Flood Frequency Analysis in the United States: Time to Update

Jery R. Stedinger, Ph.D., M.ASCE
Professor, School of Civil and Environmental Engineering, Cornell Univ., Hollister Hall, Ithaca, NY 14853-3501. E-mail: jrs5@cornell.edu

Veronica W. Griffis, Ph.D., M.ASCE
Assistant Professor, Department of Civil and Environmental Engineering, Michigan Technological University, 1400 Townsend Dr., Houghton, MI 49931-1295. E-mail: vgriffis@mtu.edu

Introduction

A large portion of the U.S. population, infrastructure, and industry is located in flood-prone areas. As a result, floods cause an average of nearly 140 deaths and cost roughly $6 billion annually excluding flooding caused by Hurricane Katrina, which cost $200 billion alone (www.usgs.gov/hazards/floods/; USGS 2006). The 1993 Midwest flooding along the Mississippi and Missouri Rivers caused $20 billion in damages. Furthermore, these estimates neglect the real costs associated with loss of personal possessions and shattered lives and communities.

With Hurricane Katrina (August 2005) and recent floods in the Northeast (June 2006), Ohio (winter 2007), Texas–Oklahoma (June 2007), and central England (July 2007), it seems like damaging floods have been in the news daily over the last few years. While engineers cannot stop floods from occurring, they should seek structural and nonstructural strategies to reduce the risk of large economic losses, social vulnerability, environmental damage, and loss of life (IFMRC 1994). Development of economically efficient and rational plans requires good estimates of the risk of flooding. In the United States, that computation is done following guidelines in Bulletin 17, for which the latest update, Bulletin 17B, was published in 1982 (IACWD 1982). That update includes the skew map published 30 years ago in the original Bulletin 17 (WRC 1976), and a list of areas needing additional research. Bulletin 17B has served the nation for over 30 years; it is a remarkable document that has withstood the test of time and use. However, given long-standing problems listed in the document, recent advances that address those problems, and the current national interest in flood risk, the time has arrived to update Bulletin 17B to maintain the statistical credibility of the guidelines and to provide accurate risk and uncertainty assessments.

Background

Accurate estimates of the magnitude and frequency of flood flows are needed for the design and operation of water-use and water-control projects, for floodplain definition and management, and for the design of transportation infrastructure such as bridges and roads. Floods and flooding are not a new problem in the United States. In the mid-1960s, it became apparent that uniform flood frequency methods were needed for the development of a national flood insurance program in the United States, as well as to facilitate coordination among government agencies and members of the private sector who participate in the management of water resource systems that affect flood risks.

The first step toward development of uniform flood-frequency techniques for federal agencies was the publication in April 1966 of Bulletin 13 “Methods of Flow Frequency Analysis” by the Interagency Committee on Water Resources (ICWR 1966); it summarized methods commonly used by U.S. federal agencies. Bulletin 15 was quickly followed by Bulletin 15 “A Uniform Technique for Determining Flood Flow Frequencies” published in December 1967 (WRC 1967); this was the first set of uniform flood frequency techniques to be employed by all federal agencies. Bulletin 15 recommended the use of the log-Pearson type 3 (LP3) distribution with a regional skew. However, uniform procedures were not specified for the treatment of low outliers or the estimation of the regional skew, and no recommendation was made for the use of historical data.

In the years following, methods were developed to address the use of regional skew information, the use of historical flood information, the identification of low outliers, and the adjustment of the flood frequency curve when low outliers were identified (Thomas 1985). These methods were formally adopted with the publication of Bulletin 17 “Guidelines for Determining Flood Flow Frequency” in March 1976, which was followed quickly by Bulletin 17A published in June 1977 (WRC 1977). Bulletin 17 was last updated with the publication of Bulletin 17B in March 1982 (IACWD 1982).

Bulletin 17B recommends using the method-of-moments (MOM) to fit a Pearson type 3 (P3) distribution to the logarithms of the flood series, thereby yielding a log-Pearson type 3 (LP3) distribution to model observed streamflow data. Estimates of the mean, standard deviation, and skew coefficient of the logarithms of the sample data are computed using traditional moment estimators. However, because the data available at a site are generally limited to less than 100 years, and are often less than 30 years, the skewness estimator can be particularly unstable. To address that concern, Bulletin 17B wisely suggests the at-site skew be weighted with a regional skewness estimator, where the recommended weights are inversely proportional to the precision of each estimator. Bulletin 17 included a regional skew map for the United States originally developed by Hardison (1974). On the back of the map reads the statement: “It is expected that Plate I [the map] will be revised as more data become available.” And in 1982 the following was added: “This generalized skew map was originally prepared for Bulletin 17 published in 1976. It has not been revised utilizing the techniques recommended in Bulletin 17B.”

Twenty-five years after the last revision of Bulletin 17B (IACWD 1982), that publication remains the methodology recommended for flood frequency analyses by U.S. federal agencies.
Despite important advances in hydrology and flood frequency analysis during the last two decades (see Griffis and Stedinger 2007a), the *Bulletin* has not been updated, nor has the skew map published in the original 1976 version of *Bulletin 17* seen an official revision. Some USGS district offices have developed regional skew estimators for their states using a wide range of statistical and ad hoc approaches (Z. Song-James and W. Thomas, personal communications, August 15, 2007). This approach has resulted in a set of maps that employed inconsistent and often inappropriate statistical procedures (Stedinger 2005). The time for a revision of *Bulletin 17B* and the associated regional skewness estimator has arrived.

**What Is Next?**

There is currently an interest among federal agencies, including the U.S. Geological Survey, the U.S. Army Corps of Engineers, and the U.S. Bureau of Reclamation, to revise the guidelines for flood frequency analysis in the United States contained in *Bulletin 17B* through the interagency Hydrologic Frequency Analysis Work Group (HFAWG) [England and Cohn (2007, http://acwi.gov/hydrology/Frequency/index.html)]. Recent studies make clear the advantage of regional skew information when weighted appropriately, and of alternative flood frequency estimation procedures that U.S. federal agencies might adopt (Stedinger and Griffis 2006; Griffis and Stedinger 2007a). For the range of regional-average flood statistics observed across the United States, Griffis and Stedinger (2007b) demonstrate that the LP3 distribution with a log-transformation of the data as recommended by *Bulletin 17B* is certainly a reasonable and flexible model of flood risk. Within this limited range of skew values, the LP3 distribution is more flexible than a two-parameter log-normal distribution, but should not deviate too far from that physically reasonable two-parameter lognormal model. Further, with two shape parameters, the LP3 density function is able to assume many shapes and covers a two-dimensional space in the L-moment ratio diagram, including negative values of the L-skewness ratio. Griffis and Stedinger (2007c) present Monte Carlo results that demonstrate that the log-space method-of-moments estimator recommended by *Bulletin 17B* is robust and performs well when employed with an informative regional skew, and is competitive with several alternatives including maximum likelihood estimators that employ regional skew information. Monte Carlo results presented by Griffis and Stedinger (2007d) illustrate the value of different weighting schemes for combining at-site and regional skew information. For reasonable values of the log-space population skew γ, weights developed in that study that provide the minimum MSE quantile estimators were found to yield only modest improvements in the MSE of quantile estimates over the MSE-skew weight recommended by *Bulletin 17B*.

**What Future Work Was Recommended?**

*Bulletin 17B* (pp. 27–28) includes a list of issues recommended for additional study, which can be summarized as follows:

1. flood-frequency distribution selection and fitting procedures;
2. identification and treatment of mixed distributions;
3. identification and treatment of outliers;
4. treatment of historical information;
5. confidence limits for the Pearson type 3 distribution;
6. use of precipitation in estimates of flood potential;
7. estimation of flood potential for ungauged sites and watersheds with limited records; and
8. estimation of flood potential for watersheds altered by urbanization and reservoirs.

Research and proposed revisions discussed below address Issues 1, 3–5, and perhaps 2.

With respect to issue 1, recently suggested revisions would also address recommended procedures for estimating regional skew. *Bulletin 17B* on pp. 10–12 has an extended discussion of how regional skews should be developed. It then recommends that, “In the absence of detailed studies, the generalized skew can be read from Plate I [the map] found in the flyleaf pocket of this guide.” Hardison (1974) describes development of that map, which has been the default for flood studies over the past 30 years. Given the current availability of an additional 30 years of data, significant and important advances for the treatment of such imprecise spatial information (Reis et al. 2005; Gruber et al. 2007), and changes in the low outlier identification procedure and in the way low outliers are handled (IACWD 1982, p. 12), it is clearly time to update the *Bulletin 17B* guidelines for the statistical analysis of regional skewness information and to provide a replacement for the 1976 skew map. A new regional skew estimator can employ basin characteristics as well as gauge location.

**Most Critical Issues to Address**

The most critical of the concerns discussed above, related to Issues 1 and 3–5, are:

1. statistically appropriate procedures for the computation of the regional skew and its precision recognizing that sample skewness estimators are relatively inaccurate themselves;
2. statistically effective and flexible methods for employing historical flood data and nonstandard measurements in flood frequency investigations;
3. consistent and straightforward treatment of low outliers, zero flows, and related problems; and
4. computation of confidence intervals for quantiles reflecting uncertainty in the skewness coefficient, and other appropriate descriptions of the statistical uncertainty in estimated quantiles, annual exceedance probabilities, and expected flood damage reduction from flood risk reduction projects.

Other issues that can be addressed in the future, related to Issues 6–8, include:

5. generation of flood records for frequency analysis from rainfall records employing conceptually based and spatially distributed watershed models and computation of associated measures of uncertainty;
6. simulation of reservoir system performance and regulated flows to support regulated-flow frequency analyses and computation of corresponding uncertainty measures; and
7. development of frequency distributions for very extreme floods needed for dam safety studies reflecting realistic physical limits on precipitation and flow in a given basin and guidelines such as the probable maximum flood and probable maximum precipitation.

**Summary of Recent Research**

This section summarizes a series of studies that evaluate *Bulletin 17B* procedures and possible extensions. Griffis and Stedinger (2007a) provide a detailed review of the motivation and history of the procedures in *Bulletin 17B*, as well as additional citations to new results.
Regional Skew

Bulletin 17B recommends weighting the station skew with a regional skew. Three procedures are recommended for estimating the regional skew, or as an alternative, a skew map provided in the Bulletin may be used to easily obtain a regional skew estimate. Griffiths and Stedinger (2007a) discuss discrepancies as to how the low outlier adjustment procedures were employed when the map was created, as well as inconsistencies that result because the skew map was not updated following changes in the low outlier identification procedures adopted with the publication of Bulletin 17B. Moreover, some studies question whether the use of a regional skew map is appropriate (e.g., Landwehr et al. 1978; McCuen 2001).

The skew map was constructed using data through water year 1973 and has yet to be updated, despite more than 30 years of additional data now available and the availability of powerful new statistical methods described by Reis et al. (2005), Gruber et al. (2007), and citations therein. Several U.S. Geological Survey studies have employed weighted least-squares regression as recommended by Tasker and Stedinger (1986) to estimate a regional skew (e.g., Rasmussen and Perry 2000; Pope et al. 2001). Those studies suggest the mean square error of good regional skew models should be 0.10 or less (corresponding to an effective record length of 60 or more years), and not 0.30 (corresponding to an effective record length of 17 years) as suggested by the skew map in the Bulletin. Given that typical record lengths are 15 to 70 years, the differences in the computed effective record length are very important in the computation of a weighted skew. Griffiths et al. (2004) and Griffiths and Stedinger (2007d) illustrate the value of an informative regional skew.

Low Outlier Adjustments and Use of Historical Information

To further refine and improve parameter estimates, the Bulletin recommends using low outlier adjustment procedures and historical information. Problems arise because the procedures for the use of historical information are performed separately from the procedures for low outlier identification and the subsequent adjustment of the frequency curve. Thus the identification of outliers and the moments of the final fitted P3 distribution are dependent on the order in which these procedures are employed.

The Bulletin provides a framework in which a conditional probability adjustment (CPA) can be used to adjust for low outliers and zero flood years that are not included in the computation of sample moments. Because an individual sample can contain both high and low outliers, the outlier identification threshold depends on sample moments, and the procedures for adjusting for low outliers and incorporating historical information are conducted separately, Bulletin 17B provides recommendations as to the order in which these procedures should be executed. The recommended steps are illustrated on page 12–3 of Bulletin 17B (also see the interpretation suggested by Griffiths and Stedinger 2007a).

Bulletin 17B employs three separate procedures to reflect historical information, to account for censored low outliers, and to introduce regional skew. Alternatively, the recently developed expected moments algorithm (EMA) for the LP3 distribution (Griffiths et al. 2004; England and Cohn 2007) combines these three steps into one consistent analysis. Furthermore, EMA makes much more effective use of historical information than the Bulletin 17B procedures as demonstrated by Cohn et al. (1997) and England et al. (2003a, b). If low outliers are identified using the moments of a systematic record, as currently recommended by Bulletin 17B, or by other means, then an EMA analysis can incorporate historical information into the analysis while representing the low outliers as censored observations that are known to be smaller than the smallest observation retained (Griffis et al. 2004; England and Cohn 2007).

It is also important to observe that the Bulletin 17B procedure for employing historical information requires that historical floods be represented by a single point estimate. Unfortunately this is not always possible. Historical floods are often defined with varying precision (England et al. 2003b; England and Cohn 2007). For example, a realistic and common case is that some threshold was NOT exceeded over an extended period. The Bulletin 17B historical flood adjustment does not make use of such information, whereas the new EMA algorithm uses threshold-exceedance and interval information very effectively.

Plotting Positions

Hirsch and Stedinger (1987) discuss problems with the plotting positions recommended in Bulletin 17B for use with historical flood information. They provide a needed generalization useful for flood records with multiple thresholds representing different historical periods, or historical information and low outliers.

Confidence Interval

Bulletin 17B recommends obtaining confidence intervals for flood discharges with specified exceedance probabilities using the noncentral t-distribution, ignoring error in the skew estimator. Stedinger (1983) demonstrates that these formulas often fail to attain the desired confidence level for skews other than zero when the skew is known, and provides the needed correction for that case. Chowdhury and Stedinger (1991) extend the formulas developed by Stedinger (1983) for the case where the skew is estimated by a weighted skewness estimator as employed by Bulletin 17B. They provide Monte Carlo results that indicate that the performance of the intervals is improved when the error in the skew is accounted for correctly.

Another issue is how to appropriately formulate confidence intervals when historical information is employed and probability adjustments are performed to account for low outliers, zero flood years, and missing observations due to a recording threshold. Cohn et al. (2001) and the extension proposed by Griffiths et al. (2004) provide formulas for the needed confidence intervals when low outliers and zero flood years are censored, and regional skew and historical flood information are employed. The Bulletin does not provide procedures to address such complex cases.

Stedinger and Griffiths (2006) report a Monte Carlo experiment that demonstrates that 90% confidence intervals for the 100-year event that ignore skew uncertainty, such as those in Bulletin 17B, do not achieve the desired level of significance. In generated samples of size 25, 50, and 100 with a population skew $\gamma=0$, the 90% confidence intervals actually contained the target quantile only 75 to 81% of the time when the mean square error of the regional skew is 0.302, and 81 to 85% of the time when the mean square error is 0.100. On the other hand, Cohn et al. (2001) showed that their properly constructed confidence intervals reflecting skew uncertainty included the true 100-year flood value 89 to 91% of the time for $\gamma=\pm 0.1$. Thus the needed methods to describe uncertainty correctly are available.

Weighting of Independent Quantile Estimates

Griffiths and Stedinger (2007a) evaluate the recommendations in Appendix 8 of Bulletin 17B to weight independent at-site and regional quantile estimates using the inverse of their respective variances as the weights. Unfortunately, the simple quantile
weighting may not yield the most accurate quantile estimate because the error in the at-site estimator is dominated by error in the sample standard deviation and skew, whereas the error in the regional estimator can be dominated by the error in the mean. Using data from South Carolina, Griffis and Stedinger (2007a) compare the standard and weighted quantile estimators in Bulletin 17B to reasonable LP3 alternatives that also make use of regional information. The alternatives considered included an index flood estimator and moment-weighted estimators that employed various combinations of at-site, regional, and weighted moment estimates. The simple weighting of at-site and regional quantile estimates recommended by Bulletin 17B performs nearly as well as more complex alternatives, and for short records provides a substantial improvement in quantile accuracy. However, when the regional standard deviation and skew are very informative, more accurate estimates would be obtained using a moment-weighted estimator wherein separate weighted estimates of the mean, standard deviation, and skew are employed.

**Why Use the LP3 Distribution with Bulletin 17 Fitting Methods**

Vijay Singh, the editor-in-chief, has challenged us to support the continued use of the LP3 distribution with the log-space moment estimators recommended in Bulletin 17B. He observed that “Bulletin 17B should be looked at a fresh, not simply updated, if we want to take full advantage of advances in the areas that the authors are emphasizing.” Bobee and Ashkar (1991, p. 76) observe that since the official adoption of the LP3 distribution in the United States and Australia, “its application to the study of floods has been both extensive and widespread.” Still a concern is whether the adopted LP3 distribution with log-space moments is a good choice.

It is perhaps unfortunate that a comprehensive study of Bulletin 17B procedures is not being pursued. The first constraint is money. While the current reexamination of Bulletin 17B by the Hydrologic Frequency Analysis Work Group began in late 2005, there has been no specific funding for the effort. Progress depends on meager internal research funds available within the U.S. Geological Survey, the U.S. Bureau of Reclamation, and the U.S. Army Corps of Engineers, along with time and effort donated by individuals and other organizations that participate in the HFAWG. The recent efforts of HFAWG and the revisions currently planned are commensurate with available funding. More comprehensive analyses and subsequent revisions may occur in the future as more funding becomes available. A major effort to reexamine the range of flood distributions and fitting methods that could be adopted across the entire United States and its territories with their different hydrologic and climate characteristics would be a very large undertaking.

It is also not clear that there would be great benefits from a major change in the Bulletin 17-LP3 flood frequency framework. In general, the differences in quantile estimators that result from fitting reasonable alternative flood distributions (3-parameter lognormal, LP3, or Generalized Extreme Value (GEV) distributions) are substantially less than the uncertainty in the flood quantile estimators themselves [e.g., see Hosking and Wallis (1997, pp. 134–38, 142); or Stedinger (1980, p. 488)]. And the true distribution will never be known. Thus we assert that the important issue is to use a reasonable distribution consistent with available data, and to fit that distribution as well as possible with one’s understanding of flood processes and the at-site and regional data that can be collected. Griffis and Stedinger (2007b) show that the LP3 distribution is a very reasonable and flexible model of flood risk within the range of parameter values consistent with U.S. flood series. In addition, Bulletin 17B’s use of a low outlier detection and censoring algorithm allows the fitting procedure to focus on the distribution of the events of interest (large floods) with protection that a long lower tail will not distort our description of the distribution of the larger events (Griffis et al. 2004).

Extensive research into fitting methods throughout the 1980s and 1990s showed that the index flood method that fit one parameter with at-site data works best with very short series; whereas for typical flood records with 25 or more observations, the best estimators are often obtained using a regional shape parameter with at-site location and scale estimators [see Stedinger and Lu (1995); Hosking and Wallis (1997, pp. 148–50); Griffis et al. 2007a; and citations therein]. Bulletin 17B does indeed use at-site (log-space) location and scale parameters, combined with a weighted average of a regional shape parameter and the at-site skew that reflects the precision of each. Griffis and Stedinger (2007c) show that the Bulletin 17B-LP3 estimator with weighted skew becomes more precise when a more precise regional skew is available, which is again motivation for developing better regional skew estimators with realistic error variance estimators.

Hosking and Wallis (1997, pp. 150–53) discuss studies by Wallis and Wood (1985) and Potter and Lettenmaier (1990) that compared the Bulletin 17B flood quantile estimators with index flood procedures, from which the overall conclusion was that index flood procedures were much more precise. However, these studies were not without their limitations. Problems with the Wallis and Wood study are discussed by Beard (1987) and Landwehr et al. (1987). In particular, the hydrologic regions Wallis and Wood developed would be very unrealistic if it were not for the fact that so little variation is introduced into the moments: in real world basins, the real-space coefficient of variation and skew increase together in a predictable fashion, as shown in Stedinger and Lu (1995, Fig. 5) where \( k \) determines the GEV skew. Potter and Lettenmaier report an extensive “bootstrap” experiment with real flood data from Wisconsin and New England. The index flood and Bulletin 17B procedures were compared based on the standardized standard deviation of the quantile estimators for each site. However, there is a trade-off between bias and variance that should be considered when choosing between an index-flood estimator that has small variance but potentially large bias, and a 2- or 3-parameter estimator such as the Bulletin 17B-LP3 estimator [see Hosking and Wallis (1997, Fig. 7.3 and 7.5)]. Potter and Lettenmaier could not evaluate the bias of their index flood estimators because they used real data whose true distribution is not known. Moreover, Bulletin 17B in its Appendix 8 includes a procedure for weighting regional estimates of flood quantiles and at-site LP3-quantile estimators. Griffis and Stedinger (2007a) illustrate the use of several regional at-site estimation procedures with the LP3 distribution; in particular, for sites with short records, the Bulletin 17B Appendix 8 weighted-quantile estimator is more precise than a LP3 index flood estimator, and both do better than the Bulletin 17B at-site LP3 estimator with a weighted skew.

Finally, the most important reason for staying with the Bulletin 17B procedure may be practical considerations. The U.S. government adopted the LP3 distribution in 1967, more than 40 years ago. Since that time, thousands of flood control structures have been designed and flood maps drawn based on use of the Bulletin’s recommendations. Federal personnel and consultants across the country are comfortable with those procedures. And people understand that as more data are collected, the Bulletin 17B flood...
risk estimators will evolve to reflect that data. A change is reasonable if new methods consistent with the old framework allow better use of historical information and threshold-exceedance data. However, the agencies have no enthusiasm for changing their flood risk estimates because a different distribution is adopted, or estimation is now to be done using L-moments. For those property owners adversely affected, a radical change in the flood estimation procedures that is not justified by a documented large gain in precision would seem unfair and arbitrary. The reality is that federal agencies need credible flood risk descriptions that will be accepted bureaucratically, politically, and legally. Bulletin 17B has earned that acceptance. What it needs is to be updated so that it can perform better with the range of data available. Given the evidence currently available, there is no obvious need to change the paradigm.

Summary

Bulletin 17B is a remarkable document containing procedures that address a wide range of situations that arise in practice. While those procedures have survived the test of time and use, the time has arrived to update low outlier and historical flood procedures, plotting positions, and confidence interval computations with more efficient and consistent procedures. It is also recommended that the regional skew map be updated to use the additional 30 years of data now available, to appropriately adjust for low outliers identified in the samples used to estimate the regional skew, and to use new and powerful statistical estimation procedures developed to use such data sets. The original map itself anticipated, if it did not recommend, such updates.

In addition, the U.S. flood management community should adopt the Expected Moments Algorithm (EMA). It provides a direct fit of the LP3 distribution using the entire data set, simultaneously employing regional skew information and a wider range of historical flood and threshold-exceedance information, while adjusting for any low outliers, missing values from an incomplete record, or zero flood years. Furthermore, adoption of EMA would allow for the use of a conceptually rigorous and accurate procedure for computing confidence intervals for quantiles computed with a given data set. There is no need to change the basic rules of the frequency analysis by use of a new distribution, or to adopt a radically different fitting procedure. The envisioned revisions address weaknesses and problems pointed out in Bulletin 17 in 1976, and for which we now have the required algorithms to incorporate historical information into the moment estimators, to correct for low outliers, and to compute confidence intervals with the target level of confidence.

A future update of Bulletin 17B might also consider procedures for regulated streams, incorporation of the impacts of urbanization, and the validity of the assumption that annual maximum floods are stationary in light of observed land-use change and identified trends in streamflows. How those issues should be addressed needs to be resolved. [See Garbrecht and Piechota (2006); Griffiths and Stedinger (2007e); and citations therein.] However, for a wide range of important and commonly occurring cases, better and well tested methods are now available. Now is the time to adopt those procedures for the traditional analyses addressed by Bulletin 17.

Acknowledgments

The authors appreciate the valuable comments and encouragement provided by Vijay Singh (Editor, Journal of Hydrologic Engineering), John England (USBR), Timothy Cohn (USGS), Will Thomas (Michael Baker, Jr., Inc.), and Beth Faber (U.S. Army Corps of Engineers).

References


