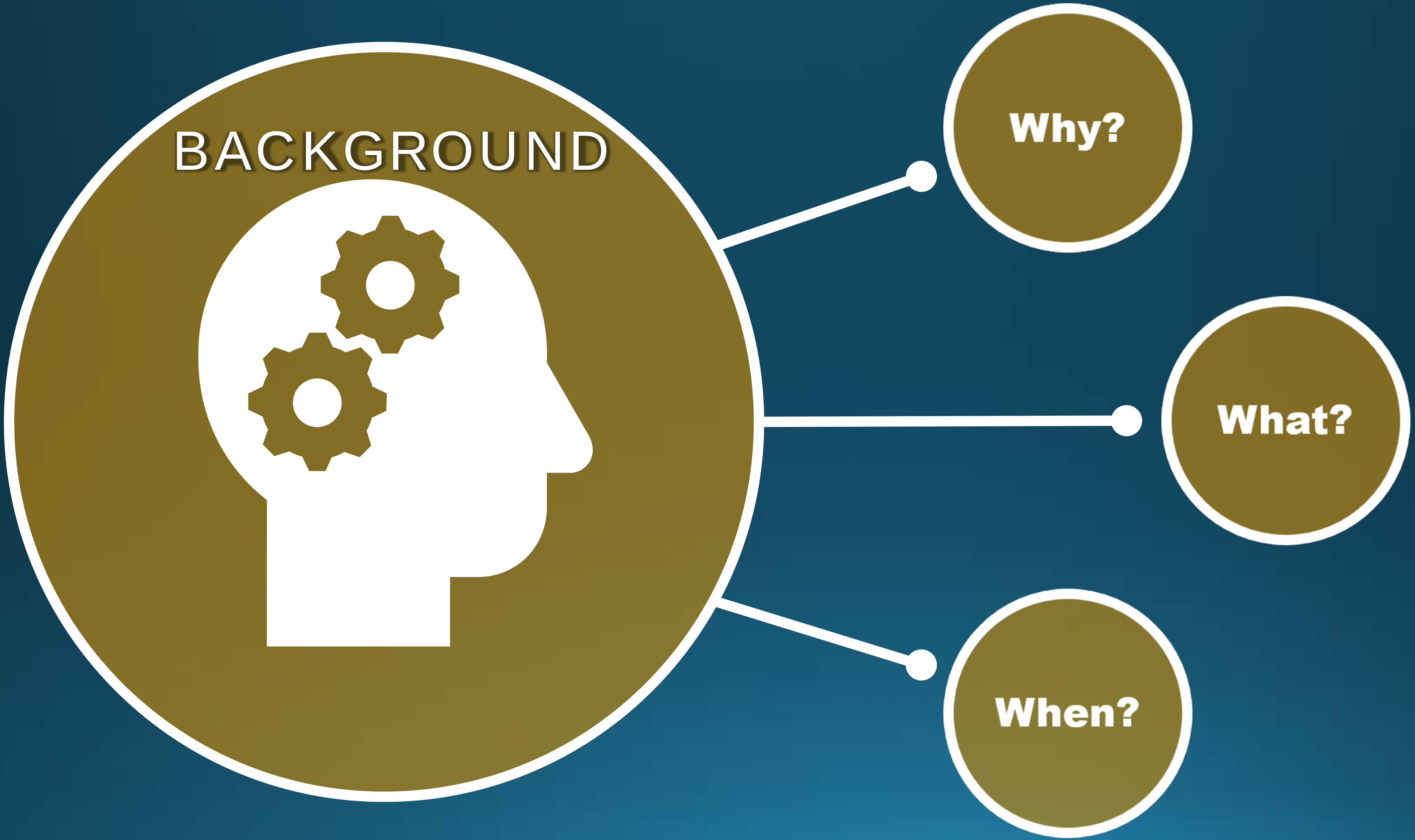
RMC-BestFit

Bayesian Estimation and Fitting Software

RMC-BestFit

BACKGROUND
USER INTERFACE
INPUT DATA
DISTRIBUTION FITTING
BAYESIAN ESTIMATION


BACKGROUND
Why?
What?
When?

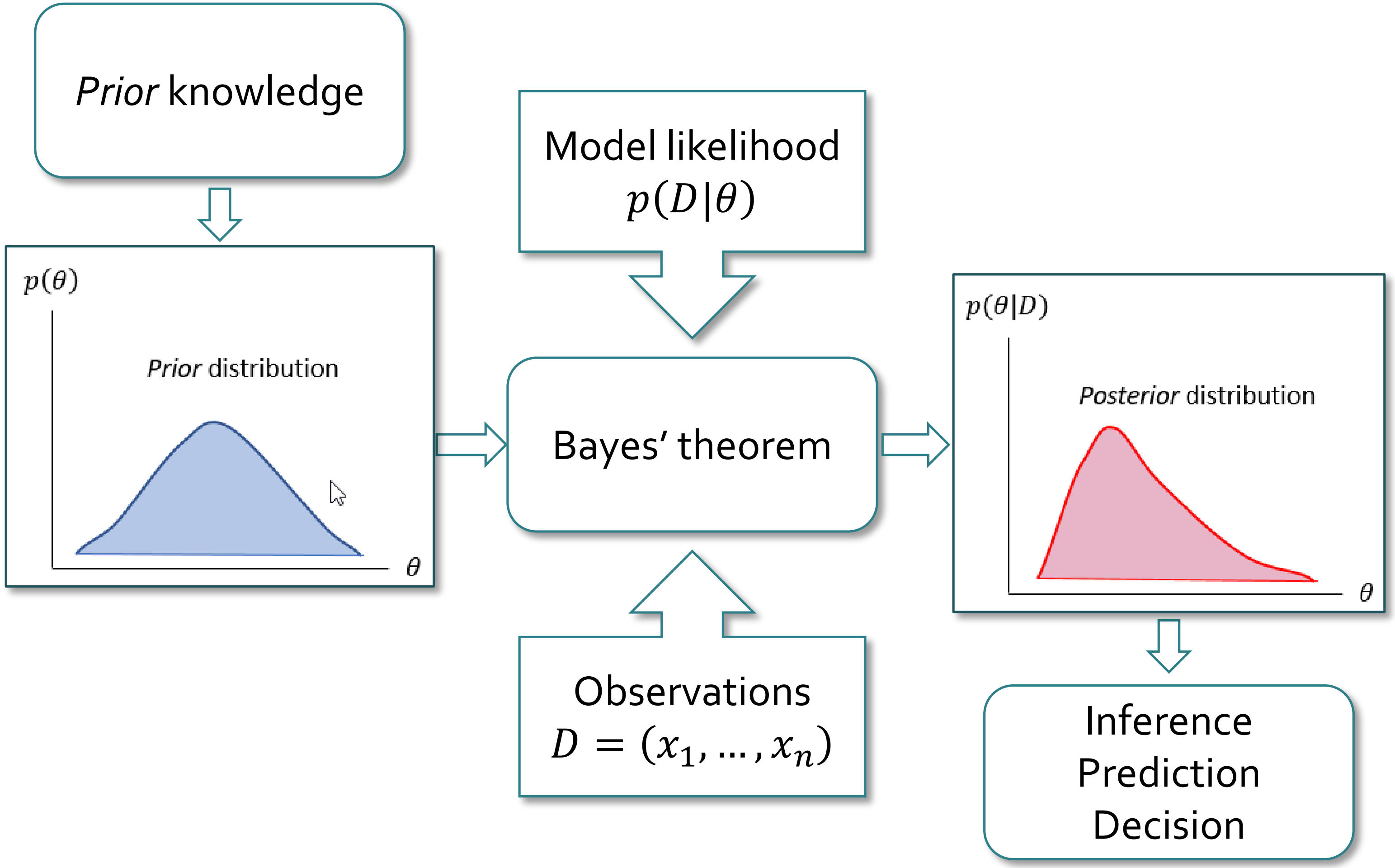
Why?

- **To enhance and expedite flood hazard assessments within the Flood Risk Management, Planning, and Dam and Levee Safety communities of practice**
  - The Bayesian method can incorporate all available sources of hydrologic information, such as paleofloods, regional rainfall-runoff results, and expert elicitation.
  - As such, it provides higher confidence in the fitted flood frequency curves and resulting reservoir stage-frequency curves
  - RMC-BestFit was developed by the RMC, in collaboration with ERDC–CHL

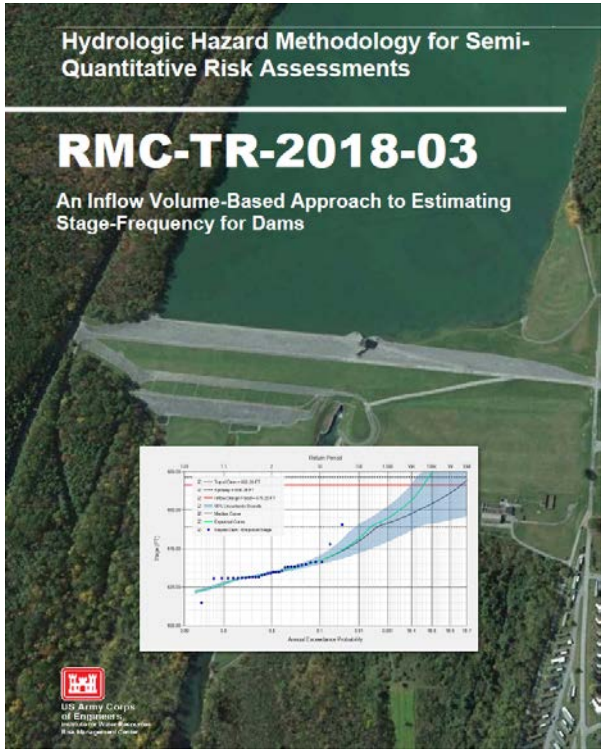
What?

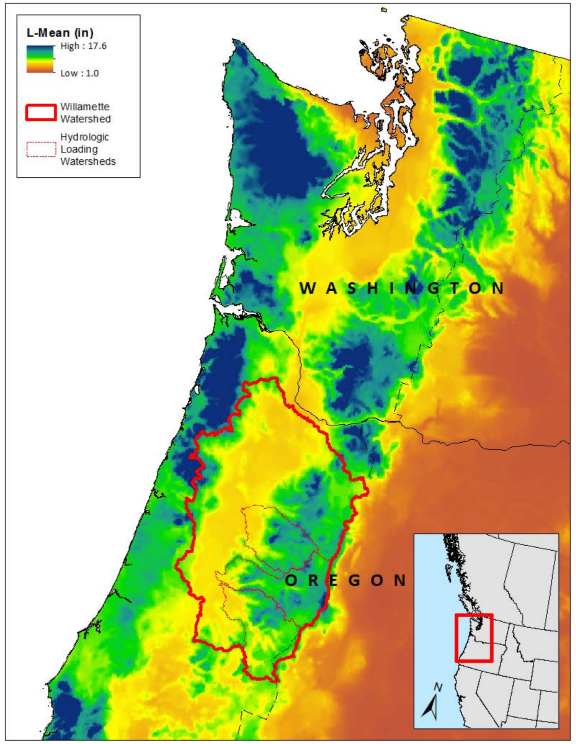
Bayes' Theorem

$$P(\theta|\mathrm{D}) = \frac{P(D|\theta)\cdot P(\theta)}{\int P(D|\theta)\cdot P(\theta)\cdot d\theta}$$

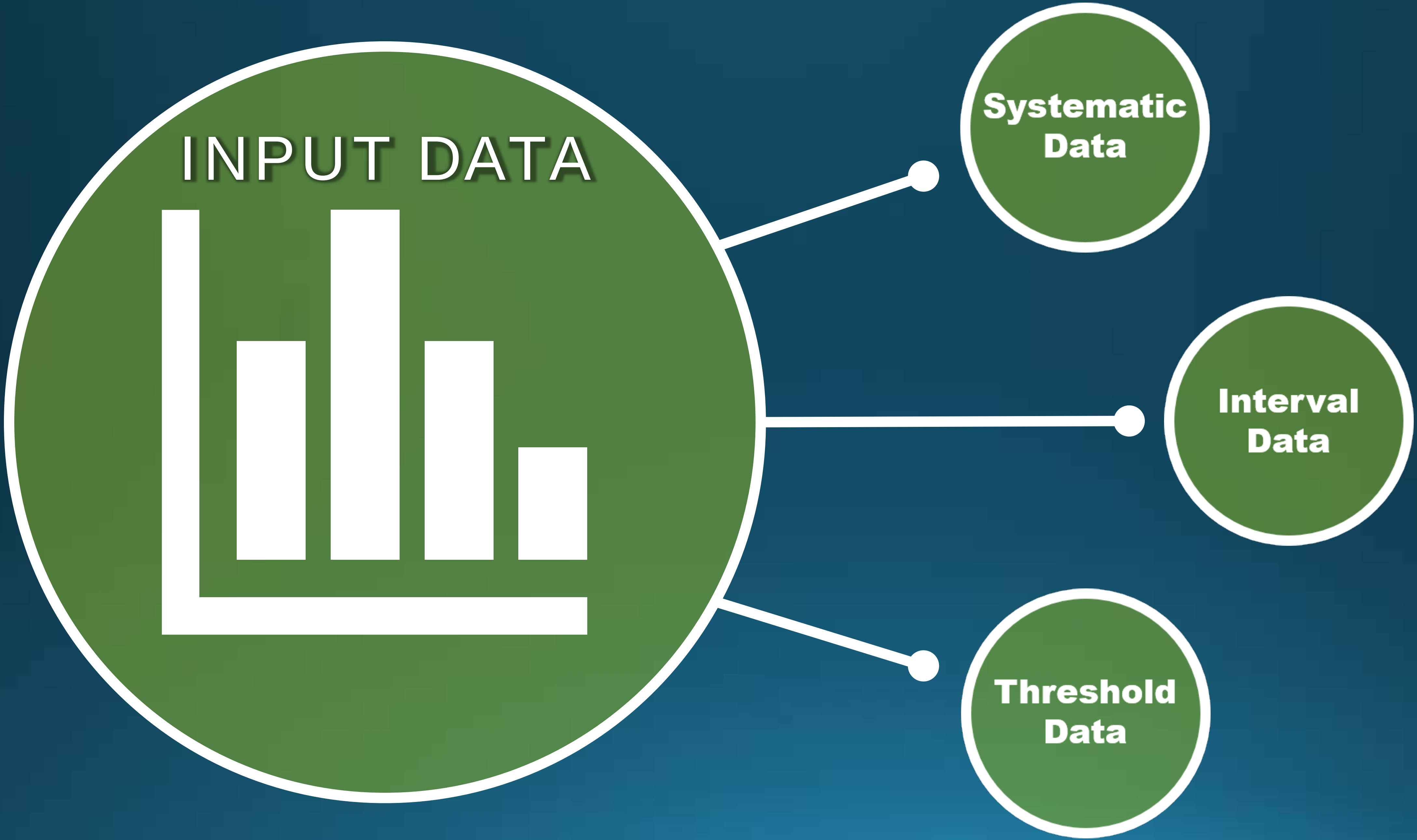


When?





- Semi-Quantitative Risk or Hazard Assessments, or higher level of effort
- Most valuable when there are multiple sources of data
- Can be used in flood and/or seismic hazard assessments and reliability analysis

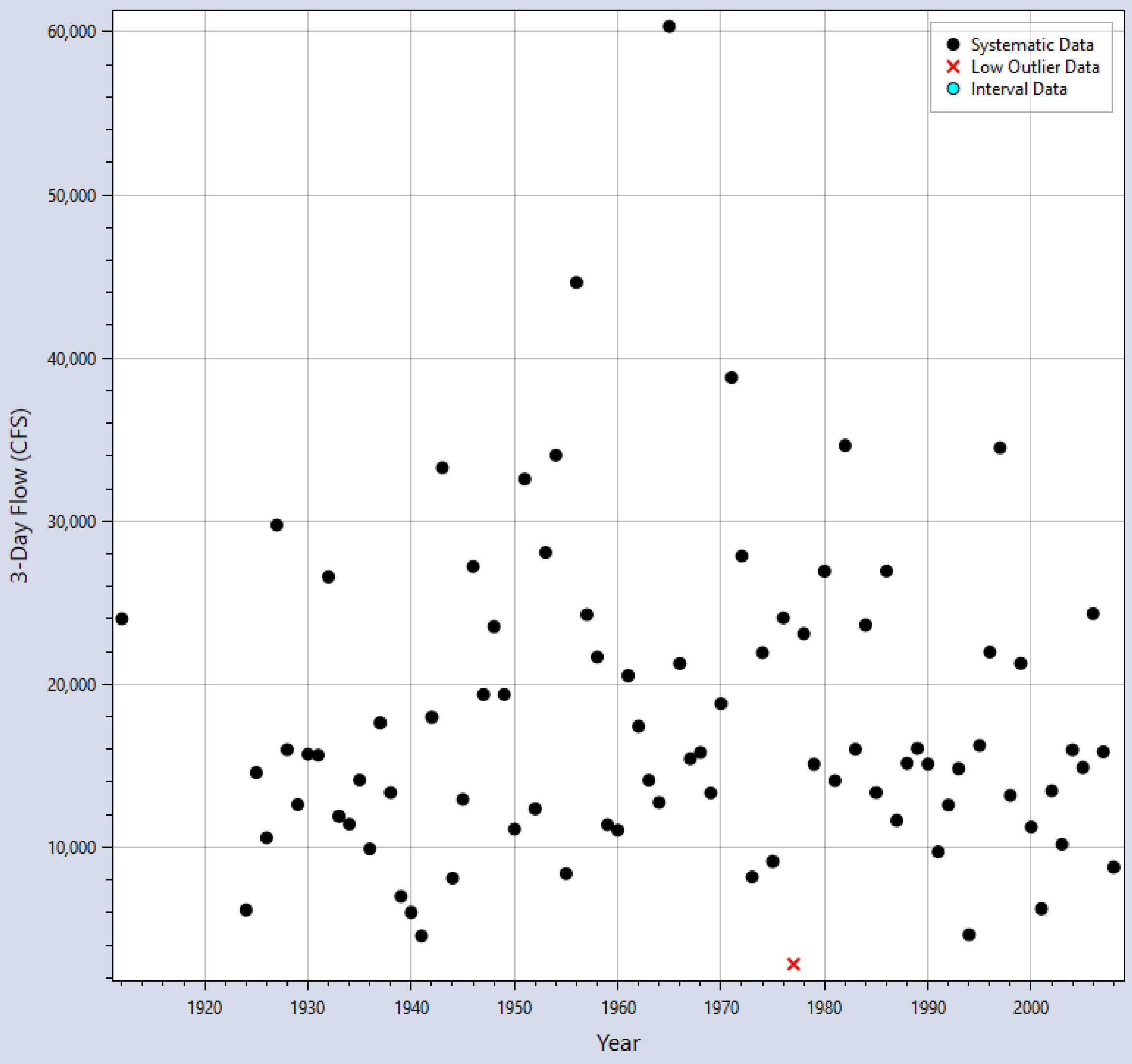

INPUT DATA
Systematic Data
Interval Data
Threshold Data

Systematic Data

Systematic

Systematic Data | Interval Data | Perception Thresholds | Summary Statistics

Year	Value	Plotting Position	Is Low Outlier
1912	24,000	0.215780	☐
1924	6,150	0.946640	☐
1925	14,580	0.575410	☐
1926	10,580	0.819030	☐
1927	29,750	0.099770	☐
1928	15,980	0.436190	☐
1929	12,610	0.703020	☐
1930	15,700	0.482600	☐
1931	15,640	0.494200	☐
1932	26,580	0.169370	☐
1933	11,900	0.737820	☐
1934	11,420	0.761020	☐
1935	14,130	0.587010	☐
1936	9,890	0.842230	☐
1937	17,630	0.378190	☐
1938	13,350	0.645010	☐
1939	6,980	0.923430	☐
1940	6,000	0.958240	☐
1941	4,560	0.981440	☐
1942	17,970	0.366590	☐
1943	33,270	0.076570	☐
1944	8,110	0.911830	☐
1945	12,930	0.679810	☐
1946	27,210	0.134570	☐
1947	19,370	0.331790	☐
1948	23,530	0.238980	☐
1949	19,370	0.343390	☐
1950	11,120	0.795820	☐
1951	32,570	0.088170	☐
1952	12,350	0.726220	☐
1953	28,080	0.111370	☐



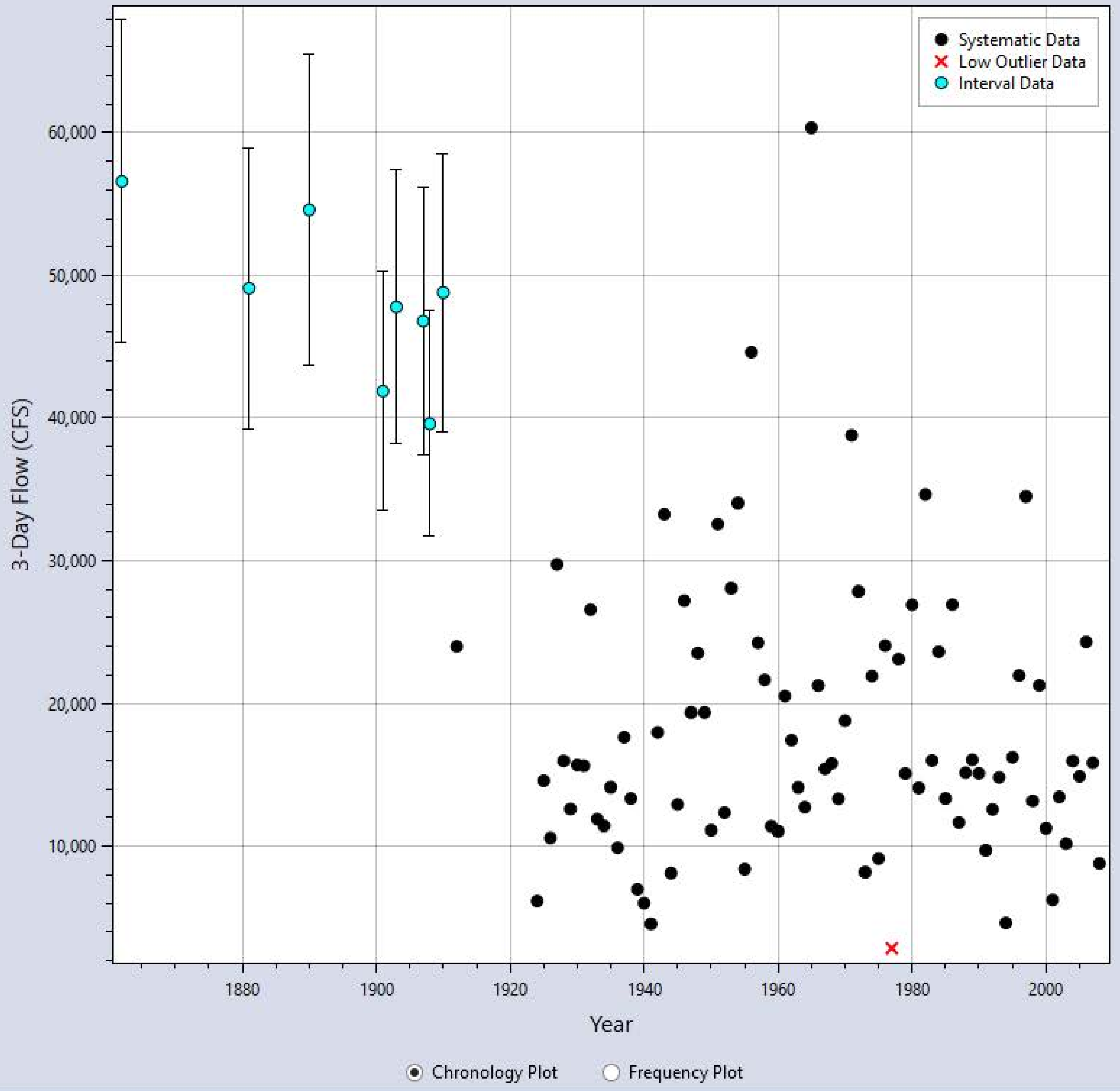
Chronology Plot | Frequency Plot

Interval Data

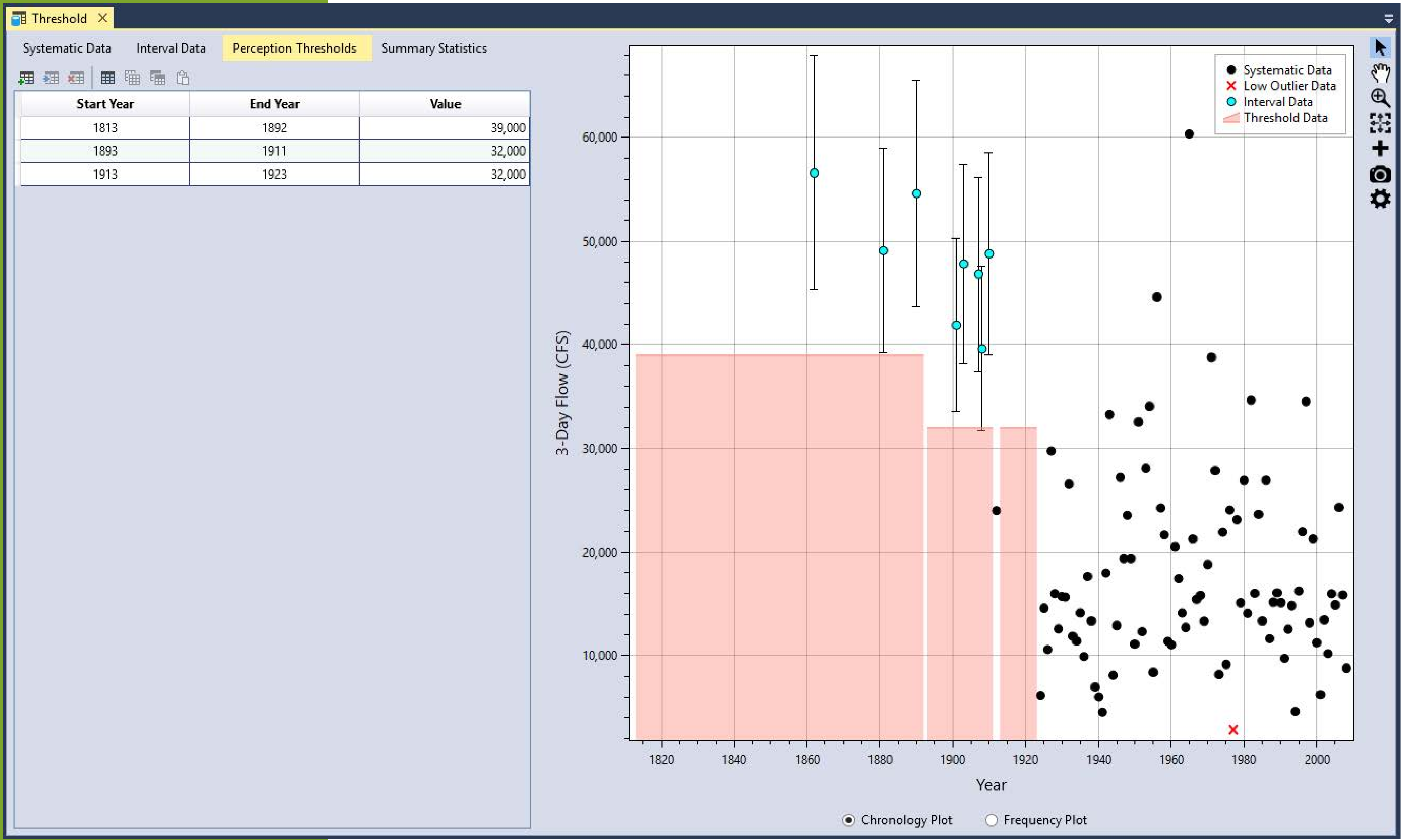
Interval

Systematic Data | Interval Data | Perception Thresholds | Summary Statistics

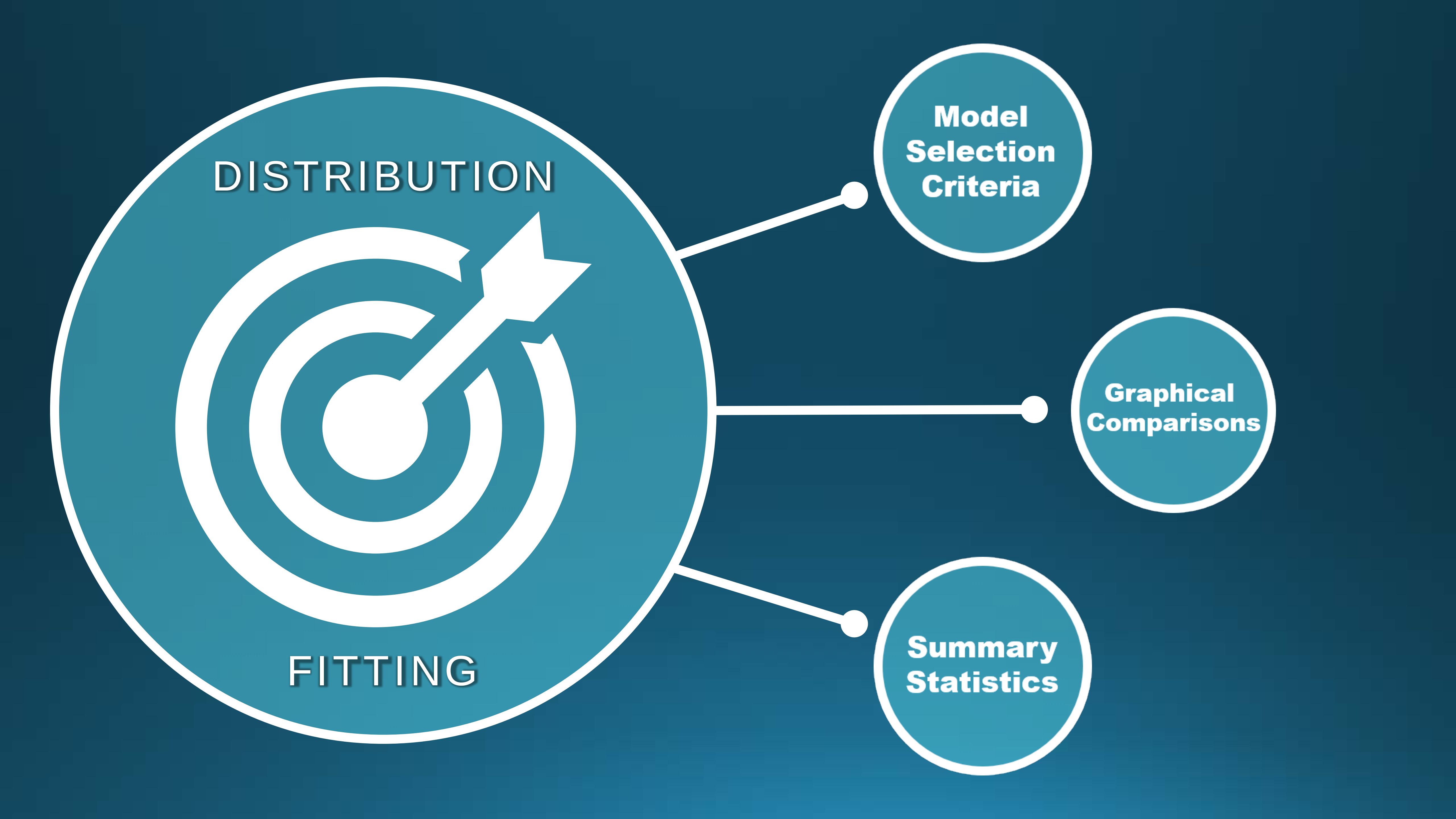
Year	Lower	Most Likely	Upper	Plotting Position
1862	45,280	56,600	67,920	0.016990
1881	39,280	49,100	58,920	0.038220
1890	43,680	54,600	65,520	0.027600
1901	33,520	41,900	50,280	0.091300
1903	38,240	47,800	57,360	0.059450
1907	37,440	46,800	56,160	0.070060
1908	31,680	39,600	47,520	0.101910
1910	39,040	48,800	58,560	0.048830



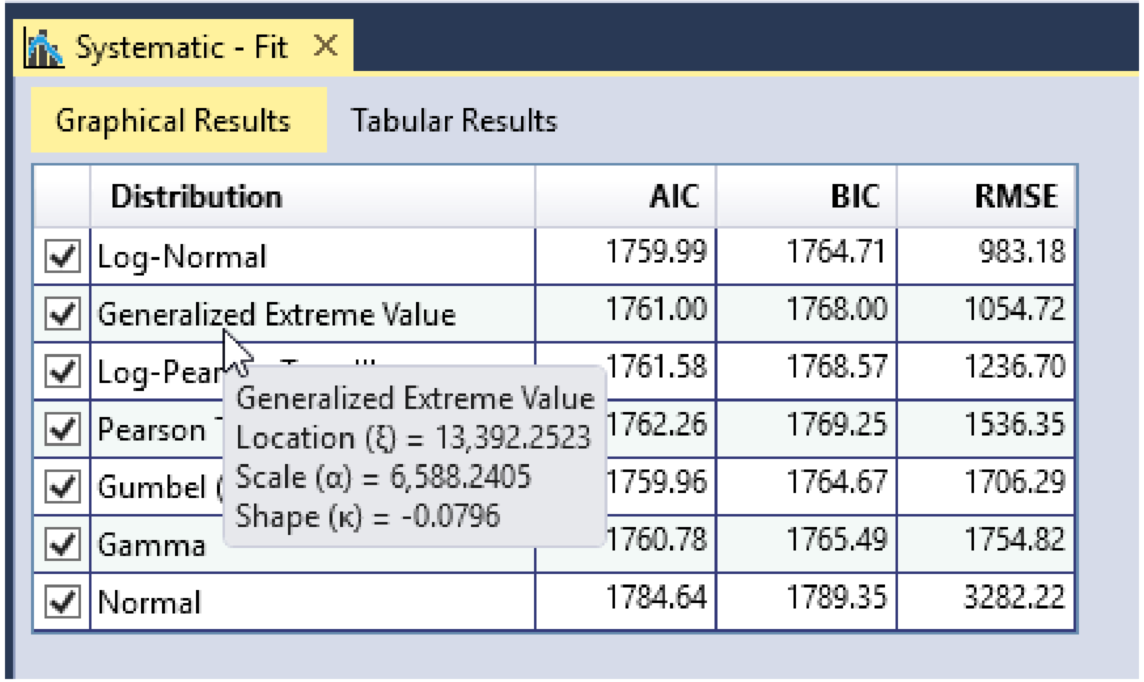
Threshold Data



Start Year	End Year	Value
1813	1892	39,000
1893	1911	32,000
1913	1923	32,000


DISTRIBUTION
FITTING
Model Selection Criteria
Graphical Comparisons
Summary Statistics

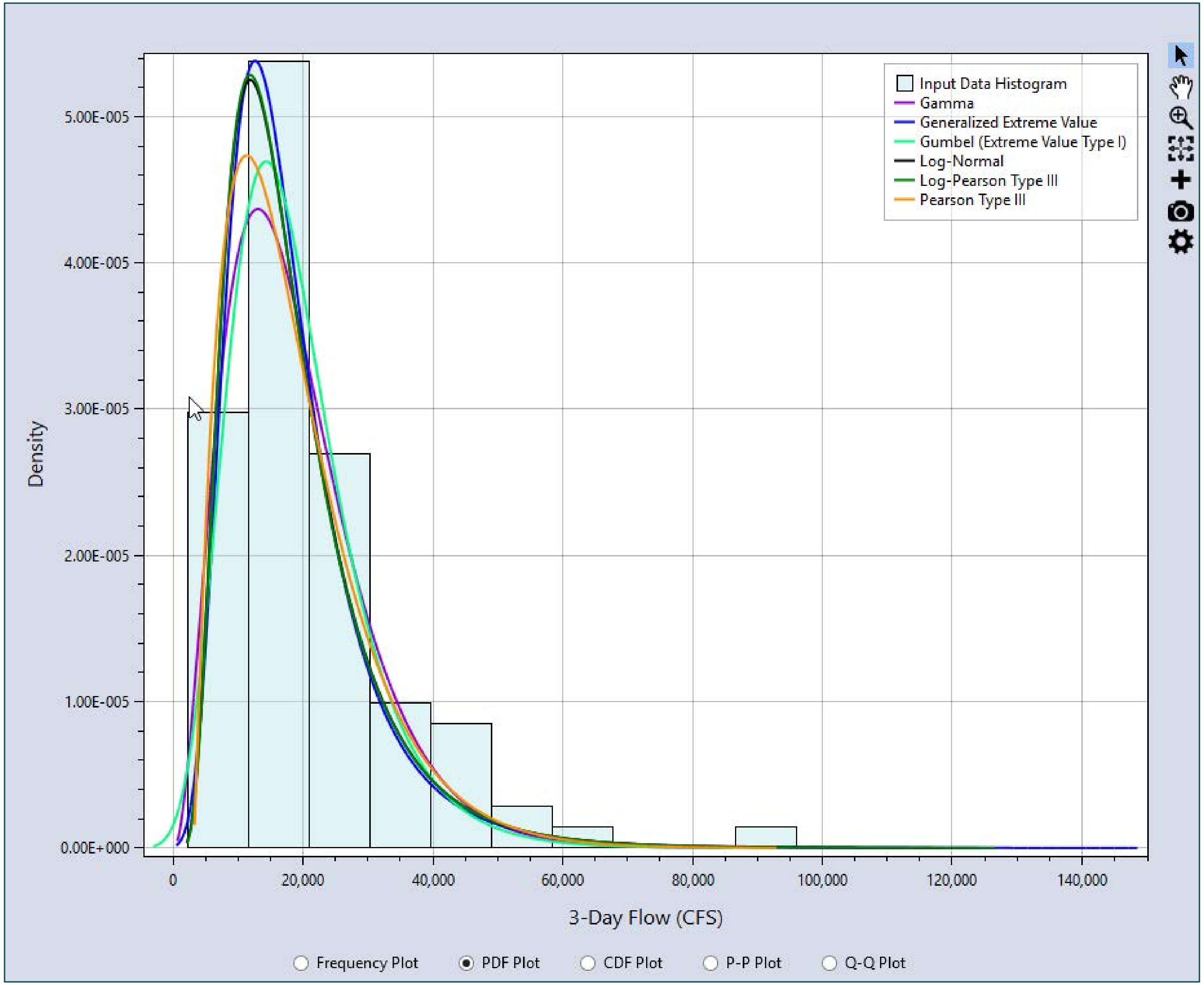
Model Selection Criteria



	Distribution	AIC	BIC	RMSE
☑	Log-Normal	1759.99	1764.71	983.18
☑	Generalized Extreme Value	1761.00	1768.00	1054.72
☑	Log-Pear	1761.58	1768.57	1236.70
☑	Pearson T	1762.26	1769.25	1536.35
☑	Gumbel (	1759.96	1764.67	1706.29
☑	Gamma	1760.78	1765.49	1754.82
☑	Normal	1784.64	1789.35	3282.22

- Three “goodness-of-fit” measures to assist with model selection:
  - Akaike Information Criterion (AIC)
  - Bayesian Information Criterion (BIC)
  - Root Mean Square Error (RMSE)

Graphical Comparisons



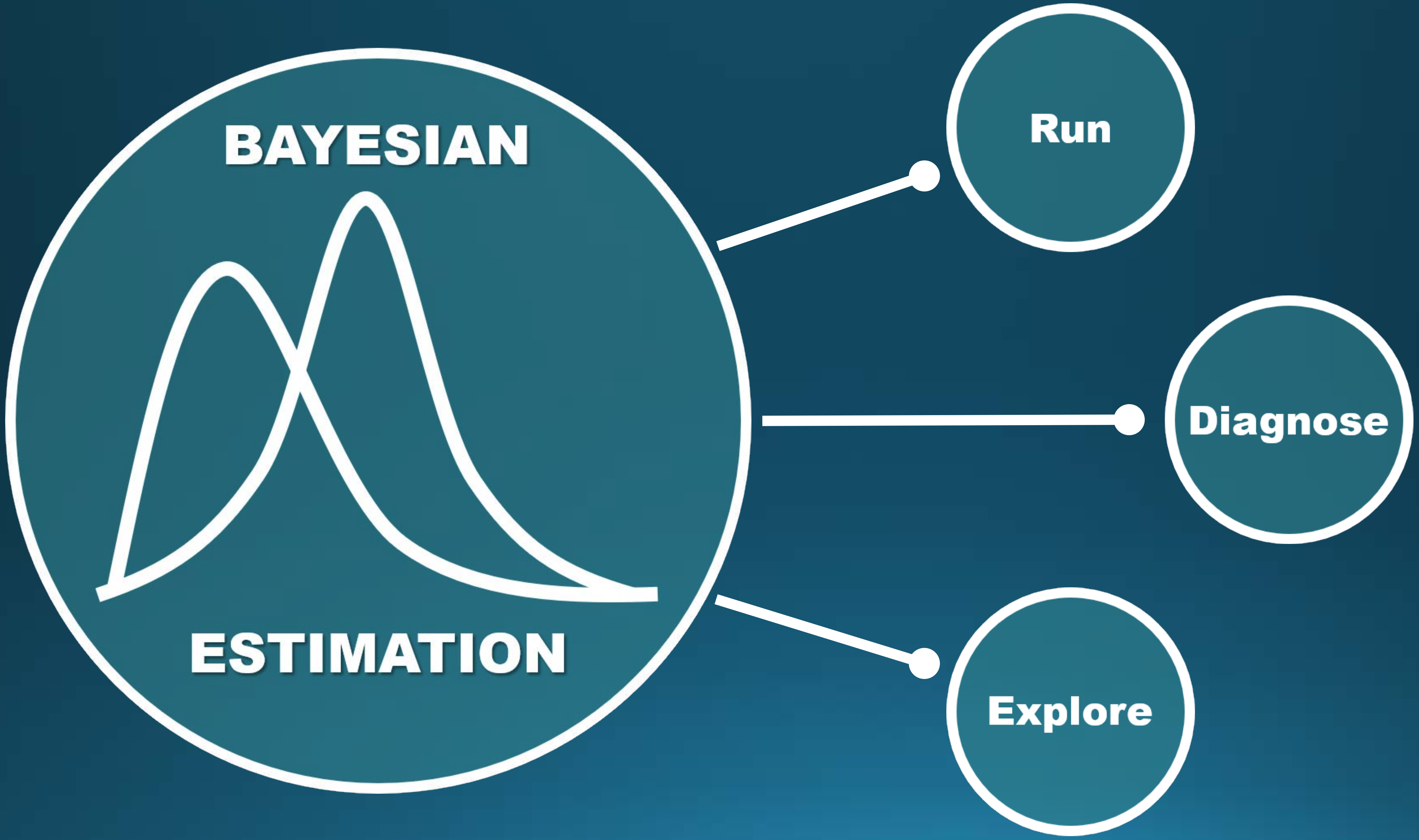
Summary Statistics

Sys, Hist, & Paleo | Sys, Hist, & Paleo - Fit

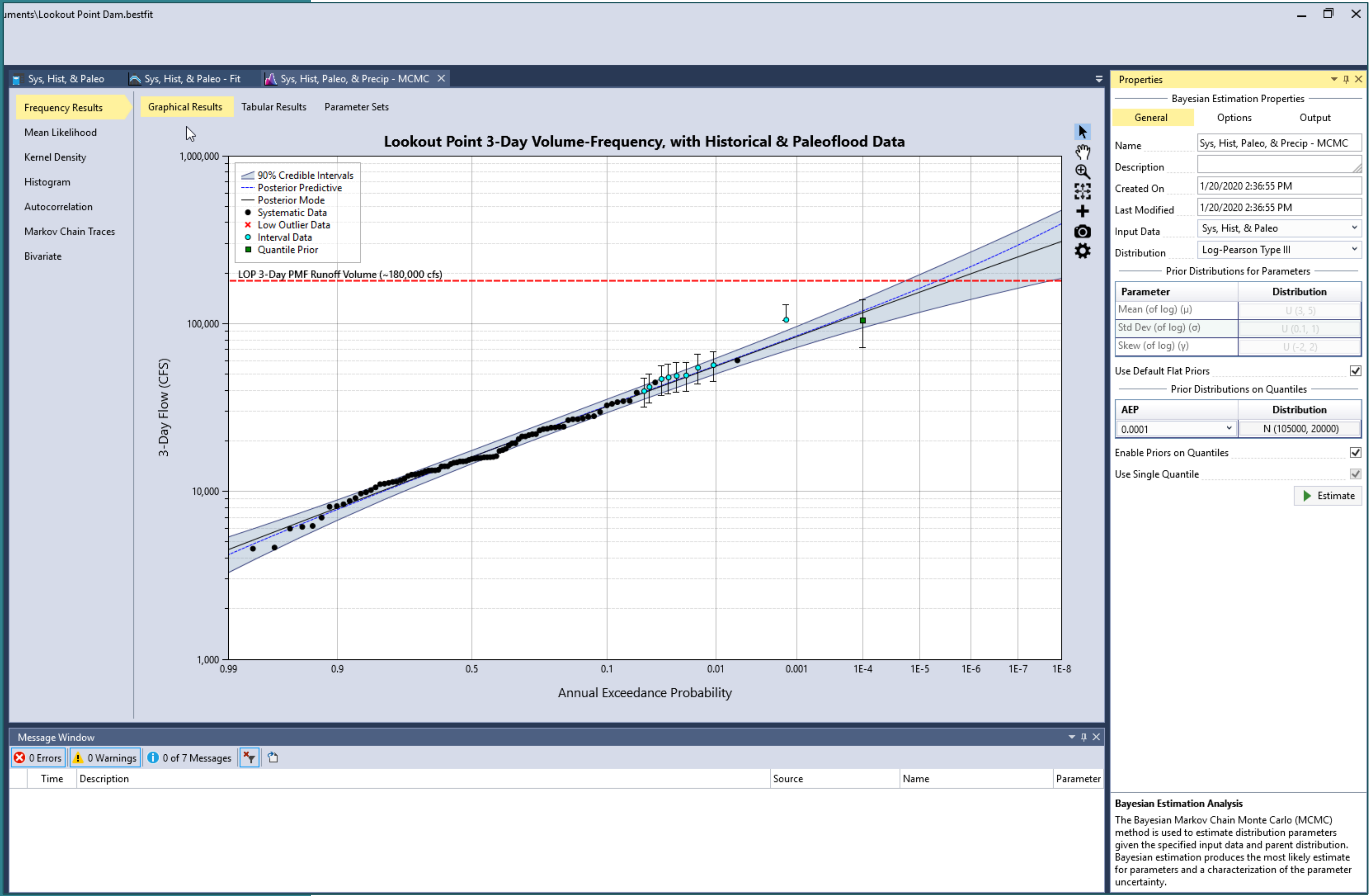
Graphical Results | Tabular Results

Measure	Gamma	Generalized Extreme Value	Gumbel (Extreme Value Type I)	Log-Normal	Log-Pearson Type III	Normal	Pearson Type III
Location	N/A	13,556.1857	14,343.3002	4.2064	4.2062	18,654.7304	18,835.8351
Scale	5,921.3132	6,906.4194	7,837.8315	0.2386	0.2382	12,640.2956	10,840.7112
Shape	3.2068	-0.1470	N/A	N/A	0.0203	N/A	1.3859
Minimum	0	-33,414	-∞	0	0	-∞	3,192
Maximum	∞	∞	∞	∞	∞	∞	∞
Mean	18,988	18,708	18,867	17,173	18,697	18,655	18,836
Std Dev	10,604	11,252	10,052	6,427	11,148	12,640	10,841
Skewness	1.1169	2.4858	1.1396	1.1752	2.0412	0.0000	1.3859
Kurtosis	4.8710	15.5189	5.4000	5.5522	11.2804	3.0000	5.8812
1E-06	116,075	324,722	122,627	219,055	226,927	78,739	130,159
2E-06	111,469	290,020	117,194	202,610	209,399	76,944	124,609
5E-06	105,346	249,252	110,012	182,105	187,627	74,489	117,248
1E-05	100,685	[illegible]	104,580	167,490	172,168	72,564	111,658
2E-05	95,995	[illegible]	99,147	153,614	157,540	70,574	106,048
5E-05	89,746	[illegible]	91,965	136,358	139,421	67,833	98,597
0.0001	84,976	148,544	86,532	124,092	126,591	65,664	92,929
0.0002	80,165	130,911	81,099	112,473	114,481	63,402	87,231
0.0005	73,731	110,195	73,916	98,063	99,520	60,248	79,645
0.001	68,797	96,275	68,481	87,845	88,955	57,716	73,858
0.002	63,795	83,700	63,045	78,183	79,001	55,036	68,020
0.005	57,053	68,916	55,851	66,218	66,725	51,214	60,208
0.01	51,831	58,966	50,398	57,737	58,060	48,060	54,208
0.02	46,472	49,953	44,926	49,706	49,887	44,615	48,107
0.05	39,100	39,279	37,623	39,705	39,753	39,446	39,829
0.1	33,207	31,978	31,981	32,521	32,507	34,854	33,330
0.2	26,870	25,147	26,100	25,539	25,494	29,293	26,502
0.3	22,824	21,244	22,424	21,454	21,406	25,283	22,257
0.5	17,055	16,157	17,216	16,084	16,047	18,655	16,416
0.7	12,360	12,292	12,888	12,058	12,043	12,026	11,927
0.8	10,003	10,382	10,613	10,129	10,128	8,016	9,804
0.9	7,285	8,135	7,806	7,954	7,970	2,456	7,511
0.95	5,480	6,558	5,744	6,515	6,543	-2,137	6,114

Generalized Extreme Value
Location (ξ) = 13,556.1857
Scale (α) = 6,906.4194
Shape (κ) = -0.1470

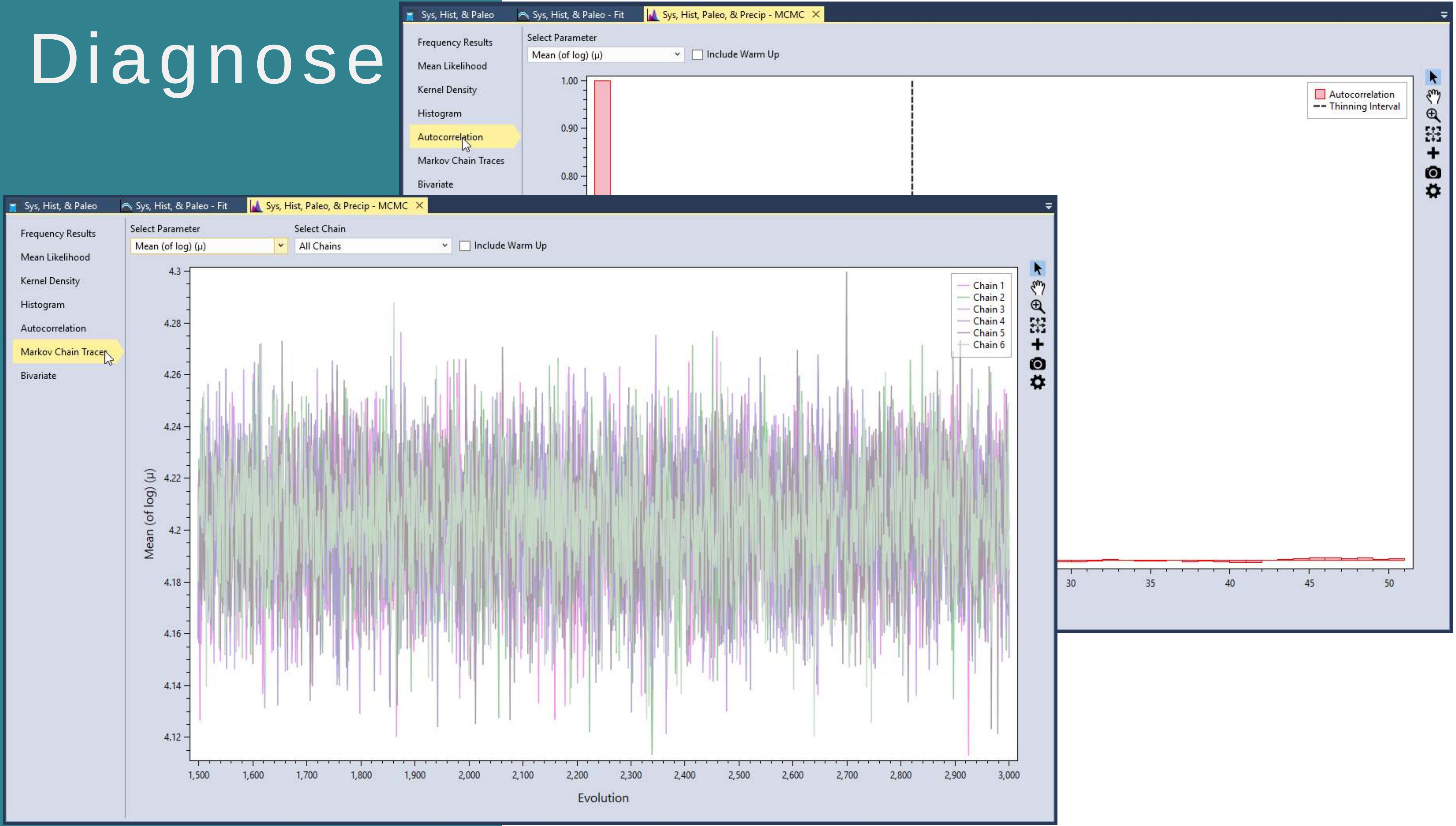

BAYESIAN
ESTIMATION
Run
Diagnose
Explore

Run

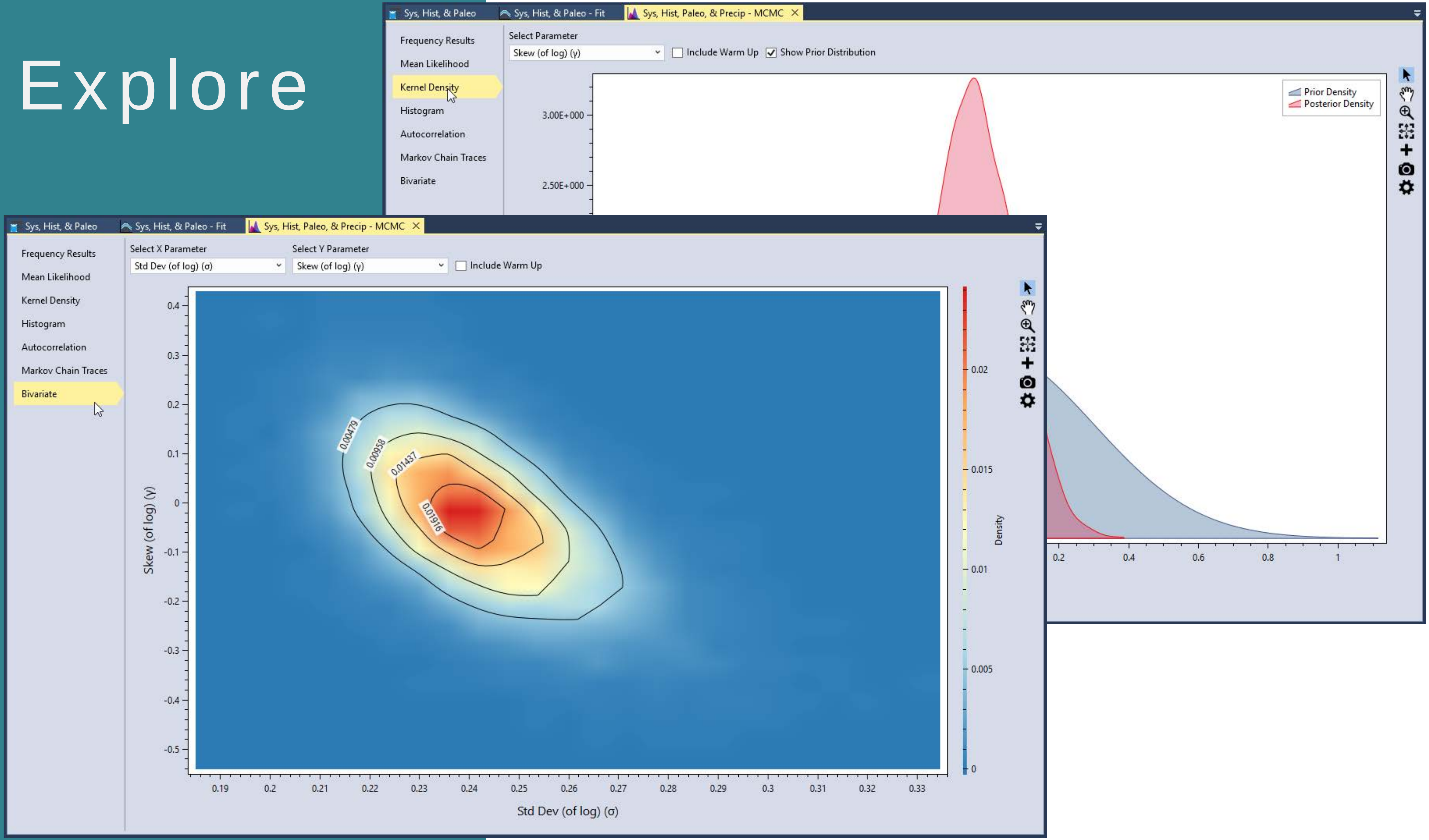


uments\Lookout Point Dam.bestfit
Sys, Hist, & Paleo
Sys, Hist, & Paleo - Fit
Sys, Hist, Paleo, & Precip - MCMC
Frequency Results
Mean Likelihood
Kernel Density
Histogram
Autocorrelation
Markov Chain Traces
Bivariate
Graphical Results
Tabular Results
Parameter Sets
Lookout Point 3-Day Volume-Frequency, with Historical & Paleoflood Data
90% Credible Intervals
Posterior Predictive
Posterior Mode
Systematic Data
Low Outlier Data
Interval Data
Quantile Prior
LOP 3-Day PMF Runoff Volume (~180,000 cfs)
3-Day Flow (CFS)
Annual Exceedance Probability
Properties
Bayesian Estimation Properties
General
Options
Output
Name
Sys, Hist, Paleo, & Precip - MCMC
Description
Created On
1/20/2020 2:36:55 PM
Last Modified
1/20/2020 2:36:55 PM
Input Data
Sys, Hist, & Paleo
Distribution
Log-Pearson Type III
Prior Distributions for Parameters
Parameter
Distribution
Mean (of log) (μ)
Std Dev (of log) (σ)
Skew (of log) (γ)
Use Default Flat Priors
Prior Distributions on Quantiles
AEP
0.0001
N (105000, 20000)
Enable Priors on Quantiles
Use Single Quantile
Estimate
Bayesian Estimation Analysis
The Bayesian Markov Chain Monte Carlo (MCMC) method is used to estimate distribution parameters given the specified input data and parent distribution. Bayesian estimation produces the most likely estimate for parameters and a characterization of the parameter uncertainty.
Message Window
0 Errors
0 Warnings
0 of 7 Messages
Time
Description
Source
Name
Parameter

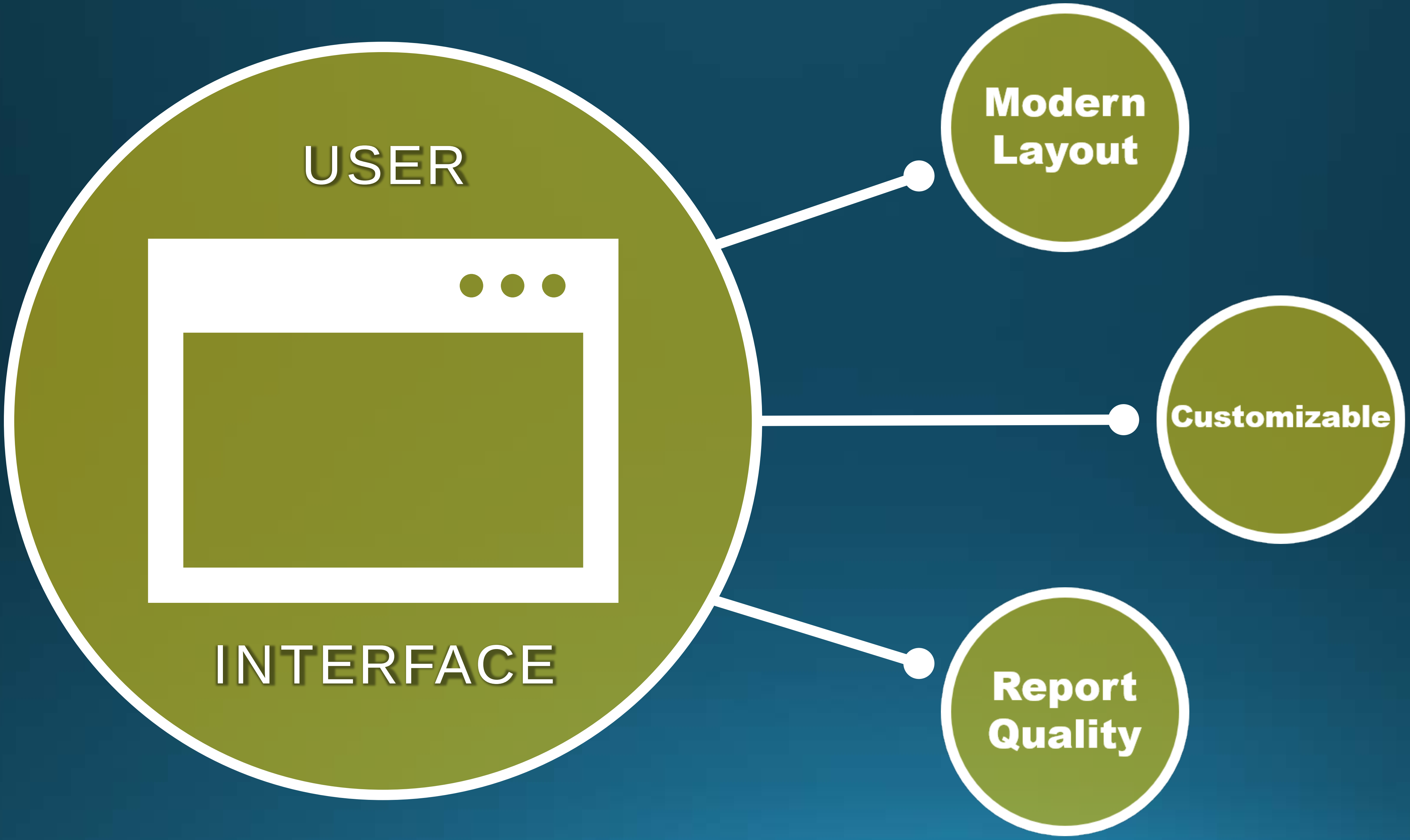
Diagnose


Sys, Hist, & Paleo
Sys, Hist, & Paleo - Fit
Sys, Hist, Paleo, & Precip - MCMC
Frequency Results
Mean Likelihood
Kernel Density
Histogram
Autocorrelation
Markov Chain Traces
Bivariate
Select Parameter
Mean (of log) (μ)
Include Warm Up
Autocorrelation
Thinning Interval
1.00
0.90
0.80
30
35
40
45
50
Sys, Hist, & Paleo
Sys, Hist, & Paleo - Fit
Sys, Hist, Paleo, & Precip - MCMC
Frequency Results
Mean Likelihood
Kernel Density
Histogram
Autocorrelation
Markov Chain Traces
Bivariate
Select Parameter
Mean (of log) (μ)
Select Chain
All Chains
Include Warm Up
Chain 1
Chain 2
Chain 3
Chain 4
Chain 5
Chain 6
Mean (of log) (μ)
4.3
4.28
4.26
4.24
4.22
4.2
4.18
4.16
4.14
4.12
1,500
1,600
1,700
1,800
1,900
2,000
2,100
2,200
2,300
2,400
2,500
2,600
2,700
2,800
2,900
3,000
Evolution

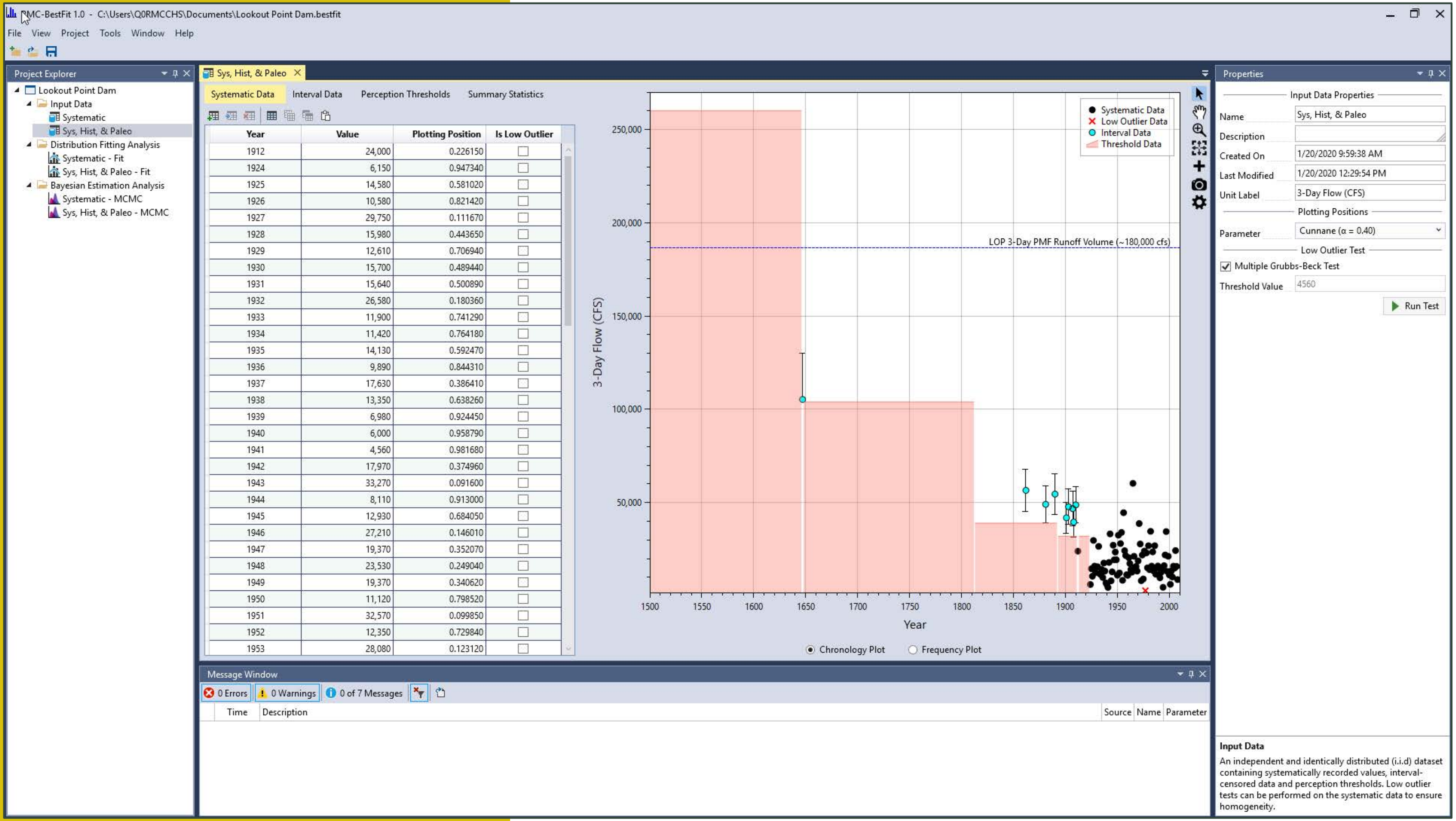
Explore



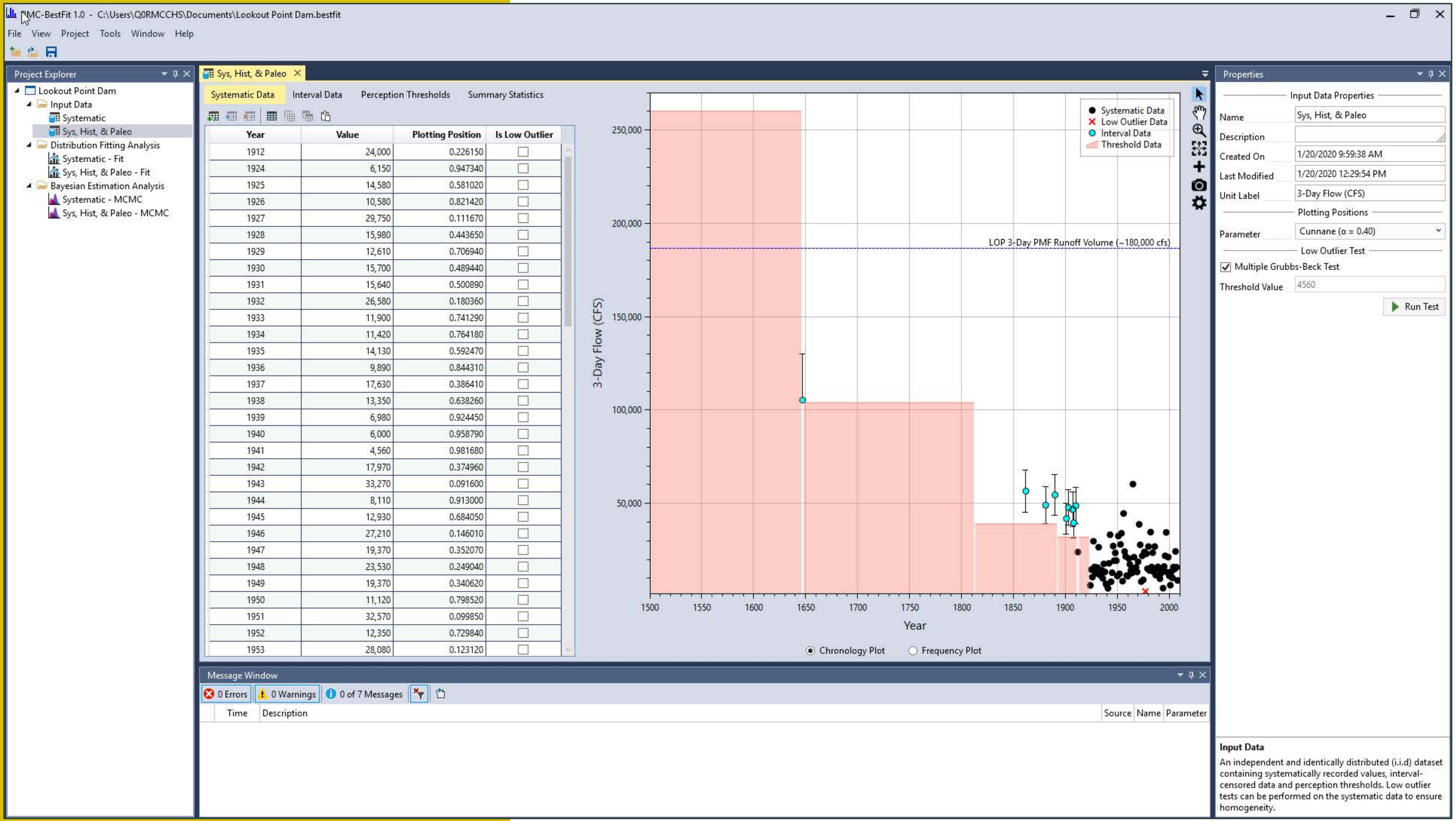
Sys, Hist, & Paleo
Sys, Hist, & Paleo - Fit
Sys, Hist, Paleo, & Precip - MCMC
Frequency Results
Mean Likelihood
Kernel Density
Histogram
Autocorrelation
Markov Chain Traces
Bivariate
Select Parameter
Skew (of log) (γ)
Include Warm Up
Show Prior Distribution
Prior Density
Posterior Density
Select X Parameter
Std Dev (of log) (σ)
Select Y Parameter
Skew (of log) (γ)
Include Warm Up
Skew (of log) (γ)
Std Dev (of log) (σ)
Density


USER
INTERFACE
Modern Layout
Customizable
Report Quality

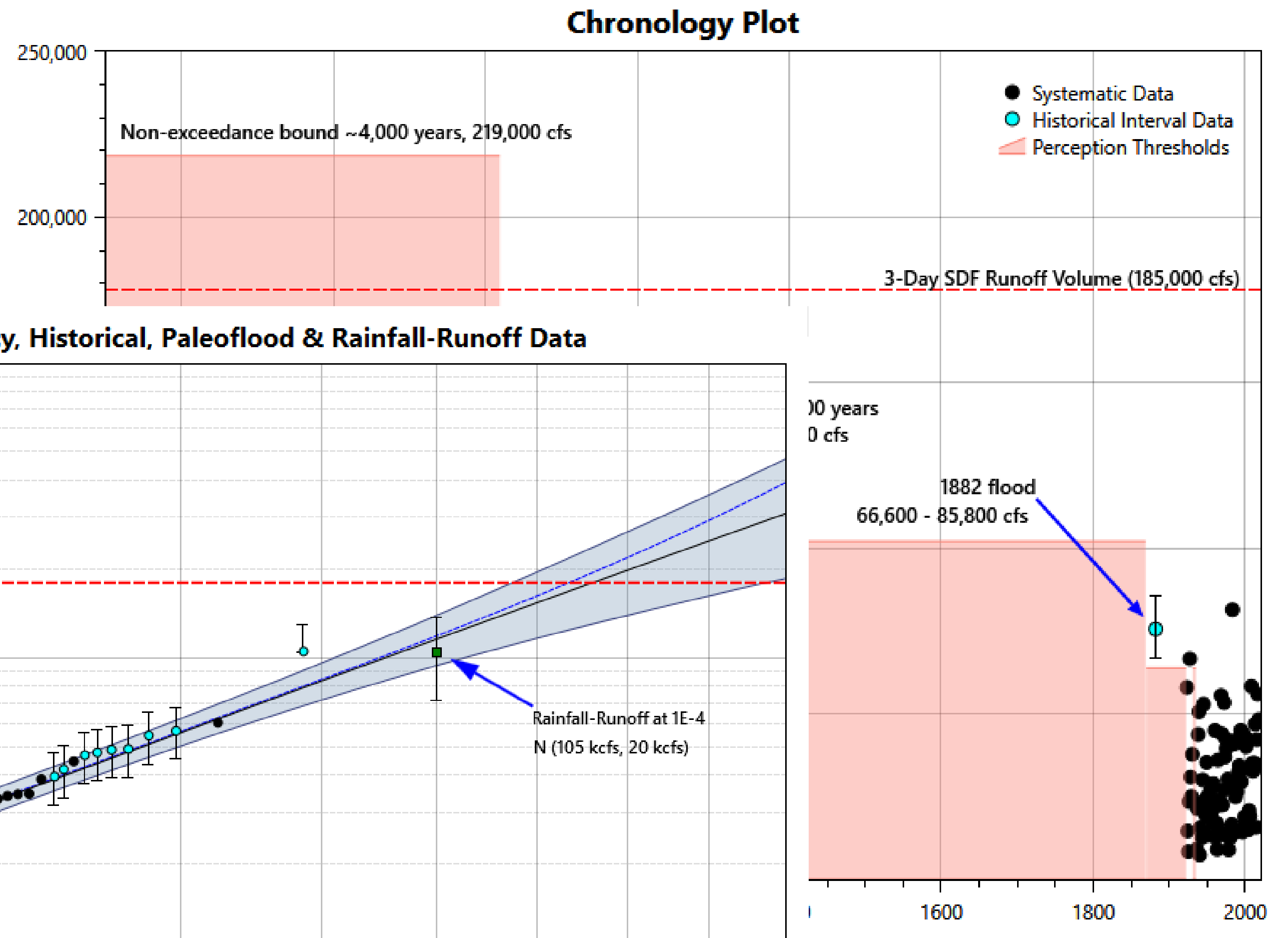
Modern Layout

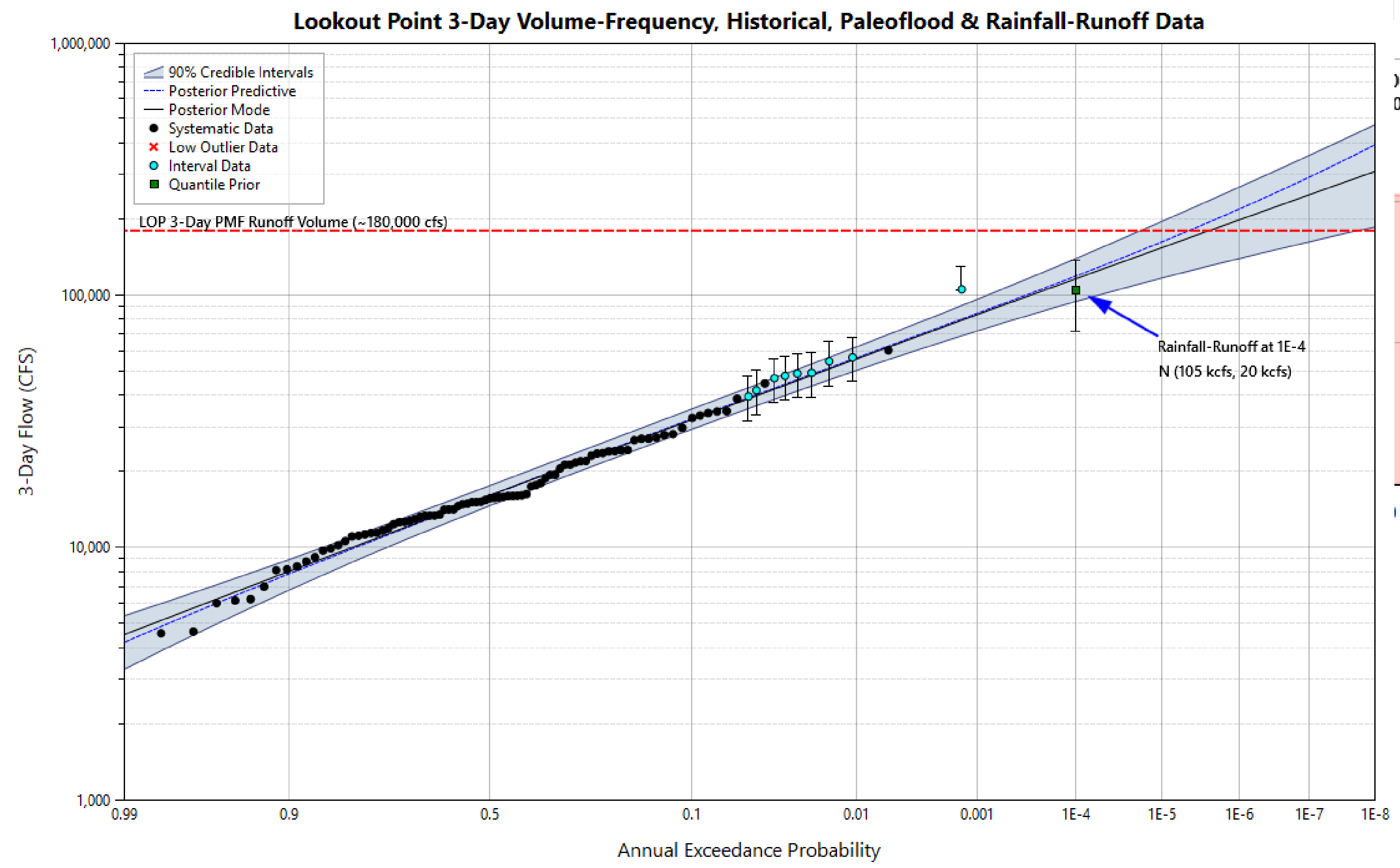


Customizable

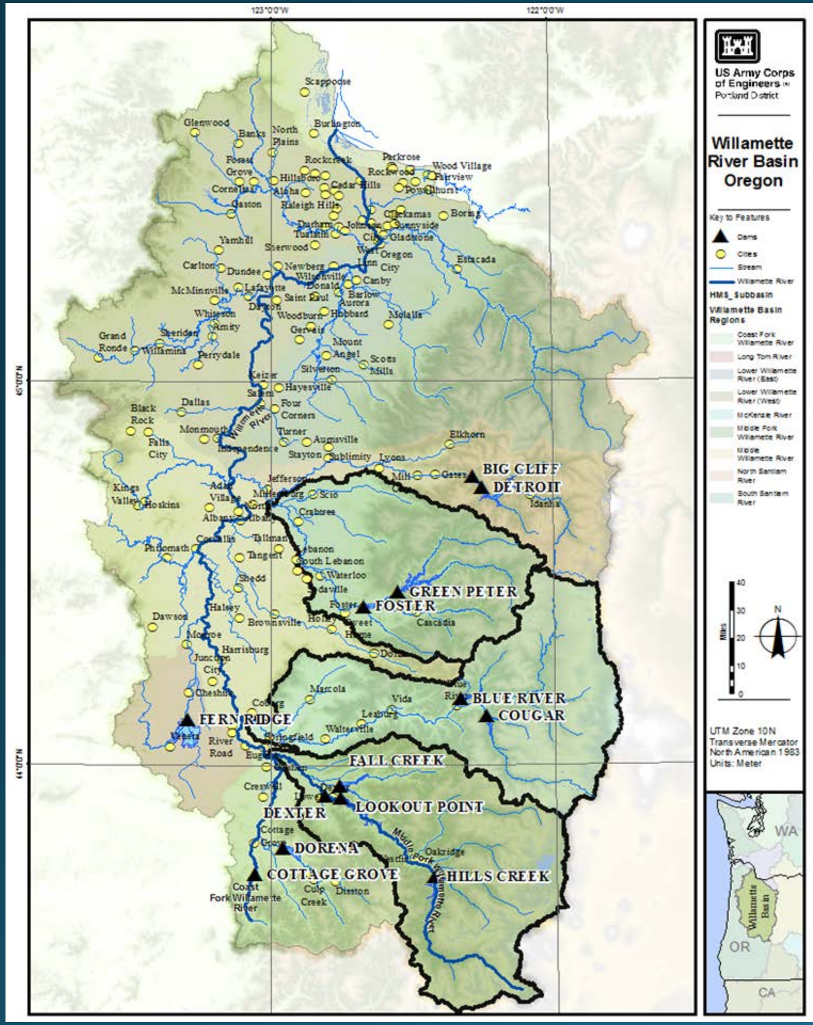


Report Quality



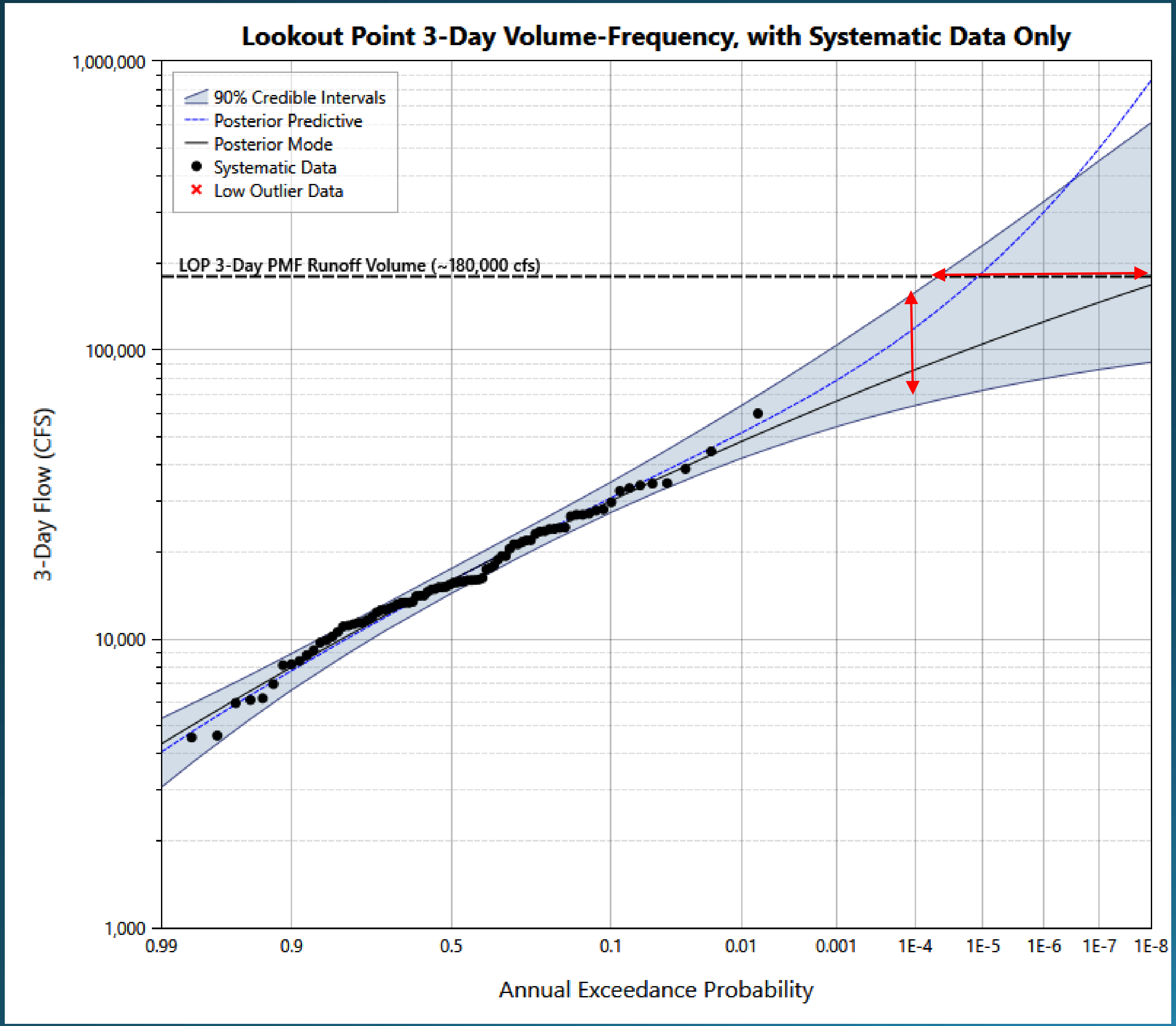


Case Study: Lookout Point Dam



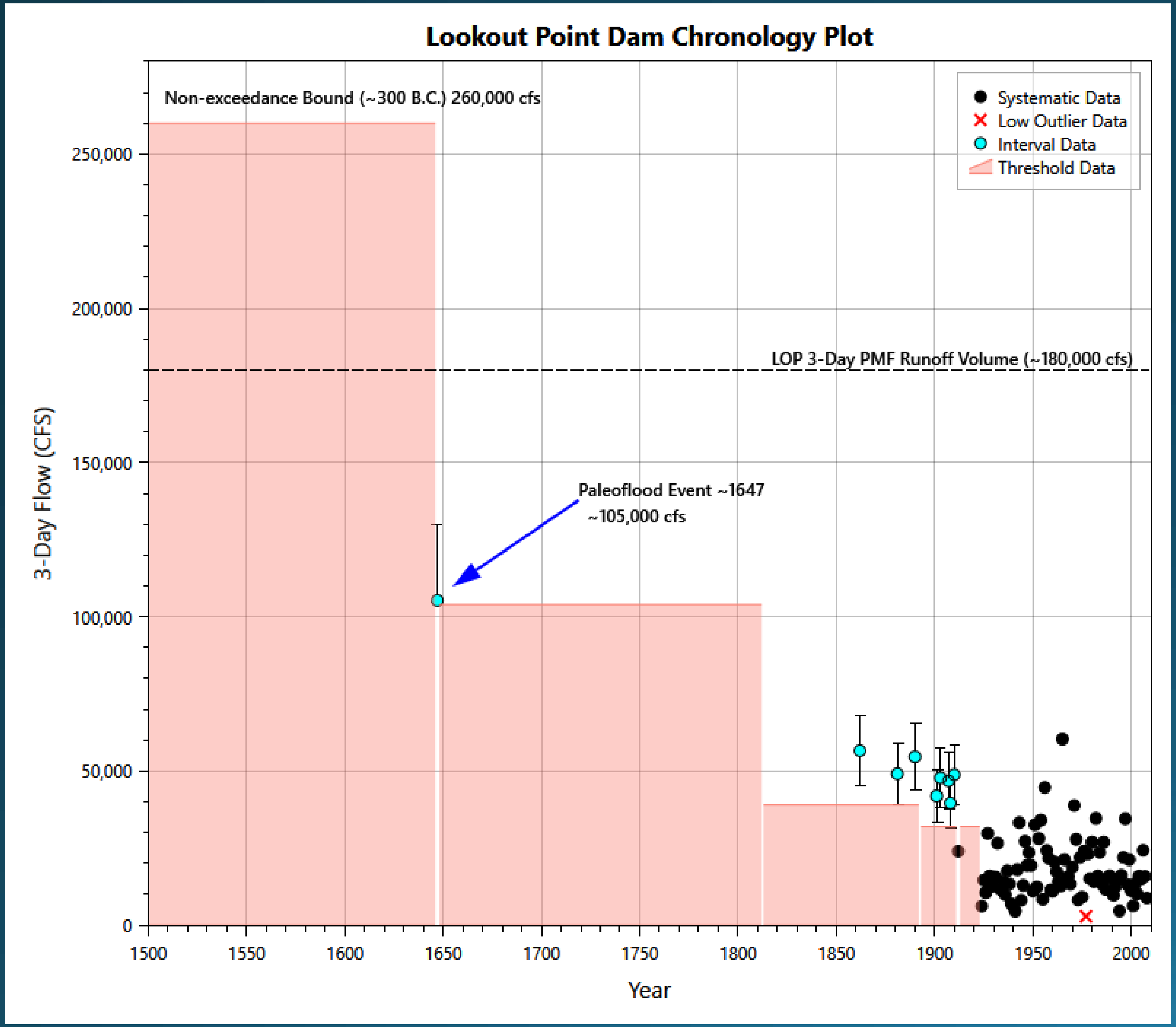
- Willamette River Basin (Oregon, USA)
  - 11,500 mi²
- Contains several high priority dams
  - Blue River
  - Cougar
  - Fall Creek
  - Foster
  - Green Peter
  - Hills Creek
  - **Lookout Point**
    - **996 mi²**
- Portland, OR downstream
- Dams operate as a complex system

Systematic Data



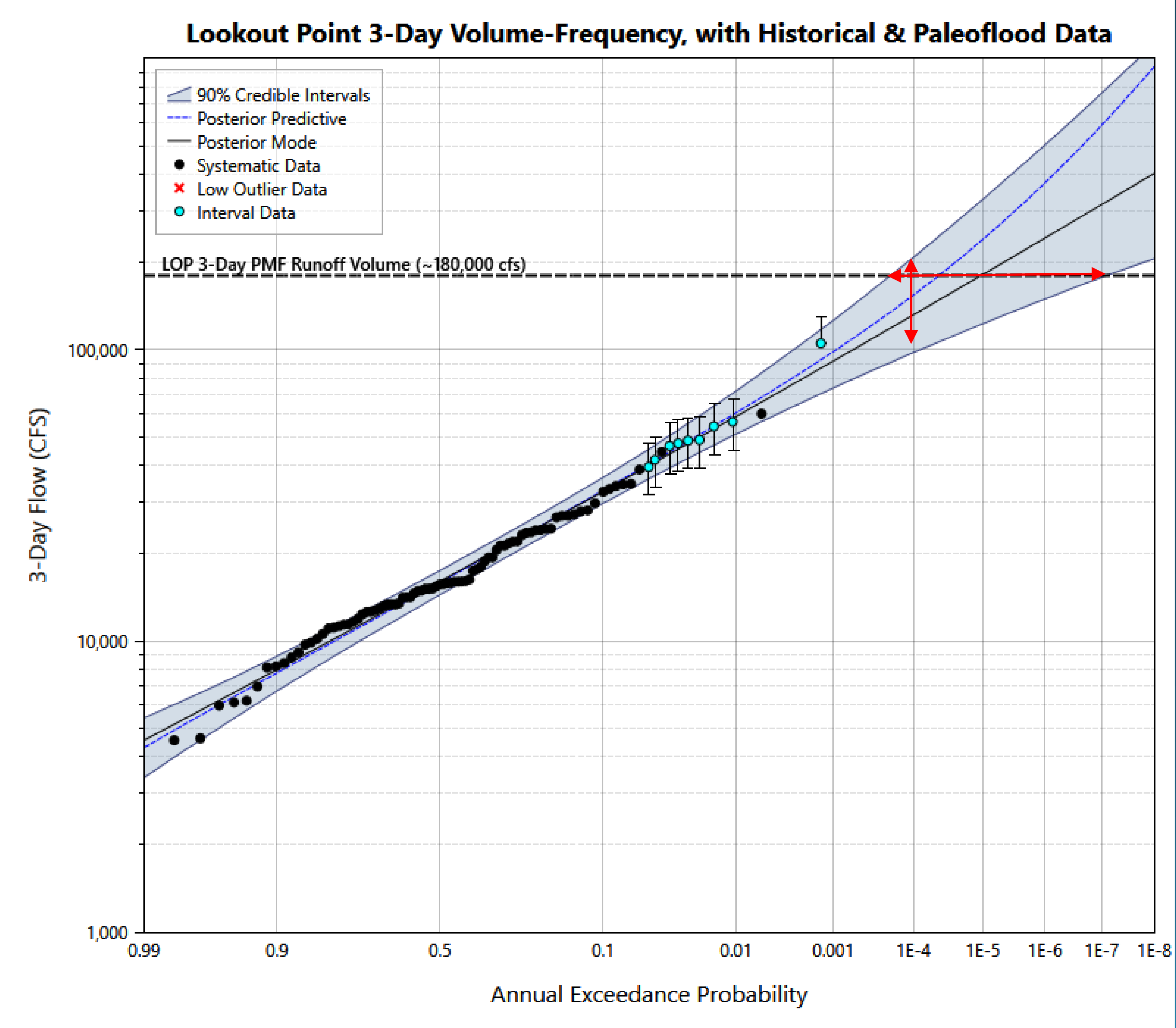
- Large uncertainty in the quantile estimate for the 1:10,000 (1E-4) AEP
- Very large uncertainty in the estimated AEP for the PMF
  - Well over 4 orders of magnitude of uncertainty

Temporal Information Expansion



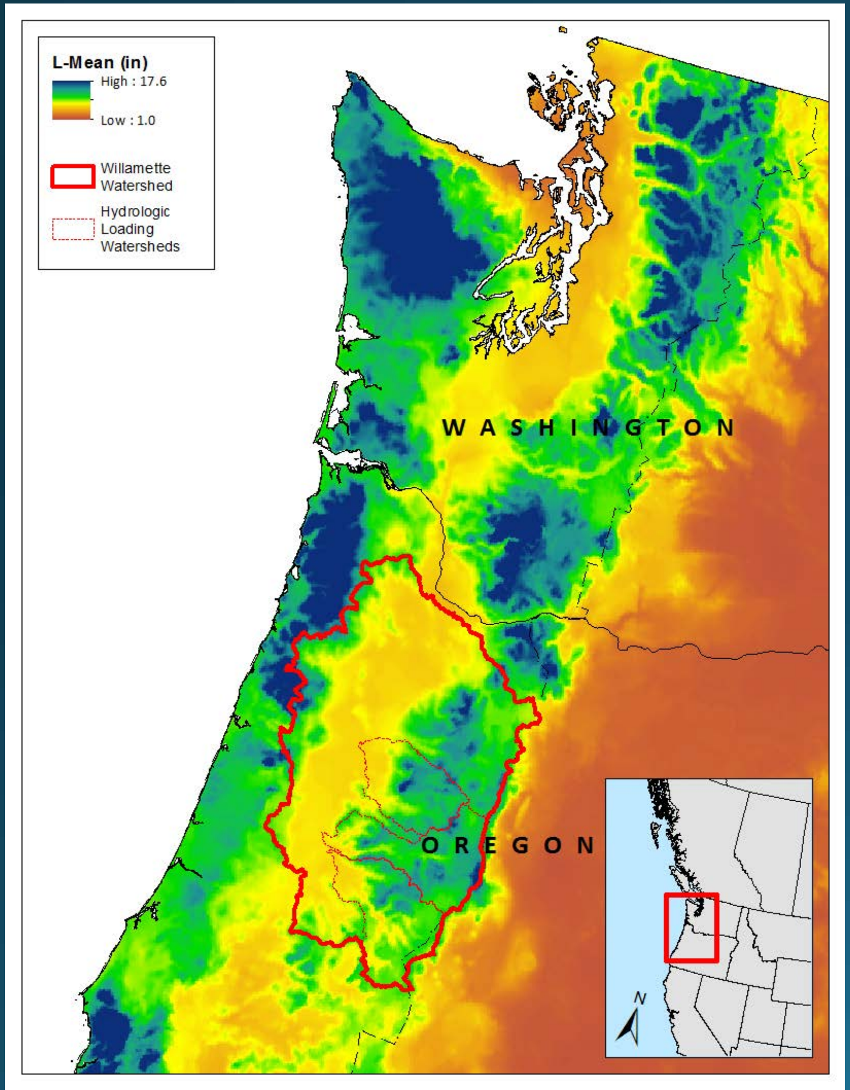
- Flood Interval
  - A paleoflood event took place approximately 370 years ago that produced a 3-day flow of approximately 105,000 cfs
- Perception Threshold
  - A 3-day flow of approximately 260,000 cfs has not been exceeded (non-exceedance bound) in the last 2,300 years.

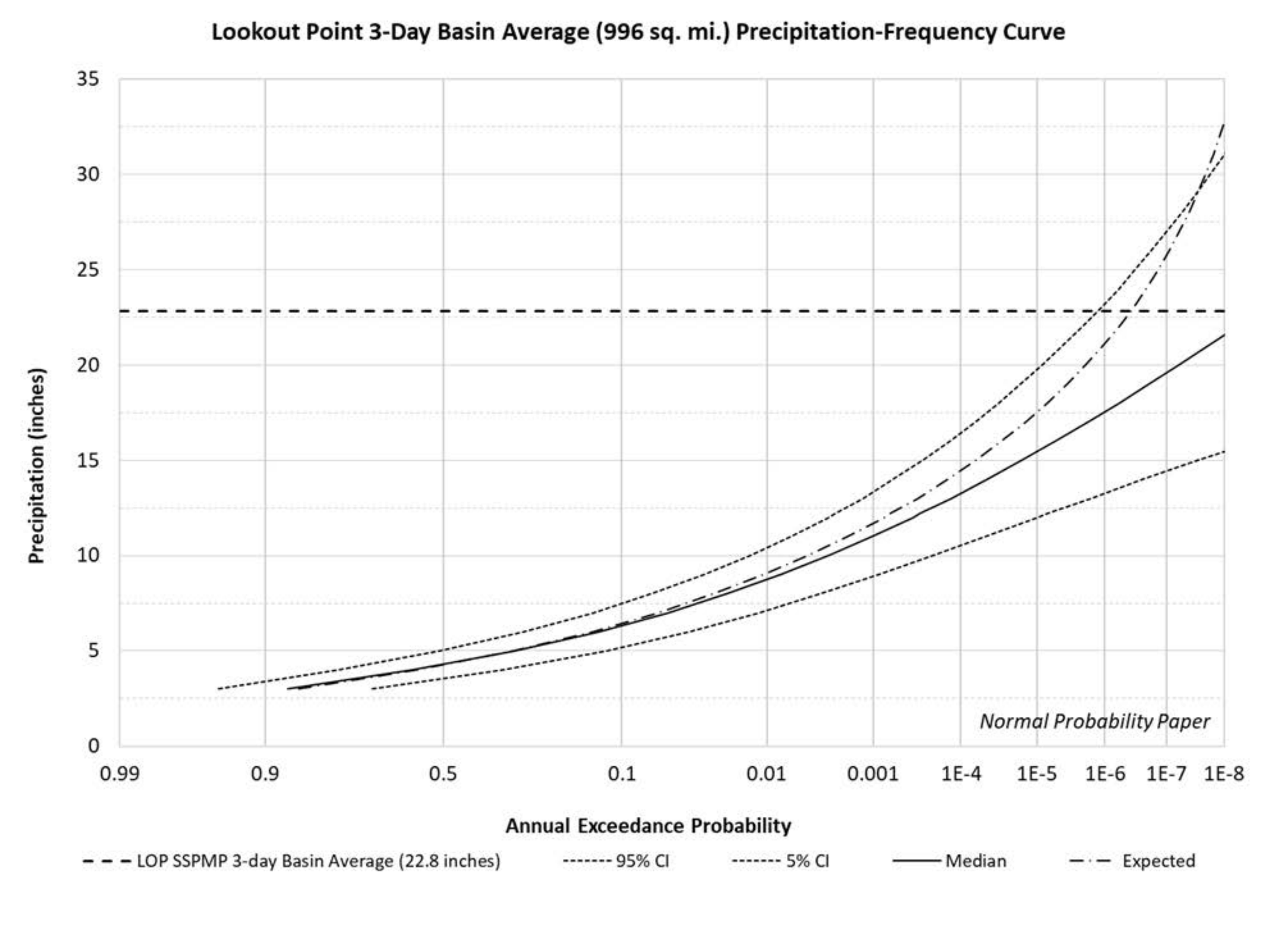
Temporal Information Expansion



- A minor reduction in uncertainty in the quantile estimate for the 1:10,000 (1E-4) AEP
  - Paleoflood increased our perception of the natural variability
- A reduction in uncertainty in the estimated AEP for the PMF
  - still over 3 orders of magnitude

Spatial & Causal Information Expansion





Spatial & Causal Information Expansion

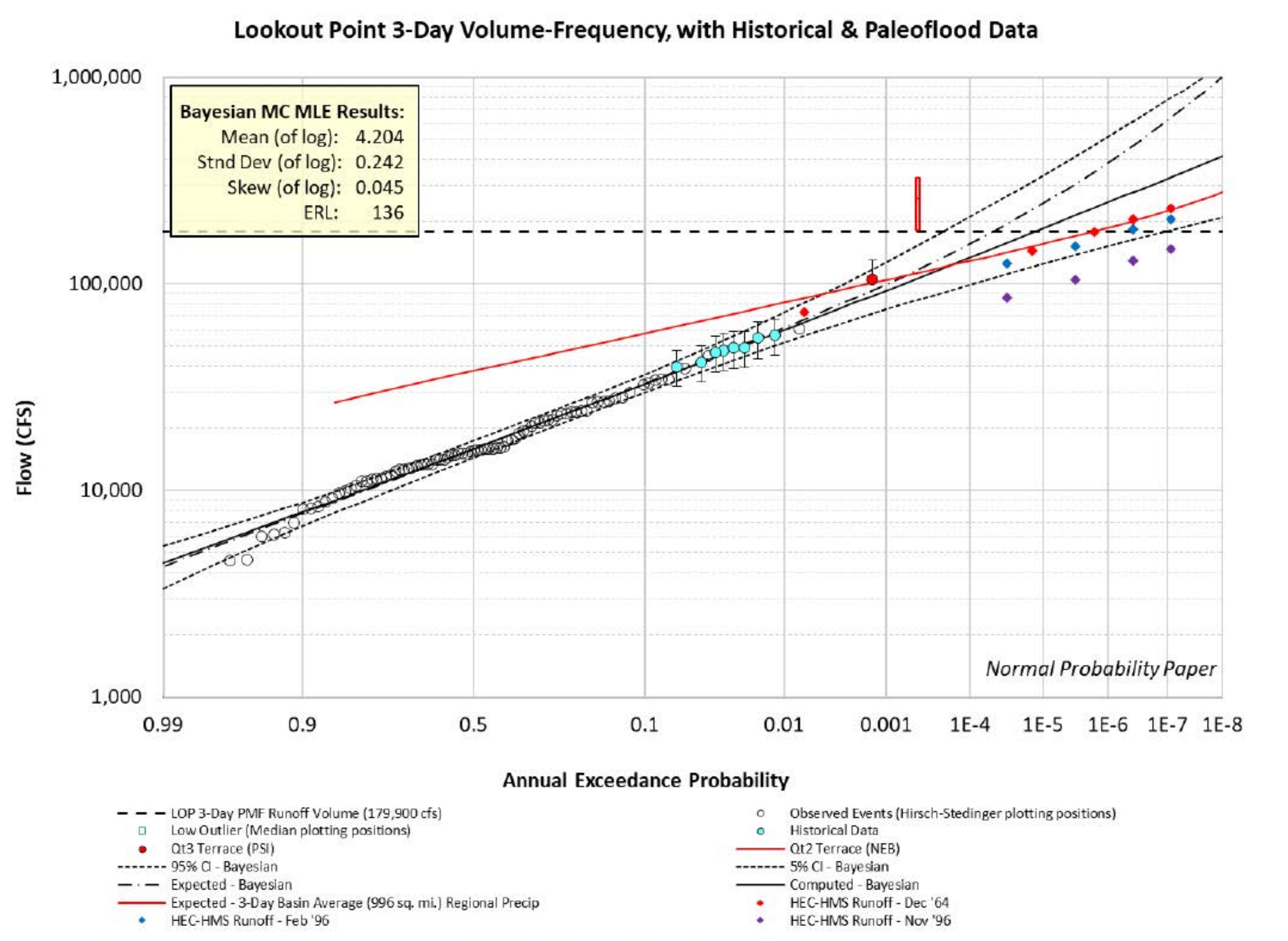
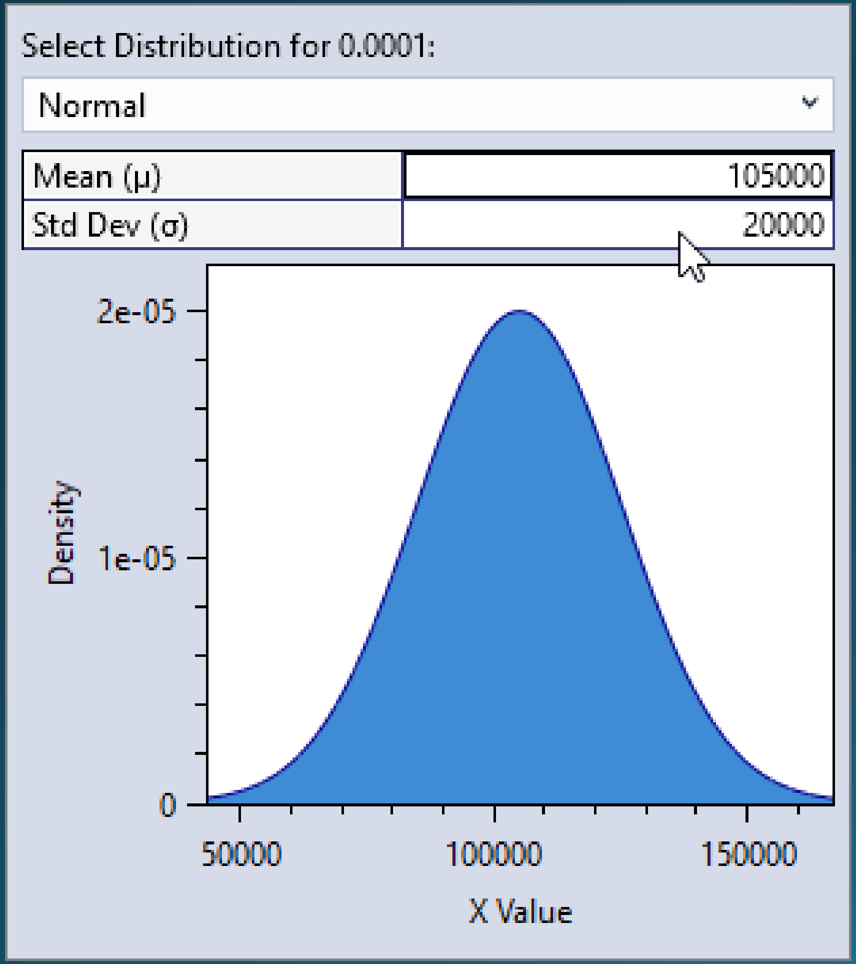


Figure 4-12: Lookout Point Dam HEC-HMS 3-Day Runoff from Precipitation-Frequency Events

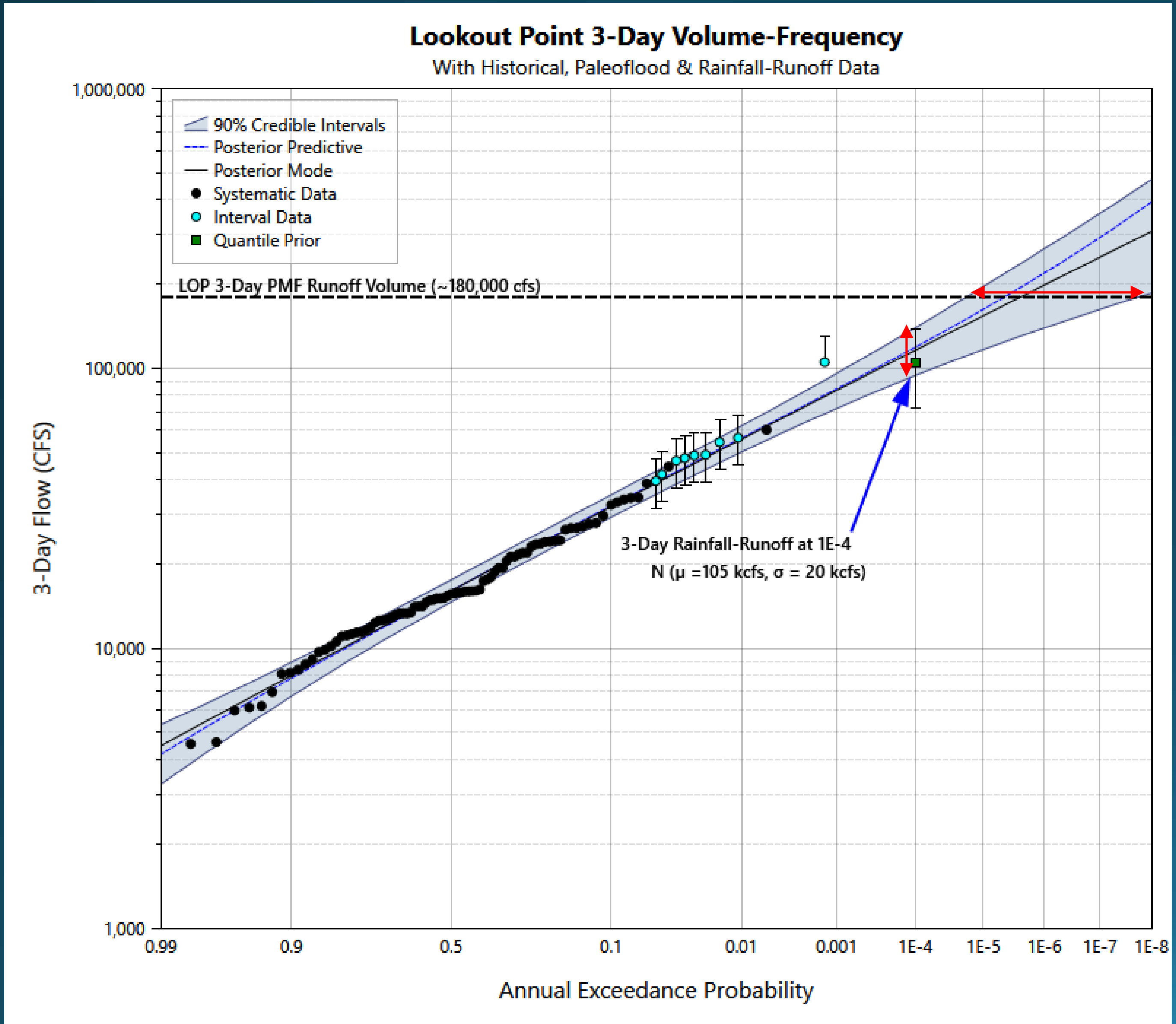
- A regional rainfall-frequency analysis was performed
- Rainfall-frequency events were routed with HEC-HMS
- Results suggest much rarer AEPs for the PMF

Spatial & Causal Information Expansion



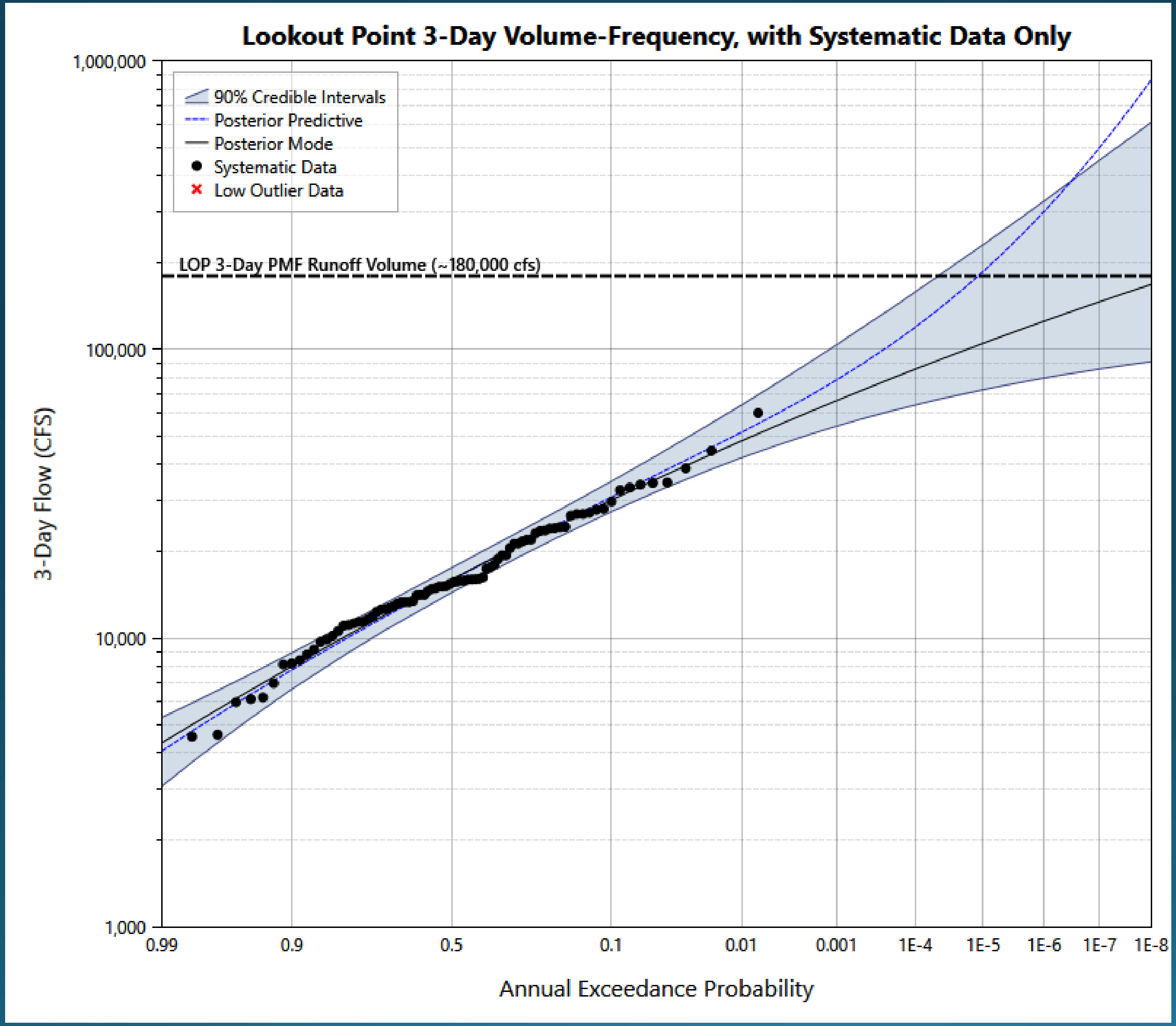
- Rainfall-Runoff at AEP of 1E-4
  - Normally distributed
  - Mean of 105,000 cfs
  - Standard Deviation of 20,000 cfs

Spatial & Causal Information Expansion

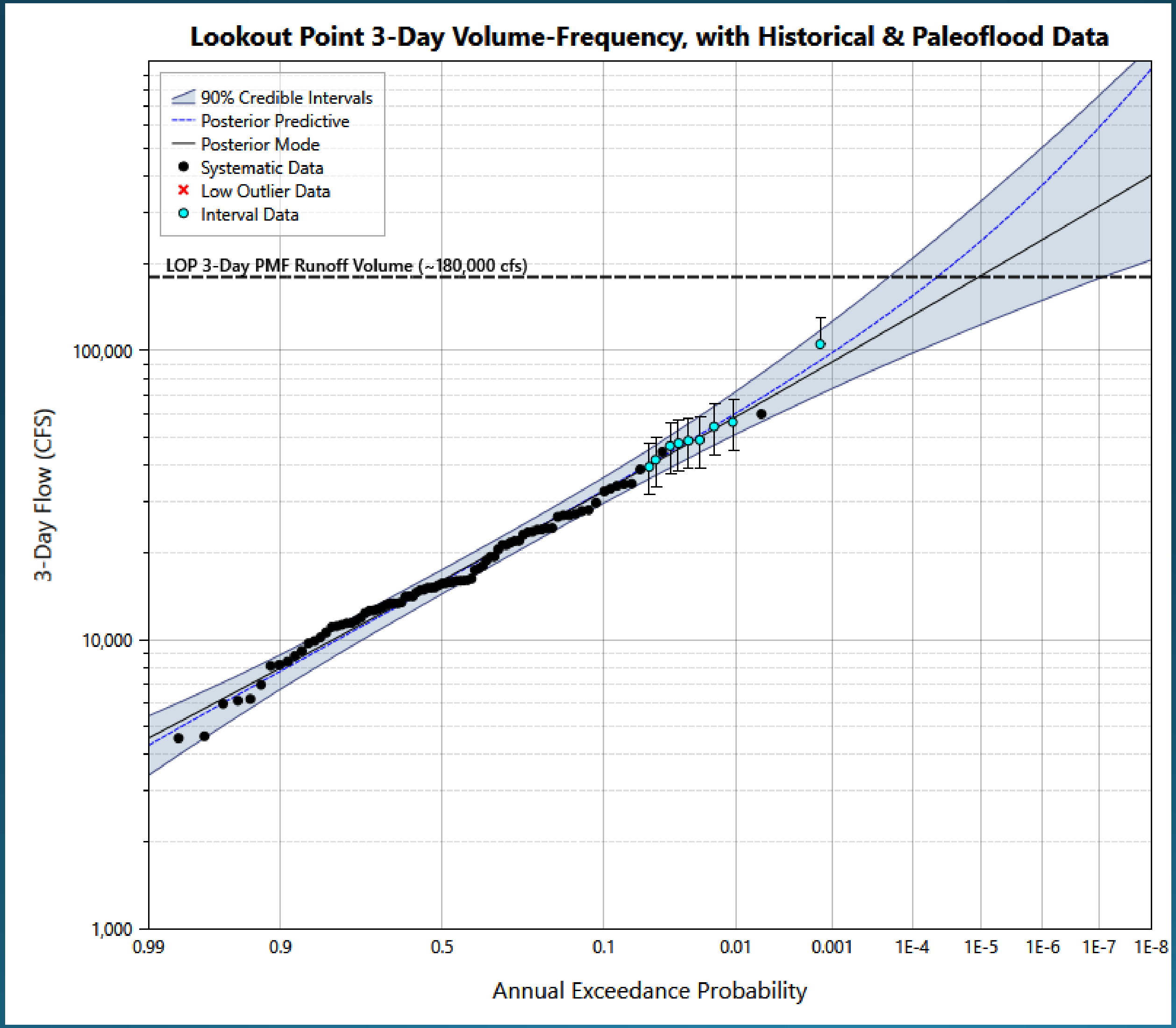


- A major reduction in uncertainty in the quantile estimate for the 1:10,000 (1E-4) AEP
- A sizeable reduction in uncertainty in the estimated AEP for the PMF
  - ~ 3 orders of magnitude
- The expected and most likely curves are much closer together

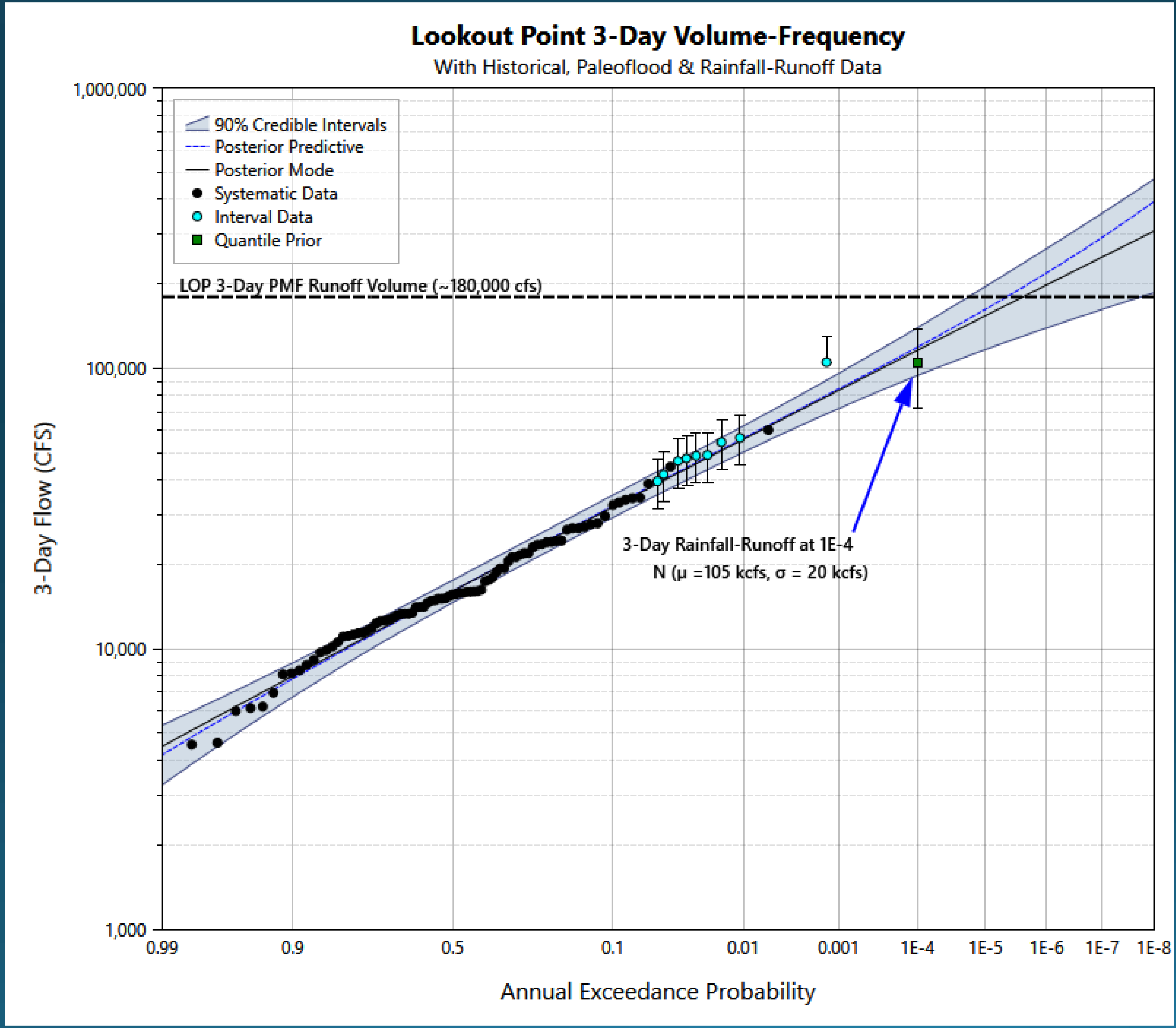
Systematic Data



Temporal Information Expansion



Spatial & Causal Information Expansion



Comparison to EMA

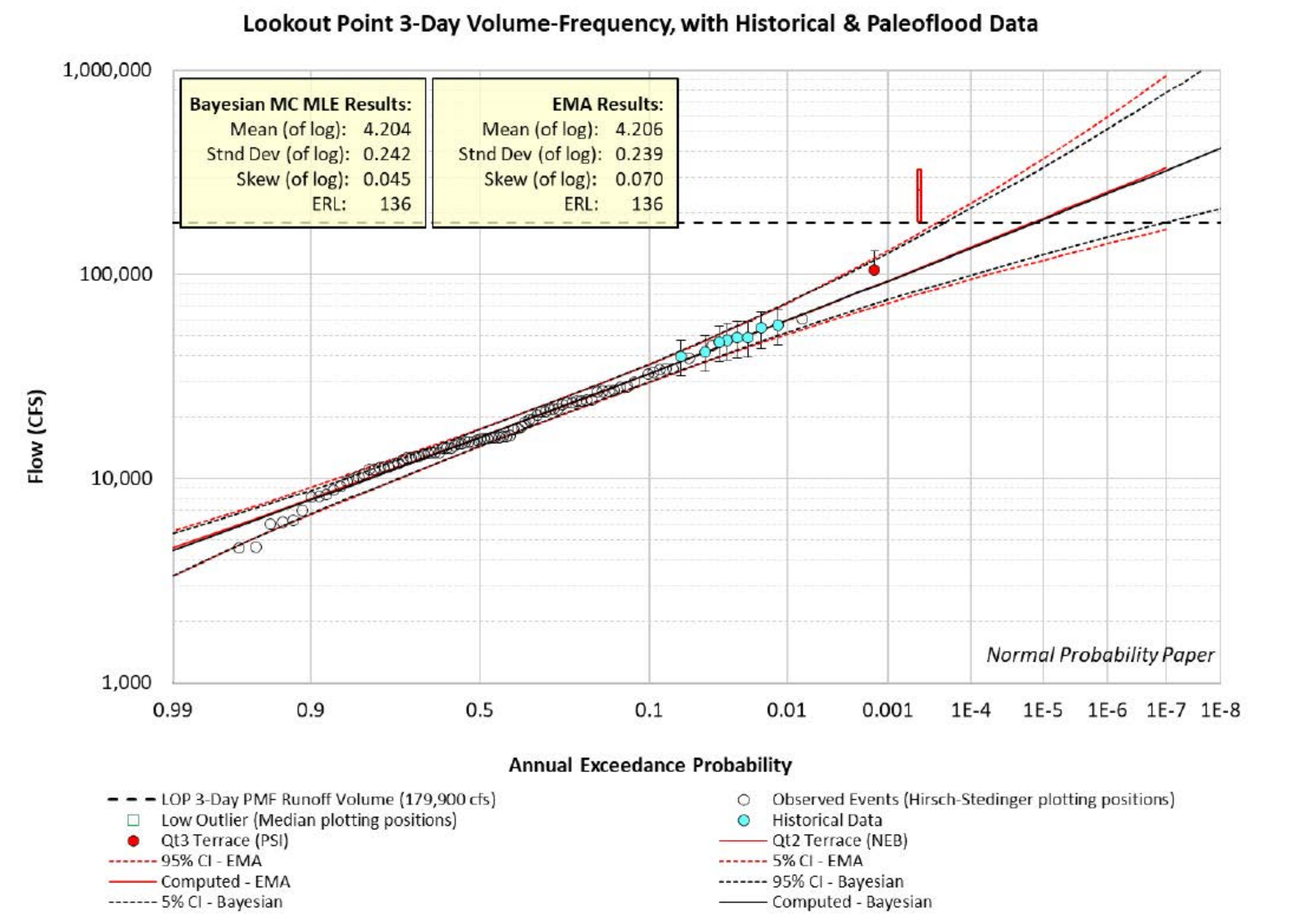


Figure D-3: Lookout Point 3-Day Volume-Frequency Curve Comparison of the Bayesian Method to EMA, with Systematic, Historical and Paleoflood Data

- Bulletin 17C recommends fitting the LPIII distribution using the Expected Moments Algorithm (EMA)
- EMA was developed as an alternative to Maximum Likelihood Estimation (MLE)
- The Bayesian approach is closely related to the MLE method.
- Both methods produce similar results given typical censored data; however, EMA is not capable of incorporating the causal rainfall-runoff information in a formal, probabilistic manner.

Conclusions

- The Bayesian flood frequency approach can incorporate all available sources of hydrologic information, such as paleofloods, regional rainfall-runoff results, and expert elicitation.
- The ability of the Bayesian approach to use all pieces of information in conjunction is a major advantage over other methods, such as EMA, and provides much better estimates of design floods with specified AEPs.
- Complementing systematic flood data with temporal, spatial, and causal information should become the standard procedure for estimating exceedance probabilities for extreme floods.

RMC-BestFit

Bayesian Estimation and Fitting Software