Summary of Gaging Station Data and Testing Methods for the EMA and Bulletin 17B Comparisons

Introduction

A data base of 80 gaging stations with systematic record lengths ranging from 37 to 110 years (generally 70 to 90 years) has been compiled (peakdata.txt). The geographic distribution of the 80 stations is shown in Figure 1. The gaging stations were selected primarily from the USGS HydroClimatic Data Network and the streamflow and watershed characteristics are not significantly affected by urbanization, deforestation, regulation, etc. These data will form the basis for the testing of Bulletin 17B and EMA methods using **observed** flood data. Some theoretical probability distributions will also be provided for testing with Monte Carlo simulation. The testing will be accomplished by Tim Cohn, U.S. Geological Survey, and John England, Bureau of Reclamation.

Description of Gaging Station Data

A spreadsheet provided as a separate file (**Data For Comparisons_08-06-07.xls**) summarizes useful information on the 80 stations such as the station name and number, drainage area, period of record and comments about unusual aspects of the data set.

The spreadsheet also contains the following information that will be needed in the EMA-Bulletin 17B testing:

Historic period in years – the length of time for which the largest peaks (historic and/or high systematic) in the data set are the highest. This is generally the time from the earliest historic peak in the data set or the time based on the "Highest Since" in the USGS NWIS File. The historic period includes the systematic record plus additional time before the systematic record began.

Historic peaks – those floods that occurred outside the systematic data collection period that are identified in USGS NWIS File with a code 7.

High systematic peaks – floods that occurred during the systematic record that are very large with respect to other systematic peaks (qualitative judgment). These floods are so large that they are clearly the highest in a period greater than the systematic record period; however, no historic information may be available. In some cases, the high systematic peaks were identified as high outliers by the Grubbs and Beck test but many times they are less than the computed threshold yet still 2-3 times the next highest annual peak.

Systematic record in years – the number of years of systematic record, that is, the years that the USGS was operating the station.

Low outlier threshold calculated in cfs – the calculated Grubbs and Beck low outlier threshold computed according to Bulletin 17B guidelines.

Peaks below low outlier threshold in cfs – a listing of the actual peaks that are below the calculated low outlier threshold.

High outlier threshold calculated in cfs – the calculated Grubbs and Beck high outlier threshold computed according to Bulletin 17B guidelines.

Historic threshold user input in cfs – the historic threshold value determined by judgment using the date of the earliest historic peak and the "Highest Since" data in the USGS NWIS File.

The gaging stations can be grouped into the following categories according to the occurrence of historic peaks, high systematic peaks, low outliers and some combination:

Stations with historic peaks = 26
Stations with low outliers based on calculated Grubbs and Beck test = 27
Stations with historic peaks and high systematic peaks = 13
Stations with high systematic peaks and no historic period = 10
Stations with historic peaks and low outliers = 9
Stations with historic peaks, high systematic peaks, and low outliers = 7
Stations with no historic peaks or outliers = 36

The results from the testing should be summarized according to the groups described above to determine if EMA and Bulletin 17B perform differently for each group.

Testing on Observed Data

Bulletin 17B and EMA methods will be applied to split samples for the 80 gaging stations. This testing will consist of:

- Selecting random samples of size 20, 40, 60 years, etc. (up to two thirds of the record length) from the full systematic record for each station. The full systematic record length will also be analyzed. If historical peaks are available, they will be added to the selected split samples. The time period from the first historic peak to the last systematic peak selected will constitute the historic period. Use part of the record (20, 40, 60 years, etc.) to estimate the frequencies and then observe exceedances of selected events in the other portion of the record and compared to the computed values. Use the expected probability exceedances approach as described in Appendix 14 of Bulletin 17B (pages 14-9 and 14-10) and in Beard's 1974 analysis.
- Summarizing the 2-, 10-, 50-, 100- and 500-year estimates for Bulletin 17B and EMA for the split sampling analyses in a table.
- Relating the relative differences for the two methods to particular conditions of the observed data sets as identified above (presence of high and low outliers, historic floods, etc.).

Monte Carlo Testing

A second part of the testing will be to perform Monte Carlo simulations. This method assumes that the **simulated** flood data comes from a particular frequency distribution, and repeatedly samples from it for some number of years of record, and fits an LPIII by Bulletin 17B and EMA methods. Beth Faber, USACE, provided an excel spreadsheet with six different frequency curves that are recommended for Monte Carlo testing. The excel spreadsheet is provided as a separate file from this document (**test curves3.xls**).

The six theoretical frequency curves are described in the separate spreadsheet and summarized below:

Curve 1 – log Pearson Type III curve with positive skew
Curve 2 – log Pearson Type III curve with negative skew
Curve 3 – Pearson Type III (no logs) with positive skew
Curve 4 – Mixed distribution of two log Pearson Type III distributions, both with positive skew
Curve 5 – Concatenation of two log Pearson Type III distributions, both with negative skew
Curve 6 – Concatenation of two GEV distributions, one with negative skew and one with positive skew

The testing would follow the gaging station analysis with repeated samples of 20, 40, 60 years, etc. and various historical periods done with Monte Carlo simulations. The "historic peaks" would be simulated by generating a large sample (greater than 100) and then randomly selecting samples of size 20, 40, 60 years, etc. The historic period for the largest peaks in the selected sample would be determined by reference to the large sample. The performance of EMA and Bulletin 17B would be compared using mean square error and bias as described by John England in his March 6, 2006 document entitled "Hydrologic Frequency Analysis Work Group Plans to Investigate Possible Improvements to Bulletin 17B" (separate document). Tim and John will provide a summary of their split sampling analyses and justify any decisions made as part of these analyses.

Will Thomas
Michael Baker, Jr.
July 31, 2007
Revised August 8, 2007

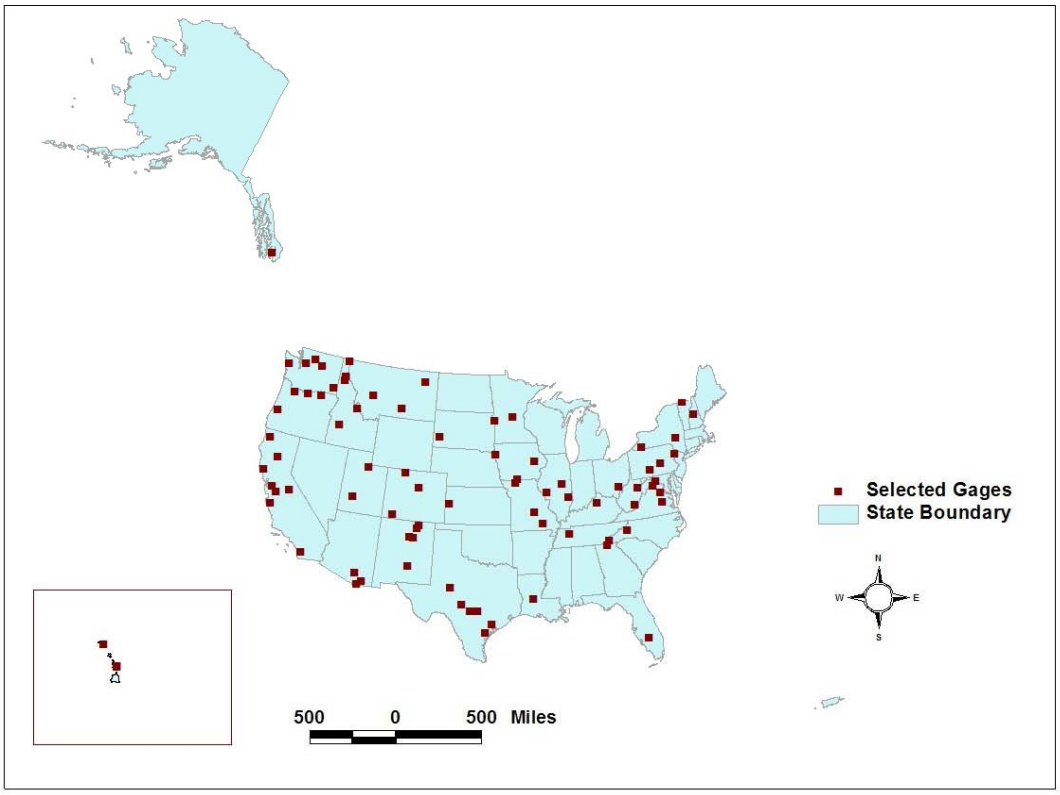

Figure 1. Map showing the geographic distribution of the gaging stations to be used in the Bulletin 17B – EMA testing.