Updating Bulletin 17B for the 21st Century *

Timothy A. Cohn[1], Nancy A. Barth[2], John England[3], Beth Faber[4], Robert R. Mason[1], and Jery R. Stedinger[1]

[1]U.S. Geological Survey, Reston, VA 20192
[2]U.S. Geological Survey, Sacramento, CA, 95818
[3]U.S. Bureau of Reclamation, Denver, CO
[4]U.S. Army Corps of Engineers, Davis, CA, 95616

8 March 2012

Unpublished Draft: Not for Citation

*The work in this report was conceived by the Hydrologic Frequency Analysis Work Group (HFAWG) under the Subcommitee on Hydrology (SOH), Advisory Committee on Water Data (ACWI). The HFAWG is currently chaired by Wilbert Thomas. A data subgroup and a testing subgroup were formed under HFAWG to examine potential changes to Bulletin 17B. The data subgroup consisted of Wilbert Thomas, Martin Becker, Don Woodward, Beth Faber, and Jerry Coffey. The HFAWG data subgroup provided a set of peak-flow data sets and suggested tests to perform in comparing Bulletin 17B and EMA.

Contents

1 Introduction **11**
1.1 Purpose of this Report (Charge) 11

2 Literature Sources: The History of Flooding and Flood Risk Estimation **14**

3 Metrics for Evaluating Flood-Frequency Estimators **15**
3.1 Simplified Criteria for Judging Estimator Performance 15
3.2 Application of Performance Criteria 16

4 Comparisons Based on Real Data at Selected Test Sites **17**
4.1 The Sites and Data 17
4.2 Software 17
4.3 Methods 18
4.4 Comparing Estimates 18
4.5 Results and Discussion 19
4.5.1 Sites with Systematic Gage Data and No Low Outliers or Historical Information 19
4.5.2 Sites with Historical Information 22
4.5.3 Sites with Systematic Gage Data and Low Outliers 24
4.5.4 Sites with Low Outlier, Historical and/or High Outliers . 30
4.6 Summary 38

5 Comparisons Based on Monte Carlo Experiments **41**
5.1 Understanding the Graphs 41
5.2 Studies with LP3 Distribution 44
5.3 Studies with LP3 Distribution and Regional Skew 46
5.4 Robustness Studies 49
5.4.1 Robustness with respect to Pearson Type 3 Population . 49
5.4.2 Robustness with respect to Population Constructed from Two LP3 Distributions 49
5.4.3 Robustness with respect to Population Constructed from Two GEV Distributions 51
5.5 Resampling Studies 56

6 Conclusions **62**

A The Test Dataset **64**

B Graphical Comparisons Between EMA and B17B at the 82 Test Sites **73**
B.1 Systematic Data Sites 73
B.2 Sites with Historical Information 97
B.3 Sites with Low Outliers 116
B.4 Sites with a Combination of Data Types 137

B.5 Studies with LP3 Distribution and Regional Skew 159
B.6 Additional Studies with Specific Frequency Curves 163

C Bulletin 17B Table of Contents, Status, and Possible Actions 166

List of Figures

1 Relative Percent Difference (RPD) for B17B/GB and EMA/MGB estimators for 10%, 1%, and 0.2% exceedance probabilities. Includes 23 sites without historical flood information where no low outliers were identified by Grubbs-Beck. 20
2 Relative Percent Difference (RPD) for B17B/GB and EMA/MGB estimators for 10%, 1%, and 0.2% exceedance probabilities. Figure represents 18 sites with historical information where no low outliers were identified by Grubbs-Beck. 23
3 Pryor Creek near Billings, MT (062165002) 25
4 Relative Percent Difference (RPD) for B17B/GB and EMA/MGB estimators for 10%, 1%, and 0.2% exceedance probabilities. Figure represents 18 sites with historical information where no low outliers were identified by Grubbs-Beck. 27
5 Percent low outliers identified using MGB test. Figure represents 20 sites without historical information. 28
6 Percent low outliers identified using MGB test. Figure represents 20 sites without historical information. 29
7 Orestimba Creek near Newman, CA, fit after application of the standard Grubbs-Beck test for low outliers. One low outlier and 6 zero flows were identified 31
8 Orestimba Creek near Newman, CA, fit after application of the Multiple Grubbs-Beck test for low outliers. Thirty-nine low outlier and 6 zero flows were identified 32
9 Wolf Creek near Wolf Point, MT 33
10 Relative Percent Difference (RPD) for B17B/GB and EMA/MGB estimators for 10%, 1%, and 0.2% exceedance probabilities. Represents 21 sites with historical flood information where low outliers were identified by Grubbs-Beck test. 35
11 Santa Cruz River near Lochiel, AZ, showing EMA/MGB estimate (Blue) and B17B/GB estimate (Red). Eight low outliers were identified by the MGB test, while only two low outliers were identified by the GB test. 36
12 Ten Mile Creek near Rimini, MT 37
13 Santa Cruz River near Lochiel, AZ, showing EMA/MGB estimate (Blue) and B17B/MGB estimate (Red). Eight low outliers were identified. 39
14 Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = 0.0$. 45

15 Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = -0.5$. 47
16 Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = 0.5$. 48
17 Results are based on 1000 replicate samples drawn from robustness test curve 3 50
18 Results are based on 1000 replicate samples drawn from robustness test curve 4 52
19 Results are based on 1000 replicate samples drawn from robustness test curve 5 53
20 Results are based on 1000 replicate samples drawn from robustness test curve 6 55
21 Results based on resampled data drawn from observed discharges at 03011020. 57
22 Results based on resampled data drawn from observed discharges at 11152000. 58
23 Results based on resampled data drawn from observed discharges at 14048000. 59
24 Results based on resampled data drawn from observed discharges at 14321000. 60
25 Results based on resampled data drawn from observed discharges at 13185000. 61
26 Site 01439500 with Systematic Data Only 74
27 Site 01555500 with Systematic Data Only 75
28 Site 01635500 with Systematic Data Only 76
29 Site 02037500 with Systematic Data Only 77
30 Site 02256500 with Systematic Data Only 78
31 Site 03183500 with Systematic Data Only 79
32 Site 05586500 with Systematic Data Only; 6 Peaks Identified as Below Gage Base Discharge 80
33 Site 06406000 with Systematic Data Only 81
34 Site 06710500 with Systematic Data Only 82
35 Site 07208500 with Systematic Data Only 83
36 Site 07382000 with Systematic Data Only 84
37 Site 08380500 with Systematic Data Only 85
38 Site 08387000 with Systematic Data Only 86
39 Site 11266500 with Systematic Data Only 87
40 Site 12134500 with Systematic Data Only 88
41 Site 12414500 with Systematic Data Only 89
42 Site 12437950 with Systematic Data Only 90
43 Site 12451000 with Systematic Data Only 91
44 Site 14021000 with Systematic Data Only 92
45 Site 14048000 with Systematic Data Only 93
46 Site 14137000 with Systematic Data Only 94
47 Site 15072000 with Systematic Data Only 95
48 Site 16587000 with Systematic Data Only 96

49 Site 01350000 with Systematic and Historical Data 98
50 Site 01562000 with Systematic and Historical Data 99
51 Site 01636500 with Systematic and Historical Data 100
52 Site 02138500 with Systematic and Historical Data 101
53 Site 03011020 with Systematic and Historical Data 102
54 Site 03051000 with Systematic and Historical Data 103
55 Site 03159500 with Systematic and Historical Data 104
56 Site 03550000 with Systematic and Historical Data 105
57 Site 03558000 with Systematic and Historical Data 106
58 Site 03606500 with Systematic and Historical Data 107
59 Site 04293500 with Systematic and Historical Data 108
60 Site 06216500 with Systematic and Historical Data 109
61 Site 06600500 with Systematic and Historical Data 110
62 Site 06898000 with Systematic and Historical Data 111
63 Site 07067000 with Systematic and Historical Data 112
64 Site 08378500 with Systematic and Historical Data 113
65 Site 09482500 with Systematic and Historical Data 114
66 Site 12413000 with Systematic and Historical Data 115
67 Site 01076500 with Systematic Data; Multiple Grubbs-Beck Test Identified 6 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 117
68 Site 01668000 with Systematic Data; Multiple Grubbs-Beck Test Identified 2 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 118
69 Site 03345500 with Systematic Data; Multiple Grubbs-Beck Test Identified 35 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 119
70 Site 05572000 with Systematic Data; Multiple Grubbs-Beck Test Identified 1 Low Outlier; Grubbs-Beck Test Identified 1 Low Outlier 120
71 Site 06176500 with Systematic Data; Multiple Grubbs-Beck Test Identified 7 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 121
72 Site 07203000 with Systematic Data; Multiple Grubbs-Beck Test Identified 6 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 122
73 Site 08133500 with Systematic Data; Multiple Grubbs-Beck Test Identified 23 Low Outliers; 20 Peaks Identified as Below Gage Base Discharge; Grubbs-Beck Test Identified 0 Low Outliers . . . 123
74 Site 08150000 with Systematic Data; Multiple Grubbs-Beck Test Identified 45 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 124
75 Site 08189500 with Systematic Data; Multiple Grubbs-Beck Test Identified 12 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 125

76 Site 09241000 with Systematic Data; Multiple Grubbs-Beck Test Identified 22 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 126
77 Site 10128500 with Systematic Data; Multiple Grubbs-Beck Test Identified 24 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 127
78 Site 10234500 with Systematic Data; Multiple Grubbs-Beck Test Identified 48 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 128
79 Site 11028500 with Systematic Data; Multiple Grubbs-Beck Test Identified 13 Low Outliers; 4 Peaks Identified as Below Gage Base Discharge; Grubbs-Beck Test Identified 1 Low Outlier 129
80 Site 11152000 with Systematic Data; Multiple Grubbs-Beck Test Identified 52 Low Outlier; Grubbs-Beck Test Identified 0 Low Outliers 130
81 Site 11274500 with Systematic Data; Multiple Grubbs-Beck Test Identified 39 Low Outliers; 11 Peaks Identified as Below Gage Base Discharge; Grubbs-Beck Test Identified 1 Low Outlier . . . 131
82 Site 11383500 with Systematic Data; Multiple Grubbs-Beck Test Identified 1 Low Outlier; Grubbs-Beck Test Identified 1 Low Outlier 132
83 Site 12307500 with Systematic Data; Multiple Grubbs-Beck Test Identified 12 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 133
84 Site 13302500 with Systematic Data; Multiple Grubbs-Beck Test Identified 45 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 134
85 Site 14321000 with Systematic Data; Multiple Grubbs-Beck Test Identified 9 Low Outliers; Grubbs-Beck Test Identified 2 Low Outliers 135
86 Site 16518000 with Systematic Data; Multiple Grubbs-Beck Test Identified 19 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 136
87 Site 03289500 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 6 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 138
88 Site 05270500 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 1 Low Outlier; Grubbs-Beck Test Identified 1 Low Outlier 139
89 Site 05291000 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 41 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 140
90 Site 05464500 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 5 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 141

91 Site 06062500 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 3 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 142
92 Site 06897000 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 6 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 143
93 Site 06933500 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 20 Low Outliers; 1 Peak Identified as Below Gage Base Discharge 144
94 Site 07138600 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 12 Low Outliers; 4 Peaks Identified as Below Gage Base Discharge 145
95 Site 08164000 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 1 Low Outlier; Grubbs-Beck Test Identified 0 Low Outliers 146
96 Site 08167000 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 23 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 147
97 Site 08171000 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 28 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 148
98 Site 09361500 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 2 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 149
99 Site 09471000 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 3 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 150
100 Site 09480000 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 8 Low Outliers; Grubbs-Beck Test Identified 2 Low Outliers 151
101 Site 11176000 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 28 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 152
102 Site 11464500 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 18 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 153
103 Site 11522500 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 1 Low Outlier; Grubbs-Beck Test Identified 1 Low Outlier 154
104 Site 12039500 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 48 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers 155
105 Site 13185000 with a Combination of Systematic and Histori- cal Data; Multiple Grubbs-Beck Test Identified 36 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier 156

106 Site 13343660 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 3 Low Outliers; 1 Peak Identified as Below Gage Base Discharge; Grubbs-Beck Test Identified 1 Low Outlier 157
107 Site 16068000 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 6 Low Outliers . . . 158
108 Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = 0.0$ Regional skew is assumed to be 0.0 with $MSE = 0.15$. 160
109 Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = -0.5$. Regional skew is assumed to be -0.5 with $MSE = 0.15$. 161
110 Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = 0.05$. Regional skew is assumed to be 0.5 with $MSE = 0.15$. 162
111 Results are based on 1000 replicate samples drawn from robustness test curve 1 164
112 Results are based on 1000 replicate samples drawn from robustness test curve 2 165

List of Tables

1 Streamflow-gaging stations used in comparing Bulletin 17B and EMA estimators 64
1 Streamflow-gaging stations used—Continued 65
1 Streamflow-gaging stations used—Continued 66
1 Streamflow-gaging stations used—Continued 67
1 Streamflow-gaging stations used—Continued 68
2 Available data and number of low outliers detected by Grubbs-Beck and Multiple Grubbs-Beck tests at streamflow-gaging stations used in comparing Bulletin 17B and EMA estimators . . . 69
2 Available data and number of low outliers – Continued 70
2 Available data and number of low outliers – Continued 71
2 Available data and number of low outliers – Continued 72
3 Assumptions Underlying Bulletin 17B, their current condition, and what to do with them 166
4 Features of Bulletin 17B, their current condition, and what to do with them 167

Executive Summary

For the past 30 years, Bulletin 17B has provided guidelines for flood-frequency analysis related to federal projects in the United States. During this period much has been learned about both hydrology and statistical methods. In keeping with the tradition of periodically updating the Guidelines in light of improved understanding, the Hydrologic Frequency Analysis Work Group (HFAWG), which reports through the Subcommittee on Hydrology (SOH) to the Advisory Committee on Water Information (ACWI), is currently considering possible updates to Bulletin 17B. The purpose of this report is to consider the impact that adoption of these changes to Bulletin 17B procedures would have on flood-frequency estimates.

The investigations reported here focuses on techniques for:

- Incorporating information related to historical flooding that occurred outside the period of systematic streamgaging;
- Addressing zero flows and low outliers;

The proposed changes are relatively modest, at least in the sense that they would not affect the main features of Bulletin 17B such as:

- Use of the log-Pearson Type III distribution;
- Fitting by Method-of-Moments;
- Identification of low outliers based on a Grubbs-Beck-type criterion.

The hydrological literature already supports the proposed changes, at least in theory. The relevant question, however, is practical: How well would the proposed methods perform in real situations with real data likely to be encountered in practice?

In order to answer this question, the HFAWG commissioned the work reported here. Several specific types of tests were developed and implemented:

- A direct comparison of results at 82 real "test sites" identified by an independent Data Group as both "typical" and "challenging" for flood frequency estimation;
- Monte Carlo simulations employing data drawn from specific LP3 populations
- Monte Carlo simulations employing data drawn from non-LP3 populations that were selected to reflect likely deviations, based on the experience of the Data Group, from the hypothesized LP3 distribution;
- Resampling with replacement of the data at the 82 sites.

It is believed that, taken together, these studies provide a reasonably comprehensive, valid and robust assessment of the properties of the proposed estimators.

A set of conclusions was reached. The proposed alternative methods:

- Perform at least as well as, and in some cases much better than, Bulletin 17B in terms of the Mean Square Error (MSE) of flood quantiles estimates;
- Allow for incorporation of more general types of flood-frequency information;
- Provide more accurate uncertainty estimates, and
- Avoid some annoying problems that arise when applying Bulletin 17B in practice.

In summary, the results generally confirm the theoretical findings reported in the hydrological literature.

1 Introduction

Flooding is the costliest natural hazard facing the United States in terms of loss of both life and property (Mileti, 1999; ACWI, 2011). While we know how to reduce flood losses through structural and non-structural mitigation, doing so efficiently requires a quantitative, uniform and consistent approach for estimating flood risk (Tasker and Thomas, 1978; Thomas Jr, 1985; Griffis and Stedinger, 2007). Recognizing this, the federal government has developed standard guidelines for performing flood-frequency analyses, and published these guidelines in a document known as "Bulletin 17B." The methods defined in Bulletin 17B inform literally millions of decisions about land use and construction, emergency response and recovery, and countless other human activities. The value of federal guidelines is not in dispute.

Bulletin 17B (IACWD, 1982, "B17B"), however, is not flawless. The document itself recognized that future work would be needed to improve some of the specified procedures (IACWD, 1982, p. 27, "Future Studies"), and much has been written about this and how things might be improved. More recently, both researchers and practicing hydrologists have called for updates to the B17B Guidelines (Stedinger and Griffis, 2008, for example).

1.1 Purpose of this Report (Charge)

The purpose of this report is to evaluate some specific modifications to the flood-frequency methods specified in Bulletin 17B. In November of 2005, the HFAWG proposed the following (ACWI, 2011, minutes from February 2005 HFAWG meeting):

> Based on recently completed research, the HFAWG proposes to investigate the following possible improvements in Bulletin 17B:

1. Evaluate and compare the performance of the Expected Moments Algorithm (EMA) (Cohn et al., 1997) to the weighted-moments approach of Bulletin 17B (Appendix 6) for analyzing data sets with historic information and paleoflood data.

    - Apply EMA and Bulletin 17B to gaging station data that include low and high outliers and historic data and those that do not. Develop criteria for determining if EMA provides more accurate and consistent flood estimates.
    - Review and evaluate the published literature for comparisons of EMA to conventional Bulletin 17B procedures.
    - Recommend improved plotting position formula when historic data are available.

2. Evaluate and compare the performance of EMA to the conditional probability adjustment of Bulletin 17B for analyzing data sets with low outliers and zero flows.

    - Apply EMA and Bulletin 17B to gaging station data that include low and high outliers and historic data and those that do not (same data set as noted above). Develop criteria for determining if EMA provides more accurate and consistent flood estimates.

3. Describe improved procedures for estimating generalized/regional skew.

    - Evaluate revisions needed in Bulletin 17B to describe improved procedures for estimating generalized/regional skew based on recently completed research.

4. Describe improved procedures for defining confidence limits.

- Evaluate revisions needed in Bulletin 17B to describe new procedures for defining confidence limits that include the uncertainty in the skew coefficient.
- Describe confidence limit procedures for EMA (if adopted).

2 Literature Sources: The History of Flooding and Flood Risk Estimation

Bulletin 17B (IACWD, 1982) defines the current method for conducting peak-flow frequency analyses in the U.S., but it descends from a long history beginning with Bulletins 13 (ICOWR, 1966), 15 (USWRC, 1967), 17 (USWRC, 1976), and 17A (USWRC, 1977), each of which includes a brief list of relevant references. The history of the various Bulletins is given in Thomas Jr (1985) and Stedinger et al. (1993); Griffis and Stedinger (2007), as are the assumptions and equations within Bulletin 17B.

AIR (AIR, 2001) provides a comprehensive chronology of the history of flood risk estimation as well as an extensive bibliography of the flood-frequency literature prior to 2000, and Stedinger et al. (1993) provide a bibliography of statistical techniques often employed in flood frequency analysis.

There is a vast literature on methods for flood-frequency estimation, and the references provided here should be regarded as no more than pointers to the larger literature. In particular, additional citations are available from Griffis and Stedinger (2007) and within the previous references.

3 Metrics for Evaluating Flood-Frequency Estimators

In practice, the benefits of more accurate flood information are site-specific. They depend on the physical, geological, social, and other characteristics of the site, as well as the decisions to be made and a host of other factors. However, in trying to assess the performance of flood-risk estimation techniques, one typically simplifies the problem by considering well-accepted criteria for judging estimator performance.

3.1 Simplified Criteria for Judging Estimator Performance

A set of simplified criteria were adopted to judge estimator performance. These can be divided into several groups:

- *Operational*
  1. *Ease of Application* Methods should be relatively easy to implement;
  2. *Applicability to Available Data* Methods should be able to make efficient use of data types at hand.
  3. *Consistency of Application* Frequency estimates should be transparent and fully reproducible;
  4. *Uniformity of Methods* Where possible, standardized methods should be used to ensure that different people performing the same analysis will obtain the same risk estimates;
- *Statistical*
  1. *Bias* On average, risk estimates should approximately equal the true risk; as more information becomes available, risk estimates should converge to the true risk;

2. *Efficiency* Estimators should extract as much *information* from the *data* as possible;

3. *Quantified uncertainty* Estimates should be accompanied by quantitative assessments of their uncertainty

- *Political, Legal and Institutional Criteria* Standard methods should be consistent, uniform, and easy to explain. They should satisfy legal requirements, and should serve institutional requirements of the federal government, and its National Flood Insurance Program, among others.

Various quantitative metrics are employed in this report to characterize the differences among the different procedures.

RPD The Relative Percent Difference (RPD) statistic is used to quantify differences among estimators.

ERL The efficiency of estimators that employ historical flood information is quantified in terms of both Effective Record Length (ERL), the amount of equivalent systematic record that would provide the precision achieved with both systematic and historical information.

AG ERL is further simplified by consideration of the Average Gain (AG) statistic, which expresses the benefit of each year of historical information in terms of an equivalent amount of systematic data.

3.2 Application of Performance Criteria

The following chapters of this report document a series of experiments that were conducted in order to determine how the existing procedures in Bulletin 17B, and some proposed alternatives, perform in light of the metrics in section 3.

4 Comparisons Based on Real Data at Selected Test Sites

4.1 The Sites and Data

Eighty-two streamflow-gaging stations were selected by the "Data Group" as a representative sample of situations that arise at real sites. The types of data found at these sites include:

1. Systematic Gage Only Data, no historical or low outlier data (23 sites);
2. Historical Data, could include high outliers (18 sites);
3. Low outliers; no historical information (20 sites);
4. Low outliers, Historical and/or High Outliers (21 sites).

All of the sites used as examples in B17B were also included in this study. While the 82 sites are not a representative sample of the many thousands of streamgage sites throughout the Nation, they are believed to cover the range of situations, and particularly the most difficult situations, that arise at these sites.

4.2 Software

The USGS PEAKFQ program was used for the B17B estimates, and the PeakfqSA v 0.95 software package was used to estimate flood quantiles and other statistics for the Expected Moments Algorithm (EMA) method. The current version of PEAKFQ (5.2) does implement EMA based on the same source code as PeakfqSA, but at present its "front end" does not permit entry of the interval data types that are under consideration (see 4.5.1 below). PEAKFQ (5.2) also does not implement the MGB (Cohn et al., 2012) multiple low-outlier test (although it can be calculated externally, as was done here, and entered manually (see 4.5.3, below)).

4.3 Methods

Estimates were computed for the 82 sites using the following methods:

1. EMA (Cohn et al., 1997) with a generalized Multiple Grubbs-Beck Test (MGB) (Cohn et al., 2012) for detecting multiple potentially influential low outliers in a flood series (EMA/MGB);

2. B17B (IACWD, 1982) with the standard Grubbs-Beck method for identifying low outliers followed by the conditional probability adjustment (CPA) (B17B/GB); and

3. B17B, where the MGB was employed to identify low outliers and the Conditional Probability Adjustment (CPA) (IACWD, 1982) was employed to fit the data (B17B/MGB).

At-site skew values were used for the comparisons between EMA and B17B methods.

4.4 Comparing Estimates

The fitted frequency curves for the 82 sites are presented in appendix B, with EMA/MGB results presented as a blue line and B17B/GB results as a red line. Note that where the two curves coincide, as is typically the case when only systematic data is available, the blue line is obscured behind the red line.
Relative percent differences (RPD) were computed to quantify the differences in estimates at the 82 sites between the three methods specified in subsection 4.3. The RPD is defined as:

$$RPD \equiv 100\left(\frac{Q_p^{EMA/MGB} - Q_p^{B17B/*}}{Q_p^{B17B/*}}\right) \tag{1}$$

where $Q_p^{B17B/*}$ is the p-th quantile flood estimate determined using either the

the B17B/GB or the B17B/MGB estimator depending on the case under consideration, and *p* refers to the probability of exceedance of the quantile under consideration (in this case 10%, 1% and 0.2%, corresponding to the 10-, 100-, and 500-year recurrence-interval events, respectively). An RPD close to zero indicates little difference between the estimates, while a large positive RPD indicates the EMA estimate is larger than the B17B estimate.

4.5 Results and Discussion

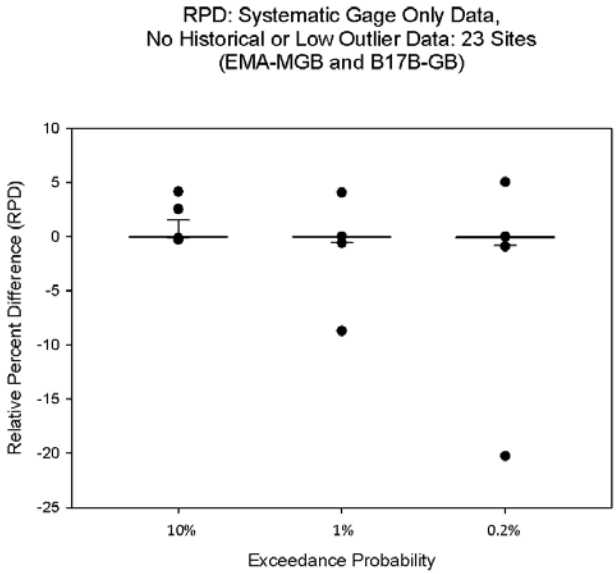
The RPD comparisons between EMA and B17B are presented in both box plots and flood-frequency plots (see appendix B). Systematic data are plotted in black on the logarithm-probability graphs using the plotting positions from the EMA-GB method. Those recorded gaged values identified as low outliers using the generalized Grubbs-Beck test (MGB) are plotted in yellow, while historical flood data are plotted in green. For those sites with atypical flood information, input between EMA and B17B were not always consistent and will be discussed in the respective categories.

4.5.1 Sites with Systematic Gage Data and No Low Outliers or Historical Information

The first category, "Gage Only" data, included 23 sites. The sites have systematic records with no historical information, no low outliers, and no below-gage-base flows identified by either using the GB or MGB tests. Almost all RPDs were zero, as expected. In this case all of the estimators are, in theory, identical[1]. As can be seen in figure 1, however, there were two sites where, in fact, the RPD was not zero.

[1]Apparent differences of less than 1% in RPD occur because PEAKFQ rounds quantile estimates to between 2 and 3 significant digits; PeakfqSA does no rounding. The estimated moments, which are not rounded, are identical in these cases

Figure 1: Relative Percent Difference (RPD) for B17B/GB and EMA/MGB estimators for 10%, 1%, and 0.2% exceedance probabilities. Includes 23 sites without historical flood information where no low outliers were identified by Grubbs-Beck.



At site 02037500, James River at Richmond, the RPD range between 3 to 5% for the 3 quantiles. This was because the 1937 peak discharge was recorded with a qualification code indicating the discharge was greater than the reported 152,000 [cfs] value. PEAKFQ, the USGS software that implements Bulletin 17B, omitted that peak altogether. This is a "known problem" with B17B, but it bears repeating: PEAKFQ is not well adapted to incorporating non-standard discharge values; in some cases the way it handles them is to ignore them altogether. As a result, the PEAKFQ estimates, which reflect B17B methods, do not properly employ the data.

EMA, the alternative method, readily accommodates interval data. Thus the 1937 peak could be correctly described as inside an interval from 152,000 to infinity. This was used in the EMA analysis to capture the additional information associated with the 1937 peak. Because EMA employs the large 1937 peak, EMA estimated higher values for the upper portion of the frequency curve.

Site 05586500, Hurricane Creek near Roodhouse, had a similar problem. It showed a RPD range from 4 to -20%. The sixth lowest recorded discharge had a qualification code indicating the recorded value was less than 70 [cfs]. PEAKFQ employed a gage base at 70 [cfs], consequently omitting five additional recorded point discharge values below 70 [cfs]. No low outliers were identified by either the GB or MGB test. Thus PEAKFQ unnecessarily truncated a portion of the left hand tail due to one qualification code. A user-supplied interval discharge range from 0 to 70 [cfs] was properly set in EMA for that single water year, and the five recorded point discharges below 70 [cfs] were included in the flood frequency analysis. Because EMA used all recorded discharges and the 0 to 70 [cfs] censored data, EMA's estimates were lower particularly at the upper end of the curve.

In both cases where the estimates differed, it was because EMA can accommo-

date data properly that PEAKFQ cannot accommodate.

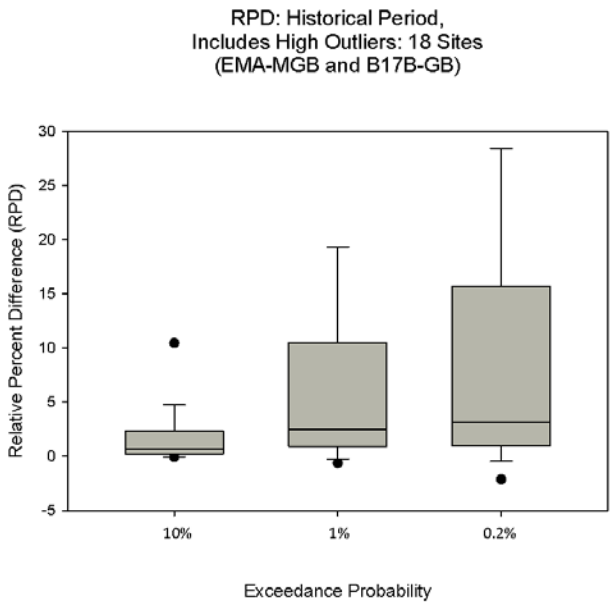
4.5.2 Sites with Historical Information

The testing for the "Historical Data" category included 18 sites, some of which included high outliers. The historical data sites illustrate the fundamental difference between EMA and B17B (Stedinger and Cohn, 1986; England et al., 2003).

EMA estimates, and therefore the RPDs, are sensitive to the historic threshold and historical period employed. All attempts were made to manually enter the same values into both software programs. However, some adjustments were made to accommodate PEAKFQ's inability to use interval discharge ranges. Many sites had one to three recorded historic peaks that exceeded the historic threshold. A few sites had recorded gage heights at or near record with missing discharges. Interval discharge values were set in EMA to accommodate these observations; because PEAKFQ has no corresponding capability, point discharge values were estimated by relating log-space discharge to gage height for those years and these were entered into PEAKFQ with the use of similar historic thresholds. Additionally, as a default, PEAKFQ sets an historic threshold at the lowest recorded historic value for a user-specified historic period. Thus all missing years of information in the historic record are effectively set to the lowest historic threshold. If a systematic record is missing any discharges (a broken systematic record) in a historical period, the missing data is set to the same historic threshold. This is not the case if there is missing data in a purely systematic record. PEAKFQ will assume no information is known about those missing years of systematic record.

Overall, the majority of sites with historical information showed a positive RPD in the flood estimates (figure 2). The interquartile distances ranged from approximately 0 to 10% for $\hat{Q}_{0.01}$. Three sites had higher RPD for $\hat{Q}_{0.01}$, ranging

Figure 2: Relative Percent Difference (RPD) for B17B/GB and EMA/MGB estimators for 10%, 1%, and 0.2% exceedance probabilities. Figure represents 18 sites with historical information where no low outliers were identified by Grubbs-Beck.



from 12 to 37%. EMA and B17B provide substantially different estimates for these three sites, all of which include historical information and high outliers. Site 06216500, Pryor Creek near Billings, had one large flood that was the largest in an extended historical period of 99 years. Figure 3 shows the frequency plots for both B17B and EMA. EMA seems to provide better visual fit to the data, most clearly to the high outlier in the right hand tail.

4.5.3 Sites with Systematic Gage Data and Low Outliers

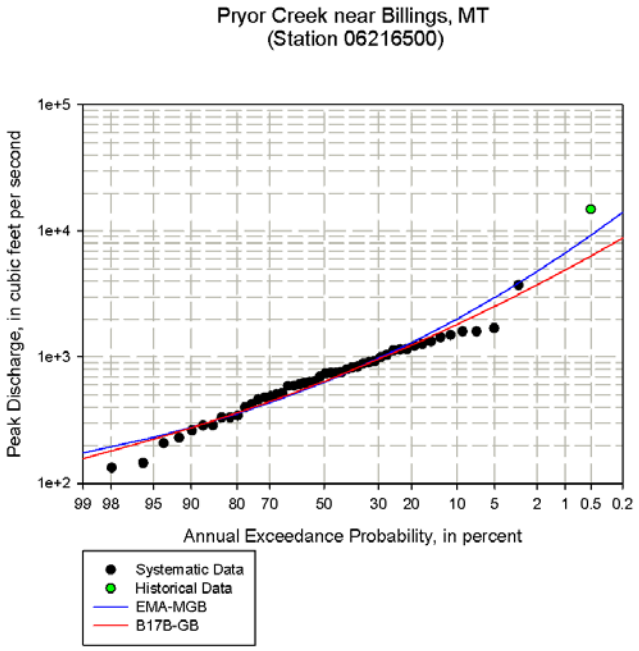
The testing for the "Low Outliers" category utilized 20 sites with low outliers identified using the standard Grubbs-Beck (GB) test recommended by B17B and a new generalized multiple Grubbs-Beck (MGB) test (Cohn et al., 2012). These sites did not include historic data. Two sets of comparisons were made between the results of EMA using only low outliers identified by the MGB (EMA/MGB) test with

1. B17B using the GB test followed by the conditional probability adjustment (CPA) (B17B/GB), and
2. B17B using the MGB identified low outlier threshold followed by the CPA (B17B/MGB).

Current PEAKFQ software does not include a MGB test option, so, where needed, the MGB threshold was computed independently and then entered into PEAKFQ as a user-supplied low outlier threshold.

The systematic flood series for site 08133500, North Concho River at Sterling City, included three peaks with a qualification code indicating the discharge was less than the reported value of 300 [cfs]. As discussed in the "Systematic Case" section (4.5.1), PEAKFQ arbitrarily set a gage base for the entire record, in this case omitting an additional 17 systematic point discharges less than 300 [cfs] that did not have remark codes. In EMA, the 3 peaks were recoded as between

Figure 3: Pryor Creek near Billings, MT (062165002)



0 and 300 [cfs]. However, the MGB test identified a low outlier threshold of 634 [cfs]. Thus EMA/MGB identified 23 low outliers and B17B/GB (with a default-set gage base of 300 [cfs]) identified 20 low outliers. The result is that the estimated flood quantiles do not differ by very much.

Flood estimate comparisons between EMA/MGB and B17B/GB for (figure 4) shows the median RPD is about -1%, or essentially zero. However, the RPD has substantial variability, indicating that the B17B/GB and EMA/MGB estimators behave differently when low outliers are present. This increased variability in RPD is attributed to the very different number of low outliers identified in the flood series and the methods used to handle low outliers in the frequency analysis, i.e., EMA's low outlier censoring versus B17B's CPA. Of the 20 sites in this low outlier category, the GB test found only 0 to 2 low outliers per site above gage base, while the MGB test found 2 to 52 low outliers (figure 5) At some sites MGB identified 50% of the recorded flows (figure 6) as "low outliers." This is the upper bound on the percentage of peaks that MGB will test.

The bserved RPDs were both positive and negative, a result of complex interactions between different low-outlier tests and different fitting procedures. However, use of the EMA/MGB method seems to produce a better fit to the upper portion of the frequency curve than is obtained with B17B/GB.

Orestimba Creek near Newman (site 11274500), which is also presented in B17B, represents an extreme case of low outliers – a flood series of 79 years including 11 zero flows. The LP3 distribution cannot describe the full range of observed flood flows at this site because the support for the LP3 distribution vanishes for $Q \leq 0$ (Cohn et al, 2011). PEAKFQ treats all zero flows as below gage base, while EMA regards them as ordinary low outliers.

When EMA is used with the ordinary Grubbs-Beck test, both it and B17B/GB employ a low-outlier threshold of 10.8 [cfs]. This results in the identification

Figure 4: Relative Percent Difference (RPD) for B17B/GB and EMA/MGB estimators for 10%, 1%, and 0.2% exceedance probabilities. Figure represents 18 sites with historical information where no low outliers were identified by Grubbs-Beck.

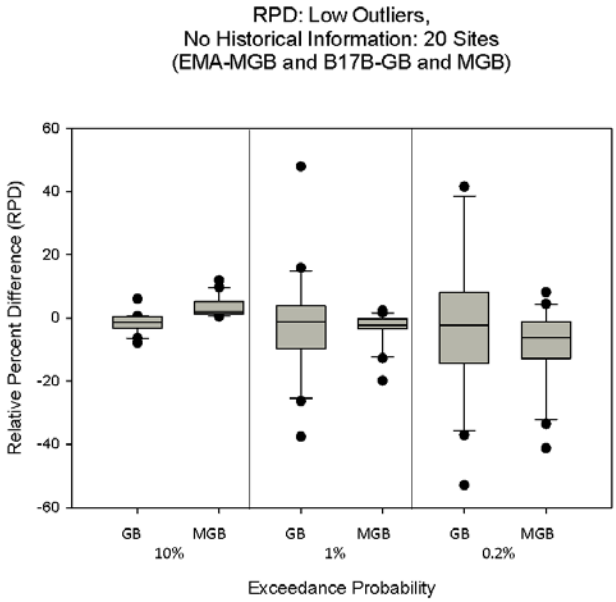


Figure 5: Percent low outliers identified using MGB test. Figure represents 20 sites without historical information.

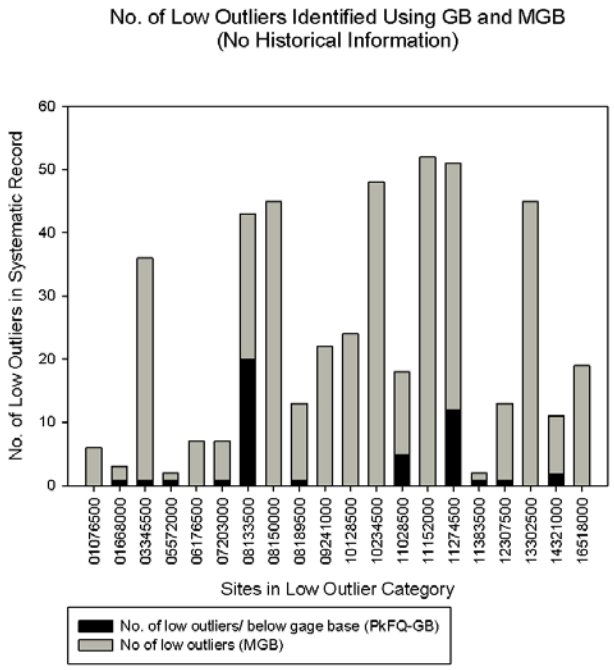
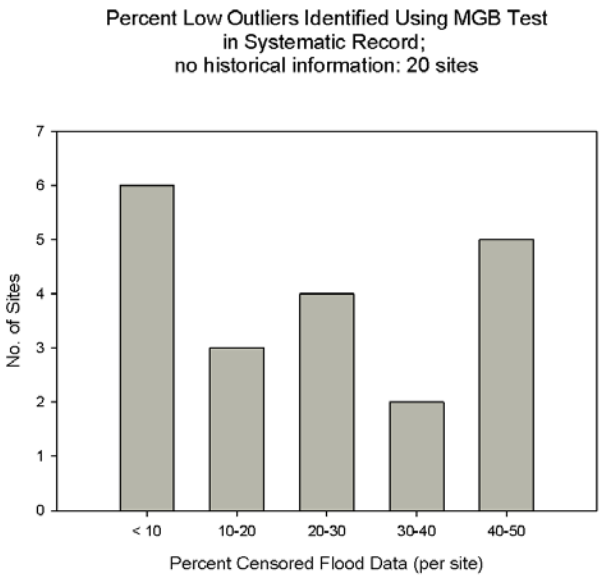


Figure 6: Percent low outliers identified using MGB test. Figure represents 20 sites without historical information.



of a single additional low outlier (4 [cfs]). However, EMA's fit to the data is very poor. All of the flows in the left-hand tail of the annual peak flow series at Orestimba are unusually small, but the single Grubbs-Beck test is really only able to identify at most one one or two low outliers. The resulting RPD between EMA/GB and B17B/GB for the flood estimates are 111% and 113%, respectively. If a poor low-outlier test is employed, EMA does a poor job fitting the frequency distribution, as is seen at Orestimba. (figure 7) The CPA, because of its complicated dependence on only the upper tail of the population, actually does a better job than EMA when using the standard Grubbs-Beck test to identify low outliers.

However, the MGB fixes this problem. The MGB test sets a low outlier threshold at 1260 [cfs], subsequently identifying 39 peaks, which represent 49% of the data, as low outliers. As a result, the RPD between EMA/MGB and B17B/GB was reduced to 2% and 7% for the 1%, and 0.2% exceedance probability estimates, respectively (figure 8). This improvement in EMA's estimation of the upper tail is a direct result of correctly identifying the multiple influential low outliers and recoding them as interval values. This greatly improves the fit in the right-hand tail.

The pattern seen at Orestimba is also observed at other sites with multiple low outliers. For example, at Wolf Creek near Wolf Point (06176500, figure 9) one sees that EMA/MGB, by treating the influential small peaks as low outliers, results in a reasonably close fit in the right-hand tail.

4.5.4 Sites with Low Outlier, Historical and/or High Outliers

The testing for the combination "Low Outlier, Historical and/or High Outlier" category included 21 sites. The methods employed in this category are the same as those in both the "Historical Information" and "Low Outlier" cases. Yet this category contains sites whose flood series have low outliers with

Figure 7: Orestimba Creek near Newman, CA, fit after application of the standard Grubbs-Beck test for low outliers. One low outlier and 6 zero flows were identified

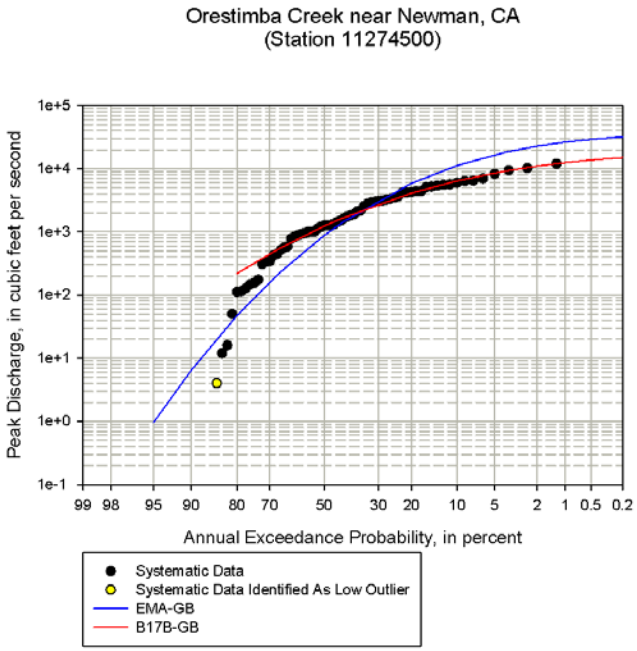


Figure 8: Orestimba Creek near Newman, CA, fit after application of the Multiple Grubbs-Beck test for low outliers. Thirty-nine low outlier and 6 zero flows were identified

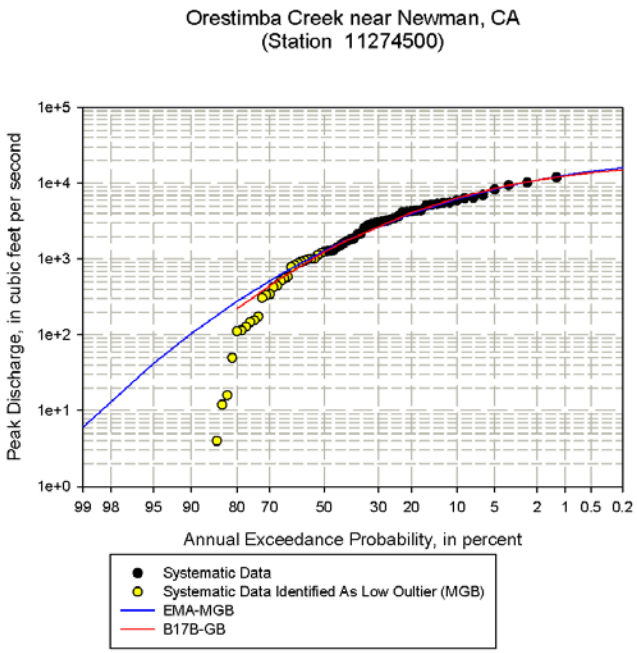
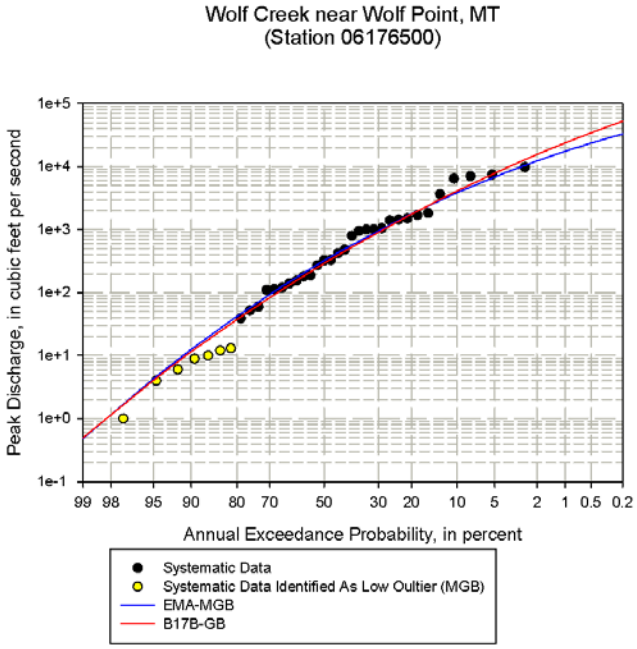


Figure 9: Wolf Creek near Wolf Point, MT



1. Only high outliers in a systematic record, or

2. High outliers in a historical period.

Nothing new appeared in these cases. The RPD between EMA/MGB and B17B/GB are similar to those found in both the "Low Outlier" and "Historical" categories. The median RPD for the estimates remained near -1%, yet the median RPD for $Q_{10\%}$ increased from approximately -1% to 2 and 8 %, respectively (figure 10). About a third of the sites in this category showed an RPD greater than 20% for the 0.2% estimates. The largest RPD between EMA/MGB and B17B/GB was found at site 09480000, Santa Cruz River near Lochiel. The RPDs for the estimates were between 58 and 112%, respectively. EMA/MGB found 8 low outliers in the systematic record (figure 11) while B17B/GB found only two low outliers above gage base. Additionally, the systematic record contains two high outliers each of which is nearly twice the magnitude of the third largest peak discharge. By censoring multiple low outliers, EMA/MGB more accurately fit the high outliers in the upper tail. A combination of B17B/GB using the CPA and a lack of historical information resulted in B17B/GB's poor fitting of the two largest recorded peaks.

The next largest RPD between EMA/MGB and B17B/GB was found at site 06062500, Tenmile Creek near Rimini. A 23 and 48% RPD difference was found for the estimates, respectively. EMA/MGB found 3 low outliers, while B17B/GB found none (figure 12).

Figures 11 and 12 illustrate an important point: Both B17B/GB and EMA/MGB often provide a good fit to the data they employ. However, even B17B/GB can be highly influenced by low outliers with the result being a poor fit at the high end of the distribution. EMA/MGB, on the other hand, avoids this problem by correctly recognizing the low outliers and recoding them so that their exact magnitudes do not distort the fit in the right-hand tail.

Figure 10: Relative Percent Difference (RPD) for B17B/GB and EMA/MGB estimators for 10%, 1%, and 0.2% exceedance probabilities. Represents 21 sites with historical flood information where low outliers were identified by Grubbs-Beck test.

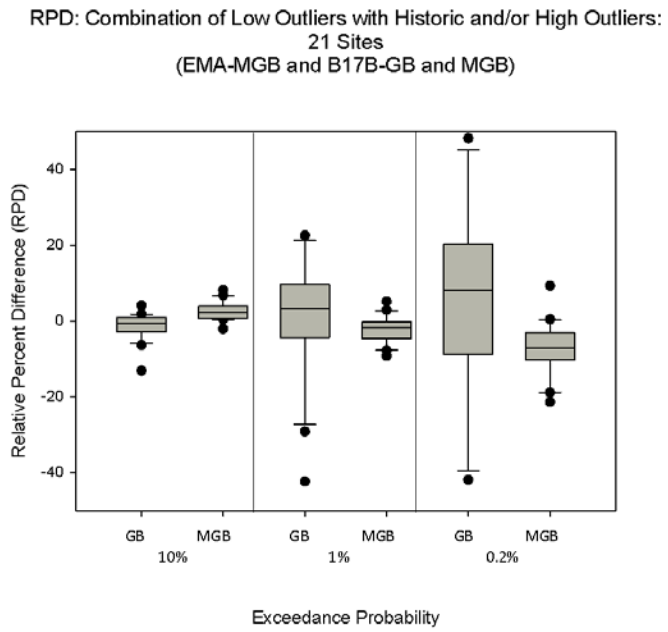


Figure 11: Santa Cruz River near Lochiel, AZ, showing EMA/MGB estimate (Blue) and B17B/GB estimate (Red). Eight low outliers were identified by the MGB test, while only two low outliers were identified by the GB test.

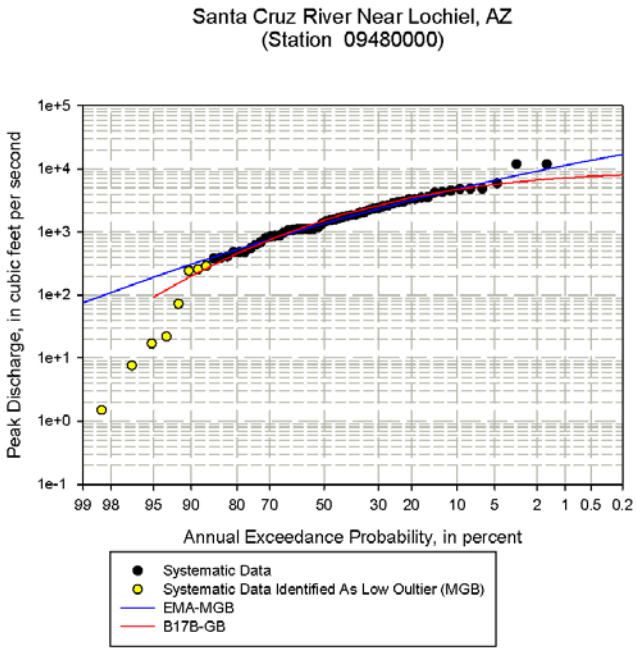
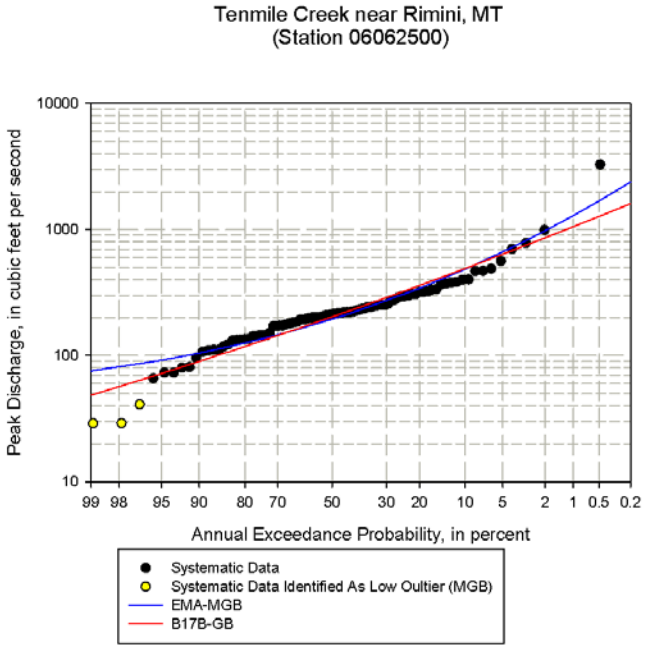


Figure 12: Ten Mile Creek near Rimini, MT



The RPD between EMA/MGB and B17B/MGB when a flood series has both low and high outlier data is similar to that found when only low outliers are present. The median value and interquartile range for the estimates was slightly positive while the median values and interquartile ranges for the estimates were negative. As expected, the variance in the RPD was minimized for all estimates when the same low outlier threshold was used. The higher estimate from B17B/MGB is illustrated by figure 13 (09480000, Santa Cruz River near Lochiel) where a -18% RPD difference was found. The MGB low outlier threshold of 380 [cfs] was used in B17B/MGB. The MGB test identified almost 15% of the peaks as low outliers. Using the same low-outlier test, based on visual inspection EMA/MGB and B17B/MGB both fit the data reasonably well.

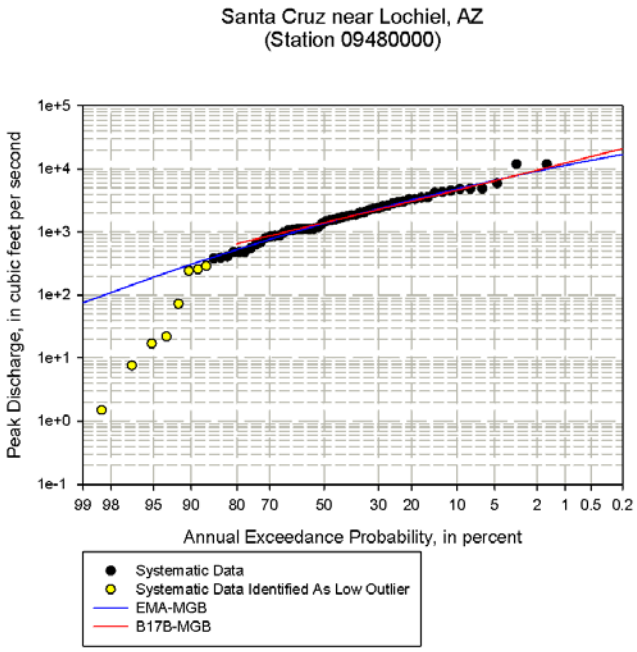
4.6 Summary

82 streamflow-gaging stations were chosen as a representative sample of long-term sites whose flood series include a variety of situations and problems that are believed to be found throughout the U.S. The flood data was divided into four categories:

1. Systematic data only with no low outliers or historical information,
2. Historical and/or high outliers,
3. Low outliers, and
4. Low outliers, with historical and/or high outliers.

The performance of EMA/MGB, B17B/GB and B17B/MGB methods were compared. A relative percent difference (RPD) was calculated corresponding to each of three exceedance probabilities to compare the EMA/MGB and B17B/GB.

Figure 13: Santa Cruz River near Lochiel, AZ, showing EMA/MGB estimate (Blue) and B17B/MGB estimate (Red). Eight low outliers were identified.



When only systematic flood data were present, identical flood estimates (RPD equal to zero) were obtained except for those cases where B17B cannot accommodate non-standard data correctly. When historical and/or high outliers were present, observed RPDs were typically positive. This is at least in part due to the 82 sites that were selected for testing, many of which included historical "high outliers." EMA tends to attach more "weight" to historical flood information than does B17B (Stedinger and Cohn, 1987), so where the historical period includes an unusually high peak, the EMA/MGB estimate will tend to be higher than the B17B/GB estimate.

Flood series that contain multiple low outliers exhibited a range of results primarily due to the very different number of low outliers identified by the GB and MGB tests and the methods used to handle low outliers in the frequency analysis, i.e., EMA's low outlier censoring versus B17B's CPA. The influence of low outliers on the upper portion of the frequency curve was greatly reduced by the EMA/MGB method. When the same low outlier threshold was used for the EMA and B17B's CPA fitting procedure, the RPDs were usually smaller and slightly negative. For those sites with low outliers with historical, and/or high outlier data, similar RPD were found to those in the historical and low outlier categories. This combination group included both low outliers and high outliers. The RPD were notably more positive for these estimates. EMA/MGB identified more low outliers, resulting in a closer fit to the larger peaks. B17B/GB identified fewer low outliers generally resulting in lower estimates of flood quantiles than were found with EMA/MGB.

5 Comparisons Based on Monte Carlo Experiments

Hundreds of distinct Monte Carlo and resampling experiments were conducted in order to determine how well the B17B/GB and EMA/MGB estimators satisfy the criteria listed in section 3.

5.1 Understanding the Graphs

In order to simplify the discussion, results are presented concisely using a standard graphical format. Each figure considers the performance of the estimators when data are drawn from a specific single population. Figure 14, for example, corresponds to an LP3 distribution with a log-skew $G = 0.0$; later graphs consider estimator performance when data are drawn from other populations. Each graphic has two parts. The top panel contains 12 boxplots, divided into four groups of three, showing the distribution of three estimators for $Q_{0.01}$, the 1 percent non-exceedance event. The box spans the 25th to 75th percentiles of the estimators' distributions. The central line indicates the median. The full range of the data is shown by the whiskers and hinges. One additional feature is present on each box, a circle with a cross which indicates the location of the sample mean.

The three estimators are defined below and in IACWD (1982); Cohn et al. (1997, 2012). Although flood quantiles other than the 1% exceedance were investigated, the results were found to be insensitive to which quantile was estimated. Because $Q_{0.01}$ is the flood of interest for many federal activities, this case is reported. All results are based on 1000 replicate samples.

The same three estimators are considered here that were considered in section 4.3:

- *B17B/GB (gray)* the Bulletin 17B estimator using the Grubbs-Beck test followed by the conditional probability adjustment;

- *B17B/MGB (green)* the Bulletin 17B estimator using the Multiple Grubbs-Beck test followed by CPA; and

- *EMA/MGB (aqua)* the EMA estimator using the Multiple Grubbs-Beck test.

Each of the four groups corresponds to $N_S = 40$ years of systematic data and a $N_H = 100$ year historical period during which, on expectation, 0, 1, 2 or 10, respectively, historic floods would have been recorded. The green slashed line indicates the true value of $Q_{0.01}$; in later plots, where $Q_{0.01}$ is estimated by interpolation between real observations, a purple line is used to indicate an estimate of $Q_{0.01}$; the true value of $Q_{0.01}$ is unknown.

The lower panel in each graph indicates the effective record length (ERL) of each estimator. ERL is defined as the ratio:

$$ERL \equiv 40\left(\frac{Var[log(\hat{Q}_{0.01}(S = 40, H = 0))]}{Var[log(\hat{Q}_{0.01}(S = 40, H = 100))]}\right) \qquad (2)$$

where s is the length of the systematic record and h is the length of the historical period. ERL expresses, in a rough way, how well an estimator exploits historical information in terms of an equivalent number of years of exclusively systematic data.

The numbers beside each point indicate the average gain (AG), in percent, associated with each year of historical flood information. The AG quantifies the relative value of an additional year of historical information to an additional year of systematic gage record, and is defined as:

$$AG \equiv ERL - 40 \qquad (3)$$

AG is the percentage increase of ERL for each year in the historical period. Where the expected number of historic floods is high – on the right side of the graph – the AG is typically relatively high. Where no historic floods can be expected – the group on the left – the average gain is zero. However, the average gain is defined in terms of the MSE of the EMA estimator, so, due to differences in treatment of low outliers, the AG for B17B/GB and B17B/MGB can differ slightly from zero[2]. Such deviations have nothing to do with the value of historical information, but rather result from differences among the estimators with respect to low outliers.

[2]Note: AG results also vary due to sampling variability. They are based on 1000 replicate samples)

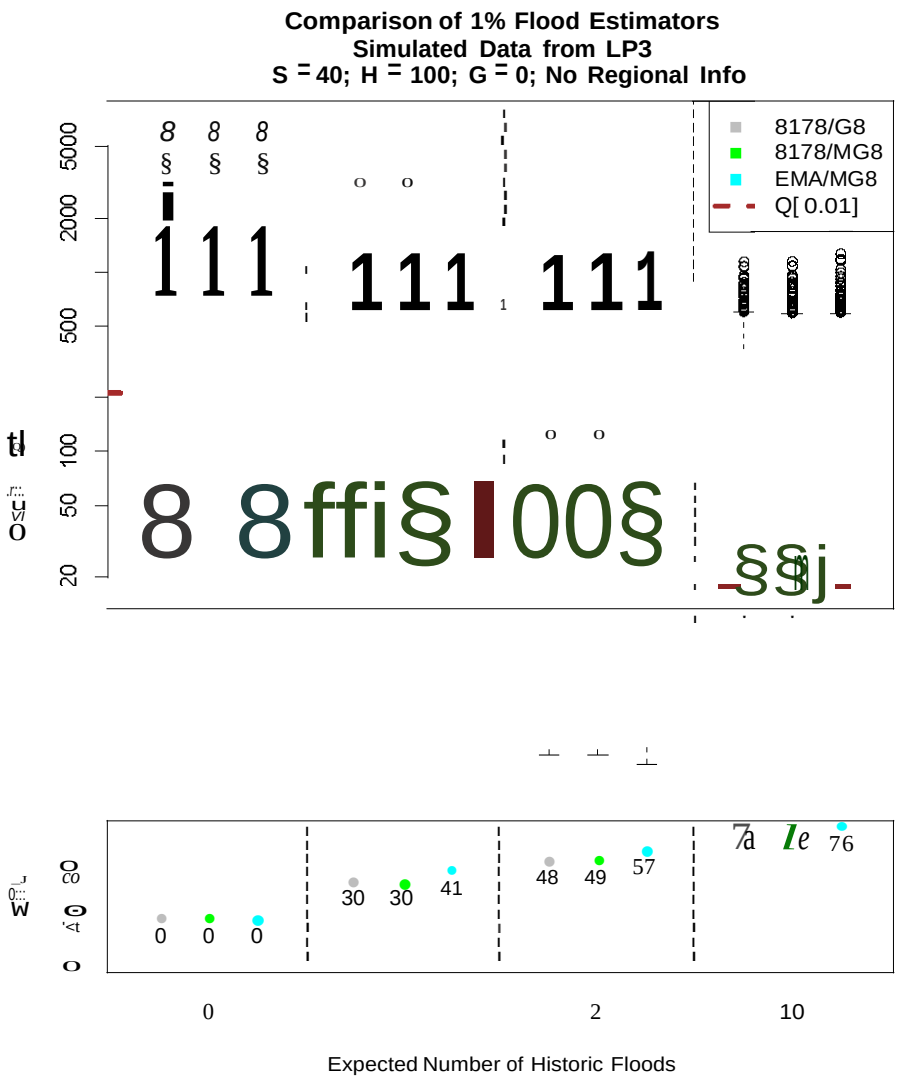
5.2 Studies with LP3 Distribution

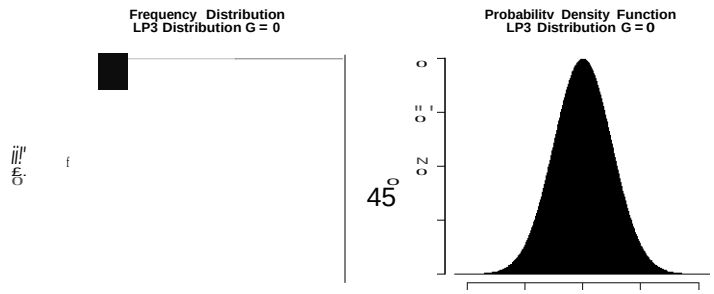
Figures 14 through 16 correspond to the case where we are fitting the LP3 distribution to LP3 data. In these cases we expect good results for all the estimators because we are applying the correct model for the population from which the data are drawn. We assume we have no regional information. The distribution of the Monte Carlo population, depicted as a frequency plot and as a probability density function of the logarithm of Q, appears below the main plot.

Figure 14, which depicts the case when the population skew is $G = 0$ and there are unlikely to be many low outliers, shows that the 3 estimators are identical when there is no historical information (the left-most three boxplots). If historical information is present with a perception threshold at approximately the 1% non-exceedance flood level (Stedinger and Cohn, 1986), the EMA/MGB (AG=41) method performs substantially better than B17B/GB (AG=30) or B17B/MGB (AG=30). The same conclusion applies when the threshold is at the 2% non-exceedance level, where the corresponding average gains are 48, 49, 57. When the threshold is at the 10% non-exceedance level, which corresponds to a very favorable situation, all of the estimators perform extremely well, with average gains above 75%.

Figure 15 depicts the case when the population skew is $G = -0.5$. In this case many low outliers are to be expected. Figure 15 reveals several interesting properties of the estimators. First, the center of the boxplots are substantially above the hashed line (the true 1% non-exceedance level), indicating that all of the estimators are biased when only systematic data are employed. This phenomenon is actually well known (Kirby, 1974; Stedinger et al., 1993): The method-of-moments estimator for the skew coefficient is biased toward zero, and thus method-of-moments quantile estimators are biased upwards for populations

Figure 14: Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = 0.0$.





99
50
10
-4
-2
Percent Exceedance Probability
Log(O)

with negative skews and downwards for populations with positive skews (see figure 16). When historical information is present, the EMA/MGB method performs substantially better than B17B/GB or B17B/MGB. This is to be expected, because the method-of-moments is highly sensitive to the smallest observations from negatively skewed populations. Thus EMA, which handles the censored small observations more efficiently, performs much better than B17B in these cases.

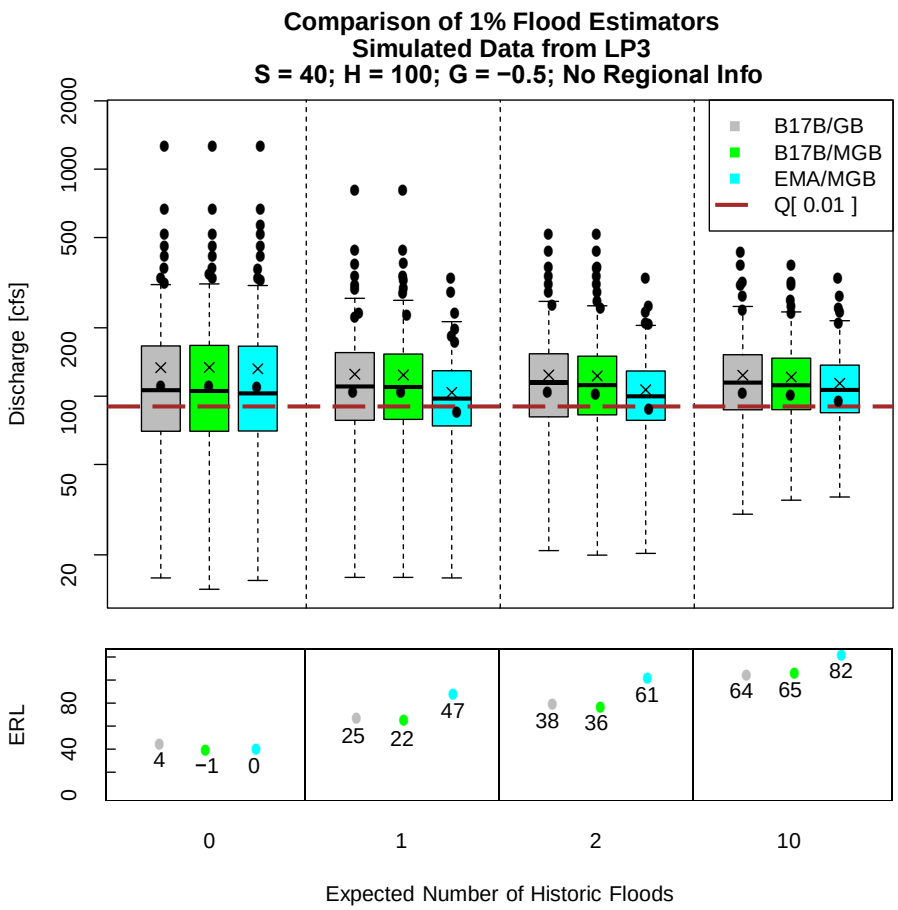
Figure 16 depicts the case when the population skew is $G = 0.5$ and no low outliers are to be expected. This case reveals the opposite bias seen in figure 15. However, in this case all of the estimators make good use of the historical information because, for positively skewed populations, the smallest values in the dataset have little leverage and it makes little difference how they are treated.

5.3 Studies with LP3 Distribution and Regional Skew

Figures 108 - 110 show the same cases as Figures 14 - 16 except that regional skew information has been added with a MSE of 0.15 – a typical value consistent with Bayesian/GLS skew maps. The addition of regional skew information substantially improves all of the estimators, including EMA/MGB and B17B/GB. Interestingly, the effect is essentially the same for all of the estimators.

In conclusion, while accurate regional skew is a valuable addition to frequency analyses, it does not need to be considered as an important factor in determining the properties of the three estimators, EMA/MGB, B17B/GB, and B17B/MGB, under consideration here.

Figure 15: Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = -0.5$.



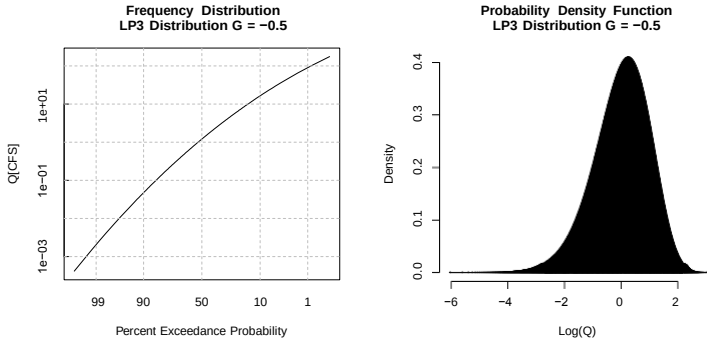
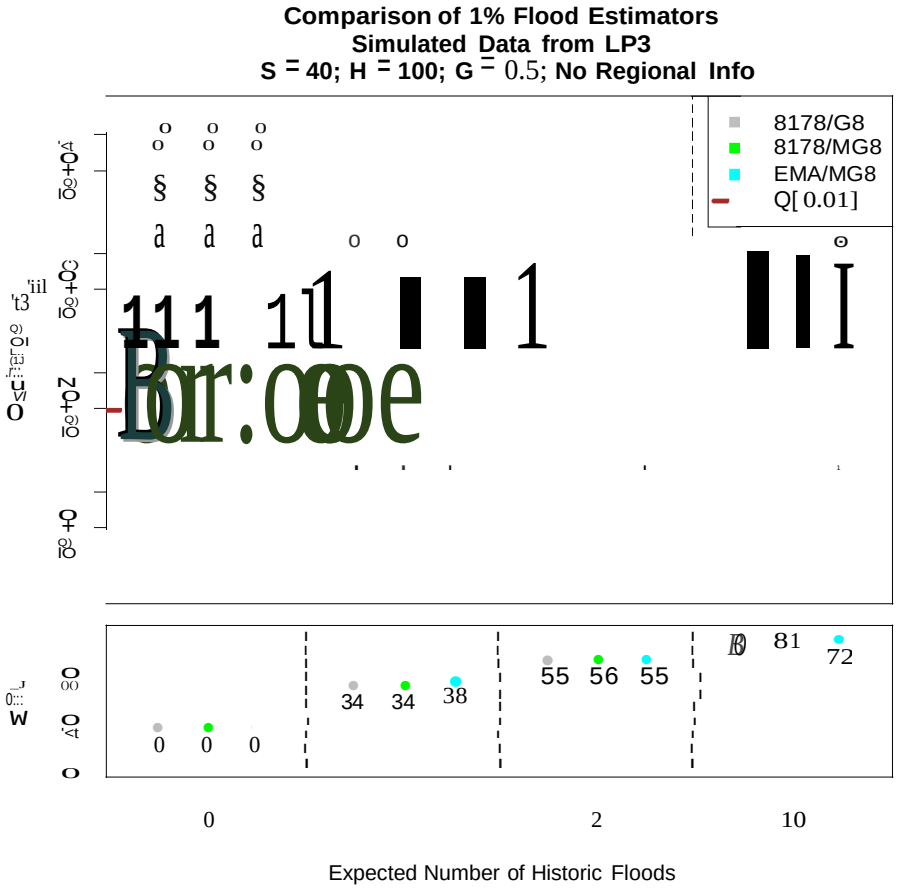
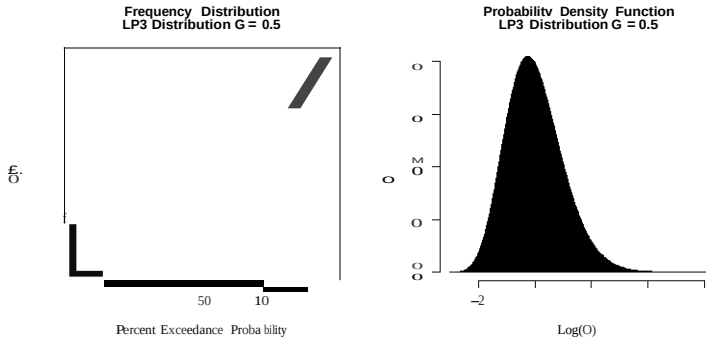


Figure 16: Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = 0.5$.





5.4 Robustness Studies

Figures 111, 112, and 17 - 20 apply the three estimators, EMA/MGB, B17B/GB and B17B/MGB, to data drawn from specific populations selected specifically to test the estimators' performance. The Cases are labeled " s test curve 1-6" in recognition of their origins.

Because figures 111 and 111 depict cases that are equivalent to cases already discussed in section 5.2, these figures have been included in the appendix.

5.4.1 Robustness with respect to Pearson Type 3 Population

Figure 17 depicts the performance of the estimators when data are drawn from a Pearson Type 3 population, not the LP3. This tests the robustness of the estimators. All of the estimators are biased when fitting this population, as is expected because it is not the assumed population. However, the EMA/MGB estimator does perform slightly better than the two other estimators when historical information is present. All three estimators are the same in the absence of historical information.

5.4.2 Robustness with respect to Population Constructed from Two LP3 Distributions

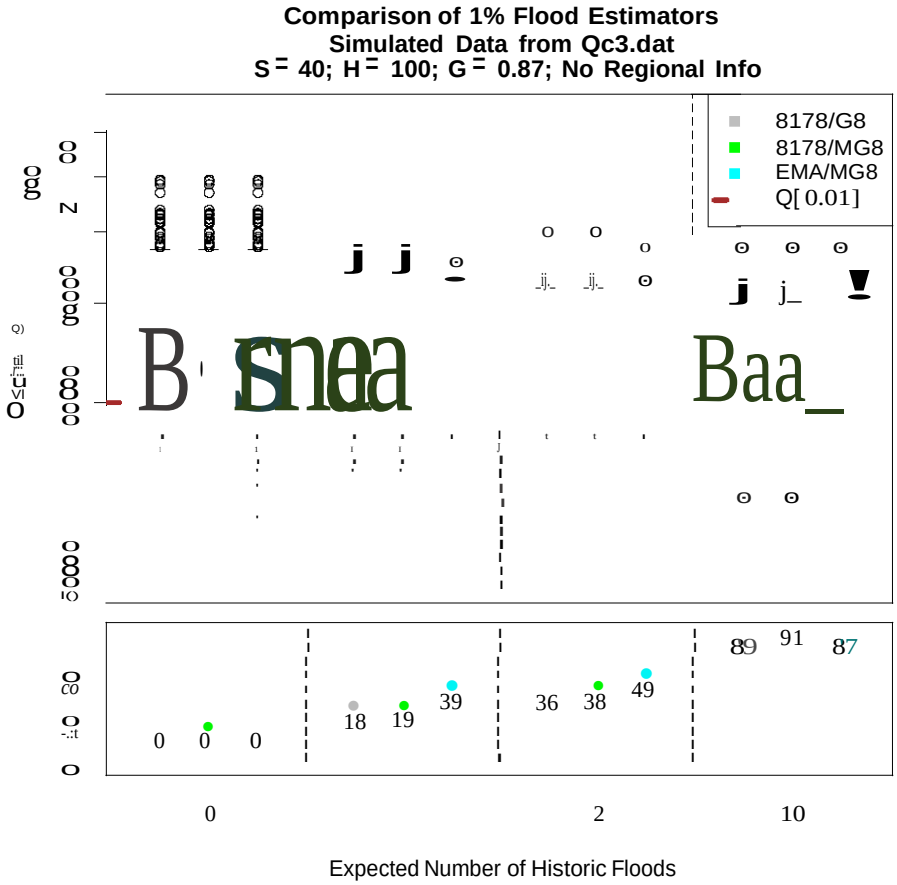
Figure 18 depicts the performance of the estimators when data are drawn from a constructed population created by choosing the minimum of observations drawn from two different LP3 distributions, the first with parameters

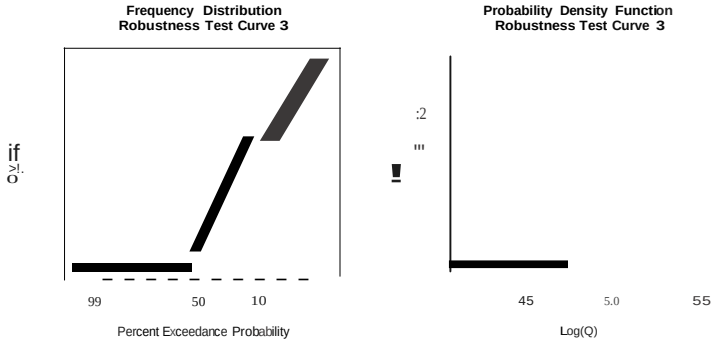
$$\{M, S^2, G\} = \{4.1212, 0.29^2, 1.00\}$$

and the second with parameters

$$\{M, S^2, G\} = \{4.0900, 0.13^2, 0.15\}$$

Figure 17: Results are based on 1000 replicate samples drawn from robustness test curve 3





Fitting data from this population tests the robustness of the estimators; low outliers are present. Without historical information, EMA/MGB and B17B/MGB are nearly identical, and both perform better than B17B/GB. When historical information is present, EMA/MGB performs much better than the other estimators.

Figure 19 depicts the performance of the estimators when data are drawn from a constructed population based on two LP3 distributions, the first with parameters

$$\{M, S^2, G\} = \{4.3438, 0.41^2, -1.00\}$$

and the second with parameters

$$\{M, S^2, G\} = \{4.3936, 0.50^2, -0.20\}$$

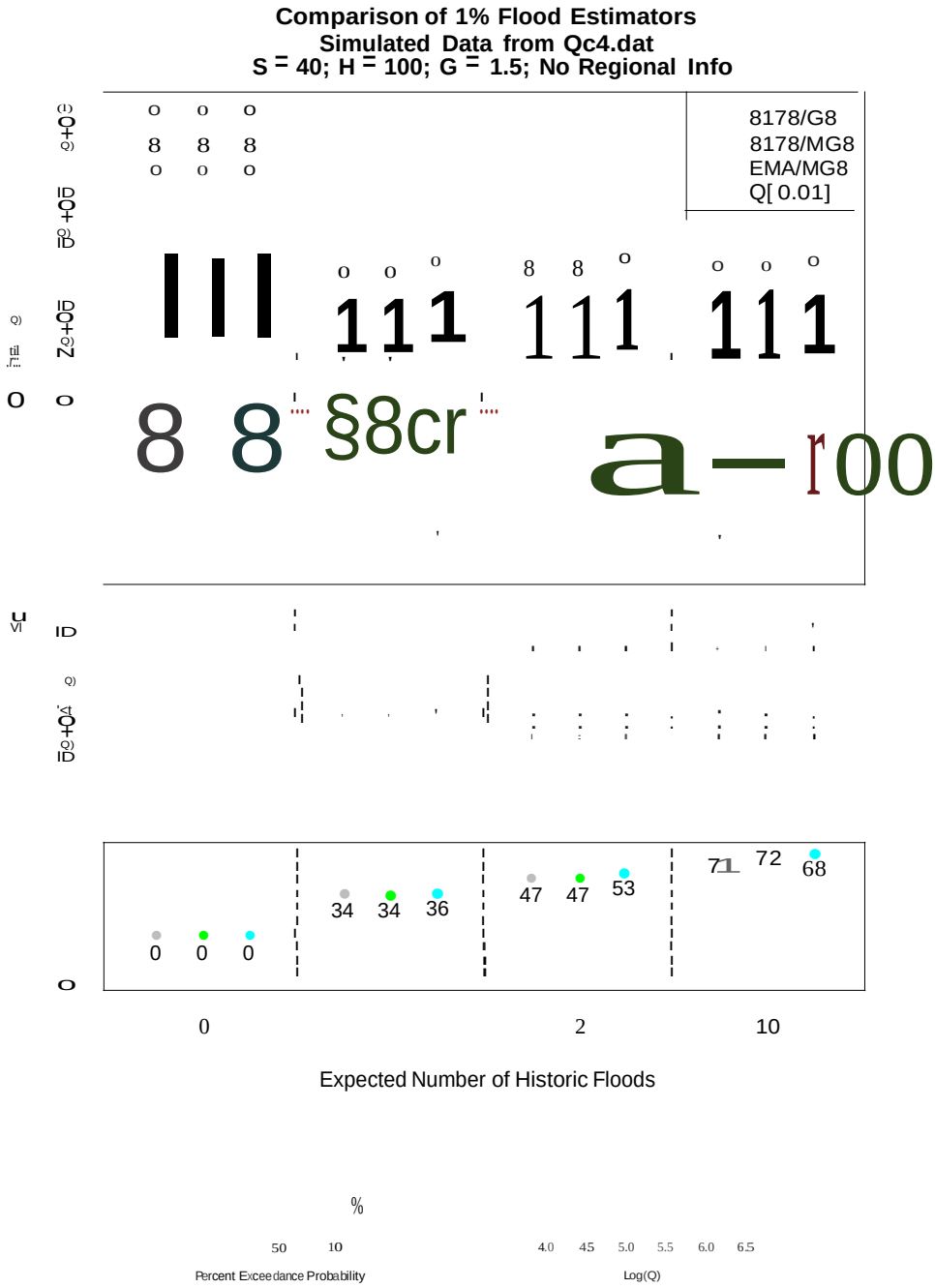
The lower half of the distribution function employs the first parameters, and the upper half is based on the second parameters. The distributions have the same median, which is where they join. Fitting data from this population tests the robustness of the estimators; low outliers are present. Without historical information, EMA/MGB and B17B/MGB are nearly identical, and both perform better than B17B/GB. When historical information is present, EMA/MGB performs much better than the other estimators.

5.4.3 Robustness with respect to Population Constructed from Two GEV Distributions

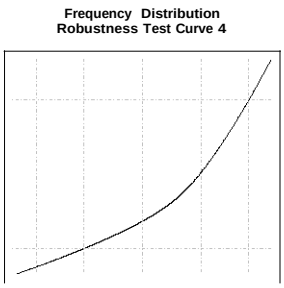
Figure 20 depicts the performance of the estimators when data are drawn from a constructed populations based on two Generalized Extreme Value distributions, the first with parameters

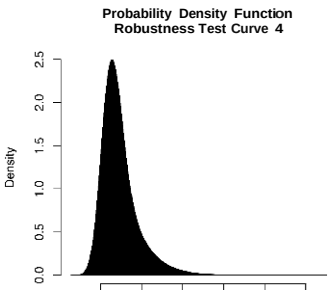
$$\{\kappa, \alpha, \xi\} = \{0.08, 24326, 6378\}$$

Figure 18: Results are based on 1000 replicate samples drawn from robustness test curve 4



if
≥.
o





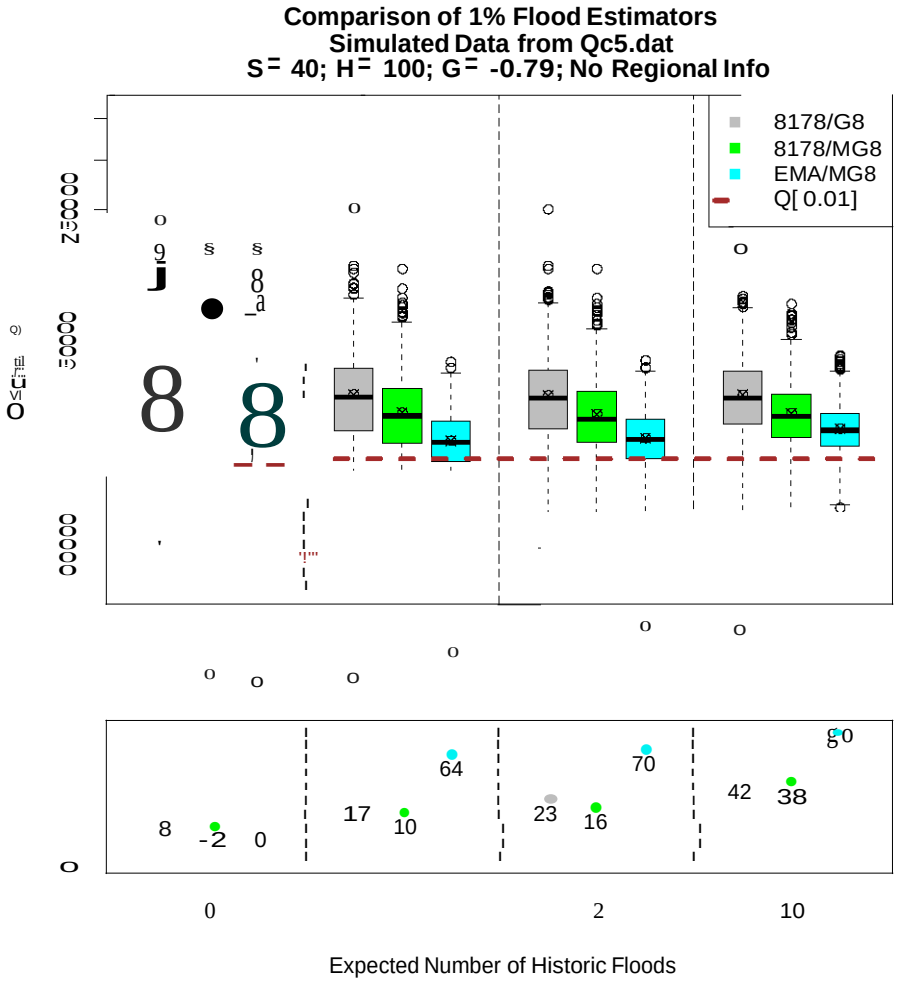
50 10

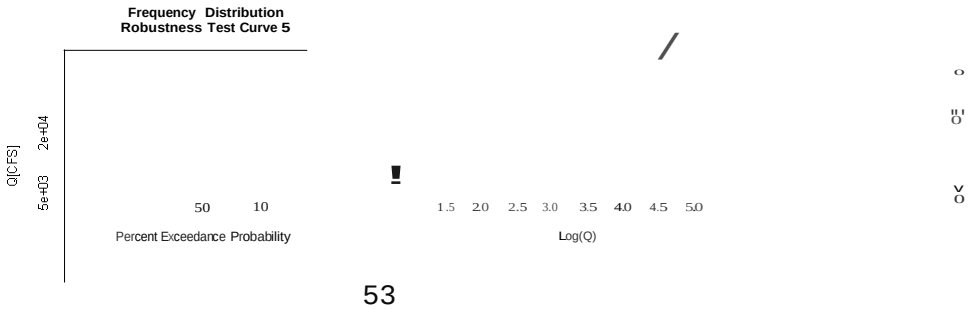
Percent Exceedance Probability

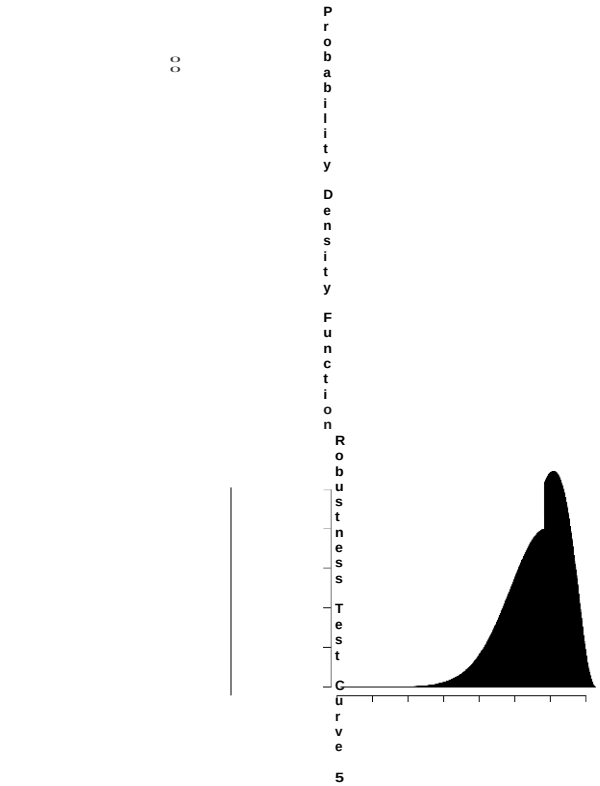
4.0 4.5 5.0 5.5 6.0 6.5

Log(Q)

Figure 19: Results are based on 1000 replicate samples drawn from robustness test curve 5





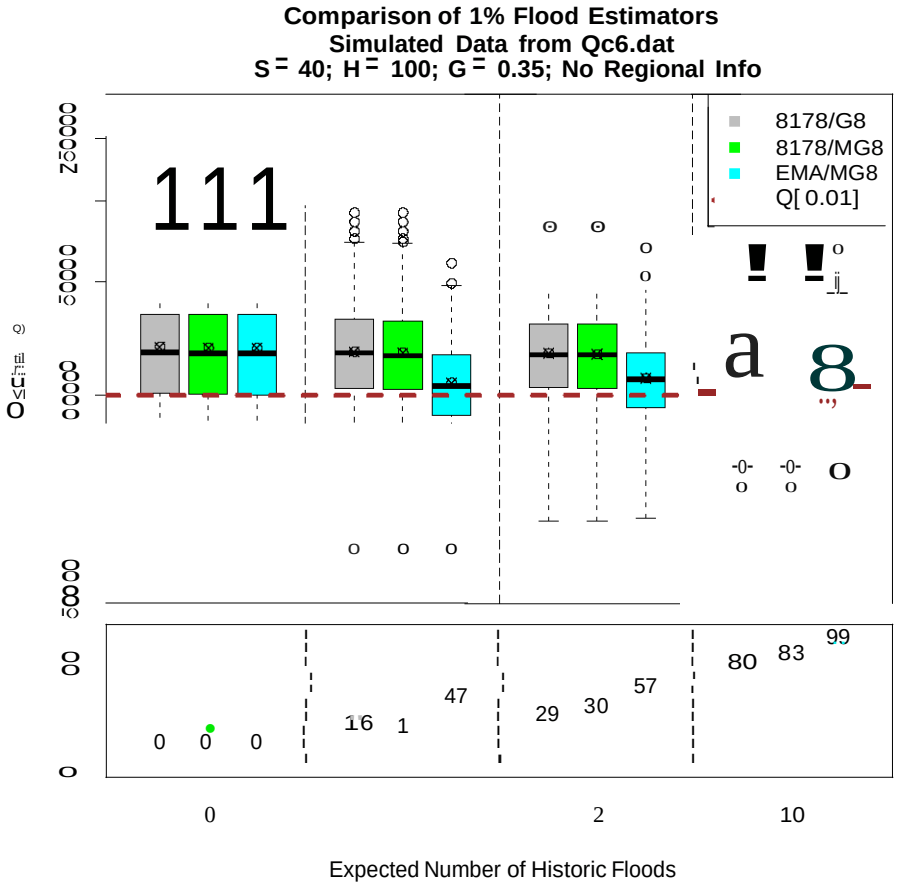

Probability Density Function
Robustness Test Curve 5
80
50
10
Percent Exceedance Probability
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Log(Q)

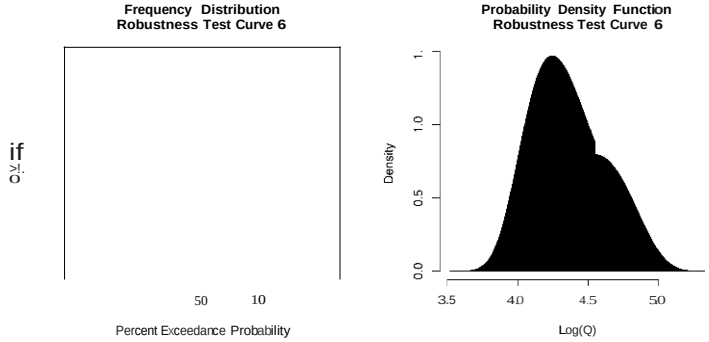
and the second with parameters

$$\{\kappa, \alpha, \xi\} = \{-0.55, 10000, 17330\}$$

The lower three quarters of the distribution function employs the second parameters, and the upper quarter is based on the first parameters. The distributions have the same upper quartile, which is where they join. For this population, all of the estimators perform about the same.

Figure 20: Results are based on 1000 replicate samples drawn from robustness test curve 6





5.5 Resampling Studies

Figures 21 - 25 depict results from applying the three estimators, EMA/MGB, B17B/GB and B17B/MGB, to resampled data from five of the longest-record (> 100 observations) sites among the 82 "test" sites considered in section 4. In these cases, we do not know what the true value of the 1% non-exceedance event is, so the figures instead employ an interpolated value based on the two largest observations in the dataset. However, this is an unreliable estimator, particularly because, as noted in section 4.6, many of the "test" sites were selected specifically because they contained high outliers. In reviewing figures 21 - 25, it is likely best to use one's judgment about the *reasonableness* of the results, possibly referring back to the test-site results, rather than trying to conjure up a strict quantitative assessment. However, it is noteworthy that, when historical information is present, the EMA/MGB estimator generally outperforms the other estimators.

Figure 21: Results based on resampled data drawn from observed discharges at 03011020.

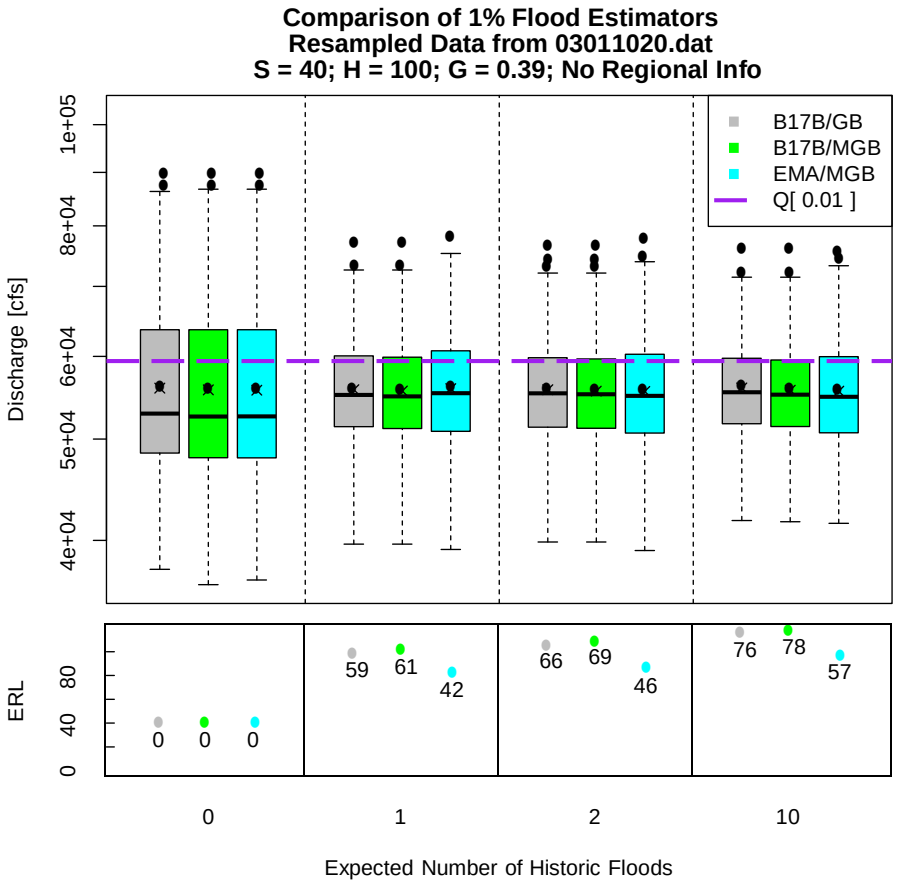


Figure 22: Results based on resampled data drawn from observed discharges at 11152000.

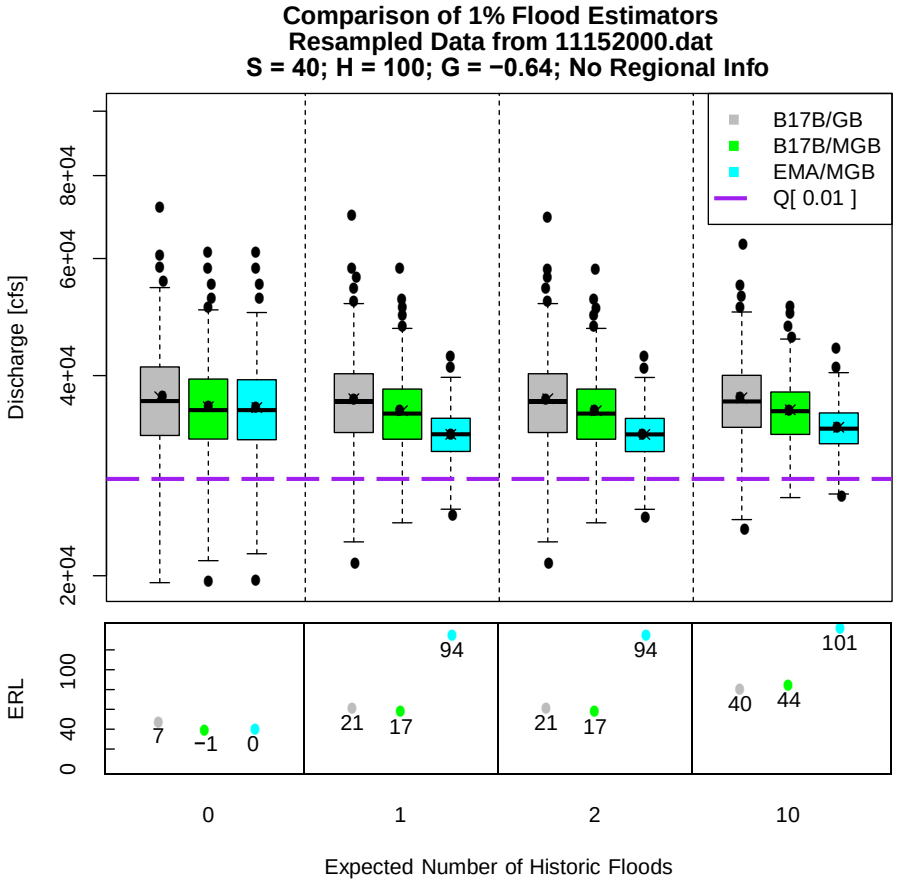


Figure 23: Results based on resampled data drawn from observed discharges at 14048000.

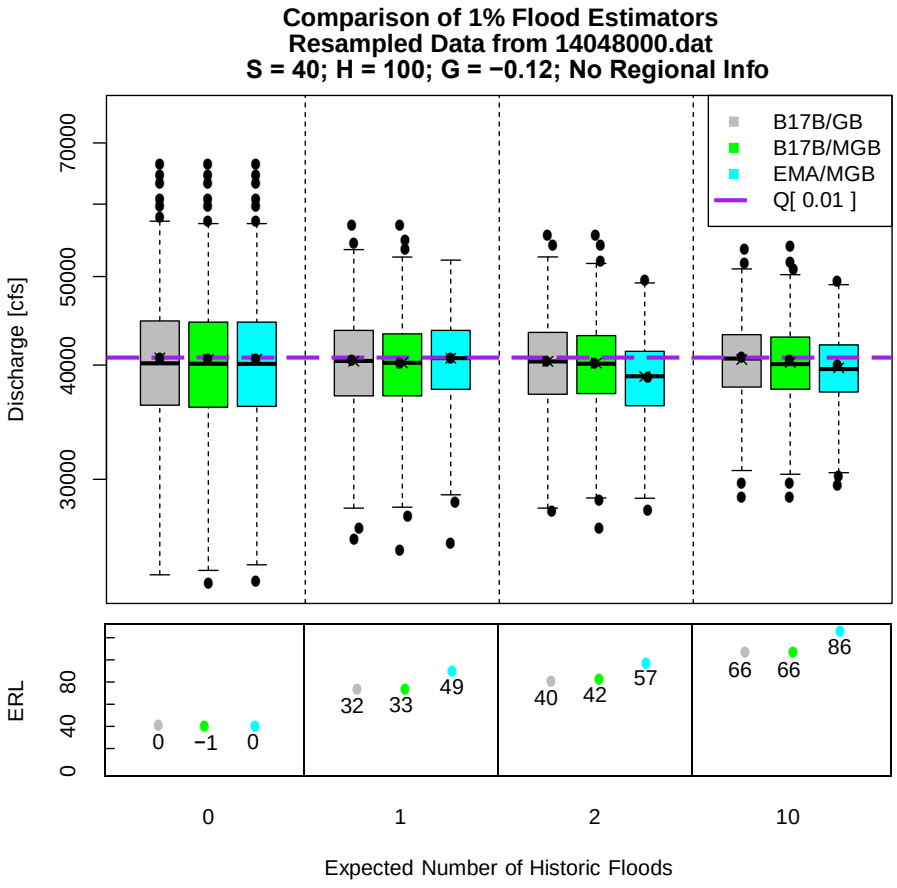


Figure 24: Results based on resampled data drawn from observed discharges at 14321000.

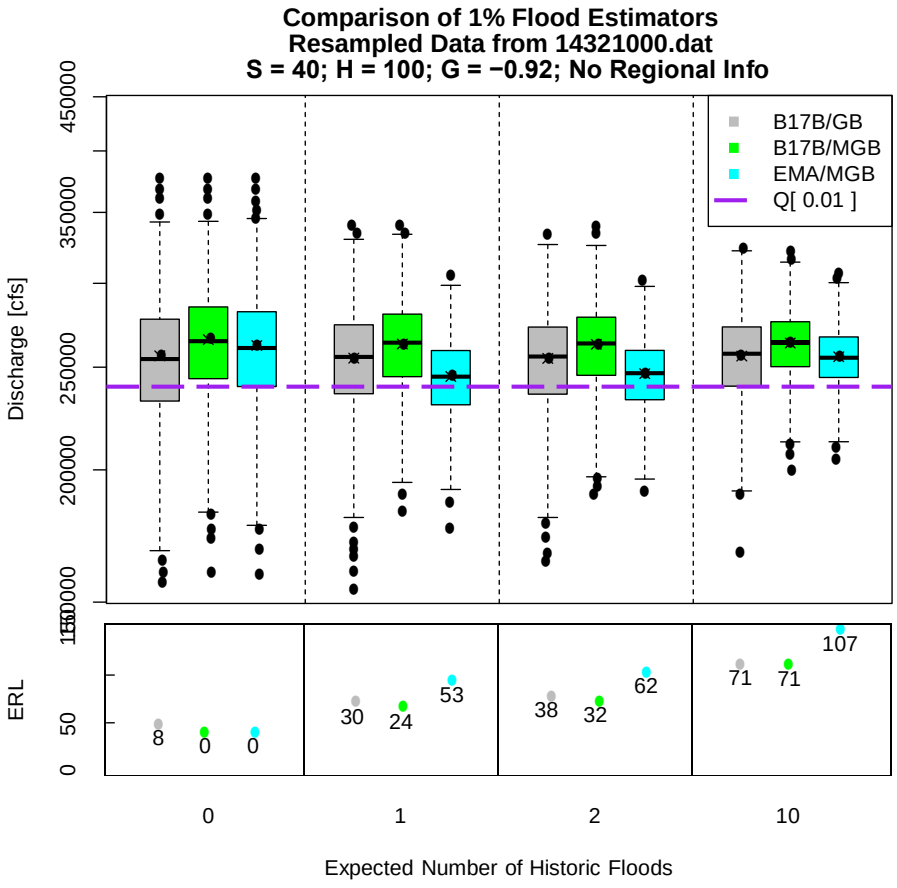
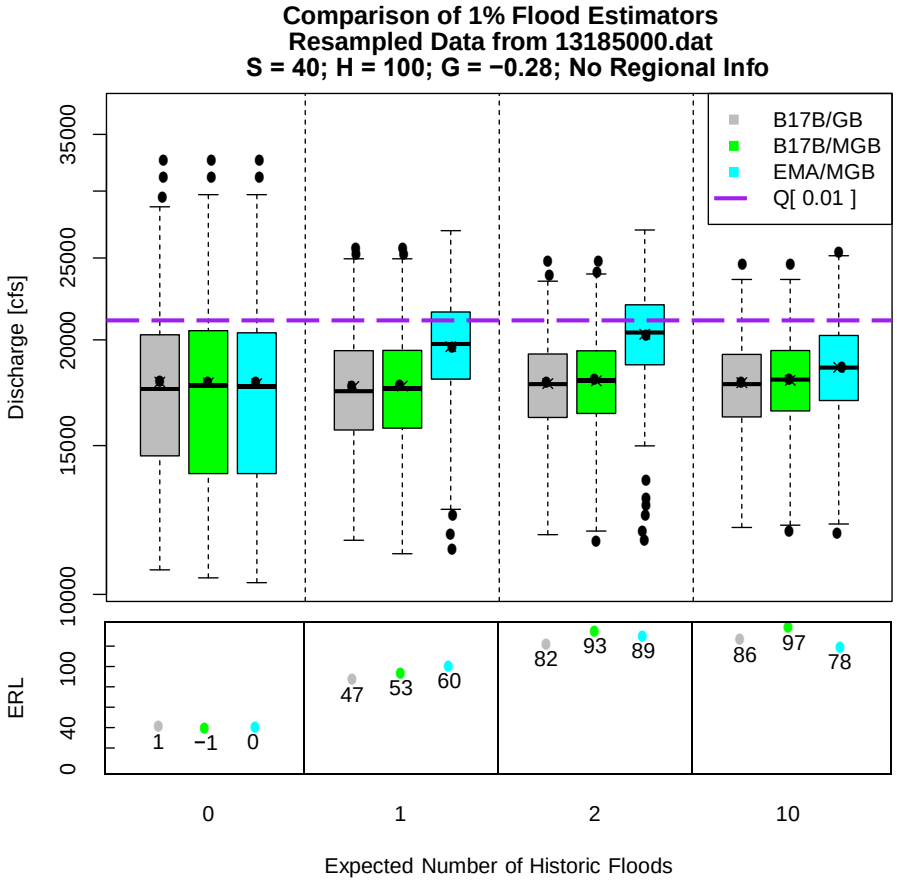


Figure 25: Results based on resampled data drawn from observed discharges at 13185000.



6 Conclusions

The study reported here was designed to determine how well some proposed changes to Bulletin 17B would perform in practice. In particular, the investigations focused on techniques for:

- Incorporating information related to historical flooding that occurred outside the period of systematic streamgaging;
- Addressing zero flows and low outliers;

In order to answer these question, datasets were selected and created, and tests were developed and conducted:

- A direct comparison of results at 82 real "test sites" identified by an independent Data Group as both "typical" and "challenging" for flood frequency estimation;
- Monte Carlo simulations employing data drawn from specific LP3 populations
- Monte Carlo simulations employing data drawn from non-LP3 populations that were selected to reflect likely deviations, based on the experience of the Data Group, from the hypothesized LP3 distribution;
- Resampling with replacement of the data at the 82 sites.

It is believed that, taken together, these studies provide a reasonably comprehensive, valid and robust assessment of the properties of the proposed estimators.

A set of conclusions was reached. The proposed alternative method, denoted EMA/MGB:

- Performs at least as well as, and in some cases much better than, Bulletin 17B in terms of the Mean Square Error (MSE) of flood quantiles estimates;

- Allows for incorporation of more general types of flood-frequency information;
- Provides more accurate uncertainty estimates, and
- Avoids some annoying problems that arise when applying Bulletin 17B in practice.

In summary, the results generally confirm the theoretical findings, reported in the hydrological literature, that EMA can provide improved flood frequency estimates.

A The Test Dataset

Table 1: Streamflow-gaging stations used in comparing Bulletin 17B and EMA estimators

Site	Station No.	Station Name	Drainage Area (mi^2)	No. Peaks	Category
1	01076500	Pemigewasset River at Plymouth, NH	622	106	gage
2	01350000	Schoharie Creek at Prattsville, NY	237	100	high
3	01439500	Bush Kill at Shoemakers, PA	117	102	gage
4	01555500	East Mahantango Creek near Dalmatia, PA	162	81	gage
5	01562000	Raystown Branch Juniata River at Saxton, PA	756	99	high
6	01635500	Passage Creek near Bucktown, VA	86.5	78	gage
7	01636500	Shenandoah River at Millville, WV	3041	97	high
8	01668000	Rappahannock River near Fredericksburg, VA	1595	100	low
9	02037500	James River at Richmond, VA	6753	76	gage
10	02138500	Linville River near Nebo, NC	66.7	89	high
11	02256500	Fisheating Creek at Palmdale, FL	311	79	low
12	03011020	Allegheny River at Salamanca, NY	1608	107	high
13	03051000	Tygart Valley River at Belington, WV	406	104	high
14	03159500	Hocking River at Athens, OH	943	78	high

Table 1: Streamflow-gaging stations used—Continued

Site	Station No.	Station Name	Drainage Area (mi^2)	No. Peaks	Category
15	03183500	Greenbrier River at Alderson, WV	1364	115	gage
16	03289500	Elkhorn Creek near Frankfort, KY	473	72	highlow
17	03345500	Embarras River at Ste. Marie, IL	1516	99	low
18	03550000	Valley River at Tomotla, NC	104	101	high
19	03558000	Toccoa River near Dial, GA	177	85	high
20	03606500	Big Sandy River at Bruceton, TN	205	70	high
21	04293500	Missisquoi River near East Berkshire, VT	479	92	high
22	05270500	Sauk River near St. Cloud, MN	1030	75	low
23	05291000	Whetstone River near Big Stone City, SD	398	83	high
24	05464500	Cedar River at Cedar Rapids, IA	6510	109	highlow
25	05572000	Sangamon River at Monticello, IL	550	101	low
26	05586500	Hurricane Creek near Roodhouse, IL	2.3	45	low
27	06062500	Tenmile Creek near Rimini, MT	30.9	94	high
28	06176500	Wolf Creek near Wolf Point, MT	251	37	gage
29	06216500	Pryor Creek near Billings, MT	440	49	high
30	06406000	Battle Creek at Hermosa, SD	178	61	gage
31	06600500	Floyd River at James, IA	886	76	high
32	06710500	Bear Creek at Morrison, CO	164	98	gage
33	06897000	East Fork Big Creek near Bethany, MO	95	53	highlow

Table 1: Streamflow-gaging stations used—Continued

Site	Station No.	Station Name	Drainage Area (mi^2)	No. Peaks	Category
34	06898000	Thompson River at Davis City, IA	701	80	high
35	06933500	Gasconade River at Jerome, MO	2840	92	highlow
36	07067000	Current River at Van Buren, MO	1667	99	high
37	07138600	White Woman Creek trib near Selkirk, KS	38	39	low
38	07203000	Vermejo River near Dawson, NM	301	76	low
39	07208500	Rayado Creek at Sauble Ranch near Cimarron, NM	65	86	gage
40	07382000	Bayou Cocodrie near Clearwater, LA	240	72	gage
41	08133500	North Concho River at Sterling City, TX	588	65	low
42	08150000	Llano River near Junction, TX	1854	91	gage
43	08164000	Lavaca River near Edna, TX	817	73	highlow
44	08167000	Guadalupe River at Comfort, TX	839	76	high
45	08171000	Blanco River at Wimberley, TX	355	87	low
46	08189500	Mission River at Refugio, TX	690	71	low
47	08378500	Pecos River near Pecos, NM	189	87	gage
48	08380500	Gallinas Creek near Montezuma, NM	84	93	gage
49	08387000	Rio Ruidoso at Hollywood, NM	120	56	high
50	09241000	Elk River at Clark, CO	216	78	high
51	09361500	Animas River at Durango, CO	692	102	highlow

Table 1: Streamflow-gaging stations used—Continued

Site	Station No.	Station Name	Drainage Area (mi^2)	No. Peaks	Category
52	09471000	San Pedro River at Charleston, AZ	1234	95	high
53	09480000	Santa Cruz River near Lochiel, AZ	82.2	62	highlow
54	09482500	Santa Cruz River at Tucson, AZ	2222	94	high
55	10128500	Weber River near Oakley, UT	162	105	gage
56	10234500	Beaver River near Beaver, UT	91	97	gage
57	11028500	Santa Maria Creek near Ramona, CA	57.6	71	highlow
58	11152000	Arroyo Seco near Soledad, CA	244	105	gage
59	11176000	Arroyo Mocho near Livermore, CA	38.2	57	highlow
60	11266500	Merced River near Yosemite, CA	321	94	gage
61	11274500	Orestimba Creek near Newman, CA	134	79	low
62	11383500	Deer Creek near Vina, CA	208	94	low
63	11464500	Dry Creek near Cloverdale, CA	87.8	40	low
64	11522500	Salmon River at Somes Bar, CA	751	86	highlow
65	12039500	Quinault River at Quinault Lake, WA	264	97	high
66	12134500	Skykomish River near Gold Bar, WA	535	82	gage
67	12307500	Moyie Creek at Eileen, ID	755	53	low
68	12413000	NF Coeur D'Alene River at Enaville, ID	895	74	high
69	12414500	St. Joe River at Calder, ID	1030	92	gage

Table 1: Streamflow-gaging stations used—Continued

Site	Station No.	Station Name	Drainage Area (mi^2)	No. Peaks	Category
70	12437950	East Fork Foster Creek trib near Bridgeport, WA	4.75	21	highlow
71	12451000	Stehekin River at Stehekin, WA	321	89	gage
72	13185000	Boise River near Twin Springs, ID	830	102	highlow
73	13302500	Salmon River at Salmon, ID	3760	96	gage
74	13343660	Smith Gulch trib near Pataha, WA	1.85	20	highlow
75	14021000	Umatilla River at Pendleton, OR	637	57	gage
76	14048000	John Day River at McDonald Ferry, OR	7580	105	gage
77	14137000	Sandy River near Marmot, OR	263	99	gage
78	14321000	Umpqua River near Elkton, OR	3683	104	low
79	15072000	Fish Creek near Ketchikan, AK	32.1	91	gage
80	16068000	EB of NF Wailua River near Lihue, Kauai, HI	6.27	95	gage
81	16518000	West Wailuaki Stream near Keanae, Maui, HI	3.66	90	gage
82	16587000	Honopou Stream near Huelo, Maui, HI	0.64	98	gage

Table 2: Available data and number of low outliers

Station No.	Record Lengths				Gage Base Discharge			No. Low Outliers	
Station No.	N_{Pk}	N_S	N_H	N_Z	Q_{GgBs}	N_{below}	Q_{crit}	N_{GB}	N_{MGB}
01076500	106	106	0	0	0	0	6316.7	0	6
01350000	100	99	106	1	0	0	2064	0	0
01439500	102	102	0	0	0	0	427.5	0	0
01555500	81	81	0	0	0	0	784.9	0	0
01562000	100	98	122	1	0	0	2977	0	0
01635500	78	78	0	0	0	0	267.5	0	0
01636500	97	95	141	2	0	0	4178.7	0	0
01668000	100	100	0	0	0	0	5209.8	1	2
02037500	76	75	0	0	0	0	4965.1	0	0
02138500	89	88	95	1	0	0	461.7	0	0
02256500	79	79	0	0	0	0	227.9	0	0
03011020	107	107	0	0	0	0	8502.9	0	0
03051000	104	103	123	1	0	0	3992.9	0	0
03159500	81	77	137	1	0	0	3655.4	0	0
03183500	115	115	0	0	0	0	190.8	0	0
03289500	72	70	94	2	0	0	3050.9	1	6
03345500	99	99	0	0	0	0	1698.2	1	35
03550000	101	100	113	1	0	0	862.1	0	0
03558000	85	84	90	1	0	0	1022.6	0	0
03606500	70	66	114	3	0	0	760.8	0	0
04293500	93	91	180	1	0	0	4328.8	0	0
05270500	75	75	0	0	0	0	152.1	1	1

Table 2: Available data and number of low outliers – Continued

Station No.	Record Lengths				Gage Base Discharge			No. Low Outliers	
Station No.	N_{Pk}	N_S	N_H	N_Z	Q_{GgBs}	N_{below}	Q_{crit}	N_{GB}	N_{MGB}
05291000	84	83	101	0	0	0	23.4	0	41
05464500	109	108	311	0	0	0	3187.3	0	5
05572000	102	101	0	0	0	0	755.6	1	1
05586500	45	45	0	0	70	6	38	0	0
06062500	94	94	103	0	0	0	28.1	0	3
06176500	37	37	0	0	0	0	0.5	0	7
06216500	49	48	99	1	0	0	101.6	0	0
06406000	61	61	0	0	0	0	1.8	0	0
06600500	76	76	0	0	0	0	282.2	0	0
06710500	98	98	0	0	0	0	19.3	0	0
06897000	53	52	77	1	0	0	376.2	1	6
06898000	85	79	126	1	0	0	1171.9	0	0
06933500	92	88	114	4	4320	1	4369.7	0	20
07067000	99	98	107	1	0	0	2536.9	0	0
07138600	52	39	0	0	0	4	2.5	0	12
07203000	76	76	0	0	0	0	59.7	1	6
07208500	87	86	0	0	0	0	7	0	0
07382000	73	72	0	0	0	0	272.6	0	0
08133500	65	65	0	0	300	20	130.7	0	23
08150000	91	91	0	0	0	0	34.4	0	45
08164000	73	72	75	1	0	0	734.7	1	1
08167000	77	72	163	1	0	0	104.9	0	23
08171000	85	84	142	1	0	0	57.6	1	28

Table 2: Available data and number of low outliers – Continued

Station No.	Record Lengths				Gage Base Discharge			No. Low Outliers	
Station No.	N_{Pk}	N_S	N_H	N_Z	Q_{GgBs}	N_{below}	Q_{crit}	N_{GB}	N_{MGB}
08189500	71	71	0	0	0	0	65.9	1	12
08378500	87	85	90	2	0	0	70.3	0	0
08380500	93	93	0	0	0	0	16.2	0	0
08387000	56	56	0	0	0	0	19.1	0	0
09241000	78	78	0	0	0	0	1025.5	0	22
09361500	102	100	113	2	0	0	1374.4	1	2
09471000	95	95	0	0	0	0	490.9	1	3
09480000	62	62	0	0	0	0	10.1	2	8
09482500	94	93	96	1	0	0	559.6	0	0
10128500	105	105	0	0	0	0	579.8	0	24
10234500	97	97	0	0	0	0	40.4	0	48
11028500	71	71	0	0	0	4	0	1	13
11152000	105	105	0	0	0	0	447.9	0	52
11176000	57	56	89	1	0	0	2.4	1	28
11266500	94	94	0	0	0	0	819.6	0	0
11274500	79	79	0	0	0	11	10.8	1	39
11383500	94	94	0	0	0	0	521.4	1	1
11464500	40	39	43	1	0	0	693.5	1	18
11522500	86	86	0	0	0	0	1923.5	1	1
12039500	97	96	101	1	0	0	6220.4	0	48
12134500	82	82	0	0	0	0	9162.2	0	0
12307500	53	53	0	0	0	0	2227.9	1	12
12413000	74	71	99	3	0	0	3271.3	0	0

Table 2: Available data and number of low outliers – Continued

Station No.	Record Lengths				Gage Base Discharge			No. Low Outliers	
Station No.	N_{Pk}	N_S	N_H	N_Z	Q_{GgBs}	N_{below}	Q_{crit}	N_{GB}	N_{MGB}
12414500	92	92	0	0	0	0	4310.5	0	0
12437950	21	21	0	0	0	0	0.8	0	0
12451000	89	89	0	0	0	0	3396.8	0	0
13185000	102	100	140	2	0	0	1840.6	1	36
13302500	96	96	0	0	0	0	2039.2	0	45
13343660	20	20	0	0	0	1	0.7	1	3
14021000	57	57	0	0	0	0	1066.7	0	0
14048000	105	105	0	0	0	0	2549.2	0	0
14137000	99	99	0	0	0	0	2894.4	0	0
14321000	104	104	0	0	0	0	8022.5	2	9
15072000	91	91	0	0	0	0	1352.2	0	0
16068000	95	95	0	0	0	0	418.8	0	6
16518000	90	90	0	0	0	0	926.2	0	19
16587000	98	98	0	0	0	0	36.7	0	0

B Graphical Comparisons Between EMA and B17B at the 82 Test Sites

B.1 Systematic Data Sites

Figure 26: Site 01439500 with Systematic Data Only

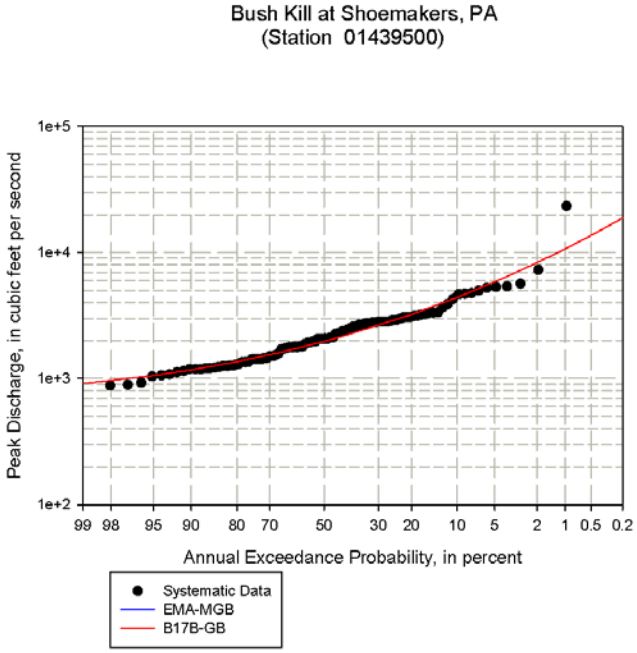


Figure 27: Site 01555500 with Systematic Data Only

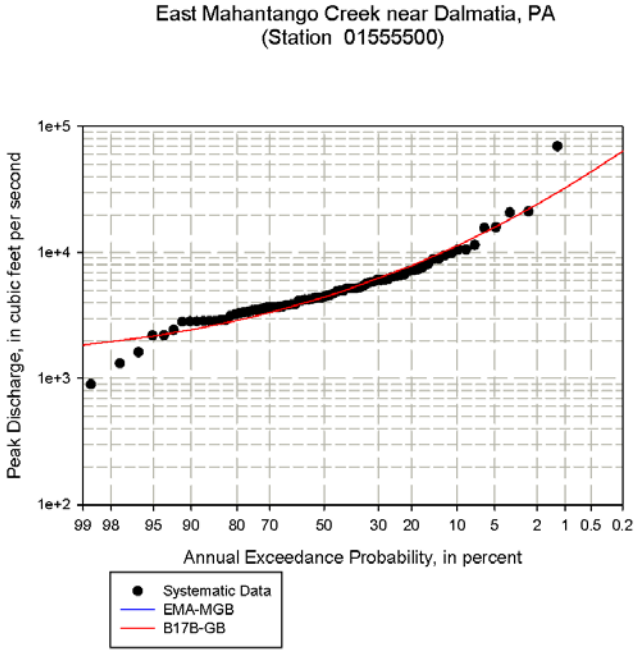


Figure 28: Site 01635500 with Systematic Data Only

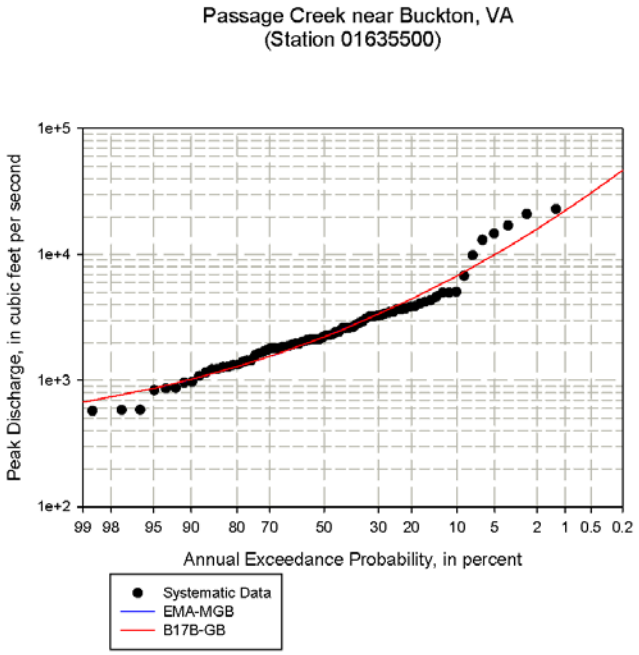


Figure 29: Site 02037500 with Systematic Data Only

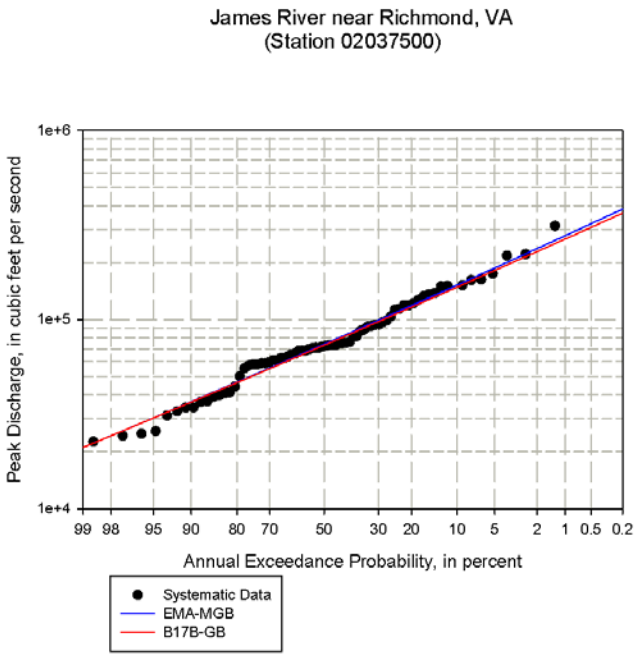


Figure 30: Site 02256500 with Systematic Data Only

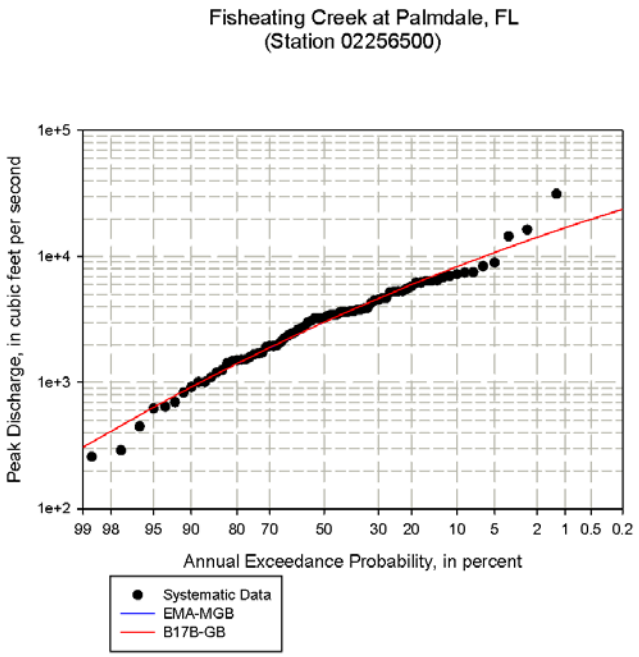


Figure 31: Site 03183500 with Systematic Data Only

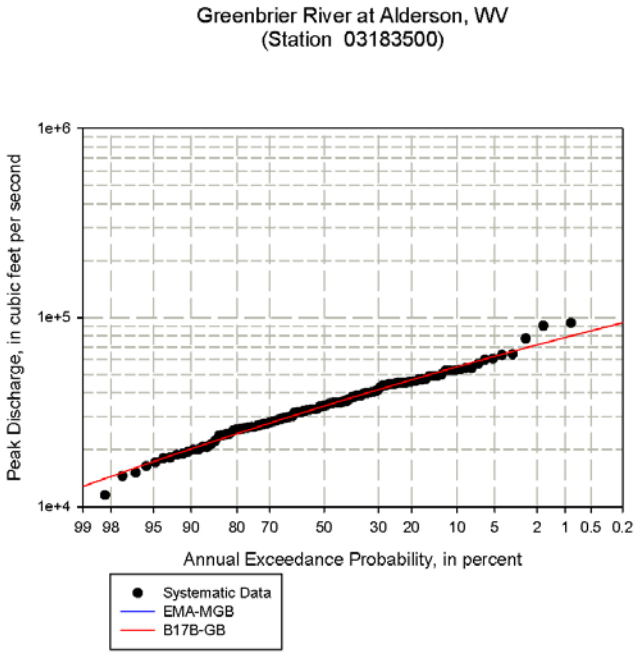


Figure 32: Site 05586500 with Systematic Data Only; 6 Peaks Identified as Below Gage Base Discharge

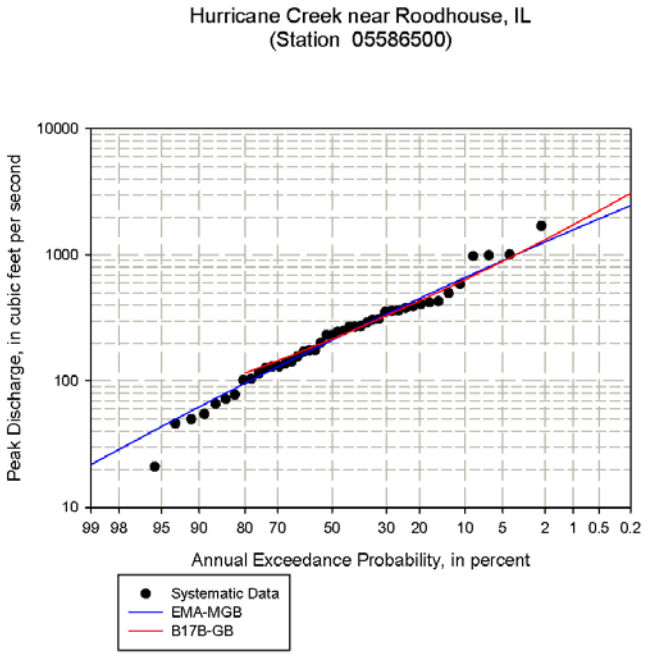


Figure 33: Site 06406000 with Systematic Data Only

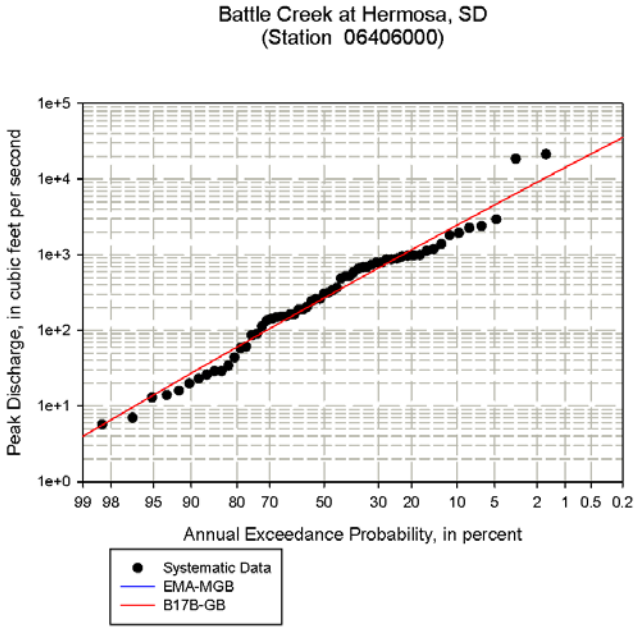


Figure 34: Site 06710500 with Systematic Data Only

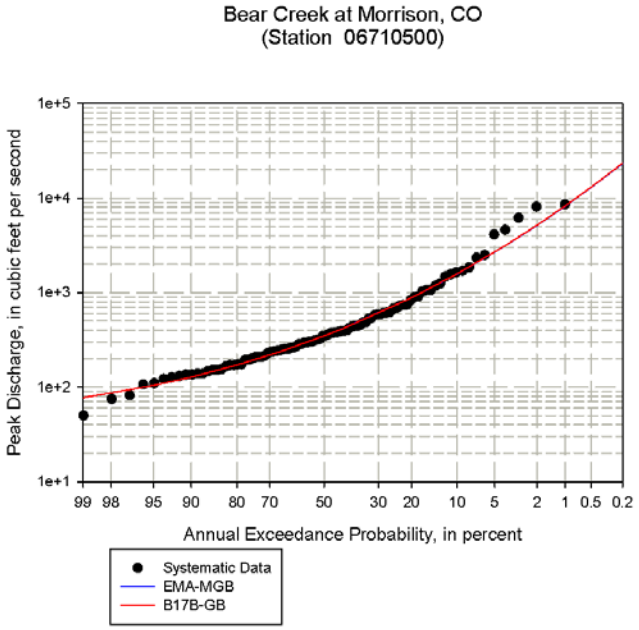


Figure 35: Site 07208500 with Systematic Data Only

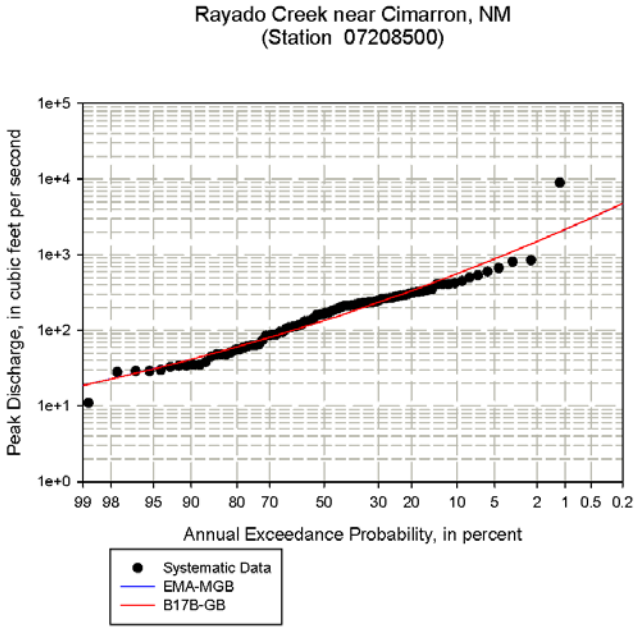


Figure 36: Site 07382000 with Systematic Data Only

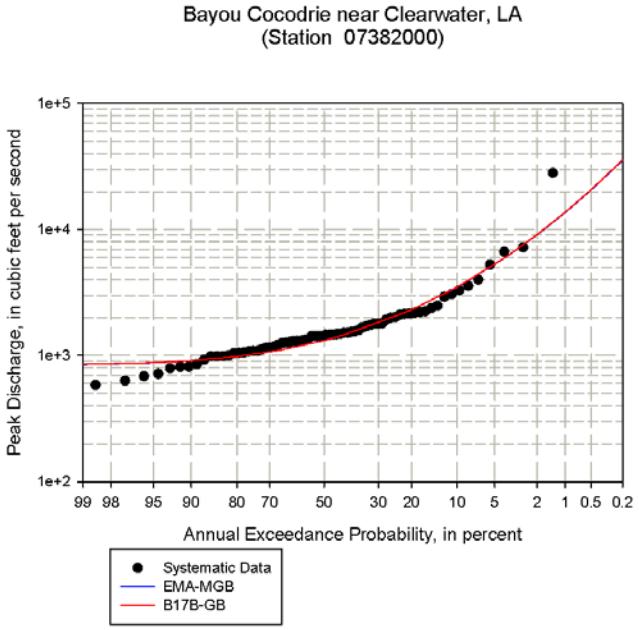


Figure 37: Site 08380500 with Systematic Data Only

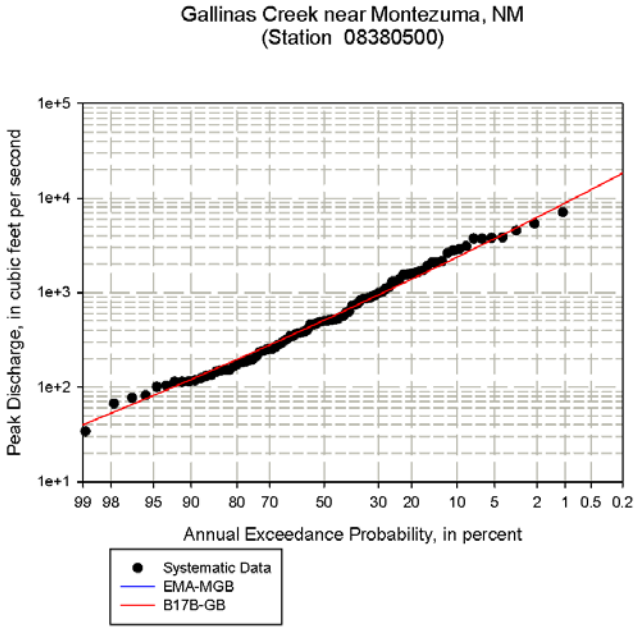


Figure 38: Site 08387000 with Systematic Data Only

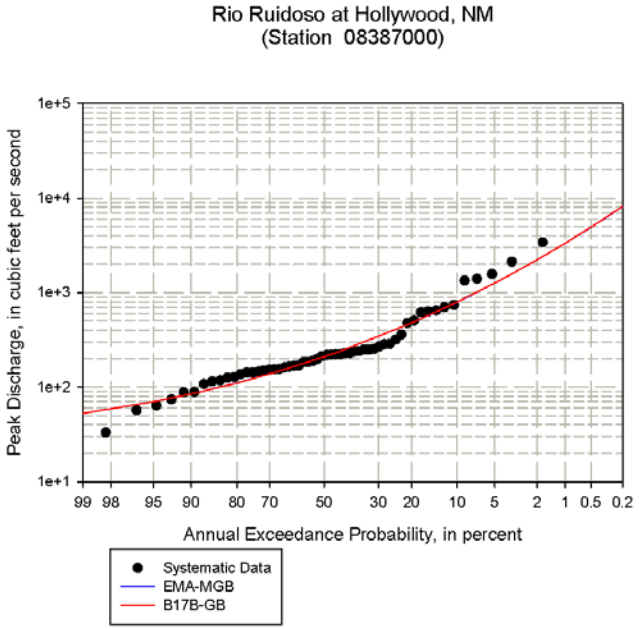


Figure 39: Site 11266500 with Systematic Data Only

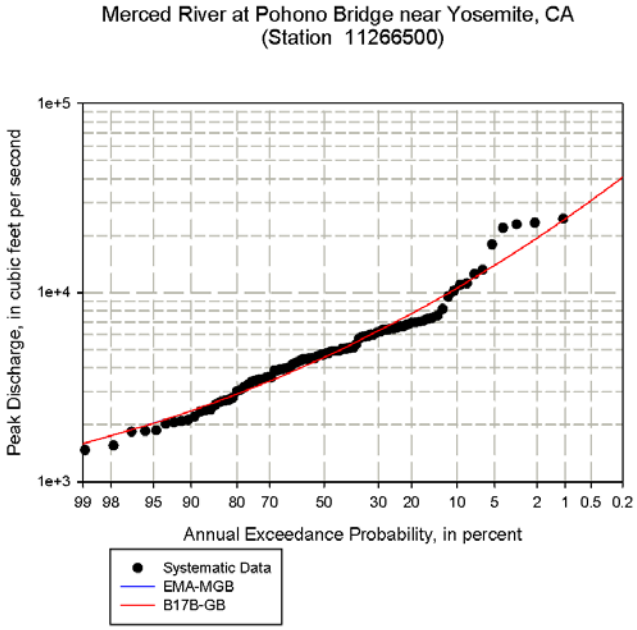


Figure 40: Site 12134500 with Systematic Data Only

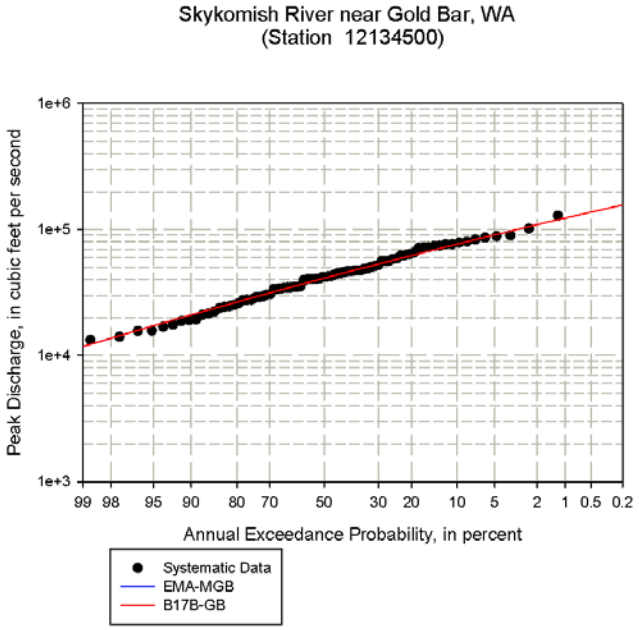


Figure 41: Site 12414500 with Systematic Data Only

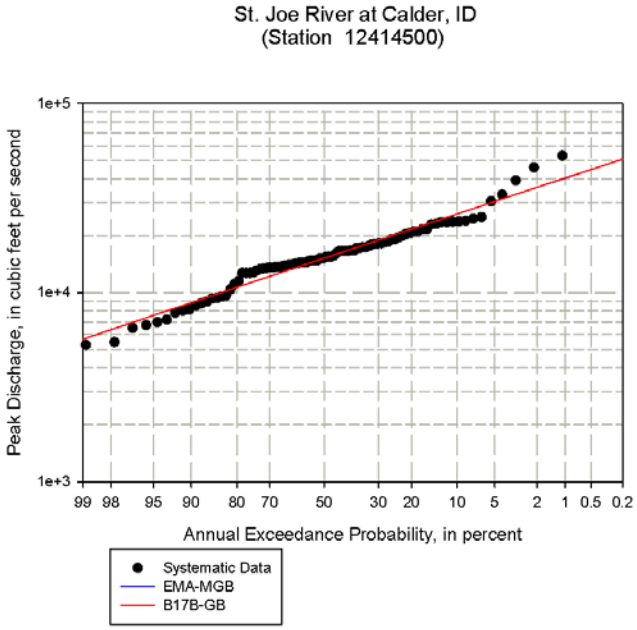


Figure 42: Site 12437950 with Systematic Data Only

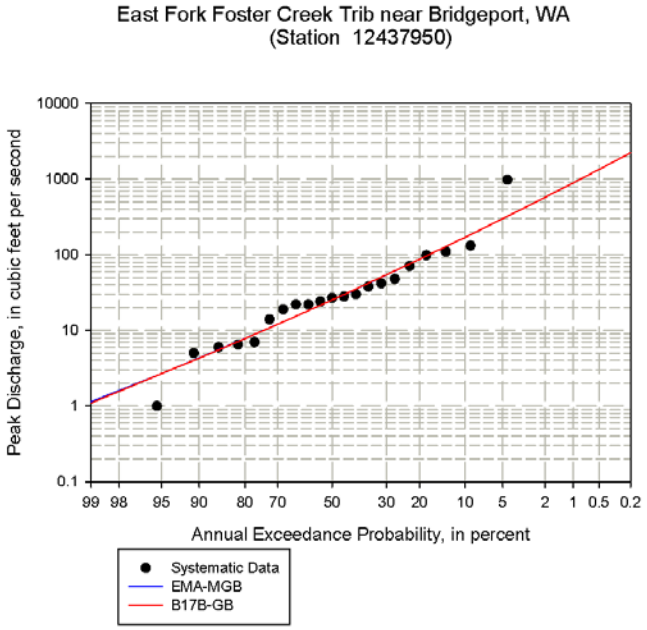


Figure 43: Site 12451000 with Systematic Data Only

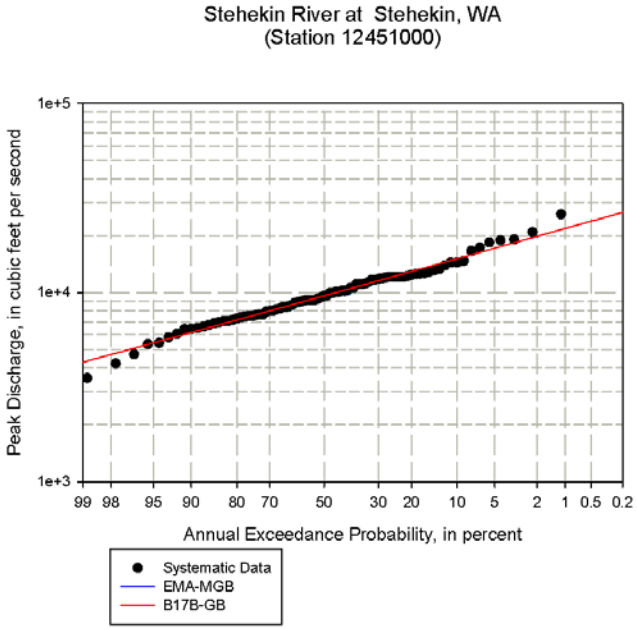


Figure 44: Site 14021000 with Systematic Data Only

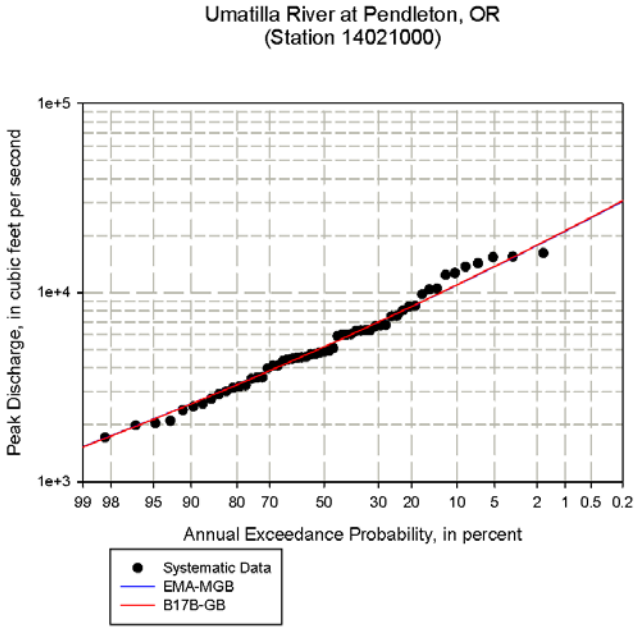


Figure 45: Site 14048000 with Systematic Data Only

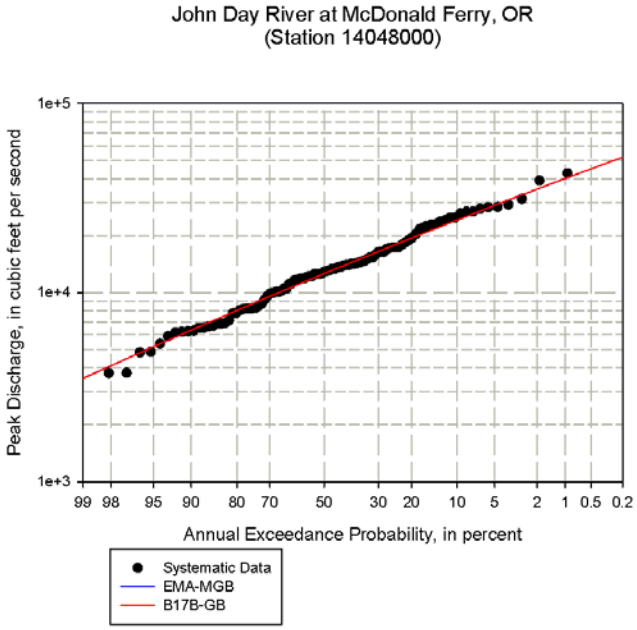


Figure 46: Site 14137000 with Systematic Data Only

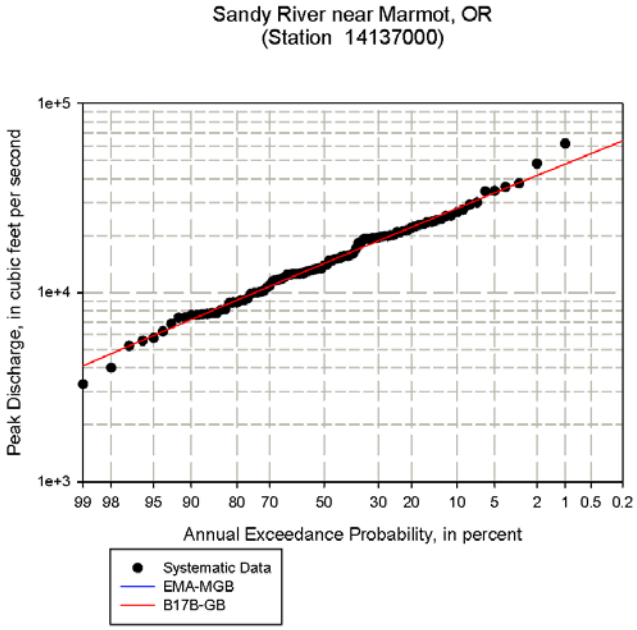


Figure 47: Site 15072000 with Systematic Data Only

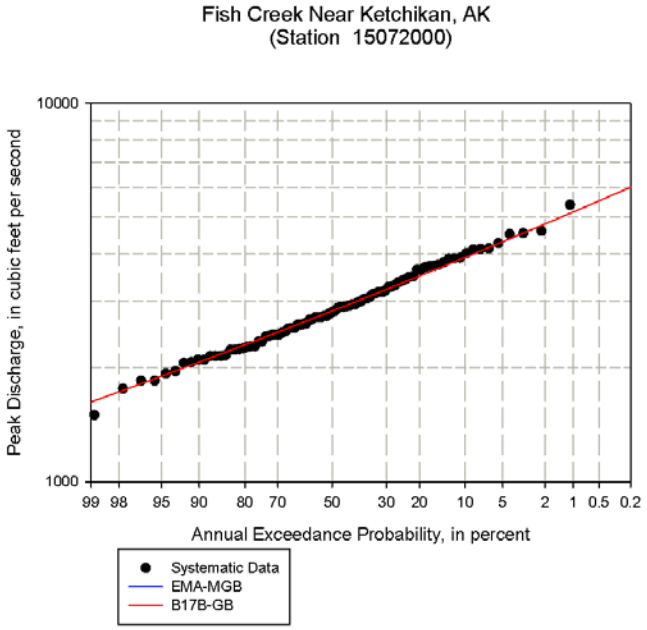
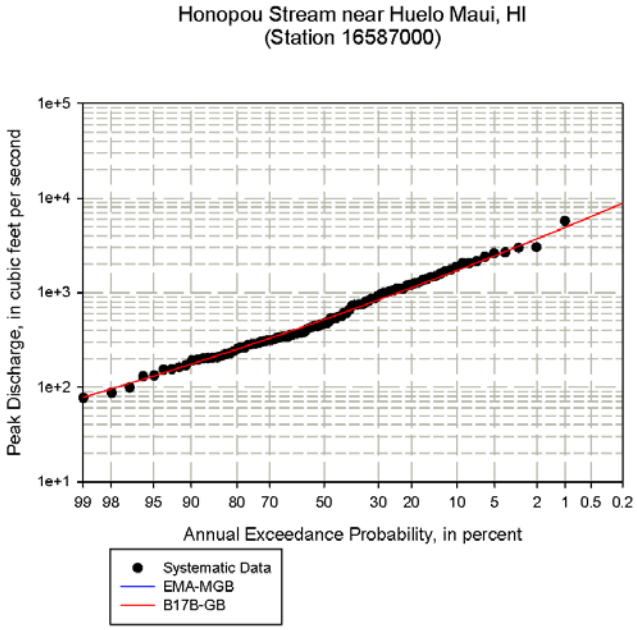


Figure 48: Site 16587000 with Systematic Data Only



B.2 Sites with Historical Information

Figure 49: Site 01350000 with Systematic and Historical Data

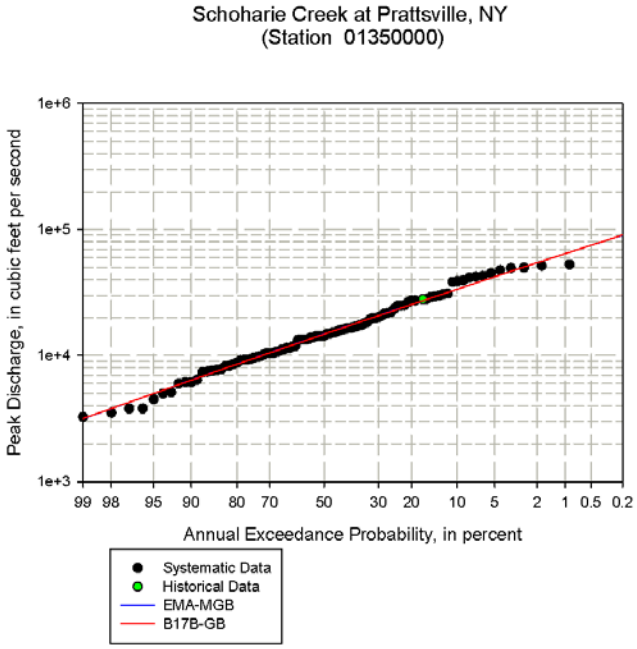


Figure 50: Site 01562000 with Systematic and Historical Data

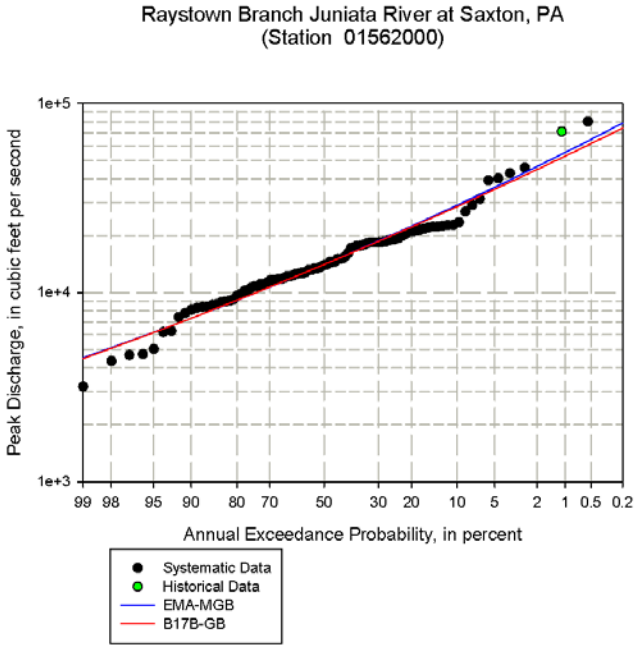


Figure 51: Site 01636500 with Systematic and Historical Data

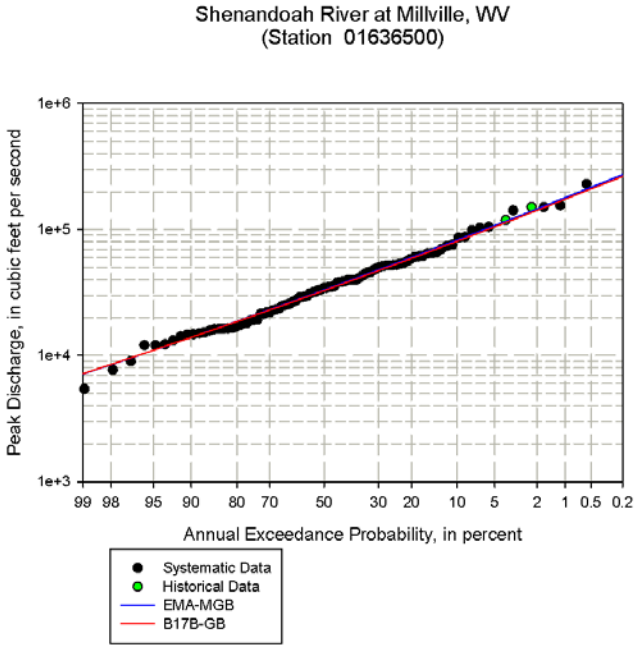


Figure 52: Site 02138500 with Systematic and Historical Data

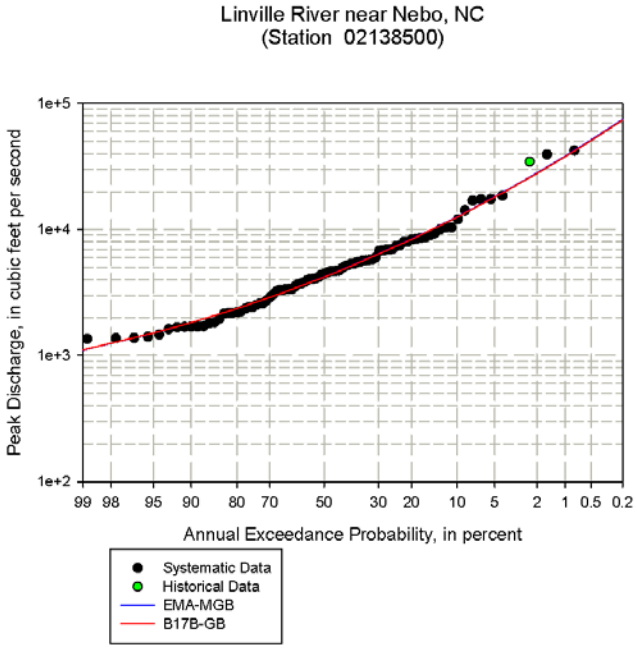


Figure 53: Site 03011020 with Systematic and Historical Data

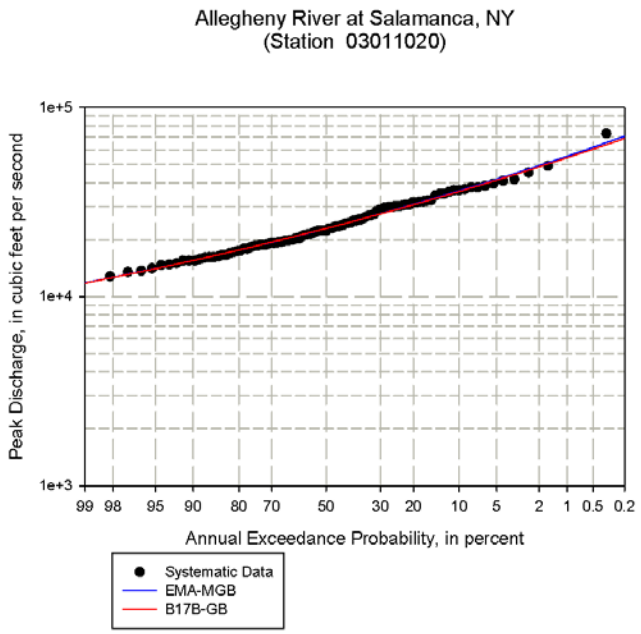


Figure 54: Site 03051000 with Systematic and Historical Data

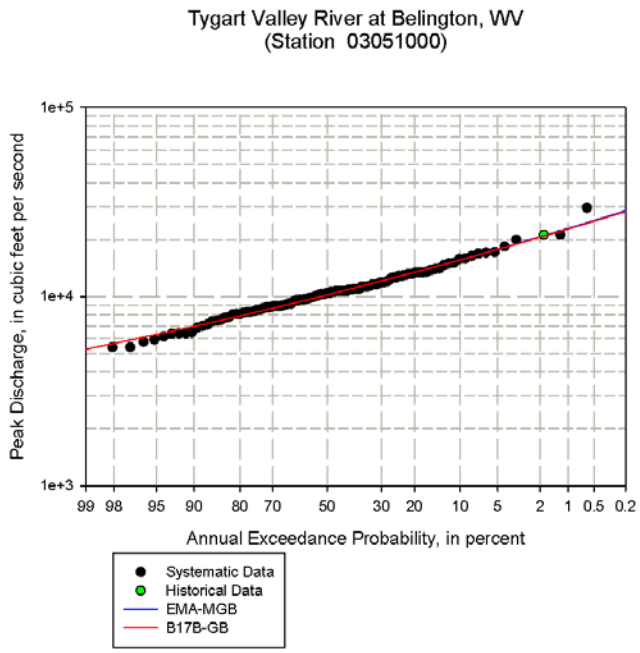


Figure 55: Site 03159500 with Systematic and Historical Data

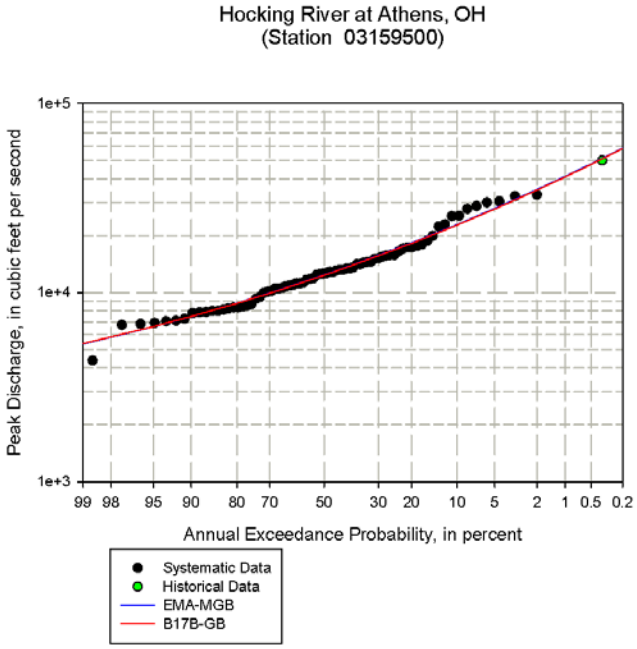


Figure 56: Site 03550000 with Systematic and Historical Data

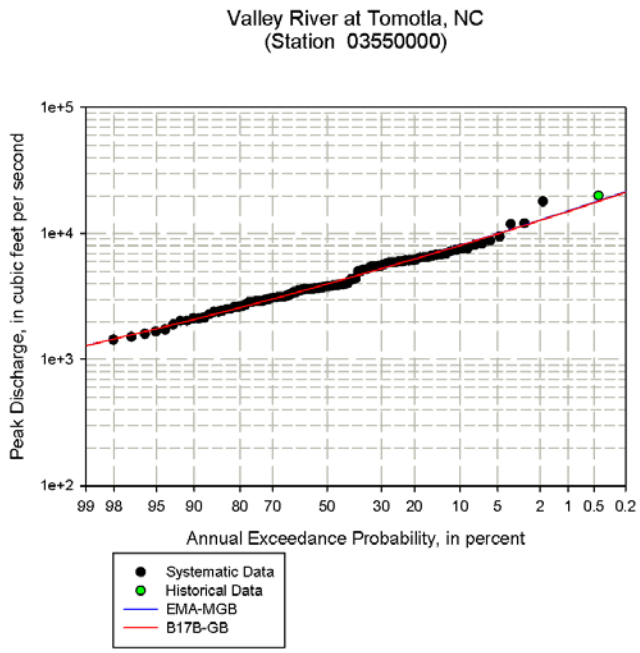


Figure 57: Site 03558000 with Systematic and Historical Data

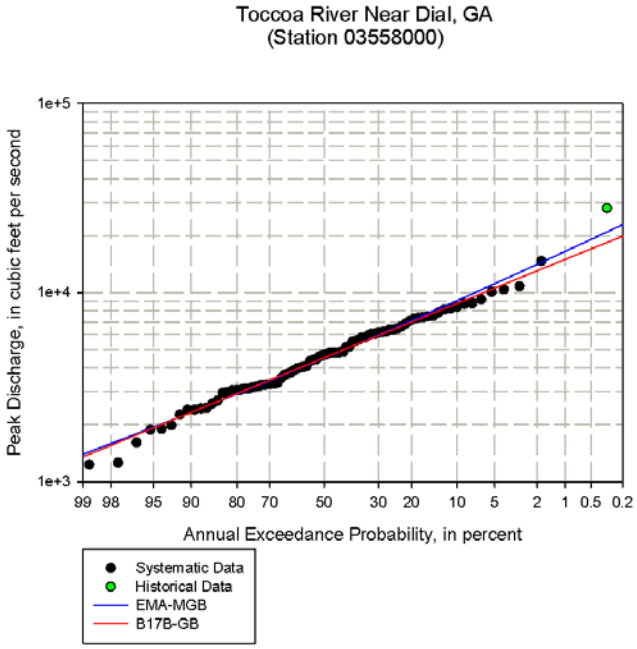


Figure 58: Site 03606500 with Systematic and Historical Data

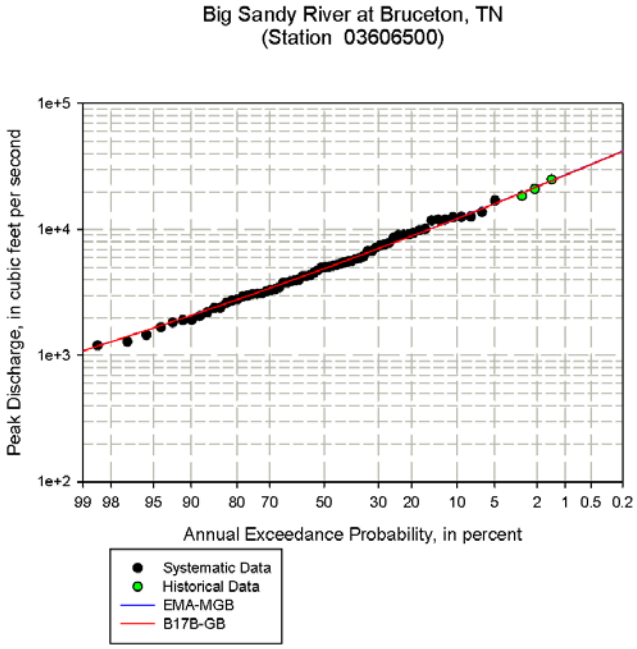


Figure 59: Site 04293500 with Systematic and Historical Data

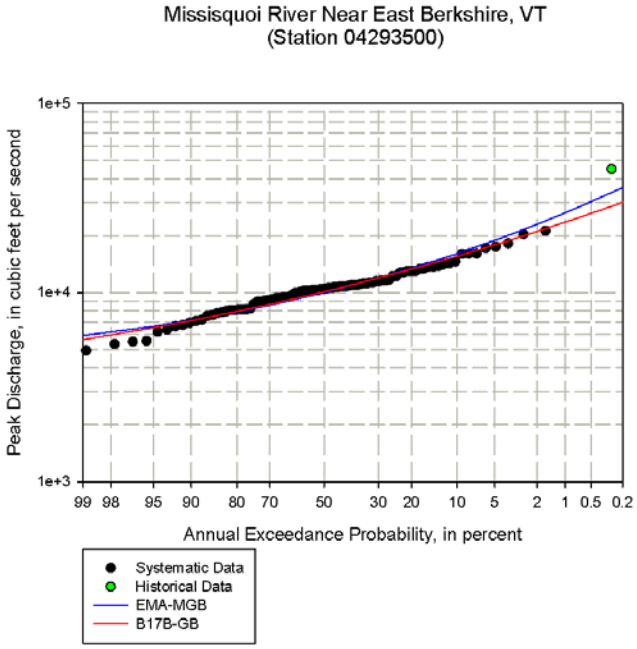


Figure 60: Site 06216500 with Systematic and Historical Data

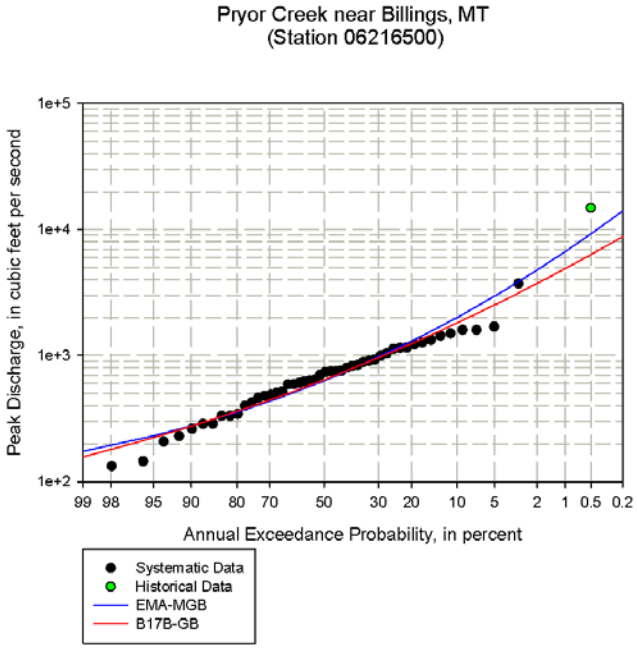


Figure 61: Site 06600500 with Systematic and Historical Data

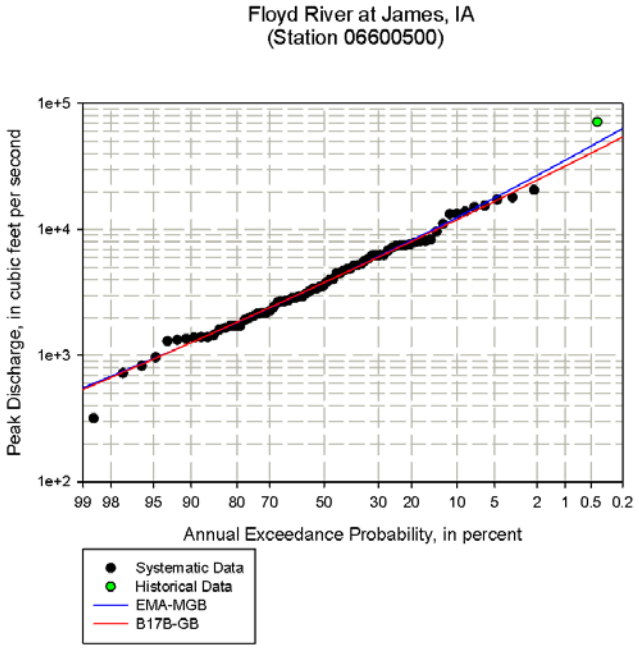


Figure 62: Site 06898000 with Systematic and Historical Data

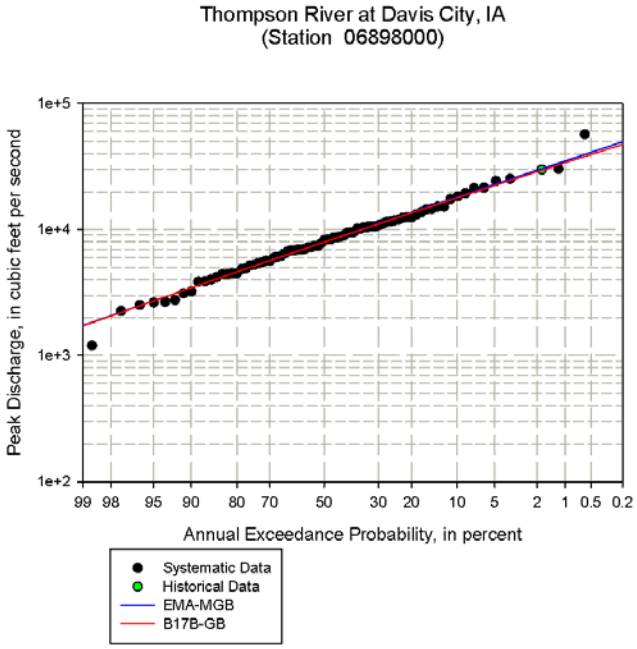


Figure 63: Site 07067000 with Systematic and Historical Data

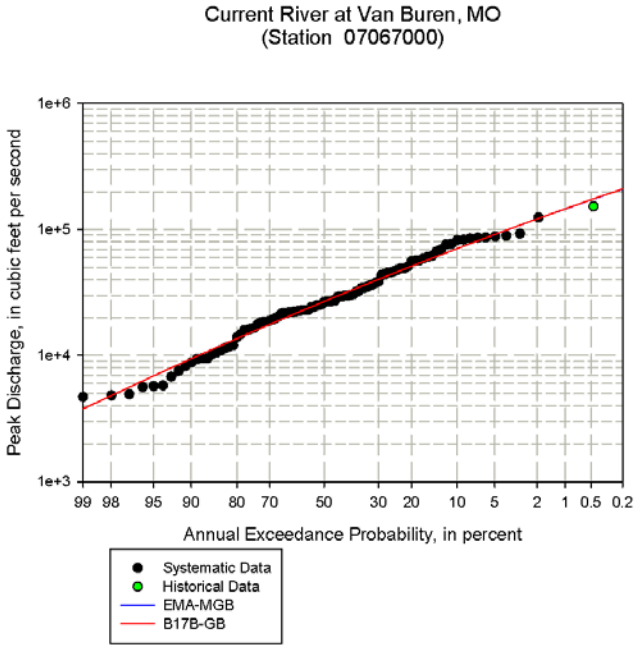


Figure 64: Site 08378500 with Systematic and Historical Data

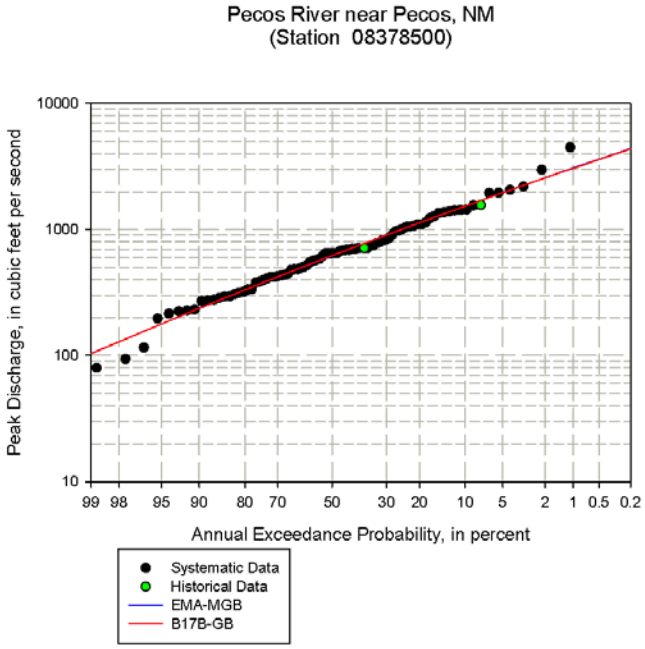


Figure 65: Site 09482500 with Systematic and Historical Data

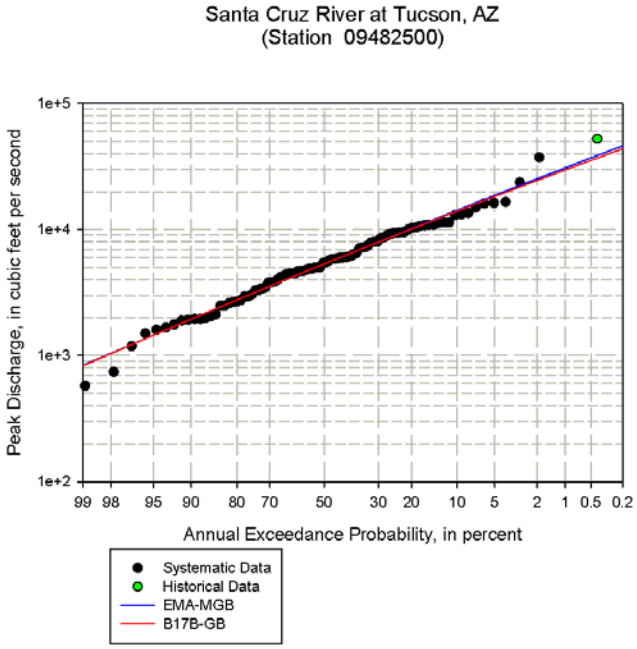
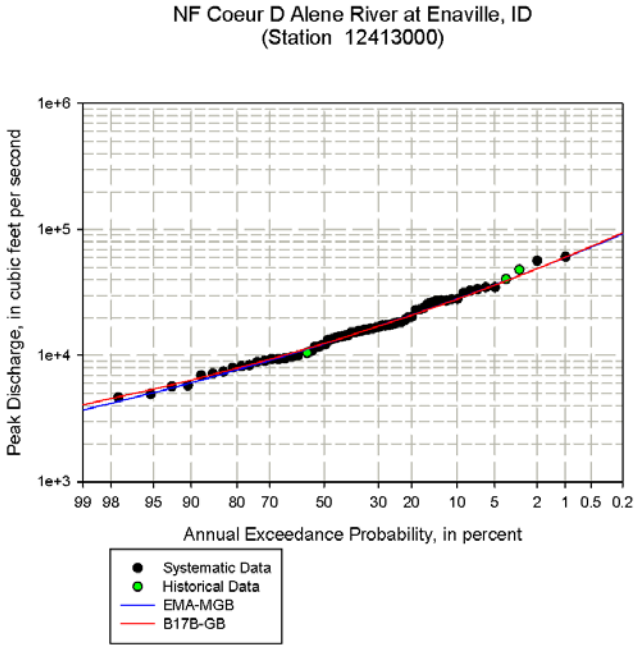


Figure 66: Site 12413000 with Systematic and Historical Data



B.3 Sites with Low Outliers

Figure 67: Site 01076500 with Systematic Data; Multiple Grubbs-Beck Test Identified 6 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers

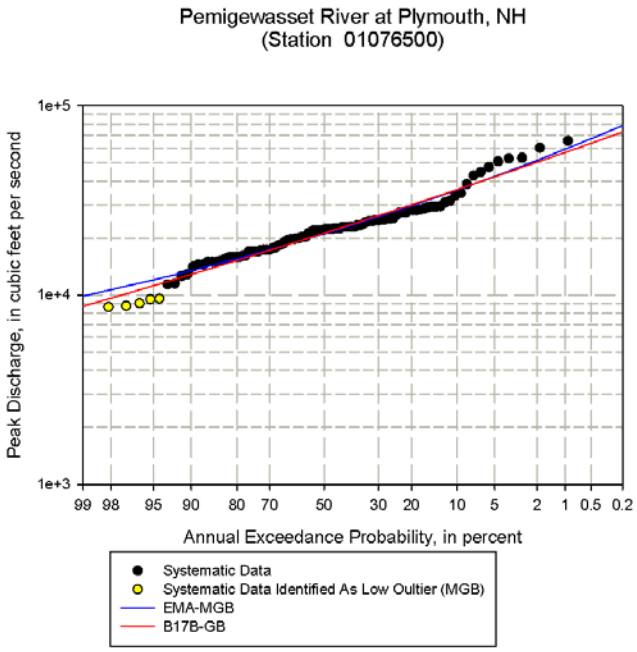


Figure 68: Site 01668000 with Systematic Data; Multiple Grubbs-Beck Test Identified 2 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

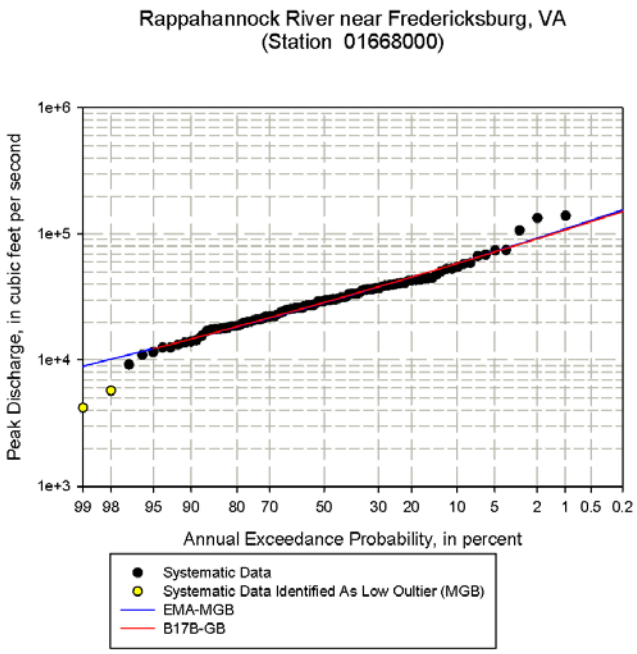


Figure 69: Site 03345500 with Systematic Data; Multiple Grubbs-Beck Test Identified 35 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

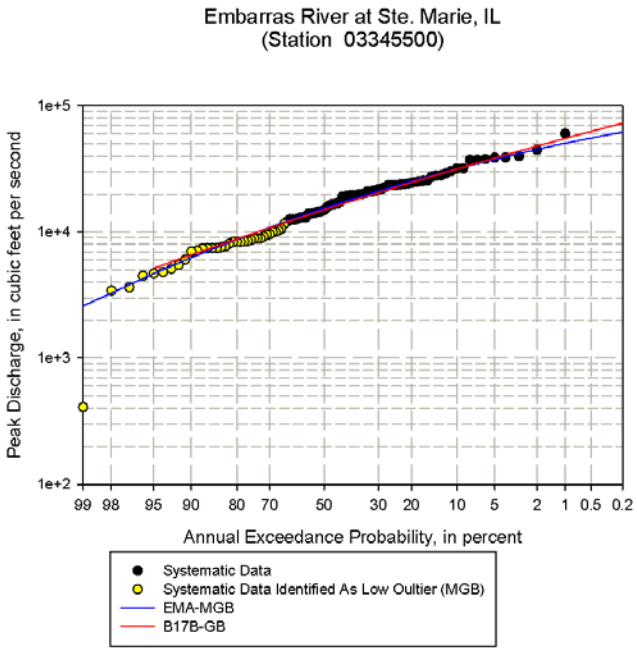


Figure 70: Site 05572000 with Systematic Data; Multiple Grubbs-Beck Test Identified 1 Low Outlier; Grubbs-Beck Test Identified 1 Low Outlier

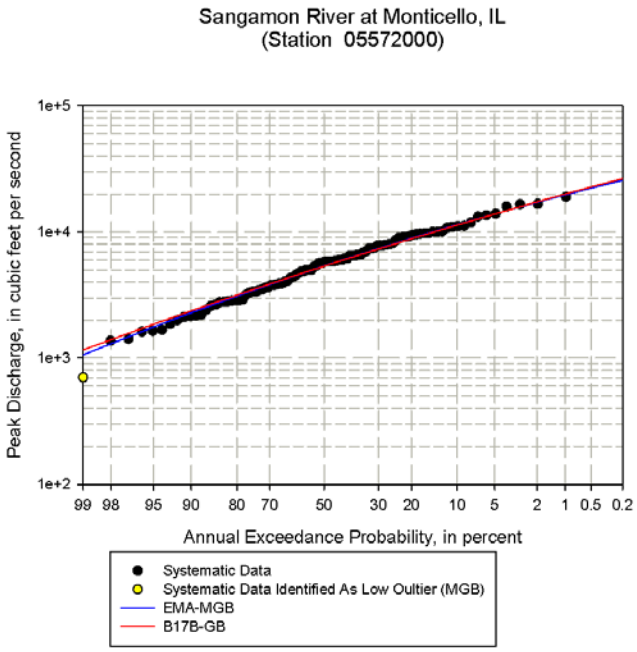


Figure 71: Site 06176500 with Systematic Data; Multiple Grubbs-Beck Test Identified 7 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers

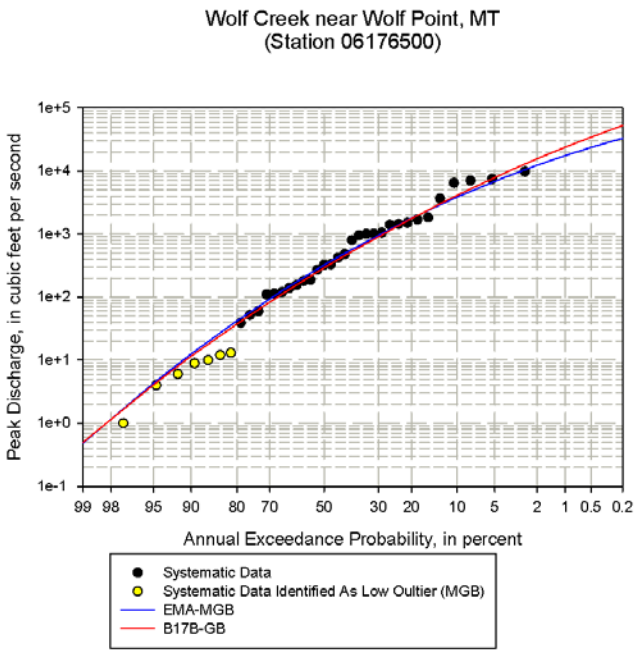


Figure 72: Site 07203000 with Systematic Data; Multiple Grubbs-Beck Test Identified 6 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

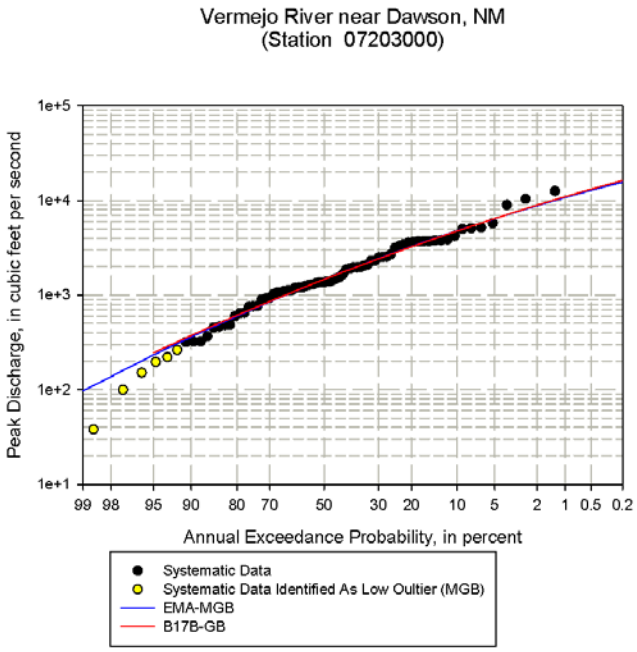


Figure 73: Site 08133500 with Systematic Data; Multiple Grubbs-Beck Test Identified 23 Low Outliers; 20 Peaks Identified as Below Gage Base Discharge; Grubbs-Beck Test Identified 0 Low Outliers

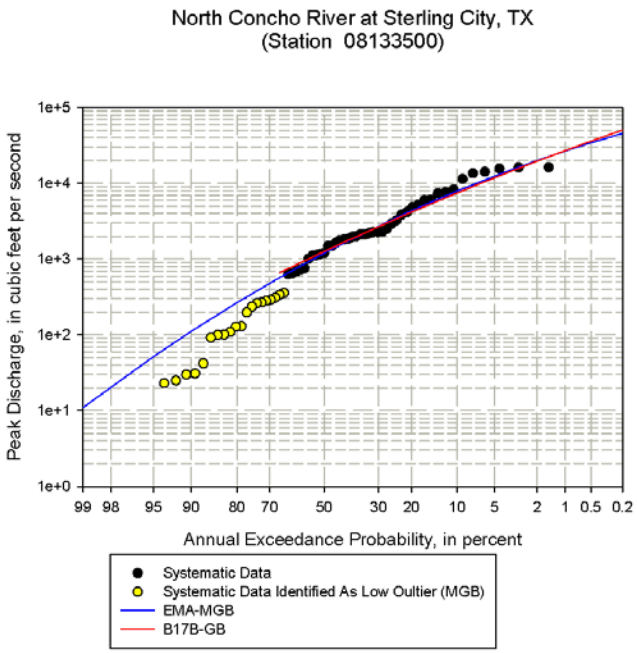


Figure 74: Site 08150000 with Systematic Data; Multiple Grubbs-Beck Test Identified 45 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers

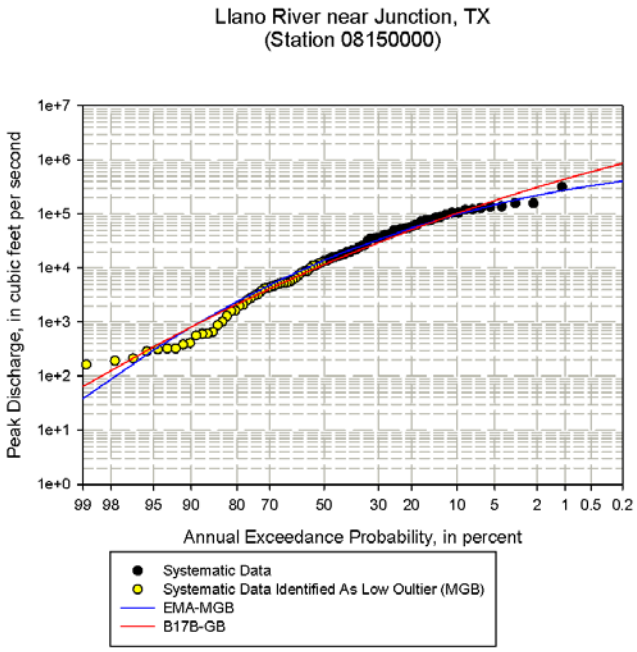


Figure 75: Site 08189500 with Systematic Data; Multiple Grubbs-Beck Test Identified 12 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

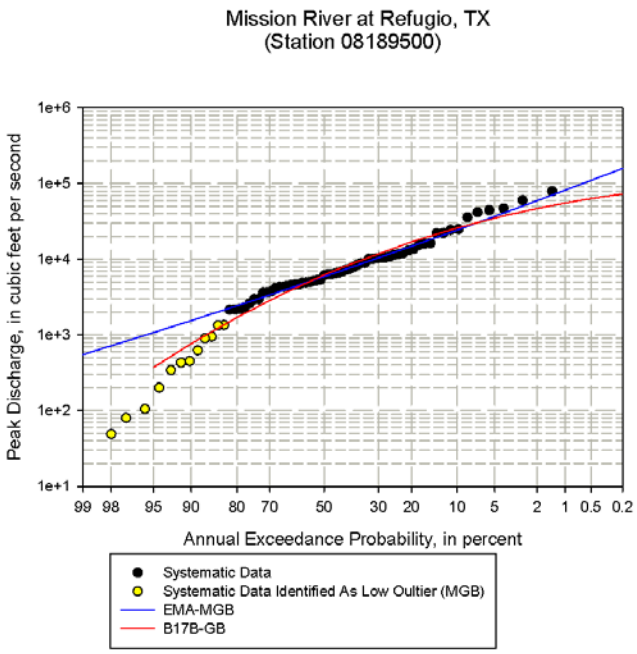


Figure 76: Site 09241000 with Systematic Data; Multiple Grubbs-Beck Test Identified 22 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers

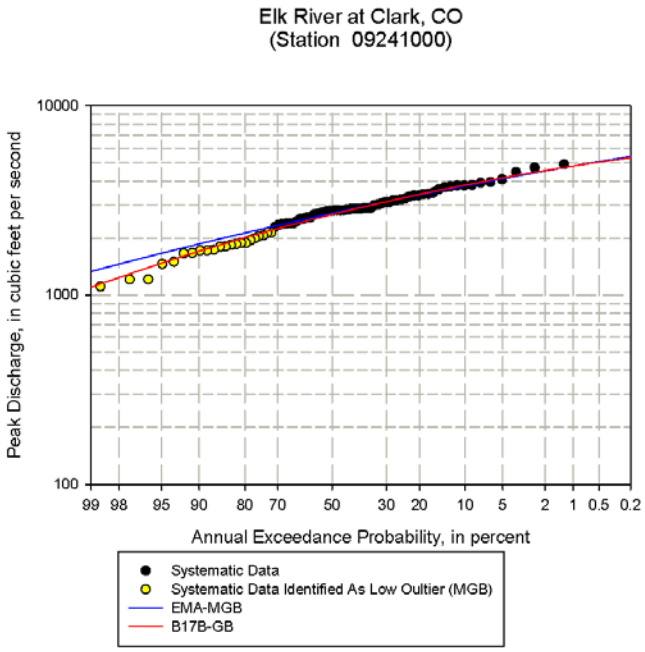


Figure 77: Site 10128500 with Systematic Data; Multiple Grubbs-Beck Test Identified 24 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers

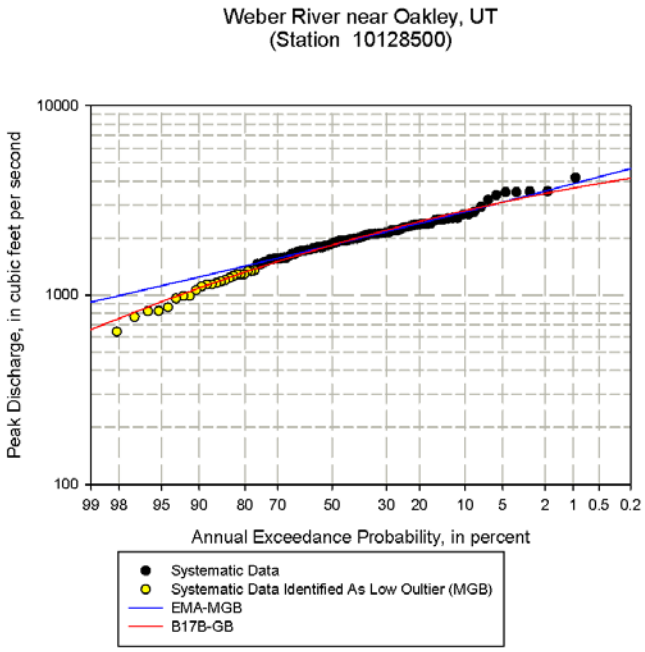


Figure 78: Site 10234500 with Systematic Data; Multiple Grubbs-Beck Test Identified 48 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers

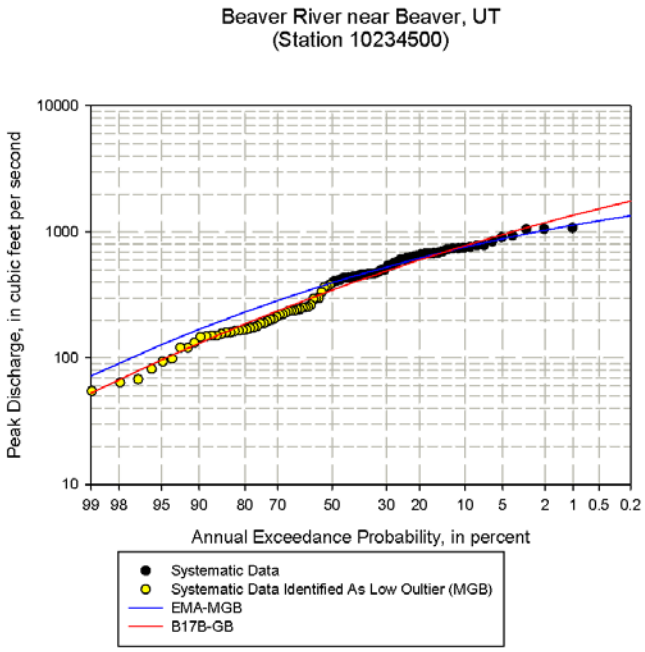


Figure 79: Site 11028500 with Systematic Data; Multiple Grubbs-Beck Test Identified 13 Low Outliers; 4 Peaks Identified as Below Gage Base Discharge; Grubbs-Beck Test Identified 1 Low Outlier

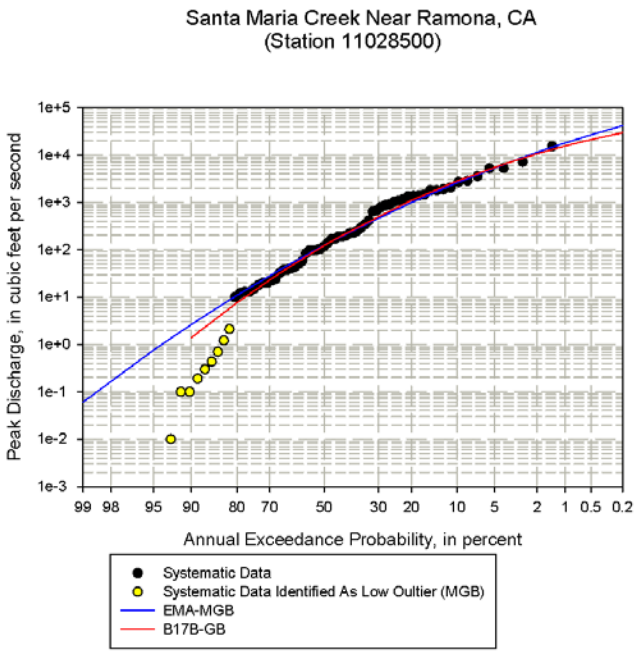


Figure 80: Site 11152000 with Systematic Data; Multiple Grubbs-Beck Test Identified 52 Low Outlier; Grubbs-Beck Test Identified 0 Low Outliers

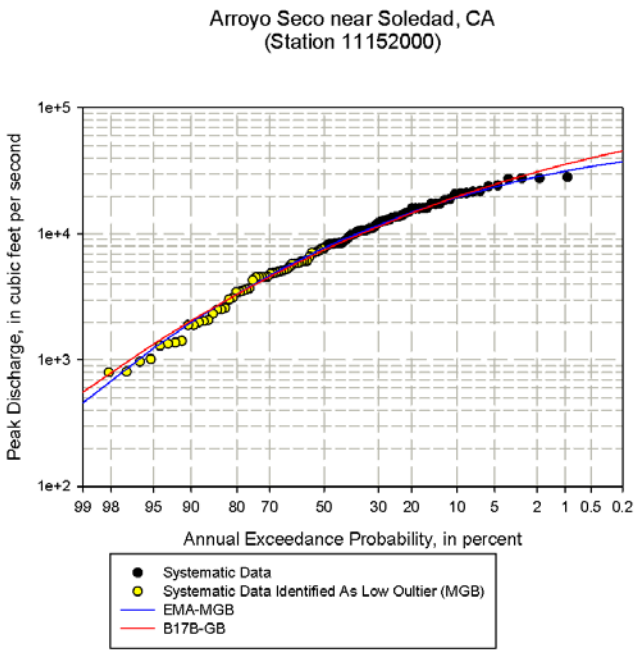


Figure 81: Site 11274500 with Systematic Data; Multiple Grubbs-Beck Test Identified 39 Low Outliers; 11 Peaks Identified as Below Gage Base Discharge; Grubbs-Beck Test Identified 1 Low Outlier

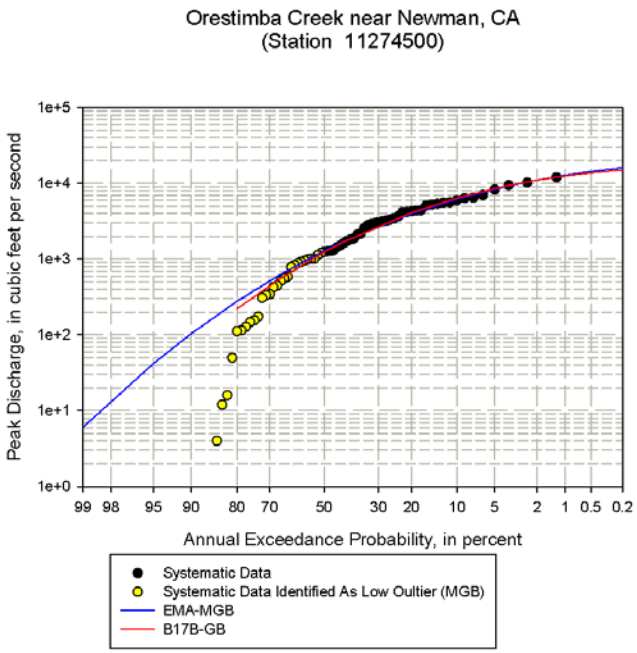


Figure 82: Site 11383500 with Systematic Data; Multiple Grubbs-Beck Test Identified 1 Low Outlier; Grubbs-Beck Test Identified 1 Low Outlier

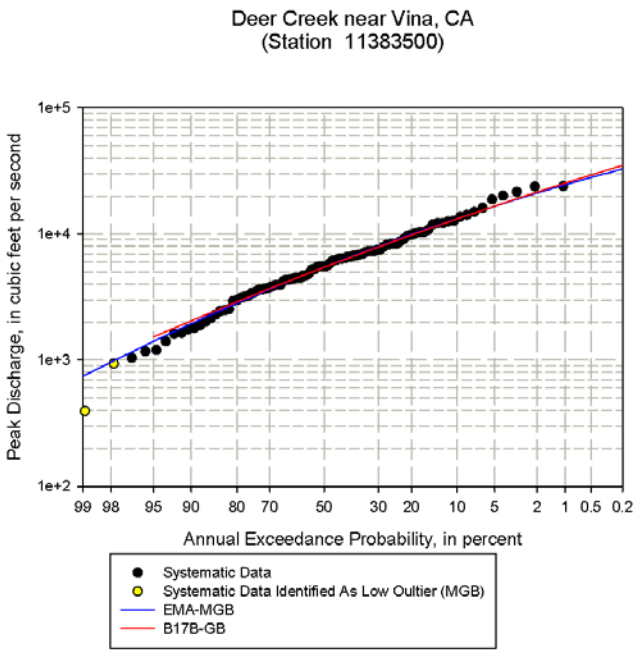


Figure 83: Site 12307500 with Systematic Data; Multiple Grubbs-Beck Test Identified 12 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

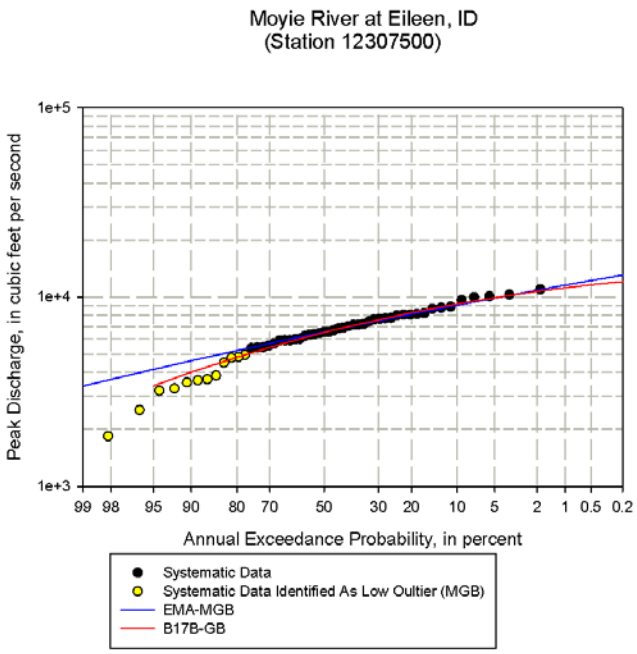


Figure 84: Site 13302500 with Systematic Data; Multiple Grubbs-Beck Test Identified 45 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers

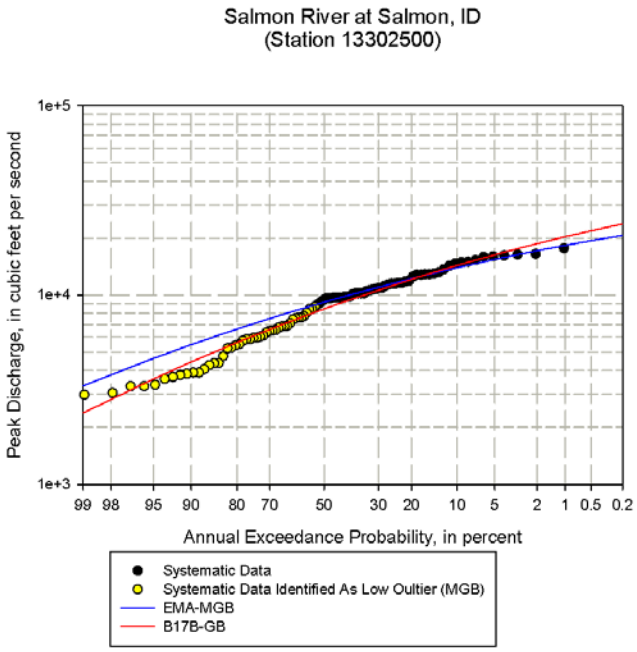


Figure 85: Site 14321000 with Systematic Data; Multiple Grubbs-Beck Test Identified 9 Low Outliers; Grubbs-Beck Test Identified 2 Low Outliers

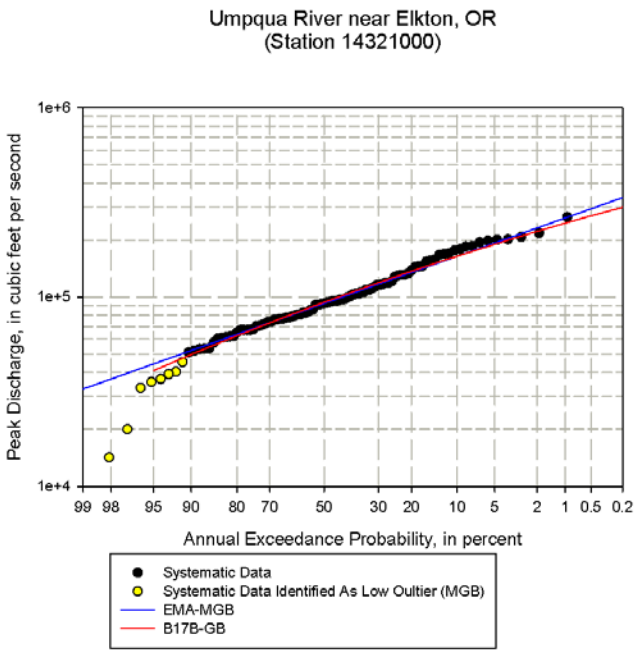
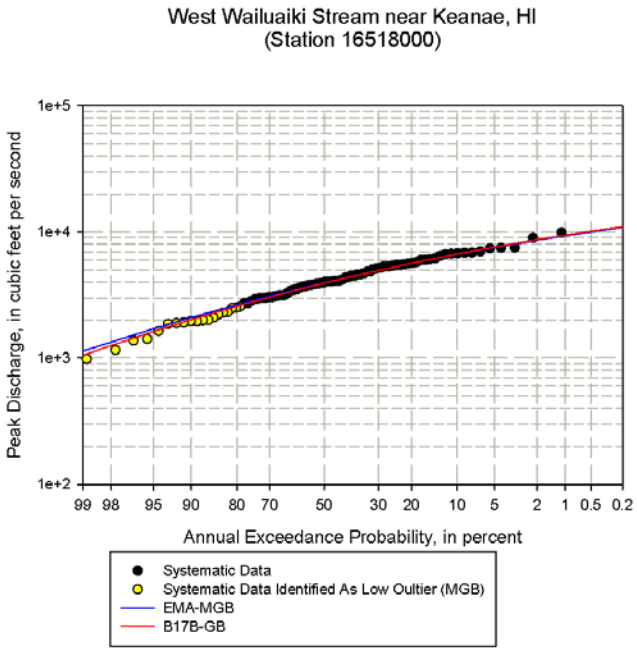


Figure 86: Site 16518000 with Systematic Data; Multiple Grubbs-Beck Test Identified 19 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers



B.4 Sites with a Combination of Data Types

Figure 87: Site 03289500 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 6 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

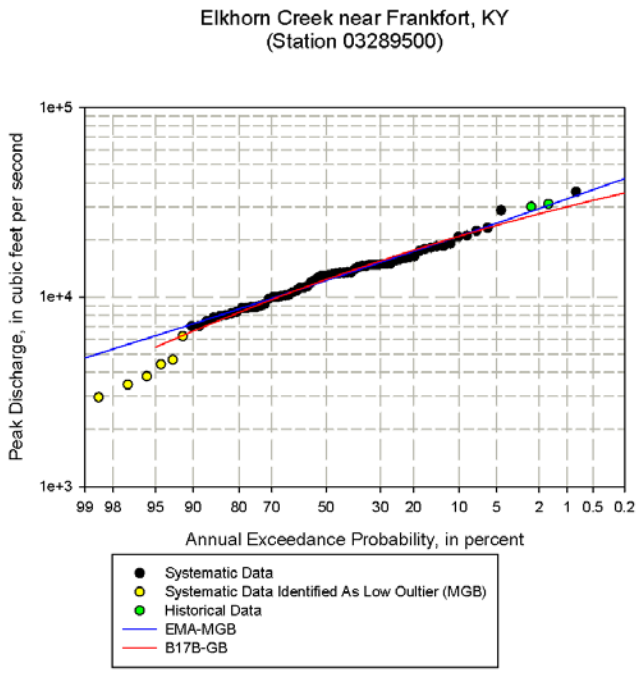


Figure 88: Site 05270500 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 1 Low Outlier; Grubbs-Beck Test Identified 1 Low Outlier

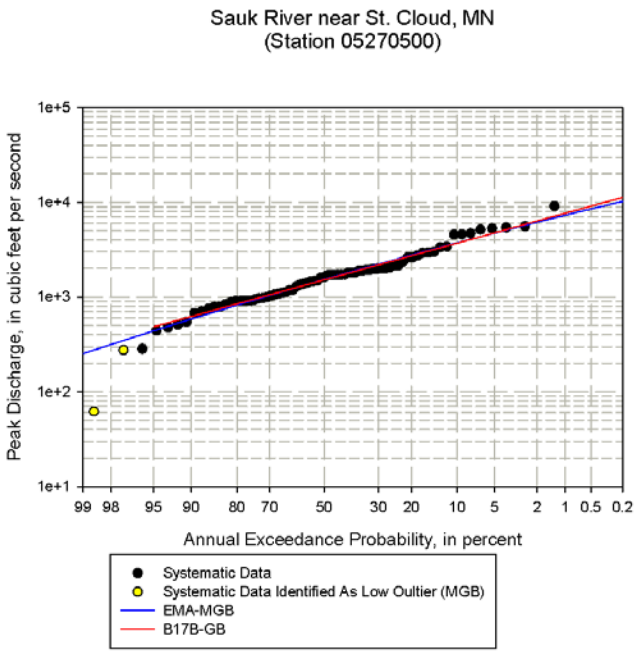


Figure 89: Site 05291000 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 41 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers

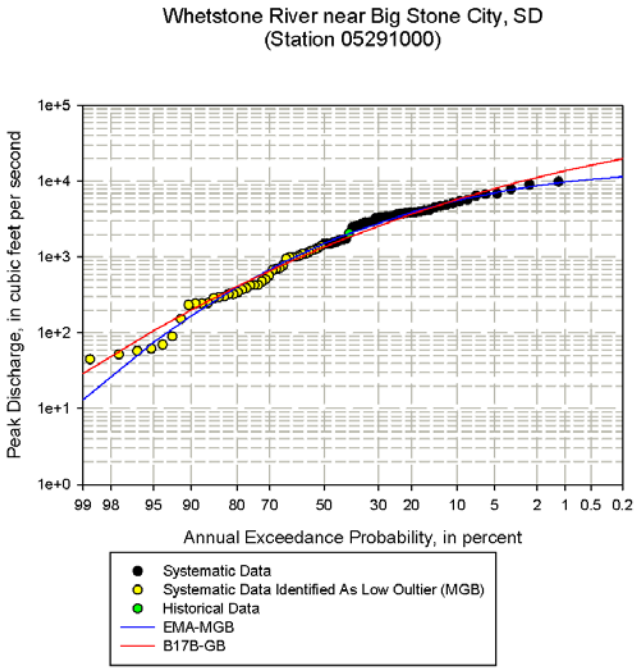


Figure 90: Site 05464500 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 5 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers

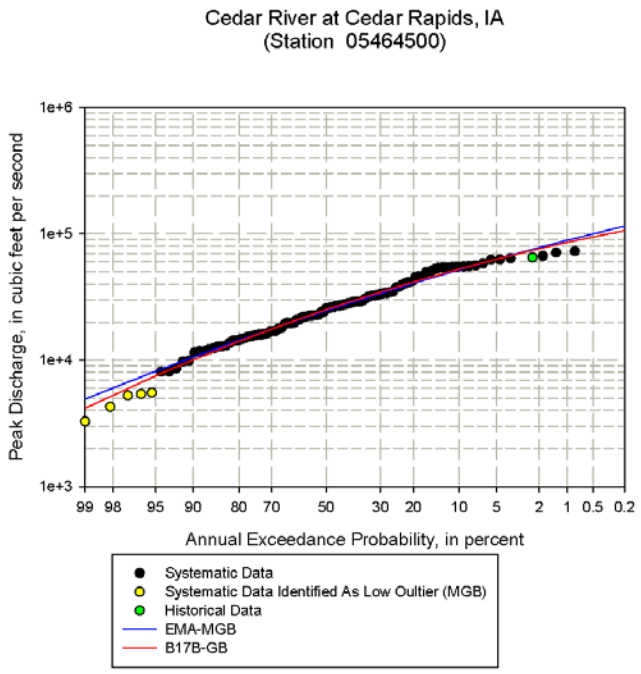


Figure 91: Site 06062500 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 3 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers

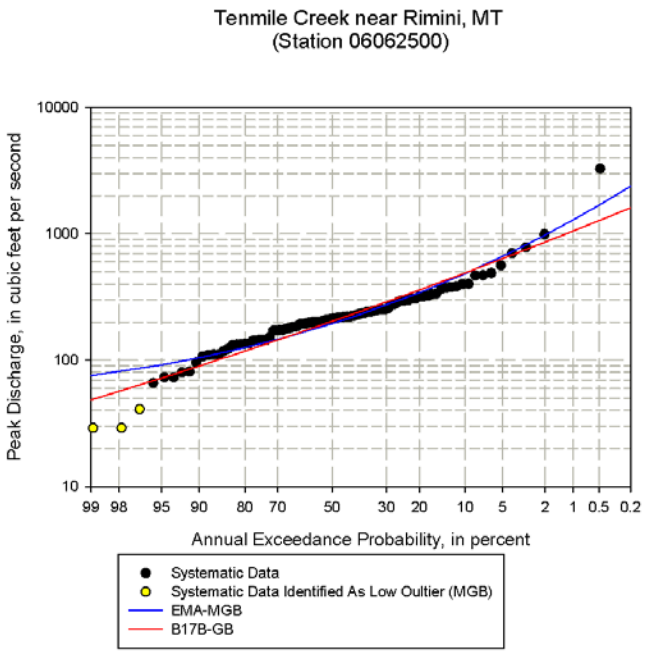


Figure 92: Site 06897000 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 6 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

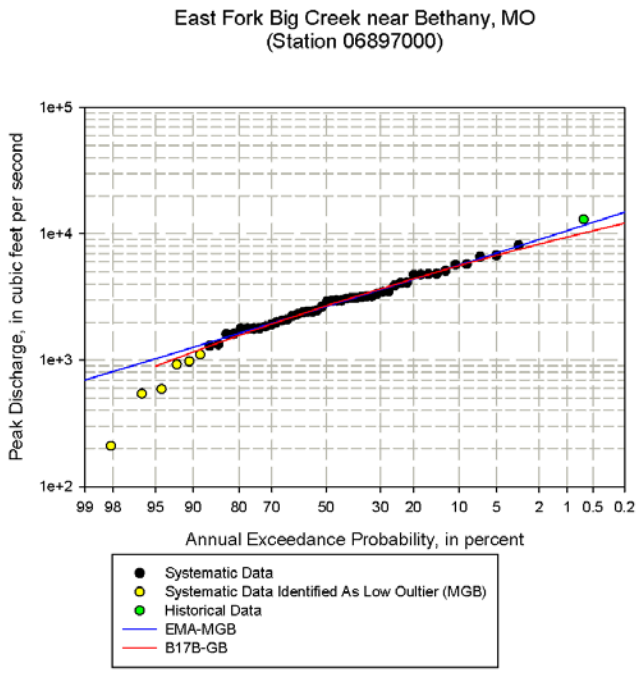


Figure 93: Site 06933500 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 20 Low Outliers; 1 Peak Identified as Below Gage Base Discharge

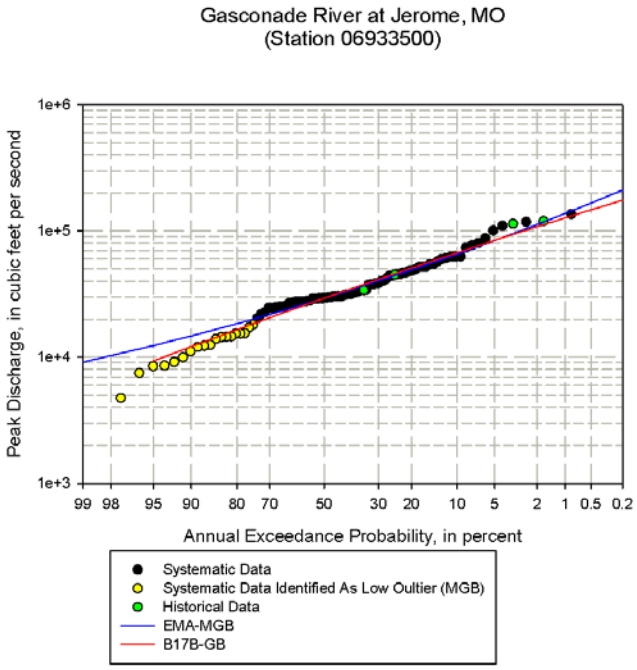


Figure 94: Site 07138600 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 12 Low Outliers; 4 Peaks Identified as Below Gage Base Discharge

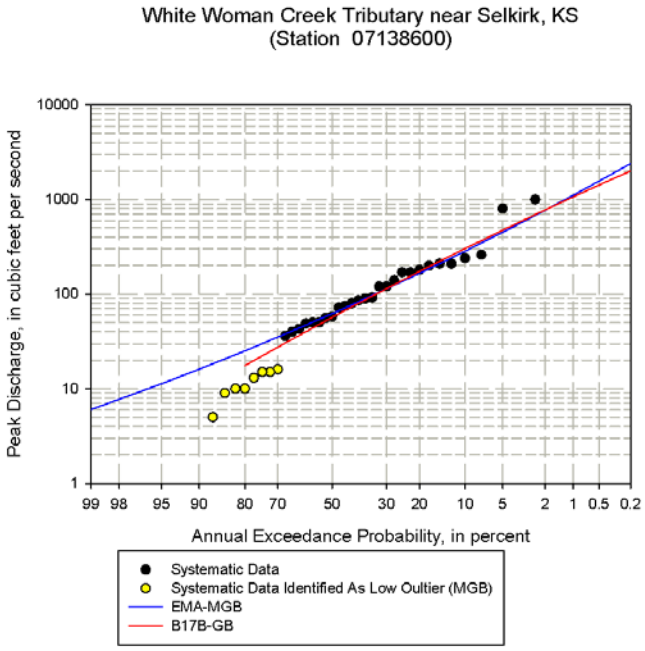


Figure 95: Site 08164000 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 1 Low Outlier; Grubbs-Beck Test Identified 0 Low Outliers

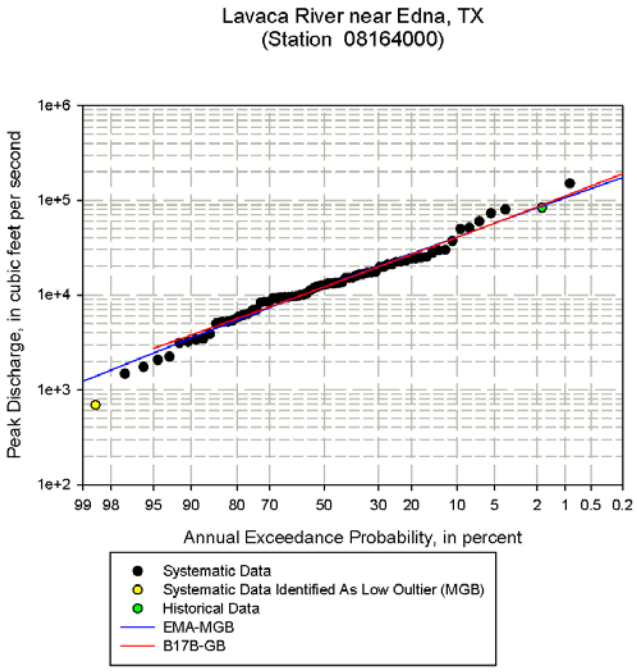


Figure 96: Site 08167000 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 23 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

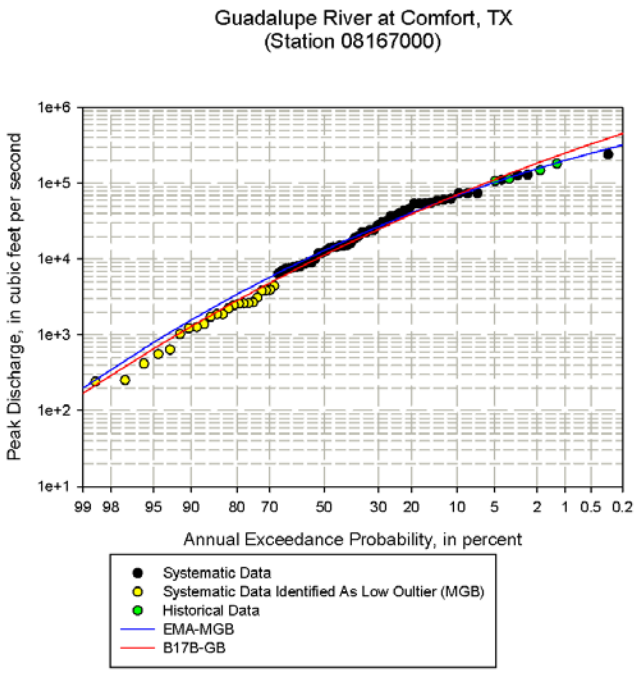


Figure 97: Site 08171000 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 28 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

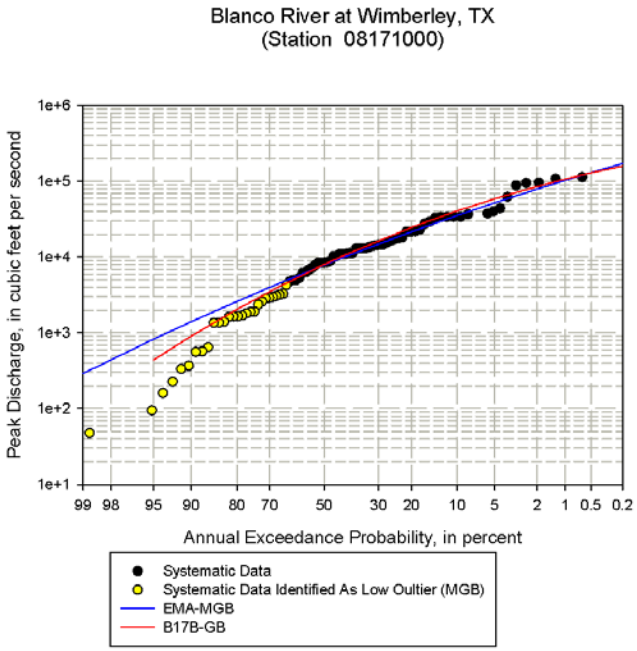


Figure 98: Site 09361500 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 2 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

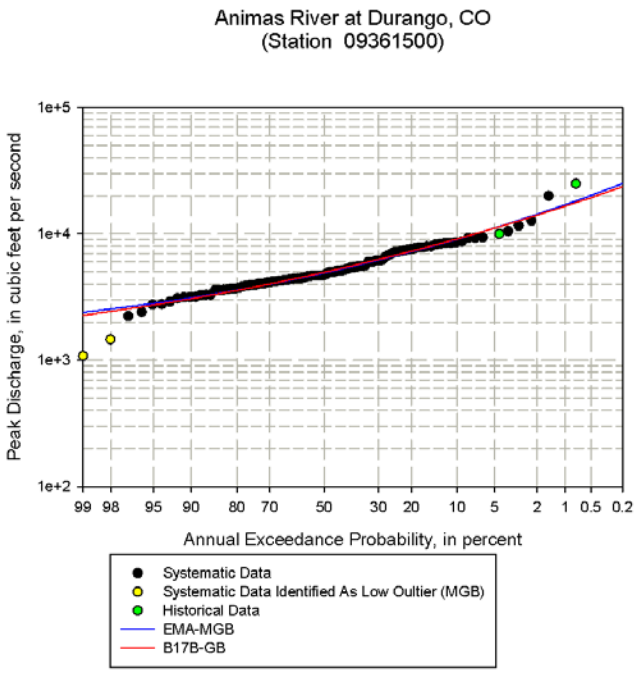


Figure 99: Site 09471000 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 3 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

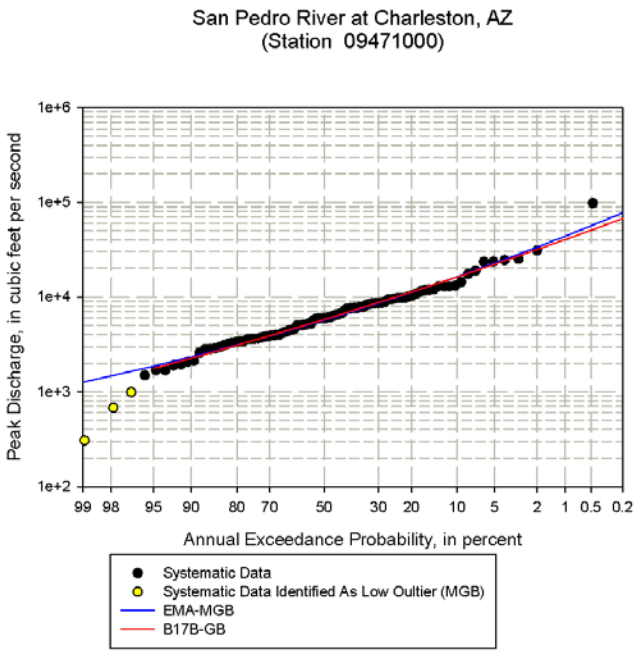


Figure 100: Site 09480000 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 8 Low Outliers; Grubbs-Beck Test Identified 2 Low Outliers

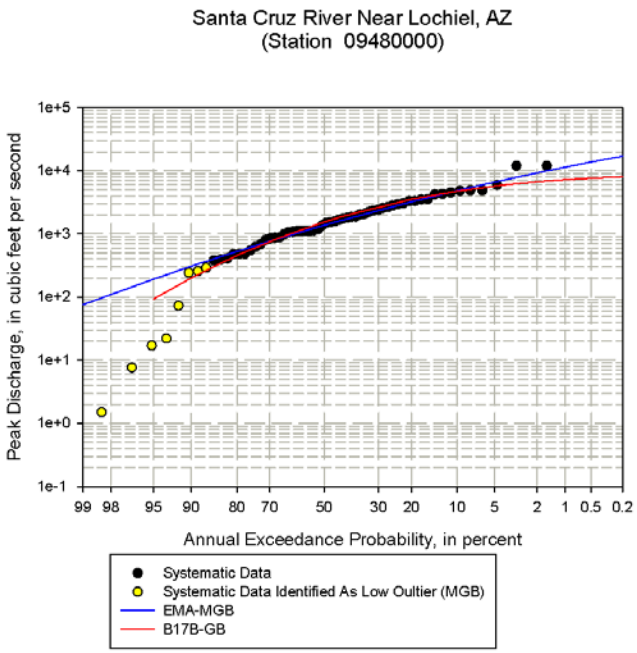


Figure 101: Site 11176000 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 28 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

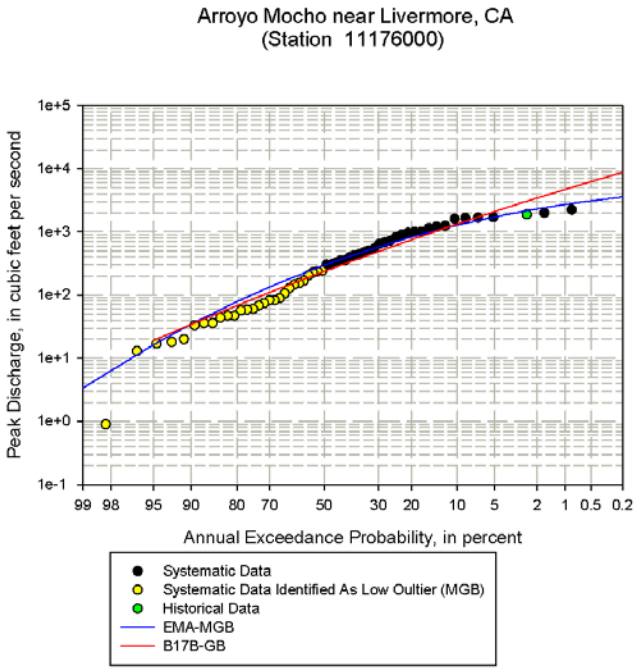


Figure 102: Site 11464500 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 18 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

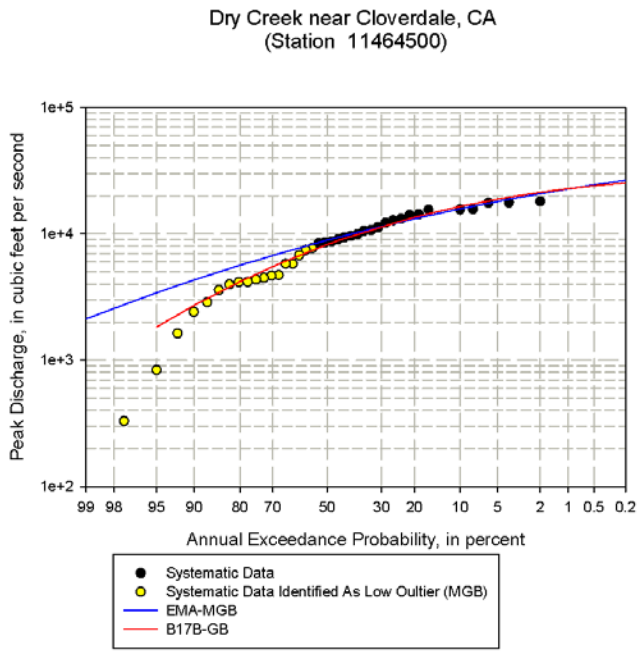


Figure 103: Site 11522500 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 1 Low Outlier; Grubbs-Beck Test Identified 1 Low Outlier

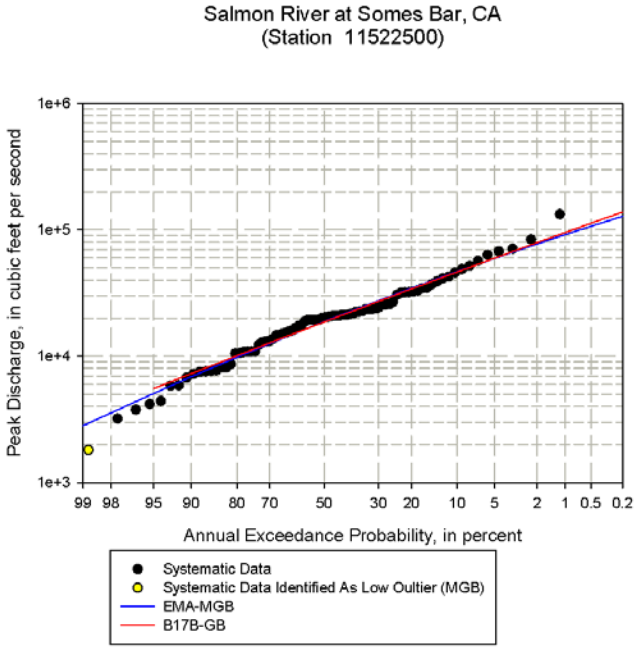


Figure 104: Site 12039500 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 48 Low Outliers; Grubbs-Beck Test Identified 0 Low Outliers

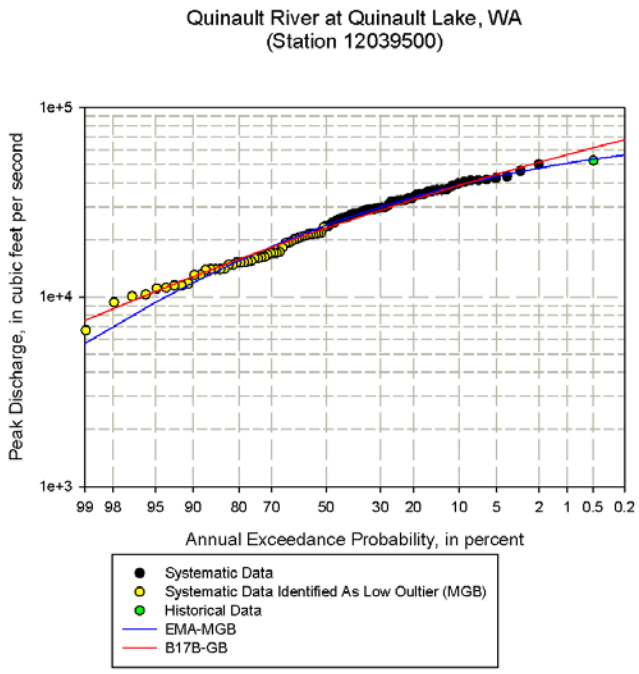


Figure 105: Site 13185000 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 36 Low Outliers; Grubbs-Beck Test Identified 1 Low Outlier

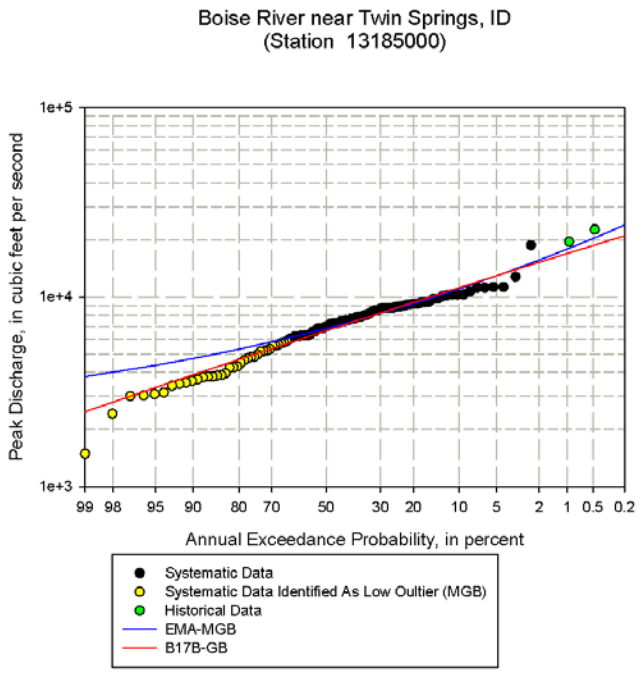


Figure 106: Site 13343660 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 3 Low Outliers; 1 Peak Identified as Below Gage Base Discharge; Grubbs-Beck Test Identified 1 Low Outlier

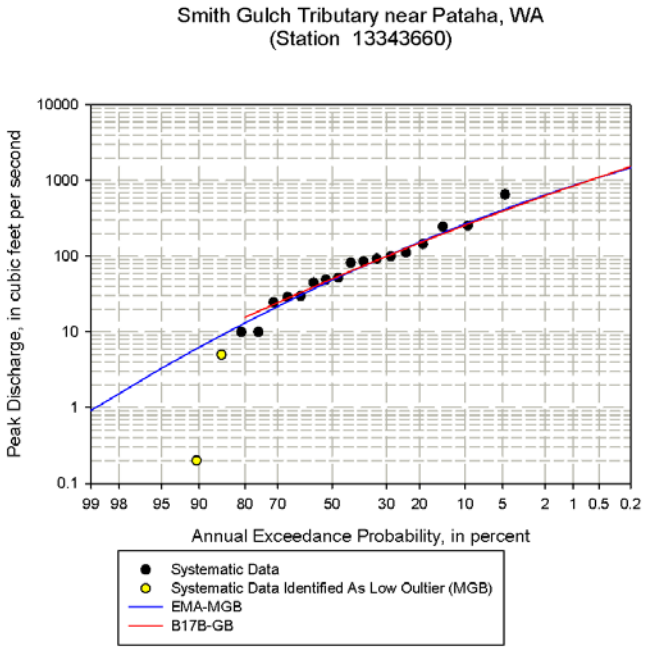
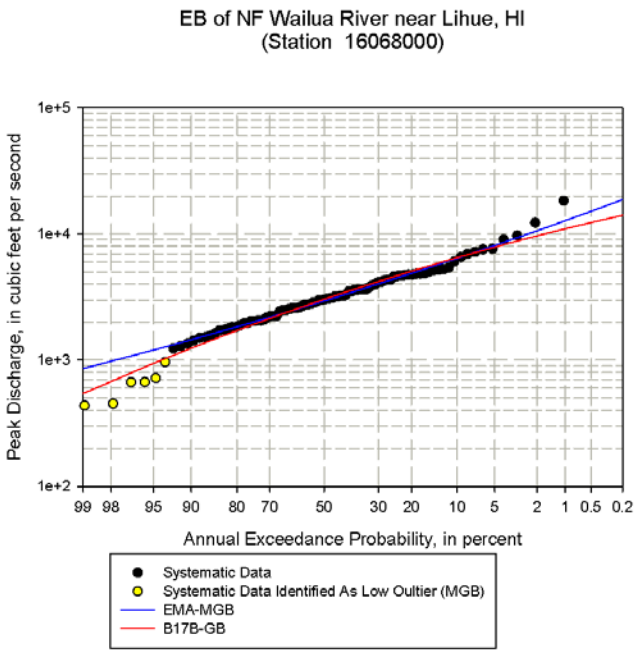


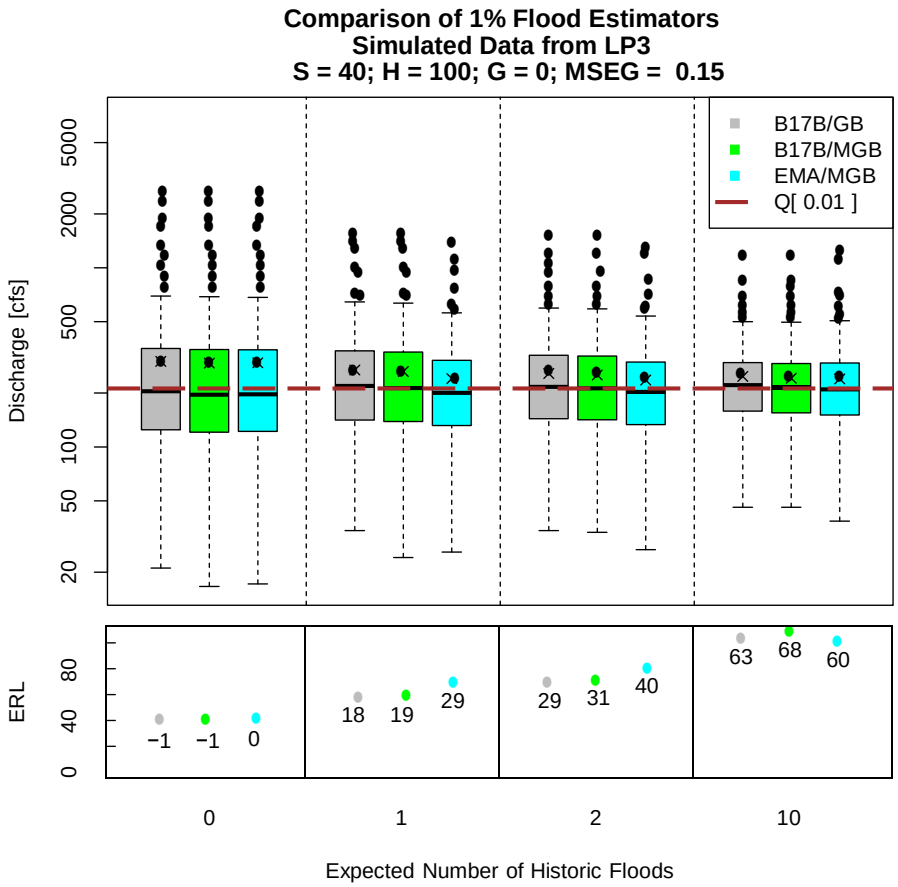
Figure 107: Site 16068000 with a Combination of Systematic and Historical Data; Multiple Grubbs-Beck Test Identified 6 Low Outliers



B.5 Studies with LP3 Distribution and Regional Skew

Figures 108 - 110 show the same cases as Figures 14 - 16 except that regional skew information has been added with a MSE of 0.15 – a typical value consistent with Bayesian/GLS skew maps. As expected, all of the estimators perform better with regional information. Aside from that, however, there is little difference between the corresponding figures.

Figure 108: Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = 0.0$ Regional skew is assumed to be 0.0 with $MSE = 0.15$.



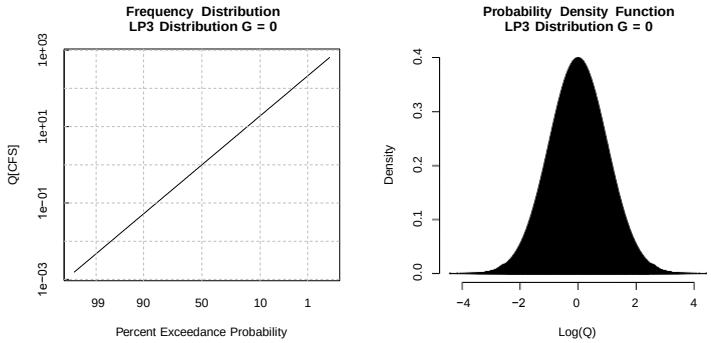
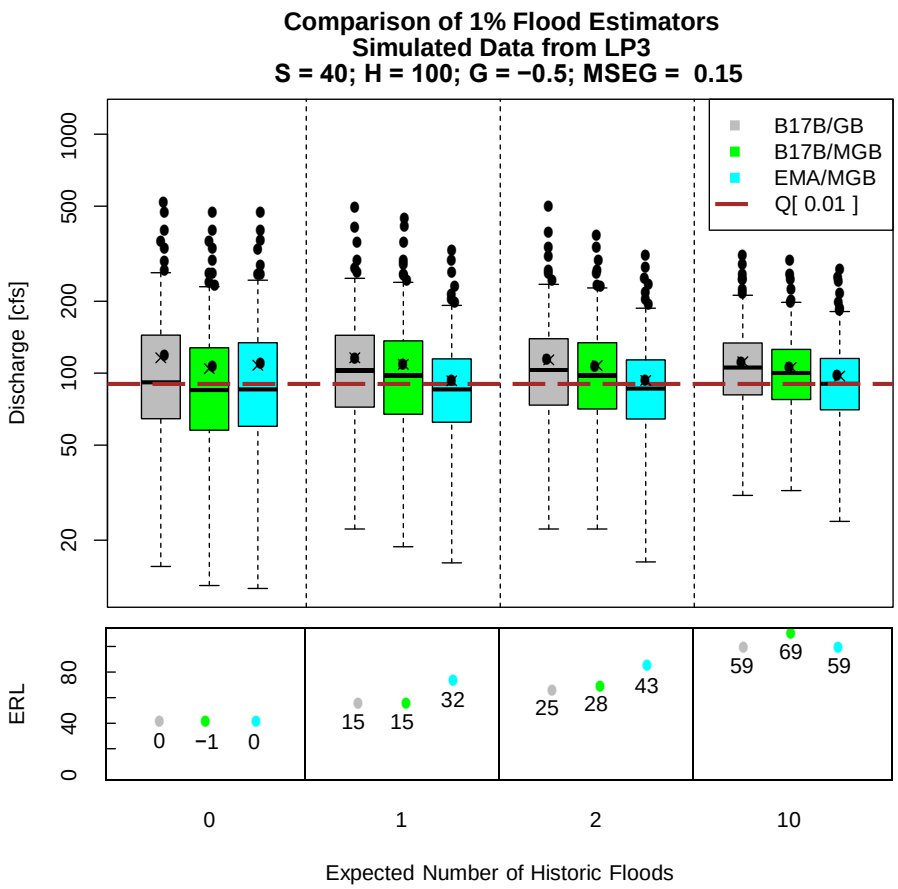


Figure 109: Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = -0.5$. Regional skew is assumed to be -0.5 with $MSE = 0.15$.



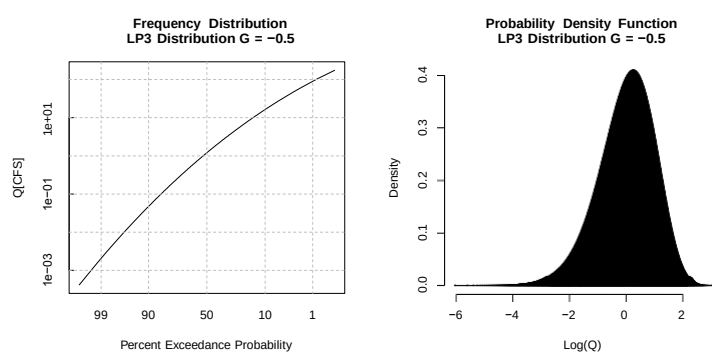
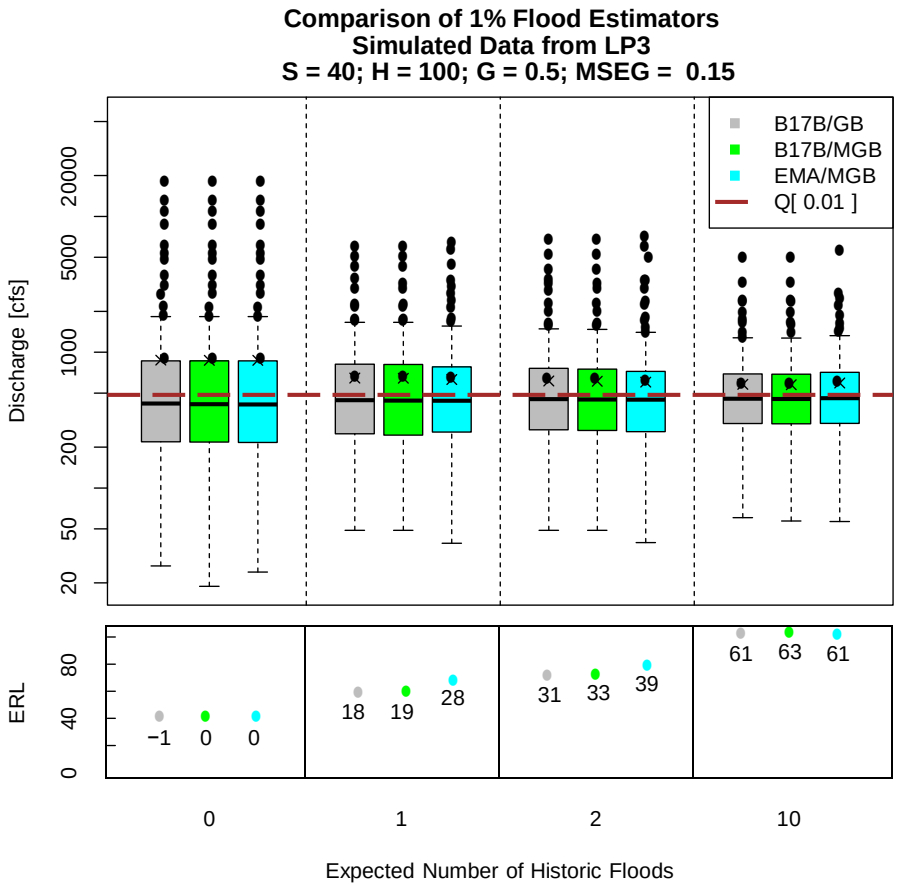
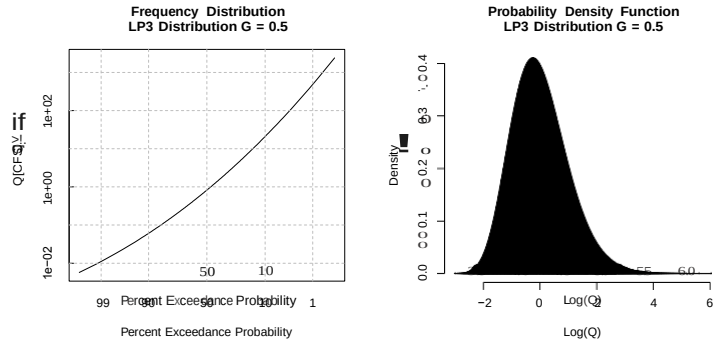


Figure 110: Results are based on 1000 replicate samples drawn from a Log-Pearson Type 3 distribution with skew $\gamma = 0.05$. Regional skew is assumed to be 0.5 with $MSE = 0.15$.





B.6 Additional Studies with Specific Frequency Curves

Figures 111 - 20 apply the three estimators, EMA/MGB, B17B/GB and B17B/MGB, to data drawn from specific populations selected to test the estimators' performance. The Cases are labeled "robustness test curve 1-6" in recognition of their origins. Figures 17 - 20 are presented in the main body of this report. Figures 111 and 112 are presented here because they do not actually deal with robustness but rather specific LP3 populations. In fact, because the estimators are all invariant with respect to location and scale, these two cases duplicate cases already considered in the report. Figure 111 depicts essentially the same case as figure 16 with a population skew of $G = 0.5$. Figure 112 depicts essentially the same case as figure 15 with a population skew of $G = -0.5$.

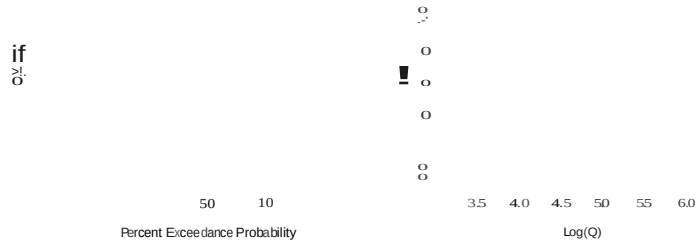
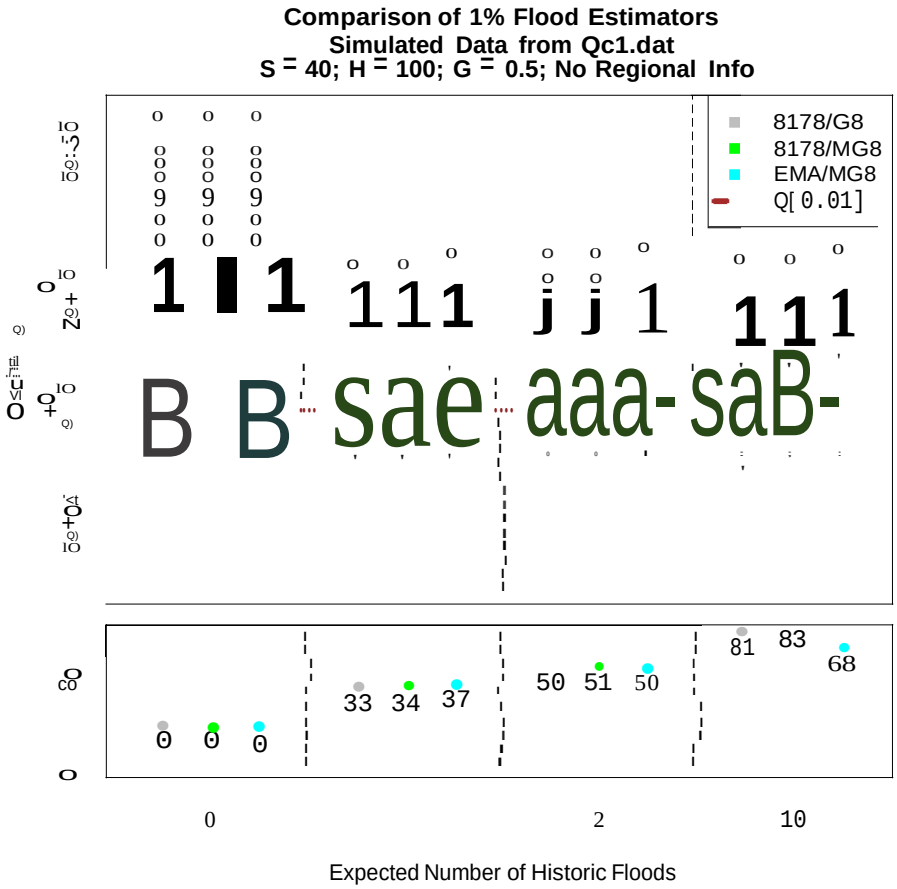
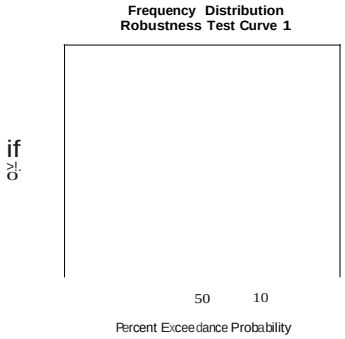


Figure 111: Results are based on 1000 replicate samples drawn from robustness test curve 1





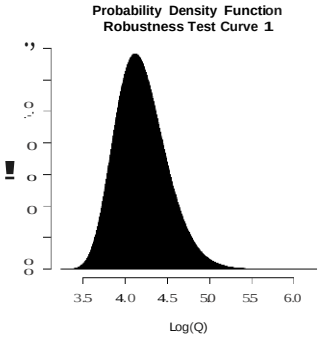
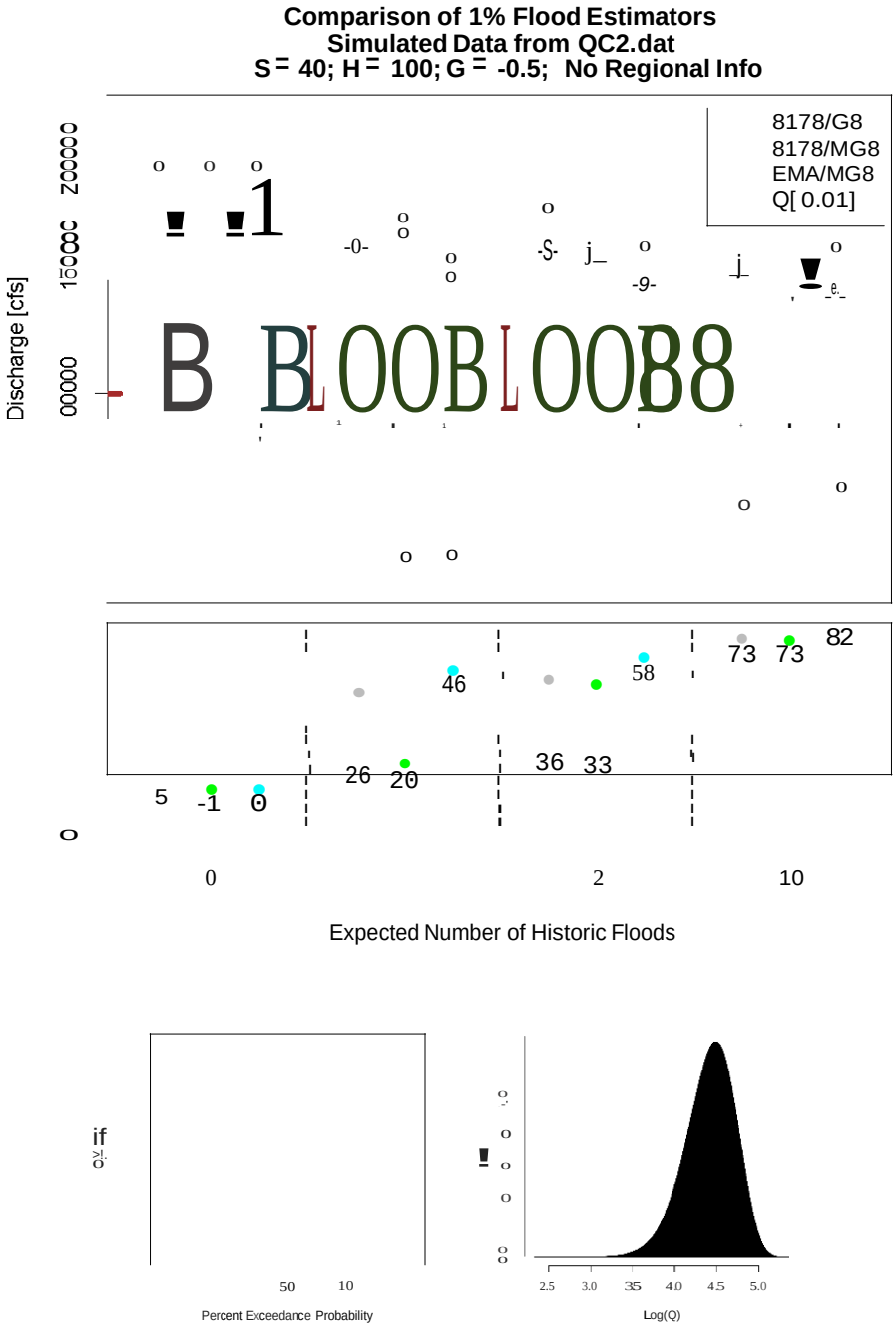
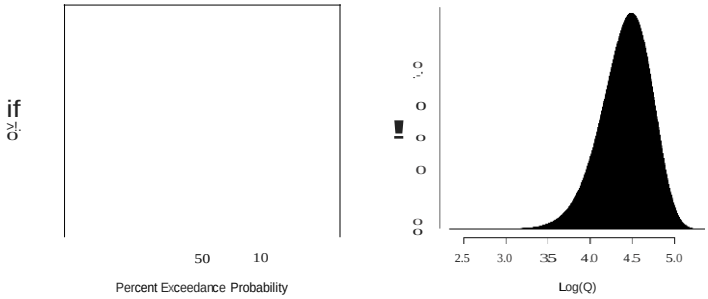


Figure 112: Results are based on 1000 replicate samples drawn from robustness test curve 2



Frequency Distribution
Robustness Test Curve 2

Probability Density Function
Robustness Test Curve 2



C Bulletin 17B Table of Contents, Status, and Possible Actions

Table 3: Assumptions Underlying Bulletin 17B, their current condition, and what to do with them

Where	Issue	Condition	Proposal
IV.A	Climatic Trends	OK	Update
IV.B	Randomness of Events	OK	Update
IV.C	Watershed Changes	OK	Keep
IV.D	Mixed Populations	OK	Keep
IV.E	Reliability of Flow Estimates	OK	Update

Table 4: Features of Bulletin 17B, their current condition, and what to do with them

Where	Statistical Procedure	Condition	Proposal
V.B-1	The Distribution (LP-III)	OK	Keep
V.B-2	Fitting the Distribution (M-o-M)	OK	Generalize
V.B-3	Estimating Generalized Skew	Narrow	Replace
V.B-4	Weighting Generalized Skew	OK	Generalize
V.B-5	Broken Record	OK	Generalize
V.B-6	Incomplete Record/CSGs	OK	Generalize
V.B-7	Zero Flood Years	OK	Generalize
V.B-8	Mixed Populations	OK	Keep
V.B-9	Grubbs-Beck test for low outliers	Limited	Generalize
V.B-10	Historic Flood Data	Limited	Replace
VI.A	Confidence Intervals	Incorrect	Replace
VI.B	Risk	OK	Keep
VI.C	Expected Probability	Incorrect	Delete
VII.A	Non-Conforming Special Situations	OK	Amend
VII.B	Plotting Positions	Limited	Generalize
VII.C	Future Studies	OK	Update
A-3	Table of K-values	Unneeded	Delete
A-4	Table of Outlier K-values	Unneeded	Delete
A-5	Conditional Probability Adjustment	Limited	Delete
A-6	Historic Weighted Moments	Limited	Delete
A-7	Two-station comparison	OK	Keep
A-8	Weighted Independent Estimates	OK	Keep
A-9	Confidence Limits	Incorrect	Replace
A-10	Risk	Awkward	Keep (?)
A-12	Trend Tests	–	New
A-11	Multiple-Threshold Plotting Pos	–	New
A-13	EMA	–	New

References

ACWI. Advisory committee on water information, subcommittee on hydrology, hydrologic frequency analysis work group, 2011. URL http://acwi.gov/hydrology/Frequency/.

AIR. A chronology of events affecting the National Flood Insurance Program. Report, American Institutes for Research, 2001.

T. A. Cohn, W. M. Lane, and W. G. Baier. An algorithm for computing moments-based flood quantile estimates when historical flood information is available. *Water Resources Research*, 33(9):2089–2096, 1997.

T. A. Cohn, J. F. England, N. Steinberger, R. R. Mason, and J. R. Stedinger. A generalized grubbs-beck test for detecting multiple potentially influential low outliers in flood series. *(to be submitted to Water Resources Research)*, 2012.

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

V.W. Griffis and J.R. Stedinger. Evolution of Flood Frequency Analysis with Bulletin 17. *Journal of Hydrologic Engineering*, 12:283, 2007.

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

ICOWR. Methods of flow frequency analysis, Bulletin No. 13. Technical report, Interagency Committee on Water Resources, Subcommittee on Hydrology, Washington, D.C., 1966.

W. Kirby. Algebraic boundedness of sample statistics. *Water Resources Research*, 10(2):220–222, 1974.

D.S. Mileti. *Disasters by design*. Joseph Henry Press Washington, DC, 1999.

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

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

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

J. R. Stedinger, Richard M. Vogel, and Efi Foufoula-Georgiou. *Frequency Analysis of Extreme Events*, chapter 18, page 99. McGraw Hill, Inc., 1993.

G. D. Tasker and W. O. Thomas. Flood frequency analysis with pre-record information. *ASCE Journal of the Hydraulics Division (HY2)*, 104(2):249–259, 1978.

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

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

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

USWRC. *Guidelines for Determining Flood Flow Frequency, Bulletin No. 17A*. U.S. Water Resources Council, Subcommittee on Hydrology, Washington, D.C., 1977.