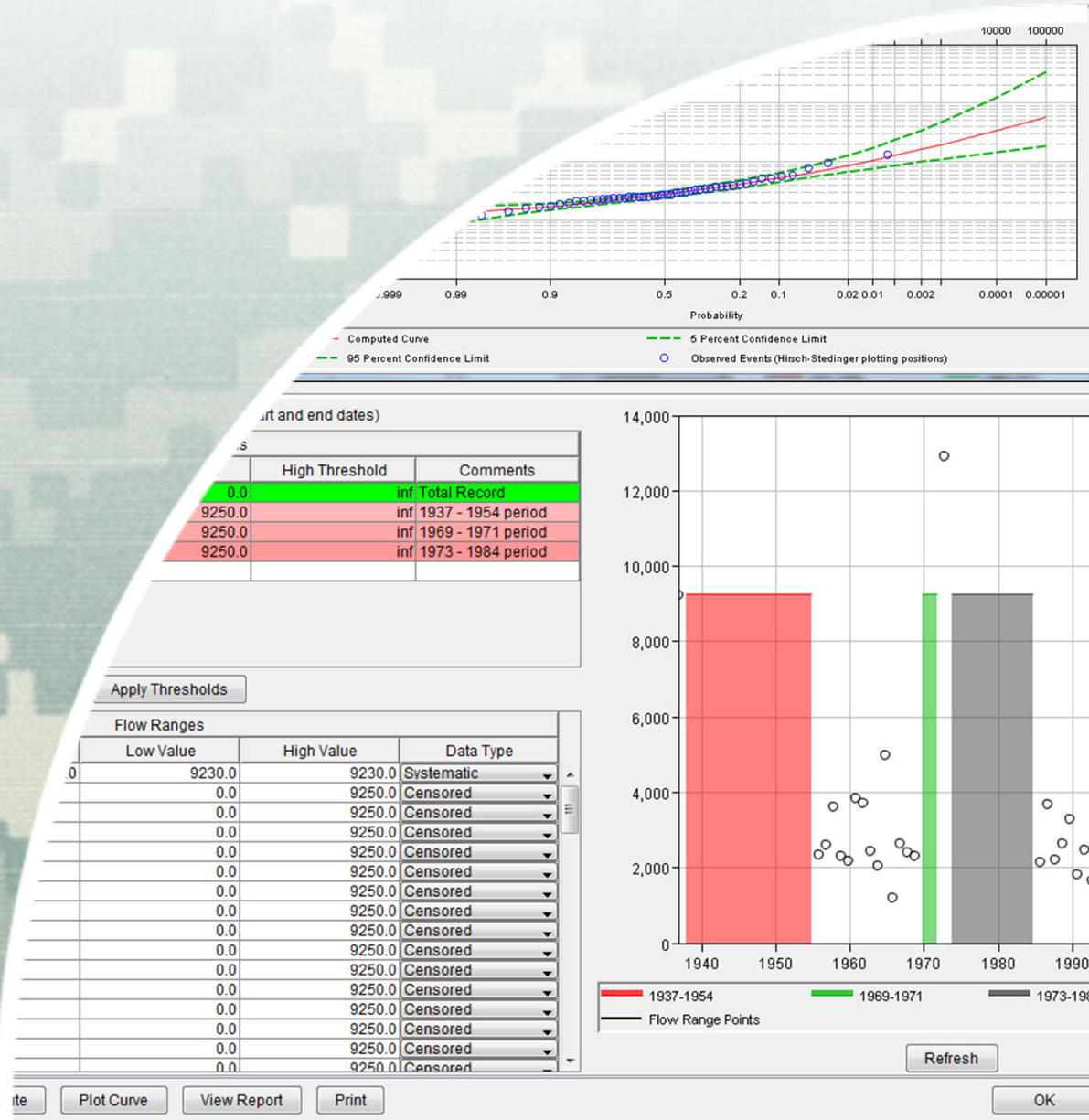
Statistical Software Package (HEC-SSP)

Mike Bartles, P.E.

Hydrologic Engineering Center

May 2022



US Army Corps of Engineers

BUILDING STRONG®

Overview

- Provide a brief history of HEC-SSP.
- Demonstration of HEC-SSP.
  - ▶ Import, inspect, and manipulate data
  - ▶ Create, compute, and visualize results of various analyses
- Detail DSS usage and conventions within HEC-SSP.

History, Status, Future…

- HEC-FFA, STATS, and REGFRQ developed by HEC in response to Corps statistical needs in 1970's
- Evolved with addition of new capabilities and platform support
- In late 1980's, HEC-FFA, STATS, and REGFRQ reconfigured for PC and UNIX

History, Status, Future…

- HEC-SSP started development in FY2005
  - Gary Brunner, Beth Faber, Jeff Harris, and Matt Fleming
- Version 1.0 Beta (Released June 2006)
  - Only computation is Bulletin 17B analysis
- Version 1.0 (Released August 2008)
  - Included General Frequency and Volume Frequency analyses
- Version 1.1 (Released April 2009)
- Version 2.0 (Released October 2010)
  - Included Duration, Coincident Frequency, and Curve Combination analyses
- Version 2.1 (Released August 2016)
  - Included B17C/EMA methodology and Balanced Hydrograph analysis
- Version 2.1.1 (Released January 2017)
  - Updated USGS Plugin and recompiled EMA Fortran code
- Version 2.2 (Released June 2019)
  - Updated EMA code, Mixed Population, and Distribution Fitting analyses

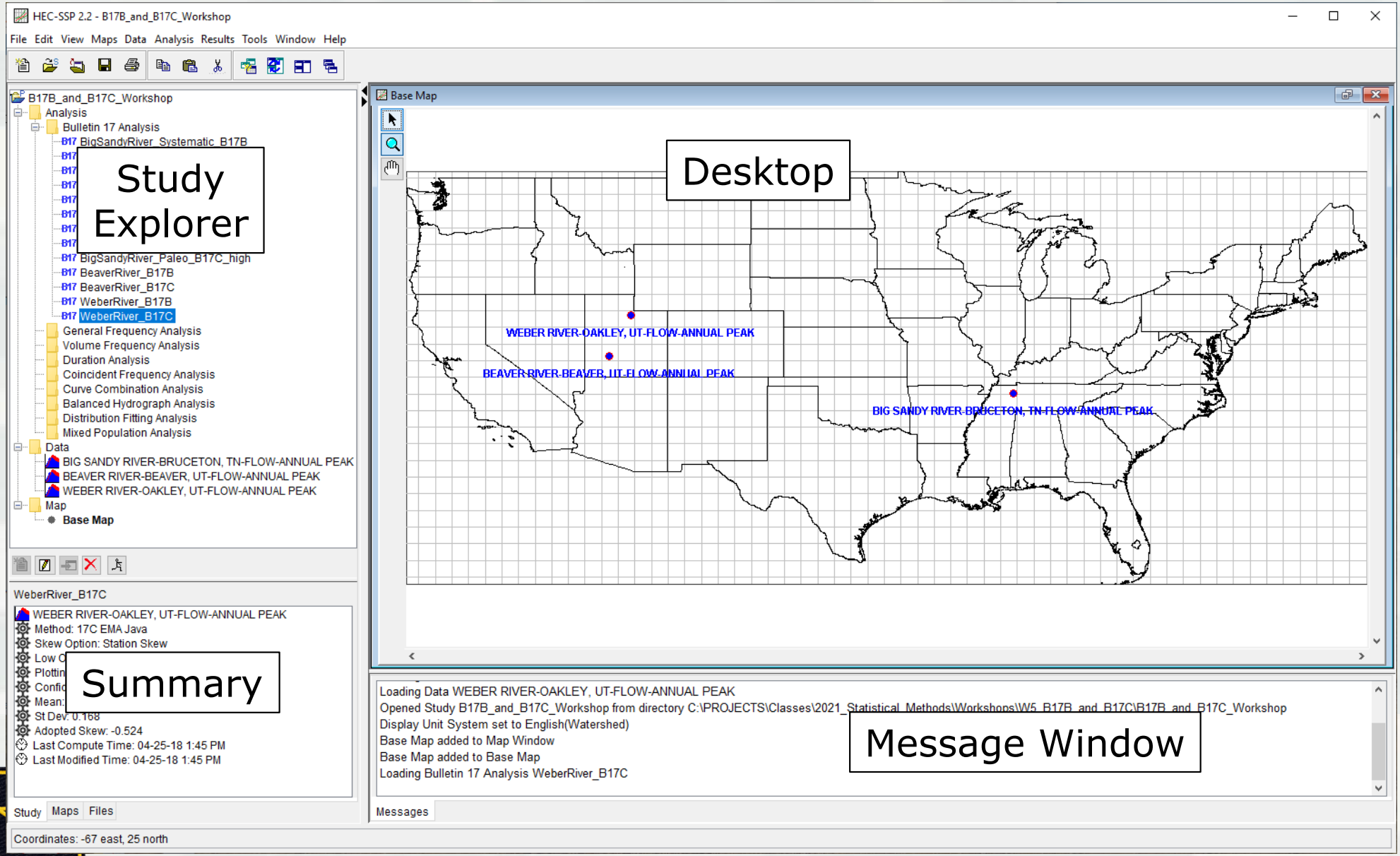
History, Status, Future…

TEAM MEMBERS	
H&S Division Lead	Matt Fleming
Lead Developer / Project Lead	Mike Bartles
Development and Application	Beth Faber
Development and Application	Greg Karlovits
Development and Application	Will Lehman
Development and Application	Haden Smith
Development and Application	John England
Development	Mark Ackerman
Development	Paul Ely
Development	Caleb DeChant
Development	Stephen Ackerman

History, Status, Future…

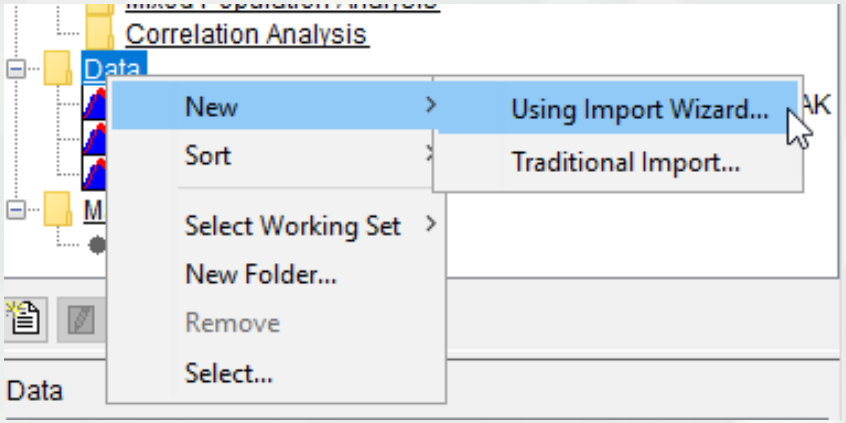
- New analytical tools to meet Corps needs
  - Updated Distribution Fitting Analysis
  - Updated Bulletin 17 Analysis
  - New Correlation Analysis
  - New Record Extension Analysis
- Improved user experience
  - Easier data input

HEC-SSP Software

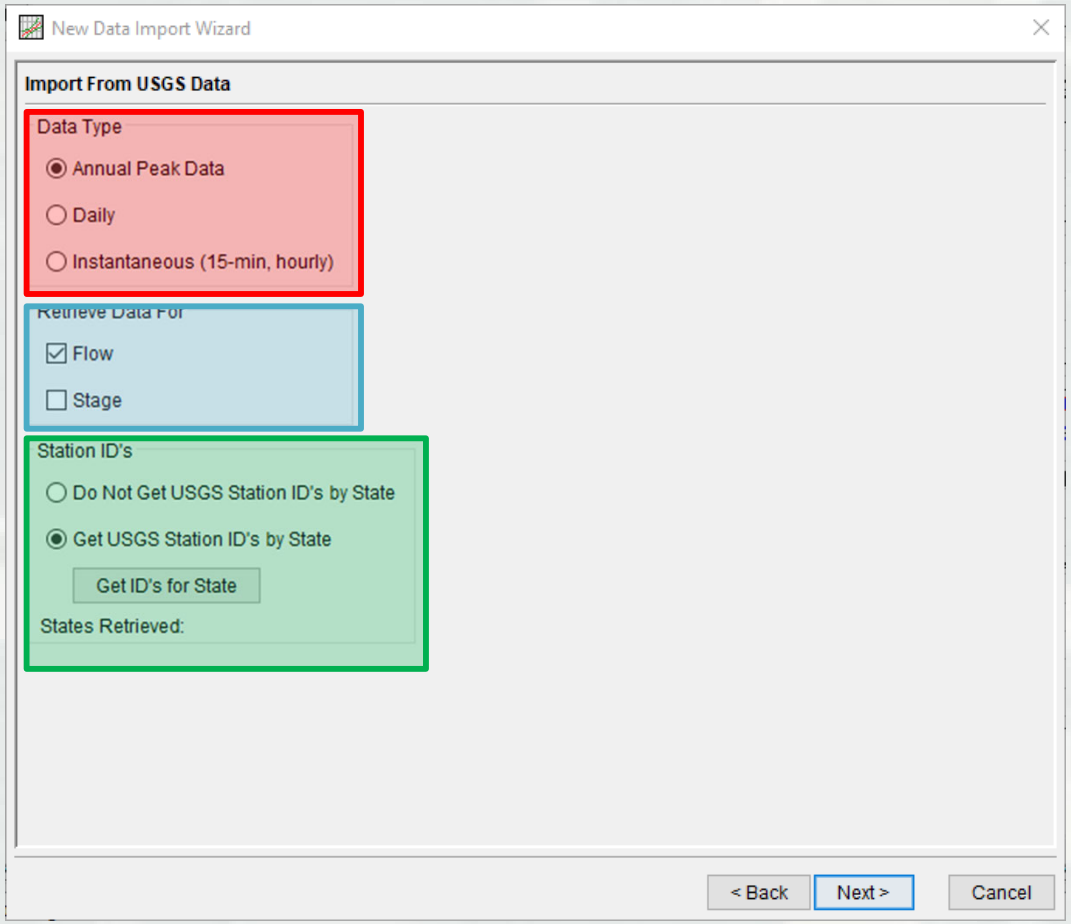


Data Importer

- New Import Wizard or Traditional Import
- Import Time Series and Paired Data
- DSS, USGS website, manual entry, Excel, and text files

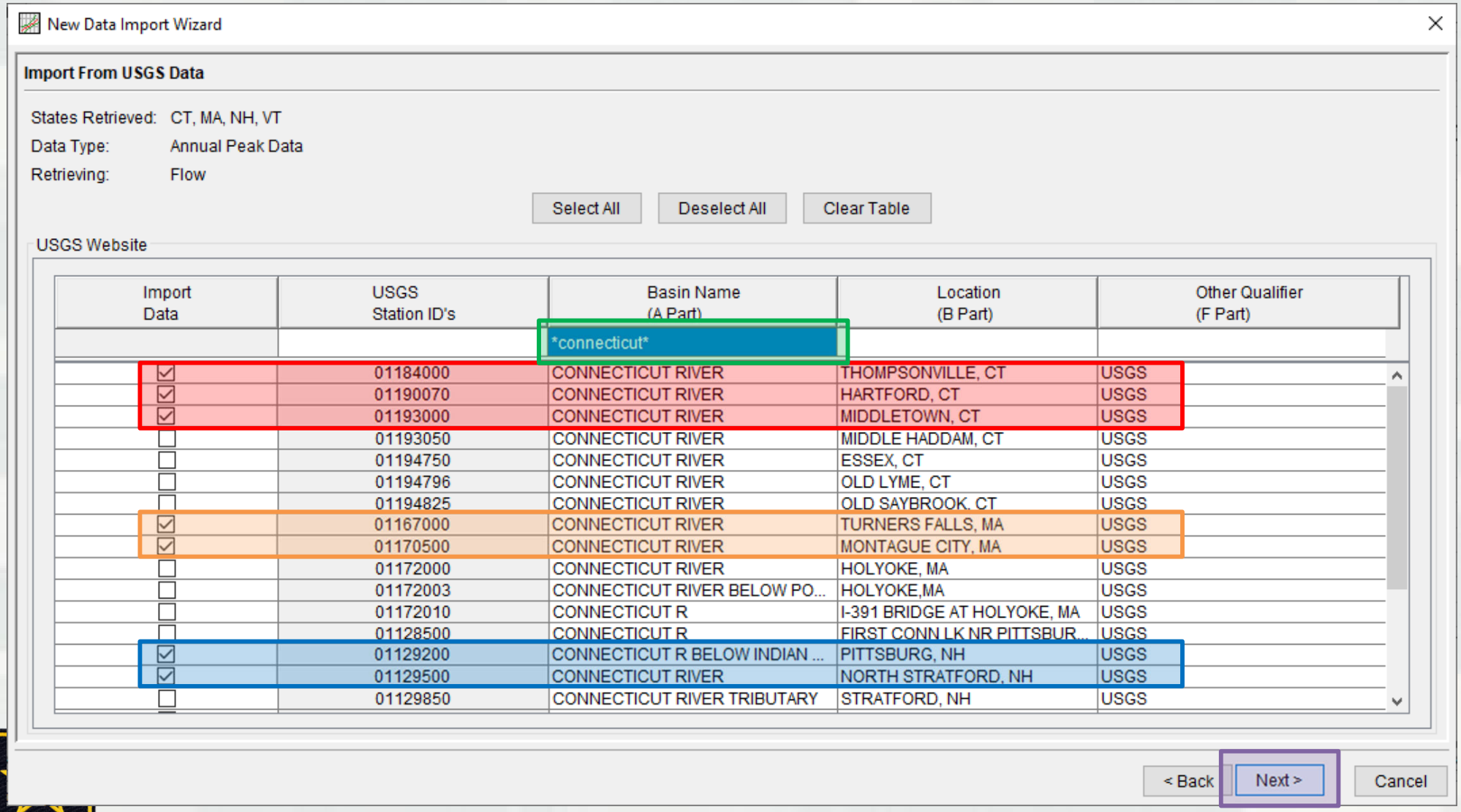


Data Importer
USGS Website



Data Importer

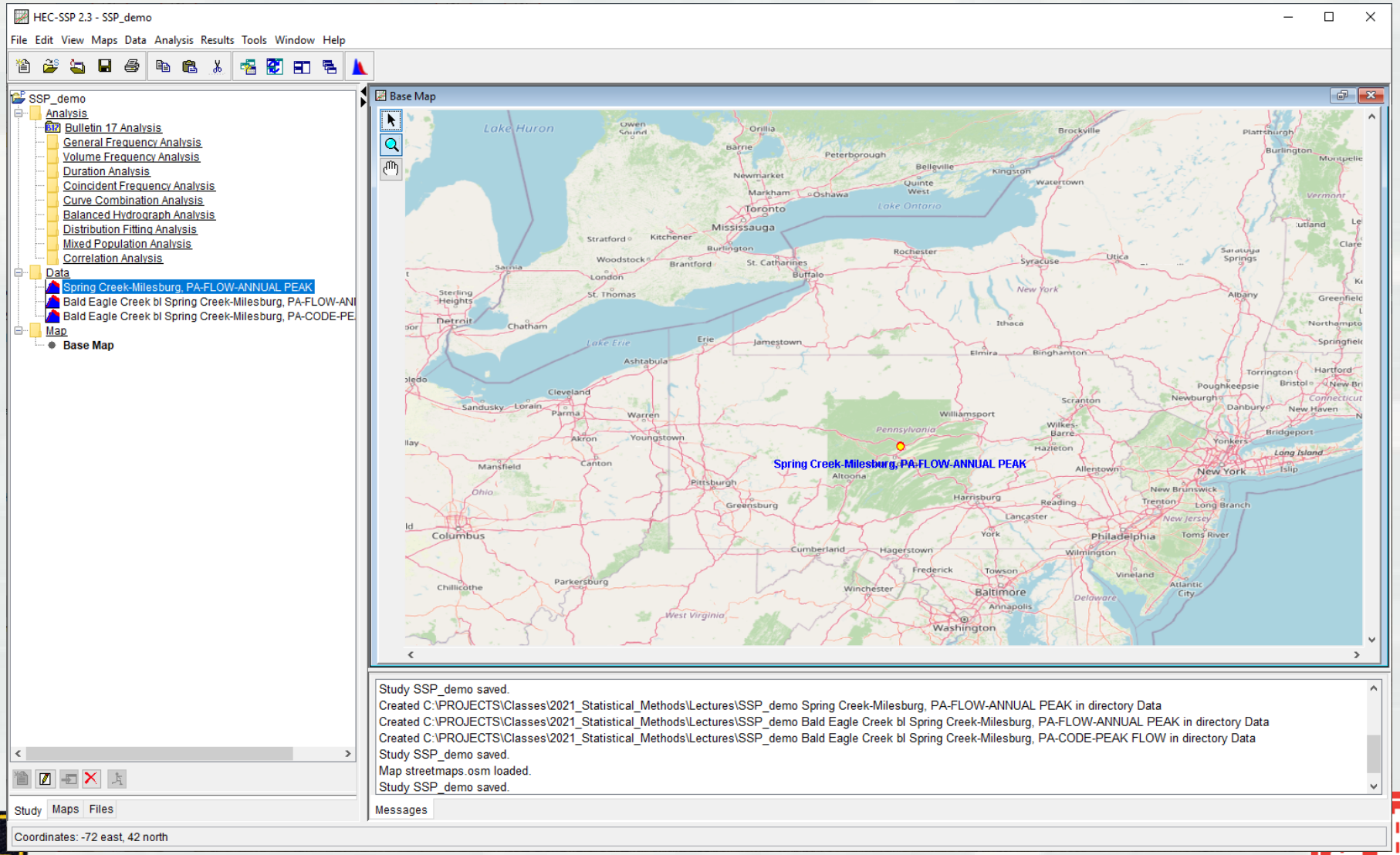
Multiple State Searches



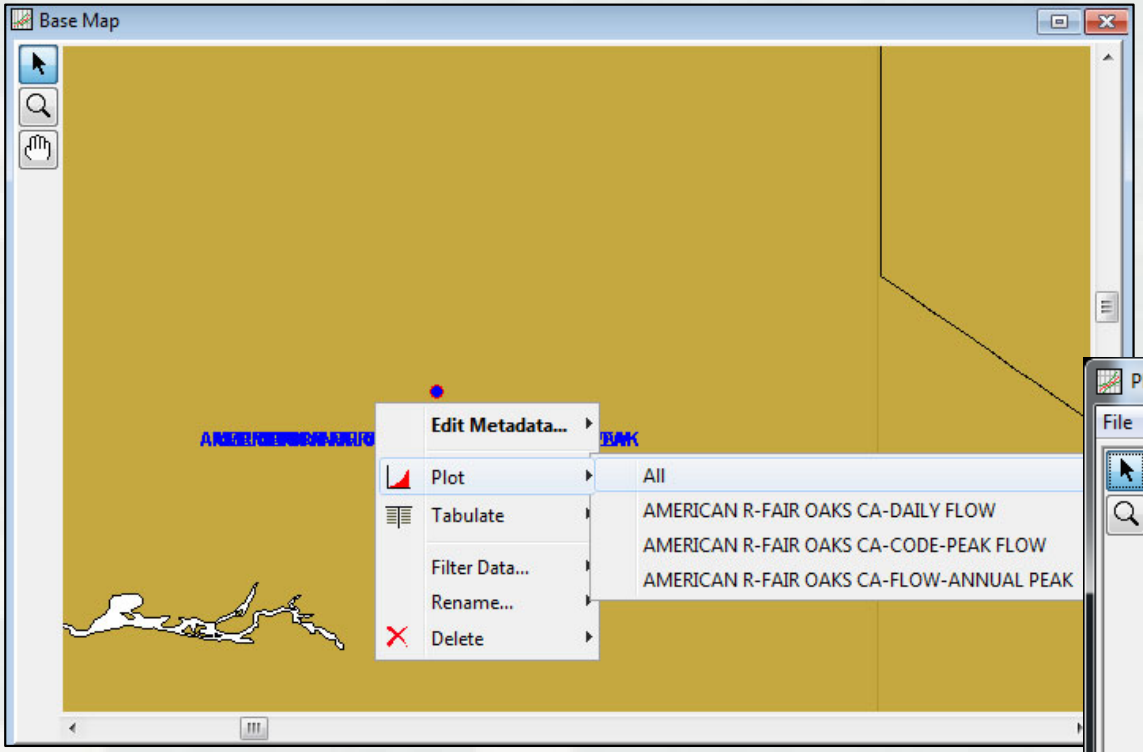
Background Maps

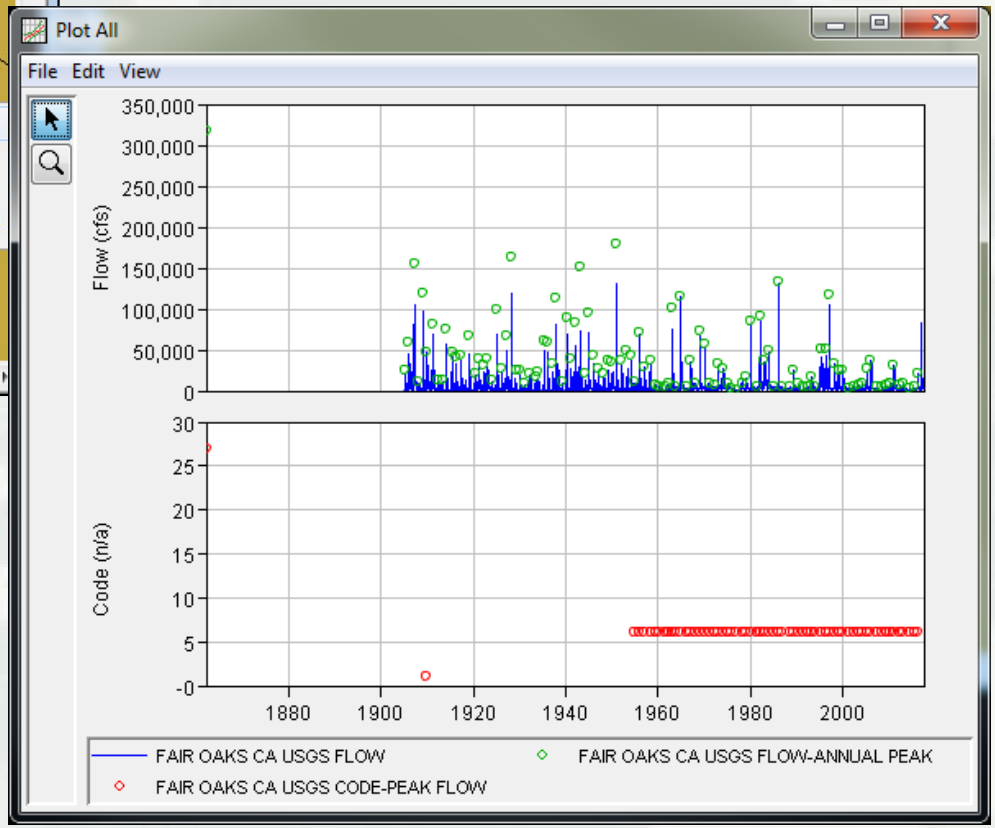
- Background Maps are Optional
- Types of Map Layers:
  - Internet Maps (Google, Bing, OSM), Shapefiles, rasters, Google Earth .kml, etc
- Gage Locations displayed on top
- Map is interactive for Editing Data and Viewing Results

Example Background Map

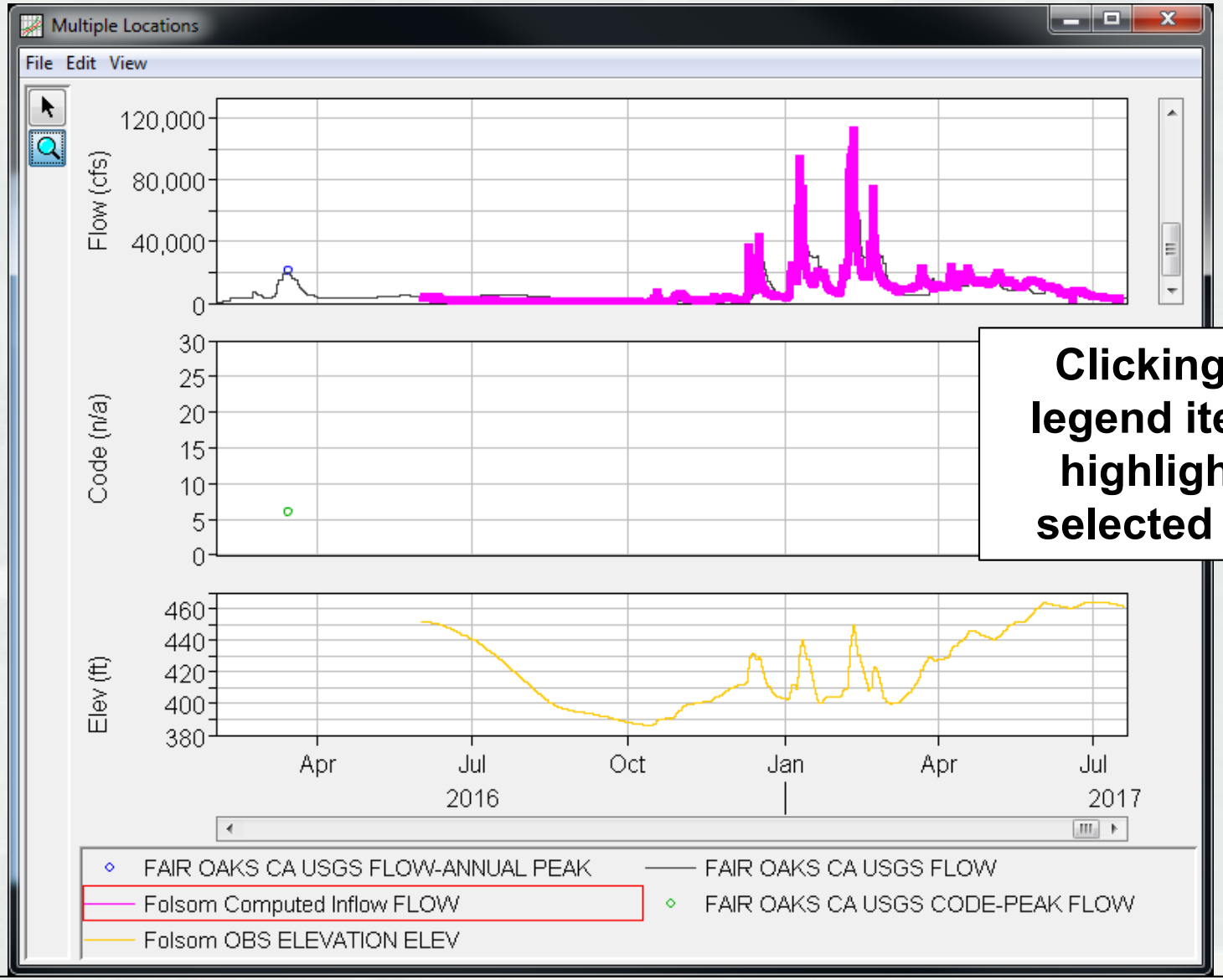


Data Visualization



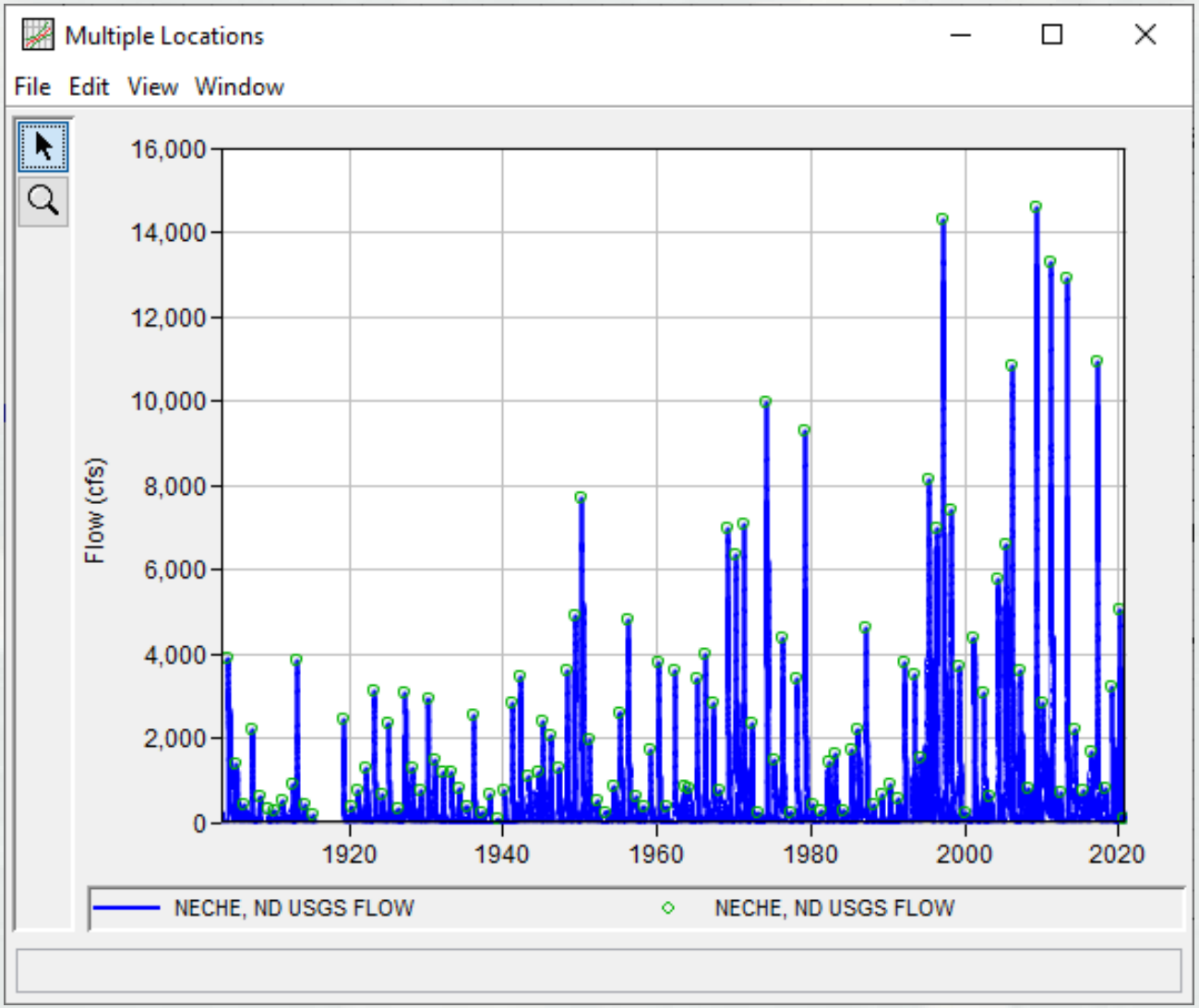


Data Visualization

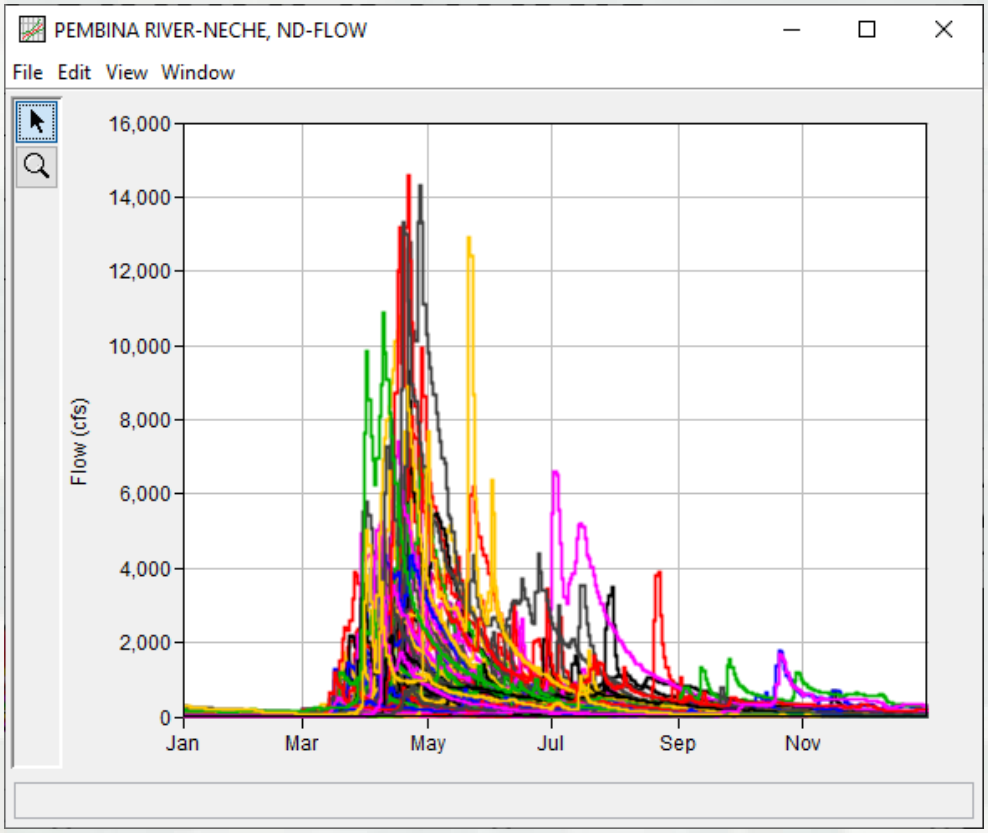


Clicking on a **legend item** will **highlight** the **selected curve**

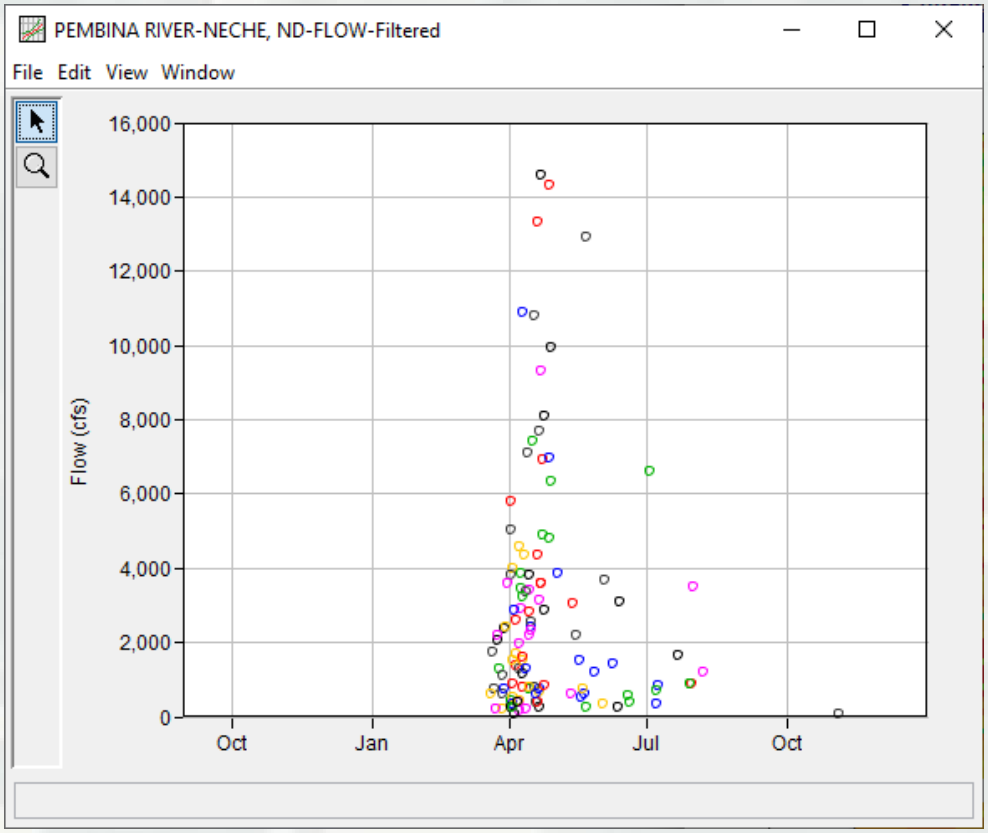
Data Visualization



Data Visualization



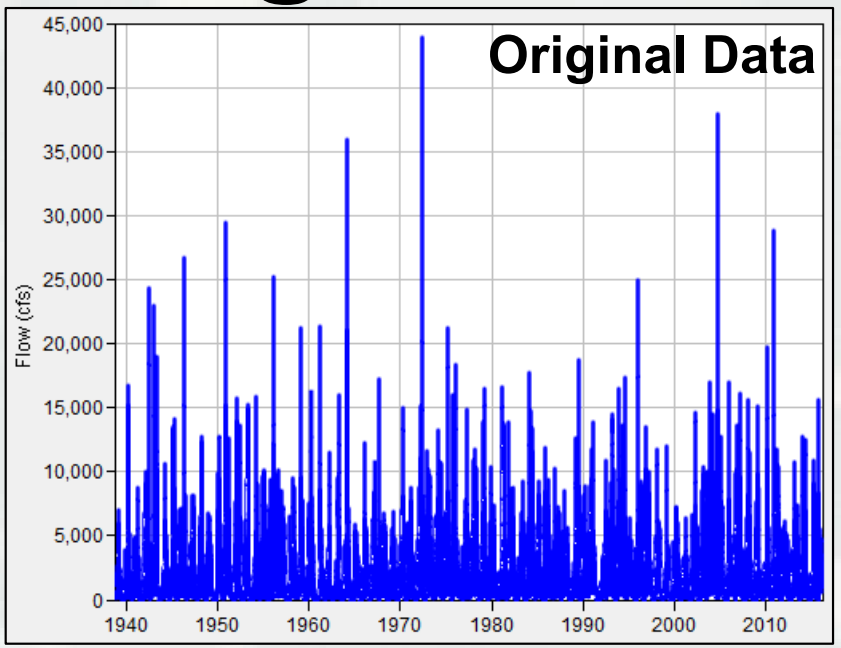
Daily Average Flow

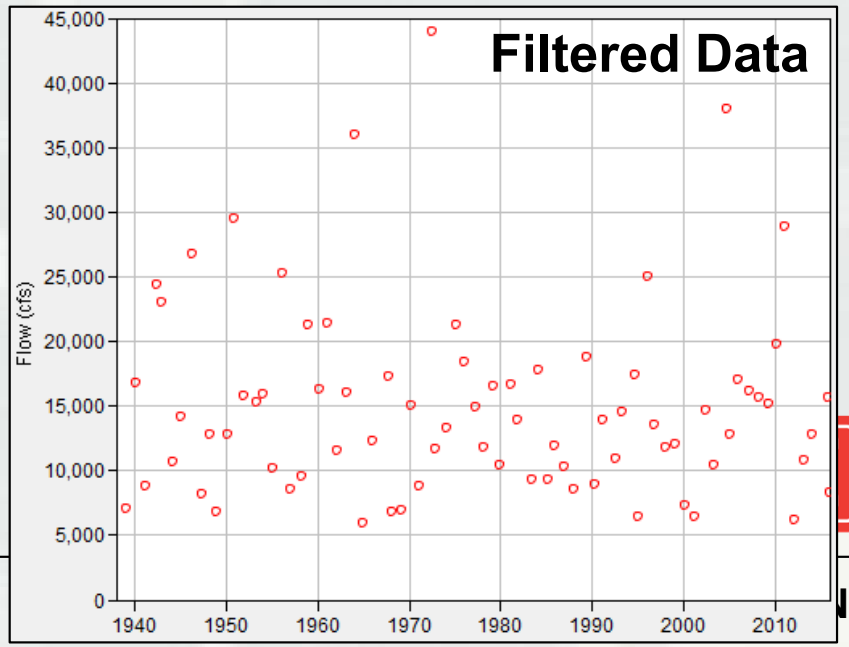


Annual Maximum Daily Average Flow

Data Filtering

- Filter data using:
  - Time Window
  - Season
  - Min/Max Threshold
  - Duration
  - Annual Maxima
  - Peaks Over Threshold
  - Starting Pool Stage/Elev





Data Storage System (DSS)

- Data is stored within the file in "*blocks*", for example:
  - Time Series (hourly data stored in months)
  - Paired Data (flow vs stage curve w/ single stage axis and multiple flow axes)
  - Gridded (single radar scan)
- Multiple blocks may make up a single "*data set*", e.g., 50 years of hourly data is one data set
- Each block is called a "*record*"
- A HEC-DSS file can have many records
- Name of a record is called a "*pathname*"
- Each pathname within a file must be unique

DSS

Time Series Data | Pathnames

- Pathname self-documents the data
- Consists of 6 parts, separated by forward slashes “/”
- Parts are labeled A – F: “/A/B/C/D/E/F/”
- Each part can be 0 to 64 characters long
- A single pathname can be up to 391 characters long
- Example:
  - /SACRAMENTO/RED BLUFF/FLOW/01MAR1972/1HOUR/OBS/

DSS

Time Series Data | Pathnames

/A/B/C/D/E/F/

Part	Description
A	Group, basin, river, region or study name
B	Location or gage name
C	Data parameter
D	Starting date for block (not 1st data)
E	Time interval (standard)
F	Version or additional information

/SACRAMENTO/RED BLUFF/FLOW/01MAR1972/1HOUR/OBS/

DSS

Conventions

- Use optional part names
- Be descriptive, but not “overly” descriptive

- Please **do not** do this:

“///FLOW/01JUN1972/1HOUR//” (i.e. no A-, B-, or F-parts)

- Instead, **do** this:

“/BALD EAGLE CREEK/SAYERS/FLOW/01JUN1972/1HOUR/COMPUTED/”

DSS

Time Series Data | Interval

- Each record contains a "*header*"
  - Data Units (e.g., FEET, CFS)
  - Data Type:
    - PER-AVER Period Average (daily average flows)
    - INST-VAL Instantaneous (15-min flows)
    - PER-CUM Period Cumulative (daily precip accumulation)
    - INST-CUM Instantaneous Cumulative (incremental precip)
  - Time offset (e.g., daily data read at 8:00 am)
- Missing data flags (-901.) are used as a place holder

DSS

Time Series Data | Regular

- Blocks are “standard size” (there are always 365 or 366 values for one year of daily data)

Interval	Block Length
1MIN, 2MIN, 3MIN, 4MIN, 5MIN, 6MIN, 10MIN, 12MIN	One day
15MIN, 20MIN, 30MIN, 1HOUR, 2HOUR, 3HOUR, 4HOUR, 6HOUR, 8HOUR, 12HOUR	One month
1DAY	One year
1WEEK, TRI-MONTH, SEMI-MONTH, 1MON	One decade
1YEAR	One century

DSS

Time Series Data | Irregular

- Same as regular-interval, except:
- Date and time store with each data value (which makes data sets much larger)
- Blocks (E parts) are:
  - IR-DAY
  - IR-MONTH
  - IR-YEAR
  - IR-DECADE
  - **IR-CENTURY**
- Block sizes are (user) variable length. Try to limit sizes between 100 and 1000 values per block

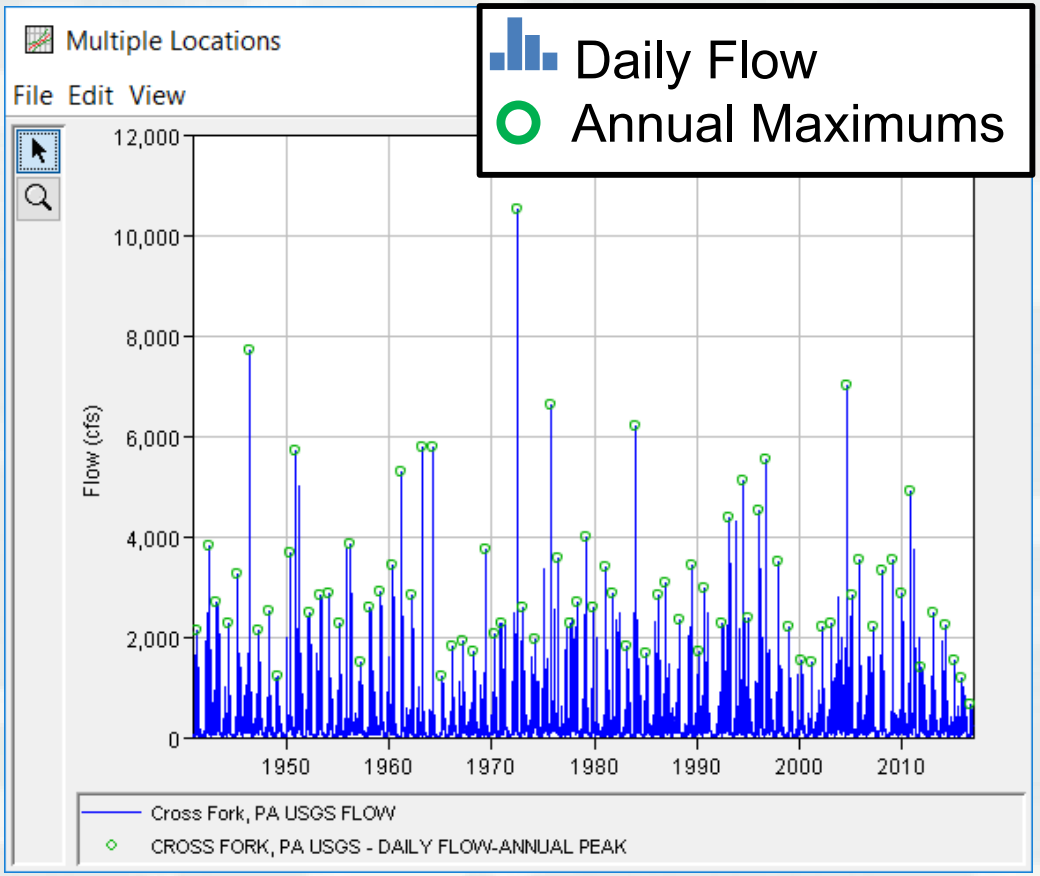
DSS Data within SSP

- **Bulletin 17** (and General Frequency) analyses require the use of **irregular** data sets
  - **Please use IR-CENTURY**
- Regular data sets will not be selectable
  - If you don't see the data set you just entered, it's because it's not irregular

- **Volume Frequency** analyses require the use of **regular** data sets
  - Use 1DAY
- Irregular data sets will not be selectable

Extracting Annual Maximum or Partial Duration Series

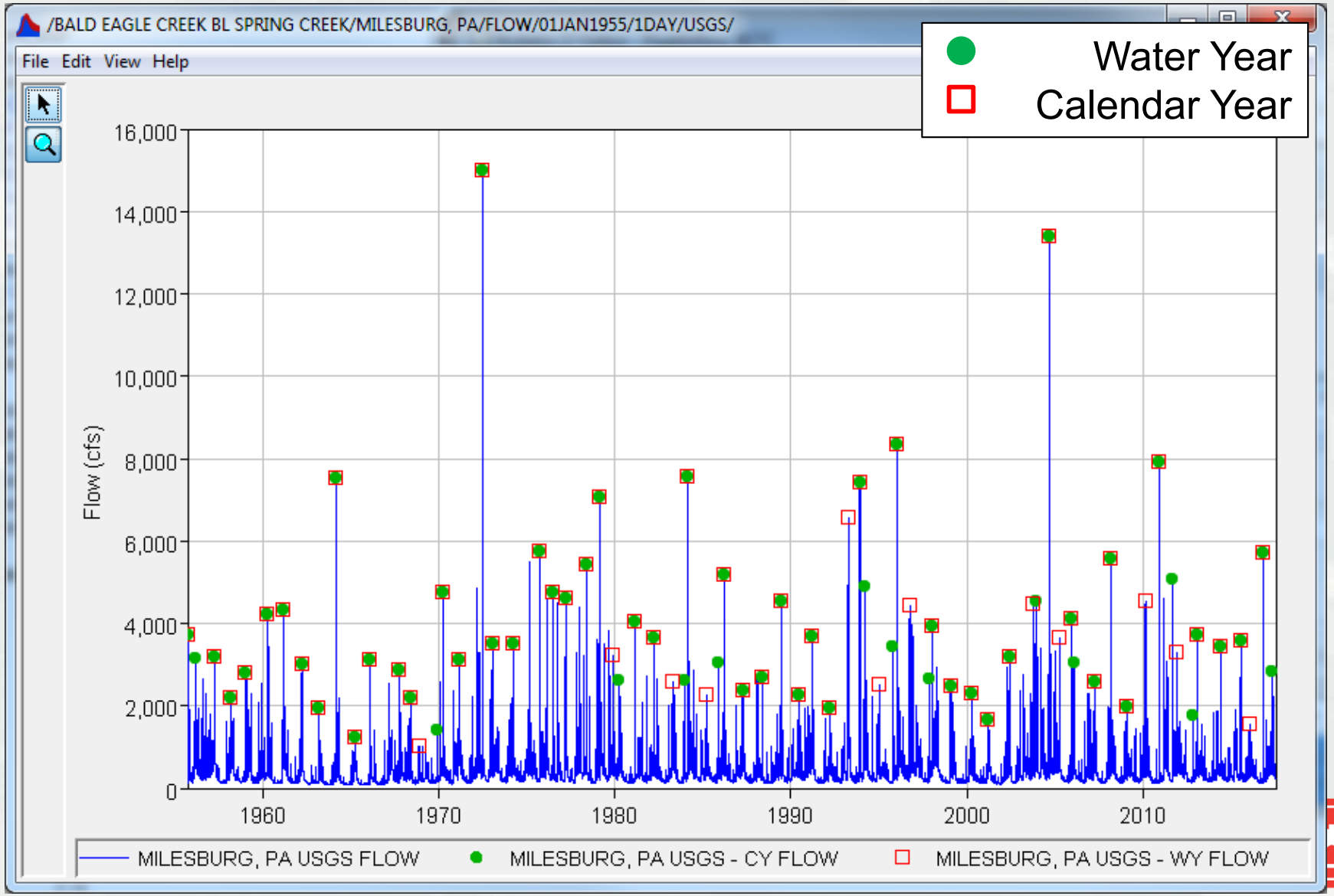
- Download data
- Right-click | Filter Data…
- Select Filter Options
  - Absolute Time Window
  - Seasonal Time Window
  - Min/Max Threshold
  - Filter to Annual Maximums
  - Filter to Partial Duration Series



Calendar Year vs. Water Year

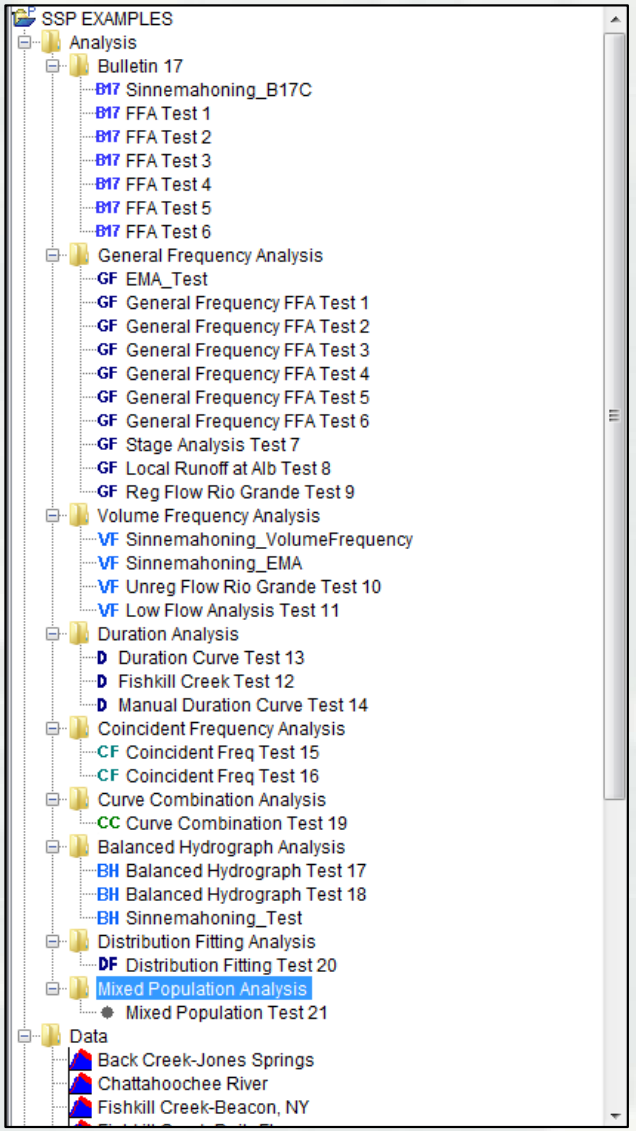
- Within SSP, Bulletin 17 analyses using EMA/B17C require that only one peak be present in any given water year
  - i.e. If the linked DSS data set contains two values in water year 1969 (01Oct1968 – 30Sep1969), **your analysis will not compute**
- If your watershed has more than one peak in a water year that must be included (i.e. partial duration) or calendar year is more appropriate to use, contact HEC for help

Calendar Year vs. Water Year



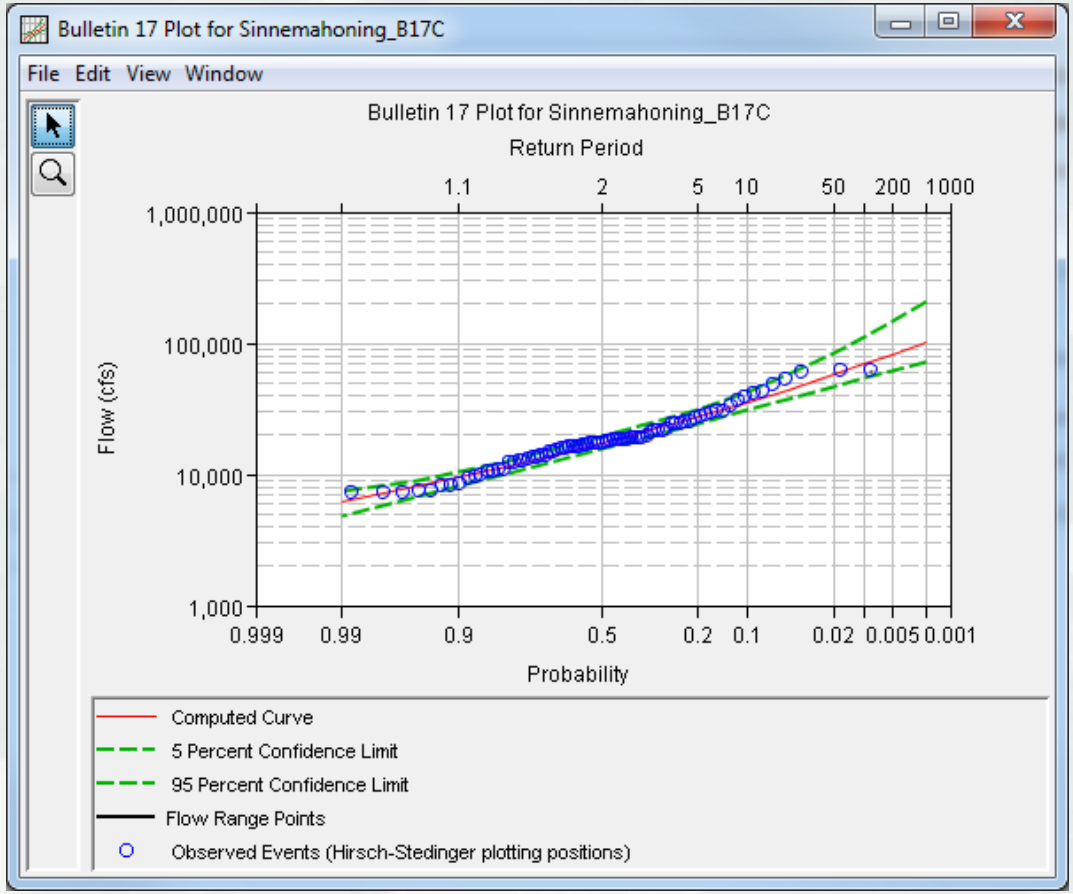
HEC-SSP Analysis Types

- Eleven Analysis Types
  - Bulletin 17
  - General Frequency
  - Volume-Frequency
  - Duration Analysis
  - Coincident Frequency
  - Curve Combination
  - Balanced Hydrograph
  - Distribution Fitting
  - Mixed Population
  - Correlation
  - Record Extension

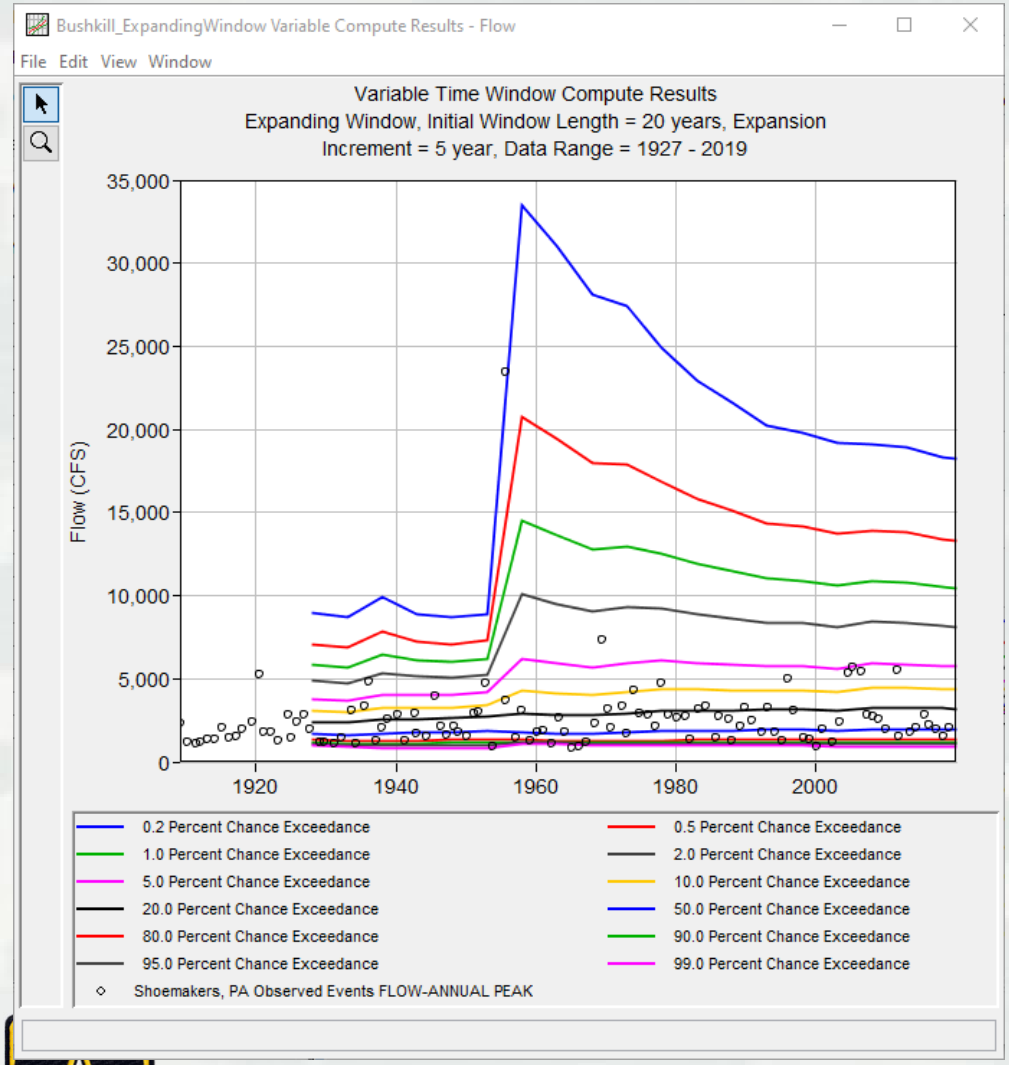


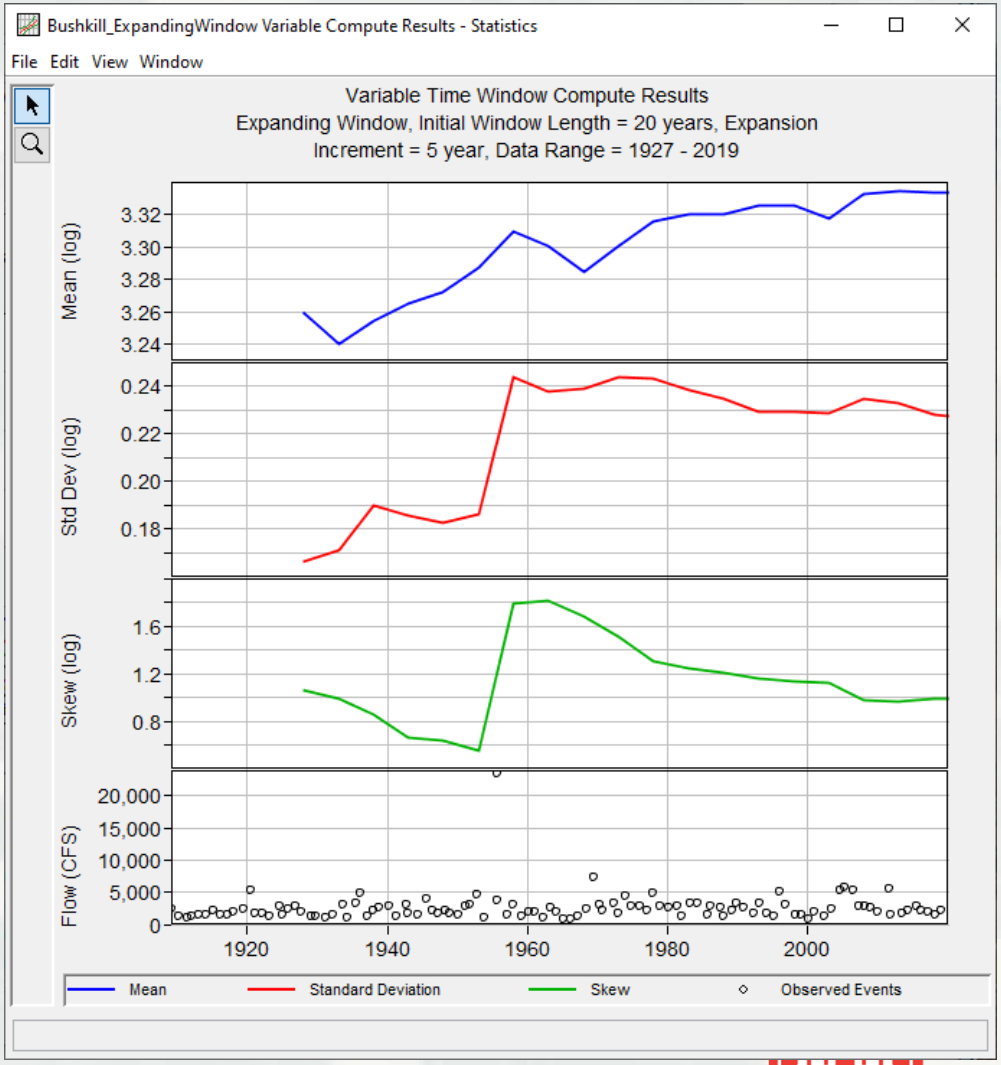
Bulletin 17 Analysis

- "**Strict**" flow-frequency analysis using either **Bulletin 17B** or **Bulletin 17C** procedures
- Can evaluate moving or expanding time windows
- **IRREGULAR** data required
  - ► i.e. IR-CENTURY



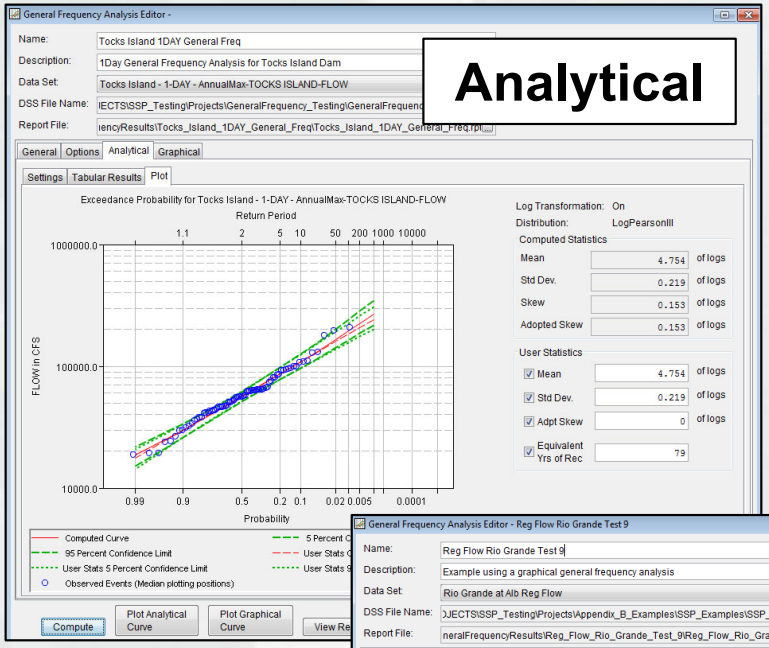
Bulletin 17 Analysis

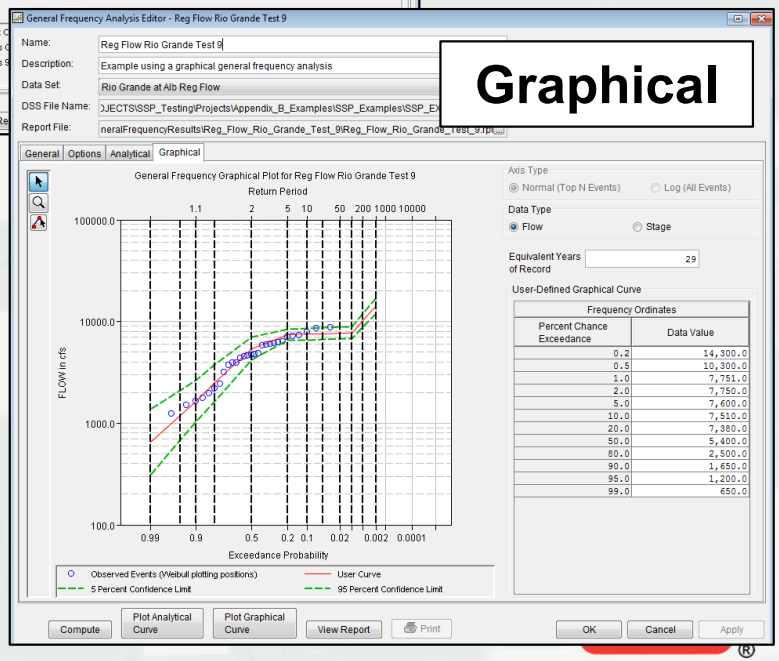




General Frequency Analysis

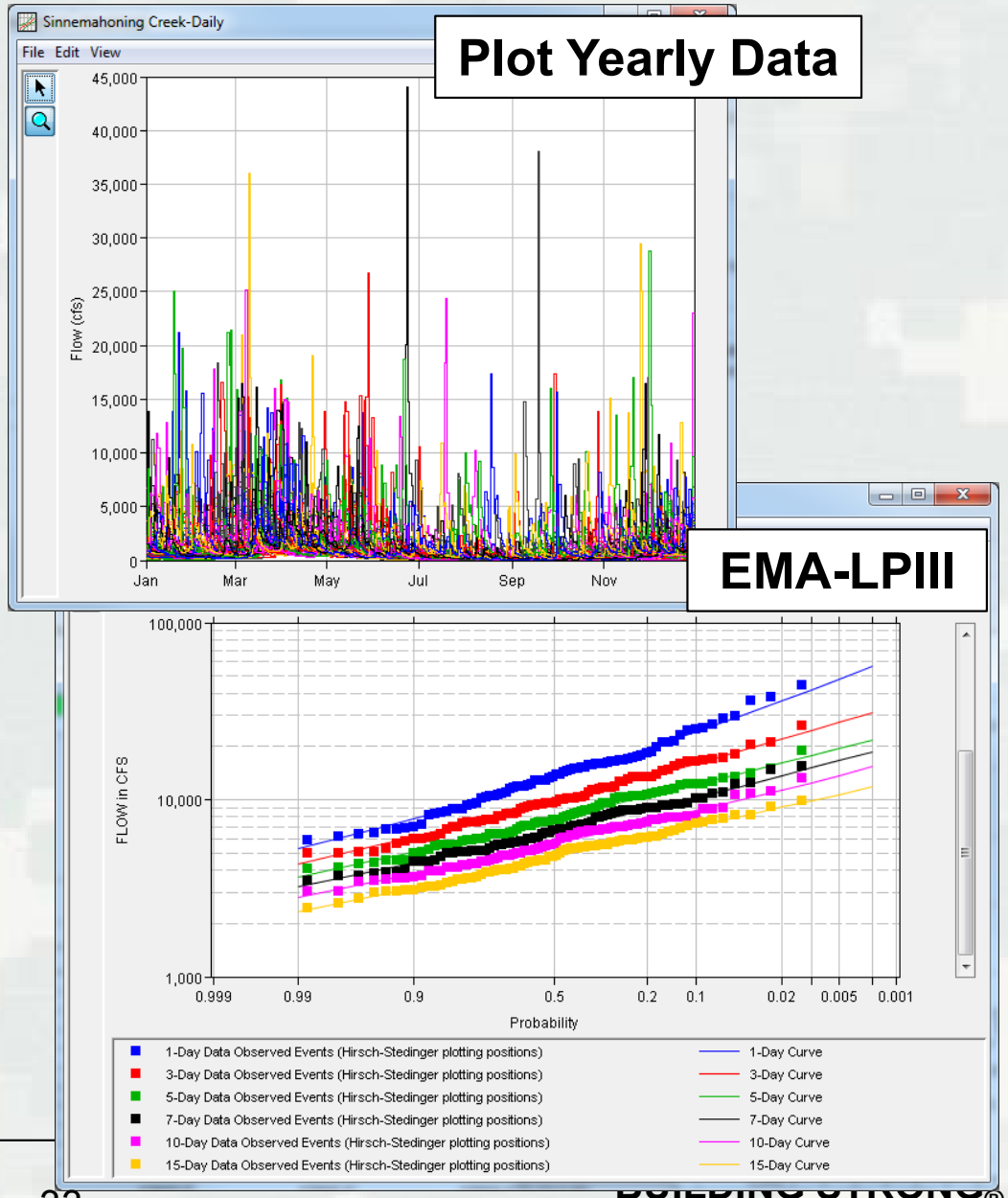
- "**Less strict**" flow-, stage-, precipitation-, etc frequency analysis
  - Mix and match procedures
- **Numerous** analytical distributions
  - Product Moments-LPIII
  - EMA-LPIII
  - Linear Moments-GEV
  - etc
- Manually define distribution parameters
- Graphical/Empirical distribution
- Annual or Partial Duration series
- **IRREGULAR** data required
  - i.e. IR-CENTURY





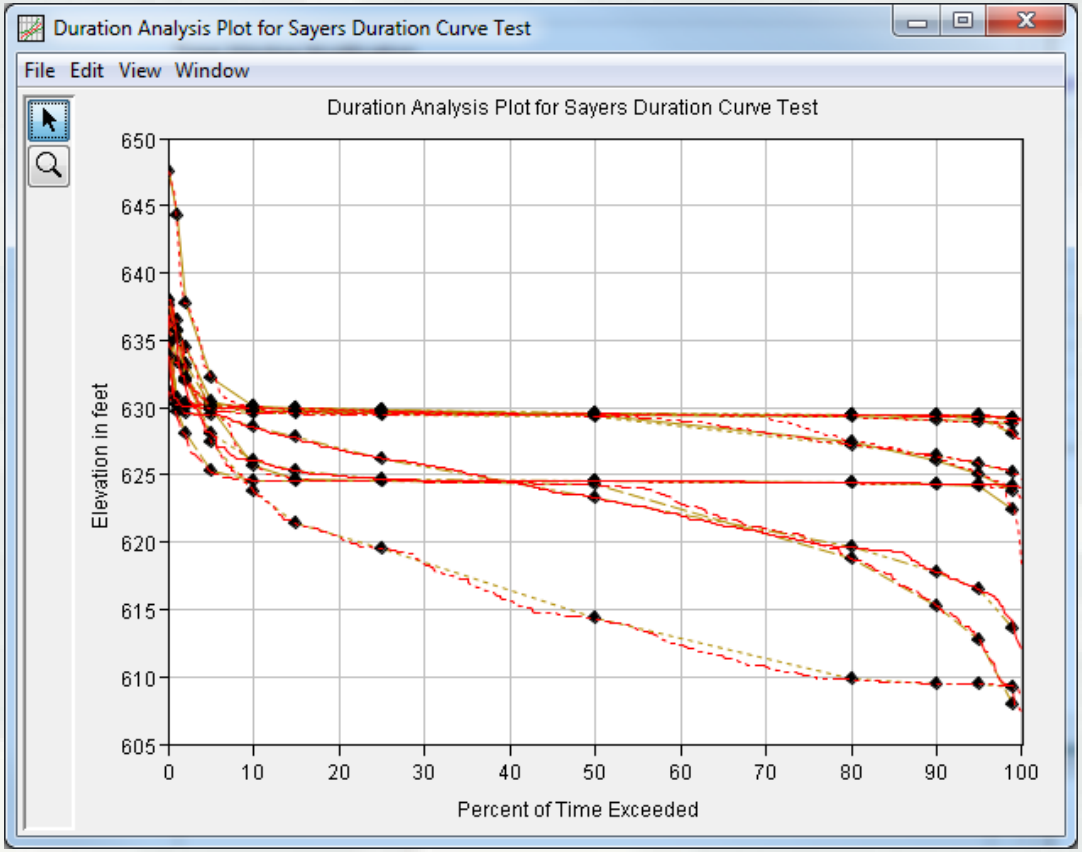
Volume Frequency Analysis

- Iterative/duplicative frequency analysis
  - Mix and match procedures
- Extract annual maximum series from input data and fit distribution
- Numerous analytical distributions
  - Product Moments-Normal
  - Product Moments-LPIII
  - EMA-LPIII
  - etc
- Manually define distribution parameters (i.e. smooth statistics)
- Graphical/Empirical distribution
- **REGULAR** data required
  - i.e. 1DAY



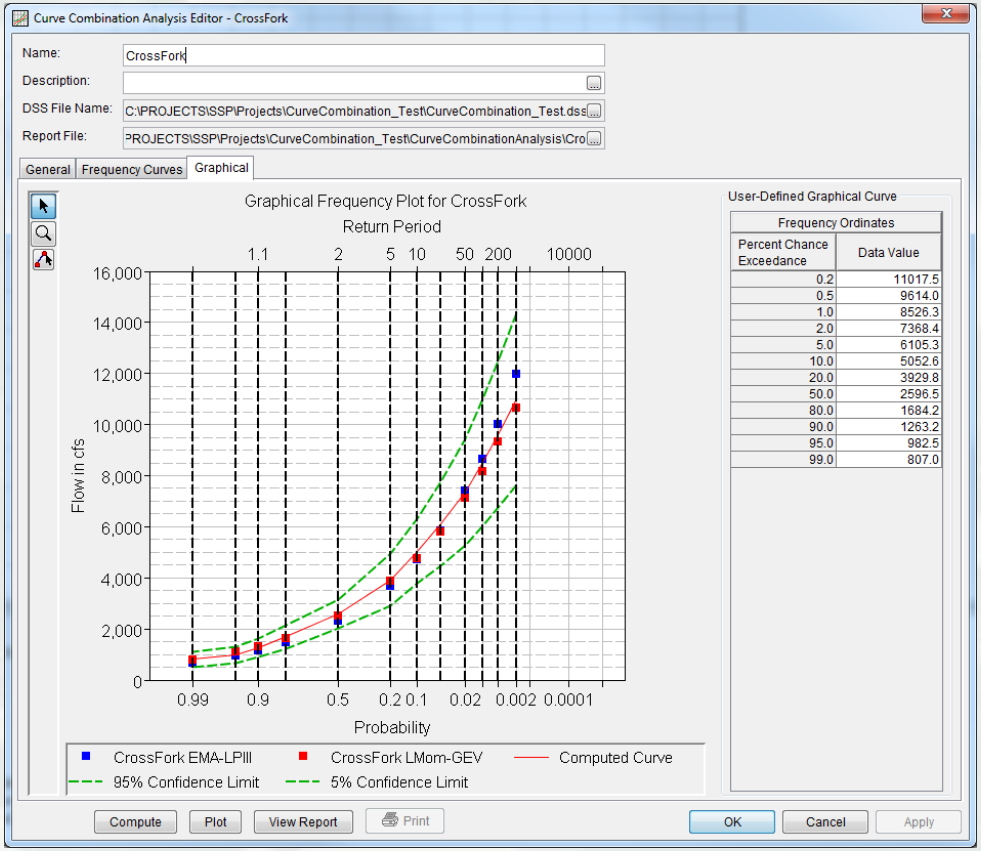
Duration Analysis

- Computes Stage- or Flow-Duration
  - i.e. percent of time stage/flow was in excess of a certain value
- Rank/Sort and STATS (i.e. bin) methods
- Annual, Quarterly, Monthly, or User-Defined Periods
- **REGULAR** data required
  - i.e. 1DAY



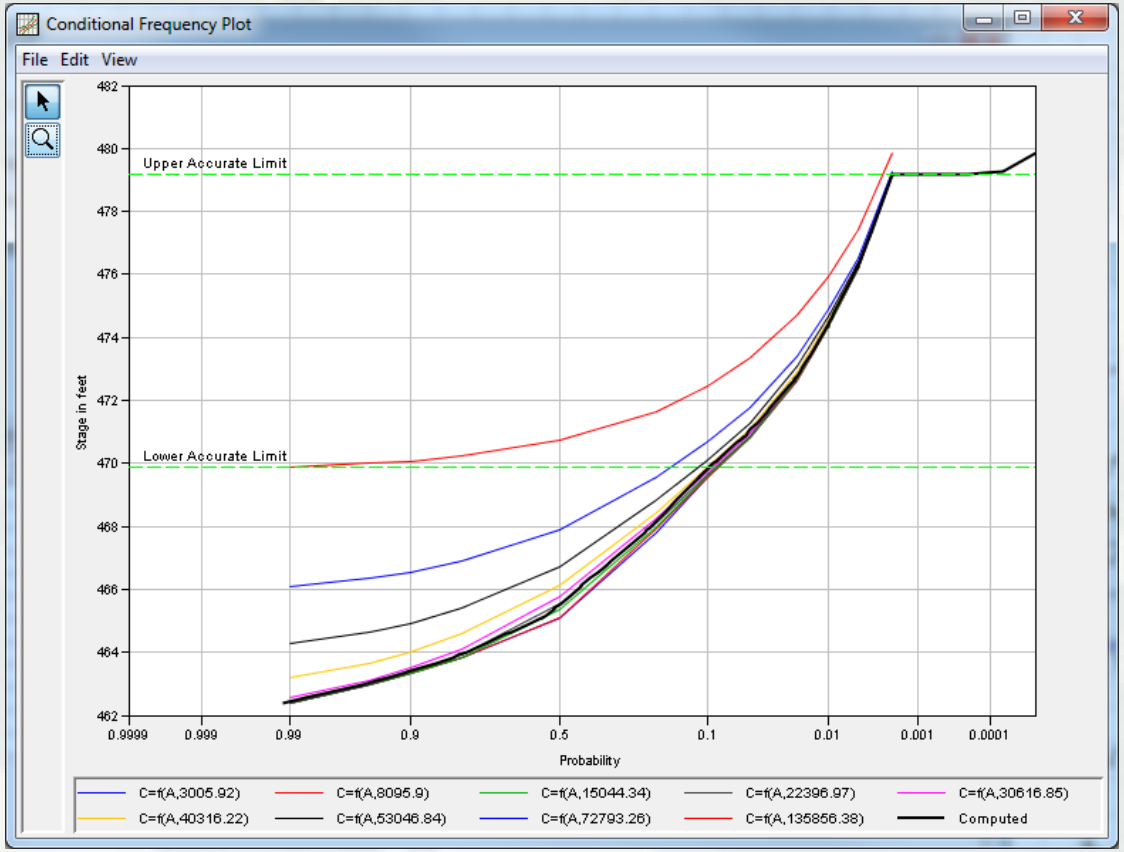
Curve Combination Analysis

- **Graphically-define** an **empirical distribution** for two or more input frequency curves
  - i.e. best-fit pool stage-frequency curve
- **Results** from other analyses **can be imported**
  - Bulletin 17, General Frequency

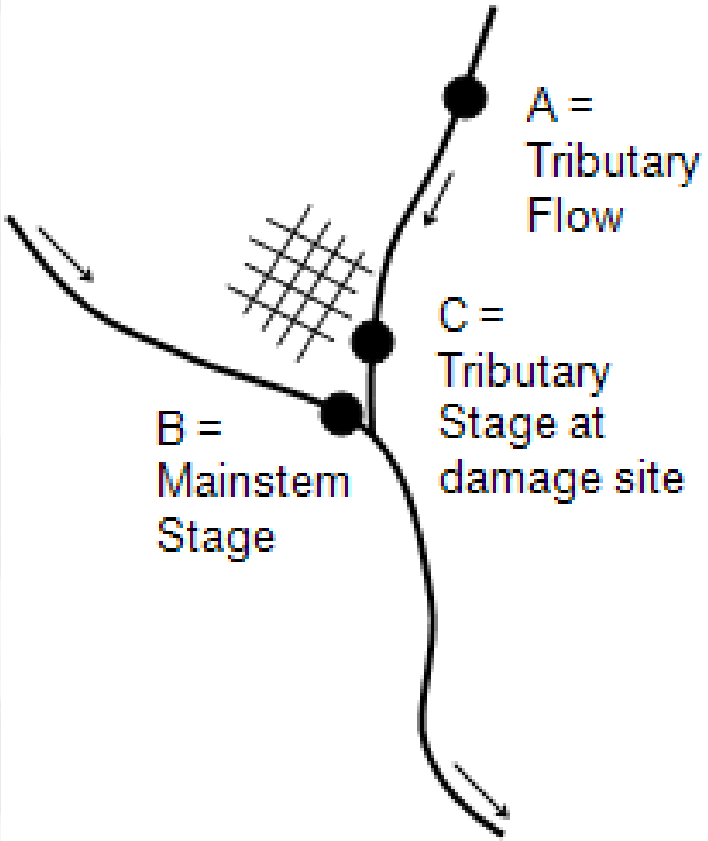


Coincident Frequency Analysis

- Uses **Total Probability Theorem** to compute a frequency curve that is a **function of two variables** (A and B)
- Two conditions are available:
  - Variable A and B are independent
  - Variable A and B are not independent
- Variable A
  - Flow- or Stage-Frequency Curve
- Variable B
  - Index Points from Flow- or Stage-Duration Curve
- Response Curves
  - Variable A results for each Variable B
  - Can have different Variable A for each Response Curve



Coincident Frequency Analysis



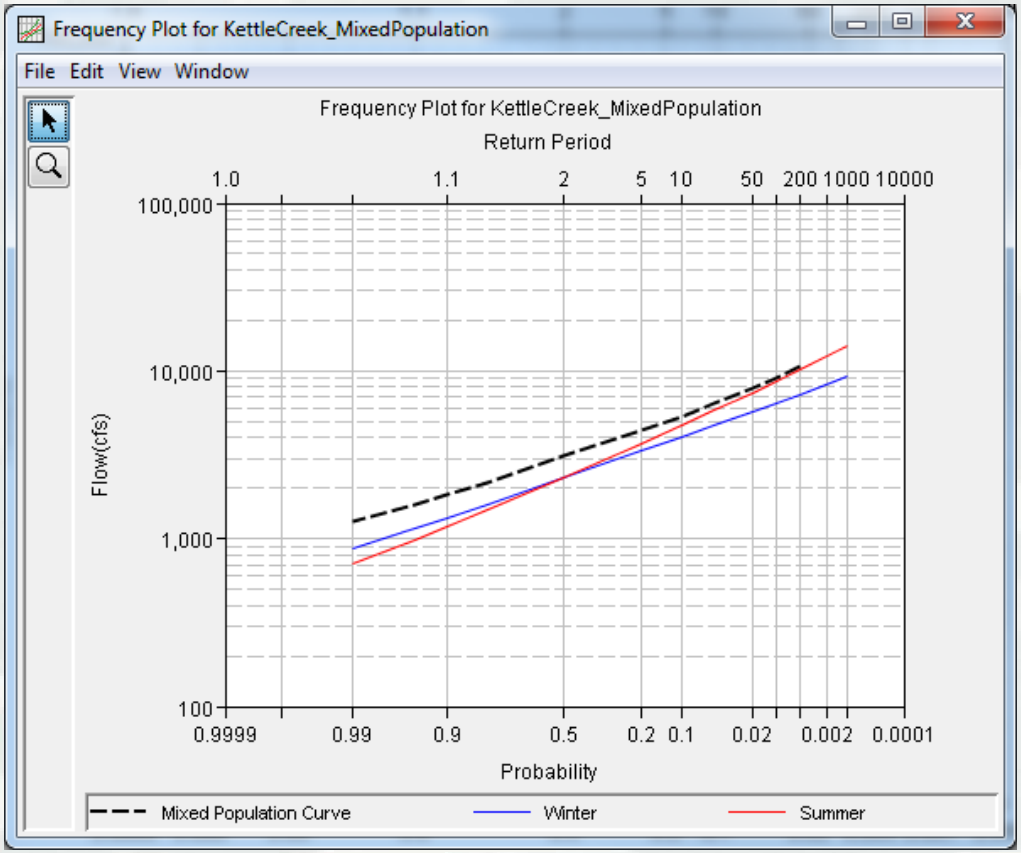
A = Tributary Flow — probability distribution "known"

B = Mainstem Stage — probability distribution "known"

C = Tributary Stage — probability distribution not known, value computed f(A,B)

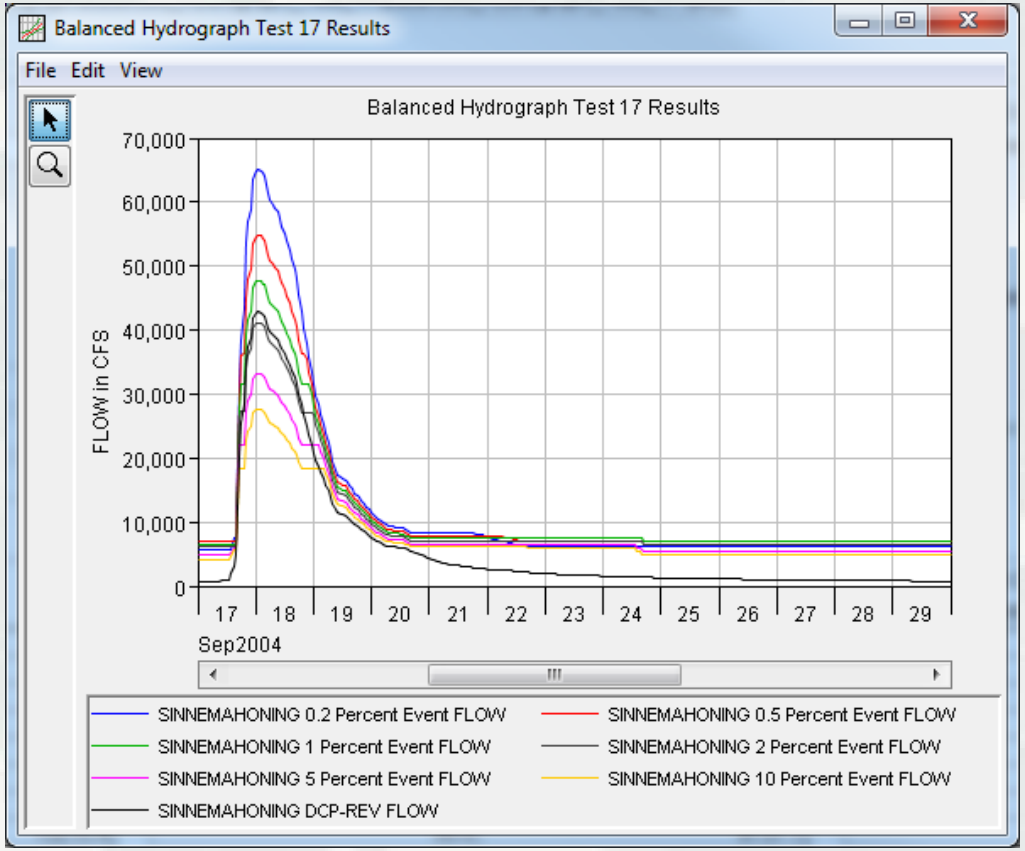
Mixed Population Analysis

- Uses **Total Probability Theorem** to compute a **frequency curve** from **two or more different runoff/causative mechanisms**
  - i.e. rainfall-only vs rain-on-snow vs snowmelt-only vs tropical storms
  - annual maximum series cannot be fit using the same analytical distribution
  - resultant empirical distribution takes into account the relative probability of a flood occurring in any year due to any of the input runoff mechanisms
- **Results** from other analyses **can be imported**
  - Bulletin 17, General Frequency



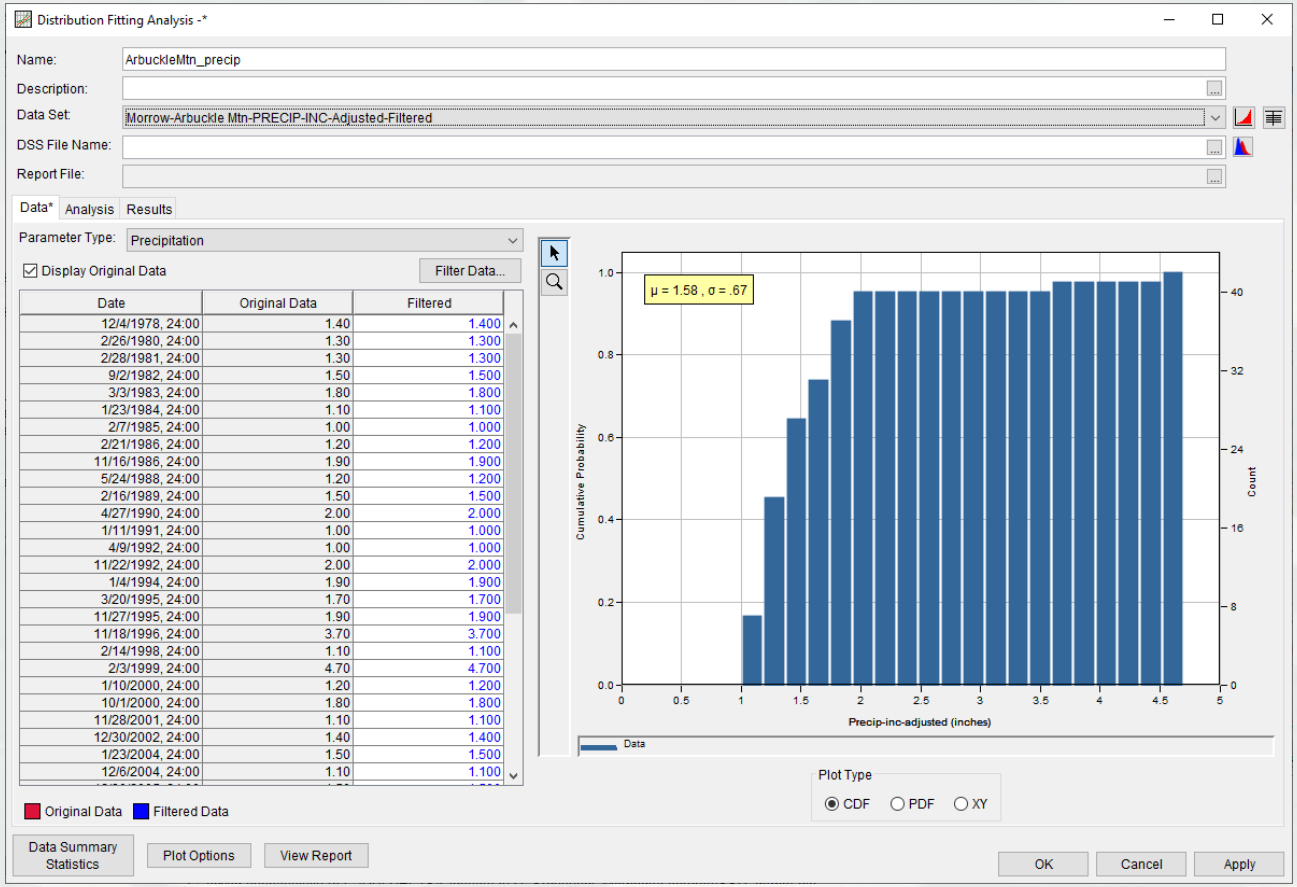
Balanced Hydrograph Analysis

- Computes **hydrograph shapes** that have been **modified** to **contain specific exceedance flow rates/volumes** across one or more **durations**
- **Results** from other analyses **can be imported**
  - Bulletin 17, General Frequency, Volume Frequency
- **REGULAR** data required
  - i.e. 1DAY



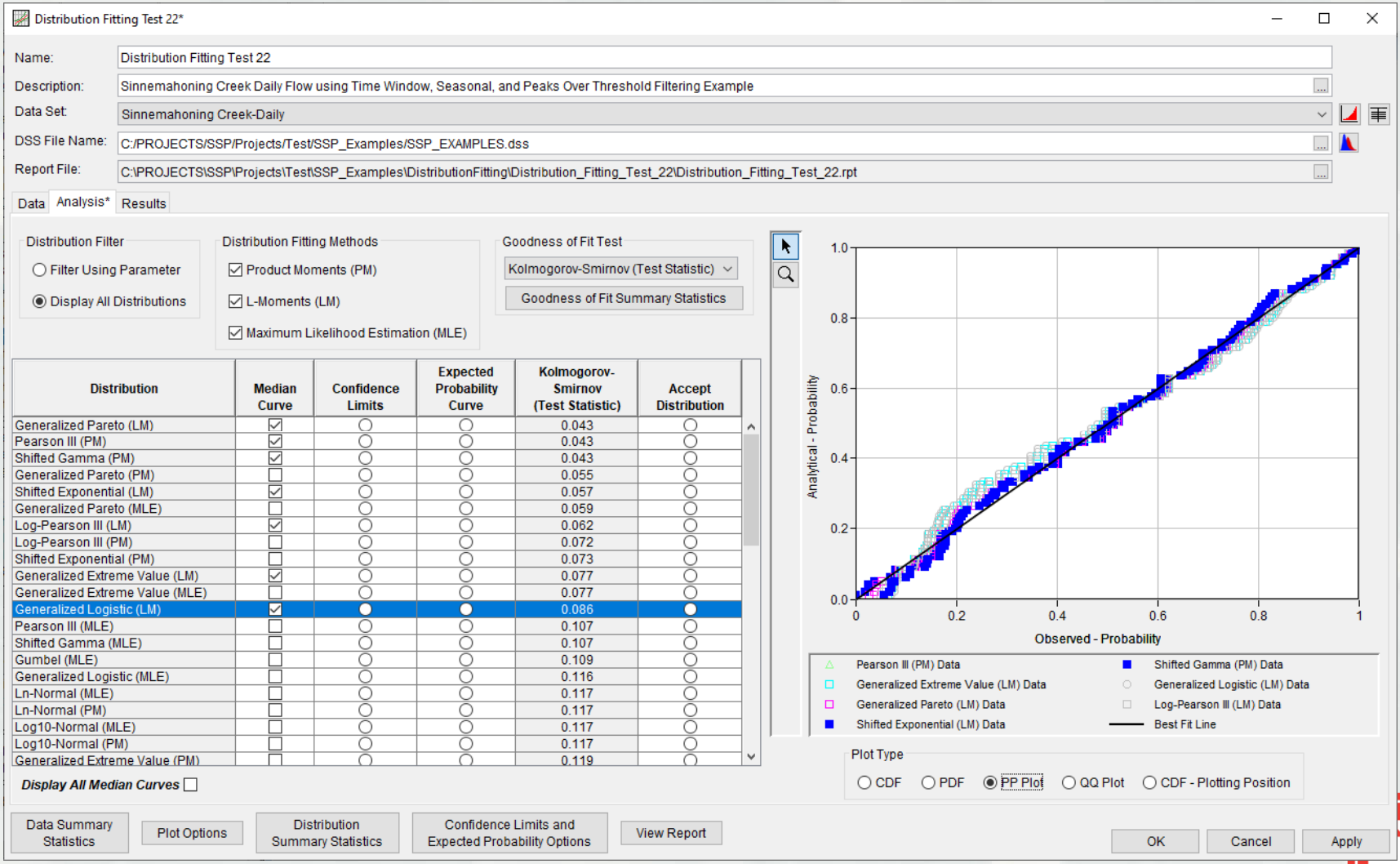
Distribution Fitting Analysis

- **Have you ever wondered what 19 analytical distributions look like when fit to the same data set?**
- **How much uncertainty is due to the choice of analytical distribution?**
- **IRREGULAR**, **REGULAR**, and **PAIRED DATA** accepted
- Can be used for flow, stage, precipitation, wind speed, wind direction, flood/event seasonality, etc



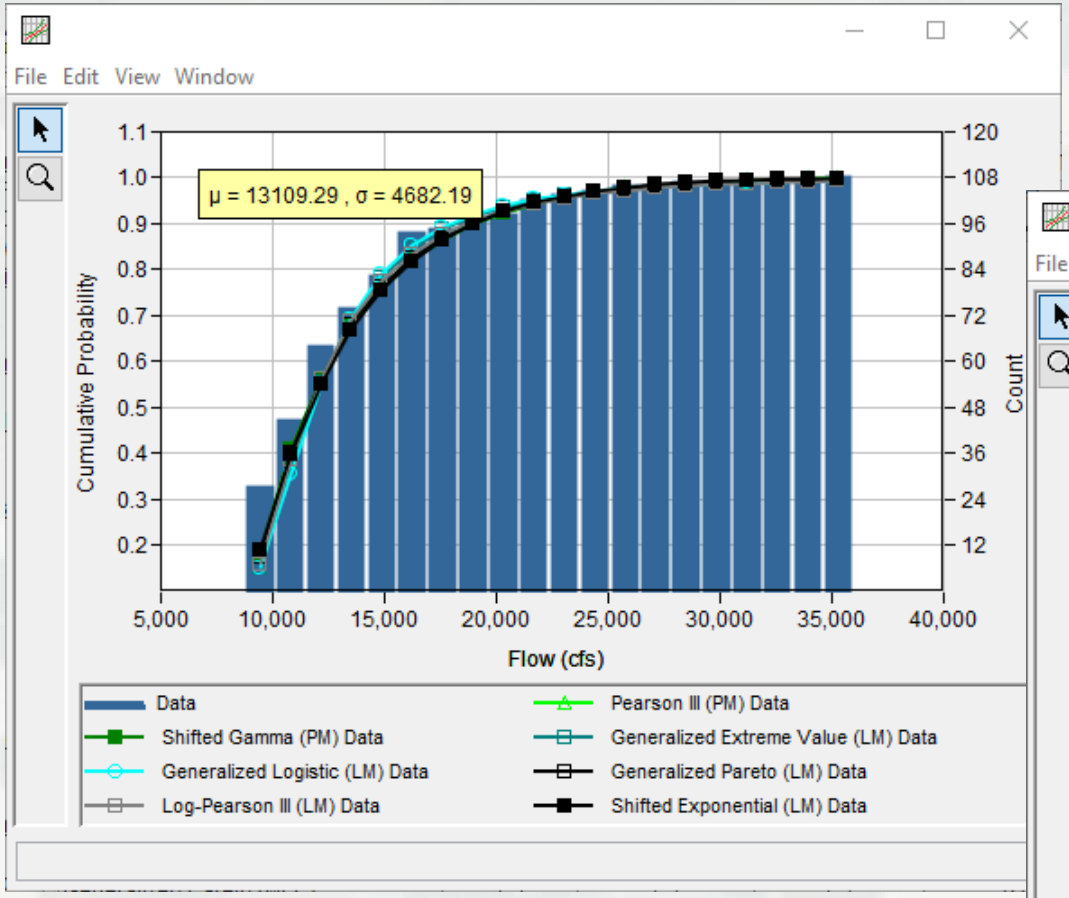
Distribution Fitting Analysis

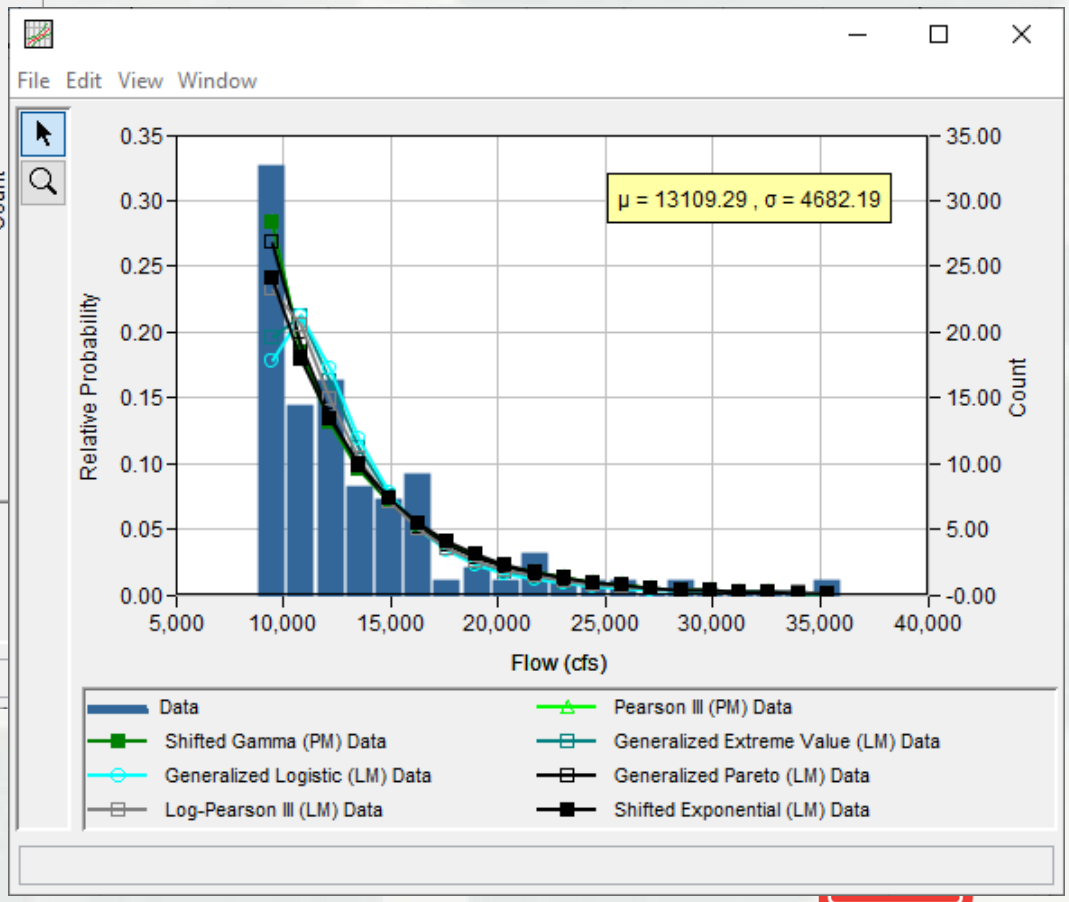
Analysis Tab



Distribution Fitting Analysis

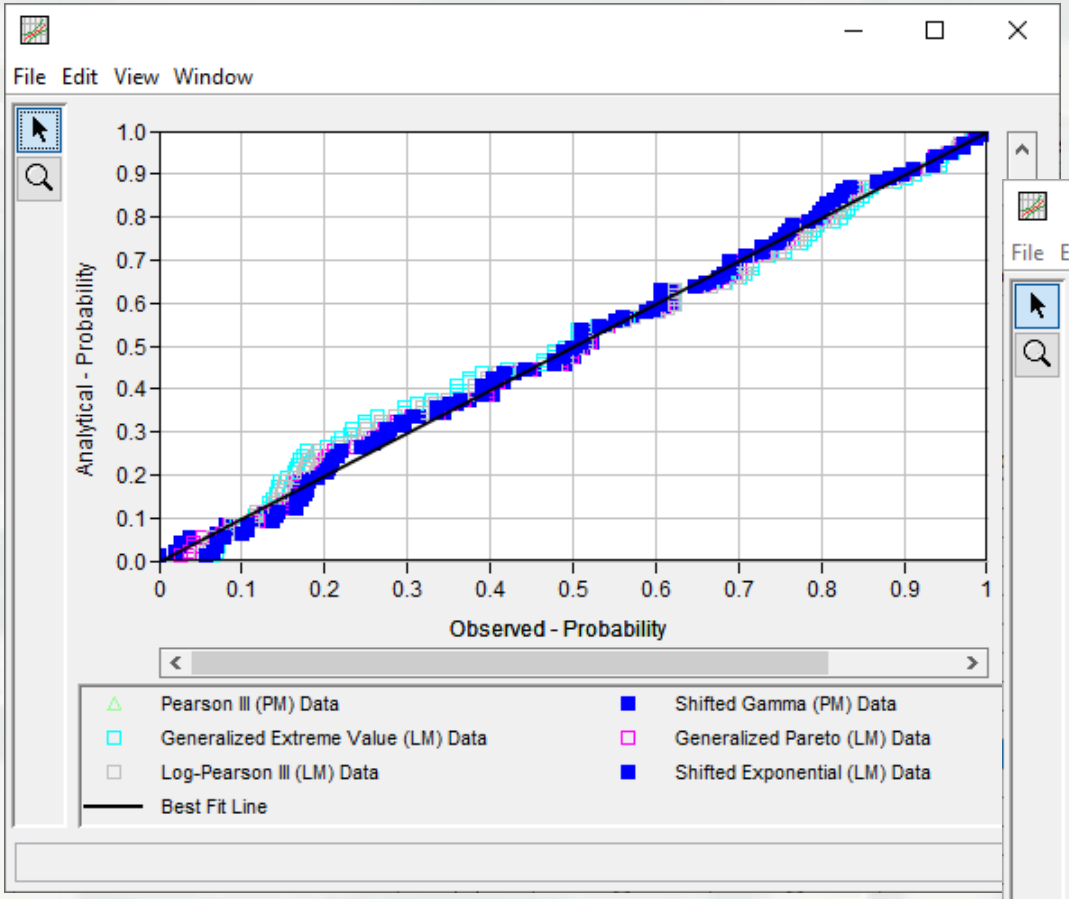
Analysis Tab

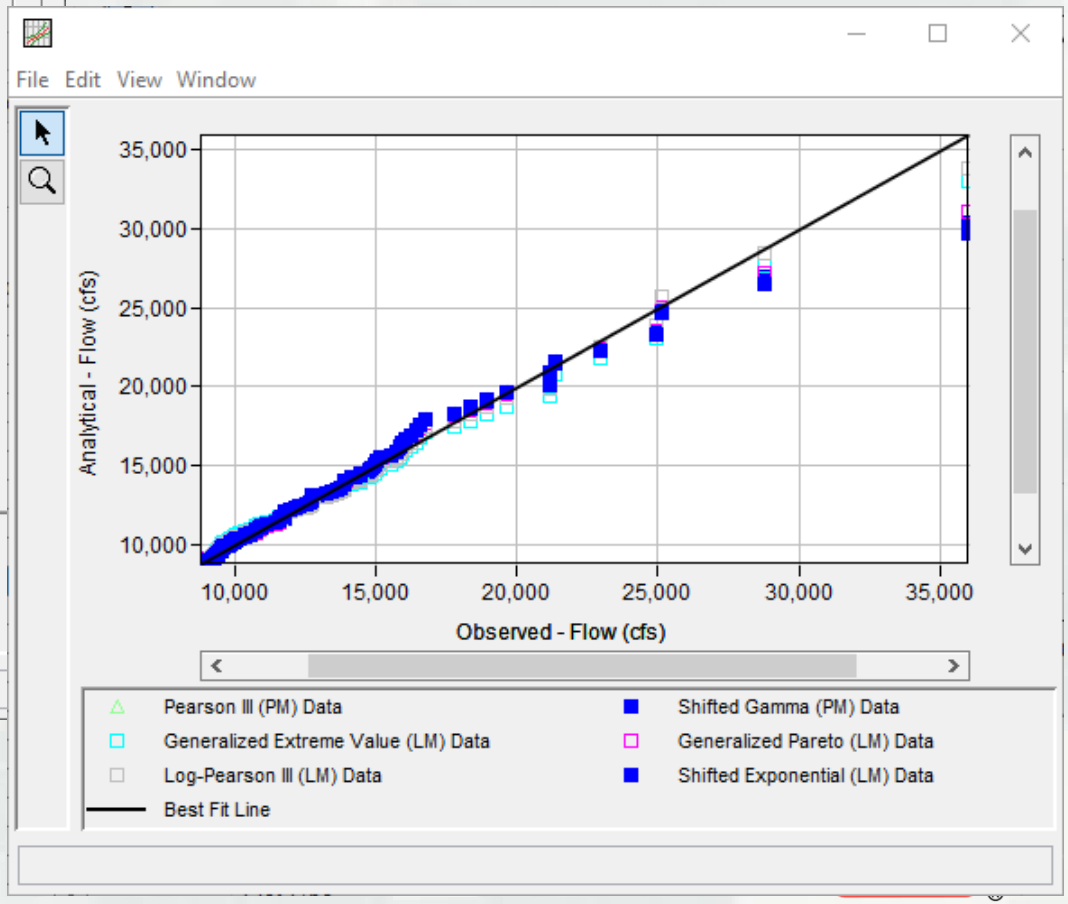




Distribution Fitting Analysis

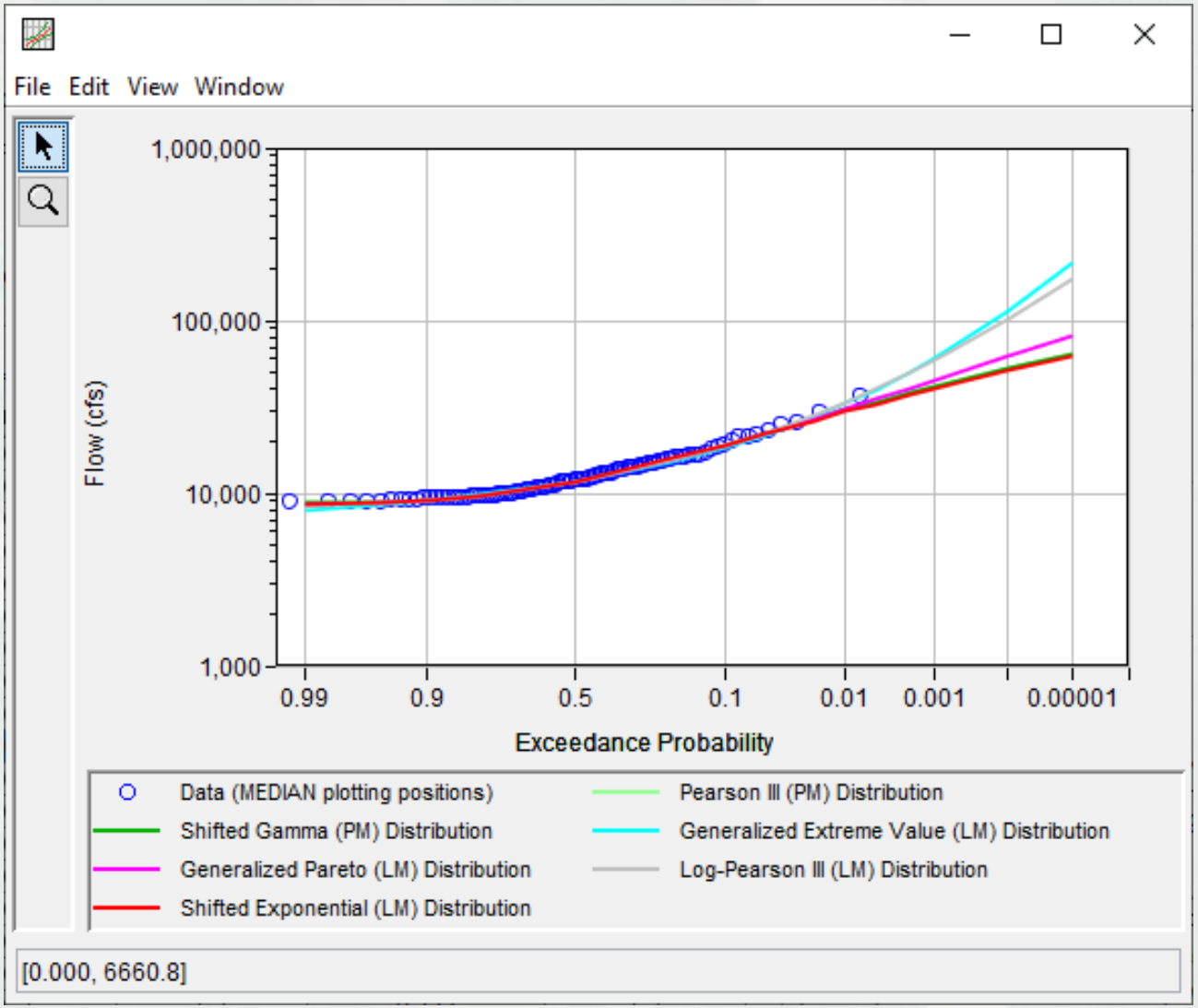
Analysis Tab





Distribution Fitting Analysis

Analysis Tab



Distribution Fitting Analysis

Results Tab

Distribution Fitting Test 22*

Name: Distribution Fitting Test 22

Description: Sinnemahoning Creek Daily Flow using Time Window, Seasonal, and Peaks Over Threshold Filtering Example

Data Set: Sinnemahoning Creek-Daily

DSS File Name: C:/PROJECTS/SSP/Projects/Test/SSP_Examples/SSP_EXAMPLES.dss

Report File: C:\PROJECTS\SSP\Projects\Test\SSP_Examples\DistributionFitting\Distribution_Fitting_Test_22\Distribution_Fitting_Test_22.rpt

Data | Analysis* | Results*

Statistics

Statistic	Value	
	Original Data	Filtered Data
Sample Size	28269	98
Min	1.400	8780.000
Max	44000.000	36000.000
Median	581.000	11750.000
Mode	1100.000	11800.000

Product Moments		
Statistic	Original Data	Filtered Data
Mean	1148.277	13109.286
Standard Dev	1749.434	4682.193
Skew	5.283	2.164
Kurtosis	55.889	6.294

L-Moment Ratios		
Statistic	Original Data	Filtered Data
L-Mean	1148.277	13109.286
L-CV	0.604	0.175
L-Skew	0.490	0.364
L-Kurtosis	0.292	0.188

Selected Distribution: Generalized Pareto (LM)

Distribution Parameters

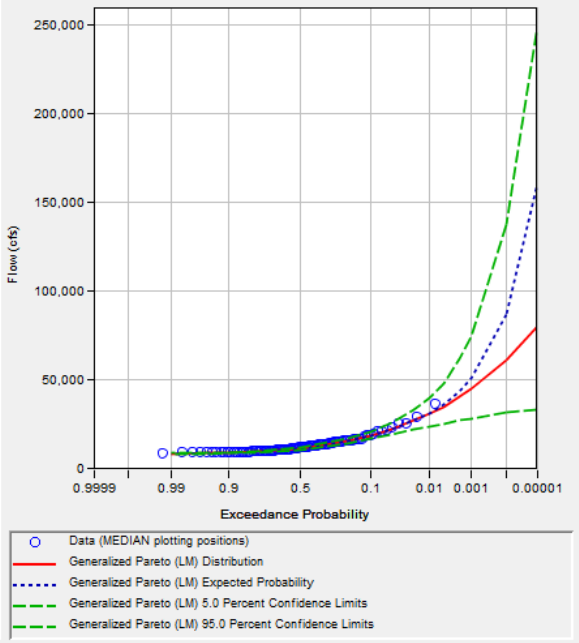
Parameter	Value
Loctn	8665.310
Scale	4146.730
Shape	-0.070

Goodness of Fit

Kolmogorov-Smirnov (Test Statistic)
0.043

Frequency Curves

Percent Chance Exceedance	Median Curve	Expected Probability	Confidence Limits Flow in cfs	
	Flow in cfs	Flow in cfs	0.05	0.95
0.001	80575.3	161302.9	250035.5	33725.1
0.01	61461.6	87041.6	137463.4	31720.0
0.1	45076.1	50823.5	74483.2	28549.5
0.2	40617.9	43776.1	61850.5	27350.8
0.5	35032.8	36155.1	48400.0	25519.7
1.0	31029.6	31360.8	40091.0	23999.8
2.0	27207.7	27182.1	33133.0	22124.5
5.0	22419.9	22311.3	25744.3	19323.0
10.0	18988.0	18928.9	21194.7	16911.1
20.0	15711.7	15693.3	17209.7	14324.7
50.0	11607.3	11592.6	12346.5	10968.1
80.0	9597.6	9588.9	9924.9	9319.8
90.0	9103.8	9099.9	9388.5	8845.3
95.0	8878.4	8863.3	9166.3	8602.0
99.0	8707.0	8592.4	9007.7	8404.7

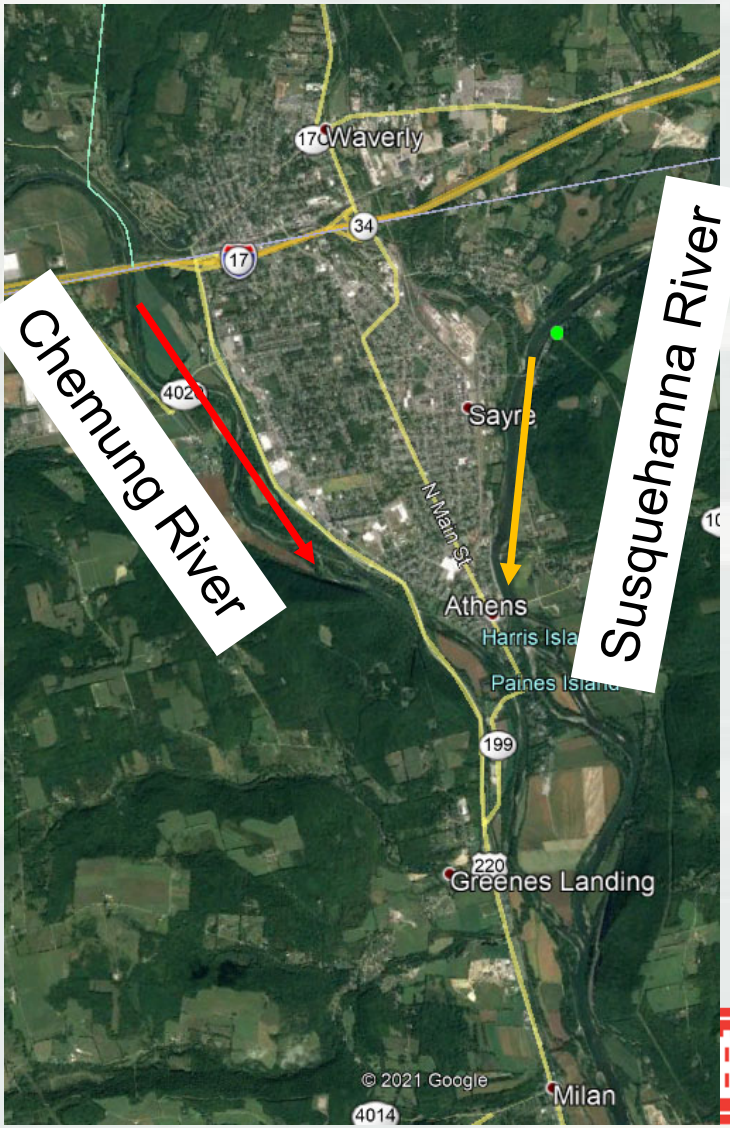


Plot Type: ○ CDF ○ PDF ◉ CDF - Plotting Position

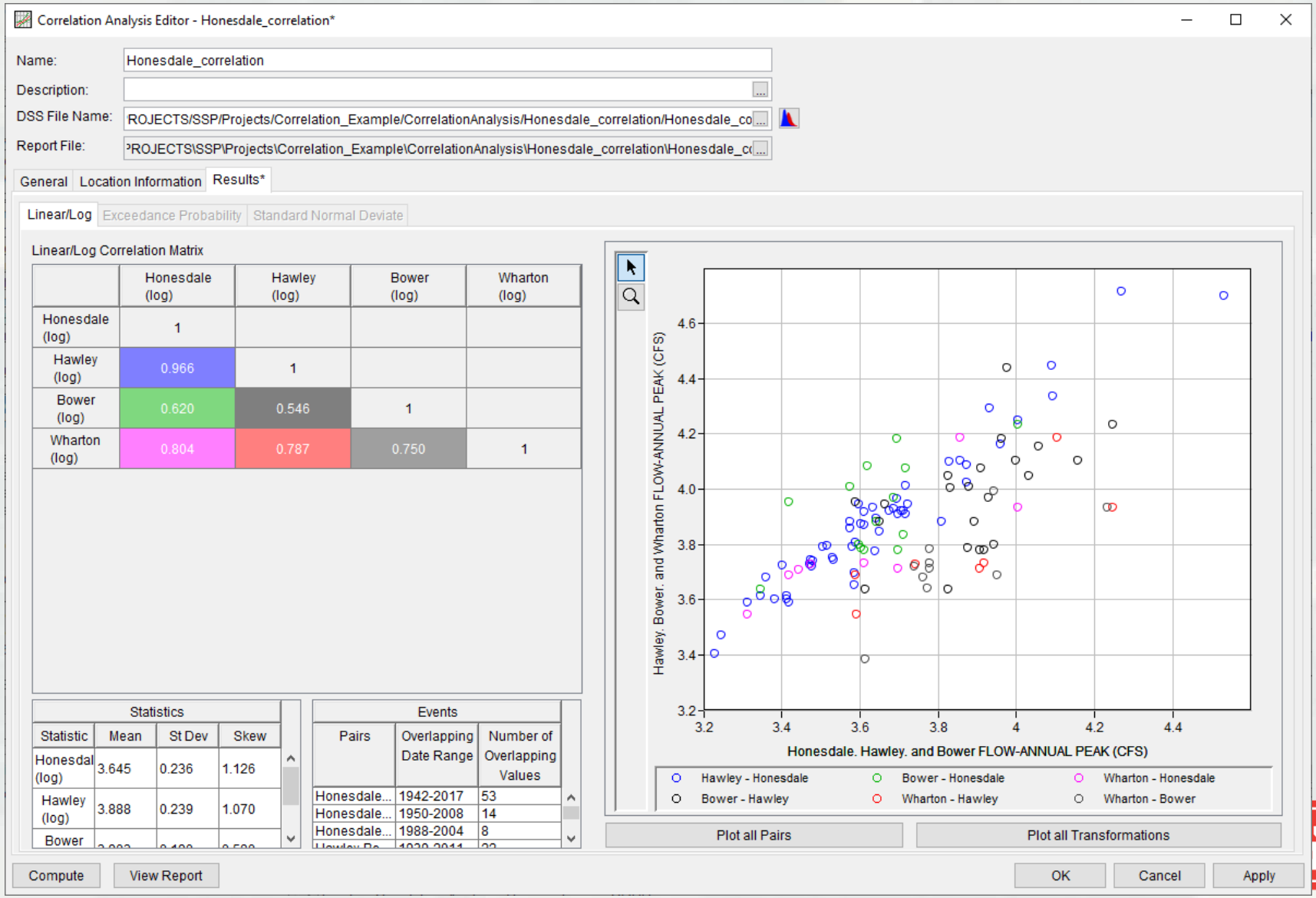
Plot Options | View Report | OK | Cancel | Apply

Correlation Analysis

- Compute the amount of correlation between various data sets
  - Tributary peak flow vs. mainstem stage
  - 3-day precipitation accumulation vs. 3-day average temperature
  - Annual maximum SWE vs. annual maximum 24-hour precipitation accumulation
- **Results** from B17 analyses **can be imported**
- **IRREGULAR** and **REGULAR** data accepted

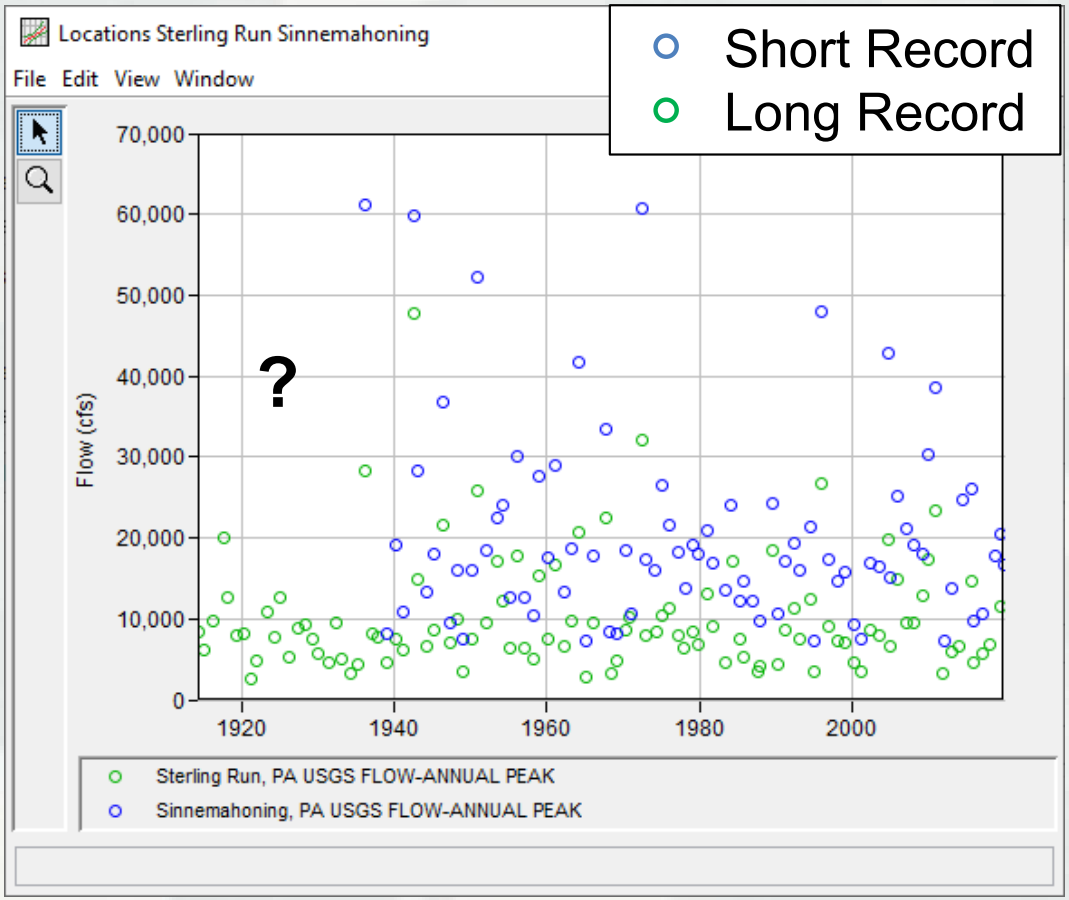


Correlation Analysis

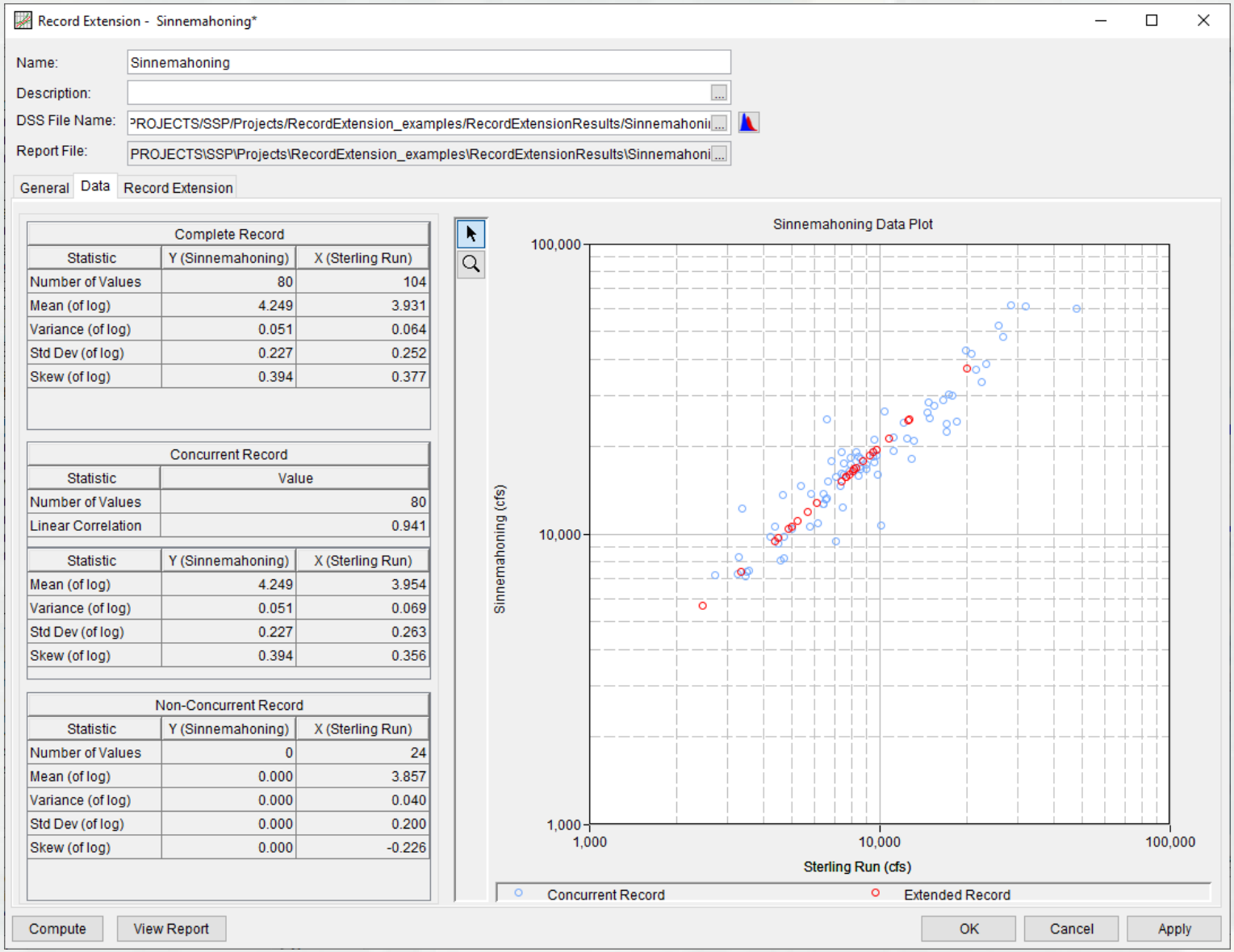


Record Extension Analysis

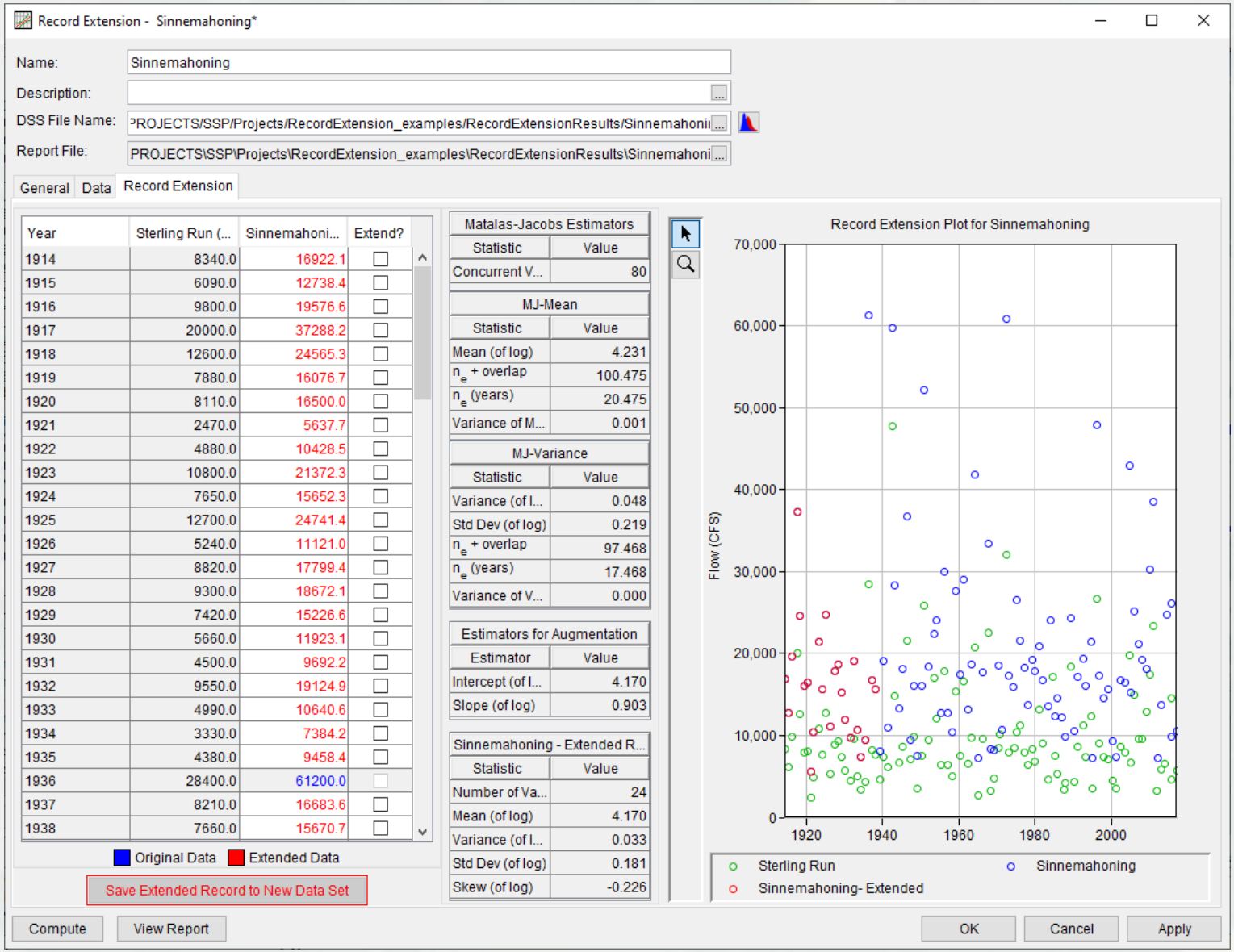
- Extend a short record using a longer record
- Multiple computational methods
  - Ordinary Least Squares
  - MOVE.1
  - MOVE.2
  - MOVE.3
  - MOVE.4
- **Results** can be used within B17 analyses to infer flow-frequency for the extended record
- **IRREGULAR** and **REGULAR** data accepted



Record Extension Analysis



Record Extension Analysis



History, Status, Future…

- New analytical tools to meet Corps needs
  - Updated Distribution Fitting analysis
  - New Bulletin 17 moving/expanding time window
  - New Correlation Analysis
  - New Record Extension Analysis
- Improved user experience
  - Improved data entry
  - Separate DSS files for each analysis
  - Button to view/open DSS file
  - Button to plot/tabulate selected data set

Summary

- Currently contains nine different statistical analyses
  - ► Future versions will include two additional analyses
- Developed primarily to meet USACE needs
- If you have ideas for future enhancements or questions about existing features, let us know