Flood-Frequency Analysis Updates for Bulletin 17B and the U.S. Geological Survey PeakFQ Program

Presented at the National Hydraulic Engineering Conference

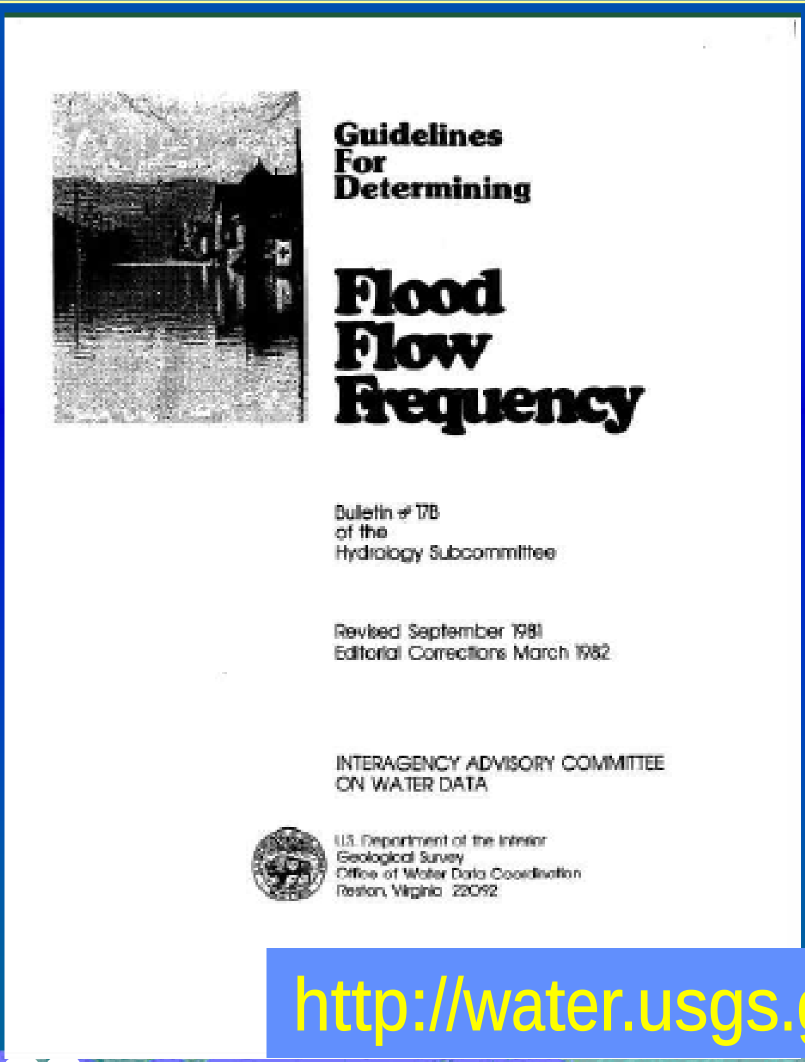
Iowa City, IA

August 22, 2014

by David Eash and Andrea Veilleux

U.S. Geological Survey
Iowa Water Science Center
Iowa City, IA 52240
319-358-3615
daeash@usgs.gov

Flood-frequency estimation at gaged sites follows guidelines in Bulletin 17B



Guidelines For Determining

Flood Flow Frequency

Bulletin # 17B
of the
Hydrology Subcommittee

Revised September 1981
Editorial Corrections March 1982

INTERAGENCY ADVISORY COMMITTEE
ON WATER DATA

U.S. Department of the Interior
Geological Survey
Office of Water Data Coordination
Reston, Virginia 22092

History:

- 1967 Bulletin 15
- 1976 Bulletin 17
- 1977 Bulletin 17A
- 1981 Bulletin 17B
- 201? Bulletin 17C

http://water.usgs.gov/osw/bulletin17b/dl_flow.pdf

USGS

Bulletin 17B [1982, p. ii]

This present revision is adopted with the knowledge and understanding that review of these procedures will continue. When warranted by experience and by examination and testing of new techniques, other revisions will be published.

- Hydrologic Frequency Analysis Work Group (HFAWG) is tasked with examining and testing flood frequency methodology and providing guidance on revisions to Bulletin 17B

http://acwi.gov/hydrology/Frequency/index.html

USGS

June 2013 HFAWG studies recommend changes in four main areas

1) Historical information and weighted-moments approach
2) Low-outlier detection and treatment
3) Procedures for estimating generalized/regional skew
4) Procedures for estimating confidence limits

http://acwi.gov/hydrology/Frequency/index.html

Key recommended revision to B17B

- Adoption of the expected moments algorithm (EMA) framework for the analysis of data sets containing zeros, outliers, interval flow estimates, multiple thresholds, or historical and/or paleoflood information as the appropriate generalization of the method-of-moments to address such situations

http://acwi.gov/hydrology/Frequency/index.html

EMA advantages

- Extension of B17B LP3 method-of-moments approach, which includes a consistent statistical framework for ALL sources of information available
  - For simple cases with only a systematic record and a regional skew: EMA estimates = B17B estimates
- EMA deals with interval and multiple threshold data that conditional probability adjustment and historical weighted moments in B17B do not.
  - CSG and historic info are best described by intervals
- Provides confidence intervals that include skew uncertainty and reflect interval observations

Other recommended revisions to B17B

- Update with the new multiple Grubbs-Beck (MGB) test for the identification of Potentially Influential Low Flows (PILFs)
- Use of Bayesian Generalized Least Squares (BGLS) procedures for the derivation of regional skews for weighting with the at-site skews
- Correct uncertainty computations and confidence intervals
- Use of a multiple-threshold plotting position formula – compatible with EMA

HFAWG testing to show new methods work

- Theoretical arguments
  - Analytical results
  - Monte Carlo results
- Evaluation using real data
  - 82 test sites
  - Complicated situations
  - Judgment
- Recommendations and testing report published on Advisory Committee on Water Information website:
  - http://acwi.gov/hydrology/Frequency/index.html

Bulletin 17B representation of peaks

- SYSTEMATIC PEAKS–
  - Recorded during one or more periods of regular data collection;
  - Can be from the continuous record or from a crest-stage gage
- HISTORIC PEAKS-
  - Records of floods that occurred outside the period of regular streamgaging
    - Discharge measurement during a flood event
    - Indirect measurement after a flood event
- Peaks represented as a point or single value

EMA representation of peaks

- Peaks are represented as intervals
- HISTORICAL – (systematic + historic peaks)
  - For every year Y, it is assumed that there was a peak discharge Q_Y, regardless of whether the discharge was recorded
- FLOW INTERVALS – $(Q_{Y,lower}, Q_{Y,upper})$
  - Describes knowledge of peak flow Q_Y for every Y
  - When peak is known with confidence, $(Q_{Y,lower}, Q_{Y,upper}) = (Q_Y, Q_Y)$
  - Censored peaks:
    - Flow greater than some value Q_Y (Code G + X):
    - $(Q_{Y,lower}, Q_{Y,upper}) = (Q_Y, inf)$
    - Flow less than some value Q_Y (Code L): $(Q_{Y,lower}, Q_{Y,upper}) = (0,Q_Y)$
  - Interval peak:
    - Flow greater than some value and less than another value

EMA representation of peaks (cont.)

- **PERCEPTION THRESHOLDS – ($T_{Y,lower}$, $T_{Y,upper}$)**
  - **Reflect range of flows that would have be measured/recorded had they occurred**
  - **Independent of actual peak discharges that have occurred**
  - **$T_{Y,lower}$ = smallest peak that would result in a recorded flow**
  - **$T_{Y,upper}$ = largest peak that would result in a documented flow**
  - **For periods of continuous, full-range streamgage record:**

    ($T_{Y,lower}$, $T_{Y,upper}$) = (0, *inf*) where $T_{Y,lower}$ = 0 is the gage-base flow
  - **Can adjust $T_{Y,lower}$ to accommodate a changing gage-base flow**

USGS

EMA representation of data: continuous-record streamgage

USGS 2011 Streamgaging Calendar, March, John A. Mazurek, April 20, 2009 http://water.usgs.gov/osw/calendar.html

- From Peak Flow File:

Date	Peak	Code
1993-03-28	276	

- Flow Interval:

$(Q_{1993,lower}, Q_{1993,upper}) = (276,276)$

- Perception Threshold:

$(T_{1993,lower}, T_{1993,upper}) = (0, inf)$

USGS

EMA representation of data: crest-stage gage

Photo courtesy of Paul Rydlund.
http://mo.water.usgs.gov/surfwat/CSGWeb/index.htm

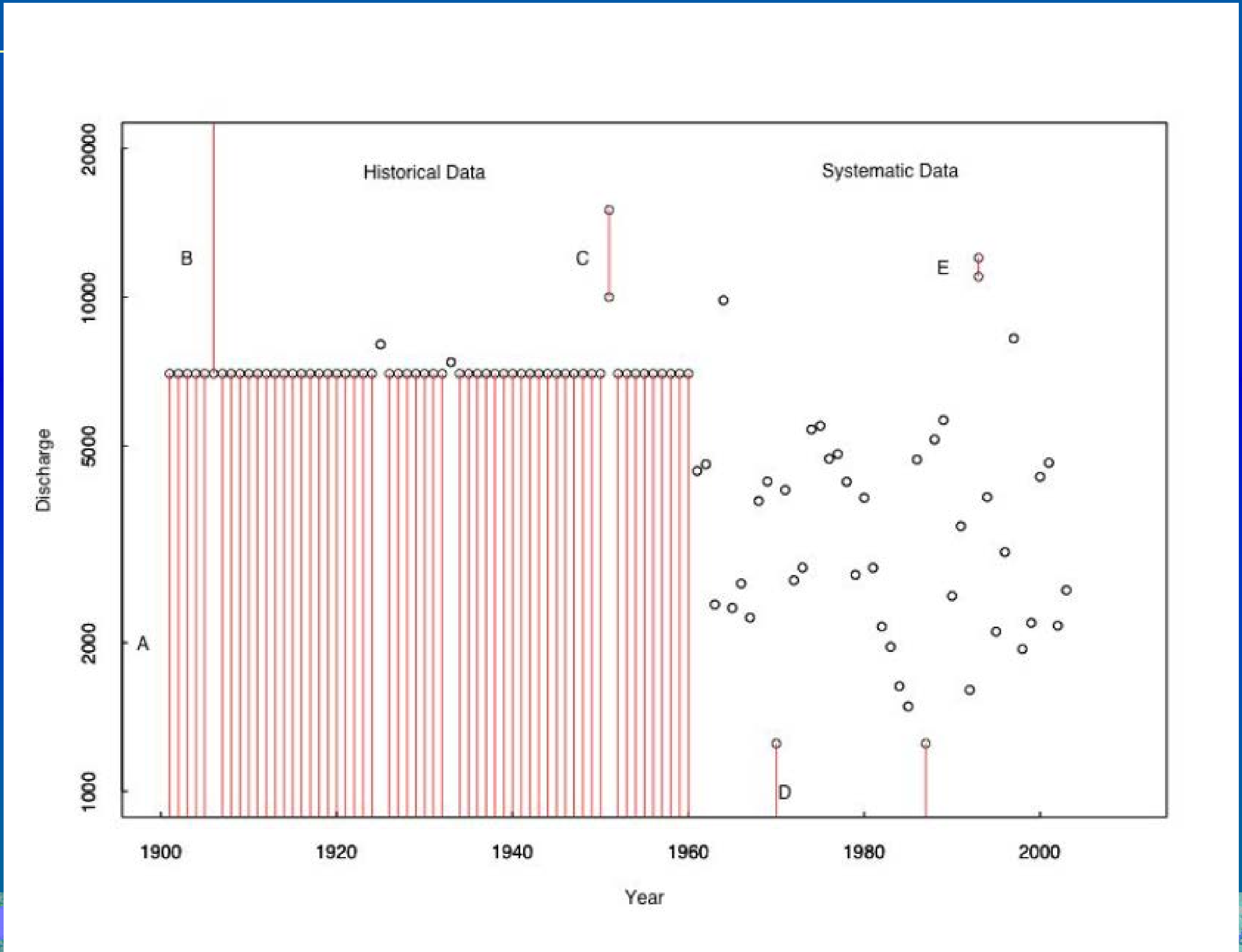
- Gage base in 1980: 20 cfs
  - CSG cannot record flow < 20 cfs
- From Peak Flow File (from USGS NWIS):

Date	Peak	Code
1980-10-20	20	4*

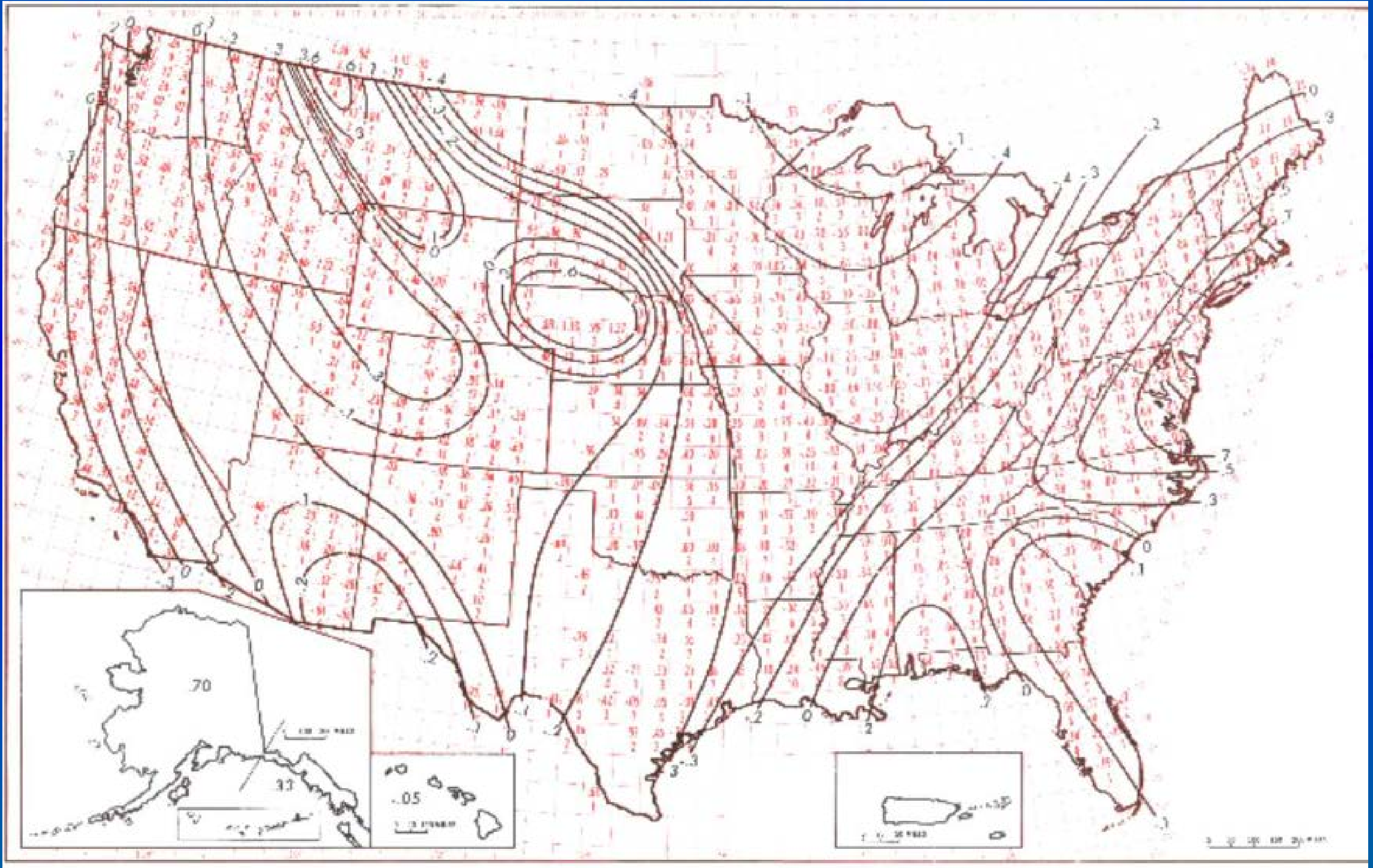
**Code 4: Discharge is less than indicated value, which is a minimum recordable discharge at this site*

- Flow Interval:
  $(Q_{1980,lower}, Q_{1980,upper}) = (0, 20)$
- Perception Threshold:
  $(T_{1980,lower}, T_{1980,upper}) = (20, \mathit{inf})$

Types of interval data



Estimating regional skew (1982 B17B)



GENERALIZED SKEW COEFFICIENTS OF LOGARITHMS OF ANNUAL MAXIMUM STREAMFLOW

AVERAGE SKEW COEFFICIENT BY ONE DEGREE QUADRANGLES

Lower number in each quadrangle is number of stream gaging stations for which the average shown above it was computed

2010 Iowa regional skew study Bayesian WLS/GLS regression

http://pubs.usgs.gov/sir/2013/5086/

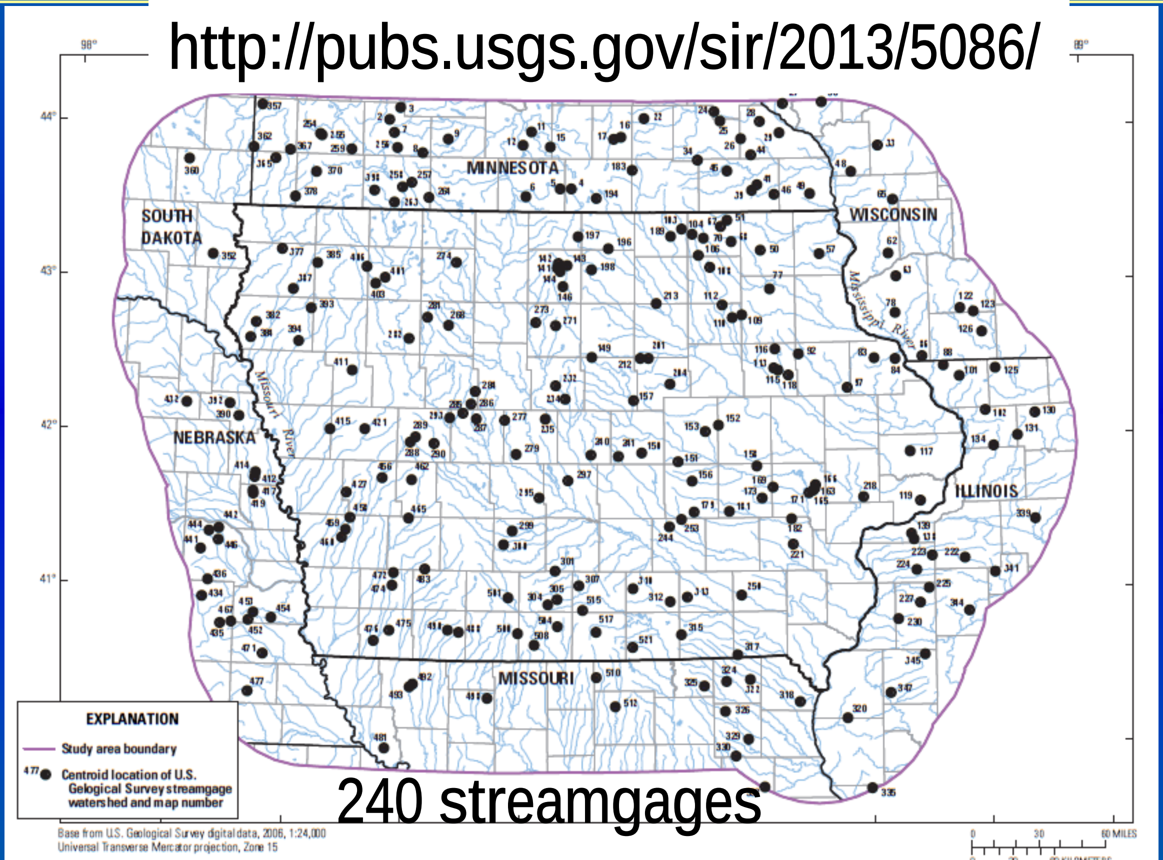


Figure 6. Location of basin centroids for 240 streamgages used for regional skew analysis for Iowa.

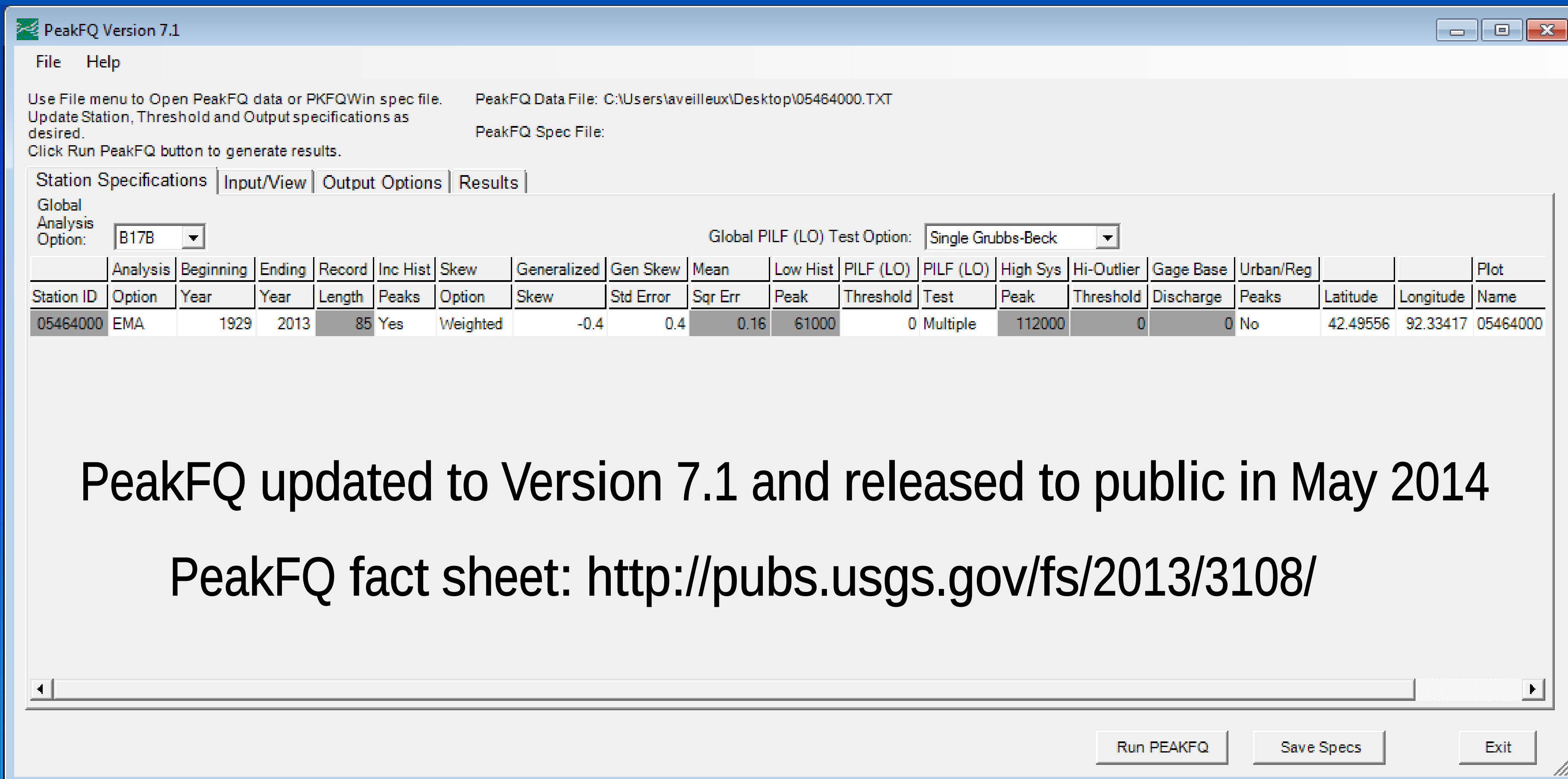
Constant Regional Skew Value for Iowa = -0.400

1982 (MSE = 0.302) → ERL = 17 yrs

2010 (MSE = 0.160) → ERL = 50 yrs

USGS

USGS PeakFQ flood-frequency analysis program (PKFQWin)



PeakFQ updated to Version 7.1 and released to public in May 2014

PeakFQ fact sheet: http://pubs.usgs.gov/fs/2013/3108/

EMA perception thresholds
(missing years 1930-32, 1934-40 in historic period)

PeakFQ Version 7.1

File Help

Use File menu to Open PeakFQ data or PKFQWin spec file.
Update Station, Threshold and Output specifications as desired.
Click Run PeakFQ button to generate results.

PeakFQ Data File: C:\Users\aveilleux\Desktop\05464000.TXT

PeakFQ Spec File:

Station Specifications | Input/View | Output Options | Results

Station: 05464000

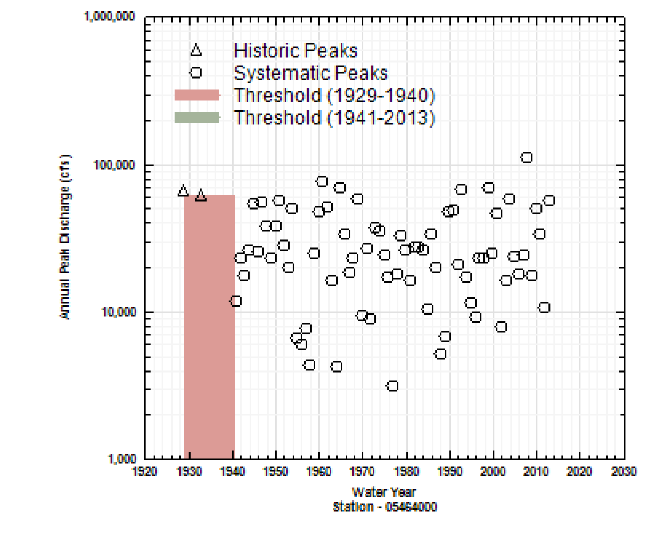
Save Input Peaks Graph as BMP - Bitm

Perception Thresholds

Start Year	End Year	Low Threshold	High Threshold	Comment (Required)
1929	2013	0	inf	Default
1929	1940	61000	inf	Historic Period
1941	2013	0	inf	Regular gaging peri+
			inf	

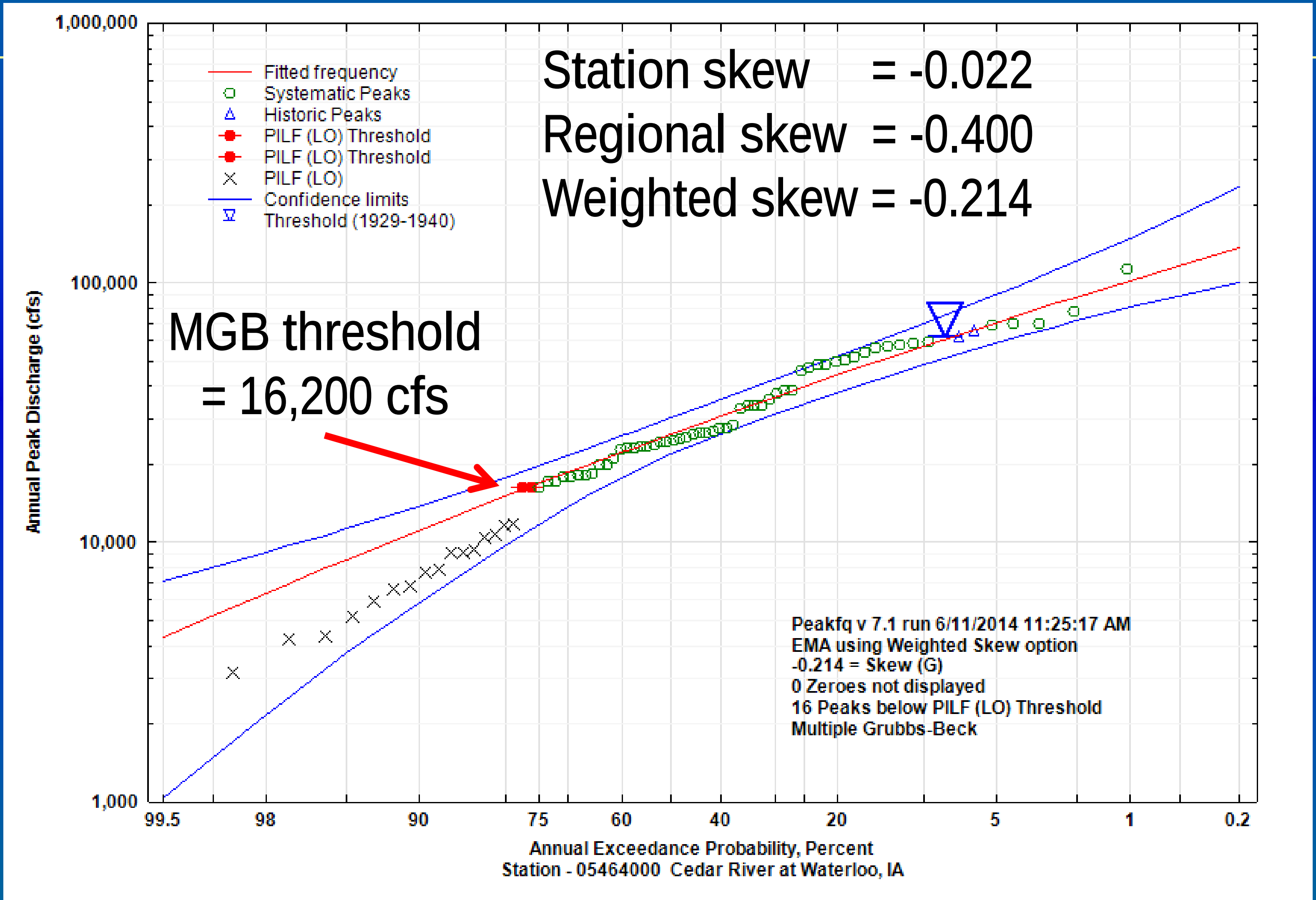
Data

Year	Peak	Remark Codes	Low Value	High Value	Comment (Required)
1929	65000	H	65000	65000	
1933	61000	H	61000	61000	
1941	11700		11700	11700	
1942	23000		23000	23000	
1943	17700		17700	17700	
1944	26400		26400	26400	
1945	53300		53300	53300	
1946	25200		25200	25200	
1947	55600		55600	55600	



Code Lookup | Run PEAKFQ | Save Specs | Exit

PeakFQ output flood-frequency curve



Effects of including or censoring low outliers

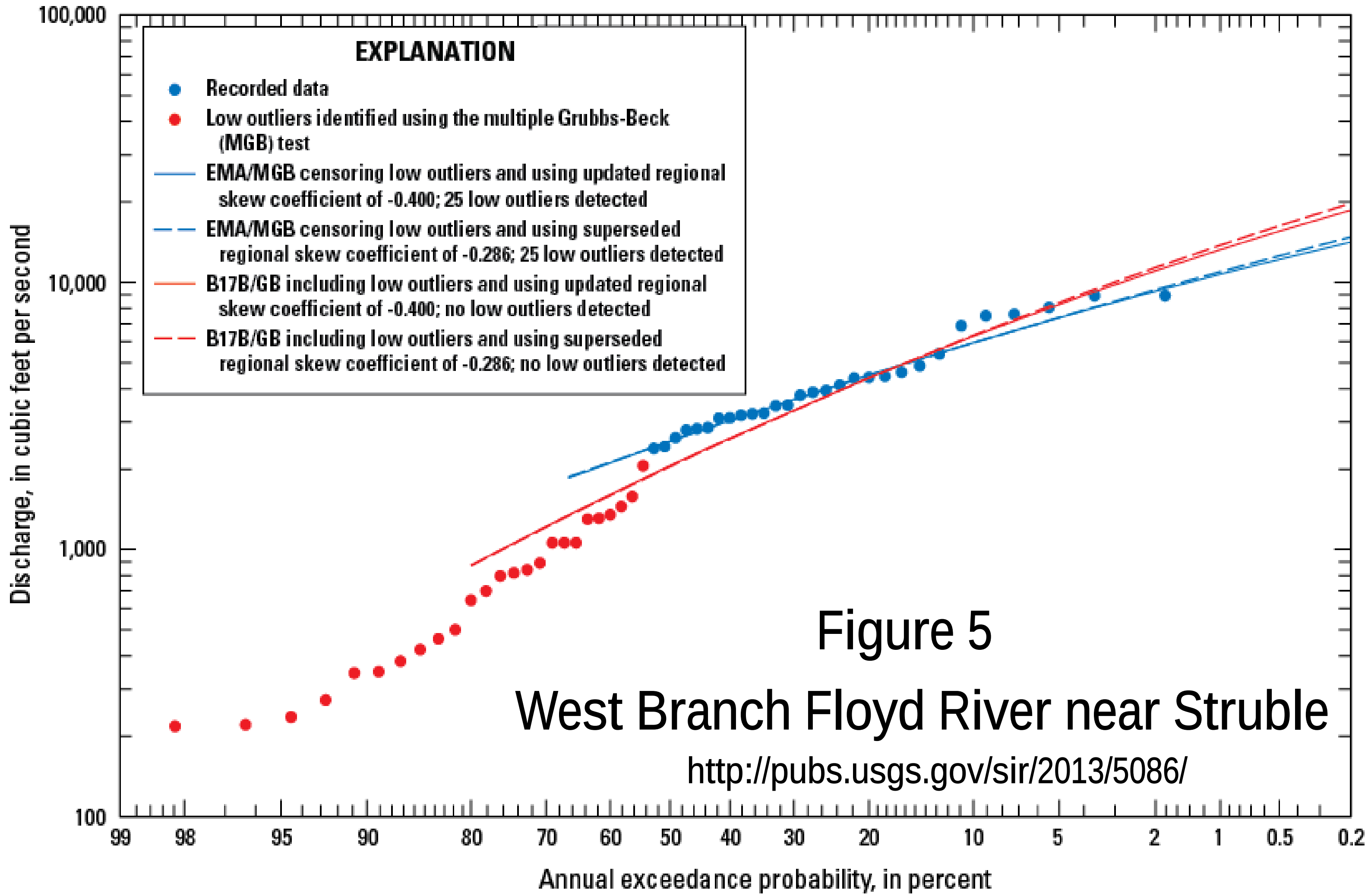


Figure 5

West Branch Floyd River near Struble

http://pubs.usgs.gov/sir/2013/5086/

USGS

PeakFQ output flood-frequency analysis results

```
Program PeakFq          U. S. GEOLOGICAL SURVEY             Seq.001.001
Version 7.1        Annual peak flow frequency analysis     Run Date / Time
3/14/2014                                                  06/11/2014 11:19

              Station - 05464000  Cedar River at Waterloo, IA


                   I N P U T   D A T A   S U M M A R Y

              Number of peaks in record              =       75
              Peaks not used in analysis             =        0
              Systematic peaks in analysis           =       73
              Historic peaks in analysis             =        2
              Beginning Year                         =     1929
              Ending Year                            =     2013
              Historical Period Length               =       85
              Generalized skew                       =   -0.400
                   Standard error                    =    0.400
                   Mean Square error                 =    0.160
              Skew option                            =   WEIGHTED
              Gage base discharge                    =      0.0
              User supplied high outlier threshold   =   --
              User supplied PILF (LO) criterion      =   --
              Plotting position parameter            =     0.00
              Type of analysis                                EMA
              PILF (LO) Test Method                          MGBT
              Perception Thresholds:
                   Begin      End        Low      High      Comment
                    1929     2013        0.0       INF       DEFAULT

                    1929     1940    61000.0       INF       HISTORIC PERIOD

                    1941     2013        0.0       INF       REGULAR GAGING PERIOD: SYSTEMATIC RECORD

              Interval Data                      =     None Specified


*********  NOTICE  --  Preliminary machine computations.        *********
*********  User responsible for assessment and interpretation.  *********

  EMA003I-PILFS (LOS) WERE DETECTED USING MULTIPLE GRUBBS-BECK TEST  16   16200.0
    THE FOLLOWING PEAKS (WITH CORRESPONDING P-VALUES) WERE CENSORED:
        3120.0     (0.2774)
        4210.0     (0.2252)
        4320.0     (0.0558)
```

MGB Threshold: 16,200 cfs

PeakFQ output flood-frequency analysis results

```
                                   Kendall's Tau Parameters

                                                             MEDIAN     No. of
                                       TAU       P-VALUE      SLOPE      PEAKS
                              ------------------------------------------------
             SYSTEMATIC RECORD      0.008       0.924        4.522      73

1

  Program PeakFq             U. S. GEOLOGICAL SURVEY                  Seq.001.002
  Version 7.1          Annual peak flow frequency analysis            Run Date / Time
  3/14/2014                                                           06/11/2014 11:19

                Station - 05464000  Cedar River at Waterloo, IA

           ANNUAL FREQUENCY CURVE PARAMETERS -- LOG-PEARSON TYPE III

                                  LOGARITHMIC
                         -------------------------------
                                    STANDARD
                            MEAN    DEVIATION      SKEW
                         -------------------------------
 EMA W/O REG. INFO         4.4074     0.2702     -0.022
 EMA W/REG. INFO           4.4037     0.2778     -0.214

 EMA ESTIMATE OF MSE OF SKEW W/O REG. INFO (AT-SITE)      0.0636
```

PeakFQ output flood-frequency analysis results (cont.)

ANNUAL FREQUENCY CURVE -- DISCHARGES AT SELECTED EXCEEDANCE PROBABILITIES

ANNUAL EXCEEDANCE PROBABILITY	EMA W/ REG INFO ESTIMATE	EMA W/O REG INFO ESTIMATE	<------ FOR EMA ESTIMATES -------> VARIANCE OF EST.	95% CONFIDENCE INTERVALS LOWER	UPPER
0.9950	4292.	5080.	0.0268	1023.0	7076.0
0.9900	5179.	5950.	0.0201	1468.0	7959.0
0.9500	8523.	9147.	0.0084	3722.0	11210.0
0.9000	11020.	11490.	0.0050	5769.0	13660.0
0.8000	14900.	15150.	0.0027	9647.0	17640.0
0.6667	19610.	19580.	0.0016	14970.0	22780.0
0.5000	25920.	25610.	0.0012	21730.0	30060.0
0.4292	29020.	28610.	0.0011	24730.0	33720.0
0.2000	43640.	43160.	0.0012	37410.0	51760.0
0.1000	56590.	56630.	0.0016	47910.0	69400.0
0.0400	73960.	75580.	0.0023	61160.0	96220.0
0.0200	87470.	91020.	0.0031	70700.0	120100.0
0.0100	101400.	107600.	0.0041	79800.0	147900.0
0.0050	115700.	125300.	0.0054	88460.0	180500.0
0.0020	135400.	150600.	0.0075	99260.0	232600.0

PeakFQ output flood-frequency analysis results (cont.)

EMPIRICAL FREQUENCY CURVES -- HIRSCH-STEDINGER PLOTTING POSITIONS

WATER YEAR	RANKED DISCHARGE	EMA ESTIMATE	INTERVALS LOW	INTERVALS HIGH
2008	112000.0	0.0103		
1961	76700.0	0.0206		
1965	69500.0	0.0309		
1999	69300.0	0.0412		
1993	68100.0	0.0515		
-1929	65000.0	0.0618		
-1933	61000.0	0.0721		
1969	58600.0	0.0956		
2004	58500.0	0.1089		
2013	57300.0	0.1222		
1951	56400.0	0.1355		
1947	55600.0	0.1488		
1945	53300.0	0.1621		
1962	51200.0	0.1754		
2010	49900.0	0.1887		
1954	49400.0	0.2020		
1991	48200.0	0.2153		
1960	48100.0	0.2286		
1990	47100.0	0.2419		
2001	45800.0	0.2552		
1950	38400.0	0.2685		
1948	38200.0	0.2818		
1973	37300.0	0.2951		
1974	35500.0	0.3084		
2011	33700.0	0.3217		
1966	33400.0	0.3483		
1986	33400.0	0.3350		
1979	32500.0	0.3616		
1952	28000.0	0.3749		
1982	27500.0	0.3882		
1983	27200.0	0.4015		
1971	26700.0	0.4148		
1944	26400.0	0.4281		
1984	26300.0	0.4414		
1980	25900.0	0.4547		
1946	25200.0	0.4680		
2000	24900.0	0.4813		
1959	24700.0	0.4946		
2007	24200.0	0.5079		
1975	24100.0	0.5212		
2005	23700.0	0.5345		
1997	23300.0	0.5611		
1998	23300.0	0.5478		
1949	23100.0	0.5744		
1942	23000.0	0.5877		
1968	22800.0	0.6010		
1992	20900.0	0.6143		
1953	19800.0	0.6276		
1987	19700.0	0.6409		
1967	18400.0	0.6542		
2006	18100.0	0.6675		
1978	18000.0	0.6808		
1943	17700.0	0.7074		
2009	17700.0	0.6941		
1976	17200.0	0.7340		
1994	17200.0	0.7207		
1981	16300.0	0.7473		
1963	16200.0	0.7739		
2003	16200.0	0.7606		
* 1941	11700.0	0.7872		
* 1995	11600.0	0.8005		
* 2012	10700.0	0.8138		
* 1985	10400.0	0.8271		
* 1970	9340.0	0.8404		
* 1996	9100.0	0.8538		
* 1972	9030.0	0.8671		
* 2002	7820.0	0.8804		
* 1957	7680.0	0.8937		
* 1989	6730.0	0.9070		
* 1955	6590.0	0.9203		
* 1956	5920.0	0.9336		
* 1988	5150.0	0.9469		
* 1958	4320.0	0.9602		
* 1964	4210.0	0.9735		
* 1977	3120.0	0.9868		

* DENOTES PILF (LO)

PeakFQ output flood-frequency analysis results (cont.)

EMA REPRESENTATION OF DATA

WATER	<----- OBSERVED	----->	<--------- EMA	-------->	<-PERCEPTION	THRESHOLDS->
YEAR	Q_LOWER	Q_UPPER	Q_LOWER	Q_UPPER	LOWER	UPPER
1929	65000.0	65000.0	65000.0	65000.0	61000.0	INF
1930	0.0	61000.0	0.0	61000.0	61000.0	INF
1931	0.0	61000.0	0.0	61000.0	61000.0	INF
1932	0.0	61000.0	0.0	61000.0	61000.0	INF
1933	61000.0	61000.0	61000.0	61000.0	61000.0	INF
1934	0.0	61000.0	0.0	61000.0	61000.0	INF
1935	0.0	61000.0	0.0	61000.0	61000.0	INF
1936	0.0	61000.0	0.0	61000.0	61000.0	INF
1937	0.0	61000.0	0.0	61000.0	61000.0	INF
1938	0.0	61000.0	0.0	61000.0	61000.0	INF
1939	0.0	61000.0	0.0	61000.0	61000.0	INF
1940	0.0	61000.0	0.0	61000.0	61000.0	INF
1941	11700.0	11700.0	0.0	16200.0	16200.0	INF
1942	23000.0	23000.0	23000.0	23000.0	16200.0	INF
1943	17700.0	17700.0	17700.0	17700.0	16200.0	INF
1944	26400.0	26400.0	26400.0	26400.0	16200.0	INF
1945	53300.0	53300.0	53300.0	53300.0	16200.0	INF
1946	25200.0	25200.0	25200.0	25200.0	16200.0	INF
1947	55600.0	55600.0	55600.0	55600.0	16200.0	INF
1948	38200.0	38200.0	38200.0	38200.0	16200.0	INF
1949	23100.0	23100.0	23100.0	23100.0	16200.0	INF
1950	38400.0	38400.0	38400.0	38400.0	16200.0	INF
1951	56400.0	56400.0	56400.0	56400.0	16200.0	INF
1952	28000.0	28000.0	28000.0	28000.0	16200.0	INF
1953	19800.0	19800.0	19800.0	19800.0	16200.0	INF
1954	49400.0	49400.0	49400.0	49400.0	16200.0	INF
1955	6590.0	6590.0	0.0	16200.0	16200.0	INF
1956	5920.0	5920.0	0.0	16200.0	16200.0	INF
1957	7680.0	7680.0	0.0	16200.0	16200.0	INF
1958	4320.0	4320.0	0.0	16200.0	16200.0	INF
1959	24700.0	24700.0	24700.0	24700.0	16200.0	INF

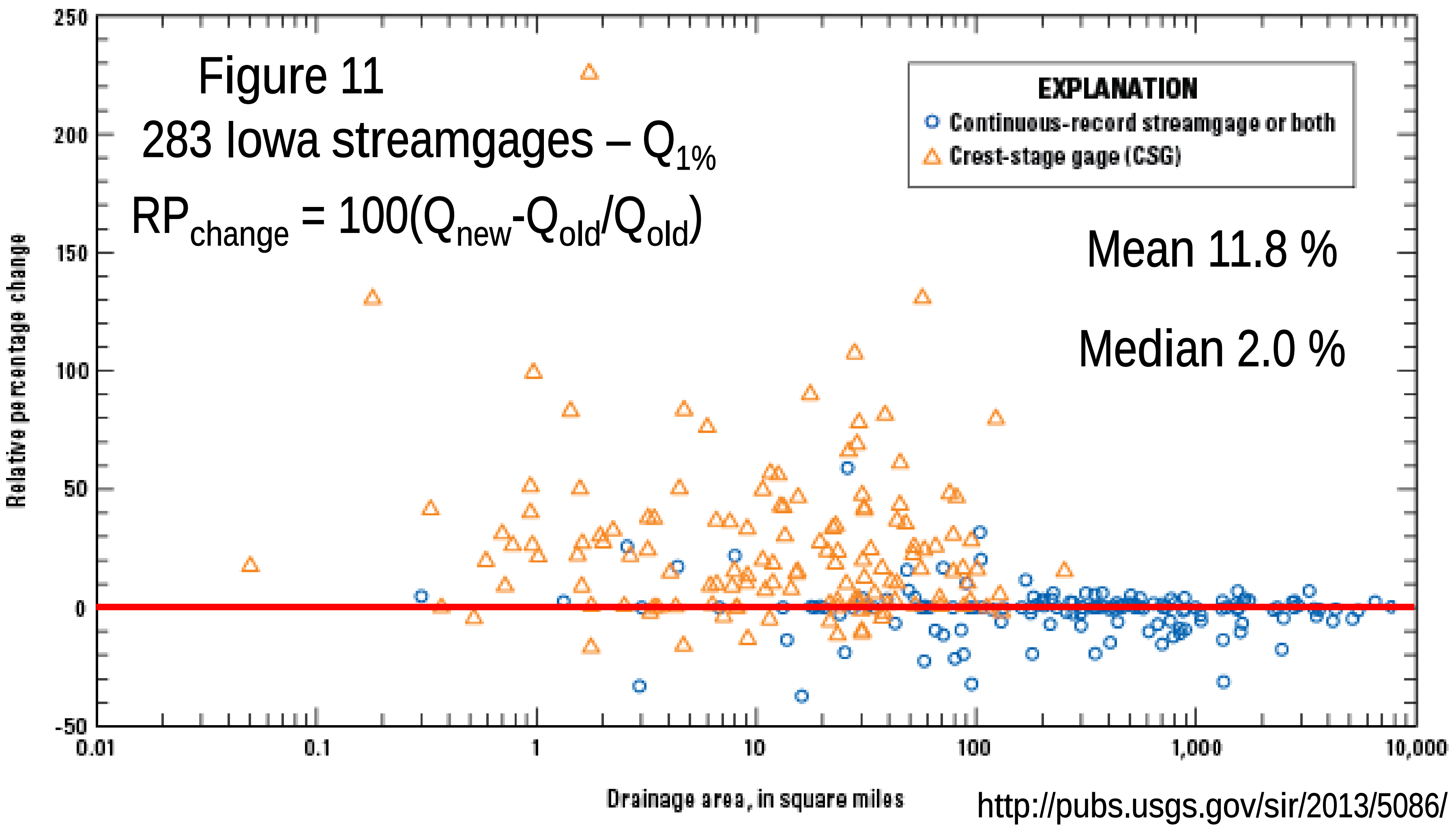
Historic Period

PILF

MGB threshold = 16,200 cfs

PILFs

Relative % change between EMA/MGB & B17B/GB using new skew for $Q_{1\%}$



http://pubs.usgs.gov/sir/2013/5086/

USGS

Final flood-frequency estimates for streamgages computed using WIE

$$\log(\hat{Q}) = \frac{\mathrm{MSE}[\log(Q_{re})]\cdot\log(Q_{as}) + \mathrm{MSE}[\log(Q_{as})]\cdot\log(Q_{re})}{\mathrm{MSE}[\log(Q_{as})] + \mathrm{MSE}[\log(Q_{re})]}$$

Weighted Independent Estimates from appendix 8 of Bulletin 17B

B17B/EMA and regional regression equation estimates are considered independent estimates

WIE program uses the variance and estimate of the B17B/EMA flood-frequency analysis and the variance and estimate of the regional regression equation to compute a weighted flood-frequency estimate and variance at a streamgage

USGS

WIE estimates for Cedar River at Waterloo streamgage 05464000

05464000				
Annual exceedance probability of instantaneous peak discharges, based on WIE method, EMA/MGB analysis computed using a historical period length of 85 years (1929-2013)				
Annual exceed-ance probability	Recur-rence interval (years)	Discharge (ft^3/s)	95% lower confi-dence interval (ft^3/s)	95% upper confi-dence interval (ft^3/s)
0.500	2	25,800	22,400	29,700
0.200	5	43,200	37,500	49,900
0.100	10	55,000	47,100	64,300
0.040	25	71,100	59,200	85,400
0.020	50	82,800	67,100	102,000
0.010	100	94,100	73,900	120,000
0.005	200	110,000	83,200	145,000
0.002	500	123,000	88,900	170,000
KENTAU statistic		0.008		
P-level		0.924		
Begin year		1941		
End year		2013		
Number of peaks		73		

EMA/MGB $Q_{1\%}$ = 101,000 cfs

variance of estimate = 0.0041

RRE $Q_{1\%}$ = 80,400 cfs

variance of estimate = 0.0090

WIE $Q_{1\%}$ = 94,100 cfs

variance of estimate = 0.0029


QUESTIONS