An Overview of High- and Low-Flow Statistics and Why Record Length Matters

Streamflow Monitoring Workshop

February 3, 2016

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

"And it never failed that during the dry years the people forgot about the rich years, and during the wet years, they lost all memory of the dry years. It was always that way." –John Steinbeck *East of Eden*

Outline

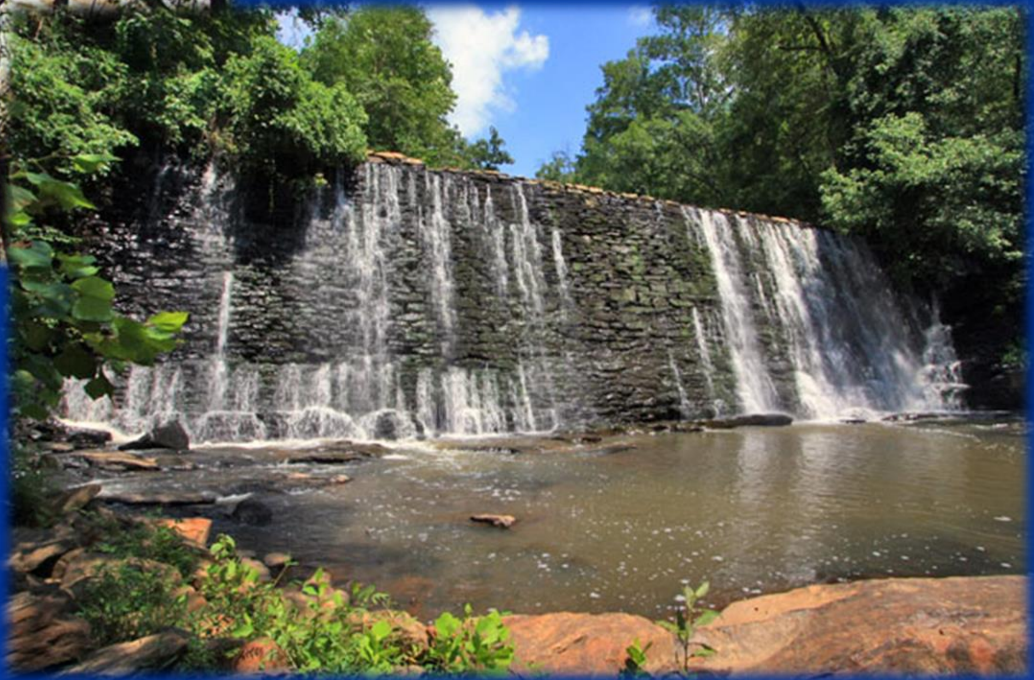
1) Basics of computing high- and low-flow frequency statistics

2) Importance of Record Length

3) Regionalization and geographical coverage

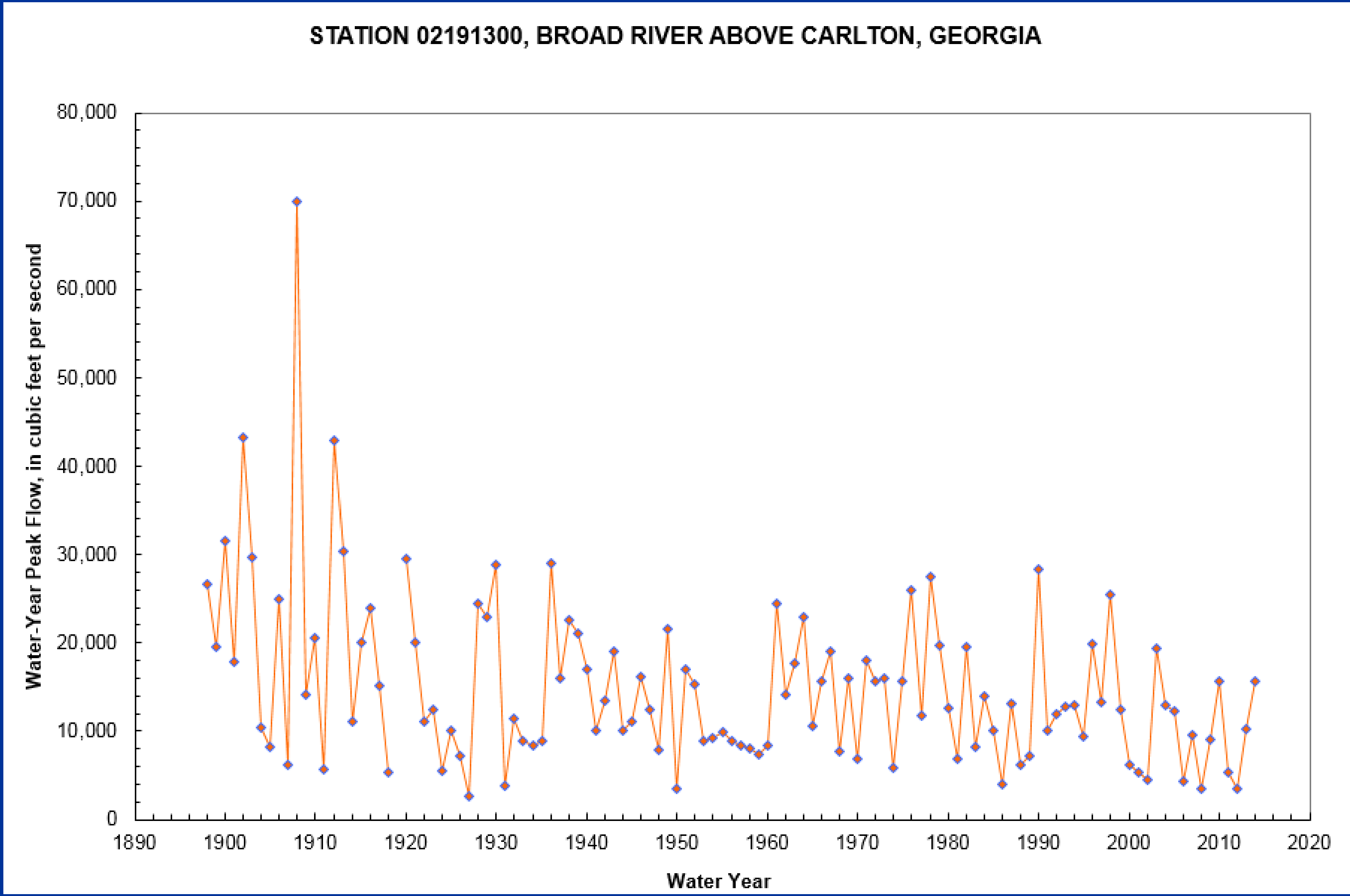
USGS

Statistical Basics for High- and Low-Flow Frequency Analyses

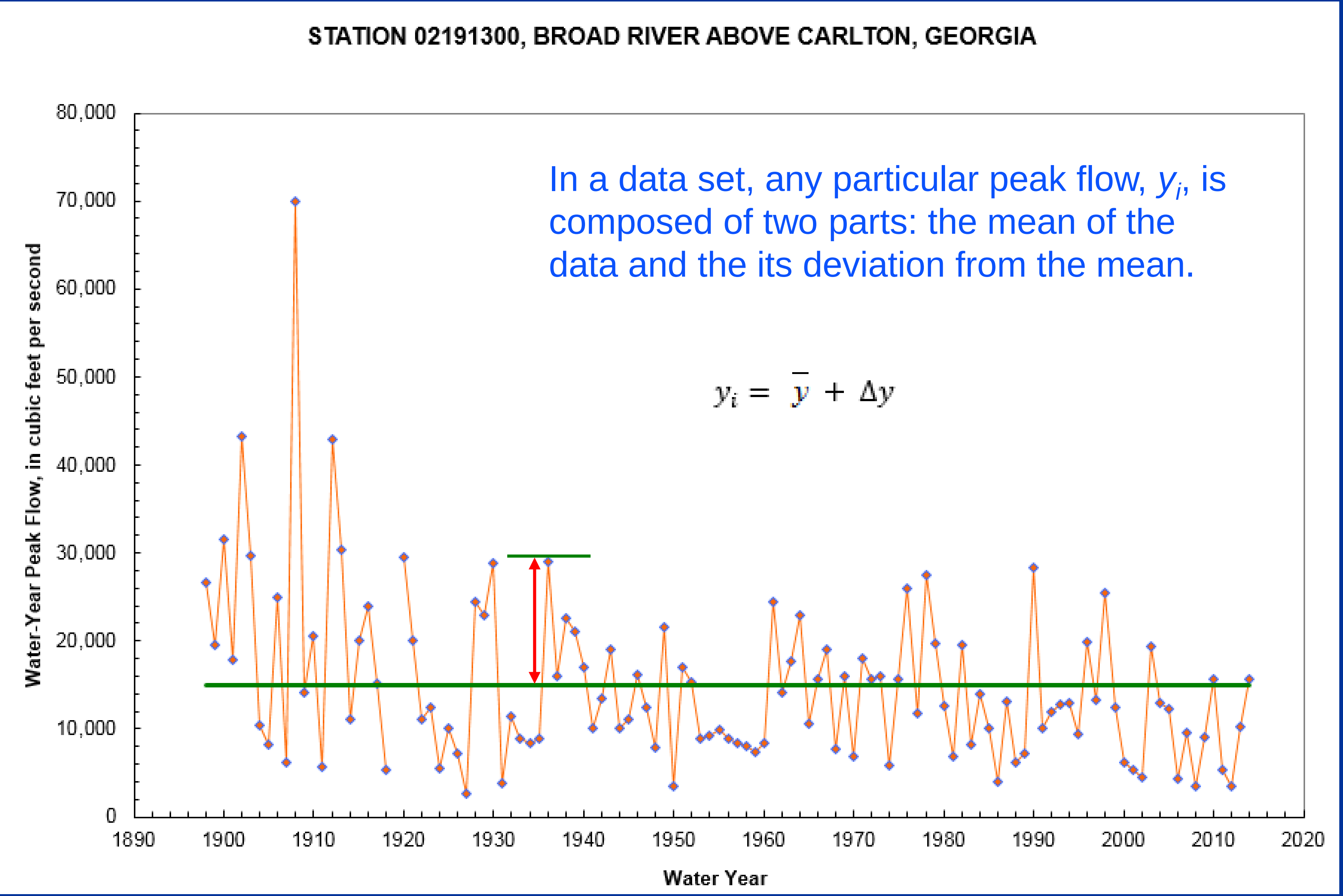


USGS

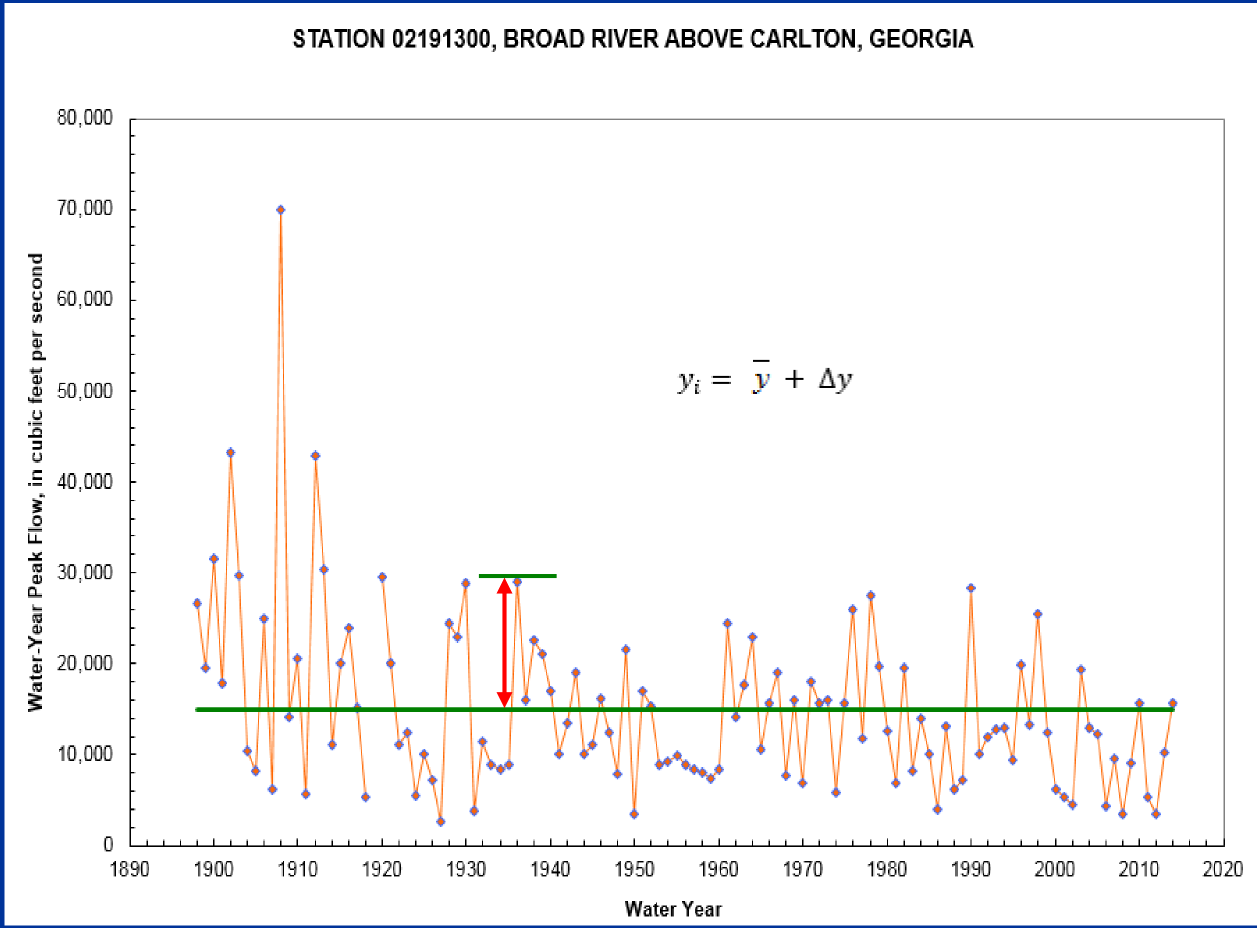
In 1951, Ven Te Chow published a paper titled "A General Formula For Hydrologic Frequency Analysis"



In 1951, Ven Te Chow published a paper titled "A General Formula For Hydrologic Frequency Analysis"

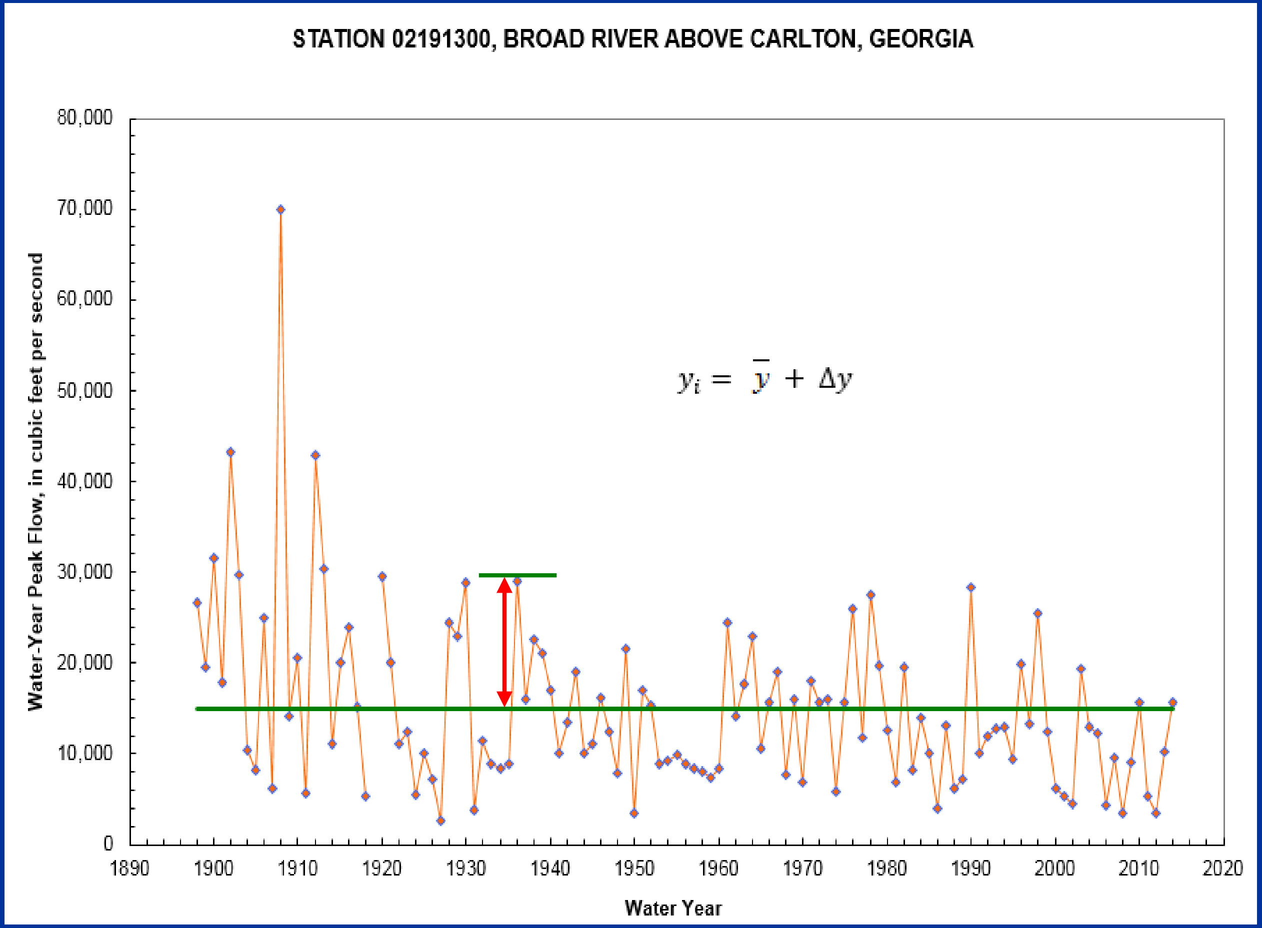


In 1951, Ven Te Chow published a paper titled “A General Formula For Hydrologic Frequency Analysis”



The departure from the mean may be positive or negative, large or small. From a statistical perspective, this quantity possesses two important properties: (1) the tendency to deviate from the mean; and (2) the frequency of occurrence.

In 1951, Ven Te Chow published a paper titled “A General Formula For Hydrologic Frequency Analysis”

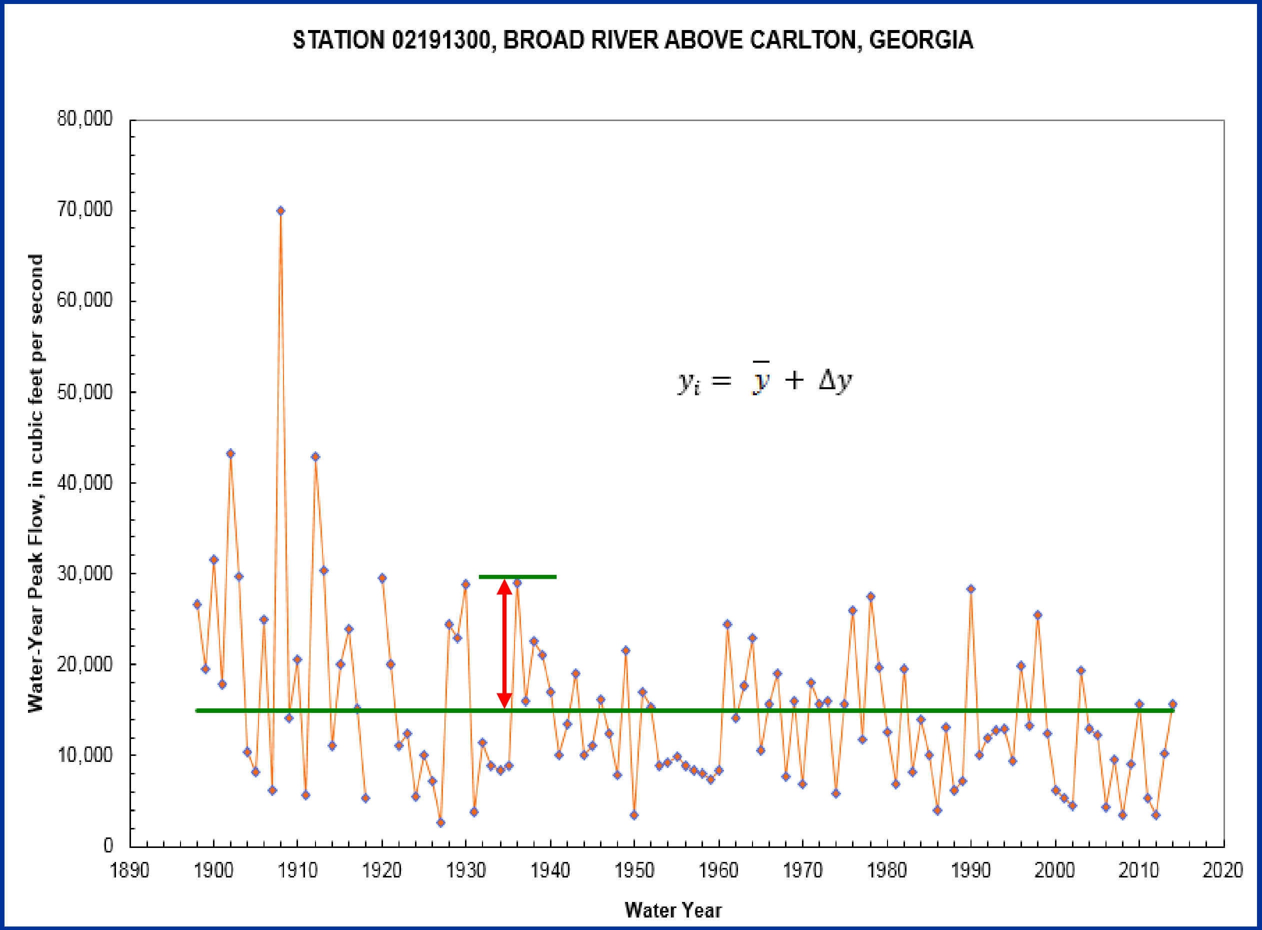


Property 1

If you sum up all the deviations from the mean, by definition, the value would be zero. Thus, an average of the deviations does not provide any useful information.

One way to overcome that problem is to square the deviations and take the average of that value. However, the resulting value is now in different units than the original data.

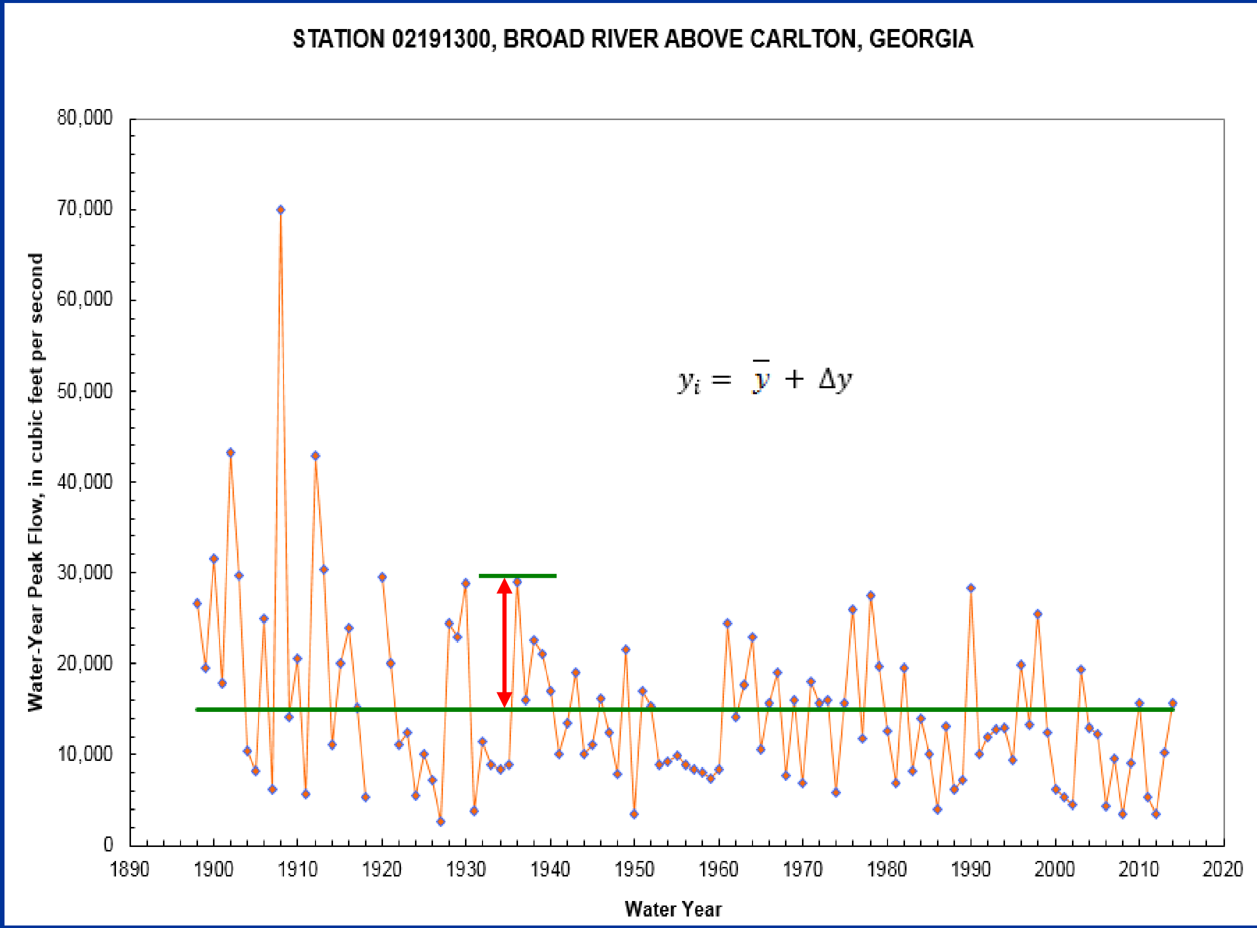
In 1951, Ven Te Chow published a paper titled "A General Formula For Hydrologic Frequency Analysis"



To get back to the original units, we can take the square root of the summed deviations.

USGS

In 1951, Ven Te Chow published a paper titled "A General Formula For Hydrologic Frequency Analysis"



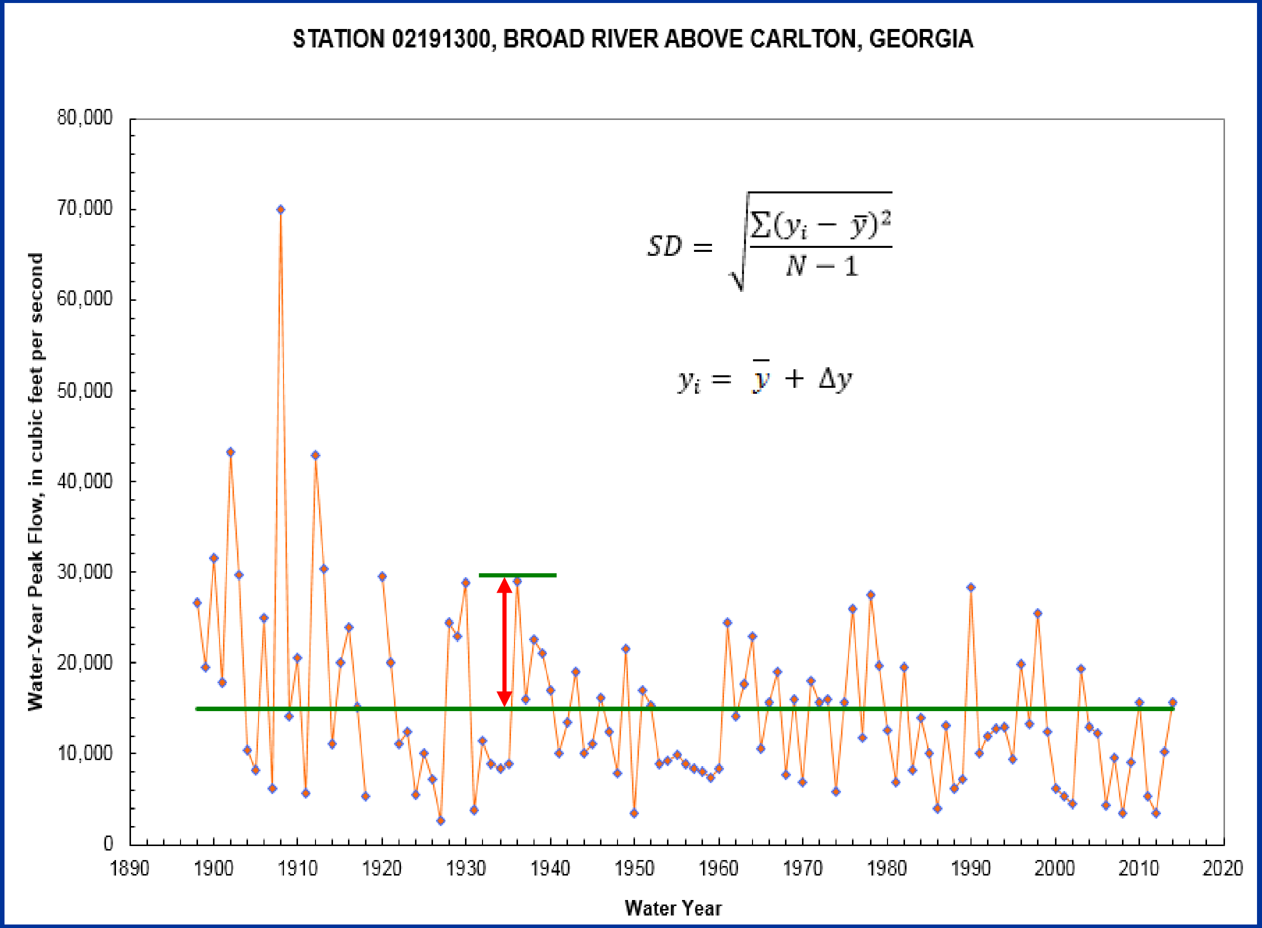
Thus, the tendency to deviate from the mean can be measured by the resulting equation, which is the standard deviation.

$$SD = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{N - 1}}$$

Consequently, knowing the mean and SD of a data set provides very useful information.

USGS

In 1951, Ven Te Chow published a paper titled "A General Formula For Hydrologic Frequency Analysis"

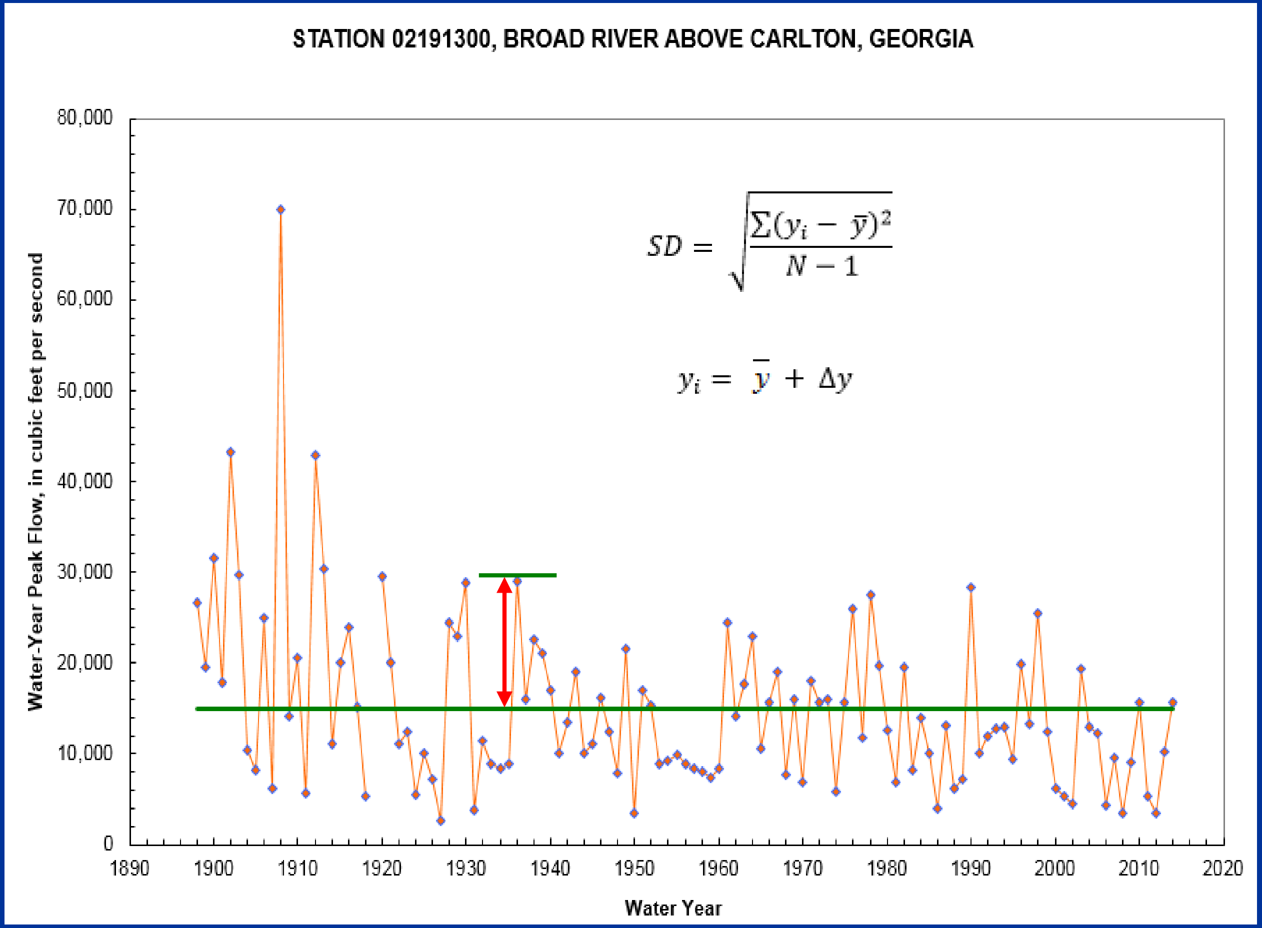


Going back to our earlier slide: Property 2

The departure from the mean may positive or negative, large or small. From a statistical perspective, this quantity possesses two important properties: (1) the tendency to deviate from the mean; and (2) the frequency of occurrence.

USGS

In 1951, Ven Te Chow published a paper titled "A General Formula For Hydrologic Frequency Analysis"



This second property is measured by a term called the frequency factor, K, which depends upon the law of occurrence of a particular hydrologic event under consideration and is a function of the return period and distribution parameters.

This all comes together in the log-Pearson Type III equation.

USGS

Log Pearson Type III Distribution

The log Q corresponding to specific exceedance probabilities, P, are calculated as:

$$\log Q = \overline{X} + KS$$

Where $\overline{X}$ = **mean of logarithms of peaks**

K = frequency factor = f(skew, P)

S = standard deviation of logs of peaks

USGS

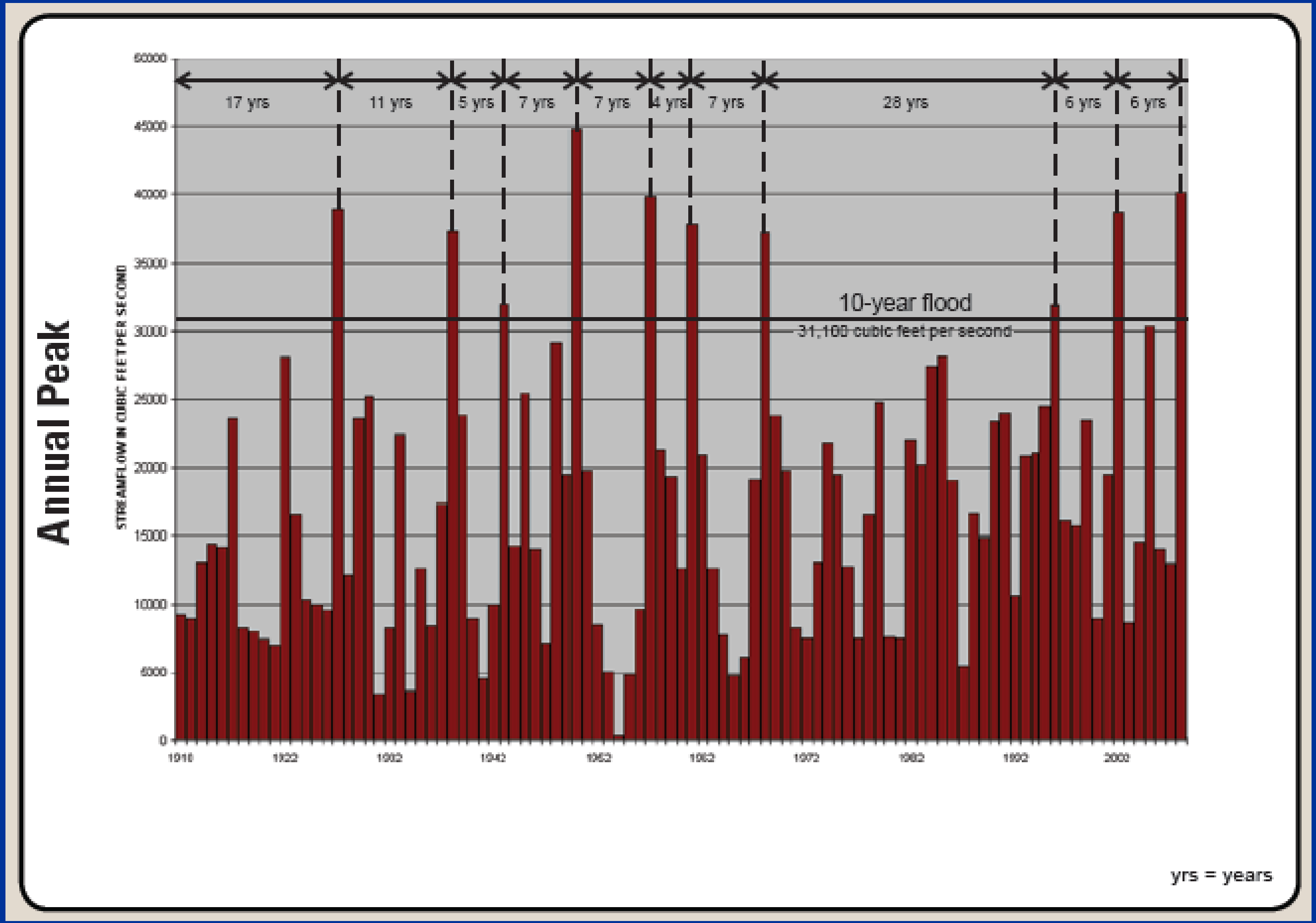
T-Year Recurrence Interval

Historically, flood-frequency flows in USGS reports were expressed as *T-year floods* based on the *recurrence interval* of the particular flood quantile.

- **Recurrence interval (or return period) is the average interval of time within which the given flood will be equaled or exceeded once.**
- **The words "recurrence" or "return" do not imply regular predictable occurrence in time.**
- **An example is the 10-year flood. If we had a 100 years of streamflow data, we would expect 10 floods of equal or greater magnitude than the "10-year flood". However, those floods would not occur at regular 10-year intervals.**

USGS

An Example For A 10-year Recurrence



USGS

Annual Exceedance Probability

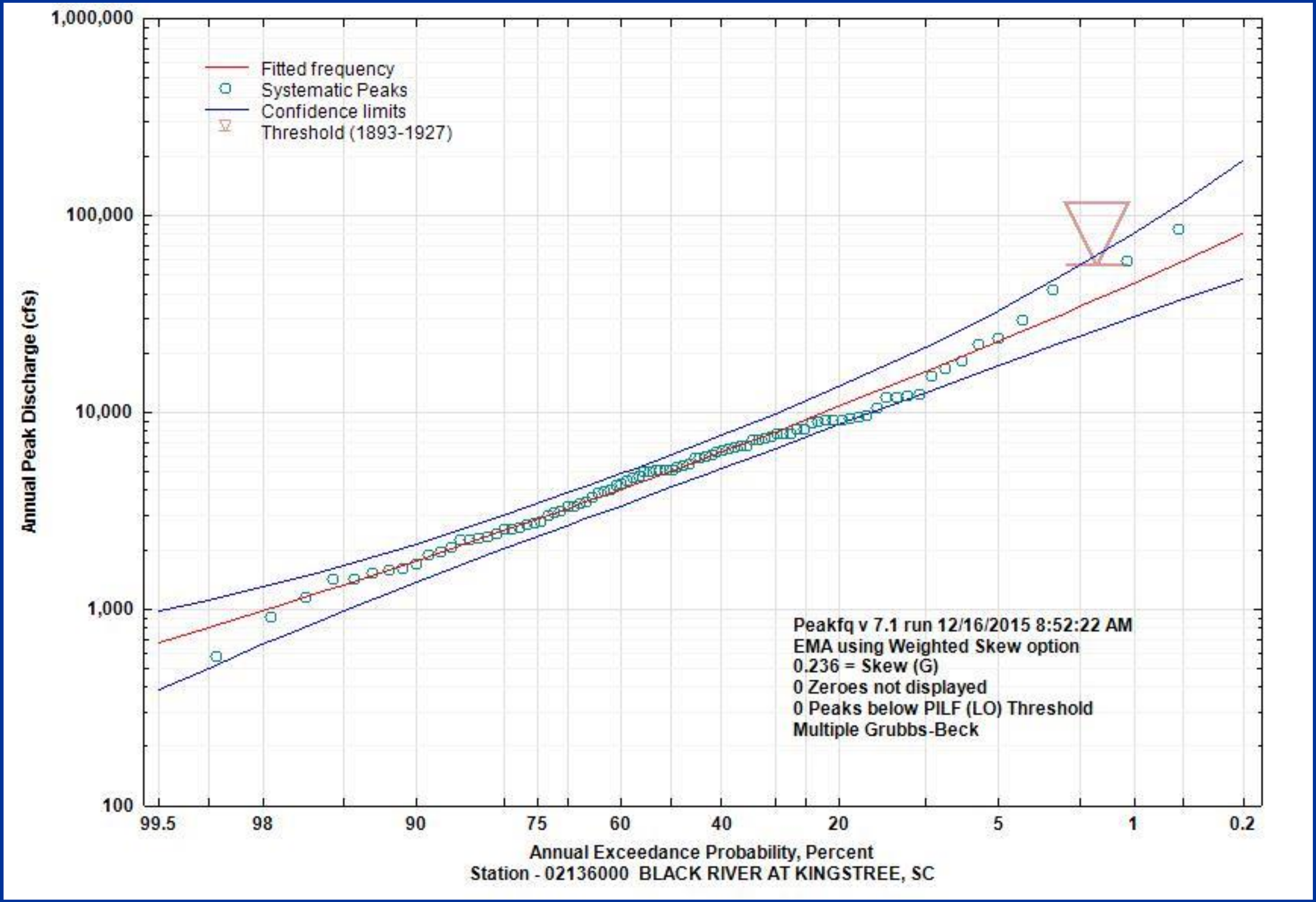
Return period (or recurrence interval),T = 1 / AEP

*P-percent chance exceedance = AEP * 100*

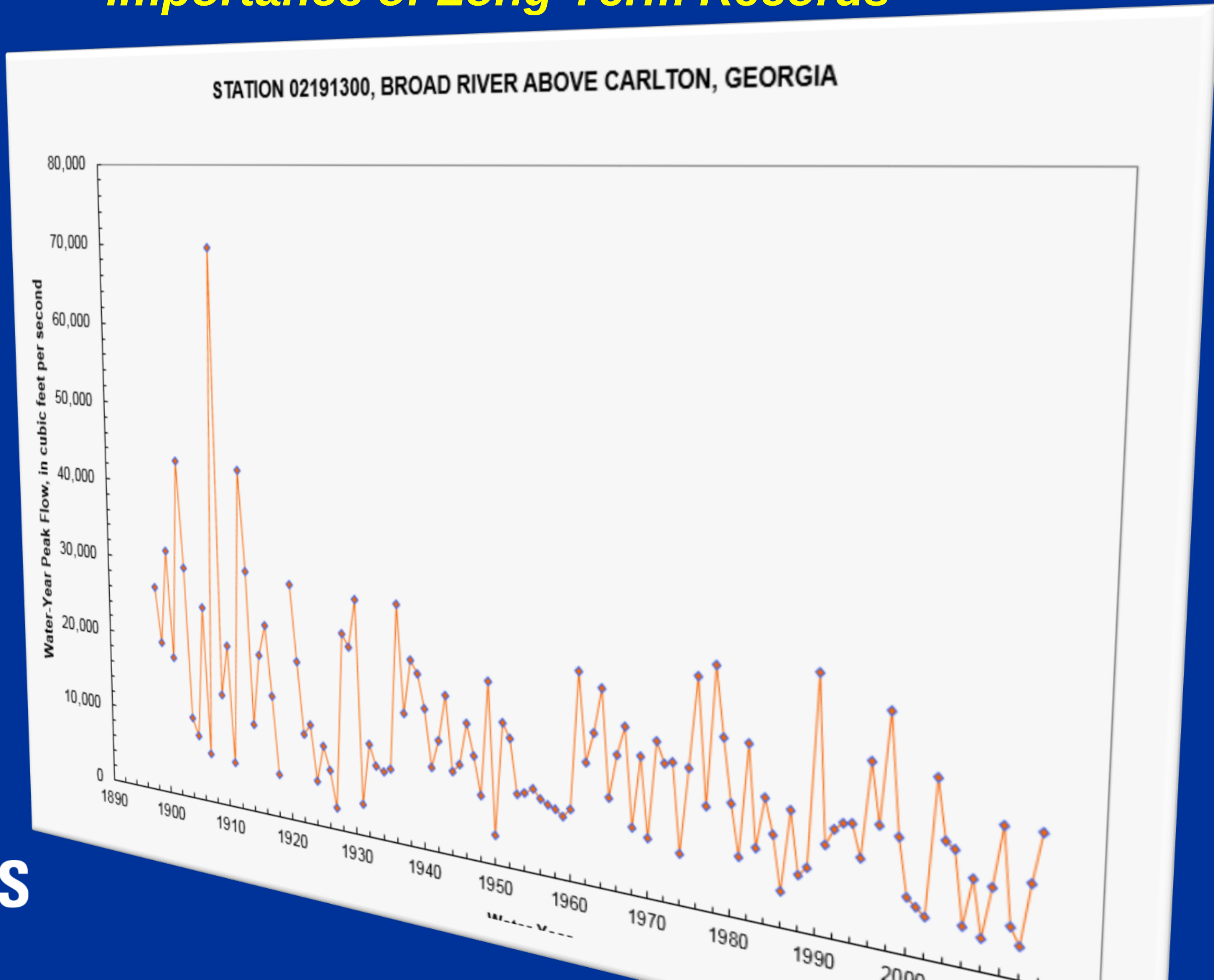
T-year recurrence interval	Annual exceedance probability	P-percent chance exceedance
2	0.5	50
5	0.2	20
10	0.10	10
25	0.04	4
50	0.02	2
100	0.01	1
200	0.005	0.5
500	0.002	0.2

USGS

Example of an LPIII Analysis



Importance of Long-Term Records



USGS

Populations versus Samples

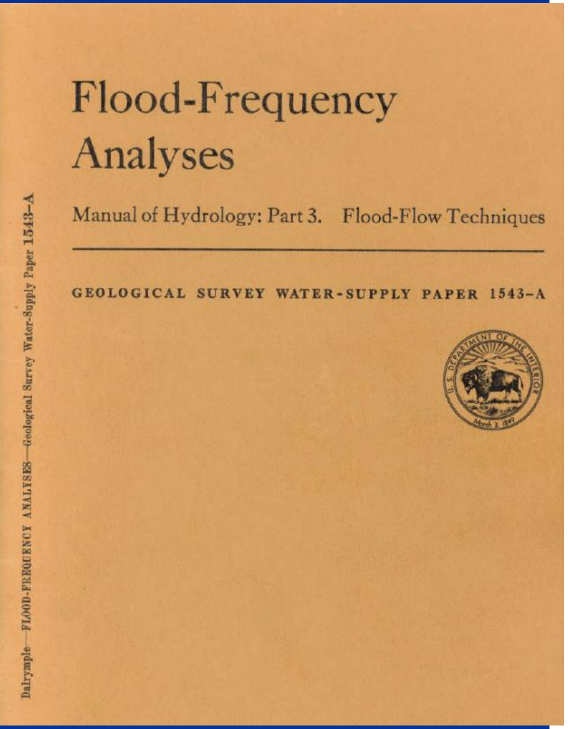
- There is a **"population"** of potential events (in our case, annual peak flows).

- All that is available to us are **samples** from the population (periods of record). We infer magnitudes and frequencies of events from the sample.

- An assumption is made that the samples are **"representative"**.

- Different samples will yield different estimates.

Populations versus Samples



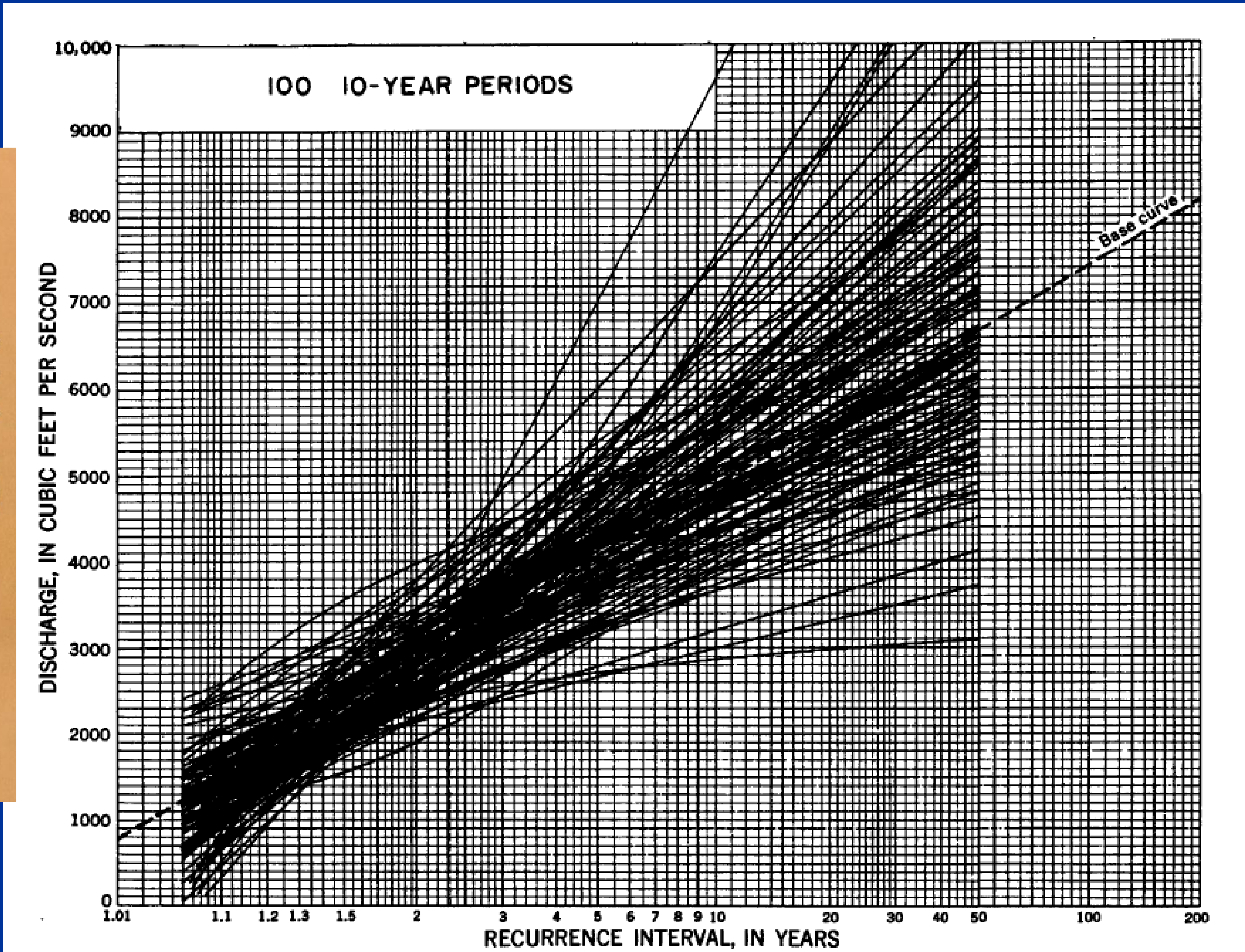


FIGURE 20.—Frequency curves for 10-year periods.

USGS

Populations versus Samples

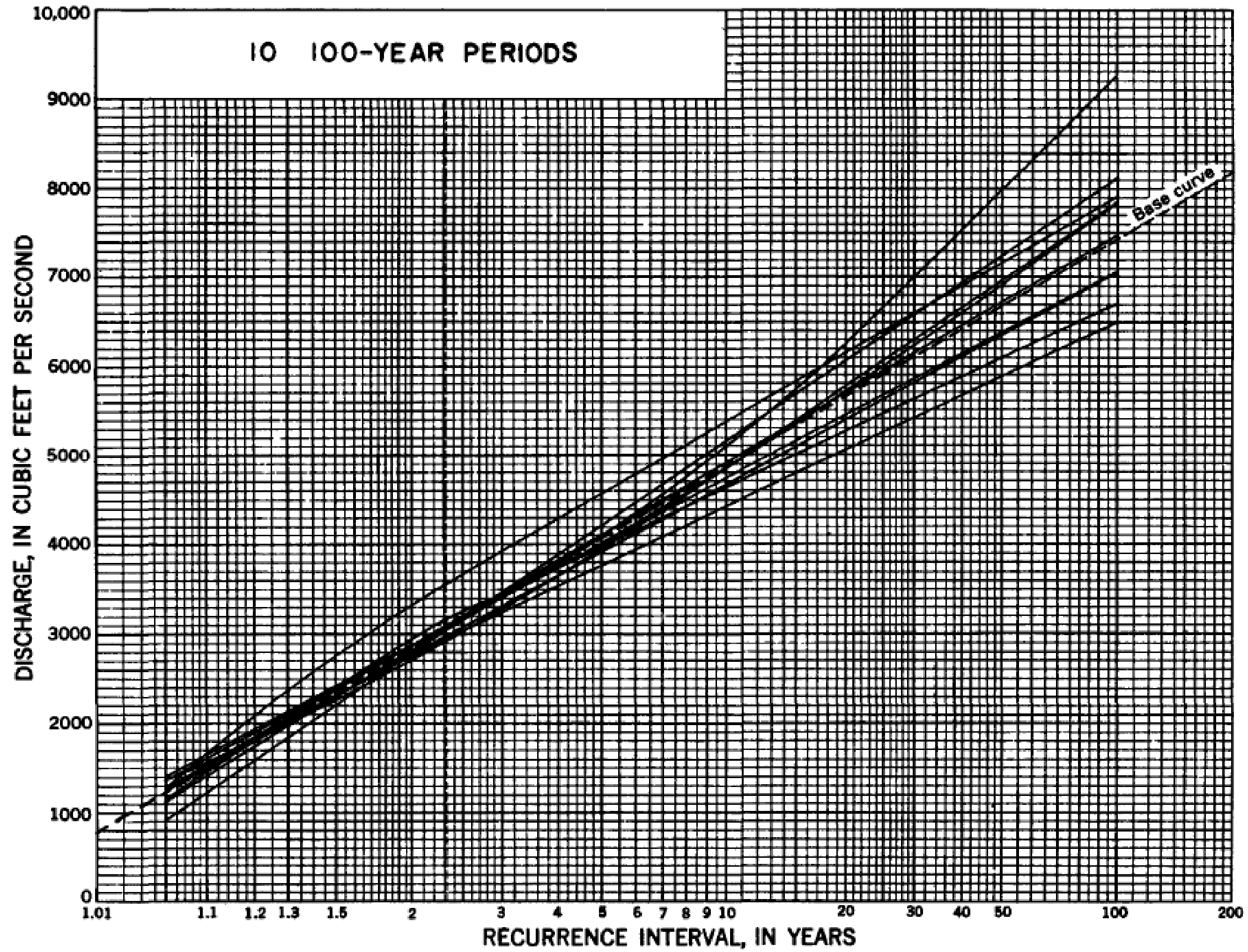
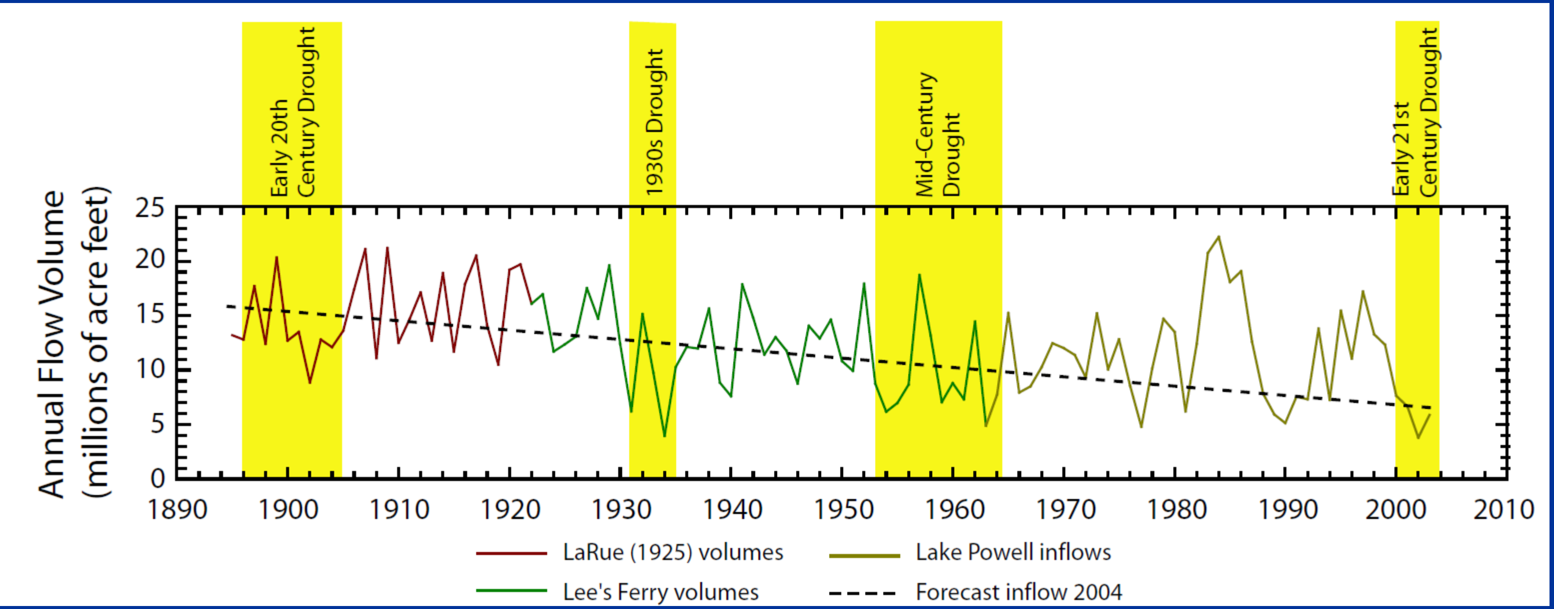


FIGURE 23.—Frequency curves for 100-year periods.

USGS

Importance of Long-Term Records

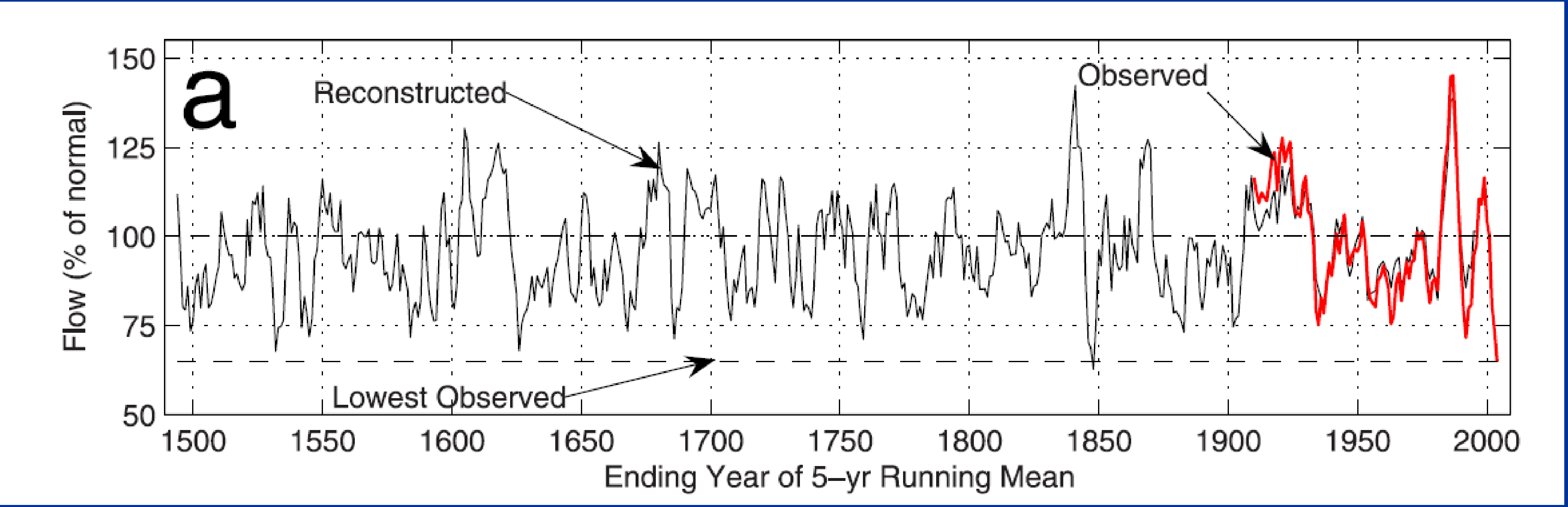
Longer records tend to capture a broader range of hydrologic conditions and will better represent the "population" value.



USGS

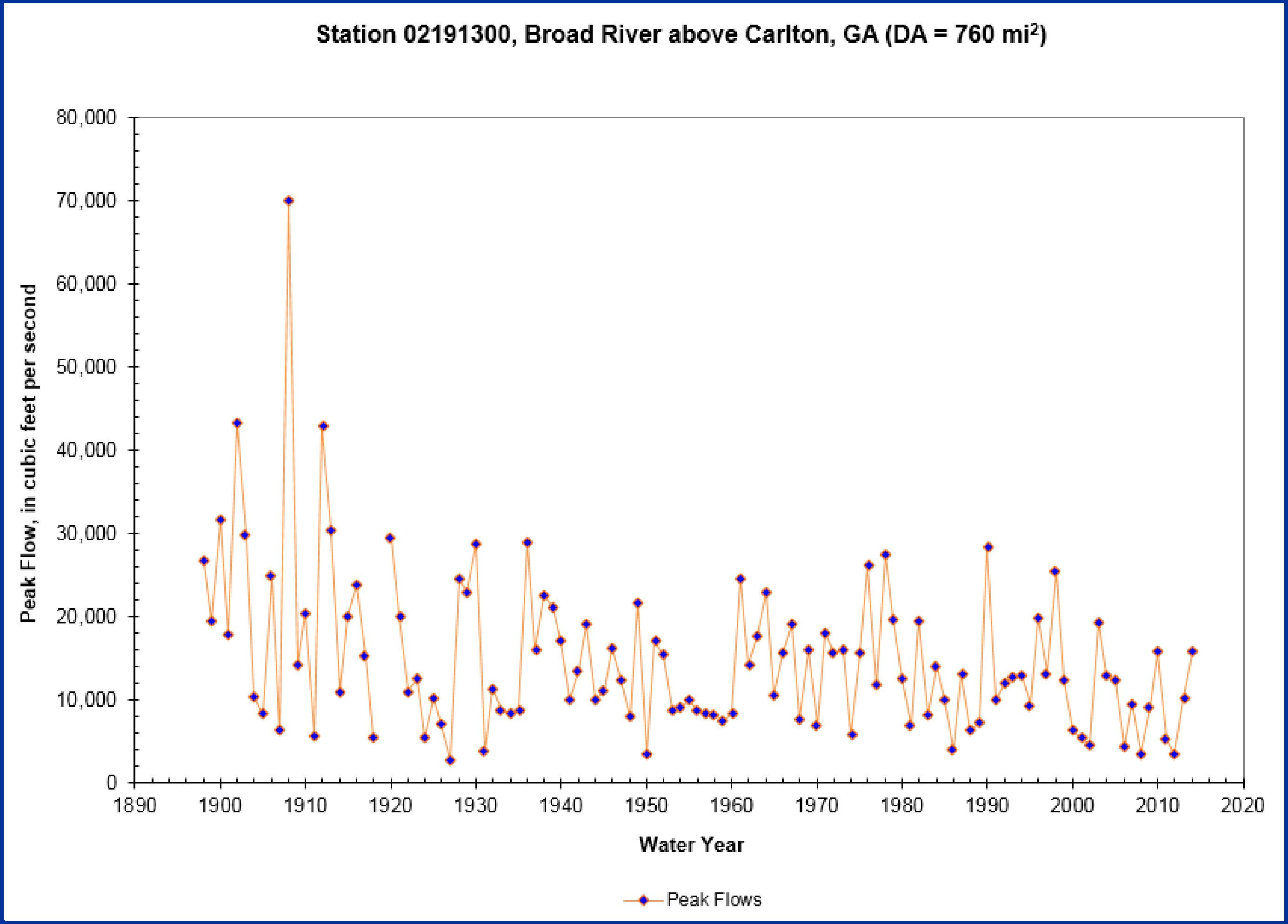
Unfortunately, the period from about 1905-1922, which time and additional research as shown was an unusually wet period, was used to estimate water production allocation under the Colorado River Compact.

Importance of Long-Term Records



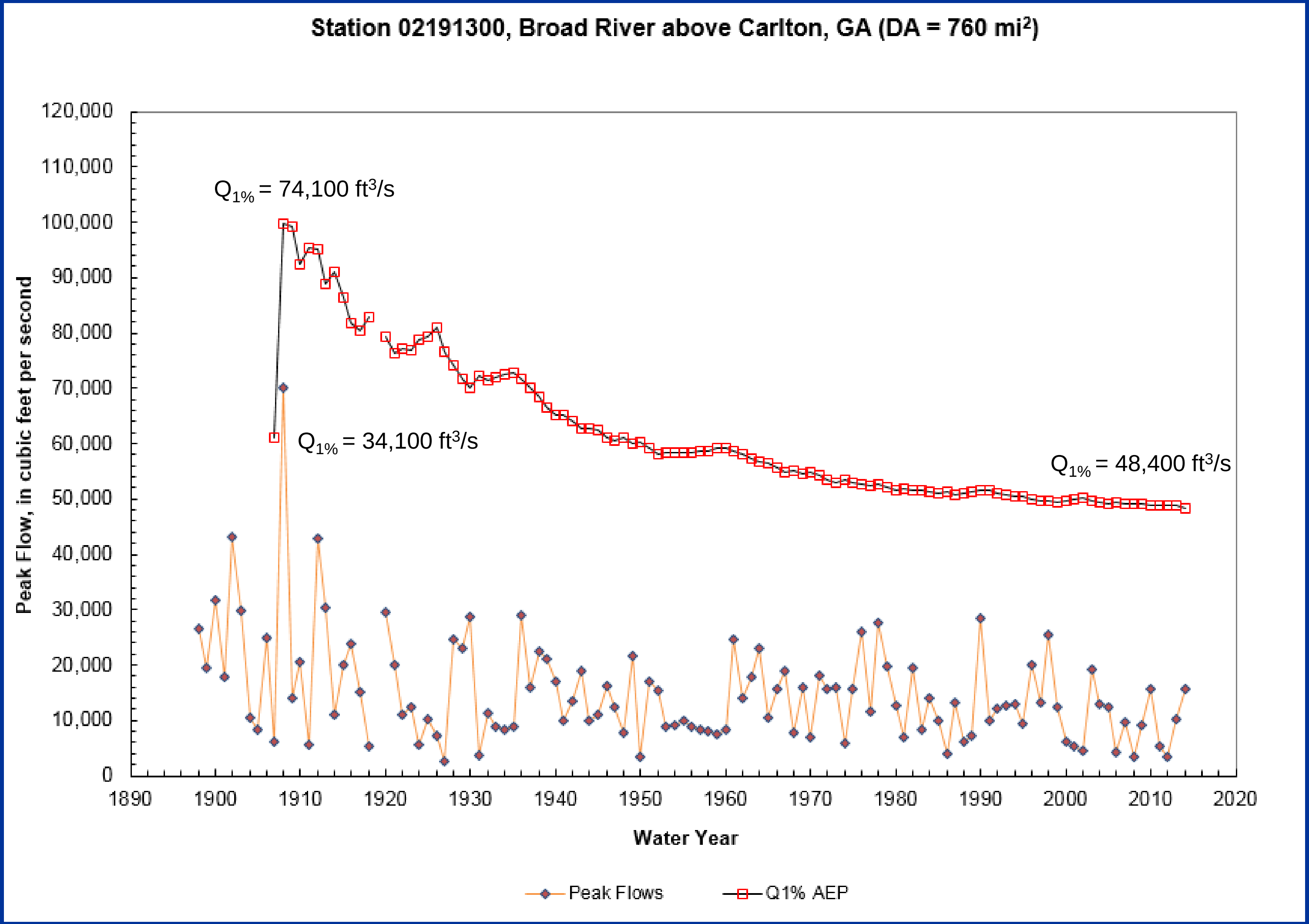
This reconstructed (tree rings) record going back to the 1400s shows how truly unfortunate it was for the Colorado River Compact water allocations were based on river flow data from the short period of record in the earlier 1900s.

Lets Look At An Example of How The 1% Chance Flood Changes Over Time Based on Record Length and Hydrologic Conditions



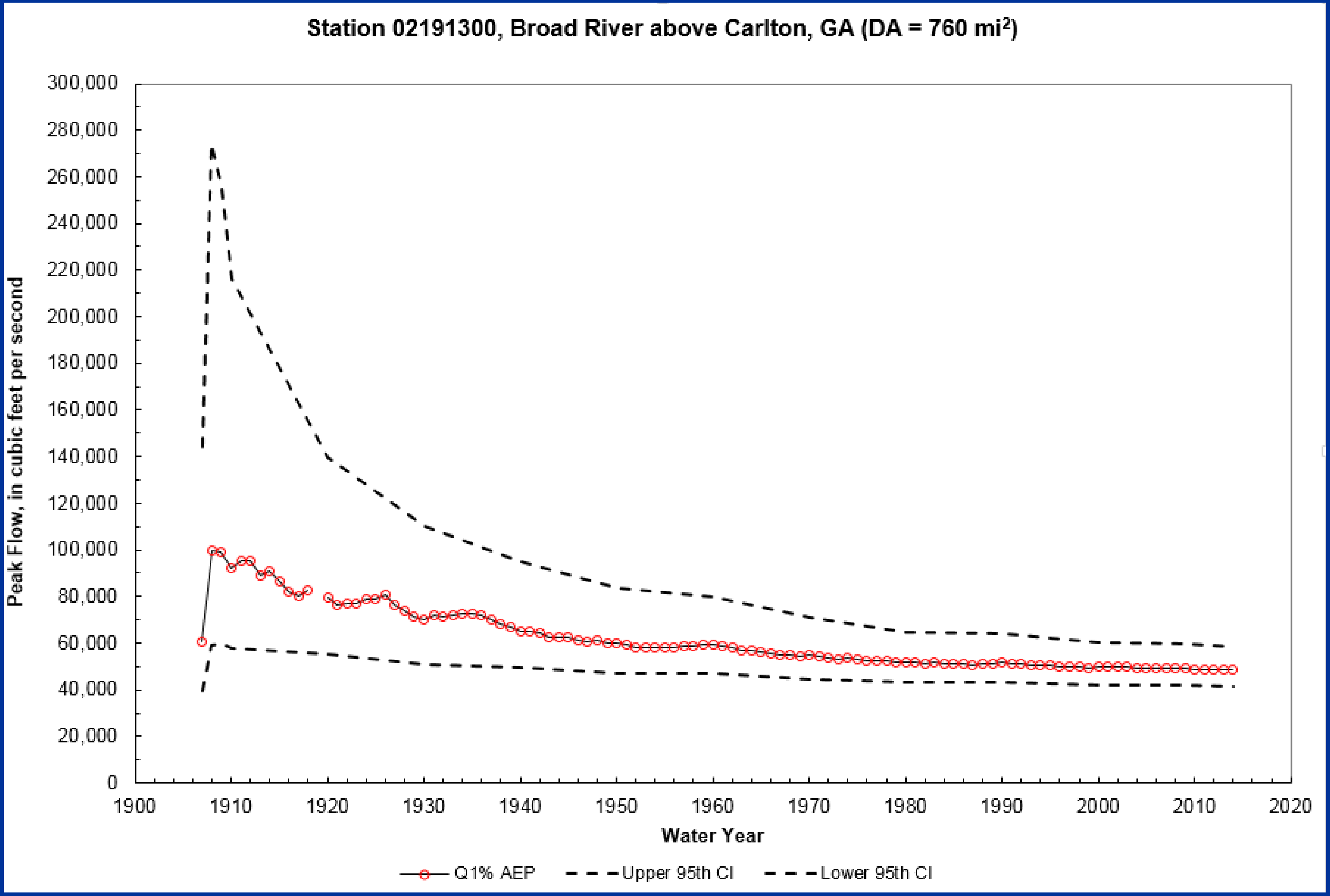
USGS

How the 1% AEP Changes With Additional Record



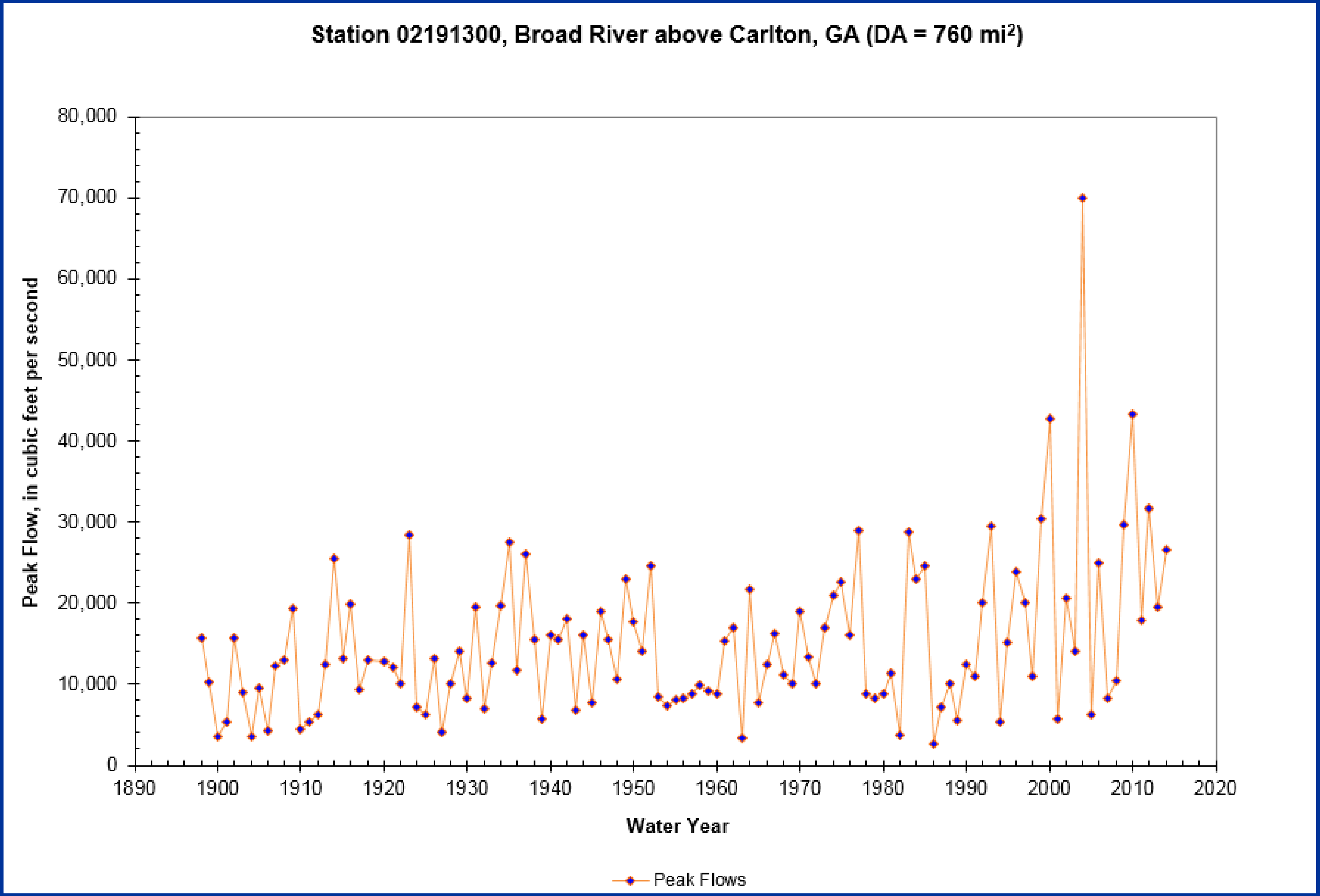
USGS

It's Important to be Aware of the Uncertainty



USGS

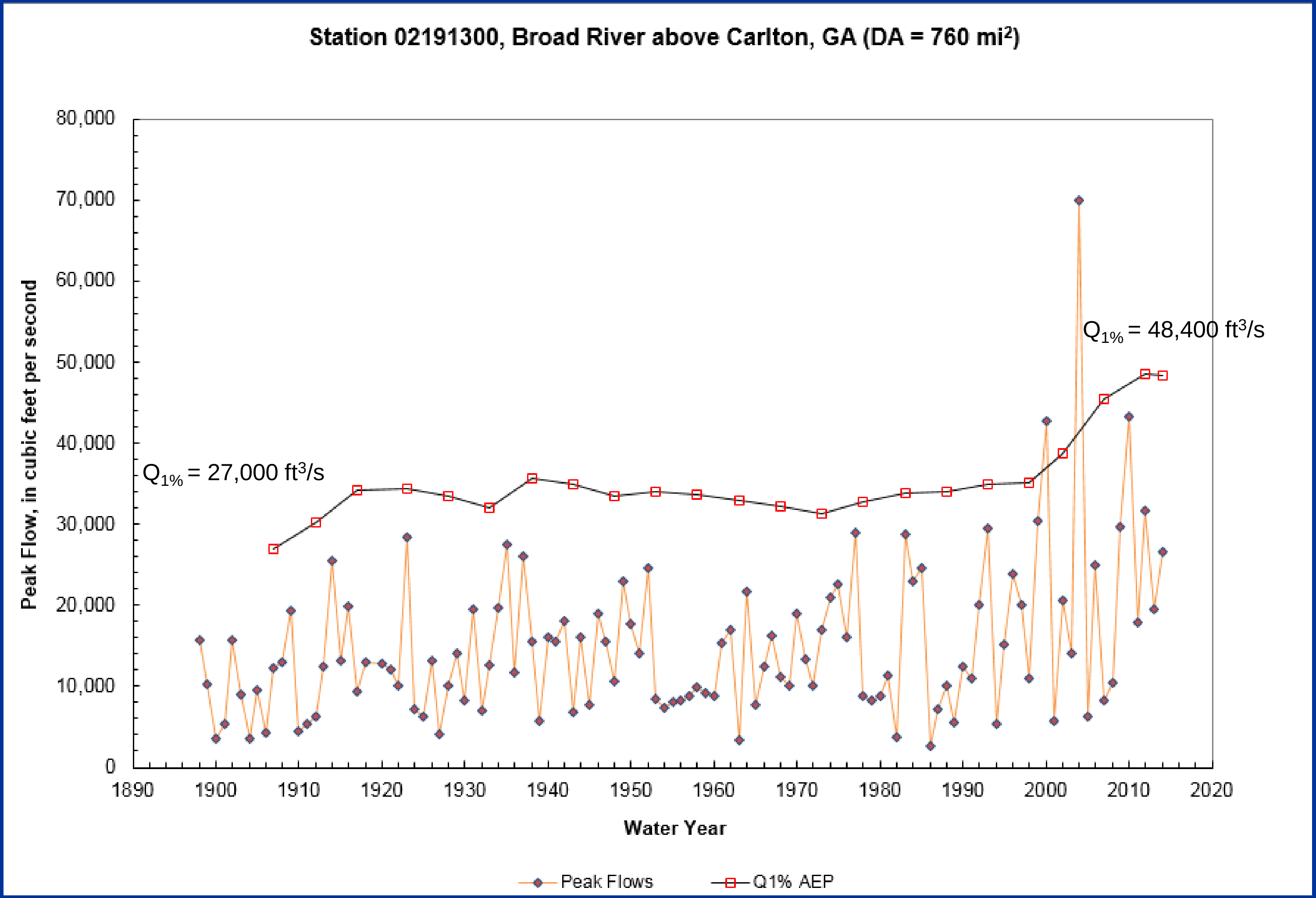
What If The Record Had Been Collected In Reverse Order?



USGS

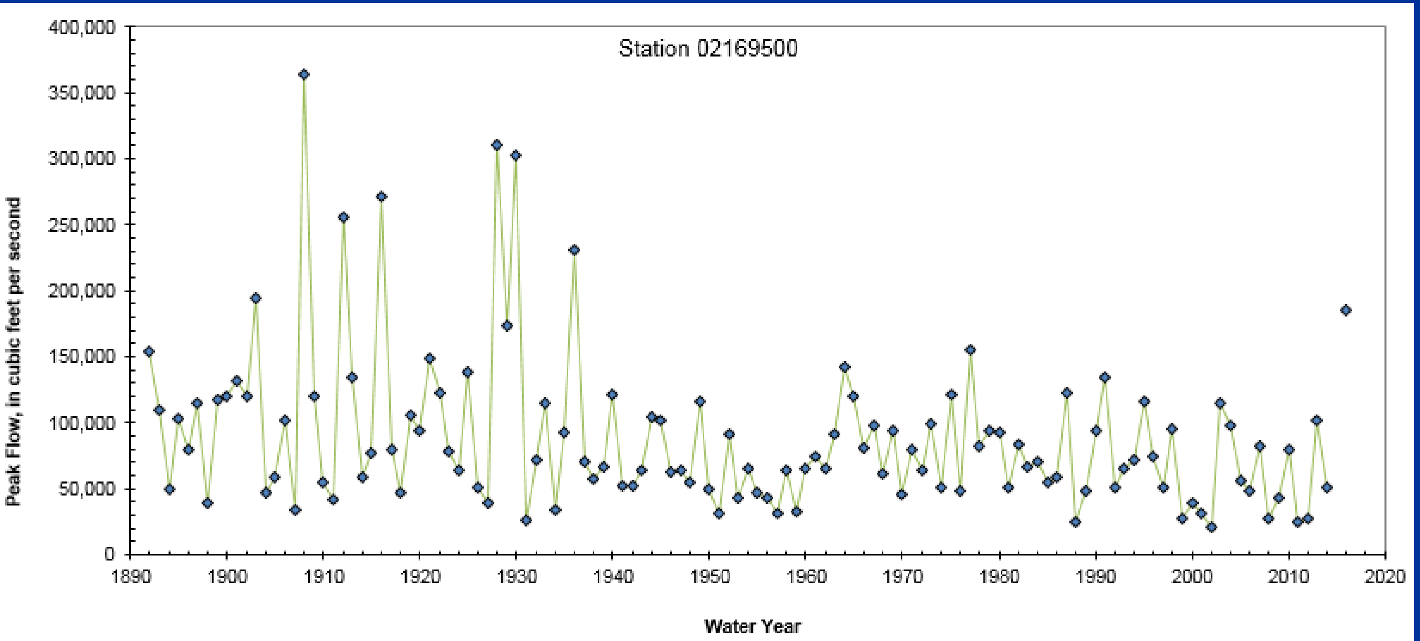
What If The Record Had Been Collected In Reverse Order?

Period of record (and hydrologic conditions captured in that record) is a big deal.



USGS

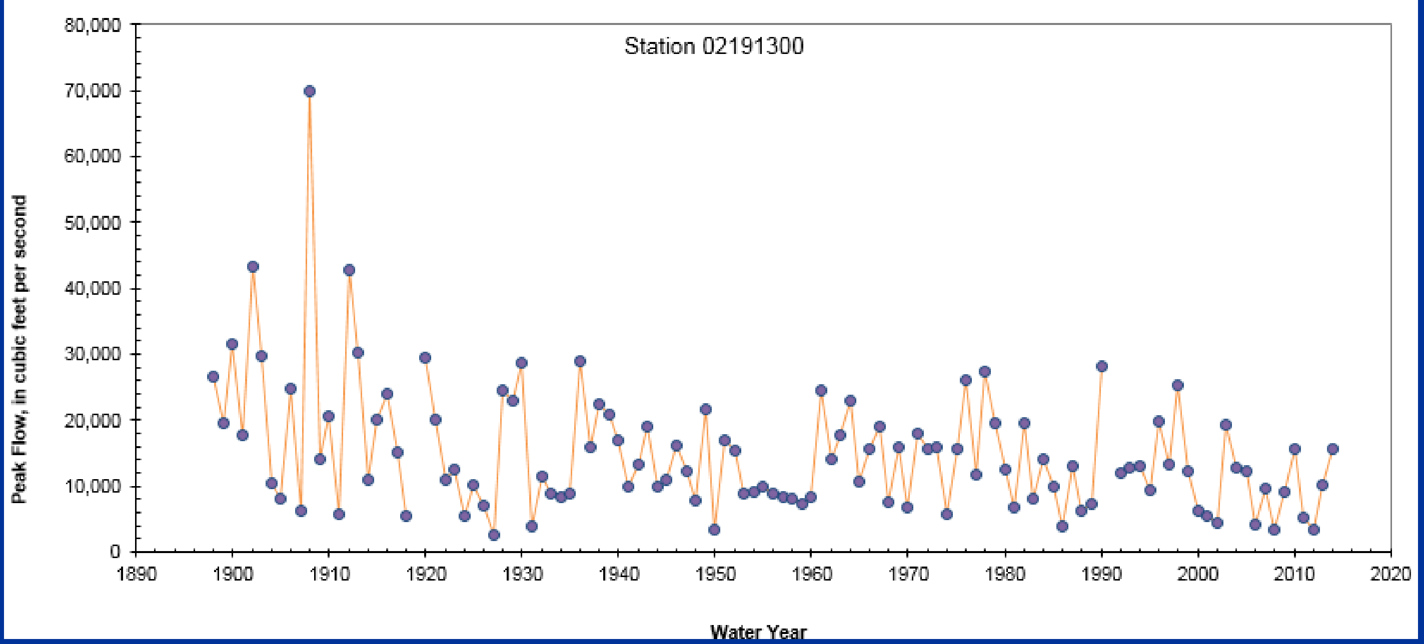
BTW



Drainage Area

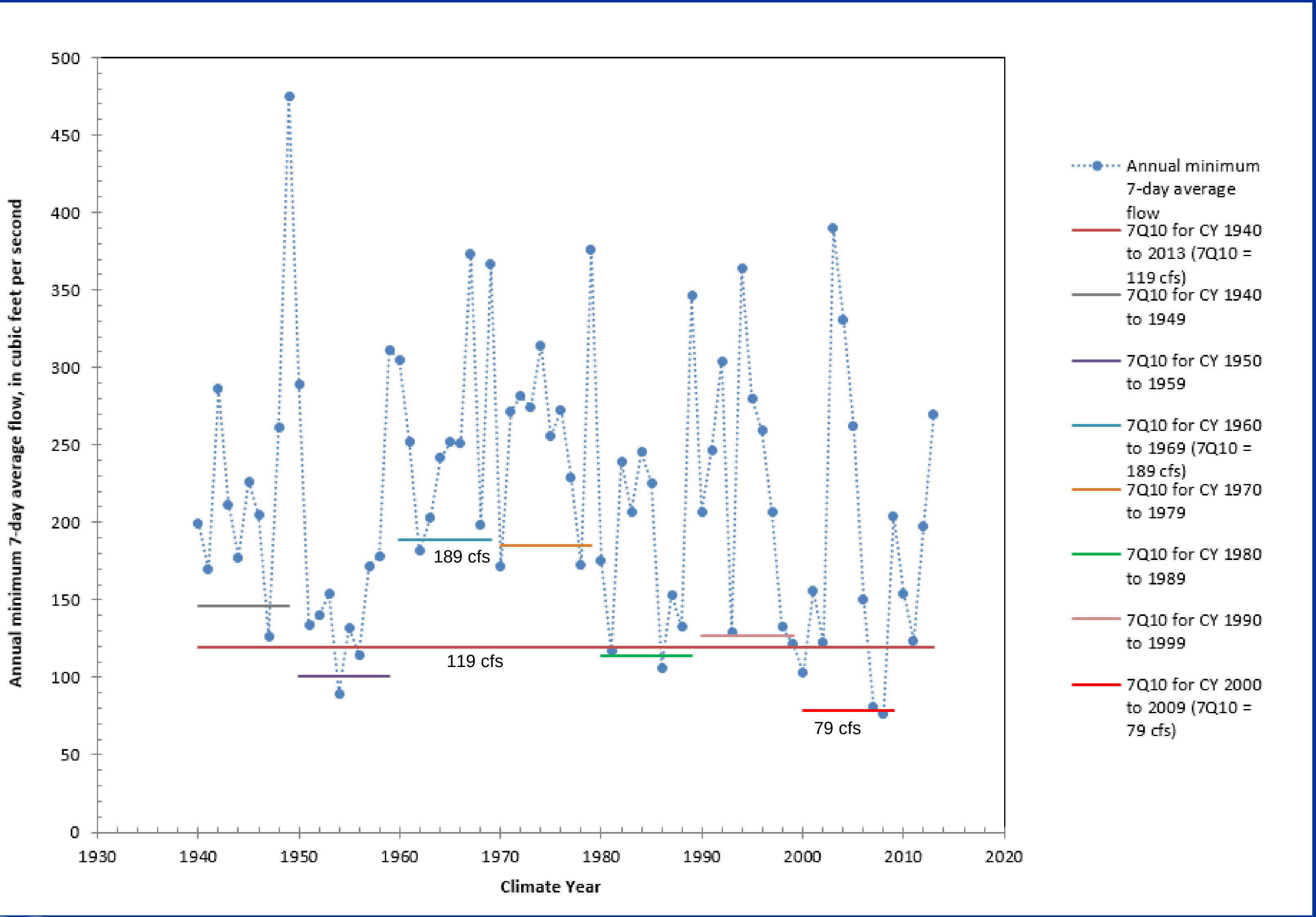
02169500 = 7,850

02191300 = 760



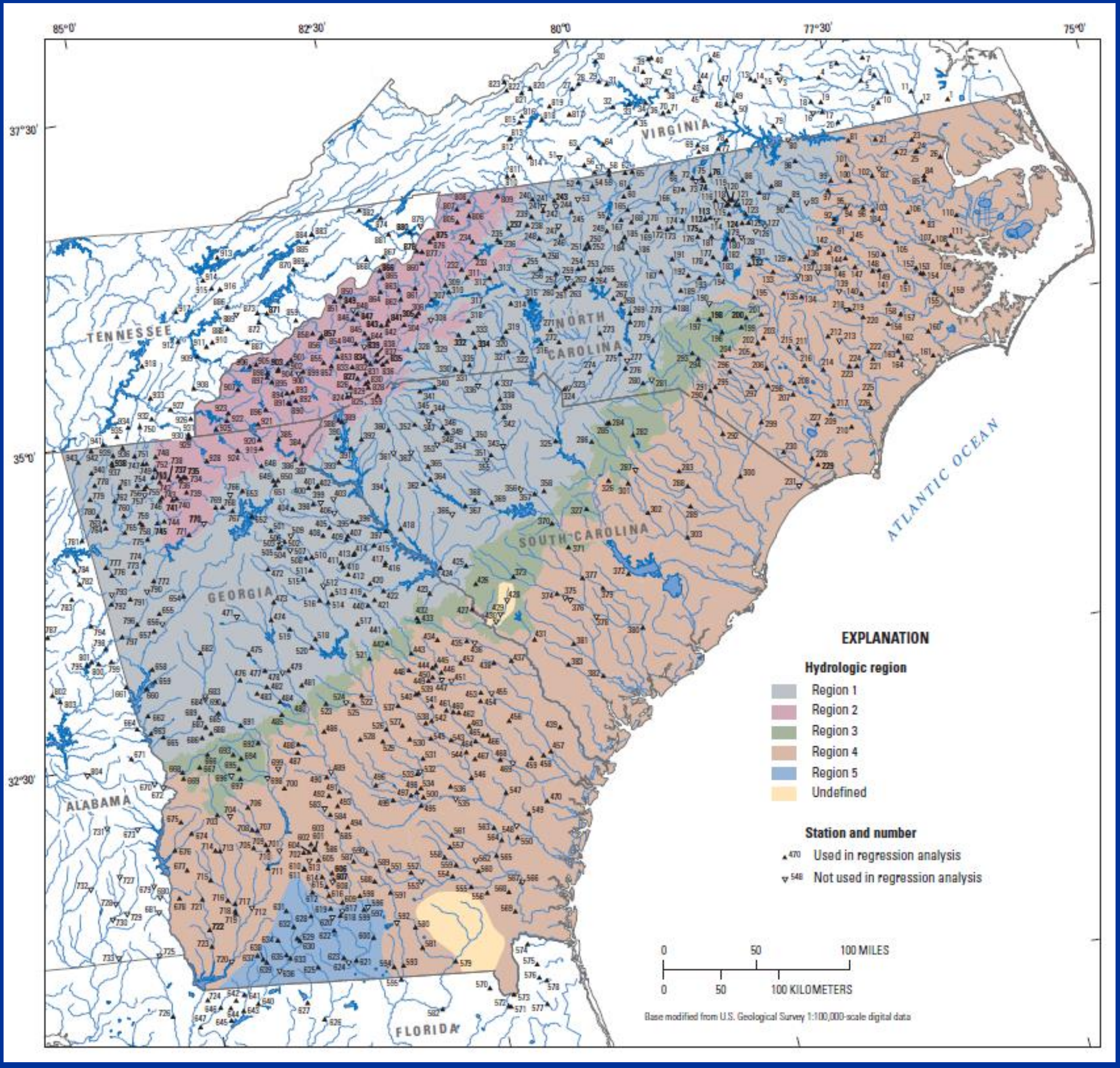
USGS

A Low-Flow Example: 7Q10



USGS

Basics of Regionalization



Murphy’s Law for Hydrology

If a region has a100 gaged streams and one ungaged stream, the most pressing need for immediate streamflow information will be for the stream with no gage…

USGS

Regionalization Analysis

Purpose:

- **Provide estimates of flow characteristics at ungaged sites using information from other gages in the area.**
- **Provide a second estimate of a flow characteristic at gaged sites to develop a weighted estimate (especially if only short records are available at the gage).**

USGS

Rural Flood Frequency: Through Water Year 1999

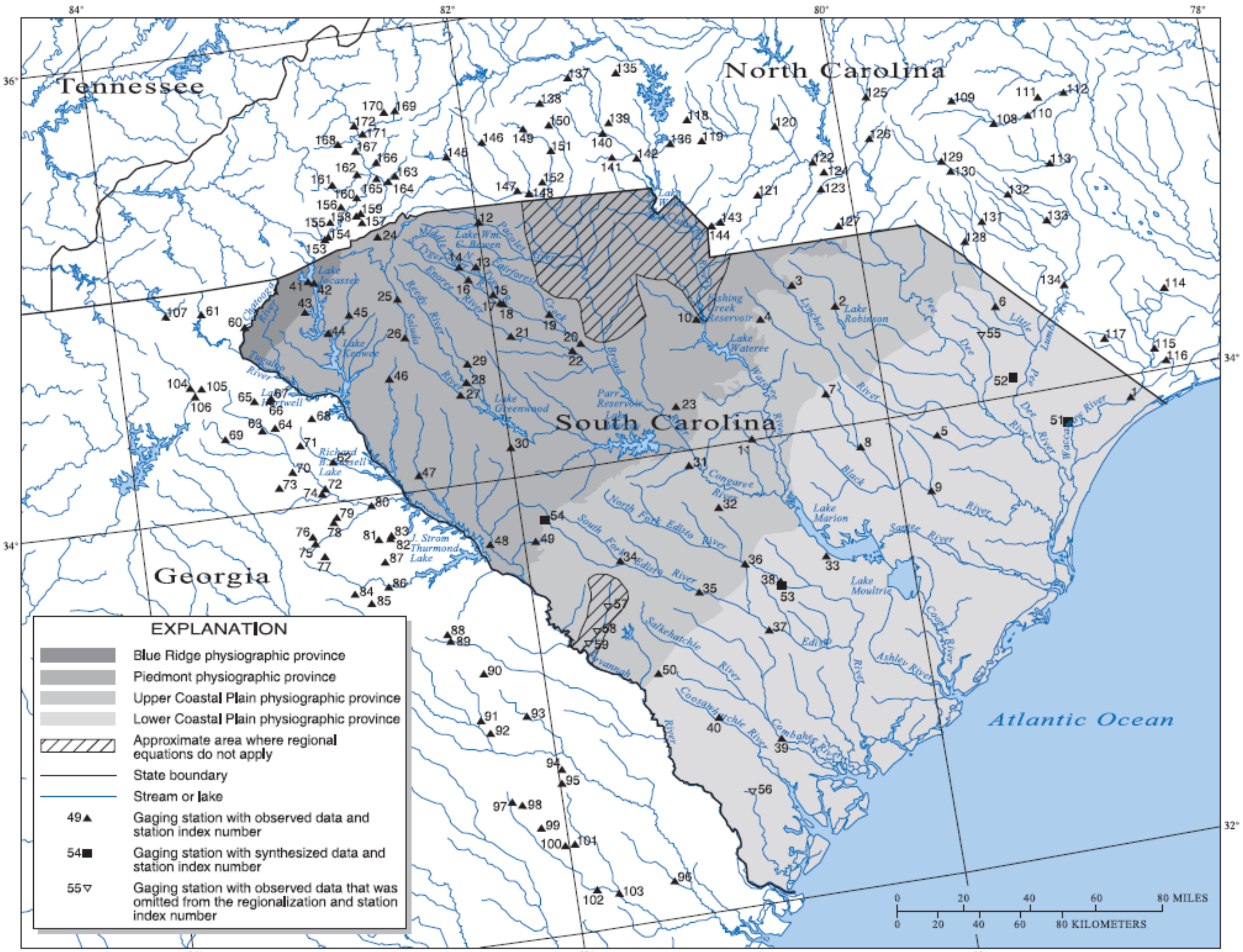


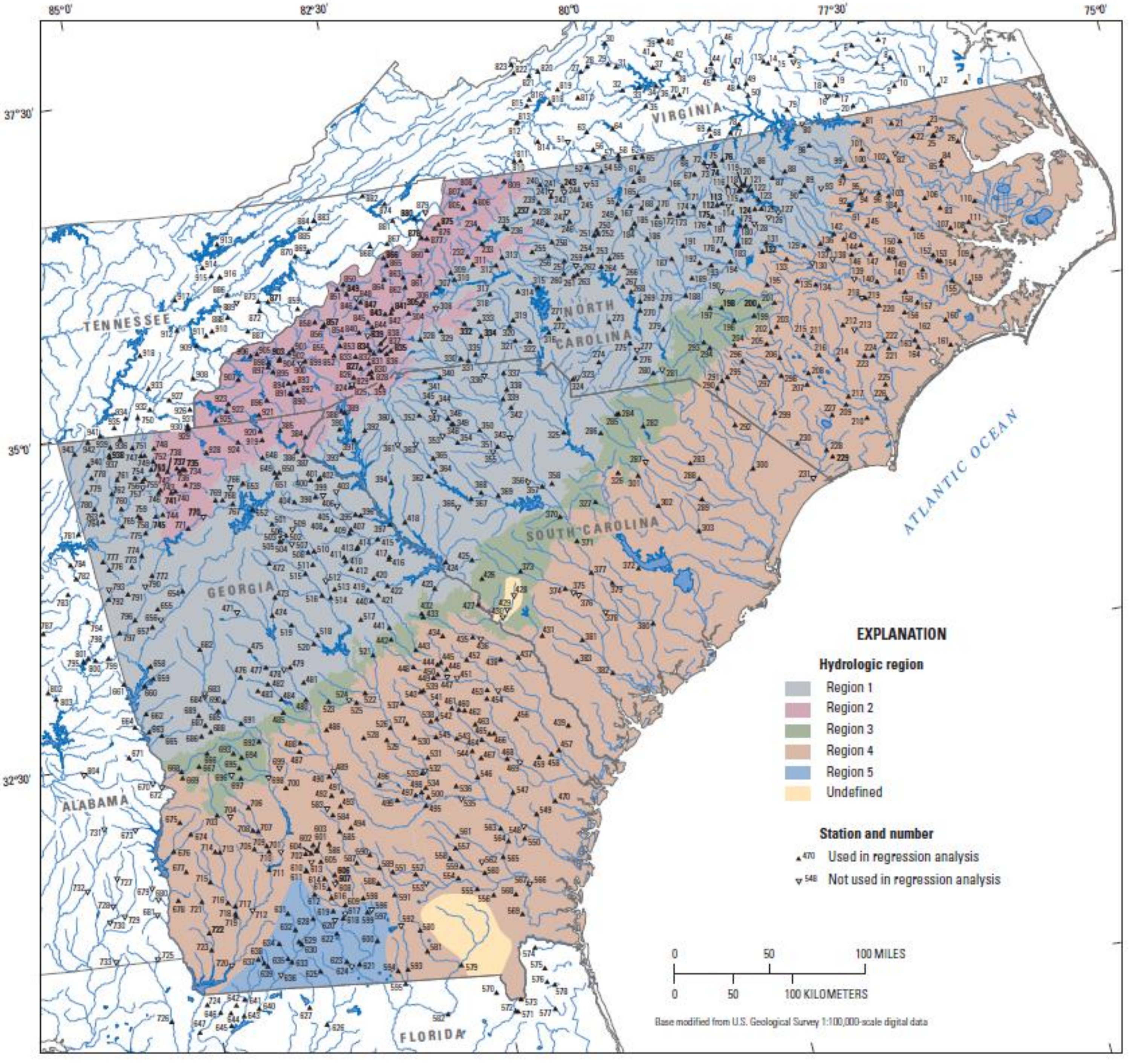
Figure 1. Physiographic provinces and locations of streamflow gaging stations on rural streams in South Carolina and parts of North Carolina and Georgia.

USGS SIR 2009-5156: Peakflows through water year 2006

GA = 310

NC = 303

SC = 64

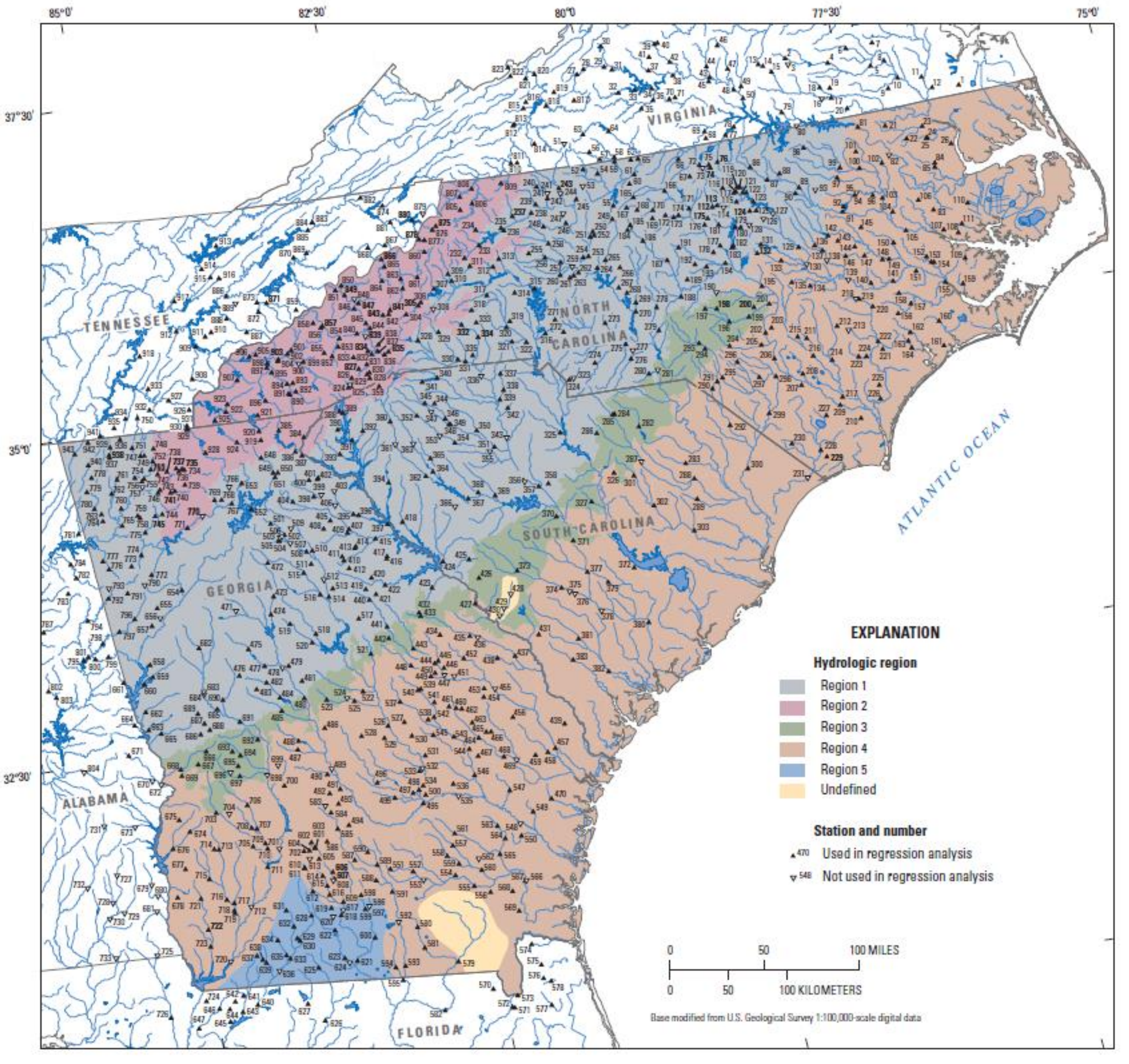


Total area divided by number of gages: (mi²/gage)

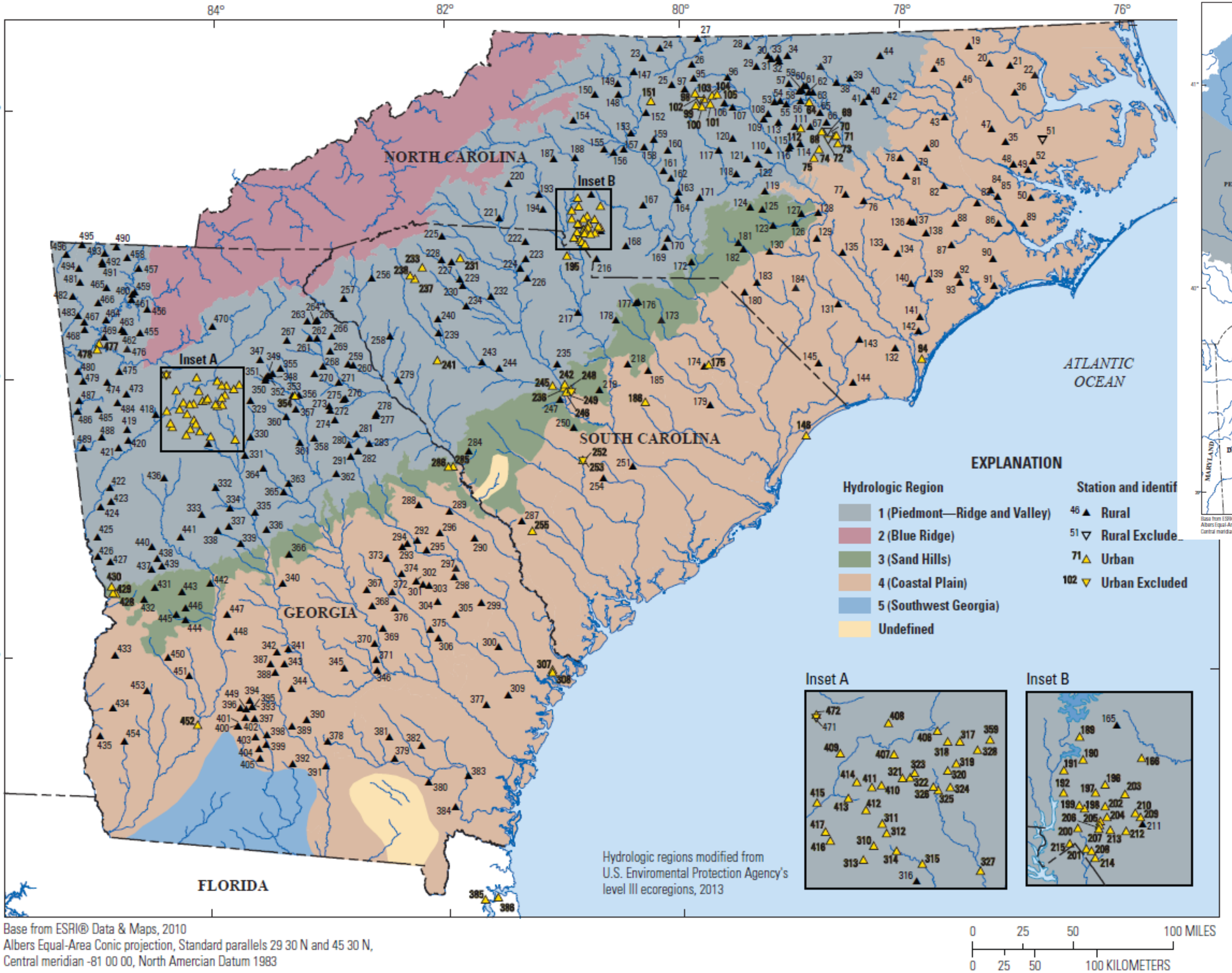
GA = 192

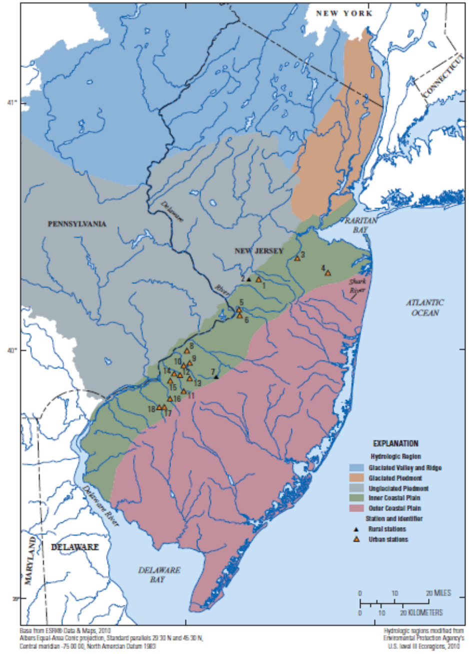
NC = 178

SC = 500



USGS





Urban and small, rural flood frequency

Toby Feaster
Hydrologist
405 College Avenue, Suite 200
Clemson, South Carolina 29631
tfeaster@usgs.gov
864-656-6747