE techniques presented by Hirsch (1982). Their MOVE.3 approach provides an estimate of the mean and variance for the short-record site that is correct for the complete record, if the correlation coefficent exceeds a critical value. The MOVE.3 approach is based on a linear relation that ensures the mean and variance of the lengthened sequence (observed plus extended record) (y_1 to $y_{n_1+n_2}$) will be preserved and equal to the Matalas and Jacobs (1964) estimators (Vogel and Stedinger, 1985). Each successive MOVE technique provides slightly more accurate estimates. The recommended approach for record extension is MOVE.3. It is called 'MOVE' in the remainder of this appendix and is described below.

MOVE Technique

Consider a peak-flow time series at a short-record site $y_i, \ldots, y_{n_1}$ with at least 10 years of data. Information from a peak-flow time series at a hydrologically-relevant longer record site $x_i, \ldots, x_{n_1}, x_{n_1+1}, \ldots, x_{n_1+n_2}$, with similar climatic and watershed characteristics, is used to extend the record at site y, where y_i and x_i are the base 10 logarithms of flows Q at sites y and x, n_1 is the length of the short record, and n_1+n_2 is the length of the long record. Note that the n_1 concurrent observations between sites y and x do not need to correspond to the first n_1 observations, nor do they need to be consecutive.

A linear regression model is used to extend the record at the short site y_i (Vogel and Stedinger, 1985):

$$\hat{y}_i = a' + b(x_i - \bar{x}_2) \tag{7.1}$$

where $\bar{x}_2$ is the mean of the longer-term station for the non-overlapping period, estimated from

$$\bar{x}_2 = \frac{1}{n_2} \sum_{i=n_1+1}^{n_2} x_i \tag{7.2}$$

the intercept a' is

$$a' = \frac{(n_1+n_2)\hat{\mu}_y - n_1\bar{y}_1}{n_2} \tag{7.3}$$

the slope is estimated from

$$b^2 = \frac{(n_1+n_2-1)\hat{\sigma}_y{}^2 - (n_1-1)s_{y1}{}^2 - n_1(\bar{y}_1-\hat{\mu}_y)^2 - n_2(a'-\hat{\mu}_y)^2}{(n_2-1)s_{x2}{}^2} \tag{7.4}$$

and $\hat{\mu}_y$ and $\hat{\sigma}_y^2$ are the Matalas and Jacobs (1964) unbiased estimators for the mean and variance of the complete extended record. The sample statistics $(\bar{y}_1, \bar{x}_1, s_{y_1}^2, s_{x_1}^2)$ of the concurrent records in equations 7.3 and 7.4 are estimated as follows:

$$\bar{y}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} y_i \tag{7.5}$$

$$\bar{x}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i \tag{7.6}$$

$$s_{y_1}{}^2 = \frac{1}{n_1-1} \sum_{i=1}^{n_1} (y_i - \bar{y}_1)^2 \tag{7.7}$$

$$s_{x_1}{}^2 = \frac{1}{n_1-1} \sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2 \tag{7.8}$$

and

$$s_{x_2}{}^2 = \frac{1}{n_2-1} \sum_{i=n_1+1}^{n_1+n_2} (x_i - \bar{x}_2)^2. \tag{7.9}$$

The Matalas-Jacobs estimators are (Matalas and Jacobs, 1964; Vogel and Stedinger, 1985):

$$\hat{\mu}_y = \bar{y}_1 + \frac{n_2}{n_1+n_2}\hat{\beta}(\bar{x}_2 - \bar{x}_1) \tag{7.10}$$

and

$$\hat{\sigma}_y^2 = \frac{1}{n_1+n_2+1}\left[(n_1-1){s_{y_1}}^2 + (n_2-1)\hat{\beta}^2{s_{x_2}}^2 + (n_2-1)\alpha^2(1-\hat{\rho}^2){s_{y_1}}^2 + \frac{n_1 n_2}{(n_1+n_2)}\hat{\beta}^2(\bar{x}_2-\bar{x}_1)^2\right] \tag{7.11}$$

where

$$\hat{\rho} = \hat{\beta}\frac{s_{x_1}}{s_{y_1}} \tag{7.12}$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n_1}(x_i-\bar{x}_1)(y_i-\bar{y}_1)}{\sum_{i=1}^{n_1}(x_i-\bar{x}_1)^2} \tag{7.13}$$

and

$$\alpha^2 = \frac{n_2(n_1-4)(n_1-1)}{(n_2-1)(n_1-3)(n_1-2)}. \tag{7.14}$$

Record extension is an appropriate technique when there is substantial improvement to the mean and variance of the short record site, based on the longer concurrent record. This occurs when the variance of the combined record $\hat{\sigma}_2^2$ is less than the variance of the short record s_{y1}^2, which occurs when (Vogel and Stedinger, 1985):

$$\rho^2 > \frac{-B \pm (B^2-4AC)^{1/2}}{2A} \tag{7.15}$$

where

$$A = \frac{(n_2+2)(n_1-6)(n_1-8)}{(n_1-5)} + (n_1-4)\left(\frac{n_1 n_2(n_1-4)}{(n_1-3)(n_1-2)} - \frac{2n_2(n_1-4)}{(n_1-3)} - 4\right) \tag{7.16}$$

$$B = \frac{6(n_2+2)(n_1-6)}{(n_1-5)} + 2(n_1^2-n_1-14) + (n_1-4)\left(\frac{2n_2(n_1-5)}{(n_1-3)} - 2(n_1-3) - \frac{2n_1 n_2(n_1-4)}{(n_1-3)(n_1-2)}\right) \tag{7.17}$$

$$C = 2(n_1+1)\frac{3(n_2+2)}{(n_1-5)} - \frac{(n_1+1)(2n_1+n_2-2)(n_1-3)}{(n_1-1)} + (n_1-4)\left(\frac{2n_2}{(n_1-3)} + 2(n_1+1) + \frac{n_1 n_2(n_1-4)}{(n_1-3)(n_1-2)}\right) \tag{7.18}$$

and thereby assures improvement in both the mean and variance. This is typically the case when $\hat{\rho} > 0.70$.

An example of applying MOVE is shown below.

MOVE Example - Suwanee Creek at Suwanee, Georgia

An example of record extension using MOVE is given for Suwanee Creek at Suwanee, Georgia (station 02334885) where the drainage area is 47.0 square miles. There are 20 years of record at station 02334885 from 1985 to 2004 which is relatively short. The watershed is located in north central Georgia as shown in Figure 7.1. The analysis of the 20 years of record for Suwanee Creek provided low estimates of the flood discharges like the 0.01 exceedance probability flood compared to other long term stations in the region. There is a nearby gaging station on the Etowah River at Canton, Georgia (station 02392000) which has 113 years of record from 1892 through 2004 and a drainage area of 613 square miles. The annual peak data for the Etowah River were used to extend the record for Suwanee Creek from 20 to 113 years to obtain more reasonable flood estimates like the 0.01 exceedance probability flood.

The concurrent annual peak data available at the Etowah River and Suwanee Creek through 2004 are given in Table 7.1. For eight of the 20 years of record, the annual maximum peak flow occurred on the same flood event for Suwanee Creek and the Etowah River. However, for 12 of the 20 years of concurrent record, the annual peak flows corresponded to different flood events. For the purposes of record extension, concurrent flood peaks are those that occurred in the same water year, not on the same flood event.

The annual peak flows for the Etowah River at Canton, Georgia, the long record station, are plotted in Figure 7.2 for 1892 to 2004. As shown in Figure 7.2, there were several large floods that were recorded on the Etowah River prior to 1985 when systematic data collection began at Suwanee Creek. The period of systematic record at Suwanee Creek for 1985 to 2004 does not include several large floods that occurred in 1892, 1916 and 1919 at the nearby Etowah River gaging station. By extending the record at Suwanee Creek, these large floods can

Table 7.1. Summary of concurrent observed annual peak data for the Etowah River and Suwanee Creek from 1985 to 2004.

Water Year	Etowah River Annual Peak Streamflow (cfs)	Suwanee Creek Annual Peak Streamflow (cfs)
1985	5030	1440
1986	3090	386
1987	12200	2150
1988	9340	948
1989	9080	1220
1990	27100	3760
1991	5940	1320
1992	7660	696
1993	10900	2540
1994	9420	1190
1995	10500	2650
1996	19500	4350
1997	11300	2360
1998	15000	2900
1999	5530	816
2000	8900	862
2001	9270	2090
2002	7100	1260
2003	13600	2940
2004	15300	3270

Figure 7.1. Location of Suwanee Creek at Suwanee, Georgia (station 02334885).

be incorporated in the Bulletin 17C frequency analysis, as well as information provided by other events in the 1892-1984 period.

The 20 years of concurrent record for Suwanee Creek and the Etowah River are plotted in Figure 7.3 on a log-log scale. As shown in Figure 7.3, the logarithms of the annual peak flows define a linear relation with a R^2 value of 0.7258. The correlation coefficient is 0.8519 which is higher than the critical value for both the mean and variance suggested by equation 7.15 for 20 concurrent years of record. This suggests that the mean and deviation from the extended record will be improved by use of the longer record. Even though the Etowah River is much larger than Suwanee Creek, there is a strong correlation in annual peak flows that facilitates record extension. The linear relation shown in Figure 7.3 is the ordinary least squares regression line computed using the logarithms of the data.

The 20 years of record at Suwanee Creek from 1985 to 2004 provides low estimates of the flood discharges like the 0.01 annual exceedance probability discharge because major floods that occurred prior to systematic data collection are not considered in the frequency analysis. The period 1985 to 2004 was a relatively dry period as compared to the period prior to 1985 as can be observed from the long term Etowah River record shown in Figure 7.2.

The flood records for Suwanee Creek were extended using MOVE and annual peak flow data for the Etowah River from 1892 to 1984. The analysis is summarized below. The annual peak flows were converted to logarithms for the analysis because, as shown in Figure 7.3, there is a strong linear relation between the logarithms of the annual peak flows.

The extended years of record Y_i for Suwanee Creek were estimated with the MOVE equations (7.1)- (7.14), with

y_i= logarithmic discharge for Suwanee Creek for year i,
x_i= logarithmic discharge for Etowah River for year i,
n_1= concurrent or overlapping period of record, 20 years (1985-2004),
$\bar{x}_1$= logarithmic mean for Etowah River for concurrent period = 3.984 log units,
$\bar{y}_1$= logarithmic mean for Suwanee Creek for concurrent period = 3.215 log units,
s_{y1}= logarithmic standard deviation for Suwanee Creek for concurrent period = 0.279 log units, and
s_{x1}= logarithmic standard deviation for Etowah River for concurrent period = 0.214 log units.

Solving equations (7.1)-(7.14), the MOVE equation in logarithmic linear form is

$$y_i = -1.954 + 1.293x_i \tag{7.19}$$

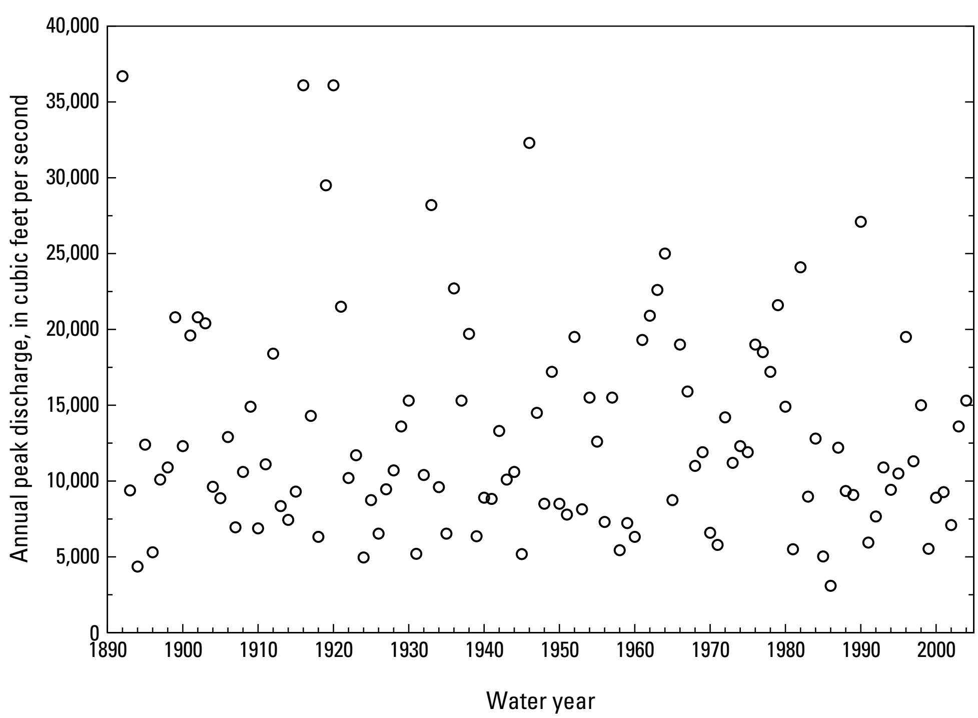

Figure 7.2. Annual peak flows for the Etowah River (station 02392000), the long record station, from 1892 to 2004.

which is as follows in exponential form

$$Q_i = 0.011(Q_{Etowah})^{1.293} \tag{7.20}$$

where Q_i is the extended discharge in cfs for Suwanee Creek and Q_{Etowah} is the discharge in cfs for the Etowah River. Equation 7.20 was used to estimate annual peak flows for Suwanee Creek for the period 1892 to 1984, thereby extending the record an additional 93 years with data from the Etowah River. Extended flow estimates for Suwanee Creek are listed in Table 7.2. Original flow estimates from the long record site (Etowah River) are listed in Table 7.3.

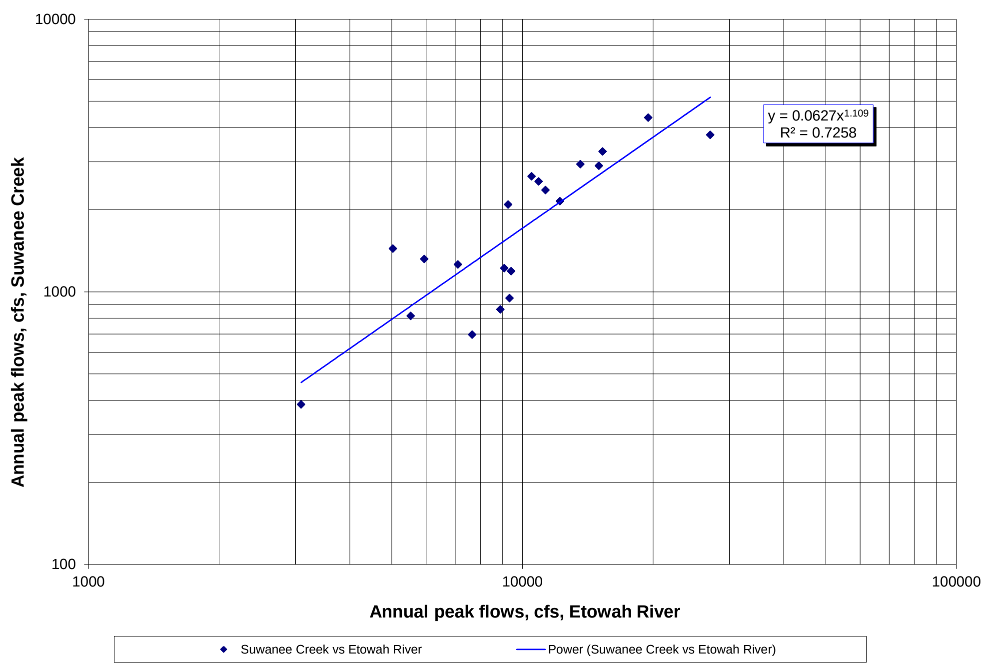

Figure 7.3. Graph of 20 concurrent years of record for Suwanee Creek and the Etowah River for the period 1985 to 2004 with the ordinary least squares regression line.

Table 7.2. MOVE extended record for 93 years (1892-1984) for Suwanee Creek at Suwanee, Georgia (station 02334885).

Water Year	Annual Peak Streamflow (cfs)	Water Year	Annual Peak Streamflow (cfs)	Water Year	Annual Peak Streamflow (cfs)
1892	8890	1923	2030	1954	2915
1893	1520	1924	668	1955	2230
1894	565	1925	1390	1956	1100
1895	2180	1926	953	1957	2915
1896	728	1927	1540	1958	753
1897	1675	1928	1805	1959	1090
1898	1850	1929	2460	1960	914
1899	4260	1930	2870	1961	3870
1900	2160	1931	710	1962	4290
1901	3950	1932	1740	1963	4750
1902	4260	1933	6320	1964	5410
1903	4160	1934	1570	1965	1390
1904	1570	1935	953	1966	3790
1905	1420	1936	4770	1967	3010
1906	2300	1937	2870	1968	1870
1907	1030	1938	3970	1969	2070
1908	1780	1939	921	1970	964
1909	2770	1940	1420	1971	816
1910	1020	1941	1410	1972	2600
1911	1890	1942	2390	1973	1910
1912	3640	1943	1675	1974	2160
1913	1310	1944	1780	1975	2070
1914	1130	1945	706	1976	3790
1915	1505	1946	7530	1977	3660
1916	8700	1947	2670	1978	3335
1917	2630	1948	1340	1979	4480
1918	914	1949	3335	1980	2770
1919	6700	1950	1340	1981	763
1920	8700	1951	1200	1982	5160
1921	4450	1952	3920	1983	1440
1922	1700	1953	1270	1984	2275

Table 7.3. Flood records for 93 years (1892-1984) for the Etowah River at Canton, Georgia (station 02335000).

Water Year	Annual Peak Streamflow (cfs)	Water Year	Annual Peak Streamflow (cfs)	Water Year	Annual Peak Streamflow (cfs)
1892	36700	1923	11700	1954	15500
1893	9380	1924	4960	1955	12600
1894	4360	1925	8740	1956	7300
1895	12400	1926	6530	1957	15500
1896	5300	1927	9460	1958	5440
1897	10100	1928	10700	1959	7230
1898	10900	1929	13600	1960	6320
1899	20800	1930	15300	1961	19300
1900	12300	1931	5200	1962	20900
1901	19600	1932	10400	1963	22600
1902	20800	1933	28200	1964	25000
1903	20400	1934	9600	1965	8740
1904	9620	1935	6530	1966	19000
1905	8870	1936	22700	1967	15900
1906	12900	1937	15300	1968	11000
1907	6950	1938	19700	1969	11900
1908	10600	1939	6360	1970	6590
1909	14900	1940	8900	1971	5790
1910	6880	1941	8820	1972	14200
1911	11100	1942	13300	1973	11200
1912	18400	1943	10100	1974	12300
1913	8350	1944	10600	1975	11900
1914	7440	1945	5180	1976	19000
1915	9300	1946	32300	1977	18500
1916	36100	1947	14500	1978	17200
1917	14300	1948	8500	1979	21600
1918	6320	1949	17200	1980	14900
1919	29500	1950	8500	1981	5500
1920	36100	1951	7790	1982	24100
1921	21500	1952	19500	1983	8970
1922	10200	1953	8140	1984	12800

Appendix 8—Weighting of Independent Estimates

Weighting of Independent Estimates

The uncertainty of peak flow statistics, such as the one-percent annual exceedance probability (*AEP*) flow at a streamgage (or site), can be reduced by combining the at-site estimate with an independent regional estimate to obtain a weighted estimate of the flow statistic at the site. The analysis assumes that the two estimators are independent, unbiased, and that their estimates of the variances are reliable and consistent. A common use of this approach is to combine at-site flood frequency analysis estimates of flood quantiles with flood quantile estimates obtained by regional regression. In that case, methods developed by federal agencies allow computation of weighted estimates using this method. In other cases, independent flood quantile estimates might be available based upon precipitation estimates with rainfall-runoff models. Alternative weighting procedures are evaluated by Griffis and Stedinger (2007a).

Weighting Method

As stated in the Section Flood Distribution, the Pearson Type III distribution with log transformation of the peak flow data should be the base method for the analysis of annual series data. Thus, the peak flow statistic Q_i (such as the 0.01 *AEP*) is transformed using base 10 logarithms:

$$X_i = log_{10}\, Q_i \tag{8.1}$$

where Q_i is the estimated peak flow statistic at site i, and X_i is the log-transformed variable. All subsequent operations are performed on the transformed variable X_i. The weighted estimate is calculated using variances as:

$$X_{weighted,i} = \frac{X_{site,i}V_{reg,i} + X_{reg,i}V_{site,i}}{V_{site,i} + V_{reg,i}} \tag{8.2}$$

where all X and V variables are in log_{10} units, $X_{weighted,i}$ is the weighted estimate for site i, $X_{site,i}$ is the at-site estimate at site i, $X_{reg,i}$ is the regional estimate at site i, $V_{site,i}$ is the variance of the at-site estimate at site i, and $V_{reg,i}$ is the variance of the regional estimate at site i.

As described in Appendix 6, the Expected Moments Algorithm (*EMA*) provides a direct fit of the log-Pearson Type III distribution, which includes an estimate of the variance $V_{site,i}$ corresponding to each computed *AEP*.

For independent $X_{site,i}$ and $X_{reg,i}$, the variance of the weighted estimate for site i is calculated (with all V variables in log_{10} units) as:

$$V_{weighted,i} = \frac{V_{site,i}V_{reg,i}}{V_{site,i} + V_{reg,i}} \tag{8.3}$$

Confidence intervals on the weighted estimated can also be calculated. For example, upper and lower 95% confidence limits on the weighted quantile estimate are calculated as:

$$95\%\_CI_i = \left[10^{\left(X_{weighted,i} - 1.96\sqrt{V_{weighted,i}}\right)}, 10^{\left(X_{weighted,i} + 1.96\sqrt{V_{weighted,i}}\right)}\right] \tag{8.4}$$

and note that $X_{weighted,i}$, $V_{weighted,i}$, and CI_i must be calculated separately for each site i for each *AEP* of interest.

Example

A flood frequency analysis at a basin (site i) using the *EMA* produces an estimate of 861 ft^3/s for the 0.01 *AEP* with a log space variance $V_{site} = 0.0281$. Based on hydrologically similar nearby basins, an independent regional estimate of the 0.01 *AEP* is 718 ft^3/s with a log space variance $V_{reg} = 0.085$. By substituting these values into the above equations, the following weighted estimates are obtained.

Using equation 8.1, the log transformed flow values are computed as:

$$X_{site} = log_{10}(861) = 2.94$$
$$X_{reg} = log_{10}(718) = 2.86.$$

Using equation 8.2, the weighted log transformed flow is computed as:

$$X_{weighted} = \frac{2.94 * 0.085 + 2.86 * 0.028}{0.028 + 0.085} = 2.92$$

and the peak flow $Q_{weighted}$ is

$$Q_{weighted} = 10^{2.92} = 832 \text{ ft}^3\text{/s}.$$

Using equation 8.3, the variance of the weighted log transformed flow is computed as:

$$V_{weighted} = \frac{0.028 * 0.085}{0.028 + 0.085} = 0.021.$$

Using equation 8.4, a 95% confidence interval on the weighted estimate is computed as

$$95\%\_CI_i = \left[10^{(2.92-1.96\sqrt{0.021})}, 10^{(2.92+1.96\sqrt{0.021})}\right]$$
$$= \left[432 \text{ ft}^3\text{/s}, 1560 \text{ ft}^3\text{/s}\right].$$

Appendix 9—Examples

Flood Frequency Examples

Some representative flood frequency examples are presented in this appendix. The main emphasis is on the data, flow intervals, and threshold inputs to *EMA*. The following flood frequency examples illustrate application of the techniques recommended in this guide. Annual flood peak data for seven stations have been selected to illustrate fitting the Log-Pearson Type III distribution when the following are present in a peak flood record at a gage site:

1. Systematic record;
2. Potentially Influential Low Floods (PILFs);
3. Broken record;
4. Historical data;
5. Crest Stage gage censored data;
6. Historic data and PILFs; and
7. Paleoflood data.

The gaging stations and types of data used in each example are listed in Table 9.1. These examples are meant to illustrate the main concepts presented in these *Guidelines*. They are not meant to be all-inclusive. It is important to note that, for the purposes of flood frequency analysis, water years are used in these examples to define the years in which annual peak flows occur. A water year is defined as the 12 month period from October 1 to September 30. The water year is designated by the calendar year in which it ends and which includes 9 of the 12 months. The U.S. Geological Survey *PeakFQ* program (Section Software and Examples) is used for most of the examples; *PeakfqSA* is used for the historical and paleoflood examples. Input and output files from software used to create the examples are also available on the HFAWG web page at http://acwi.gov/hydrology/Frequency/b17c/.

Weighted skew was only used in Example 1: Systematic data; it was not used in Examples 2-7. In order to clearly illustrate how the *EMA* and *MGBT* screening for PILFs are used in flood frequency analysis, only the at-site skew at each station was used.

Table 9.1. Summary of flood frequency examples.

Example No.	Type	USGS Station No.	Station Name	Systematic	Historical	Broken	Censored	PILF
1	Systematic	01134500	Moose River at Victory, VT	X				
2	PILF	11274500	Orestimba Creek near Newman, CA	X				X
3	Broken record	01614000	Back Creek near Jones Springs, WV	X		X		
4	Historical data	07099500	Arkansas River at Pueblo, CO	X	X	X	X	
5	Crest Stage (censored)	05489490	Bear Creek at Ottumwa, IA	X			X	
6	Historic data + PILFs	09480000	Santa Cruz River at Lochiel, AZ	X	X			X
7	Paleoflood data	11446500	American River at Fair Oaks, CA	X	X	X	X	

Systematic Record Example - Moose River at Victory, Vermont

This example illustrates the use of *EMA* and the *MGBT* to perform a flood frequency analysis on a gage site with a record comprised of systematic annual flood peaks.

For this example, USGS gage 01134500 Moose River at Victory, Vermont is used. The Moose River is located in the northeastern part of the state and flows mostly from north to south through very hilly terrain. The Moose River basin is approximately 75 square miles of nearly all forest (Olson, 2014). Historically, it was an important logging area and some logging still continues today. Attempts at farming in the basin have generally failed due to the presence of shallow rocky soil. There are a small number of villages in the basin, but overall it is sparsely populated with only a few miles of paved roadway. There is also a large bog approximately a third of a mile upstream from the gage. The bog is part of the 5,000 acre Victory Basin Wildlife management area. While there is no streamflow regulation in the basin, the bog attenuates peaks in the basin.

Gage 01134500 has an annual peak record consisting of 68 peaks beginning in 1947 and ending in 2014. The annual peaks are listed in Table 9.2 and can be downloaded from USGS NWIS at http://nwis.waterdata.usgs.gov/nwis/peak/?site_no=01134500&agency_cd=USGS&.

EMA Representation of Peak Flow Data for Flood Frequency Analysis

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, when using *EMA* the annual peak flow for every water year during the historical period is described by a flow interval ($Q_{Y,lower}$, $Q_{Y,upper}$) for each water year Y. For peaks whose values are known and are not censored, the flow interval can be described as ($Q_{Y,lower} = Q_Y$, $Q_{Y,upper} = Q_Y$). For example, as shown in Table 9.2, the peak for the 1947 water year is recorded as 2,080 cfs. This peak is known and is not censored, thus the flow interval for the 1947 water year is ($Q_{1947,lower}$ = 2,080, $Q_{1947,upper}$= 2,080). In this example, the flow values are known for all the years where the gage was in operation. Table 9.3 contains the *EMA* flow intervals for each water year in the record for gage 01134500.

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, *EMA* distinguishes among sampling properties by employing perception thresholds denoted ($T_{Y,lower}$, $T_{Y,upper}$) for each year Y, 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 in water year Y. For most peaks at most gages, $T_{Y,upper}$ is assumed to be infinite, as bigger floods that might exceed the measurement capability of the streamgage are determined through study of highwater marks and other physical evidence of the flood. For periods of continuous, full-range peak flow record, the perception threshold is represented by ($T_{Y,lower} = 0$, $T_{Y,upper} = \infty$), where $T_{Y,lower} = 0$ is the gage-base discharge. Table 9.4 contains the *EMA* perception thresholds for each water year in the record for Gage 01134500.

The annual peaks as well as their corresponding *EMA* flow intervals and perception thresholds can be displayed graphically. Figure 9.1 contains a graphical representation of the recorded annual peaks, *EMA* flow intervals and *EMA* perception thresholds. This graph of the data is simple for Gage 01134500, as each year in the record has a recorded peak and the perception threshold for the entire period of record spans from ($T_{Y,lower} = 0$, $T_{Y,upper} = \infty$), thus indicating that all peaks were able to be recorded.

Results from Flood Frequency Analysis

A flood frequency analysis at USGS Gage 01134500 was performed using the *EMA* flow intervals and perception thresholds as shown in Table 9.3 and Table 9.4. The output from an at-site flood frequency analysis

Table 9.2. USGS gage 01134500 Moose River at Victory, VT annual peak flow record consisting of 68 peaks from 1947 to 2014. This table contains the date of the annual peak recorded at the gage, the water year of the annual peak and the corresponding annual peak in cubic feet per second (cfs).

Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)	Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)	Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)
1947-04-13	1947	2080	1970-04-25	1970	3010	1993-04-17	1993	1900
1948-03-28	1948	1670	1971-05-04	1971	1490	1994-04-17	1994	2760
1949-03-28	1949	1480	1972-05-05	1972	2920	1995-08-06	1995	4536
1950-04-21	1950	2940	1973-07-01	1973	4940	1996-04-24	1996	2160
1950-12-05	1951	1560	1973-12-22	1974	2550	1996-12-02	1997	1860
1952-06-02	1952	2380	1975-04-20	1975	1250	1998-03-31	1998	2680
1953-03-27	1953	2720	1976-04-02	1976	2670	1999-09-18	1999	1540
1954-04-23	1954	2860	1977-03-31	1977	2020	2000-05-11	2000	2110
1955-04-15	1955	2620	1978-05-10	1978	1460	2001-04-25	2001	2950
1956-04-30	1956	1710	1979-03-26	1979	1620	2002-04-14	2002	2410
1957-04-22	1957	1370	1980-04-10	1980	1460	2003-03-30	2003	2230
1957-12-21	1958	2180	1981-02-21	1981	1570	2003-10-28	2004	1980
1959-04-04	1959	1160	1982-04-18	1982	2890	2005-04-04	2005	1610
1959-11-29	1960	2780	1983-05-04	1983	1840	2005-10-17	2006	2640
1961-04-24	1961	1580	1984-05-31	1984	2950	2007-04-24	2007	1930
1962-04-08	1962	2110	1985-04-17	1985	1380	2008-04-20	2008	1940
1963-04-22	1963	2160	1986-03-31	1986	2350	2009-04-04	2009	1810
1964-04-15	1964	2750	1987-03-31	1987	4180	2010-03-24	2010	1900
1965-06-14	1965	1190	1988-04-06	1988	1700	2010-10-01	2011	3140
1966-03-26	1966	1560	1989-04-06	1989	2200	2012-03-20	2012	1370
1967-04-03	1967	1800	1990-03-18	1990	3430	2013-04-20	2013	2180
1968-03-24	1968	1600	1990-12-24	1991	2270	2014-04-16	2014	4250
1969-04-29	1969	2400	1992-04-23	1992	2180			

Table 9.3. USGS Gage 01134500 *EMA* flow intervals for the systematic period from 1947 to 2014. This table contains the water year of the annual peak and the corresponding flow interval defined by lower bound, $Q_{Y,lower}$, and upper bound, $Q_{Y,upper}$, in cubic feet per second (cfs) for each water year Y.

Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments	Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments
1947	2080	2080		1981	1570	1570	
1948	1670	1670		1982	2890	2890	
1949	1480	1480		1983	1840	1840	
1950	2940	2940		1984	2950	2950	
1951	1560	1560		1985	1380	1380	
1952	2380	2380		1986	2350	2350	
1953	2720	2720		1987	4180	4180	
1954	2860	2860		1988	1700	1700	
1955	2620	2620		1989	2200	2200	
1956	1710	1710		1990	3430	3430	
1957	1370	1370		1991	2270	2270	
1958	2180	2180		1992	2180	2180	
1959	1160	1160		1993	1900	1900	
1960	2780	2780		1994	2760	2760	
1961	1580	1580		1995	4536	4536	
1962	2110	2110		1996	2160	2160	
1963	2160	2160		1997	1860	1860	
1964	2750	2750		1998	2680	2680	
1965	1190	1190		1999	1540	1540	
1966	1560	1560		2000	2110	2110	
1967	1800	1800		2001	2950	2950	
1968	1600	1600		2002	2410	2410	
1969	2400	2400		2003	2230	2230	
1970	3010	3010		2004	1980	1980	
1971	1490	1490		2005	1610	1610	
1972	2920	2920		2006	2640	2640	
1973	4940	4940		2007	1930	1930	
1974	2550	2550		2008	1940	1940	
1975	1250	1250		2009	1810	1810	
1976	2670	2670		2010	1900	1900	
1977	2020	2020		2011	3140	3140	
1978	1460	1460		2012	1370	1370	
1979	1620	1620		2013	2180	2180	
1980	1460	1460		2014	4250	4250	

Table 9.4. USGS Gage 01134500 Moose River at Victory, VT *EMA* perception thresholds for the systematic period from 1947 to 2014. This table contains the water year ranges to which each perception threshold applies, $T_{Y,lower}$ the lower bound of the perception threshold (in cfs) for water year Y, $T_{Y,upper}$, the upper bound of the perception threshold in cfs for water year Y, and a comment describing the threshold.

		EMA Perception Threshold		
Start Year	**End Year**	$T_{Y,lower}$	$T_{Y,upper}$	**Comments**
1947	2014	0	infinity	continuous systematic record

using *EMA* with the Multiple Grubbs-Beck test to screen for potentially influential low floods (PILFs) is shown below. Note that for the analysis described below weighted skew was used. As described in the Section Estimating Regional Skew, an improved estimate of skew can be computed by weighting the station skew with a regional skew (see the Section Weighted Skew Coefficient Estimator for details). The regional skew used for this station is 0.44 with a corresponding standard error of 0.28 (MSE=0.078) (Olson, 2014). The at-site skew estimate is 0.397 with a MSE=0.10. The estimated peak flow for selected annual exceedance probabilities can be found in Table 9.5, while the fitted frequency curve is displayed in Figure 9.2.

The results of the above analysis were generated using weighted skew. In order to demonstrate the potential impact of the weighted skew on a flood frequency analysis, here we present the results for the same analysis using solely the station skew and compare the results to those previously obtained using the weighted skew. As shown in Figure 9.3, the confidence intervals when using only the station skew are wider for the smaller exceedance probabilities as compared to those in Figure 9.2 when the weighted skew is used. It is important to

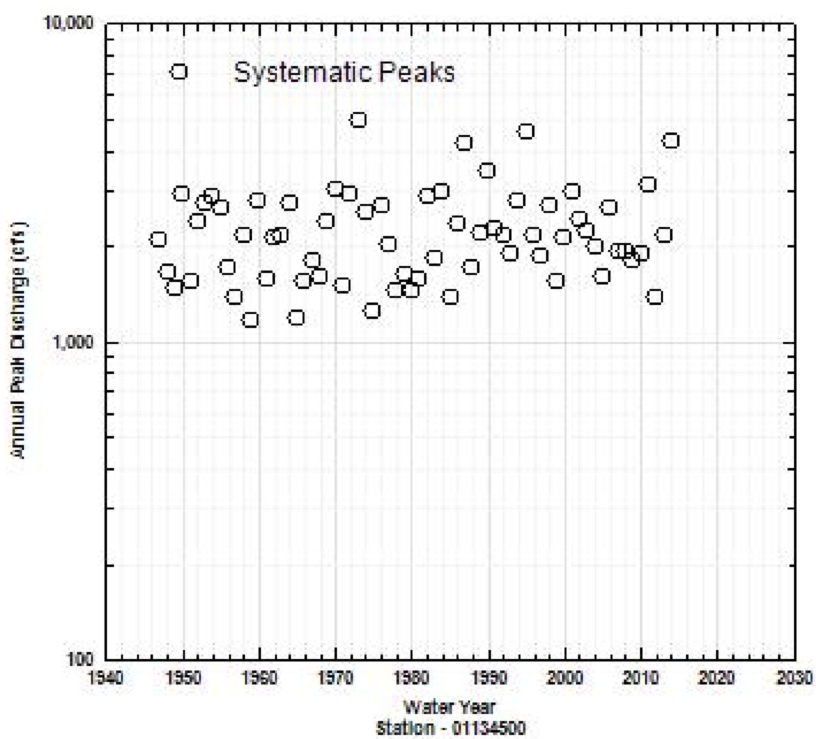

Figure 9.1. USGS gage 01134500 Moose River at Victory, VT annual peak flow time series consisting of 68 peaks from 1947 to 2014. Open circles represent recorded systematic peaks.

Table 9.5. Peak-flow quantiles in cubic feet per second for USGS Gage 01134500 based on flood frequency analysis using *EMA* with *MGBT*; variance of estimate shown in log space.

Annual Exceedance Probability	*EMA* Estimate	Variance of Estimate	Lower 5% Confidence Limit	Upper 95% Confidence Limit
0.1	3262	0.0007	2931	3795
0.04	3920	0.0012	3437	3795
0.02	4440	0.0017	3813	5742
0.01	4985	0.0024	4187	6785
0.005	5560	0.0031	4565	7979
0.002	6374	0.0043	5070	9832

note that this is just one example of the effect of weighted skew on a flood frequency analysis. The impact could be more significant or less significant than shown above depending on the peak flow data at the station, as well as the value of the station's corresponding regional skew and the accuracy of that regional skew.

The fitted frequency curve computed using *EMA* with *MGBT* is displayed in red in Figure 9.2. Because the annual peak flow record contains only systematic peaks with no historic information, no censored peaks and no PILFs identified by the *MGBT*, the fitted frequency curve using these flood frequency *Guidelines* is the same as that from Bulletin 17B.

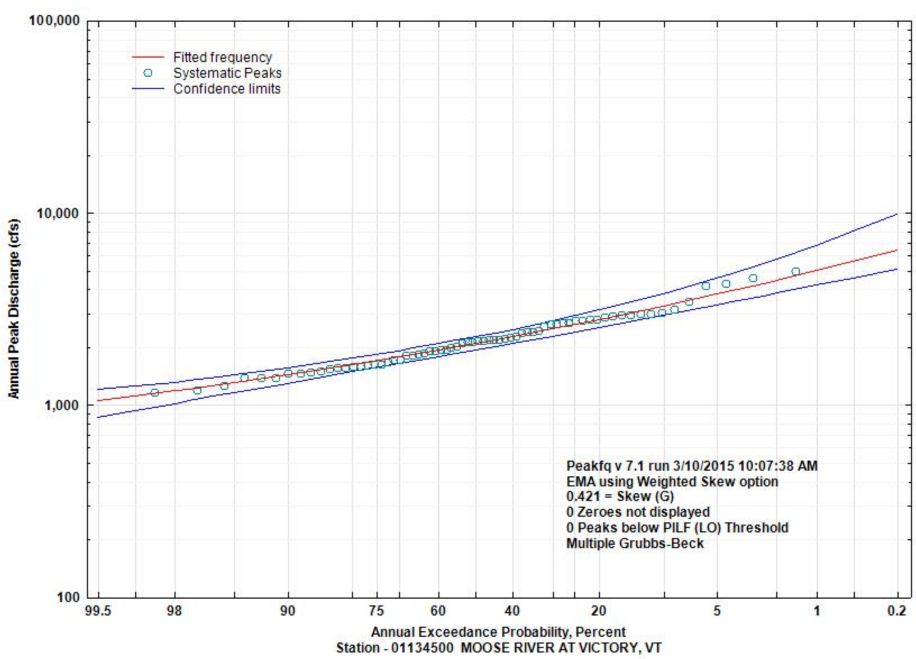

Figure 9.2. Annual Exceedance Probability Plot for USGS Gage 01134500 Moose River at Victory, VT based on flood frequency analysis using *EMA* with *MGBT* and weighted skew. The red line is the fitted log Pearson Type III frequency curve, the blue lines are the upper and lower bounds of the confidence limits, and the green circles are the systematic peaks.

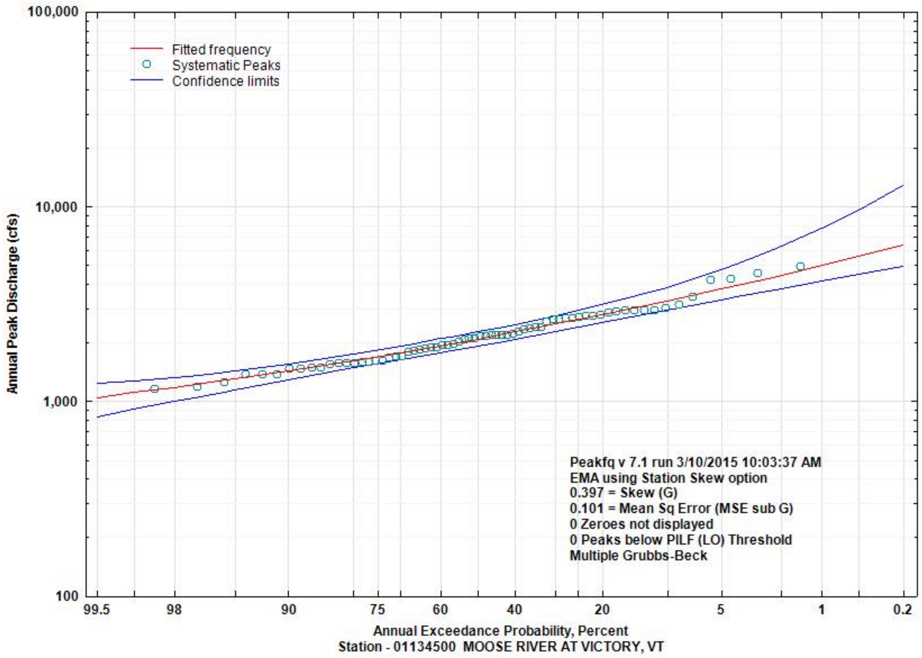

Figure 9.3. Annual Exceedance Probability Plots for USGS Gage 01134500 Moose River at Victory, VT based on flood frequency analysis using *EMA* with *MGBT* and station skew only. The red line is the fitted log Pearson Type III frequency curve, the blue lines are the upper and lower bounds of the confidence limits, and the green circles are the systematic peaks.

PILF Example - Orestimba Creek near Newman, California

This example demonstrates how the Expected Moments Algorithim (*EMA*) and the Multiple Grubbs-Beck test (*MGBT*) can be used to perform a flood frequency analysis on a gage site with a record comprised of systematic annual flood peaks when Potentially Influential Low Floods (PILFs) are present.

For this example, USGS gage 11274500 Orestimba Creek near Newman, CA is used (Parrett et al., 2011; Gotvald et al., 2012). Orestimba Creek is a tributary to the San Joaquin River, whose 134 mi^2 drainage area lies on the eastern slope of the Diablo Range section of the Coast Range Mountains of California (U.S. Army Corps of Engineers, 2008). The drainage basin has an average basin elevation of 1,551 feet with peak flows usually occurring in late winter. Orestimba Creek is one of the few tributaries in the area to maintain a definite stream channel from the foothills to the San Joaquin River (U.S. Army Corps of Engineers, 2008). Some additional details about this gage are in Gotvald et al. (2012).

Gage 11274500 has an annual peak record consisting of 82 peaks beginning in 1932 and ending in 2013. The annual peaks are listed in Table 9.6 (downloaded from USGS NWIS: http://nwis.waterdata.usgs.gov/nwis/peak/?site_no=11274500&agency_cd=USGS&) and shown in Figure 9.4. Of the 82 annual peaks, there are 12 years for which the annual peak is 0 ft^3/s.

EMA Representation of Peak Flow Data for Flood Frequency Analysis

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, when using EMA the annual peak flow for every water year during the historical period is described by a flow interval $(Q_{Y,lower}, Q_{Y,upper})$ for each water year Y. For peaks whose values are known and are not censored, the flow

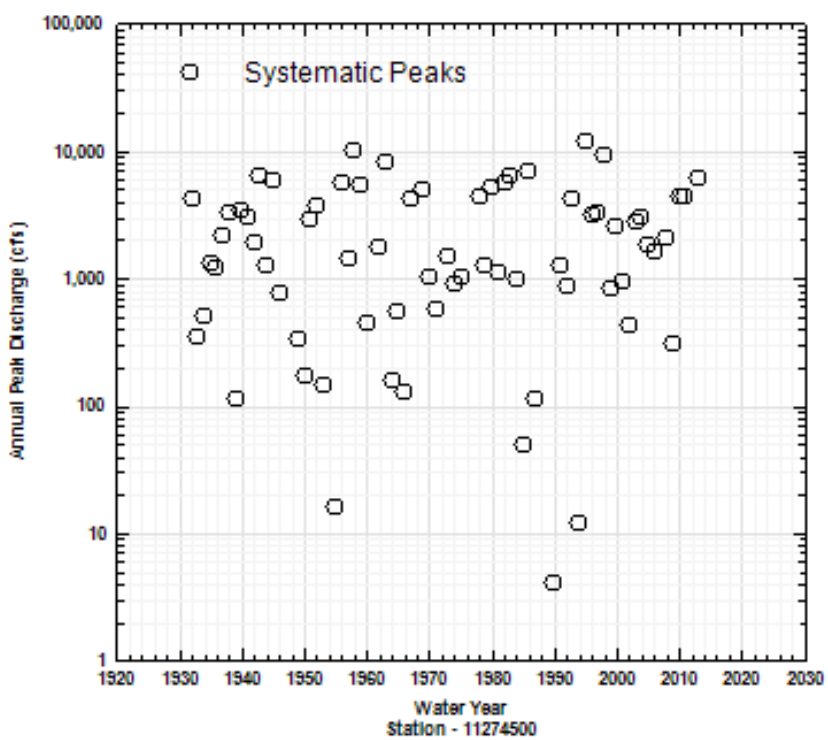

Figure 9.4. USGS gage 11274500 annual peak flow time series consisting of 82 peaks from 1932 to 2013.

Table 9.6. USGS gage 11274500 annual peak flow record consisting of 82 peaks from 1932 to 2013. This table contains the date of the annual peak recorded at the gage, the water year of the annual peak and the corresponding annual peak in cubic feet per second (cfs).

Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)	Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)	Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)
1932-02-08	1932	4260	1960-02-10	1960	448	1988-00-00	1988	0
1933-01-29	1933	345	1961-00-00	1961	0	1989-00-00	1989	0
1934-01-01	1934	516	1962-02-15	1962	1740	1990-05-28	1990	4
1935-04-08	1935	1320	1963-02-01	1963	8300	1991-03-24	1991	1260
1936-02-13	1936	1200	1964-01-22	1964	156	1992-02-15	1992	888
1937-02-13	1937	2180	1965-01-06	1965	560	1993-01-13	1993	4190
1938-02-11	1938	3230	1965-12-30	1966	128	1994-02-20	1994	12
1939-03-09	1939	115	1967-01-24	1967	4200	1995-03-10	1995	12000
1940-02-27	1940	3440	1968-00-00	1968	0	1996-02-19	1996	3130
1941-04-04	1941	3070	1969-01-25	1969	5080	1997-01-23	1997	3320
1942-01-24	1942	1880	1970-03-01	1970	1010	1998-02-03	1998	9470
1943-01-21	1943	6450	1970-12-21	1971	584	1999-02-09	1999	833
1944-02-29	1944	1290	1972-00-00	1972	0	2000-02-14	2000	2550
1945-02-02	1945	5970	1973-02-11	1973	1510	2001-03-05	2001	958
1945-12-25	1946	782	1974-03-03	1974	922	2002-01-03	2002	425
1947-00-00	1947	0	1975-03-08	1975	1010	2002-12-16	2003	2790
1948-00-00	1948	0	1976-00-00	1976	0	2004-02-25	2004	2990
1949-03-12	1949	335	1977-00-00	1977	0	2005-02-16	2005	1820
1950-02-05	1950	175	1978-01-17	1978	4360	2006-01-02	2006	1630
1950-12-03	1951	2920	1979-02-21	1979	1270	2007-00-00	2007	0
1952-01-12	1952	3660	1980-02-16	1980	5210	2008-01-25	2008	2110
1952-12-07	1953	147	1981-01-29	1981	1130	2009-02-17	2009	310
1954-00-00	1954	0	1982-01-05	1982	5550	2010-01-20	2010	4400
1955-01-19	1955	16	1983-01-24	1983	6360	2011-03-24	2011	4440
1955-12-23	1956	5620	1983-12-25	1984	991	2012-00-00	2012	0
1957-02-24	1957	1440	1985-02-09	1985	50	2012-12-24	2013	6250
1958-04-02	1958	10200	1986-02-19	1986	6990			
1959-02-16	1959	5380	1987-03-06	1987	112			

interval can be described as ($Q_{Y,lower} = Q_Y$, $Q_{Y,upper} = Q_Y$). For example, as shown in Table 9.6, the peak for the 1932 water year is recorded as 4260 ft^3/s. This peak is known and is not censored, thus the flow interval for the 1932 water year is ($Q_{1932,lower} = 4260$, $Q_{1932,upper} = 4260$). Table 9.7 contains the EMA flow intervals for each water year in the record for gage 11274500.

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, EMA distinguishes among sampling properties by employing perception thresholds denoted ($T_{Y,lower}$, $T_{Y,upper}$) for each year Y, 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 in water year Y. For most peaks at most gages, $T_{Y,upper}$ is assumed to be infinite, as bigger floods that might exceed the measurement capability of the streamgage are determined through study of highwater marks and other physical evidence of the flood. For periods of continuous, full-range peak flow record, the perception threshold is represented by ($T_{Y,lower} = 0$, $T_{Y,upper} = \infty$), where $T_{Y,lower} = 0$ is the gage-base discharge. Table 9.8 contains the EMA perception thresholds for each water year in the record for Gage 11274500.

Results from Flood Frequency Analysis

A flood frequency analysis at USGS Gage 11274500 was performed using the EMA flow intervals and perception thresholds as shown in Table 9.7 and Table 9.8. The output from an at-site flood frequency analysis using EMA with the Multiple Grubbs-Beck test to screen for potentially influential low floods (PILFs) is shown below. Note that station skew was used, thus allowing the focus to be on the at-site data. The fitted frequency curve is displayed in Figure 9.5 with estimates provided in Table 9.9.

As shown in Figure 9.5, the PILF threshold T_{PILF} established by the MGBT is 782 ft^3/s. Thus, all 30 annual peaks less than 782 ft^3/s are censored and re-coded in the framework of EMA with flow intervals of ($Q_{Y,lower} = 0$, $Q_{Y,upper} = 782$). The MGBT threshold also has the effect of adjusting the lower bound of the perception threshold. Thus for the entire historical period from 1932 to 2013, the perception threshold based on T_{PILF} is ($T_{Y,lower} = 782$, $T_{Y,upper} = \infty$). As shown in Figure 9.5, by censoring the 30 smallest peaks in the record, the smallest annual exceedance probability peaks are well fit by the frequency curve (red line).

Table 9.7. USGS Gage 11274500 EMA flow intervals for the systematic period from 1932 to 2013. This table contains the water year of the annual peak and the corresponding flow interval defined by lower bound, $Q_{Y,lower}$, and upper bound, $Q_{Y,upper}$, in cubic feet per second (cfs) for each water year Y.

Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments	Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments
1932	4260	4260		1973	1510	1510	
1933	345	345		1974	922	922	
1934	516	516		1975	1010	1010	
1935	1320	1320		1976	0	0	zero flow
1936	1200	1200		1977	0	0	zero flow
1937	2180	2180		1978	4360	4360	
1938	3230	3230		1979	1270	1270	
1939	115	115		1980	5210	5210	
1940	3440	3440		1981	1130	1130	
1941	3070	3070		1982	5550	5550	
1942	1880	1880		1983	6360	6360	
1943	6450	6450		1984	991	991	
1944	1290	1290		1985	50	50	
1945	5970	5970		1986	6990	6990	
1946	782	782		1987	112	112	
1947	0	0	zero flow	1988	0	0	zero flow
1948	0	0	zero flow	1989	0	0	zero flow
1949	335	335		1990	4	4	
1950	175	175		1991	1260	1260	
1951	2920	2920		1992	888	888	
1952	3660	3660		1993	4190	4190	
1953	147	147		1994	12	12	
1954	0	0	zero flow	1995	12000	12000	
1955	16	16		1996	3130	3130	
1956	5620	5620		1997	3320	3320	
1957	1440	1440		1998	9470	9470	
1958	10200	10200		1999	833	833	
1959	5380	5380		2000	2550	2550	
1960	448	448		2001	958	958	
1961	0	0	zero flow	2002	425	425	
1962	1740	1740		2003	2790	2790	
1963	8300	8300		2004	2990	2990	
1964	156	156		2005	1820	1820	
1965	560	560		2006	1630	1630	
1966	128	128		2007	0	0	zero flow
1967	4200	4200		2008	2110	2110	
1968	0	0	zero flow	2009	310	310	
1969	5080	5080		2010	4400	4400	
1970	1010	1010		2011	4440	4440	
1971	584	584		2012	0	0	zero flow
1972	0	0	zero flow	2013	6250	6250	

Table 9.8. USGS Gage 11274500 EMA perception thresholds for the systematic period from 1932 to 2013. This table contains the water year ranges to which each perception threshold applies, $T_{Y,lower}$ the lower bound of the perception threshold (in cfs) for water year Y, $T_{Y,upper}$, the upper bound of the perception threshold in cfs for water year Y, and a comment describing the threshold.

		EMA Perception Threshold		
Start Year	End Year	$T_{Y,lower}$	$T_{Y,upper}$	Comments
1932	2003	0	infinity	continuous systematic record

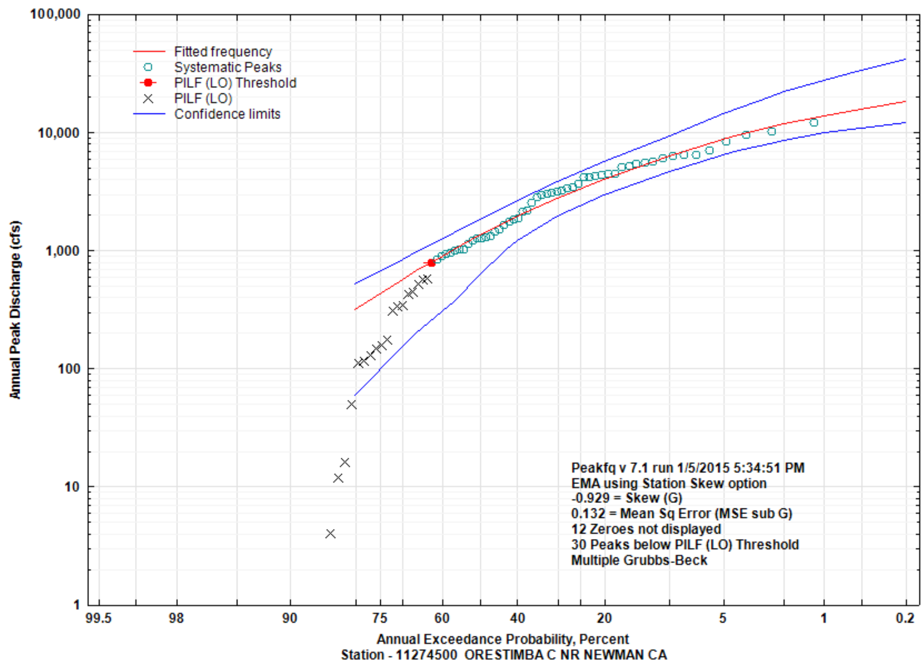

Figure 9.5. Annual Exceedance Probability Plot for USGS Gage 11274500 based on flood frequency analysis using EMA with MGBT. The red line is the fitted log Pearson Type III frequency curve, the blue lines are the upper and lower bounds of the confidence limits, the green circles are the systematic peaks, the solid red circle with a line through it is the potentially influential low floods (PILFs) thresholds as identified by the MGBT, and the black x's are the PILFs identified by the MGBT.

Table 9.9. Peak-flow quantiles in cubic feet per second for USGS Gage 11274500 based on flood frequency analysis using EMA with MGBT; variance of estimate shown in log space.

Annual Exceedance Probability	EMA Estimate	Variance of Estimate	Lower 5% Confidence Limit	Upper 95% Confidence Limit
0.5	1339	0.0078	620.6	1840
0.2	4026	0.0045	2965	5595
0.1	6328	0.0049	4686	9394
0.04	9426	0.0061	6944	16110
0.02	11690	0.0073	8489	21920
0.01	13820	0.0088	9813	27730
0.005	15800	0.0106	10910	33500
0.002	18150	0.0135	12040	41610

Broken Record Example - Back Creek near Jones Springs, WV

This example illustrates the use of *EMA* for a broken record, as described in the Section Broken, Incomplete and Discontinued Records. For this example, USGS gage 01614000 Back Creek near Jones Springs, West Virginia is used. Back Creek is a tributary to the Potomac River; the 235 square mile watershed lies within the Valley and Ridge province in West Virginia (Wiley and Atkins, 2010).

Gage 01614000 has an annual peak record consisting of 56 peaks beginning in 1929 and ending in 2012. There are three "broken record" periods where the gage was discontinued: 1932-1937, 1976-1991, and 1999-2003. Thus, there are 28 years of missing data at this gage during the period 1929-2012. There is a historic flood that occurred outside the period of gaging record on March 17, 1936. This flood is noted in the USGS Annual Water Data Report for this gage, available in the peak-flow file, and there is historical information available for this large flood (Grover, 1937). The annual peaks are listed in Table 9.10 and shown in Figure 9.6. Of the 56 annual peaks, the October 1942 flood slightly exceeds the March 1936 historic flood peak. Based on the historical flood information in Grover (1937) for the 1936 flood, and the large regional floods and historical floods described by Wiley and Atkins (2010) in West Virginia for the period 1888-1996, information from the March 1936 flood is used as a perception threshold to represent the 28 years of missing information.

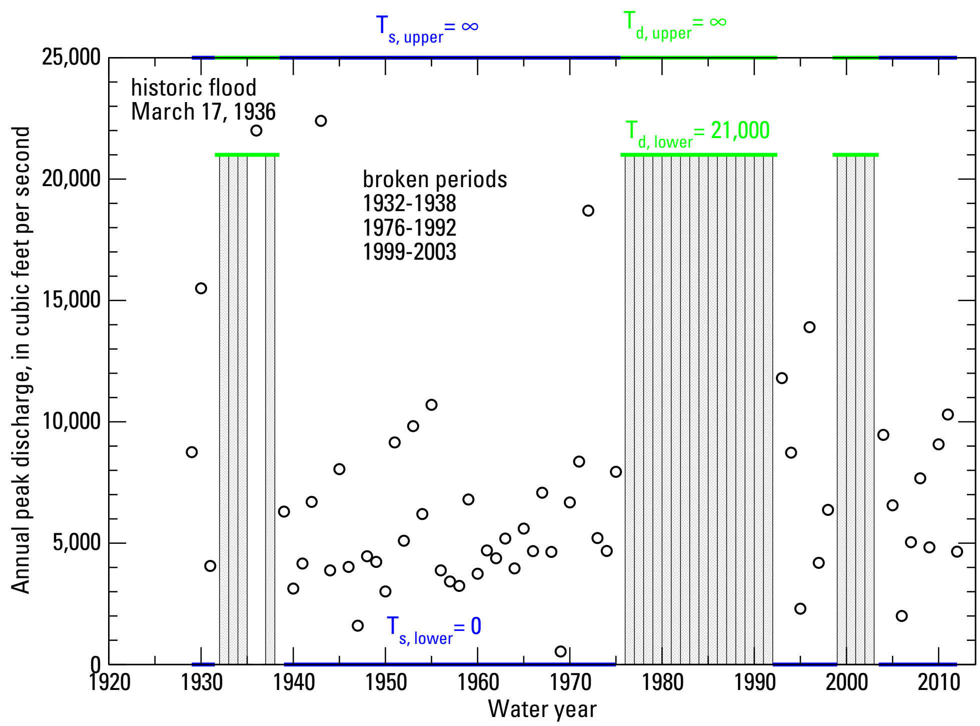

Figure 9.6. USGS gage 01614000 annual peak flow time series consisting of 56 peaks from 1929 to 2012. Flood intervals are shown as black vertical bars with caps that represent lower and upper flow estimates. The grey shaded areas represents floods of unknown magnitude less than the perception threshold $T_{d,lower}$ during the broken record periods. The green lines represent the range in which floods would have been measured or recorded for the broken record periods 1932-1938, 1976-1992, and 1999-2003, with lower and upper perception thresholds $T_{d,lower}$ (21,000 cfs) and $T_{d,upper}$ estimated from the March 1936 historic flood. The perceptible range for the systematic (gage) periods $T_{s,lower}, T_{s,upper}$ $(0, \infty)$ is shown as blue lines.

EMA Representation of Peak Flow Data for Flood Frequency Analysis

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, when using EMA the annual peak flow for every water year during the historical period is described by a flow interval ($Q_{Y,lower}$, $Q_{Y,upper}$) for each water year Y. For peaks whose values are known and are not censored, the flow interval can be described as ($Q_{Y,lower} = Q_Y$, $Q_{Y,upper} = Q_Y$). In this example, the flow values are known for all the years where the gage was in operation. Table 9.11 contains the EMA flow intervals for each water year in the record for gage 01614000. Missing years are described by perception thresholds.

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, EMA distinguishes among sampling properties by employing perception thresholds denoted ($T_{Y,lower}$, $T_{Y,upper}$) for each year Y, 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 in water year Y. For most peaks at most gages, $T_{Y,upper}$ is assumed to be infinite, as bigger floods that might exceed the measurement capability of the streamgage are determined through study of highwater marks and other physical evidence of the flood. For periods of continuous, full-range peak flow record, the perception threshold is represented by ($T_{Y,lower} = 0$, $T_{Y,upper} = \infty$), where $T_{Y,lower} = 0$ is the gage-base discharge. In this example, there are missing years that are described by the 1936 historical flood magnitude. Based on the March 1936 large historical flood (Grover, 1937) and the regional historical flood information available for the largest floods in West Virginia (Wiley and Atkins, 2010), it is known that floods at this location would have been estimated (or recorded), had they exceeded approximately 21,000 cfs. Table 9.12 contains the EMA perception thresholds for each water year in the record, including missing periods, for Gage 01614000.

Results from Flood Frequency Analysis

A flood frequency analysis at USGS Gage 01614000 was performed using the EMA flow intervals and perception thresholds as shown in Table 9.11 and Table 9.12. The output from an at-site flood frequency analysis using EMA with the Multiple Grubbs-Beck test to screen for potentially influential low floods (PILFs) is shown below. Note that station skew was used, thus allowing the focus to be on the at-site data. The fitted frequency curve is displayed in Figure 9.7 with estimates provided in Table 9.13.

As shown in Figure 9.7, there are two floods that exceed the historical threshold (21,000 cfs): the March 1936 flood and the October 1942 flood. Using MGBT, one PILF was identified, with a threshold equal to 1,600 ft^3/s. One annual peak less than 1,600 ft^3/s (equal to 536 cfs) is censored and re-coded in the framework of EMA with flow intervals of ($Q_{Y,lower} = 0$, $Q_{Y,upper} = 1600$). The MGBT threshold also has the effect of adjusting the lower bound of the perception threshold. Thus for the entire historical period from 1929 to 2012, with the exception of the missing years, the perception threshold is ($T_{Y,lower} = 1600$, $T_{Y,upper} = \infty$). For the broken-record years covered by historical information, the lower threshold $T_{Y,lower} = 21000$ (Table 9.12). As shown in Figure 9.7, by censoring the one smallest peak in the record, the remaining smallest annual exceedance probability peaks and the largest floods are well fit by the frequency curve (red line).

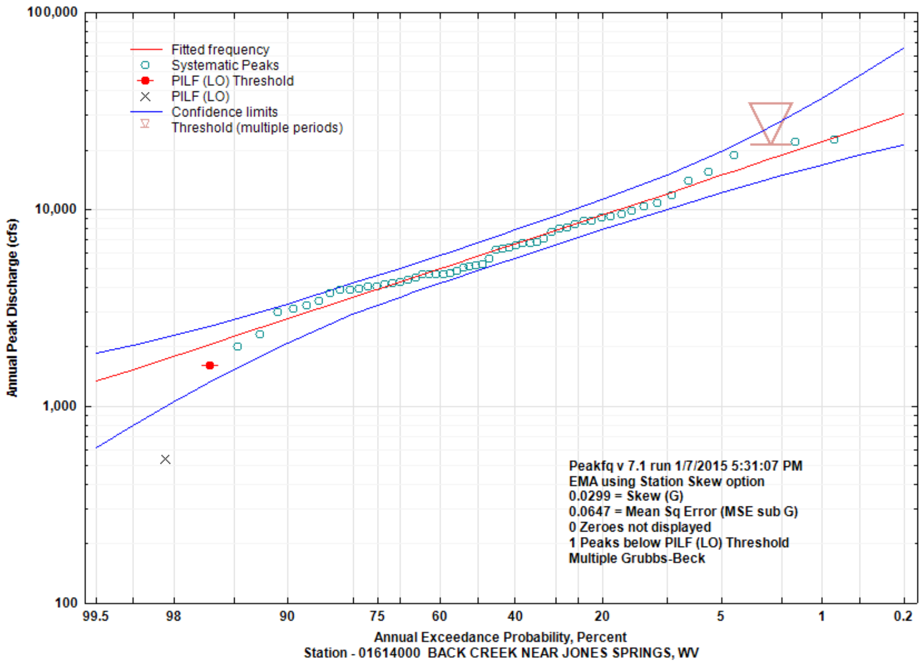

Figure 9.7. Annual Exceedance Probability Plot for USGS Gage 01614000 based on flood frequency analysis using EMA with MGBT. The red line is the fitted log Pearson Type III frequency curve, the blue lines are the upper and lower bounds of the confidence limits, the green circles are the systematic peaks, the solid red circle with a line through it is the potentially influential low floods (PILFs) thresholds as identified by the MGBT, and the black x's are the PILFs identified by the MGBT. The red triangle with the horizontal line represents the lower limit of the historical perception threshold (21,000 cfs).

Table 9.10. USGS gage 01614000 annual peak flow record consisting of 56 peaks from 1929 to 2012, including the 1936 historical flood. This table contains the date of the annual peak recorded at the gage, the water year of the annual peak and the corresponding annual peak in cubic feet per second (cfs). Horizontal lines indicate broken-record years.

Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)	Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)	Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)
1929-04-17	1929	8750	1954-03-02	1954	6200	1972-12-09	1973	5210
1929-10-23	1930	15500	1955-08-19	1955	10700	1973-12-27	1974	4680
1931-05-08	1931	4060	1956-03-15	1956	3880	1975-03-20	1975	7940
1936-03-17	1936	22000	1957-02-10	1957	3420	1993-03-05	1993	11800
1939-02-04	1939	6300	1958-03-27	1958	3240	1994-05-08	1994	8730
1940-04-20	1940	3130	1959-06-03	1959	6800	1995-01-16	1995	2300
1941-04-06	1941	4160	1960-05-09	1960	3740	1996-01-19	1996	13900
1942-05-22	1942	6700	1961-02-19	1961	4700	1996-11-09	1997	4190
1942-10-15	1943	22400	1962-03-22	1962	4380	1998-03-21	1998	6370
1944-03-24	1944	3880	1963-03-20	1963	5190	2004-09-29	2004	9460
1945-09-18	1945	8050	1964-01-10	1964	3960	2005-03-29	2005	6560
1946-06-03	1946	4020	1965-03-06	1965	5600	2005-11-30	2006	2000
1947-03-15	1947	1600	1966-09-21	1966	4670	2007-04-16	2007	5040
1948-04-14	1948	4460	1967-03-08	1967	7080	2008-04-21	2008	7670
1948-12-31	1949	4230	1968-03-17	1968	4640	2009-05-05	2009	4830
1950-02-02	1950	3010	1969-02-02	1969	536	2010-03-14	2010	9070
1950-12-05	1951	9150	1970-07-10	1970	6680	2011-04-17	2011	10300
1952-04-28	1952	5100	1970-11-13	1971	8360	2012-03-01	2012	4650
1952-11-22	1953	9820	1972-06-22	1972	18700			

Table 9.11. USGS Gage 01614000 EMA flow intervals for the systematic period from 1929 to 2012. This table contains the water year of the annual peak and the corresponding flow interval defined by lower bound, $Q_{Y,lower}$, and upper bound, $Q_{Y,upper}$, in cubic feet per second (cfs) for each water year Y. Horizontal lines indicate broken-record years.

Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments	Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments
1929	8750	8750		1963	5190	5190	
1930	15500	15500		1964	3960	3960	
1931	4060	4060		1965	5600	5600	
1936	22000	22000	historic flood	1966	4670	4670	
1939	6300	6300		1967	7080	7080	
1940	3130	3130		1968	4640	4640	
1941	4160	4160		1969	536	536	
1942	6700	6700		1970	6680	6680	
1943	22400	22400		1971	8360	8360	
1944	3880	3880		1972	18700	18700	
1945	8050	8050		1973	5210	5210	
1946	4020	4020		1974	4680	4680	
1947	1600	1600		1975	7940	7940	
1948	4460	4460		1993	11800	11800	
1949	4230	4230		1994	8730	8730	
1950	3010	3010		1995	2300	2300	
1951	9150	9150		1996	13900	13900	
1952	5100	5100		1997	4190	4190	
1953	9820	9820		1998	6370	6370	
1954	6200	6200		2004	9460	9460	
1955	10700	10700		2005	6560	6560	
1956	3880	3880		2006	2000	2000	
1957	3420	3420		2007	5040	5040	
1958	3240	3240		2008	7670	7670	
1959	6800	6800		2009	4830	4830	
1960	3740	3740		2010	9070	9070	
1961	4700	4700		2011	10300	10300	
1962	4380	4380		2012	4650	4650	

Table 9.12. USGS Gage 01614000 EMA perception thresholds for the systematic period from 1929 to 2012. This table contains the water year ranges to which each perception threshold applies, $T_{Y,lower}$ the lower bound of the perception threshold (in cfs) for water year Y, $T_{Y,upper}$, the upper bound of the perception threshold in cfs for water year Y, and a comment describing the threshold.

		EMA Perception Threshold		
Start Year	End Year	$T_{Y,lower}$	$T_{Y,upper}$	Comments
1929	1931	0	infinity	continuous systematic record
1932	1938	21000	infinity	missing record with historical information
1939	1975	0	infinity	continuous systematic record
1976	1992	21000	infinity	missing record with historical information
1993	1998	0	infinity	continuous systematic record
1999	2003	21000	infinity	missing record with historical information
2004	2012	0	infinity	continuous systematic record

Table 9.13. Peak-flow quantiles in cubic feet per second for USGS Gage 01614000 based on flood frequency analysis using EMA with MGBT.

Annual Exceedance Probability	EMA Estimate	Lower 5% Confidence Limit	Upper 95% Confidence Limit
0.500	5714	4845	6730
0.200	9272	7839	11100
0.100	11960	9972	14840
0.040	15710	12730	21360
0.020	18750	14770	27990
0.010	22000	16760	36440
0.005	25470	18690	47130
0.002	30430	21150	65650

Historical Record Example - Arkansas River at Pueblo, CO

This example illustrates the use of *EMA* for a historical record with several large floods (described in the Section Historical Flood Information) and paleoflood information. The Arkansas River at Pueblo Dam near Pueblo, Colorado is presented to illustrate the use of *EMA* with extensive historical information, paleoflood information, and the ability to place the record June 1921 flood in longer time context. The largest historic floods are described as interval data, and multiple thresholds are needed to effectively extend the discontinued streamgaging record after the dam was built. Paleoflood data are also included for this Reclamation dam safety application. Details of this example are presented in England et al. (2006) and England et al. (2010). Peak discharge probability estimates were made at four paleoflood sites on the Arkansas River at Pueblo State Park, Parkdale, at Loma Linda and at Adobe Park. We focus on the Pueblo State Park site flood frequency in this example; flood frequency results for other locations as well as regional frequency results are presented in England et al. (2006).

For this example, peak discharge estimates on the Arkansas River at Pueblo State Park are combined from USGS gaging stations at Portland (07097000) (years 1975-1976), near Portland (07099200) (1974), and near Pueblo (07099500) (years 1864-1973), and are used with Pueblo reservoir records (years 1977-2004) in order to gain a complete record of all large floods that exceeded approximately 10,000 ft^3/s for the period of record. These gaging stations were previously analyzed by England et al. (2006); see also England et al. (2010).

The annual peaks are listed in Table 9.14 and shown in Figure 9.8. Of the 85 annual peaks, including historical information, the June 3, 1921 peak (Follansbee and Jones, 1922; Munn and Savage, 1922) is the largest. The total combined gage record length, excluding historical data, is 110 years (1895-2004) (Figure 9.8). The largest peak discharge estimates from these gages were unaffected by upstream regulation. Reviews of available historical information (Follansbee and Jones, 1922; Munn and Savage, 1922; Follansbee and Sawyer, 1948) indicated there was historical flood information at the site for frequency analysis. The historical record was estimated to begin in 1859, resulting in a 146-year period (1859-2004). Three historical floods were included: June 1864, July 1893, and May 1894. The magnitudes of these floods were large relative to the floods in the gaging record; estimates within a range were based on Follansbee and Sawyer (1948) and included in the flood frequency analysis. These estimates have relatively large uncertainties as compared to the smaller floods in the gage record. A paleohydrologic bound of about 840 years (before water year 2004) was estimated at this site for inclusion in the flood frequency curve. The estimate is based on three soils pits, two radiocarbon ages, and hydraulic modeling of a 7,500 foot reach (England et al., 2006). No estimates of individual paleofloods were made at this site, due to the relatively wide channel geometry and the lack of apparent stratigraphic evidence of large paleofloods during the limited field study (England et al., 2010). Peak discharge, historical flood and nonexceedance bound data synthesis for flood frequency shows that these historical floods are the largest in the record, and combined with the paleoflood data result in a substantially longer time series (Figure 9.8).

EMA Representation of Peak Flow Data for Flood Frequency Analysis

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, when using EMA the annual peak flow for every water year during the historical period is described by a flow interval $(Q_{Y,lower}, Q_{Y,upper})$ for each water year Y. For peaks whose values are known and are not censored, the flow interval can be described as $(Q_{Y,lower} = Q_Y, Q_{Y,upper} = Q_Y)$. In this example, the flow values are known for all the years where the gage was in operation. Table 9.15 contains the EMA flow intervals for each water year in the record for gage 07099500. The historical period is described by a perception threshold, as is the period after the gage was discontinued (1977-2004).

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, EMA

Table 9.14. USGS gage 07099500 (and others) Arkansas River annual peak flow record consisting of 85 peaks from 1864 to 1976. This table contains the water year of the annual peak and the corresponding annual peak in cubic feet per second (cfs).

Water Year	Annual Peak Streamflow (cfs)	Water Year	Annual Peak Streamflow (cfs)	Water Year	Annual Peak Streamflow (cfs)
1864	>41000	1921	>80000	1950	8700
1893	>20000	1922	8850	1951	9300
1894	>35000	1923	25600	1952	4740
1895	6100	1924	6510	1953	6770
1896	16500	1925	4930	1954	10200
1897	4300	1926	4520	1955	11100
1898	7500	1927	12400	1956	8010
1899	8800	1928	7800	1957	9070
1900	7600	1929	10500	1958	4540
1901	11100	1930	6050	1959	2820
1902	30000	1931	3560	1960	5260
1903	10500	1932	4380	1961	5760
1904	8500	1933	8630	1962	3540
1905	8000	1934	2580	1963	8360
1906	11000	1935	9880	1964	2840
1907	6600	1936	11200	1965	23500
1908	7600	1937	9300	1966	10600
1909	5800	1938	11200	1967	5870
1910	8400	1939	2910	1968	5190
1911	3700	1940	3860	1969	6620
1912	10500	1941	7560	1970	6300
1913	7800	1942	10300	1971	3360
1914	7500	1943	3320	1972	3360
1915	17000	1944	5980	1973	6760
1916	8900	1945	9290	1974	5440
1917	6800	1946	7050	1975	10200
1918	9600	1947	7280	1976	12800
1919	6300	1948	10900		
1920	8500	1949	12800		

Table 9.15. Arkansas River at Pueblo EMA flow intervals for the period from 1864 to 1976.

Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments	Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments
1864	41000	60000	historical flood	1935	9880	9880	
1893	20000	25000	historical flood	1936	11200	11200	
1894	35000	40000	historical flood	1937	9300	9300	
1895	6100	6100		1938	11200	11200	
1896	16500	16500		1939	2910	2910	
1897	4300	4300		1940	3860	3860	
1898	7500	7500		1941	7560	7560	
1899	8800	8800		1942	10300	10300	
1900	7600	7600		1943	3320	3320	
1901	11100	11100		1944	5980	5980	
1902	30000	30000		1945	9290	9290	
1903	10500	10500		1946	7050	7050	
1904	8500	8500		1947	7280	7280	
1905	8000	8000		1948	10900	10900	
1906	11000	11000		1949	12800	12800	
1907	6600	6600		1950	8700	8700	
1908	7600	7600		1951	9300	9300	
1909	5800	5800		1952	4740	4740	
1910	8400	8400		1953	6770	6770	
1911	3700	3700		1954	10200	10200	
1912	10500	10500		1955	11100	11100	
1913	7800	7800		1956	8010	8010	
1914	7500	7500		1957	9070	9070	
1915	17000	17000		1958	4540	4540	
1916	8900	8900		1959	2820	2820	
1917	6800	6800		1960	5260	5260	
1918	9600	9600		1961	5760	5760	
1919	6300	6300		1962	3540	3540	
1920	8500	8500		1963	8360	8360	
1921	80000	103000	historical flood	1964	2840	2840	
1922	8850	8850		1965	23500	23500	
1923	25600	25600		1966	10600	10600	
1924	6510	6510		1967	5870	5870	
1925	4930	4930		1968	5190	5190	
1926	4520	4520		1969	6620	6620	
1927	12400	12400		1970	6300	6300	
1928	7800	7800		1971	3360	3360	
1929	10500	10500		1972	3360	3360	
1930	6050	6050		1973	6760	6760	
1931	3560	3560		1974	5440	5440	
1932	4380	4380		1975	10200	10200	
1933	8630	8630		1976	12800	12800	
1934	2580	2580					

distinguishes among sampling properties by employing perception thresholds denoted ($T_{Y,lower}$, $T_{Y,upper}$) for each year Y, 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 in water year Y. For most peaks at most gages, $T_{Y,upper}$, is assumed to be infinite, as bigger floods that might exceed the measurement capability of the streamgage are determined through study of highwater marks and other physical evidence of the flood. For periods of continuous, full-range peak flow record, the perception threshold is represented by ($T_{Y,lower} = 0$, $T_{Y,upper} = \infty$), where $T_{Y,lower} = 0$ is the gage-base discharge. Based on the historical floods and reservoir records, it is known that floods at this location would have been estimated (or recorded), had they exceeded approximately 20,000 cfs. Table 9.16 contains the EMA perception thresholds for each water year in the record, including the historical and paleoflood period, for Gage 07099500.

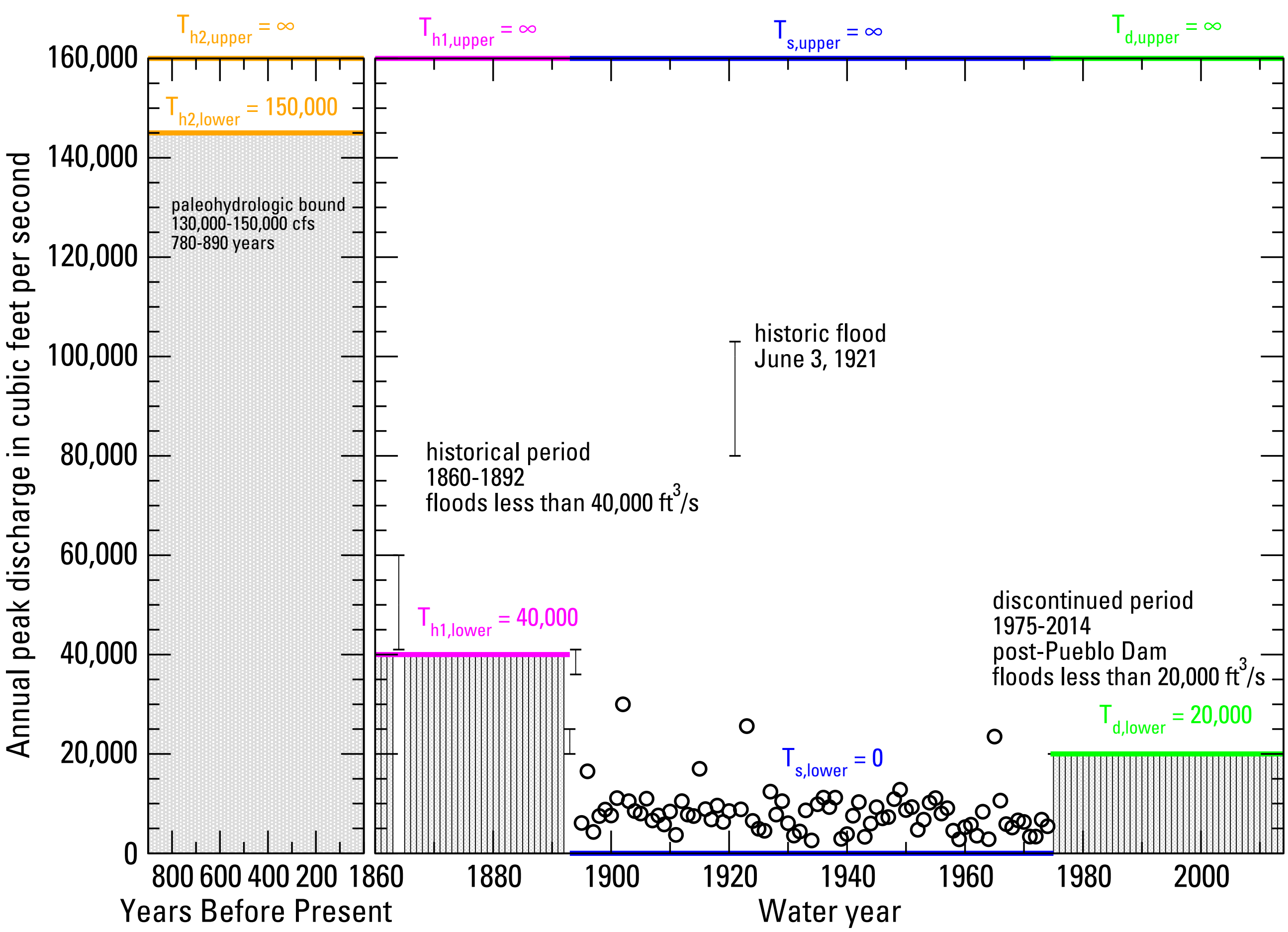

Figure 9.8. Peak discharge, historical and paleoflood estimates, Arkansas River at Pueblo State Park. A scale break is used to separate the gage and historical data from the longer paleoflood record. Flood intervals are shown as black vertical bars with caps that represent lower and upper flow estimates, including unobserved estimates in the historical period and historical floods in 1864, 1893, 1894 and 1921. The grey shaded areas represents floods of unknown magnitude less than the perception thresholds for the paleoflood period $T_{h2,lower}$, the historical period $T_{h1,lower}$, and the discontinued period $T_{d,lower}$. Perception threshold ranges are shown as orange lines for the paleoflood period, magenta lines for the historical period, blue lines for the systematic period, and green lines for the discontinued period.

Table 9.16. USGS Gage 07099500 EMA perception thresholds for the historical and systematic period from 1165 to 2004. This table contains the water year ranges to which each perception threshold applies, $T_{Y,lower}$ the lower bound of the perception threshold (in cfs) for water year Y, $T_{Y,upper}$, the upper bound of the perception threshold in cfs for water year Y, and a comment describing the threshold.

		EMA Perception Threshold		
Start Year	End Year	$T_{Y,lower}$	$T_{Y,upper}$	Comments
1165	1858	150000	infinity	paleoflood nonexceedance bound
1859	1892	40000	infinity	1864 historical information
1893	1894	19900	infinity	1893 historical information
1895	1976	0	infinity	continuous systematic record
1977	2004	20000	infinity	post-reservoir bound

Table 9.17. Peak-flow quantiles in cubic feet per second for USGS Gage 07099500 based on flood frequency analysis using EMA with MGBT; variance of estimate shown in log space.

Annual Exceedance Probability	EMA Estimate	Variance of Estimate	Lower 5% Confidence Limit	Upper 95% Confidence Limit
0.5	7100	0.000960	6300	8000
0.2	11900	0.001280	10400	13700
0.1	16400	0.001650	14100	19300
0.04	23800	0.002900	19600	29600
0.02	31000	0.004630	24300	40900
0.01	39800	0.007170	29500	56800
0.005	50600	0.010610	35600	79400
0.002	68800	0.016660	44800	124100
0.001	86300	0.022430	53000	174400
0.0001	177300	0.049590	88700	545300

Results from Flood Frequency Analysis

A flood frequency analysis for the Arkansas River at Pueblo was performed using the EMA flow intervals and perception thresholds as shown in Table 9.15 and Table 9.16. The output from an at-site flood frequency analysis using EMA with the Multiple Grubbs-Beck test is shown below; no PILFs were identified. Note that station skew was used, thus allowing the focus to be on the at-site data. The fitted frequency curve is displayed in Figure 9.9 with estimates provided in Table 9.17. The flood frequency results (Figure 9.9) indicate the LP-III model fits the bulk of the data well, including most of the large floods, but underfits the largest flood (June 1921) because of the paleoflood data influence. The paleoflood nonexceedance bound data at Pueblo State Park increases the peak discharge record length substantially to about 840 years, and has an effect on the upper end of the extrapolated frequency curve principally by reducing the skewness coefficient. One can observe the large positive skew and relatively steep transition between snowmelt-dominant floods to rainfall-dominant floods greater than about 10,000 ft^3/s. These large rainfall floods are responsible for the shape of the upper portion of the frequency curve. The *AEP* of the largest flood on record (June 1921) is about 1 in 270 from the exceedance-based plotting position, and about 1 in 1,600 from the LP-III model.

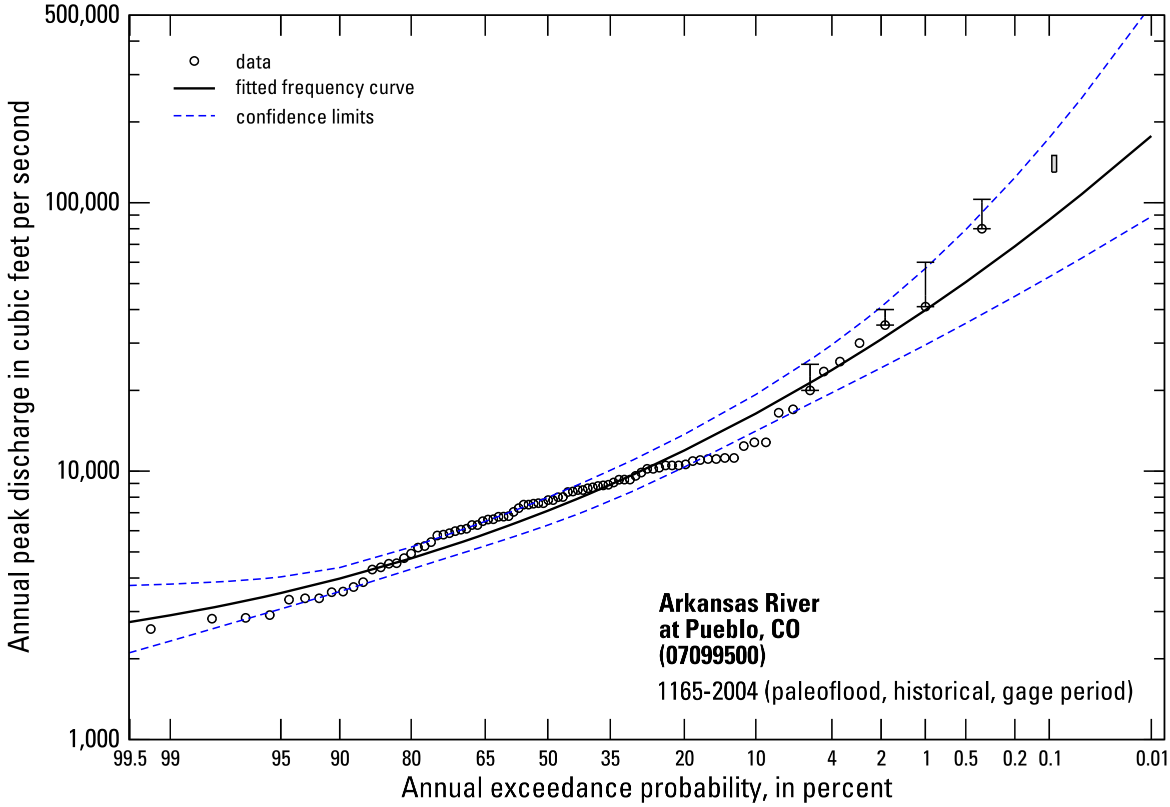

Figure 9.9. Peak discharge frequency curve, Arkansas River at Pueblo State Park, including gage, historical and paleoflood data. Peak discharge estimates from the gage are shown as open circles; vertical bars represent estimated data uncertainty for some of the largest floods. Paleoflood nonexceedance bound shown as a grey box.

Crest Stage Gage Example - Bear Creek at Ottumwa, IA

This example demonstrates how the Expected Moments Algorithm (*EMA*) and Multiple Grubbs-Beck test (*MGBT*) can be used to correctly perform a flood frequency analysis when censored data are present with variable perception thresholds from a crest stage gage.

A crest stage gage (CSG) is a simple, reliable device used to obtain the elevation of the flood peak of a stream. Most commonly, a CSG consists of a vertical metal pipe containing a wood or aluminum staff held in a fixed position with relation to a datum reference. At the bottom of the pipe is a perforated cap containing regranulated cork. When the water in the stream reaches and exceeds the height of the bottom cap (commonly referred to as the gage base), water is able to enter the pipe. As the water rises up the pipe, the cork floats on the waters surface and as the water reaches its peak and starts to recede, the cork adheres to the staff thereby retaining the crest stage of the flood (Sauer and Turnipseed, 2010). Thus, CSGs provide a censored record of peak flows, as no annual peak flow that results in a flood stage below the bottom cap of the pipe will be recorded. This example demonstrates how the Expected Moments Algorithm (EMA) with the Multiple Grubbs-Beck test for potentially influential low flows (PILFs) can correctly represent these censored annual peak records from CSGs in a flood frequency analysis.

For this example, USGS gage 05489490 Bear Creek at Ottumwa, IA is used. This gage is a CSG and has a drainage area of 22.9 square miles. It is located in southeast Iowa in the Southern Iowa Drift Plain land-form region which is characterized by rolling hills and deeply carved stream channels (Prior, 1991). The stream banks and channel bed are comprised of sand, silt, and clay materials that are prone to shifting from hydrologic events. The floodplain areas contain a combination of wooded areas, pasture, and row-crop fields.

Gage 05489490 has an annual peak record consisting of 49 peaks beginning in 1965 and ending in 2014 (Eash et al., 2013, Table 1). The annual peaks are listed in Table 9.18 (downloaded from USGS NWIS: http://nwis.waterdata.usgs.gov/nwis/peak/?site_no=05489490&agency_cd=USGS&) and shown in Figure 9.10.

EMA Representation of Peak Flow Data for Flood Frequency Analysis

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, when using EMA the annual peak flow for every water year during the historical period is described by a flow interval ($Q_{Y,lower}$, $Q_{Y,upper}$) for each water year Y. For peaks whose values are known and are not censored, the flow interval can be described as ($Q_{Y,lower} = Q_Y$, $Q_{Y,upper} = Q_Y$). For example, as shown in Table 9.18, the peak for the 1965 water year is recorded as 4000 cfs. This peak is known and is not censored, thus the flow interval for the 1965 water years is ($Q_{1965,lower} = 4000$, $Q_{1965,upper} = 4000$).

As shown in Table 9.18, there are 6 censored peaks occurring in 1966, 1971, 1975, 1988, 1997, and 2006. Five of these water years (1966, 1971, 1975, 1997, and 2006) have censored peaks due to the stage of the annual peak not reaching the gage base of the CSG. These peak can be described by flow intervals in which $Q_{Y,lower} = 0$ and $Q_{Y,upper} =$ CSG gage base. Similarly, the annual peak in water year 1988 is censored, however in this case the censoring is due to issues related to backwater. The CSG recorded an annual peak of 899 cfs, but it is known that the peak was affected by backwater due to ice causing the recorded peak to be larger than the actual peak. Thus, since there is no further information pertaining to the 1988 peak, it can be represented as a flow interval in which $Q_{1988,lower} = 0$ and $Q_{1988,upper} = 899$ cfs. Table 9.19 contains the EMA flow intervals for each water year in the record for Gage 05489490.

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, EMA distinguishes among sampling properties by employing perception thresholds denoted ($T_{Y,lower}$, $T_{Y,upper}$) for each year Y, 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

Table 9.18. USGS gage 05489490 annual peak flow record consisting of 49 peaks from 1965 to 2014. This table contains the date of the annual peak recorded at the gage, the water year of the annual peak and the corresponding annual peak in cubic feet per second (cfs).

Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)	Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)	Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)
1965-09-21	1965	4000	1982-07-03	1982	4030	1998-10-05	1999	2840
1966-00-00	1966	$<$ 1180	1982-10-08	1983	2180	2000-06-23	2000	3520
1967-06-09	1967	2880	1984-06-08	1984	1780	2001-05-15	2001	2430
1967-10-15	1968	1310	1985-03-04	1985	1610	2002-05-11	2002	2670
1968-10-15	1969	1420	1986-09-19	1986	1910	2003-06-26	2003	560
1970-06-24	1970	3130	1987-05-31	1987	990	2004-08-27	2004	3000
1971-00-00	1971	$<$ 1180	1988-02-20	1988	$<$ 899	2005-04-12	2005	859
1972-05-08	1972	1620	1989-09-09	1989	1820	2006-00-00	2006	$<$ 710
1973-01-19	1973	1570	1990-05-25	1990	3120	2007-08-23	2007	2390
1974-05-19	1974	2060	1991-04-18	1991	1850	2008-05-11	2008	3160
1975-00-00	1975	$<$ 705	1992-09-15	1992	1840	2009-08-27	2009	2520
1976-04-24	1976	3340	1993-05-07	1993	2410	2010-08-09	2010	3750
1977-08-07	1977	3530	1994-06-23	1994	1400	2011-06-14	2011	2600
1978-07-21	1978	2010	1995-04-11	1995	1560	2012-04-14	2012	1450
1979-03-29	1979	1830	1996-05-28	1996	3130	2013-05-28	2013	3850
1980-08-17	1980	2240	1997-00-00	1997	$<$ 714	2014-09-10	2014	1200
1981-07-04	1981	2770	1998-06-18	1998	1940			

Table 9.19. USGS Gage 05489490 EMA flow intervals for the systematic period from 1965 to 2014. This table contains the water year of the annual peak and the corresponding flow interval defined by lower bound, $Q_{Y,lower}$, and upper bound, $Q_{Y,upper}$, in cubic feet per second (cfs) for each water year Y.

Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments	Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments
1965	4000	4000		1990	3120	3120	
1966	0	1180	Peak $<$ gage base	1991	1850	1850	
1967	2880	2880		1992	1840	1840	
1968	1310	1310		1993	2410	2410	
1969	1420	1420		1994	1400	1400	
1970	3130	3130		1995	1560	1560	
1971	0	1180	Peak $<$ gage base	1996	3130	3130	
1972	1620	1620		1997	0	714	Peak $<$ gage base
1973	1570	1570		1998	1940	1940	
1974	2060	2060		1999	2840	2840	
1975	0	705	Peak $<$ gage base	2000	3520	3520	
1976	3340	3340		2001	2430	2430	
1977	3530	3530		2002	2670	2670	
1978	2010	2010		2003	560	560	
1979	1830	1830		2004	3000	3000	
1980	2240	2240		2005	859	859	
1981	2770	2770		2006	0	710	Peak $<$ gage base
1982	4030	4030		2007	2390	2390	
1983	2180	2180		2008	3160	3160	
1984	1780	1780		2009	2520	2520	
1985	1610	1610		2010	3750	3750	
1986	1910	1910		2011	2600	2600	
1987	990	990		2012	1450	1450	
1988	0	899	Peak affected by backwater	2013	3850	3850	
1989	1820	1820		2014	1200	1200	

peak discharges that have occurred. The lower bound, $T_{Y,lower}$, represents the smallest peak flow that would result in a recorded flow in water year Y. Thus, for a CSG, $T_{Y,lower}$ can be adjusted to accommodate a changing gage-base discharge. Table 9.20 contains the EMA perception thresholds for each water year in the record for Gage 05489490.

The annual peaks as well as their corresponding EMA flow intervals and perception thresholds can be displayed graphically. Figure 9.10 shows a representation of the recorded annual peaks, EMA flow intervals and EMA perception thresholds. The flow intervals whose lower bound is equal to the upper bound are represented by black circles, while the green lines represent the interval flood estimates for those peaks that were not able to be recorded as they were below gage base. The solid colored blocks represent the many perception thresholds applied to the record. The colored areas represent flows which would be unable to be recorded as they are smaller than the lower bound of the perception threshold $T_{Y,lower}$. The white space above the colored areas represents flow ranges for which annual peaks were able to be recorded had they occurred. For example, in Figure 9.10, the left-most light blue colored block represents a perception threshold from 1965 to 1972 where $T_{Y,lower}$=1180 cfs, $T_{Y,upper} = \infty$. The light blue colored block spans from 0 cfs to 1180 cfs signifying that no annual peak less than 1180 cfs could be measured during the time period from 1965 to 1972.

Results from Flood Frequency Analysis

A flood frequency analysis at USGS Gage 05489490 was performed using the EMA flow intervals and perception thresholds as shown in Table 9.19 and Table 9.20. The output from an at-site flood frequency analysis using EMA with the MGBT to screen for PILFs is shown below. Note that station skew was used, thus allowing the focus to be on the at-site data. The fitted frequency curve is displayed in Figure 9.11 with estimates provided in Table 9.21.

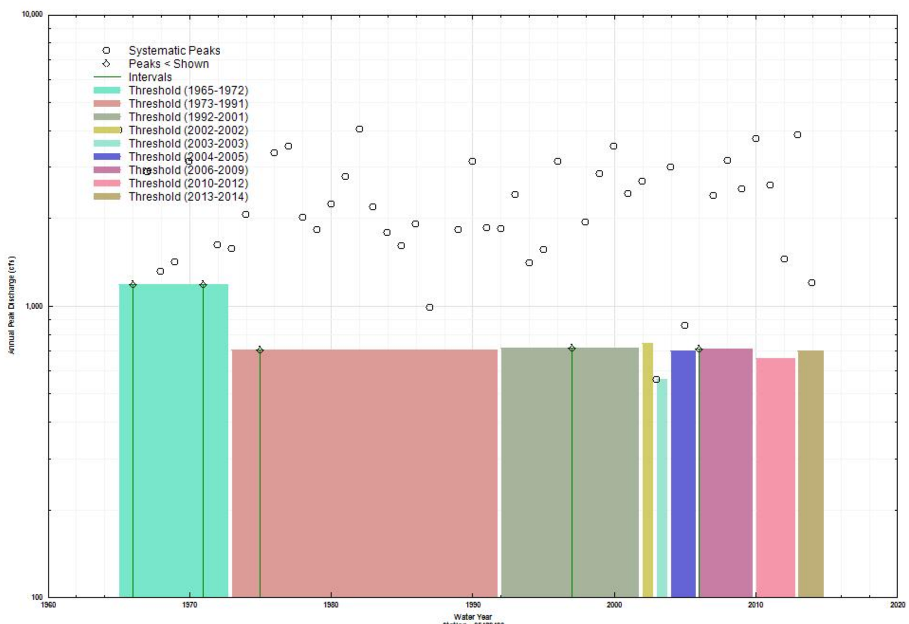

Figure 9.10. USGS gage 05489490 annual peak flow time series consisting of 49 peaks from 1965 to 2014. The black, open circles are the systematic peaks, the green lines with black, open triangles represent the interval flood estimates, and the solid rectangle blocks are the perception thresholds.

Table 9.20. USGS Gage 05489490 EMA perception thresholds for the systematic period from 1965 to 2014. This table contains the water year ranges to which each perception threshold applies, $T_{Y,lower}$ the lower bound of the perception threshold (in cfs) for water year Y, $T_{Y,upper}$, the upper bound of the perception threshold in cfs for water year Y, and a comment describing the threshold.

		EMA Perception Threshold		
Start Year	End Year	$T_{Y,lower}$	$T_{Y,upper}$	Comments
1965	1972	1180	infinity	initial gage base of CSG = 1180 cfs
1973	1991	705	infinity	gage base lowered
1992	2001	714	infinity	gage base raised as a result of spring thaw
2002	2002	743	infinity	gage base raised as a result of spring thaw
2003	2003	560	infinity	gage base lowered as a result of routine site visit (HWM)
2004	2005	700	infinity	gage base raised as a result of spring thaw
2006	2009	710	infinity	gage base raised as a result of spring thaw
2010	2012	661	infinity	gage base lowered as a result of spring thaw
2013	2014	700	infinity	gage base raised as a result of spring thaw

Table 9.21. Peak-flow quantiles in cubic feet per second for USGS Gage 05489490 based on flood frequency analysis using EMA with MGBT; variance of estimate shown in log space.

Annual Exceedance Probability	EMA Estimate	Variance of Estimate	Lower 5% Confidence Limit	Upper 95% Confidence Limit
0.5	2061	0.0012	1702	2406
0.2	3004	0.0009	2611	3444
0.1	3507	0.0008	3080	4064
0.04	4021	0.001	3543	4856
0.02	4329	0.0013	3779	5454
0.01	4586	0.0018	3942	6074
0.005	4802	0.0024	4057	6746
0.002	5036	0.0033	4160	7750

As shown in the example above, EMA correctly represents the censored annual peak data through the use of flow intervals and perception thresholds. The EMA flow intervals provide a straightforward approach to appropriately represent the censored flows, while the perception thresholds accommodate the changing gage base. Special thanks to Jon Nania and David Eash of the USGS Iowa WSC for providing data and insight relating to USGS Gage 05489490 Bear Creek at Ottumwa, IA.

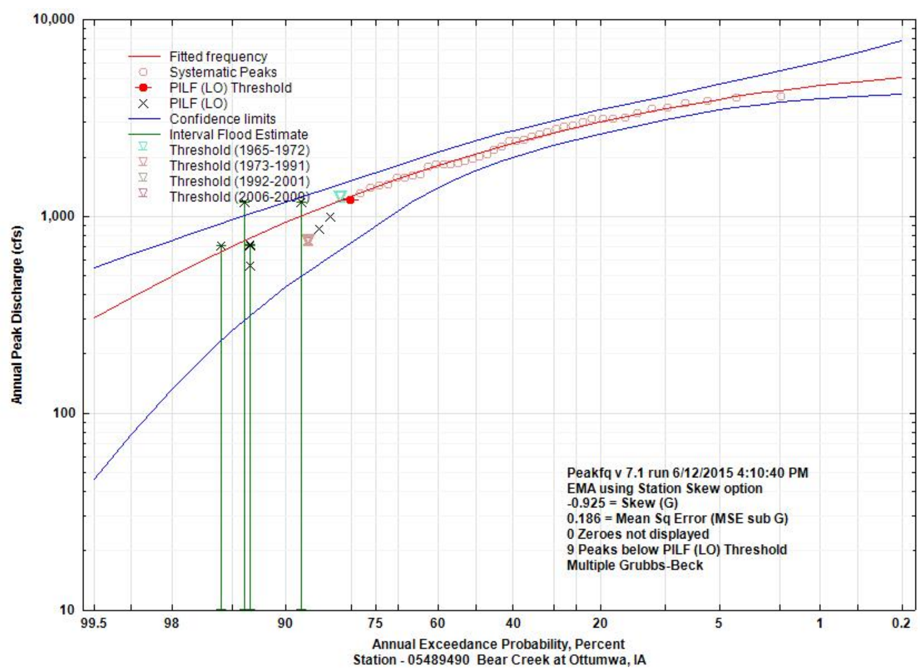

Figure 9.11. Annual Exceedance Probability Plot for USGS Gage 05489490 based on flood frequency analysis using EMA with MGBT. The red line is the fitted log Pearson Type III frequency curve, the blue lines are the upper and lower bounds of the confidence limits, the red circles are the systematic peaks, the green lines represent the interval flood estimates, the solid red circle with a line through it is the potentially influential low floods (PILFs) thresholds as identified by the MGBT, and the black x's are the PILFs identified by the MGBT.

Historical and PILF Example - Santa Cruz River near Lochiel, AZ

This example illustrates the use of *EMA* for a historical record with one large flood (described in the Section Historical Flood Information) and a number of PILFs (described in Section Zero Flows and Potentially-Influential Low Floods).

For this example, USGS gage 09480000 Santa Cruz River near Lochiel, Arizona is used. The Santa Cruz River is a tributary to the Gila River; the 82.2 square mile watershed lies within the Basin and Range province in Arizona (Paretti et al., 2014a). This gaging station was previously analyzed by Cohn et al. (2014) and Paretti et al. (2014a, Figure 21).

Gage 09480000 has an annual peak record consisting of 65 peaks beginning in 1949 and ending in 2013. There is a historic flood that occurred within the period of gaging record on October 9, 1977. This flood is noted in the USGS Annual Water Data Report for this gage, available in the peak-flow file, and there is historical information available for this large flood (Aldridge and Eychaner, 1984), that indicates this flood is the largest since 1927. The annual peaks are listed in Table 9.22 and shown in Figure 9.12. Of the 65 annual peaks, the August 15, 1984 flood is equal to the October 1977 historic flood peak. Based on the historical flood information in Aldridge and Eychaner (1984) for the 1977 flood, information from the October 1977 flood is used as a perception threshold to represent the 22 years of missing information from 1927-1946.

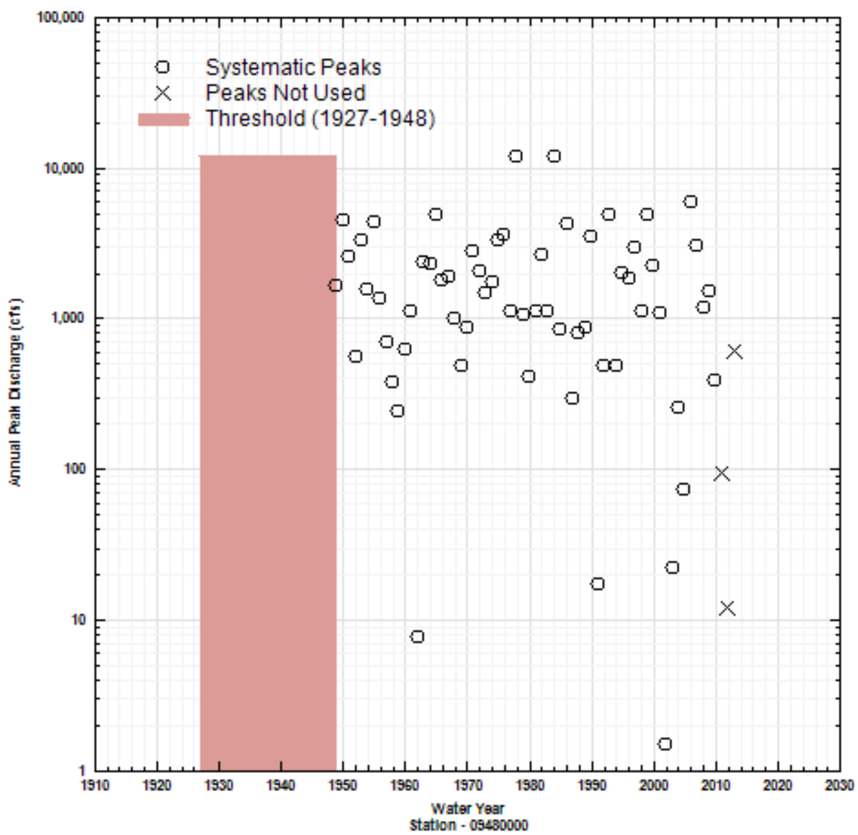

Figure 9.12. USGS gage 09480000 annual peak flow time series consisting of 65 peaks from 1949 to 2013. The historical period is shown in red, with perception threshold (12,000 cfs) estimated from the October 1977 historic flood.

Table 9.22. USGS gage 09480000 annual peak flow record consisting of 65 peaks from 1949 to 2013. This table contains the date of the annual peak recorded at the gage, the water year of the annual peak and the corresponding annual peak in cubic feet per second (cfs).

Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)	Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)	Date of Peak Streamflow	Water Year	Annual Peak Streamflow (cfs)
1949-09-13	1949	1650	1971-08-10	1971	2830	1993-01-18	1993	4880
1950-07-30	1950	4520	1972-07-16	1972	2070	1994-08-30	1994	478
1951-08-02	1951	2560	1973-06-30	1973	1490	1995-07-12	1995	2020
1952-08-16	1952	550	1974-08-04	1974	1730	1996-07-10	1996	1860
1953-07-14	1953	3320	1975-07-22	1975	3330	1997-09-11	1997	2970
1954-07-22	1954	1570	1976-07-22	1976	3540	1998-07-07	1998	1110
1955-08-06	1955	4300	1977-09-05	1977	1130	1999-07-28	1999	4870
1956-07-17	1956	1360	1977-10-09	1978	12000	2000-08-06	2000	2240
1957-08-09	1957	688	1979-01-25	1979	1060	2000-10-22	2001	1080
1958-08-07	1958	380	1980-06-30	1980	406	2002-03-04	2002	1.5
1959-08-14	1959	243	1981-07-15	1981	1110	2003-08-14	2003	22
1960-07-30	1960	625	1982-08-11	1982	2640	2004-08-05	2004	256
1961-08-08	1961	1120	1983-03-04	1983	1120	2005-08-23	2005	73
1962-07-29	1962	7.6	1984-08-15	1984	12000	2006-08-08	2006	5940
1963-08-25	1963	2390	1985-07-19	1985	850	2007-07-19	2007	3060
1964-09-09	1964	2330	1986-08-29	1986	4210	2008-07-23	2008	1180
1965-09-12	1965	4810	1987-08-10	1987	291	2009-07-20	2009	1530
1966-08-18	1966	1780	1988-08-23	1988	804	2010-07-31	2010	392
1967-08-03	1967	1870	1989-08-04	1989	871	2011-08-13	2011	95
1967-12-20	1968	986	1990-07-17	1990	3510	2012-07-28	2012	12
1969-08-05	1969	484	1991-07-26	1991	17	2013-09-08	2013	612
1970-08-03	1970	880	1992-08-01	1992	483			

EMA Representation of Peak Flow Data for Flood Frequency Analysis

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, when using EMA the annual peak flow for every water year during the historical period is described by a flow interval $(Q_{Y,lower}, Q_{Y,upper})$ for each water year Y. For peaks whose values are known and are not censored, the flow interval can be described as $(Q_{Y,lower} = Q_Y, Q_{Y,upper} = Q_Y)$. In this example, the flow values are known for all the years where the gage was in operation. Table 9.23 contains the EMA flow intervals for each water year in the record for gage 09480000. The historical period is described by a perception threshold.

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, EMA distinguishes among sampling properties by employing perception thresholds denoted $(T_{Y,lower}, T_{Y,upper})$ for each year Y, 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 in water year Y. For most peaks at most gages, $T_{Y,upper}$ is assumed to be infinite, as bigger floods that might exceed the measurement capability of the streamgage are determined through study of highwater marks and other physical evidence of the flood. For periods of continuous, full-range peak flow record, the perception threshold is represented by $(T_{Y,lower} = 0, T_{Y,upper} = \infty)$, where $T_{Y,lower} = 0$ is the gage-base discharge. Based on the October 1977 large historical flood (Aldridge and Eychaner, 1984), it is known that floods at this location would have been estimated (or recorded), had they exceeded approximately 12,000 cfs. Table 9.24 contains the EMA perception thresholds for each water year in the record, including the historical period, for Gage 09480000.

Results from Flood Frequency Analysis

A flood frequency analysis at USGS Gage 09480000 was performed using the EMA flow intervals and perception thresholds as shown in Table 9.23 and Table 9.24. The output from an at-site flood frequency analysis using EMA with the Multiple Grubbs-Beck test to screen for potentially influential low floods (PILFs) is shown below. Note that station skew was used, thus allowing the focus to be on the at-site data. The fitted frequency curve is displayed in Figure 9.13 with estimates provided in Table 9.25.

As shown in Figure 9.13, there are two floods that exceed the historical threshold (12,000 cfs): the October 1977 flood and the August 1984 flood. Using MGBT, eight PILFs were identified, with a threshold equal to 380 ft^3/s. Thus, all 8 annual peaks less than 380 ft^3/s are censored and re-coded in the framework of EMA with flow intervals of $(Q_{Y,lower} = 0, Q_{Y,upper} = 380)$. The MGBT threshold also has the effect of adjusting the lower bound of the perception threshold. Thus for the systematic period from 1949 to 2013, the perception threshold is $(T_{Y,lower} = 380, T_{Y,upper} = \infty)$. For the historical information, the lower threshold $T_{Y,lower} = 12000$ (Table 9.24). As shown in Figure 9.13, by censoring the eight smallest peaks in the record, the remaining smallest annual exceedance probability peaks and the largest floods are well fit by the frequency curve (red line).

Table 9.23. USGS Gage 09480000 EMA flow intervals for the systematic period from 1949 to 2013. This table contains the water year of the annual peak and the corresponding flow interval defined by lower bound, $Q_{Y,lower}$, and upper bound, $Q_{Y,upper}$, in cubic feet per second (cfs) for each water year Y.

Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments	Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments
1949	1650	1650		1982	2640	2640	
1950	4520	4520		1983	1120	1120	
1951	2560	2560		1984	12000	12000	
1952	550	550		1985	850	850	
1953	3320	3320		1986	4210	4210	
1954	1570	1570		1987	291	291	
1955	4300	4300		1988	804	804	
1956	1360	1360		1989	871	871	
1957	688	688		1990	3510	3510	
1958	380	380		1991	17	17	
1959	243	243		1992	483	483	
1960	625	625		1993	4880	4880	
1961	1120	1120		1994	478	478	
1962	7.6	7.6		1995	2020	2020	
1963	2390	2390		1996	1860	1860	
1964	2330	2330		1997	2970	2970	
1965	4810	4810		1998	1110	1110	
1966	1780	1780		1999	4870	4870	
1967	1870	1870		2000	2240	2240	
1968	986	986		2001	1080	1080	
1969	484	484		2002	1.5	1.5	
1970	880	880		2003	22	22	
1971	2830	2830		2004	256	256	
1972	2070	2070		2005	73	73	
1973	1490	1490		2006	5940	5940	
1974	1730	1730		2007	3060	3060	
1975	3330	3330		2008	1180	1180	
1976	3540	3540		2009	1530	1530	
1977	1130	1130		2010	392	392	
1978	12000	12000	historic flood	2011	95	95	
1979	1060	1060		2012	12	12	
1980	406	406		2013	612	612	
1981	1110	1110					

Table 9.24. USGS Gage 09480000 EMA perception thresholds for the historical and systematic period from 1927 to 2013. This table contains the water year ranges to which each perception threshold applies, $T_{Y,lower}$ the lower bound of the perception threshold (in cfs) for water year Y, $T_{Y,upper}$, the upper bound of the perception threshold in cfs for water year Y, and a comment describing the threshold.

		EMA Perception Threshold		
Start Year	**End Year**	$T_{Y,lower}$	$T_{Y,upper}$	**Comments**
1927	1946	12000	infinity	historical information
1949	2013	0	infinity	continuous systematic record

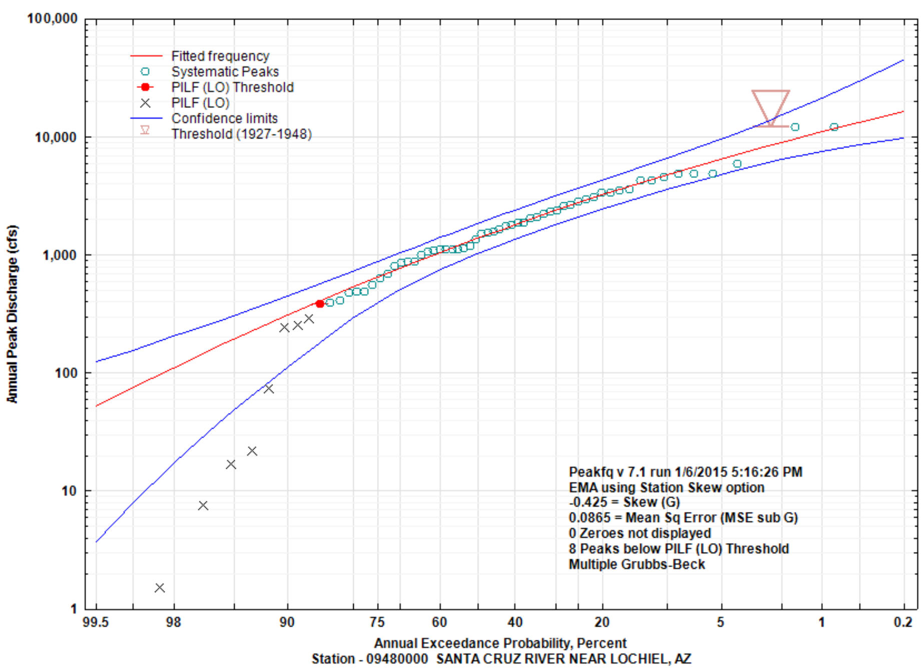

Figure 9.13. Annual Exceedance Probability Plot for USGS Gage 09480000 based on flood frequency analysis using EMA with MGBT. The red line is the fitted log Pearson Type III frequency curve, the blue lines are the upper and lower bounds of the confidence limits, the black circles are the systematic peaks, the solid red circle with a line through it is the potentially influential low floods (PILFs) thresholds as identified by the MGBT, and the black x's are the PILFs identified by the MGBT.

Table 9.25. Peak-flow quantiles in cubic feet per second for USGS Gage 09480000 based on flood frequency analysis using EMA with MGBT; variance of estimate shown in log space.

Annual Exceedance Probability	EMA Estimate	Variance of Estimate	Lower 5% Confidence Limit	Upper 95% Confidence Limit
0.5	1279	0.0042	936.1	1719
0.2	3079	0.0040	2314	4138
0.1	4652	0.0042	3481	6394
0.04	6982	0.0056	5119	10460
0.02	8914	0.0076	6337	14780
0.01	10970	0.0105	7474	20570
0.005	13150	0.0144	8509	28280
0.002	16170	0.0211	9719	42560

Paleoflood Record Example - American River at Fair Oaks, CA

This example illustrates the use of *EMA* for a historical record with several large floods (described in the Section Historical Flood Information) and detailed paleoflood data (described in the Section Paleoflood and Botanical Information), utilizing multiple censoring and interval data for Reclamation's Folsom Dam (Bureau of Reclamation, 2002). For this example, USGS gage 11446500 American River at Fair Oaks, California is used. In 1986 and 1997 floods on the American River and in central California heightened concerns about the hydrologic risk at Folsom Dam. In part, these concerns led to two National Research Council panels to evaluate American River flood hazards (National Research Council, 1995, 1999). These panels reviewed flood control and floodplain management issues, focusing on estimating floods with *AEPs* greater than 0.005 (1 in 200), specifically 1 in 100 (0.01). For dam safety, the primary concern is floods with very small *AEPs* generally in the range of 0.001 to 0.0001 (1 in 1,000 to 1 in 10,000). For this example, these estimates are made using gage, historical and paleoflood data.

Bureau of Reclamation (2002) conducted a paleoflood and flood frequency study to investigate these issues. The primary objective of the study was to develop an estimate of peak discharge frequency of the American River at Folsom Dam in the above annual probability range. The peak discharge frequency information was subsequently combined with historical hydrographs to develop probabilistic hydrographs based on paleoflood information. Paleoflood information for the Bureau of Reclamation (2002) study was based on geomorphic, stratigraphic, and geochronologic information collected from four sites in the American River basin: 1) South Fork American River near Kyburz, 2) South Fork American River near Lotus, 3) North Fork of the American River at Ponderosa Bridge, and 4) lower American River near Fair Oaks. Two main types of paleoflood data were collected from the four sites to evaluate the flood hazard for Folsom Dam: 1) paleoflood magnitude and age estimates for the South Fork near Kyburz and Lotus, and the lower American River, and 2) a single paleohydrologic bound for the North Fork. Stratigraphic information from 14 sites provides evidence for late Holocene paleofloods that are preserved at or above the peak stage of the largest historical floods. The age of these paleofloods is constrained by 38 radiocarbon ages, published archaeological age correlations, and published obsidian hydration age estimates.

For this example, peak-flow data from the lower American River at Fair Oaks, California (Gage 11446500) are used. There are 77 peaks beginning in 1905 and ending in 1997, with several years with very low floods or missing values (1910, 1912-13, 1918, 1929, 1977). Large historical floods occurred in 1997, 1986, and 1862, and are described in National Research Council (1999) and Bureau of Reclamation (2002). The paleoflood period covers the past 2,000 years, from year 1 to 1847, the historical period begins in 1848, and the gaging period begins in 1905. The annual peaks, historical floods and paleofloods are listed in Table 9.26 and shown in Figure 9.14. Perception thresholds are estimated based on the March 1907 flood, the January 1862 flood, and paleofloods.

EMA Representation of Peak Flow Data for Flood Frequency Analysis

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, when using EMA the annual peak flow for every water year during the historical period is described by a flow interval $(Q_{Y,lower}, Q_{Y,upper})$ for each water year Y. For peaks whose values are known and are not censored, the flow interval can be described as $(Q_{Y,lower} = Q_Y, Q_{Y,upper} = Q_Y)$. In this example, the flow values are known for all the years where the gage was in operation. Table 9.27 contains the EMA flow intervals for each water year in the record for gage 11446500. The historical and paleoflood periods are described by perception thresholds.

As described in the Data Representation using Flow Intervals and Perception Thresholds Section, EMA distinguishes among sampling properties by employing perception thresholds denoted $(T_{Y,lower}, T_{Y,upper})$ for

Table 9.26. USGS gage 11446500 American River at Fair Oaks annual peak flow record consisting of 77 peaks from 1905 to 1997, with historical floods and paleofloods. Horizontal lines indicate breaks in data.

Water Year	Annual Peak Streamflow (cfs)	Water Year	Annual Peak Streamflow (cfs)	Water Year	Annual Peak Streamflow (cfs)
650	>600000	1933	16500	1961	8000
1437	>400000	1934	22600	1962	40000
1574	>400000	1935	60900	1963	240000
1711	>400000	1936	58300	1964	24000
1862	>262000	1937	33000	1965	260000
1905	24200	1938	114000	1966	6500
1906	59700	1939	10900	1967	46000
1907	156000	1940	89200	1968	30000
1908	10300	1941	38800	1969	120000
1909	119000	1942	83200	1970	122000
1911	81300	1943	152000	1971	48000
1914	74100	1944	20100	1972	12000
1915	47900	1945	94400	1973	69000
1916	40700	1946	42200	1974	55000
1917	42300	1947	27900	1975	46000
1919	67500	1948	21000	1976	15000
1920	20100	1949	37500	1978	40000
1921	39200	1950	34400	1979	33000
1922	31600	1951	180000	1980	175000
1923	39000	1952	37200	1981	20000
1924	14000	1953	49700	1982	152000
1925	99500	1954	42600	1983	93000
1926	27400	1955	10800	1984	88000
1927	67700	1956	219000	1985	17000
1928	163000	1957	42000	1986	259000
1930	24400	1958	54000	1997	298000
1931	9900	1959	20000		
1932	21100	1960	75000		

Table 9.27. USGS Gage 11446500 EMA flow intervals for the historical and systematic period from 650 to 1997.

Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments	Water Year	$Q_{Y,lower}$	$Q_{Y,upper}$	Comments
650	600000	850000	paleoflood	1946	42200	42200	
1437	400000	550000	paleoflood	1947	27900	27900	
1574	400000	550000	paleoflood	1948	21000	21000	
1711	400000	550000	paleoflood	1949	37500	37500	
1862	262000	300000	historical flood	1950	34400	34400	
1905	24200	24200		1951	180000	180000	
1906	59700	59700		1952	37200	37200	
1907	156000	156000		1953	49700	49700	
1908	10300	10300		1954	42600	42600	
1909	119000	119000		1955	10800	10800	
1911	81300	81300		1956	219000	219000	
1914	74100	74100		1957	42000	42000	
1915	47900	47900		1958	54000	54000	
1916	40700	40700		1959	20000	20000	
1917	42300	42300		1960	75000	75000	
1919	67500	67500		1961	8000	8000	
1920	20100	20100		1962	40000	40000	
1921	39200	39200		1963	240000	240000	
1922	31600	31600		1964	24000	24000	
1923	39000	39000		1965	260000	260000	
1924	14000	14000		1966	6500	6500	
1925	99500	99500		1967	46000	46000	
1926	27400	27400		1968	30000	30000	
1927	67700	67700		1969	120000	120000	
1928	163000	163000		1970	122000	122000	
1930	24400	24400		1971	48000	48000	
1931	9900	9900		1972	12000	12000	
1932	21100	21100		1973	69000	69000	
1933	16500	16500		1974	55000	55000	
1934	22600	22600		1975	46000	46000	
1935	60900	60900		1976	15000	15000	
1936	58300	58300		1978	40000	40000	
1937	33000	33000		1979	33000	33000	
1938	114000	114000		1980	175000	175000	
1939	10900	10900		1981	20000	20000	
1940	89200	89200		1982	152000	152000	
1941	38800	38800		1983	93000	93000	
1942	83200	83200		1984	88000	88000	
1943	152000	152000		1985	17000	17000	
1944	20100	20100		1986	259000	259000	
1945	94400	94400		1997	298000	298000	

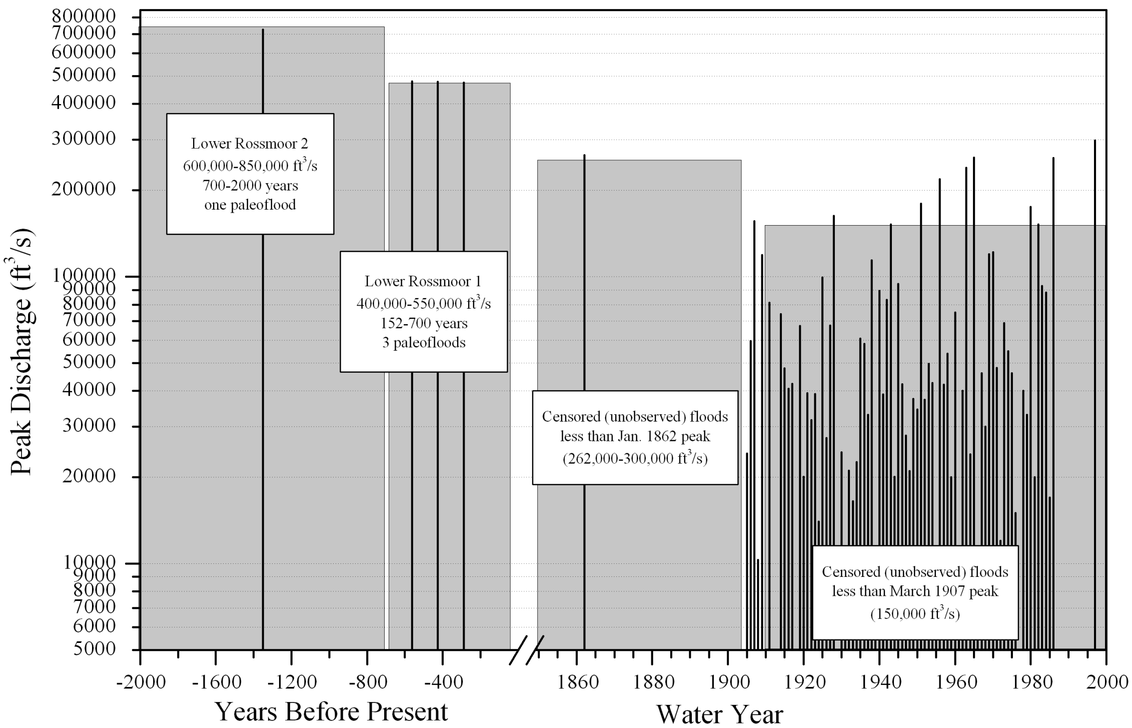

Figure 9.14. Approximate unregulated peak discharge and paleoflood estimates, with historical and paleoflood exceedance thresholds, American River at Fair Oaks. A scale break is used to separate the gaging station data from the much longer paleoflood record. Mean values of paleofloods and threshold age and discharge data are plotted for simplicity.

each year Y, 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 in water year Y. For most peaks at most gages, $T_{Y,upper}$, is assumed to be infinite, as bigger floods that might exceed the measurement capability of the streamgage are determined through study of highwater marks and other physical evidence of the flood. For periods of continuous, full-range peak flow record, the perception threshold is represented by ($T_{Y,lower} = 0$, $T_{Y,upper} = \infty$), where $T_{Y,lower} = 0$ is the gage-base discharge. Based on the March 1907 large historical flood (Bureau of Reclamation, 2002), it is known that floods at this location would have been estimated (or recorded), had they exceeded approximately 150,000 cfs. Table 9.28 contains the EMA perception thresholds for each water year in the record, including the historical period, for Gage 11446500.

Results from Flood Frequency Analysis

A flood frequency analysis at USGS Gage 11446500 was performed using the EMA flow intervals and perception thresholds as shown in Table 9.27 and Table 9.28. The output from an at-site flood frequency analysis

Table 9.28. USGS Gage 11446500 EMA perception thresholds for the historical and systematic period from 1927 to 2013. This table contains the water year ranges to which each perception threshold applies, $T_{Y,lower}$ the lower bound of the perception threshold (in cfs) for water year Y, $T_{Y,upper}$, the upper bound of the perception threshold in cfs for water year Y, and a comment describing the threshold.

		EMA Perception Threshold		
Start Year	End Year	$T_{Y,lower}$	$T_{Y,upper}$	Comments
1	1301	599000	infinity	Lower Rossmoor Terrace 1
1302	1847	399000	infinity	Lower Rossmoor Terrace 2
1848	1904	261000	infinity	1862 historical threshold
1905	1909	0	infinity	gage record
1910	1910	150000	infinity	March 1907 Low floods and Missing
1911	1911	0	infinity	gage record
1912	1913	150000	infinity	March 1907 Low floods and Missing
1914	1917	0	infinity	gage record
1918	1918	150000	infinity	March 1907 Low floods and Missing
1919	1928	0	infinity	gage record
1929	1929	150000	infinity	March 1907 Low floods and Missing
1930	1976	0	infinity	gage record
1977	1977	150000	infinity	March 1907 Low floods and Missing
1978	1986	0	infinity	gage record
1987	1996	150000	infinity	March 1907 Low floods and Missing
1997	1997	0	infinity	gage record
1998	2000	150000	infinity	March 1907 Low floods and Missing

using EMA with the Multiple Grubbs-Beck test to screen for potentially influential low floods (PILFs) is shown below. Note that station skew was used, thus allowing the focus to be on the at-site data. The fitted frequency curve is displayed in Figure 9.15 with estimates provided in Table 9.29. Peak discharge estimates for the interval floods are shown in the figure with estimated uncertainty. Peak discharge probabilities are estimated using Cunnane's plotting position with the threshold-exceedance formula that includes paleoflood data. The results indicate that the LP-III model provides an adequate fit to the gage and paleoflood data.

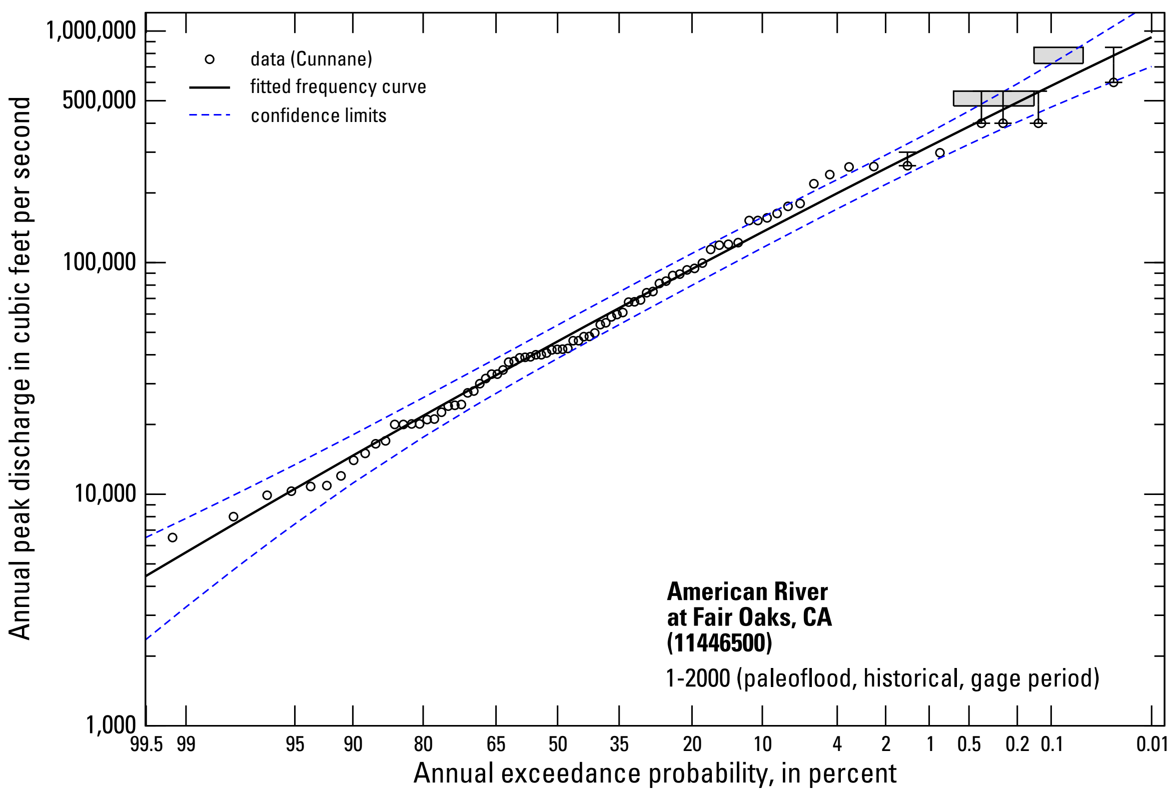

Figure 9.15. Approximate unregulated peak discharge frequency curve, American River at Fair Oaks, including gage, historical and paleoflood data. Peak discharge estimates from the gage are shown as open circles; vertical bars represent estimated data uncertainty for some of the largest floods. Paleoflood nonexceedance bounds shown as grey boxes.

Table 9.29. Peak-flow quantiles in cubic feet per second for USGS Gage 11446500 based on flood frequency analysis using EMA with MGBT; variance of estimate shown in log space.

Annual Exceedance Probability	EMA Estimate	Variance of Estimate	Lower 5% Confidence Limit	Upper 95% Confidence Limit
0.5	45700	0.001890	38600	53700
0.2	93800	0.001730	79800	109600
0.1	135500	0.001590	115800	157000
0.04	199400	0.001460	170700	228300
0.02	255000	0.001450	217500	291100
0.01	317500	0.001570	268800	364500
0.005	387300	0.001830	324600	451400
0.002	491600	0.002440	404700	591900
0.001	580200	0.003110	469300	720800
0.0001	941200	0.006810	702800	1325300

Glossary

Acronyms

ACWI — Advisory Committee on Water Information.

AEP — Annual Exceedance Probability.

AMS — Annual Maximum Series.

B-GLS — Bayesian Generalized Least Squares.

CDF — Cumulative Distribution Function.

CSG — Crest-stage gage.

EMA — Expected Moments Algorithm.

FEMA — Federal Emergency Management Agency.

GLS — Generalized Least Squares.

HFAWG — Hydrologic Frequency Analysis Work Group.

HWM — High-water mark.

LP-III — Log-Pearson Type III distribution.

MGBT — Multiple Grubbs-Beck Test.

MOVE — Maintenance of Variance Extension.

MSE — Mean-Square Error.

NRCS — Natural Resources Conservation Service.

NWIS — USGS National Water Information System.

NWS — National Weather Service.

OLS — Ordinary Least Squares.

PDS — Partial-Duration Series.

PILF — Potentially-Influential Low Flood.

PSI — Paleostage indicator.

Reclamation — Bureau of Reclamation.

RFC — River Forecast Center (NWS).

SOH — Subcommittee on Hydrology.

USACE — U.S. Army Corps of Engineers

WLS — Weighted Least Squares.

Symbols

a — Plotting position parameter that is dependent on an assumed distribution ($0 \leq a \leq 0.5$); $a = 0 \equiv$ Weibull; $a = 0.5 \equiv$ Hazen (eq. 2).

e_s — Number of floods/records that exceed a censoring level, such as a historical threshold T_h or low-outlier threshold T_{PILF}, during the systematic record n_s ($e_s \leq k; e_s < n_s$).

e_h — Number of floods/records that exceed the historical threshold T_h during the historical/paleoflood period n_h ($e_h \leq k; e_h < n_h$).

g — Total number of known flood (observations) during the entire period of observation record n ($g = n_s + k - e_s = n_s + e_h$).

$\hat{\gamma}$ — At-site (station) sample skew coefficient (in log space).

G — Regional sample skew coefficient (in log space).

$\tilde{G}$ — Weighted skew coefficient.

k — Total number of floods/records that exceed a censoring level, such as a historical threshold T_h or low-outlier threshold T_{PILF}, during the entire period of observation record n ($k = e_s + e_h$).

$\hat{\mu}$ — At-site (station) sample mean (in log space).

$\hat{\sigma}$ — At-site (station) sample standard deviation (in log space).

n — Total peak-flow period of record (years), including systematic n_s and historical n_h periods, as available, where $n = n_s + n_h$.

n_h — Length of the historical period (years); possibly includes a paleoflood period ($n_h < n$).

n_s — Length of the peak-flow systematic (gaging) record (years) ($n_s \leq n$).

p — Annual exceedance probability (AEP), $p = 1 - q$.

q — Cumulative probability, $q = 1 - p$.

Q — Flood discharge.

Q_b — Base discharge. Can be a constant, or vary with each year at a gaging station or CSG.

Q_p — Discharge quantile for annual exceedance probability p.

Q_q — Discharge quantile for cumulative probability q, equivalent to Q_p.

Q_Y — Flood discharge estimate in year Y.

$Q_{Y,lower}$ — Discharge lower bound for year Y in EMA.

$Q_{Y,upper}$ — Discharge upper bound for year Y in EMA.

T_h — Perception threshold for a historical period n_h.

T_{PILF} — PILF censoring threshold from the *MGBT*.

$T_{Y,lower}$ — Perception threshold lower bound for Year Y in EMA; represents the smallest peak flow that would result in a recorded flow.

$T_{Y,upper}$ — Perception threshold upper bound for Year Y in EMA; represents the largest peak flow that would result in a recorded flow.

X_h — Base-10 logarithm of a perception threshold for a historical period n_h.

X_l — Base-10 logarithm of the PILF censoring threshold from the *MGBT*.

$X_{Y,lower}$ — Base-10 logarithm of discharge lower bound $Q_{Y,lower}$ for Year Y in EMA.

$X_{Y,upper}$ — Base-10 logarithm of discharge upper bound $Q_{Y,upper}$ for Year Y in EMA.

Y — Year.

Definitions

annual exceedance probability (*AEP*) — The probability that flooding will occur in any given year considering the full range of possible annual floods.

annual flood — The highest instantaneous peak discharge in each year of record. Practically, this is the highest alue observed in the record of 15 minute or 60 minute values, depending on the recording interval of the device. Sometimes the maximum mean daily discharge is used on larger rivers.

annual flood series — A list of annual maximum floods.

annual series — A general term for a set of any kind of data in which each item is the maximum or minimum in a year.

autocorrelation — The presence of autocorrelation indicates that the data in the time series are not random. Rather, future values are correlated with past values. Autocorrelation is calculated as the correlation between the values in a time series and the values in that same time series lagged by one or more timesteps (i.e., the correlation between X_i and X_{i+k} where i is the timestep and k is the lag). Also known as serial correlation.

base discharge (for peak discharge) — A discharge value, determined for selected stations, above which peak discharge data are published. The base discharge at each station is selected so that an average of about three peak flows per year will be published (Langbein and Iseri, 1960).

binomial censored data — Floods that exceeded a threshold, where one knows only that a flood was larger than some level, and does not know the magnitude of the flood (Russell, 1982; Stedinger and Cohn, 1986).

broken record — A systematic record which is divided into separate continuous segments because of deliberate discontinuation of recording for significant periods of time. This typyically occurs when a gage is shut off due to funding, prioritization, other hydrological or management reasons, then reestablished at a later time (several years, rather than weeks or months) at the same location.

censored data — In a sample size of n, a known number of observations is missing at either end or at both ends (David, 1981; Cohen, 1991).

coefficient of skewness — A numerical measure or index of the lack of symmetry in a frequency distribution. Function of the third moment of magnitudes about their mean, a measure of asymmetry. Also called *coefficient of skew* or *skew coefficient*.

confidence limits — Computed values on both sides of an estimate of a parameter or quantile that show for a specified probability the range in which the true value of the parameter or quantile lies.

crest-stage gage (CSG) — A simple, economical, reliable, and easily installed device for obtaining the elevation of the flood crest of streams (Sauer and Turnipseed, 2010). These gages are nonrecording and consist of a partial streamflow record. Flow intervals and perception thresholds are needed to describe each year of the flood record.

cross-correlation — A measure of similarity, interdependence or relationship between two time series of observations in space at the same point or lagged points in time.

exceedance — Knowledge that the magnitude (discharge or stage) of a flood was larger than some level or threshold. For example, the flood exceeded the bridge deck.

exceedance frequency — The percentage of values that exceed a specified magnitude, 100 times exceedance probability.

exceedance probability — Probability that a random event will exceed a specified magnitude in a given time period, usually one year unless otherwise indicated.

Expected Moments Algorithm (*EMA*) — A generalized method of moments procedure to estimate the P-III distribution parameters using the entire data set, simultaneously employing regional skew information and a wide range of historical flood and threshold-exceedance information, while adjusting for any potentially influential low floods, missing values from an incomplete record, or zero flood years.

extraordinary flood — Those floods that are the largest magnitude at a gaging station or miscellaneous site and that substantially exceed the other flood observations (Costa and Jarrett, 2008).

gage base — The minimum stage or discharge level at a gaging station, below which observations are not recorded or published. Also called *base discharge*.

gaging record — Streamflow data collected at streamflow-gaging stations. A gaging record can consist of systematic data and historical flood data.

gaging station — A selected site on a stream equipped and operated to furnish basic data from which continuous, systematic records of stage and discharge may be obtained (Grover and Harrington, 1943; Rantz and Others, 1982).

generalized skew coefficient — See *regional skew coefficient*.

high-water mark (HWM) — Typically recent (hours to weeks) physical evidence of the (approximate) maximum flood stage (Jarrett and England, 2002). The physical evidence generally is of three types: (1) deposits along channel margins and in vegetation that consist of very light, floatable material such as pine needles, seeds, small twigs, grasses, and very fine sediments; (2) damage to vegetation such as bent or matted grasses, twigs, and branches, stripped leaves or bark; and (3) small erosional features such as scour lines. Benson and Dalrymple (1967) discuss identification and rating of high-water marks. The HWM evidence is typically short-lived (weeks), but woody debris may last from several years to several decades in arid and semi-arid climates (Baker, 1987), and geomorphic evidence can be preserved for millennia. Physical evidence, such as marks on buildings and other structures, is also long-lived. This stage may represent a maximum discharge when a single-valued relationship exists between stage and discharge; Costa and Jarrett (2008) describe other hydraulic situations. See also *paleostage indicator*.

historical data — Broad category of data collected by humans prior to establishing systematic protocols; it generally consists of diaries, written accounts of settlements, folklore, and descriptions that may document periods where extreme weather and/or floods have occurred. It may also be used to infer times when there have been no large floods. These accounts were recorded in a manner that was preserved well enough that we know about it today.

historical floods — Flood events which were directly observed by humans, generally in a non-systematic manner by non-hydrologists (Baker, 1987). These events usually occurred and were described in some qualitative and/or quantitative fashion prior to the systematic record. Information about the floods was recorded and preserved well enough so that we know about it today.

homogeneity — Records from the same populations. Floods may be from different populations because they occurred before the building of a dam and after the building of a dam, or before the watershed was urbanized and after it became urbanized, or because some are generated by summer storms and others by snowmelt, or because some were generated in El Nino years and some were in other years. It may be difficult in some cases to definitively say if the flood record is homogeneous.

incomplete record — A streamflow record in which some peak flows are missing because they were too low or high to record or the gage was out of operation for a short period because of flooding.

interval data — Floods whose magnitude are not known exactly, but are known to fall within a range or interval (Stedinger et al., 1988; Cohn et al., 1997)

level of significance — The probability of rejecting a hypothesis when it is in fact true. At a "10-percent" level of significance the probability is 1/10.

low outlier — See *outlier*.

mean-square error — Sum of the squared differences between the true and estimated values of a quantity divided by the number of observations. It can also be defined as the bias squared plus the variance of the quantity (Stedinger et al., 1993).

method of moments — A standard statistical computation for estimating the parameters of a distribution from the moments of the sample data.

Multiple Grubbs-Beck Test (*MGBT*) — A statistical test used to identify multiple potentially-influential low flood observations in an annual maximum time series.

nonexceedance — Knowledge that the magnitude (discharge or stage) of a flood was less than some level or threshold.

outlier — Outliers (extreme events) are observations that are exceedingly low or high compared to the distributional properties of the vast majority of the data. When plotted, along with a resonably fitted CDF to the data, the outlier values plot far from the fitted line at the low or high ends of the distribution. A CDF, such as the LP-III, may not fit data sets with outliers, and the fitted curve usually fails to fit the bulk of the data as well as the outliers. Low outliers are outliers at the low end of the data set, near zero, at least in comparison with the rest of the data. On a log-probability plot, the low outliers impart a strong downward curvature and a downward-drooping lower tail to the frequency curve. In comparison with the lower tail, the upper tail of the low-outlier-affected curve may appear relatively flat.

paleoflood data — Physical evidence of past floods and their ages as observed from the geologic record or from botanical evidence. Paleoflood data typically consists of observations on individual past floods such as those derived from slackwater deposits, boulder bars, silt lines, or botanical information, that are collected as part of a paleoflood hydrology study (Benito and O'Connor, 2013). It can also consist of periods of landscape stability that can be used to place limits on flood magnitude over time, such as paleohydrologic bounds (Levish, 2002). Paleoflood data are distinguished from historical flood data, as a separate line of evidence, by the use of applied field geology techniques to examine and describe the geomorphic and stratigraphic context of extreme floods. In some cases, there is overlap between historical and paleoflood data, as historical and cultural artifacts such as barbed wire, beer cans (House and Baker, 2001) or pottery may be observed and used in dating and estimation of floods.

paleoflood hydrology — The study of past or ancient floods which occurred prior to the time of human observation or direct measurement by modern hydrologic procedures (Baker, 1987). Paleoflood hydrology has also been defined as "the study of the movements of water and sediment in channels before the time of continuous hydrologic records or direct measurements" (Costa, 1986).

paleohydrologic bound — A time interval during which a given discharge has not been exceeded (Levish, 2002). The term is sometimes shortened to *bound*. The paleohydrologic bound represents stages and discharges that have not been exceeded since the geomorphic surface stabilized. Bounds are appropriate for paleohydrologic information and are not dependent on human observation of a particular event, but on the physical setting (hydraulic and geomorphic). (alt: *paleoflood bound*).

paleostage indicator (PSI) — An erosional or depositional feature that recorded the near peak stage of an individual flood (Jarrett and England, 2002) prior to human observation. Indirect evidence of the stage of past floods includes botanical evidence and sedimentological deposits (Jarrett, 1991). Large floods, especially in high gradient channels, can transport and deposit coarse material (gravel, boulders, and woody debris, etc.) that may be interpreted as HWMs. The primary differences between PSIs and HWMs are: (1) HWMs represent events which occurred "more recent" in time due to their relatively short preservation length as compared to PSIs; and (2) some PSIs may not represent the exact peak stage.

Partial-Duration Series (PDS) — A list of all flows (such as flood peaks) that exceed a chosen base stage or discharge, regardless of the number of peaks occurring in a year. Also called basic-stage flood series, or floods above a base (Langbein and Iseri, 1960).

percent chance — A probability multiplied by 100.

perception threshold — The stage or flow above which it is estimated a source would provide information on the flood peak in any given year. Perception thresholds ($T_{Y,lower}$; $T_{Y,upper}$) reflect the range of flows that would have been measured/recorded had they occurred. If an event magnitude had occurred in a specific year, there is information indicate it would have been "recorded" in a manner that we could perceive it today. Perception thresholds describe the range of measurable potential discharges and are independent of the actual peak discharges that have occurred. They are used to provide a rank and record length for each reported flood peak (Gerard and Karpuk, 1979). Perception thresholds are used for historical data, when the information provided is based on human observation. They are also used to describe a paleoflood period and paleoflood data. In addition, perception thresholds are used to properly accommodate unrecorded floods below a “gage base”. A perception threshold is allocated to each information source for each year Y of the flood record. Perception thresholds may involve a significant amount of judgment on the part of the scientist and/or historian regarding, for any given year, what is the smallest event that would have been recorded (in a physical or textural manner) such that we would actually know about it today (alt: *threshold*).

population — The entire (usually infinite) number of data from which a sample is taken or collected. The total number of past, present, and future floods at a location on a river is the population of floods for that location even if the floods are not measured or recorded.

potentially-influential low flood (PILF) — In an annual maximum flood series, small-magnitude flows (including zeros) that do not represent the physical processes that cause the largest flood observations. These “PILFs” can exert high leverage and influence on the flood frequency distribution.

quantile — Estimate of the flood magnitude Q for exceedance probability p from a fitted distribution.

record augmentation — A procedure to improve the accuracy of the moments (mean and variance) of a short-record flood series by using information from longer records at nearby locations with high cross-correlation (Matalas and Jacobs, 1964; Stedinger et al., 1993).

record extension — The creation of a longer flood-flow record (individual floods) at a site with a short record, by using flood observations at a long-record site with high cross-correlation. The technique can also be used to fill in missing observations (Hirsch et al., 1993; Stedinger et al., 1993).

regional skew coefficient — A skew coefficient derived by a procedure which integrates values obtained at many locations.

robustness — In flood frequency, a procedure that is reasonably efficient when the assumed characteristics of the flood distribution are true, while not doing poorly when those assumptions are violated (Kuczera, 1982; Cohn et al., 2013).

sample — An element, part, or fragment of a “population.” Every hydrologic record is a sample of a much longer record.

serial correlation — See *autocorrelation*.

skew coefficient — See *coefficient of skewness*.

standard deviation — A measure of the dispersion or precision, of a series of statistical values such as precipitation or streamflow. It is the square root of the sum of squares of the deviations from the arithmetic mean divided by the number of values or events in the series. It is standard practice to divide by the number of values minus one in order to get an unbiased estimate of the variance from the sample data.

standard error — An estimate of the standard deviation of a statistic. Often calculated from a single set of observations. Calculated like the standard deviation but differing from it in meaning.

systematic data — Data that are collected at regular, prescribed intervals under a defined protocol. In the context of streamflow, systematic data consist of discharge and stage data collected at regular, prescribed intervals, typically at gaging stations. (syn. systematic record).

threshold — See *perception threshold*.

variance — A measure of the amount of spread or dispersion of a set of values around their mean, obtained by calculating the mean value of the squares of the deviations from the mean, and hence equal to the square of the standard deviation.

weighted means — A value obtained by multiplying each of a series of values by its assigned weight and dividing the sum of those products by the sum of the weights.

Prepared by John F. England, Jr., William H. Asquith and the USGS Publishing Service Center
Edited by USGS
Illustrations by USGS
Layout and typesetting by John F. England using LaTeX
Final portable document format preparation by USGS

For information regarding this report , contact:

U.S. Geological Survey
415 National Center
12201 Sunrise Valley Drive
Reston, Virginia 20192
(703) 648-5301
http://water.usgs.gov/osw/