Partial Duration Series and Peaks Over Threshold

Flood Frequency Analysis PROSPECT

May 2022

Gregory S. Karlovits, *P.E., PH, CFM*

Statistical Hydrologist

US Army Corps of Engineers

Hydrologic Engineering Center

US Army Corps of Engineers®

Purpose

- Understand a different approach for estimating flood frequency
- Identify when applying peaks over threshold can be beneficial

US Army Corps of Engineers®

Outline

1. Peaks over threshold approach
2. Models for partial duration series
3. Computing AEP using partial duration series
4. Why PDS?

Peaks Over Threshold Approach

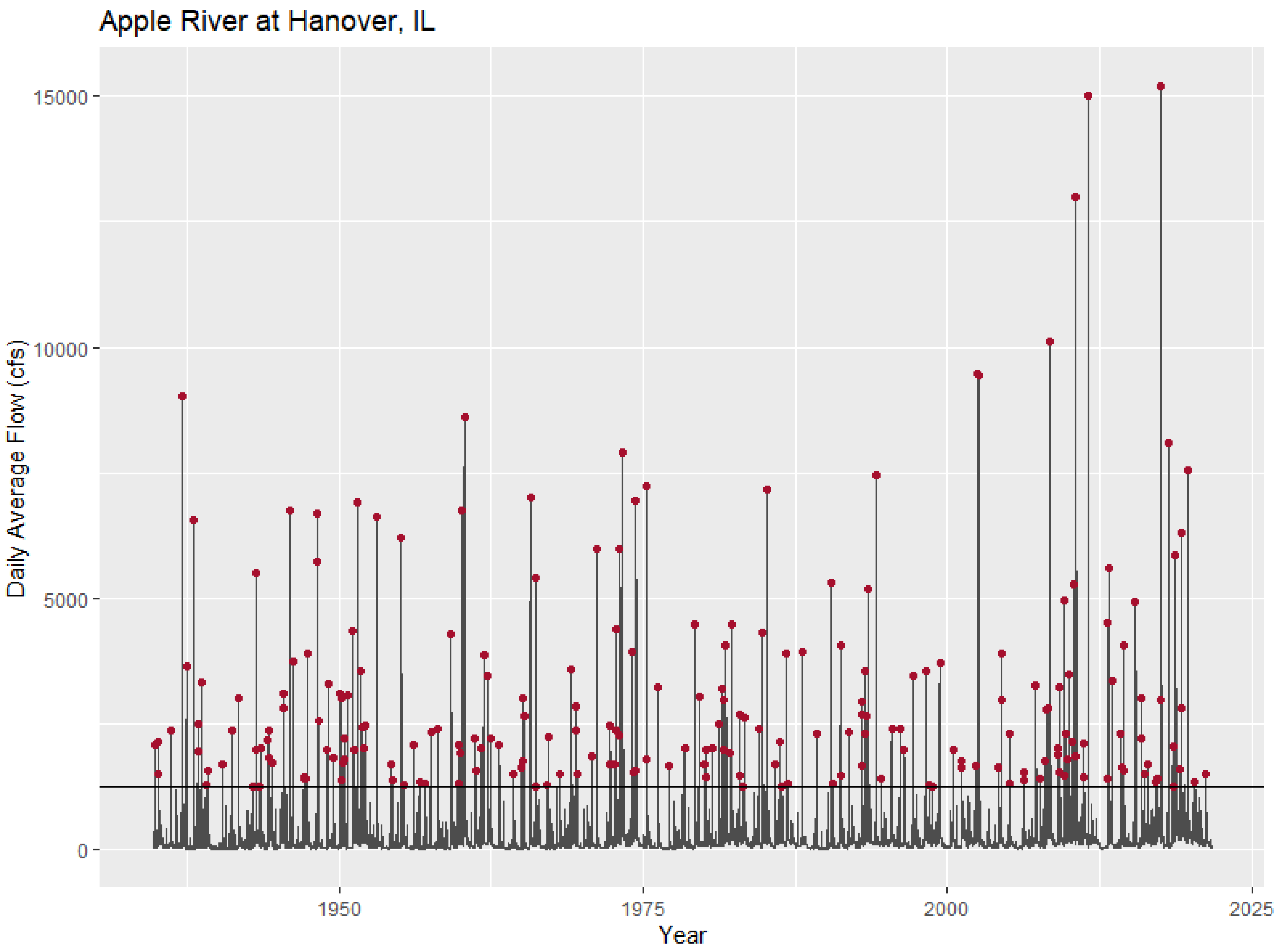
US Army Corps of Engineers®

Peaks Over Threshold

- Choose a high flow value as a threshold
- Keep **all independent** flood peaks greater than the threshold
- Varying number of floods per year
  - 0 or more

US Army Corps
of Engineers®

Peaks Over Threshold



US Army Corps of Engineers®

Partial Duration Series

- The collection of independent peaks over a sufficiently high threshold is called the *partial duration series (PDS)*
- Number of events can be different than in the annual maximum series (AMS)

US Army Corps of Engineers®

Inference Assumptions

- Sample is representative of population
- Sample values are independent
- Sample values are identically-distributed

I.I.D.

Representative Sample

- High threshold ensures sample made only of floods
- Too high of a threshold limits sample size

- Violation: model for population doesn’t actually represent it

US Army Corps
of Engineers®

Identical Distribution

- High threshold ensures sample made only of floods
- May still be a mixture of flood-generating mechanisms

- Violation: population model misrepresents extremes

US Army Corps of Engineers®

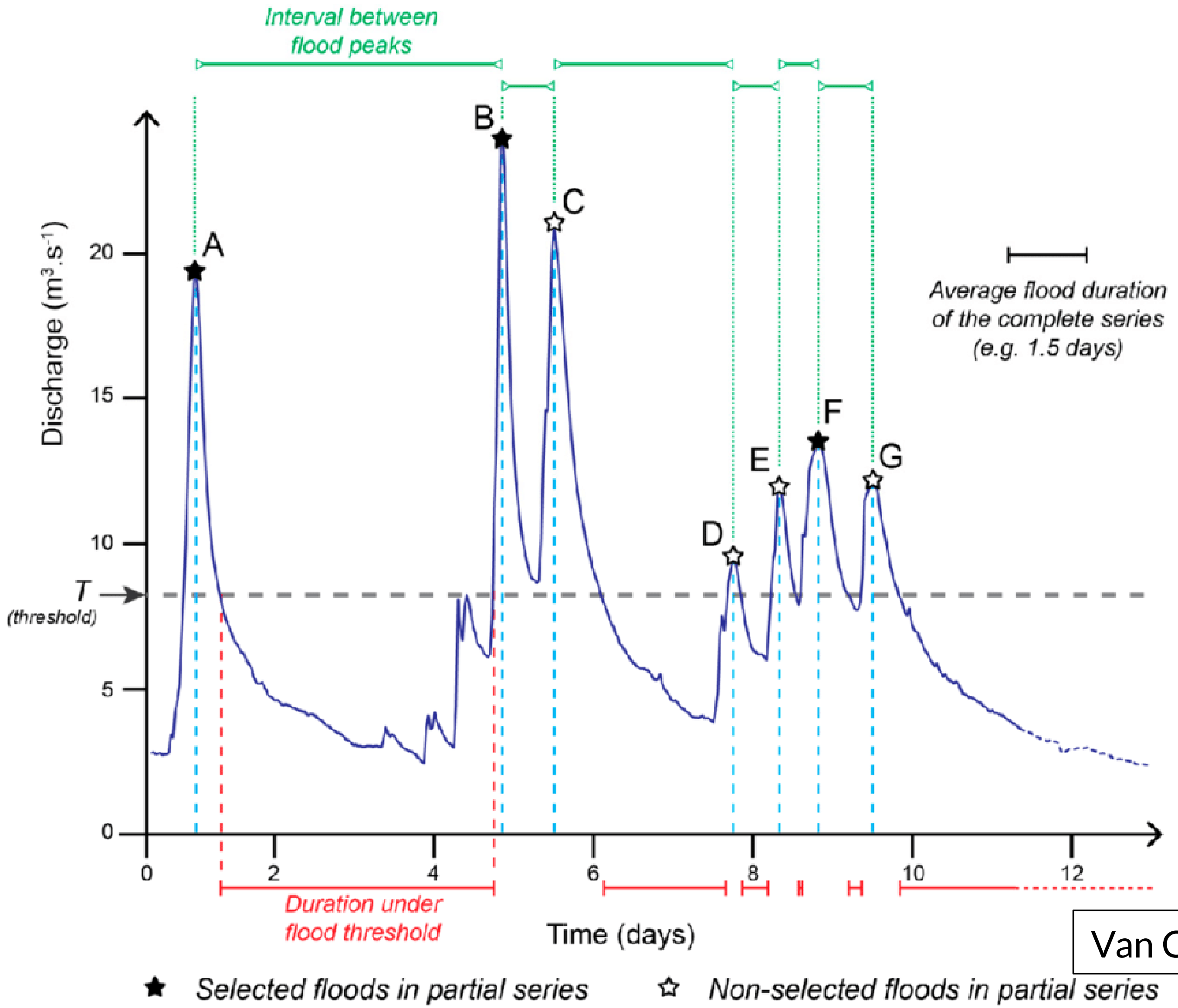
Independence

- Floods you keep should be independent of others
- Enough time passes between floods
- Hydrograph sufficiently recedes between floods

- Violation: we think large floods happen more often than they actually do

US Army Corps of Engineers®

Independence

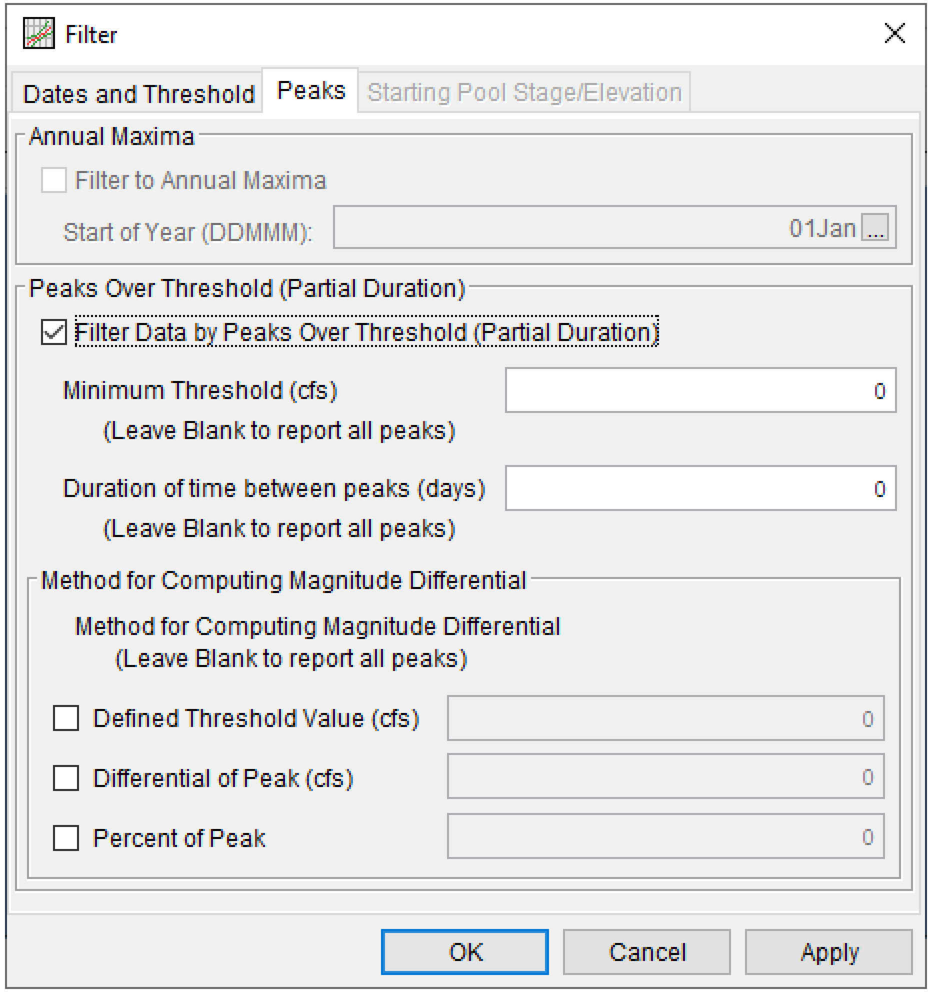


Van Campenhout, et al. (2020)

US Army Corps of Engineers®

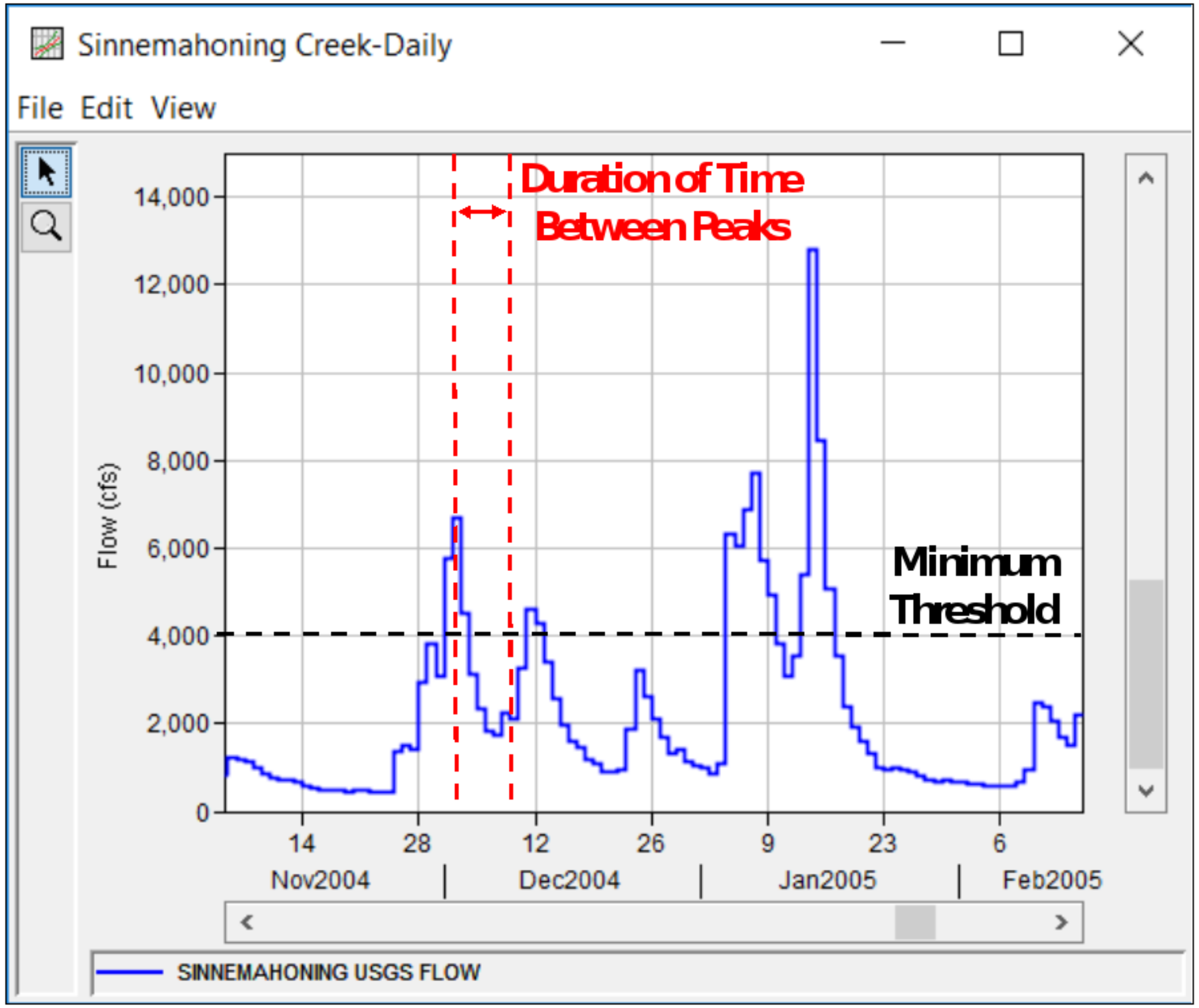
Filtering Peaks in HEC-SSP

- Threshold
- Time between peaks
- Magnitude difference
  - Recede below a threshold
  - Recede by x cfs
  - Recede by y% of peak



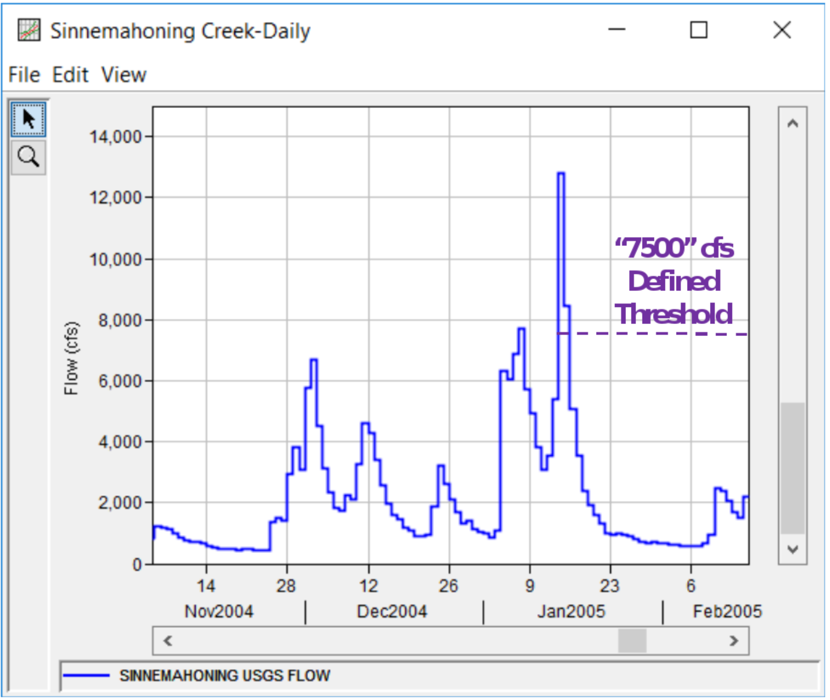
US Army Corps of Engineers®

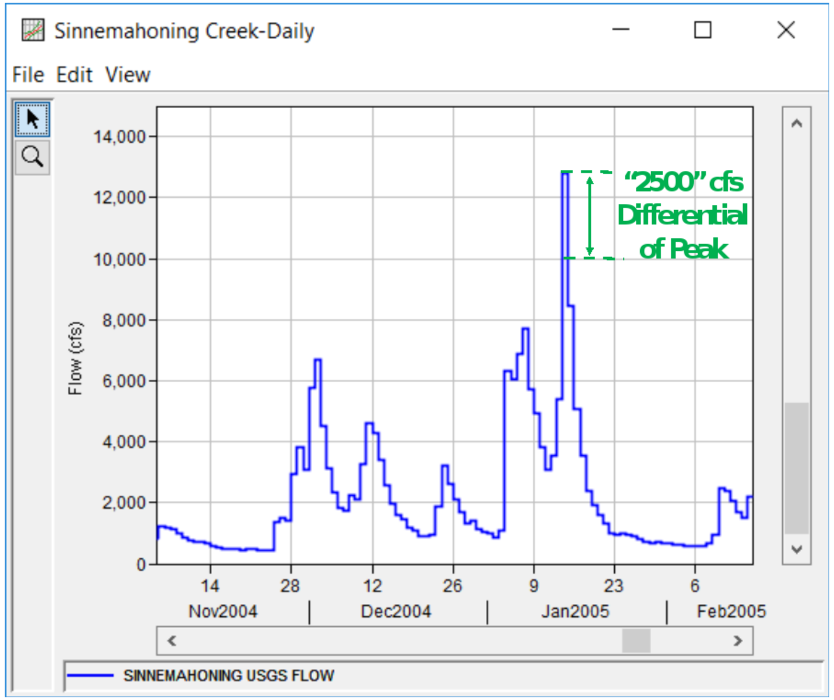
Filtering Peaks

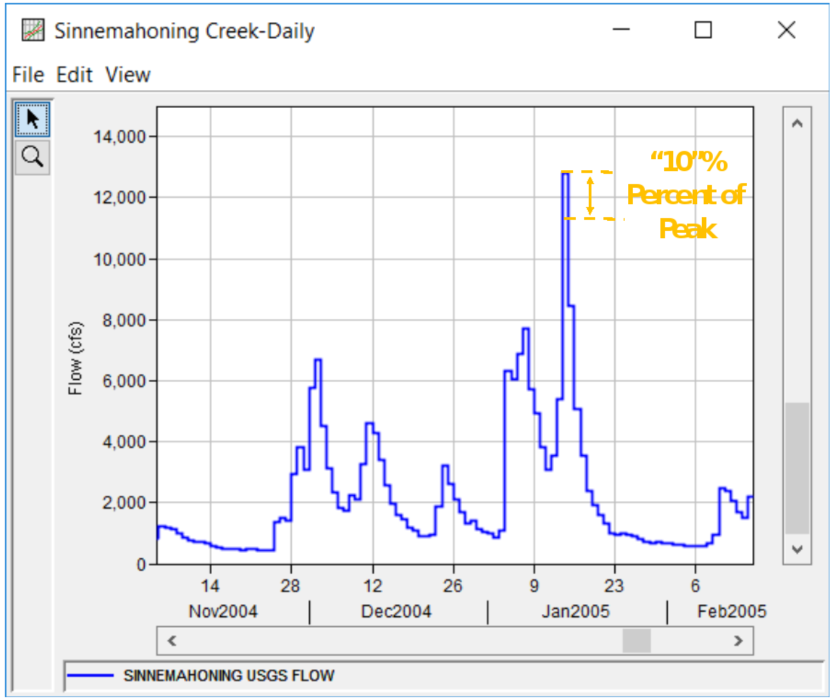


US Army Corps of Engineers®

Filtering Peaks



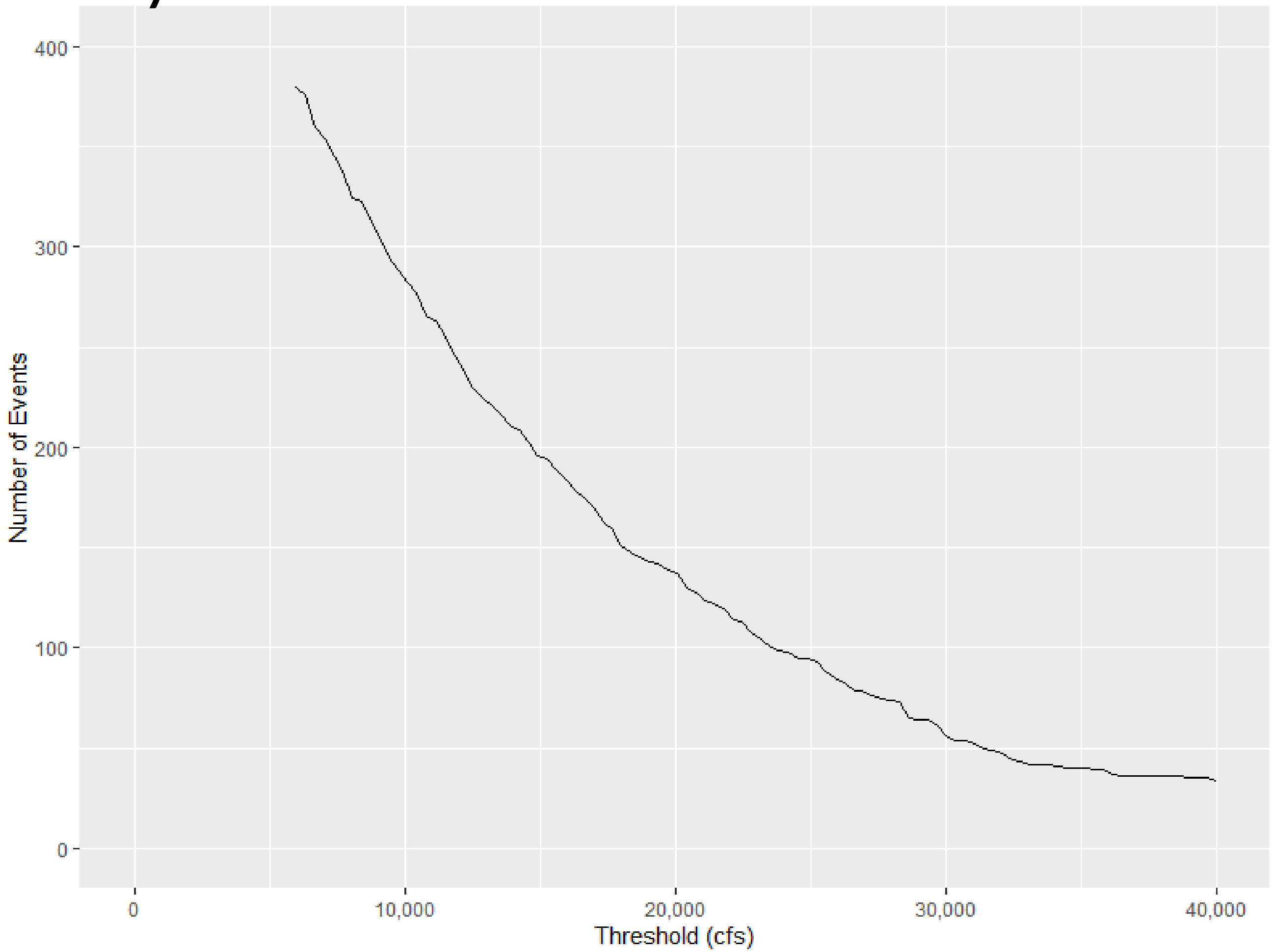




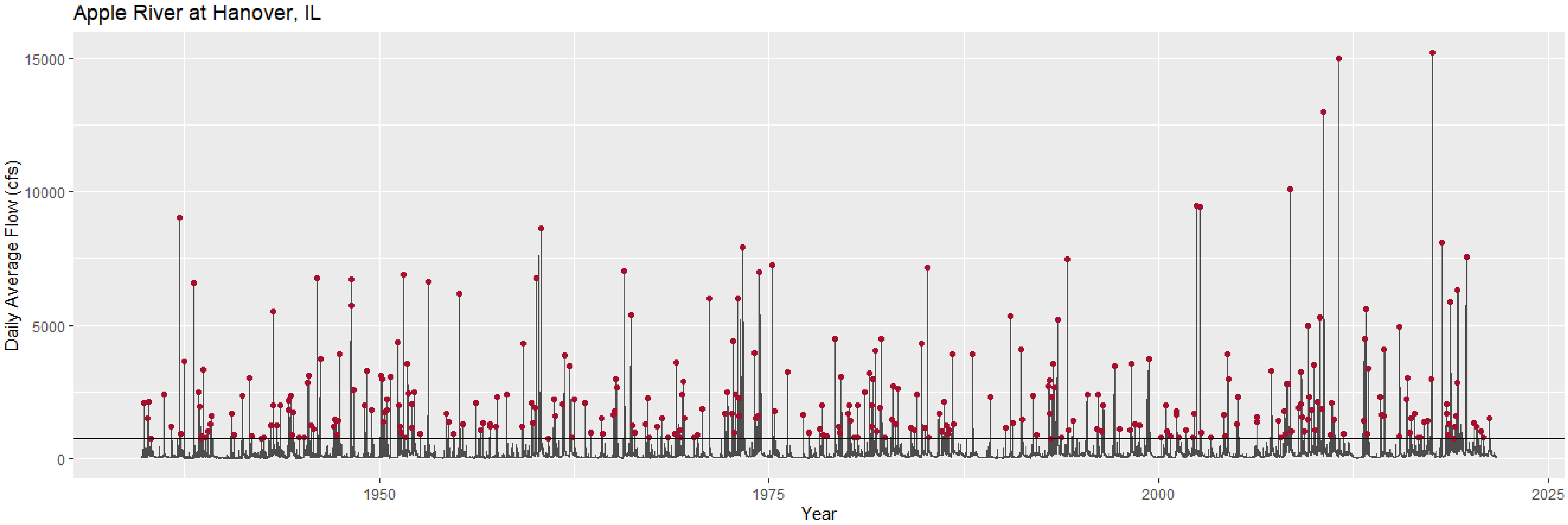
US Army Corps of Engineers®

Choosing a Threshold

- Sufficiently high threshold necessary
  - Helps with assumptions
  - Makes our model work
- Too high a threshold results in small sample sizes

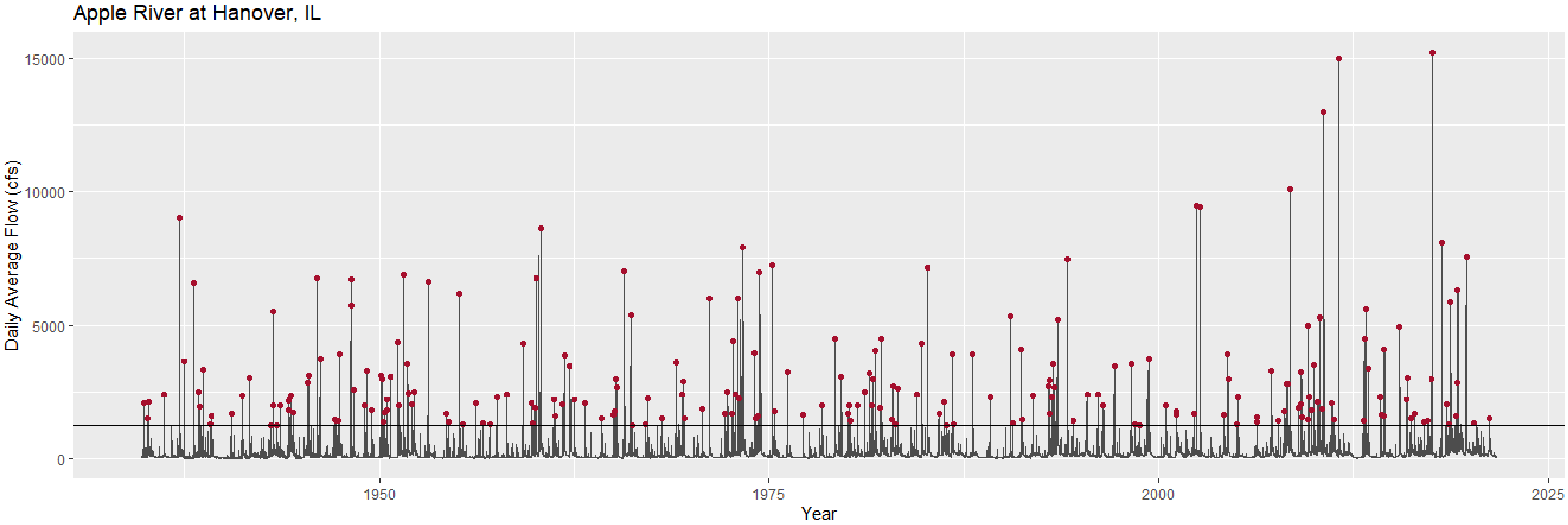


US Army Corps of Engineers®



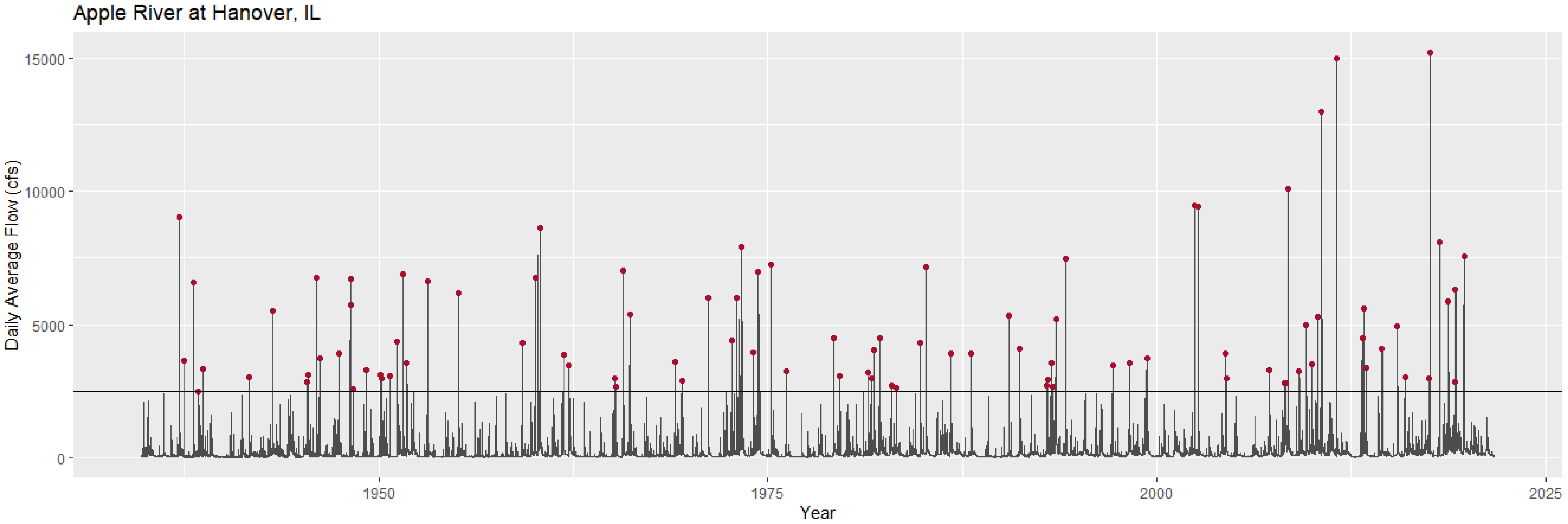
US Army Corps of Engineers®

750 cfs
11 days minimum separation
250 cfs minimum recession
318 events



US Army Corps of Engineers®

1,250 cfs
11 days minimum separation
250 cfs minimum recession
220 events



US Army Corps of Engineers®

2,500 cfs
11 days minimum separation
250 cfs minimum recession
93 events

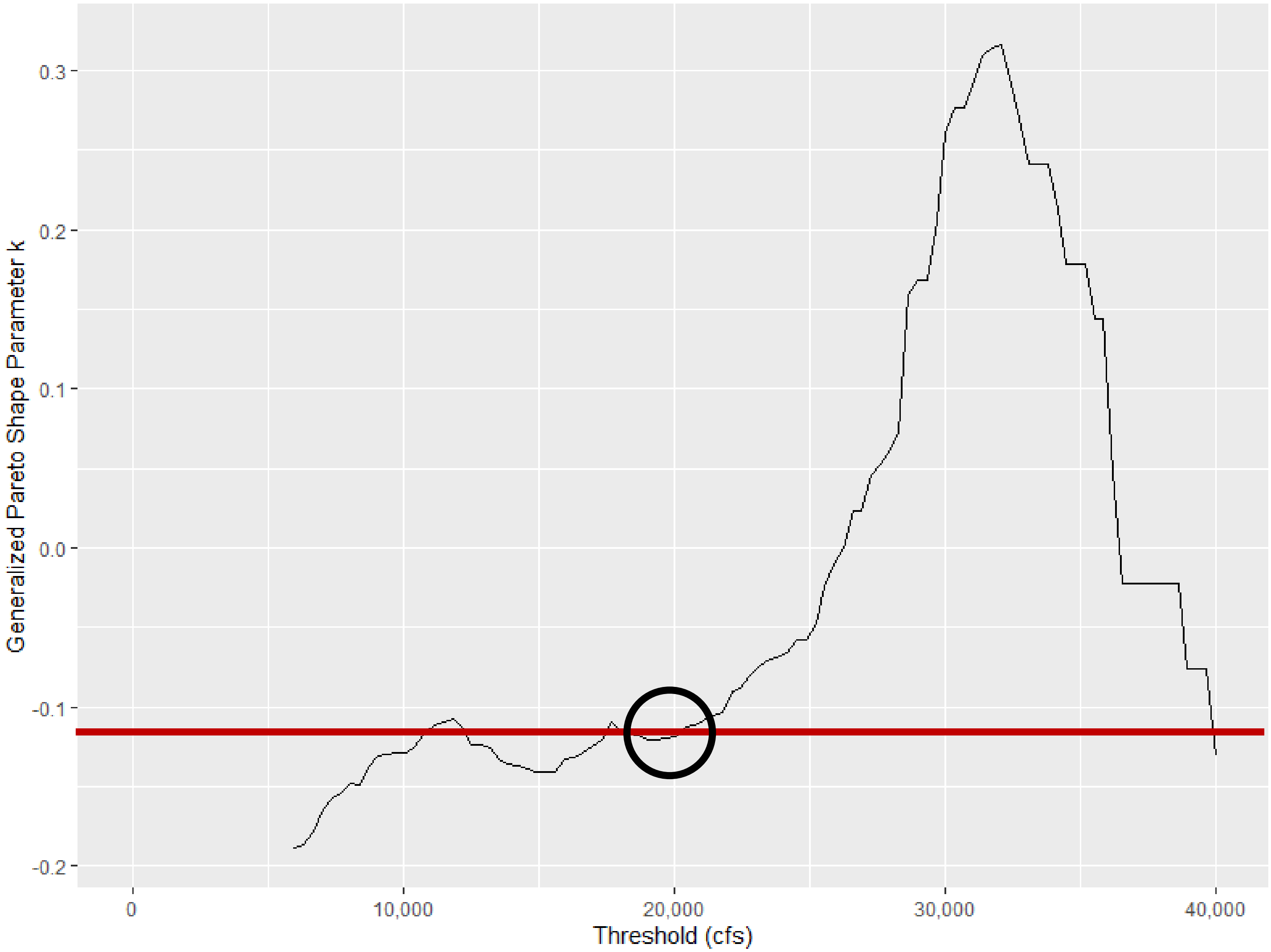
Choosing a Threshold: Ideas

- Minimum annual maximum value
- Bankfull discharge
- Mean rate = 1 event / year (# events = # years)
- More "advanced" diagnostics
  - Pareto shape plot
  - Mean excess plot

US Army Corps of Engineers®

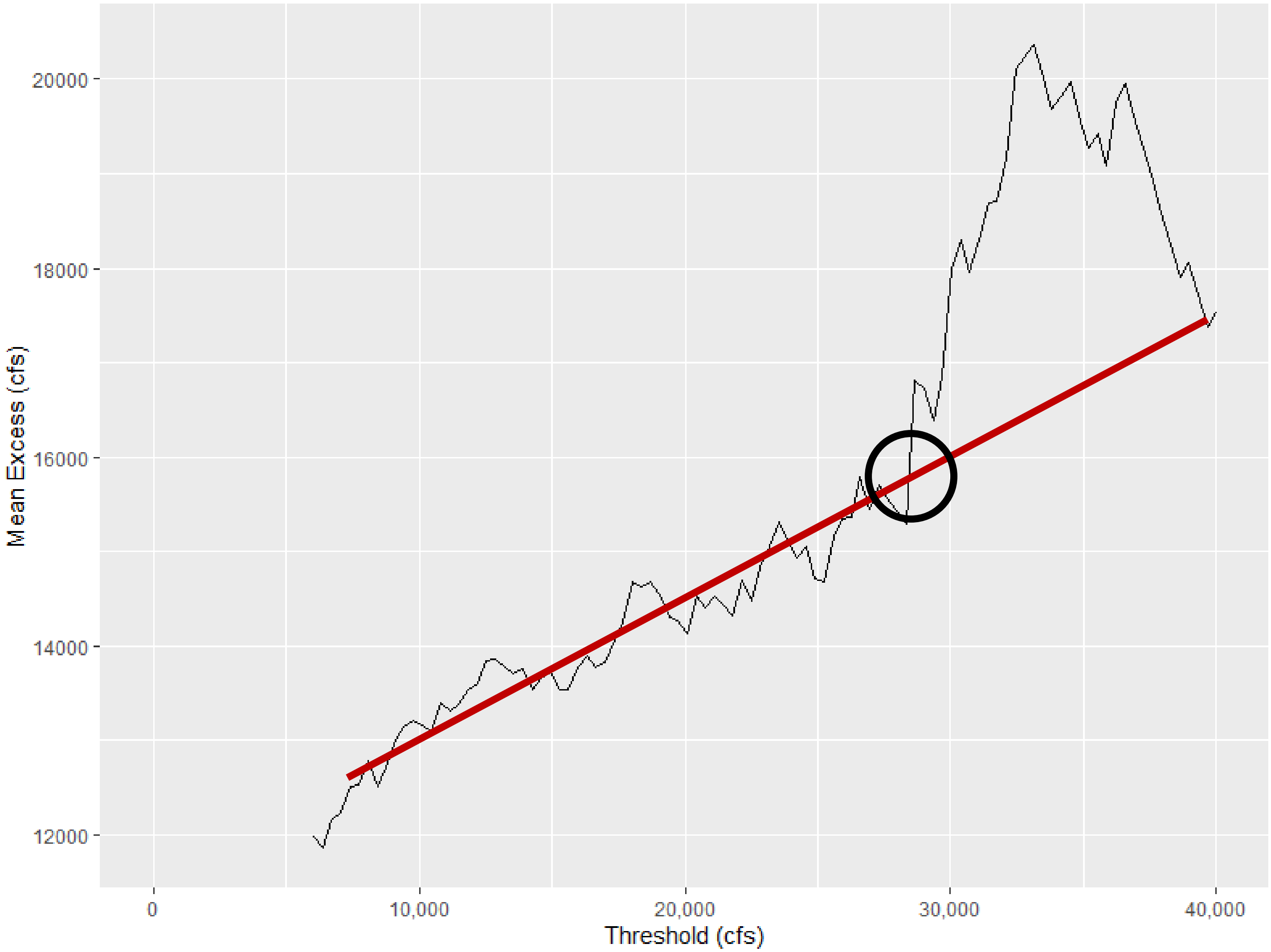
Other Diagnostics: Pareto Shape Plot

- Compute PDS for varying thresholds
- Fit generalized Pareto distribution to each
- Plot threshold vs. Pareto shape parameter
- Check for flat spot or change in shape



Other Diagnostics: Mean Excess Plot

- Compute PDS for varying thresholds
- Compute mean amount peaks exceed threshold
- Plot threshold vs. mean
- Check for linearity or change in shape



US Army Corps of Engineers®

Result

- Collection of IID flood events
- Variable number of events per year
- All events greater than a specified threshold
- ***Partial duration series (PDS)***

US Army Corps
of Engineers®

Models for Partial Duration Series

Flood Frequency Guidance: Review

Annual maximum series (AMS) +

Log-Pearson type III distribution (LP3) +

Expected moments algorithm (EMA) =

Bulletin 17C

US Army Corps
of Engineers®

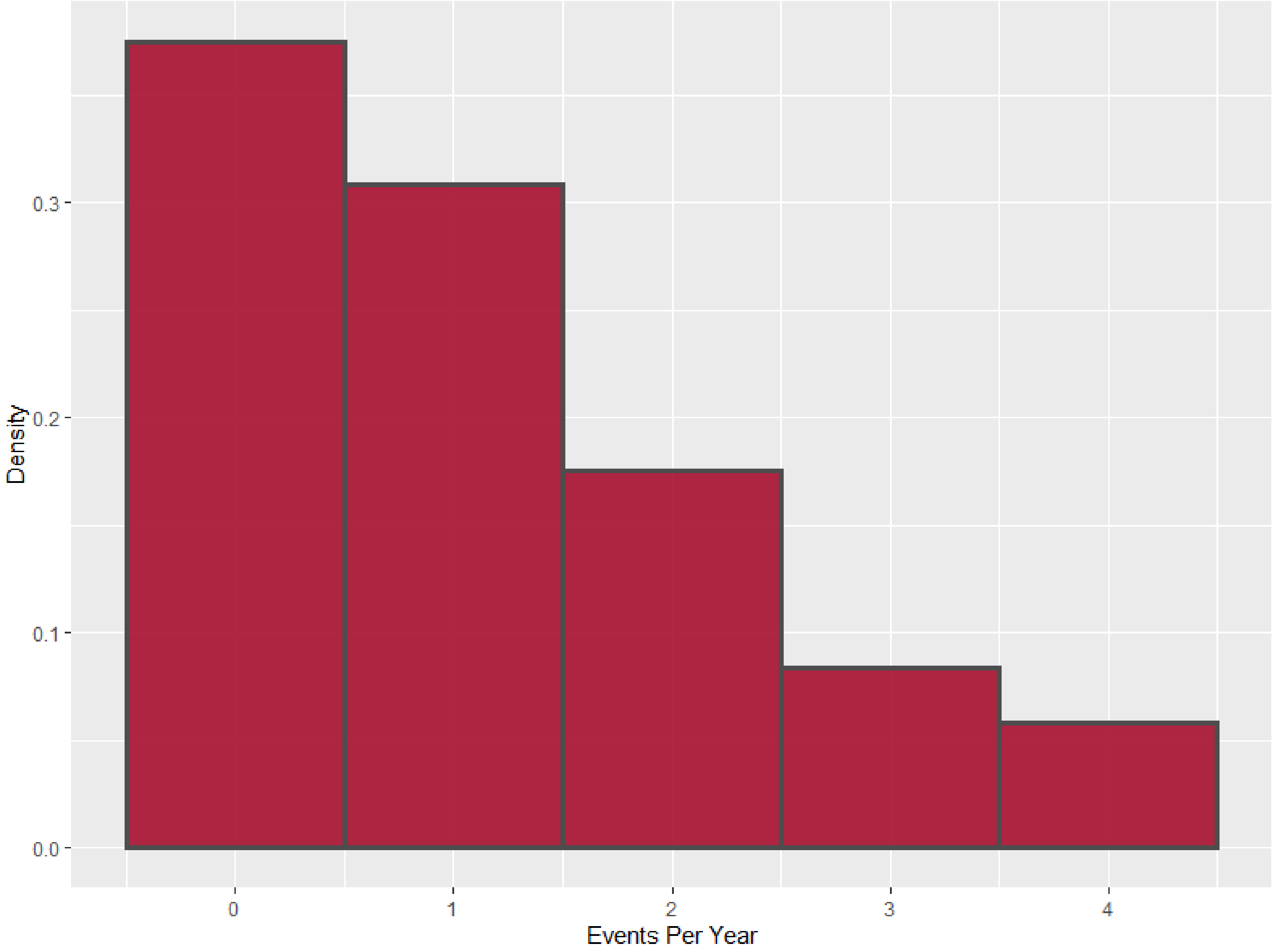
Partial Duration Series

- Distribution of independent values exceeding a threshold
- *Probability of exceeding a value, given it's already exceeded threshold*
  - $P(X > x | X > x_0)$
- Need two things:
  - Distribution of those values
  - Distribution of counts per year

US Army Corps
of Engineers®

Counts

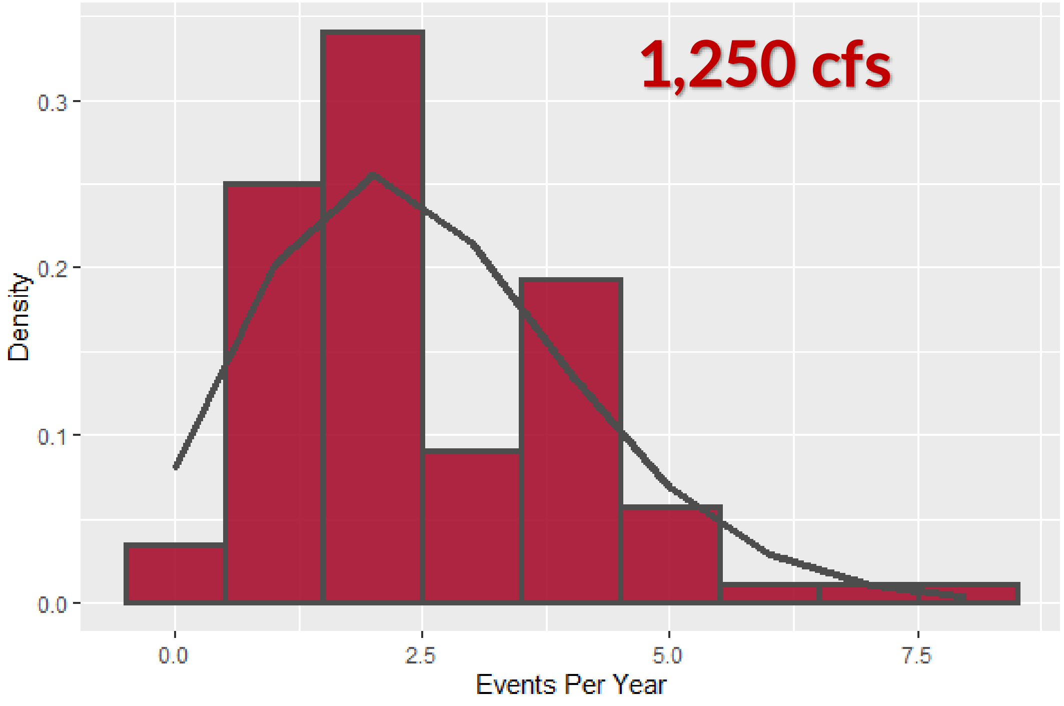
- 0 or more events per year in PDS
- Two key questions:
  - Probability of at least one event in a year?
  - Typical number of events in a year?

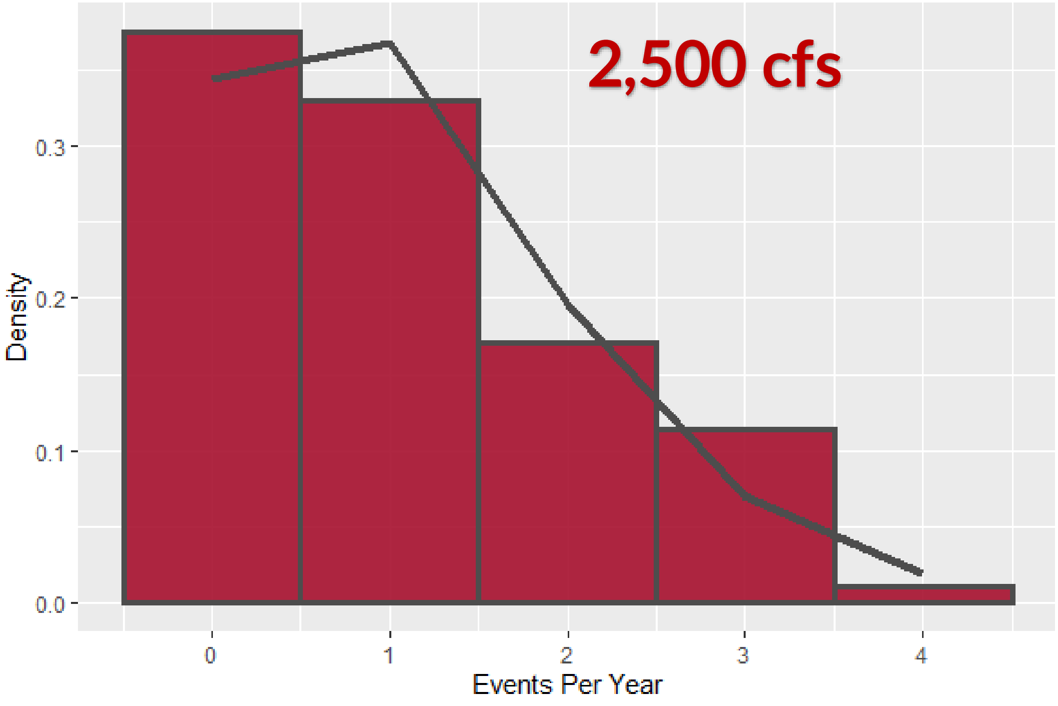


US Army Corps of Engineers®

Counts

- Threshold choice affects count distribution
- Usually assumed to have a *Poisson Distribution*

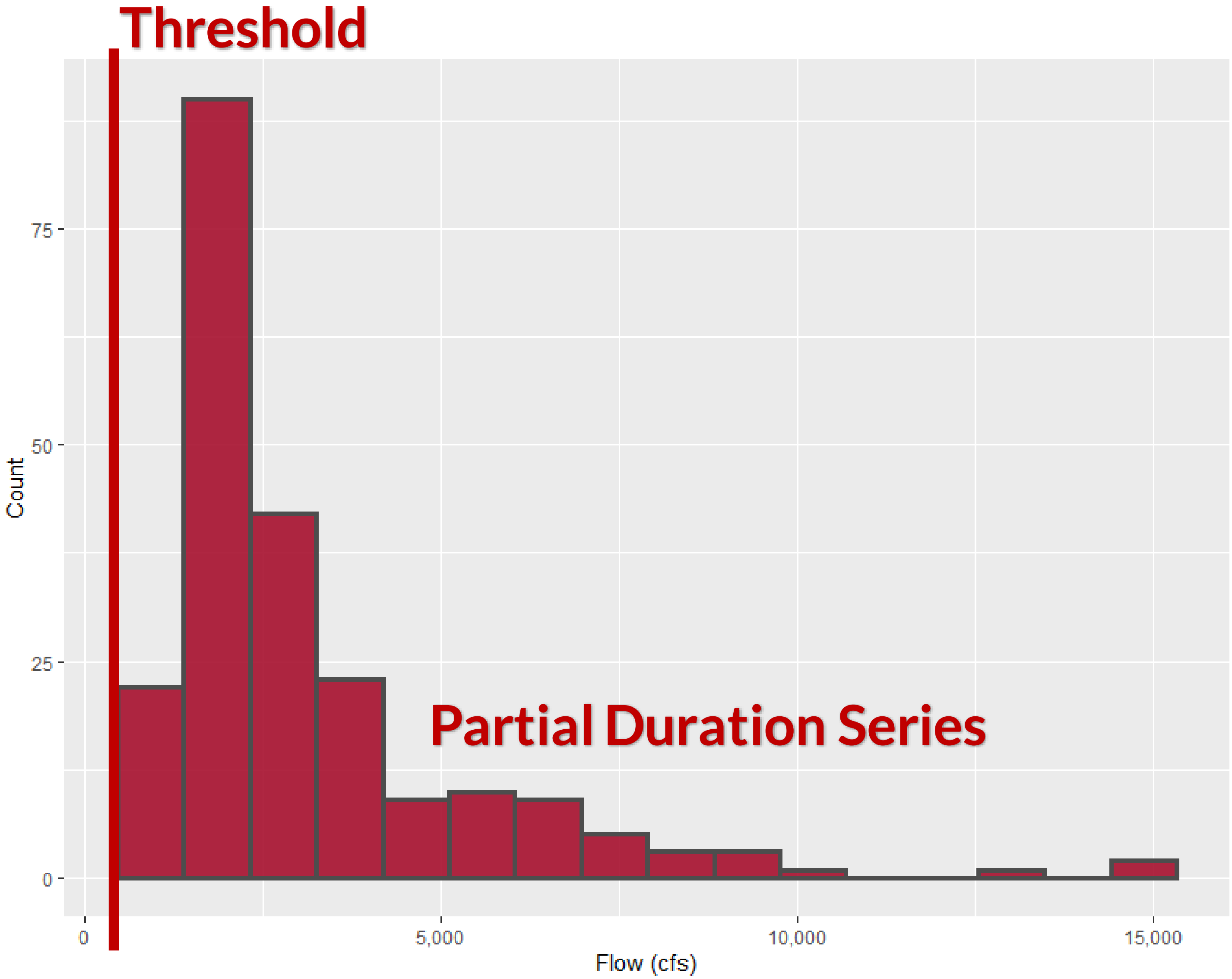




US Army Corps of Engineers®

Distribution of Values

- Model magnitude of values greater than threshold
- Tend to cluster close to threshold

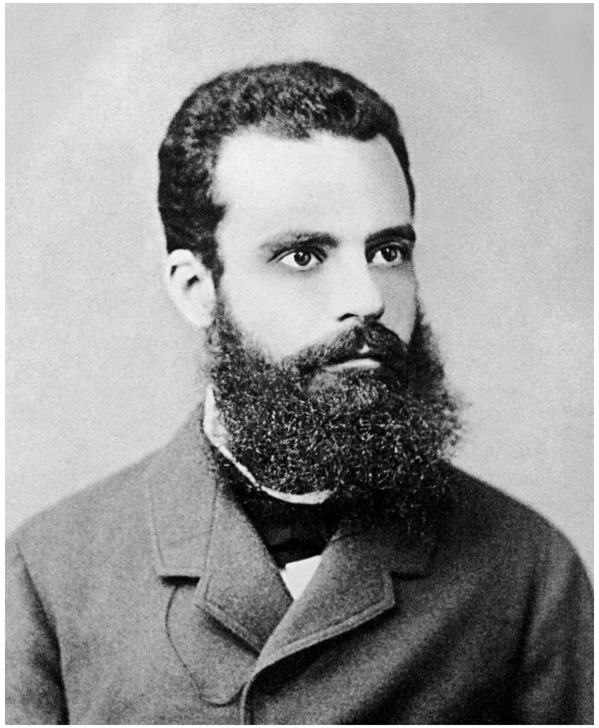


US Army Corps of Engineers®

Distribution of Values

- Lean on extreme value theory
  - We have a model for this situation!
- Generalized Pareto Distribution (GPD)
  - 3 parameters: ξ (location) α (scale) κ (shape)

For more information, come to the Statistical Methods class!

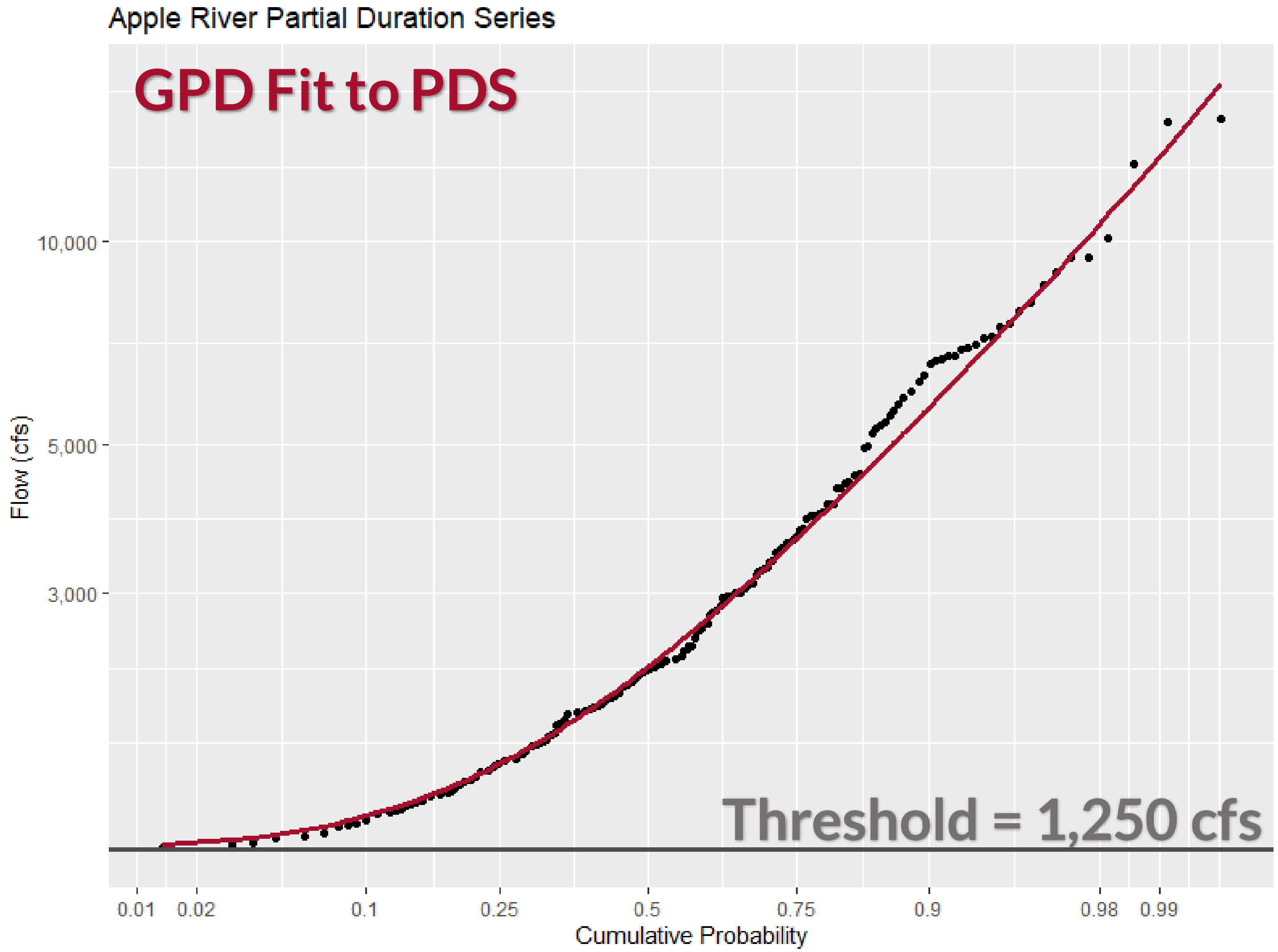


Vilfredo Pareto

US Army Corps of Engineers®

Distribution of Values

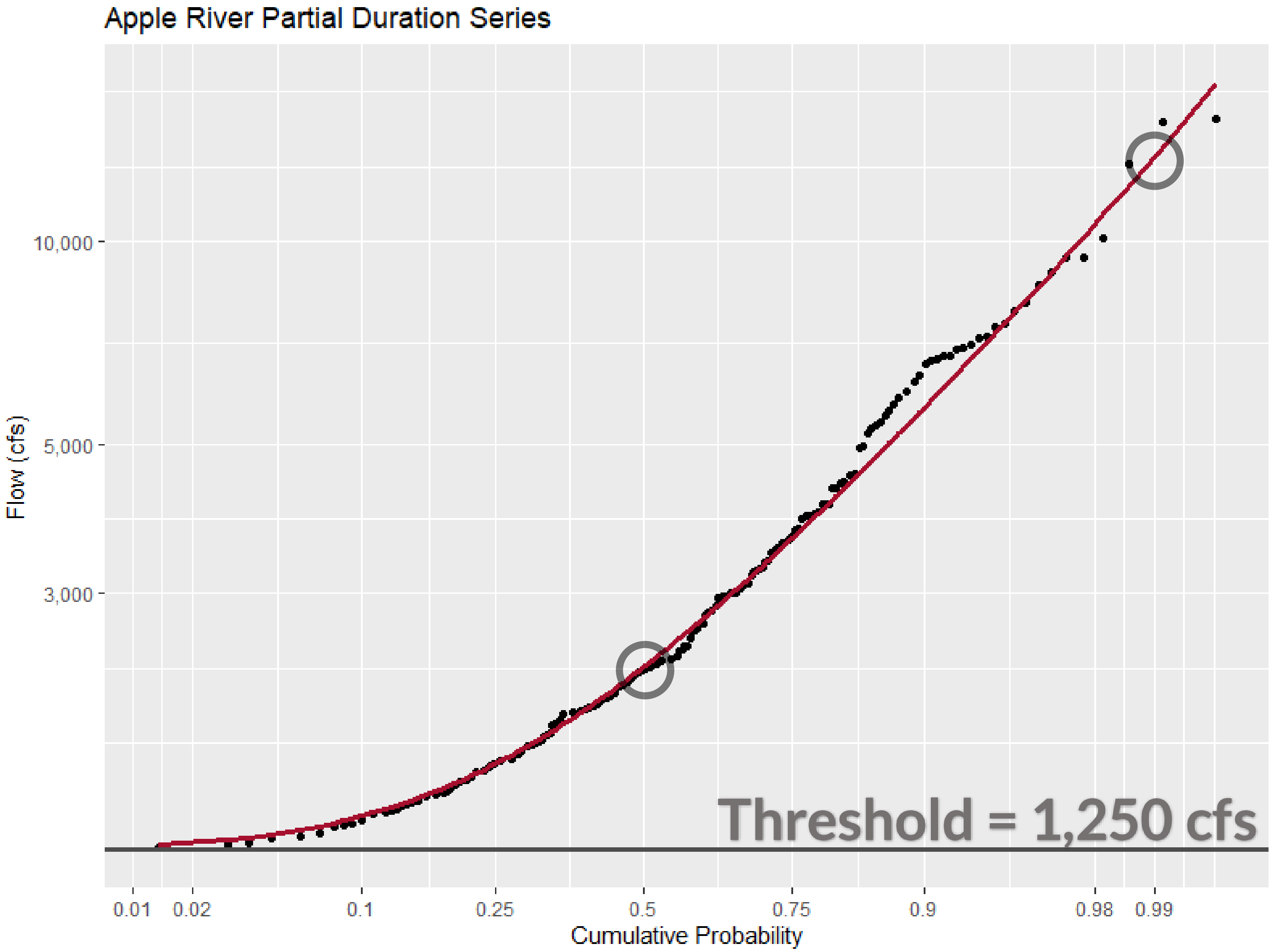
- Cumulative probability, given the values exceed the threshold
- ***NOT AN AEP***



US Army Corps of Engineers®

Interpreting the Distribution

- 50% of floods that exceed 1,250 cfs are greater than 2,330 cfs
- 1% of floods that exceed 1,250 cfs are greater than 13,400 cfs



US Army Corps of Engineers®

Computing AEP from PDS

US Army Corps
of Engineers®

Annual Exceedance Probability

- AMS analysis provides AEP events
- PDS analysis provides conditional exceedance events

- We need AEP for things like annualized damages!

US Army Corps
of Engineers®

Computing AEP from PDS

Two general approaches:

- Empirical
- Analytical

- “Annualization”

US Army Corps
of Engineers®

Empirical Annualization

- Langbein (1949) and Takeuchi (1984)
- Assume storm per year counts have a Poisson distribution
  - One parameter λ
  - PDS CDF $G(x)$, AMS CDF $F(x)$
- $F(x) = e^{[-\lambda\{1-G(x)\}]}$
- $G(x) = \frac{ln[F(x)]+\lambda}{\lambda}$

US Army Corps of Engineers®

Applying Empirical Annualization

- Get AEPs you want to compute
- Convert them to PDS CDF values
- Plug them into your GPD distribution quantile function

- Can also be used to plot observed peaks
  - Convert the plotting positions

US Army Corps of Engineers®

Analytical Annualization

- Stedinger et al. (1993), Madsen et al. (1997)
  - Generalized Pareto + Poisson → Generalized Extreme Value
- Have:
  - Poisson parameter λ
  - Generalized Pareto parameters ξ, α, κ
- Get:
  - GEV parameters ξ^*, α^*, κ^*
- $\xi^* = \xi + \frac{\alpha}{\kappa}(1 - \lambda^{-\kappa})$ $\quad \alpha^* = \alpha\lambda^{-\kappa}$ $\quad \kappa^* = \kappa$

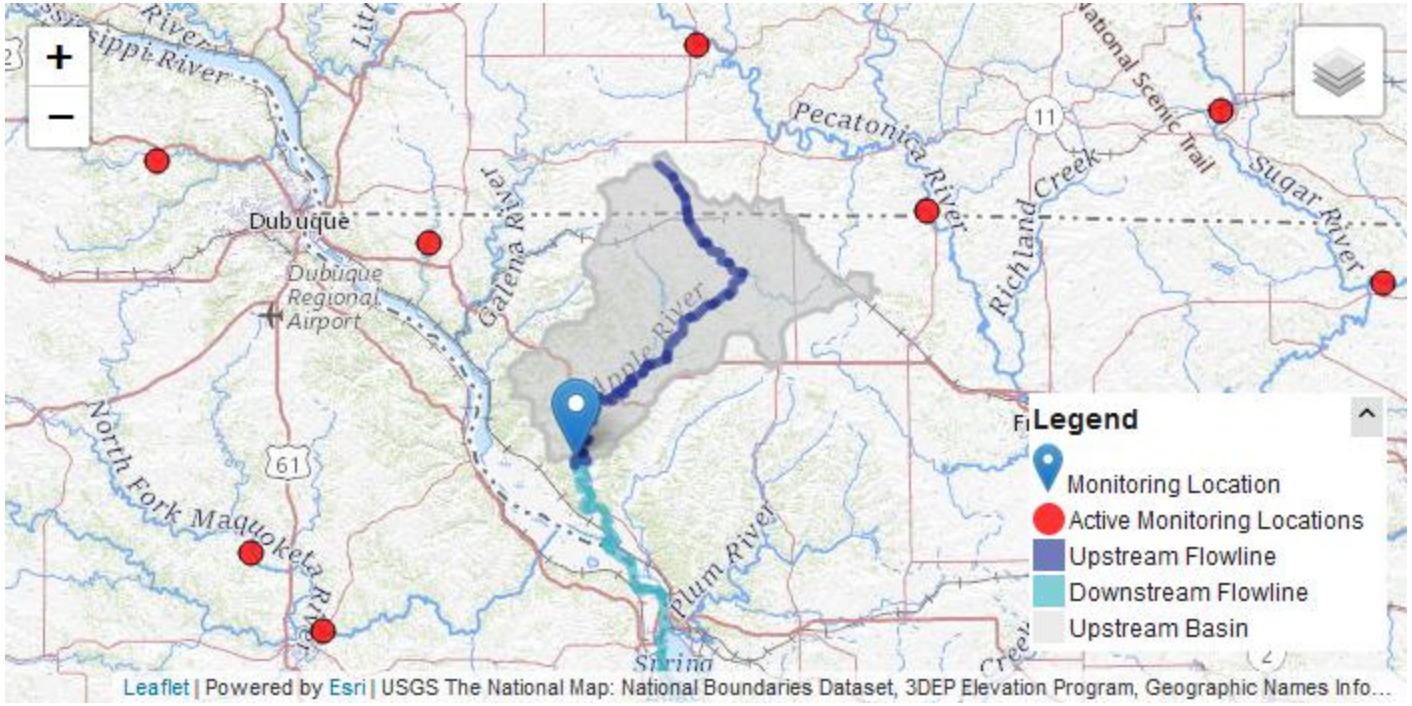
US Army Corps of Engineers®

Applying Analytical Annualization

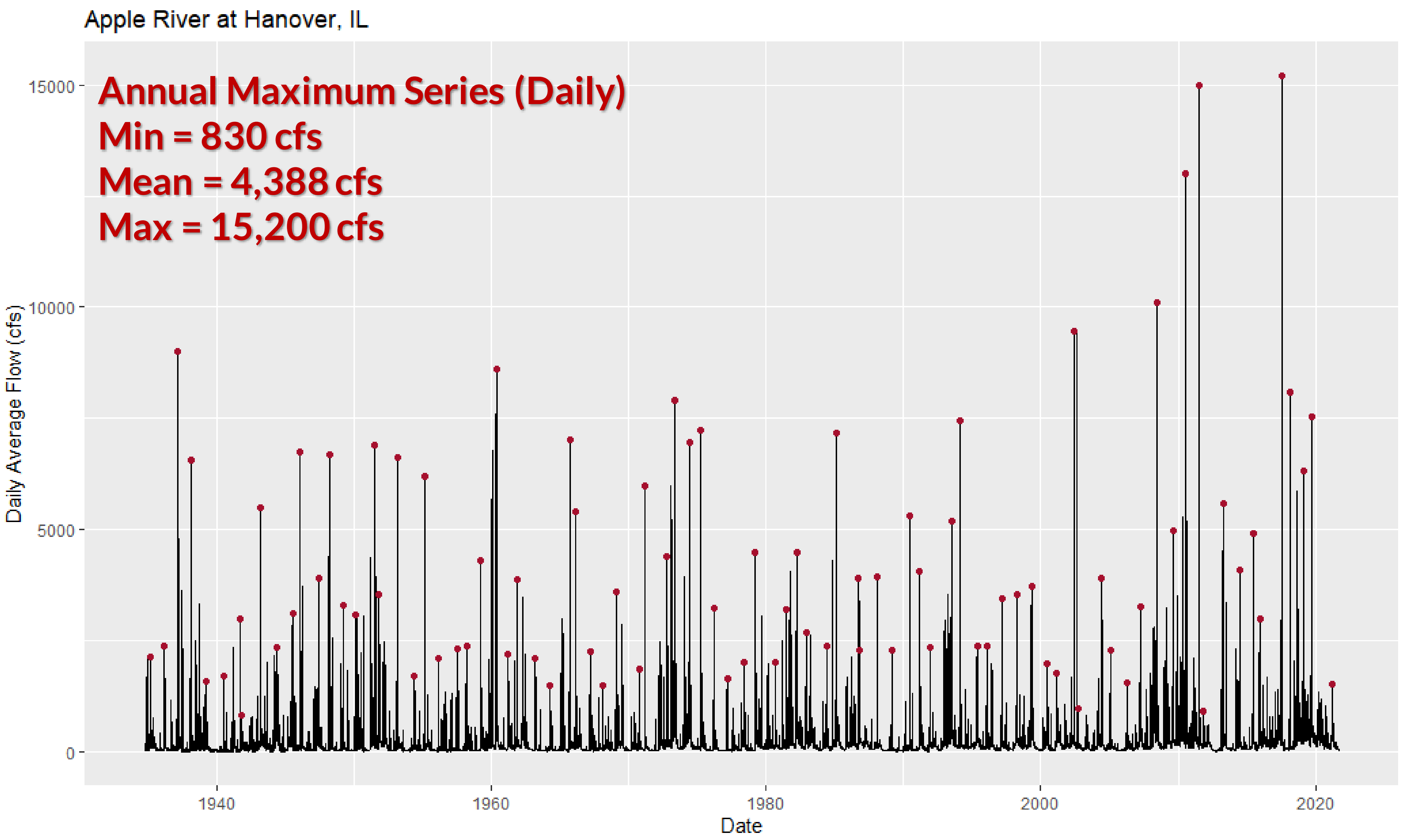
- Compute equivalent GEV parameters from GPD/Poisson
- Plug AEP values into GEV quantile function

Fit and Annualization Example

- Apple River at Hanover, IL
  - 246 mi^2
  - WY1935-2021
  - Daily average flows
  - Unregulated
  - Little urbanization

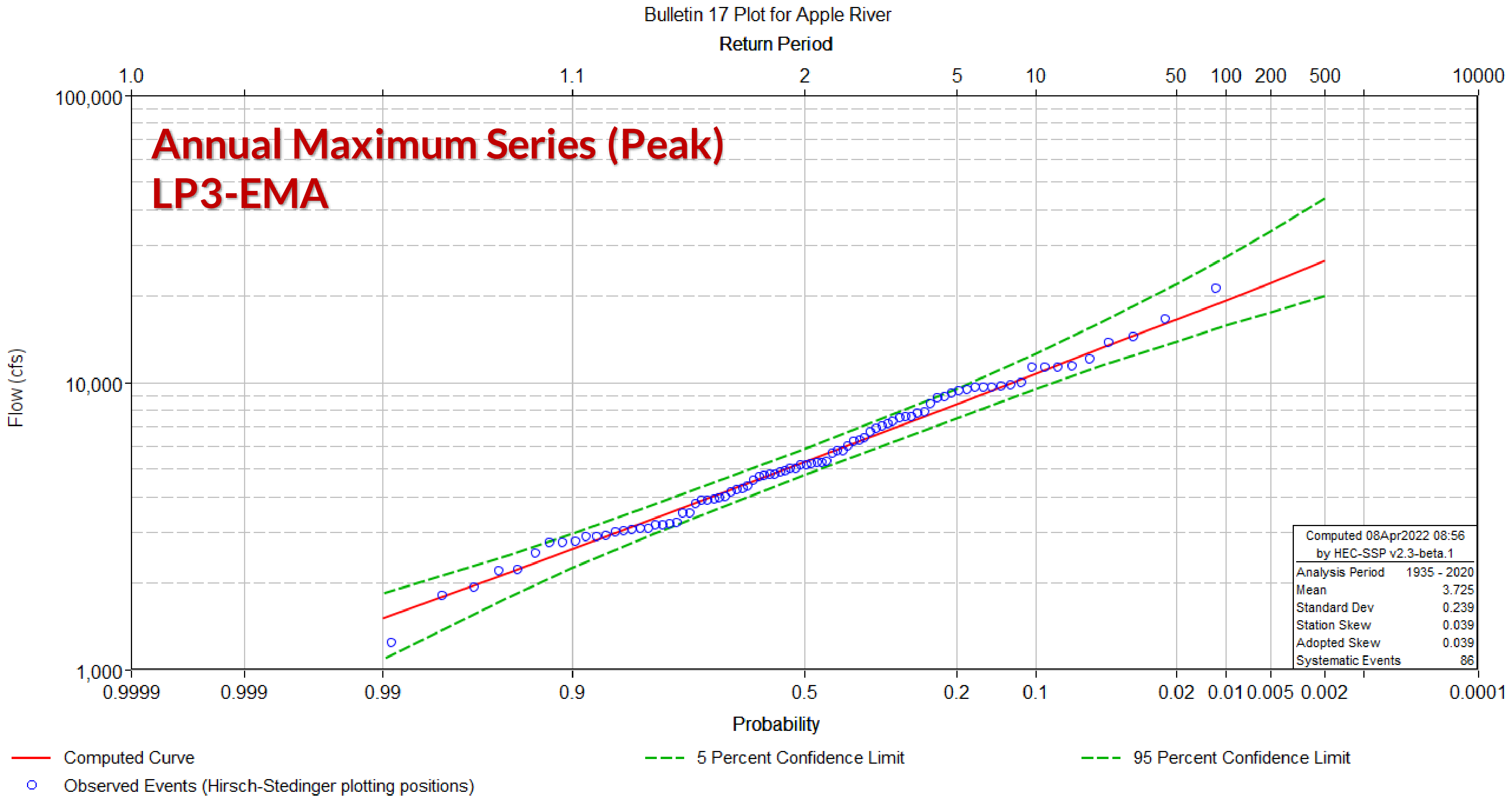


US Army Corps of Engineers®



US Army Corps
of Engineers®

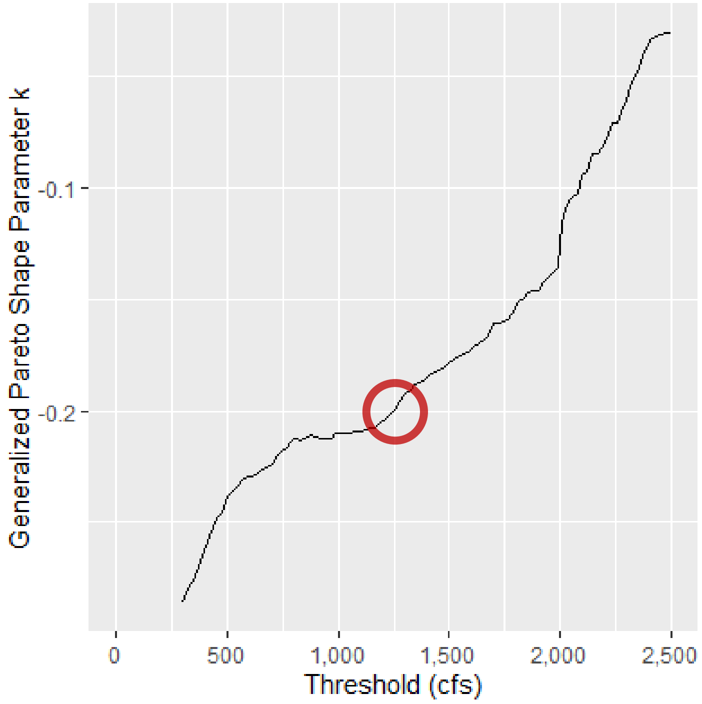
Annual Maximum Series (Peak)
LP3-EMA

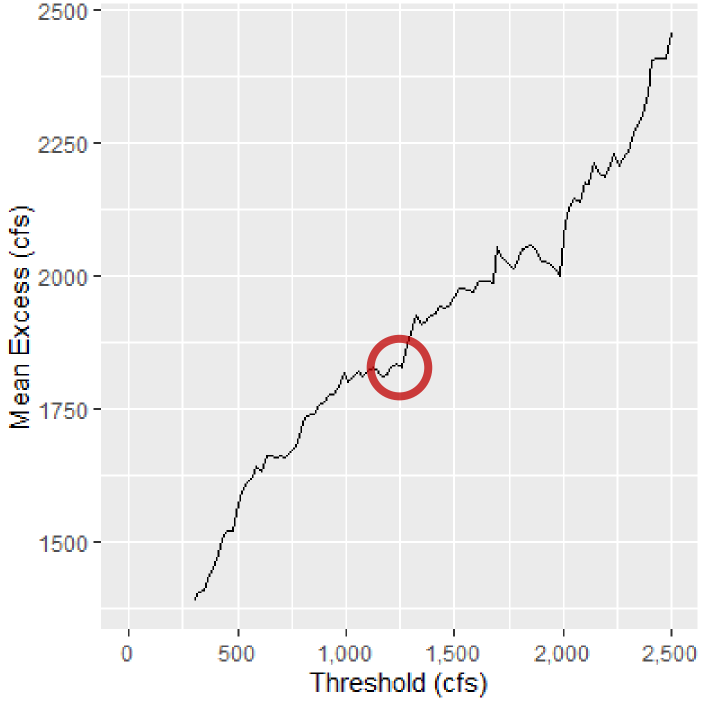


US Army Corps of Engineers®

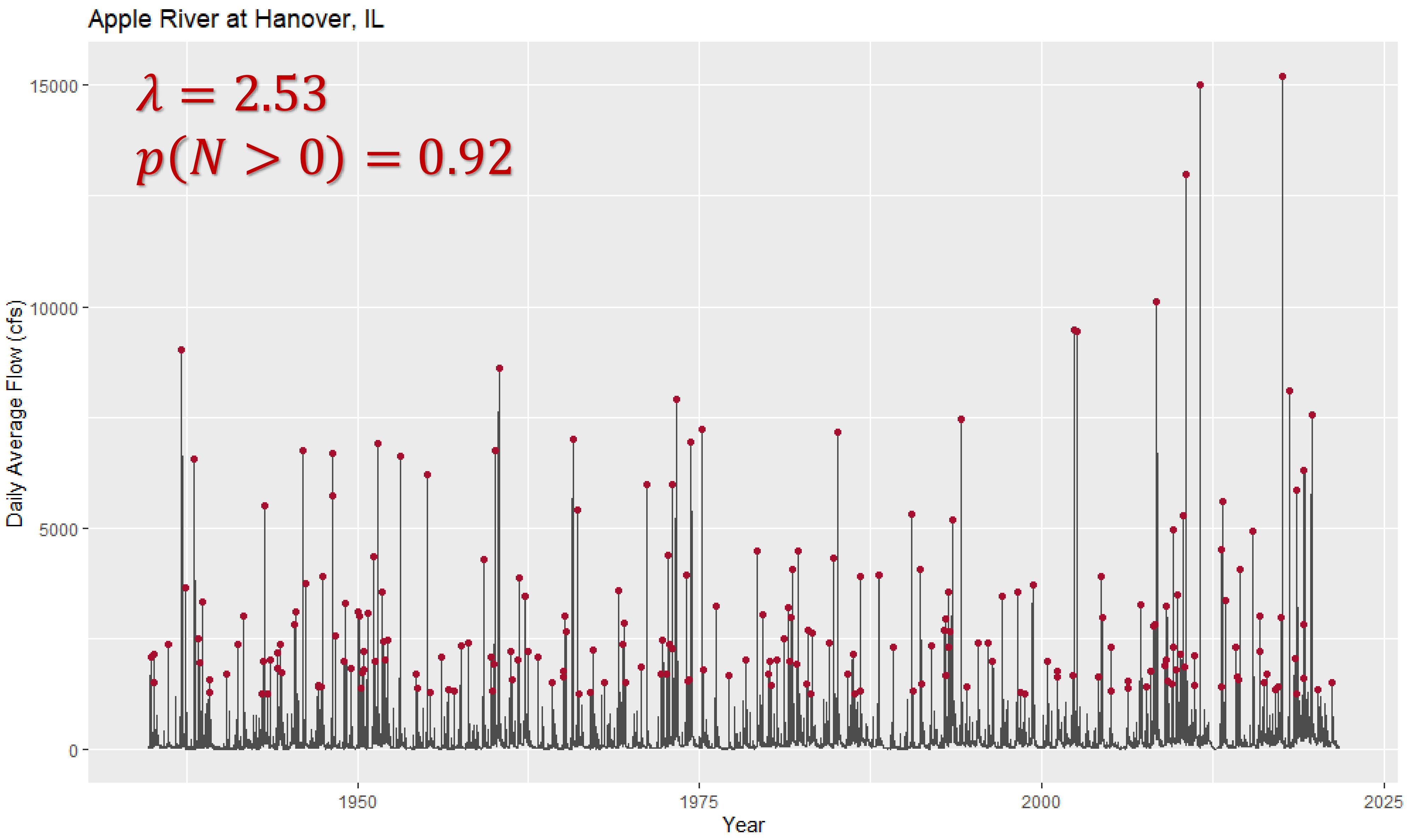
Filtering for PDS

- Threshold = 1,250 cfs
  - Based on shape plot, mean excess plot
- Time between events 11 days
  - $[5 + \ln(DA)]$ guideline
- Recession threshold 250 cfs
  - Flow must go below threshold between events

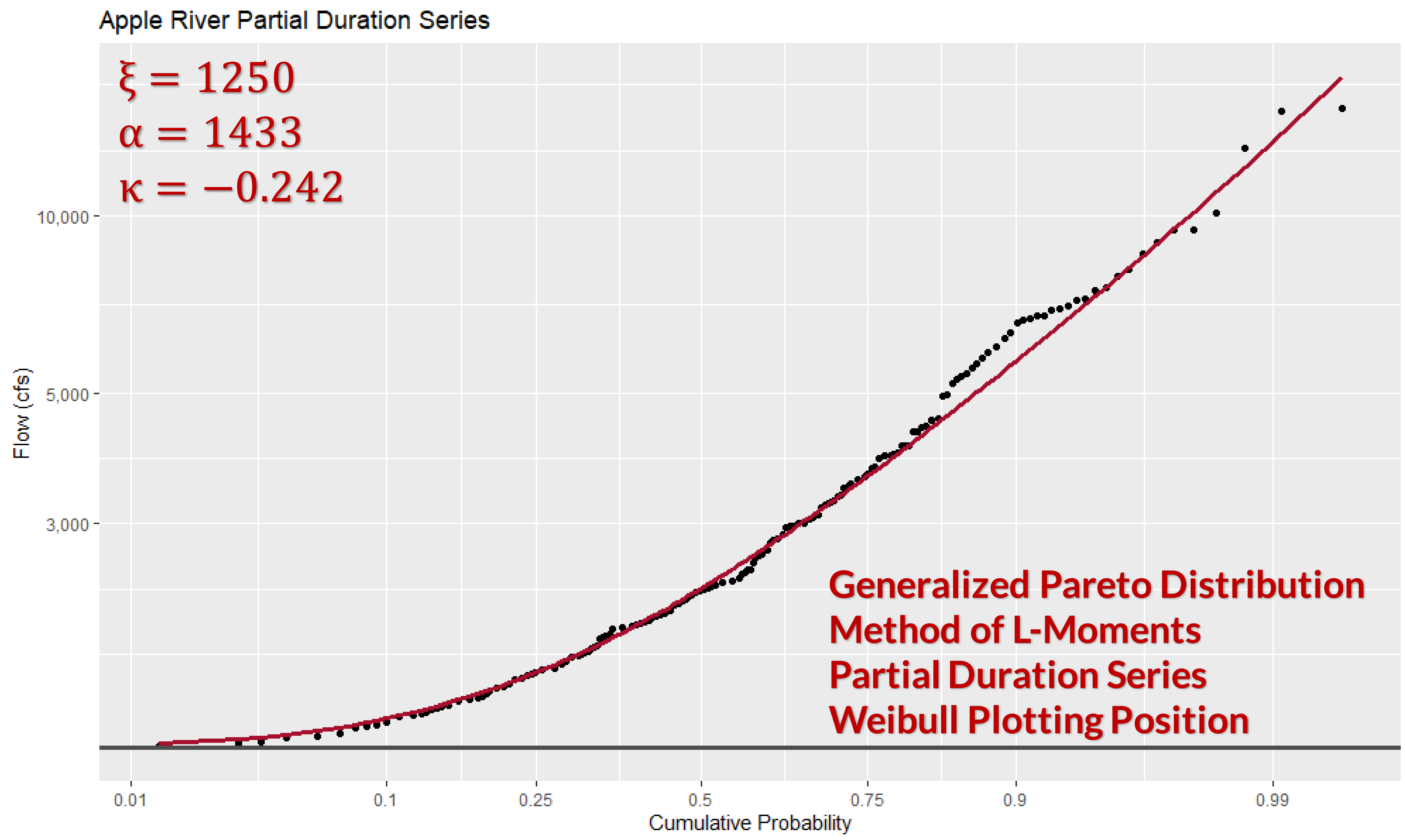




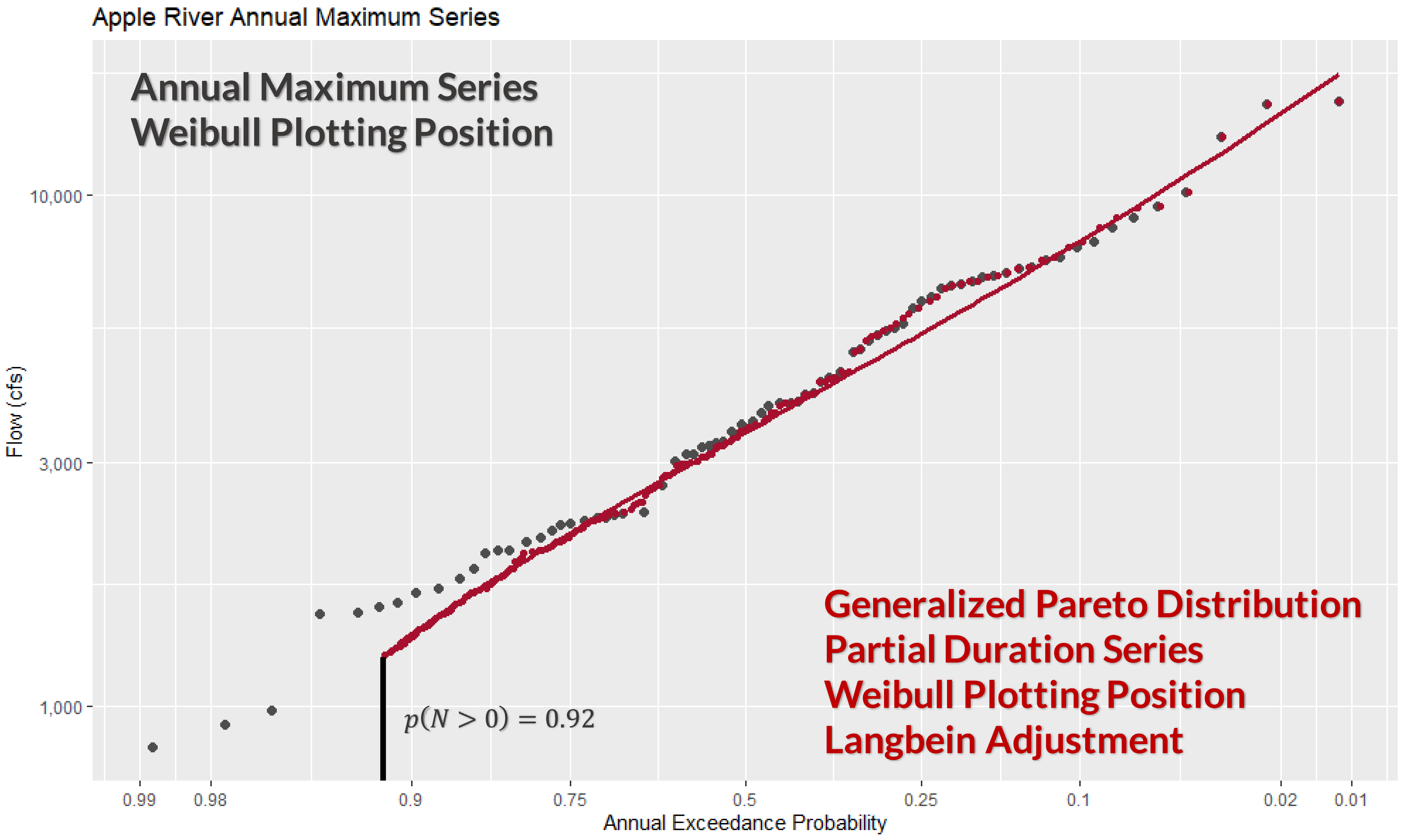
US Army Corps of Engineers®



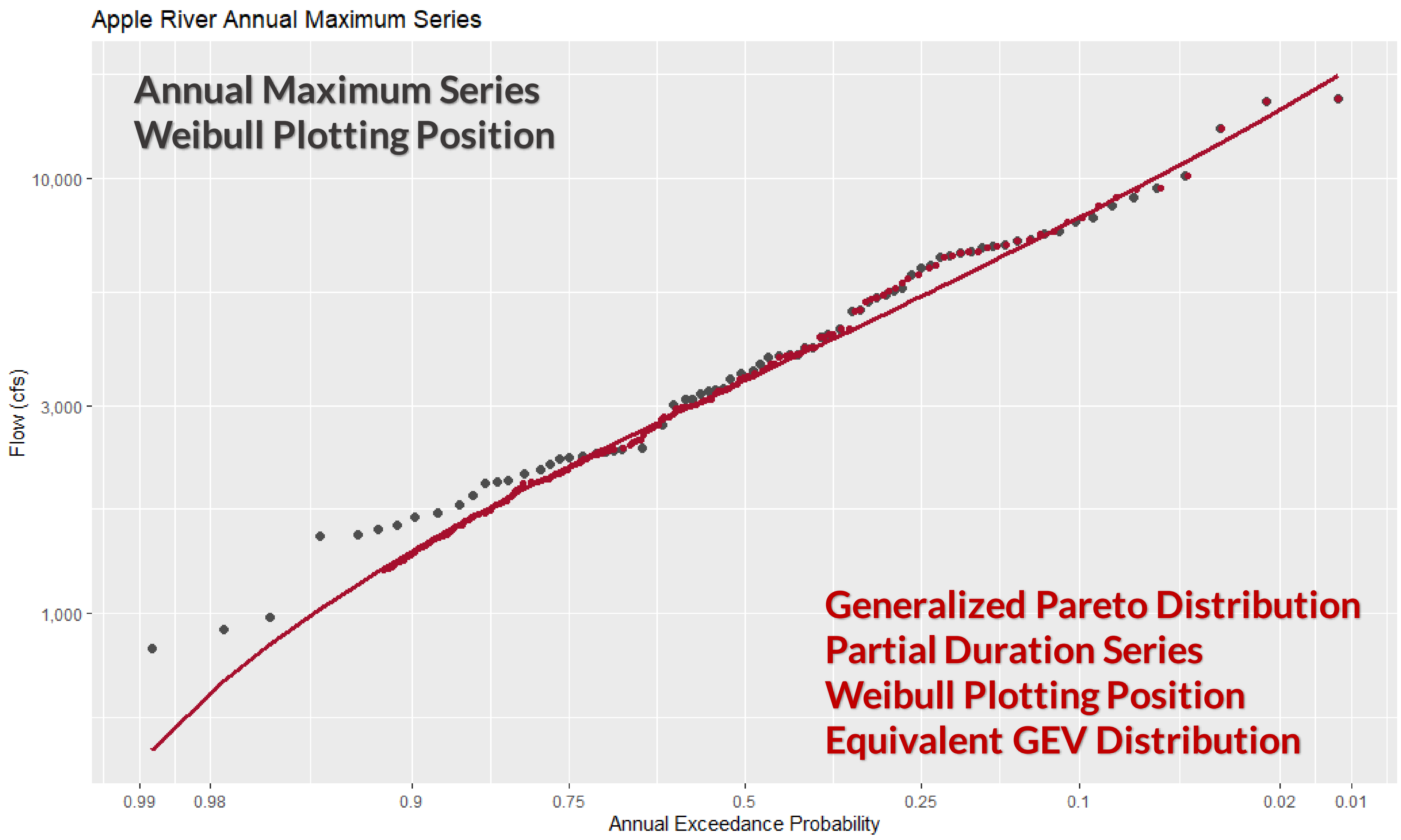
US Army Corps of Engineers®



US Army Corps of Engineers®



US Army Corps of Engineers®



US Army Corps of Engineers®

Why PDS for Streamflow?

US Army Corps of Engineers®

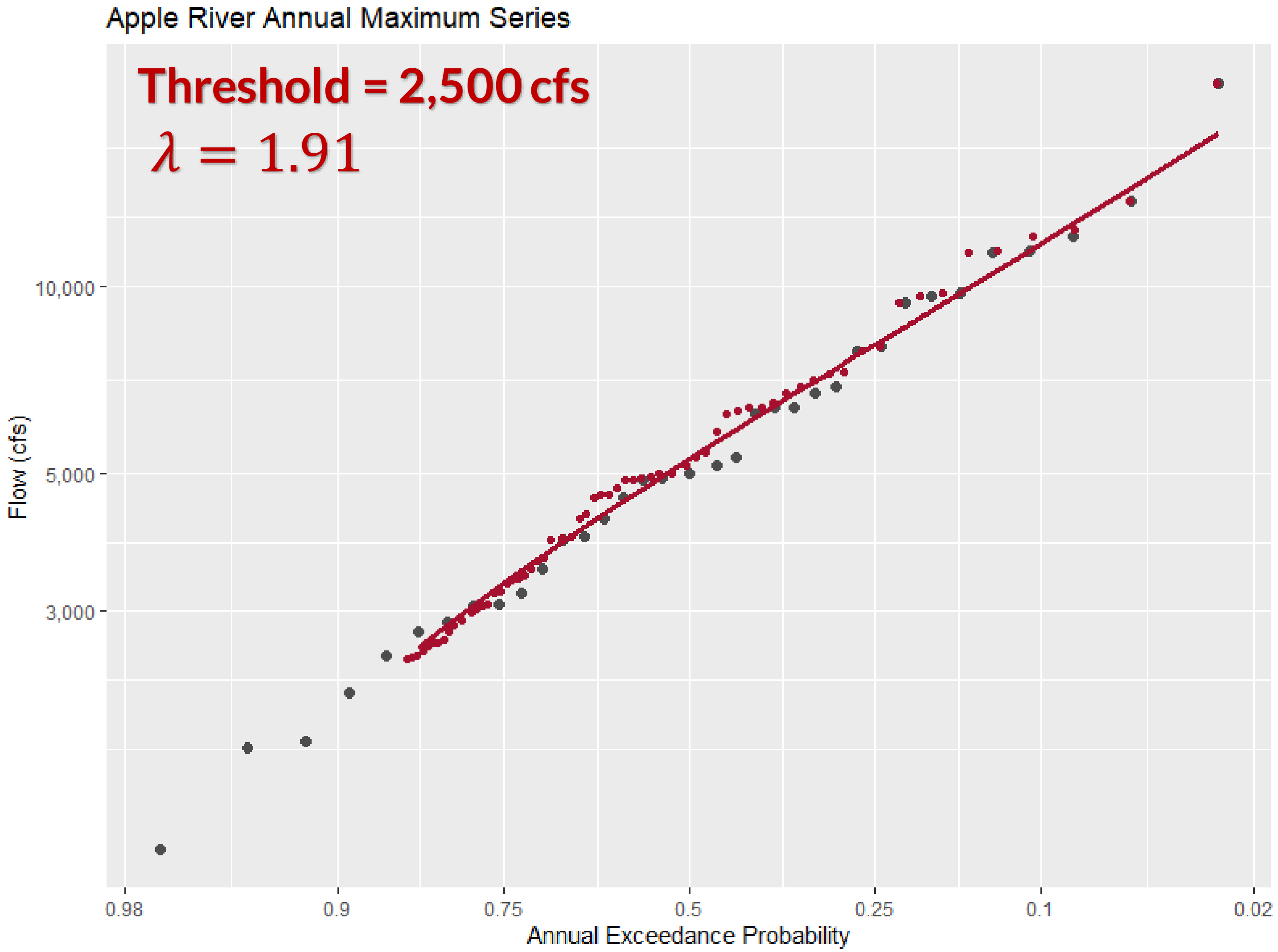
Partial Duration Series Limitations

- Usually has to be done on daily average data
  - Instantaneous data have a short record
  - Requires approximation of peaks
- Takes extra work to ensure independence
- Some subjectivity involved
  - Choosing a threshold and separation criteria can be hard!

US Army Corps of Engineers®

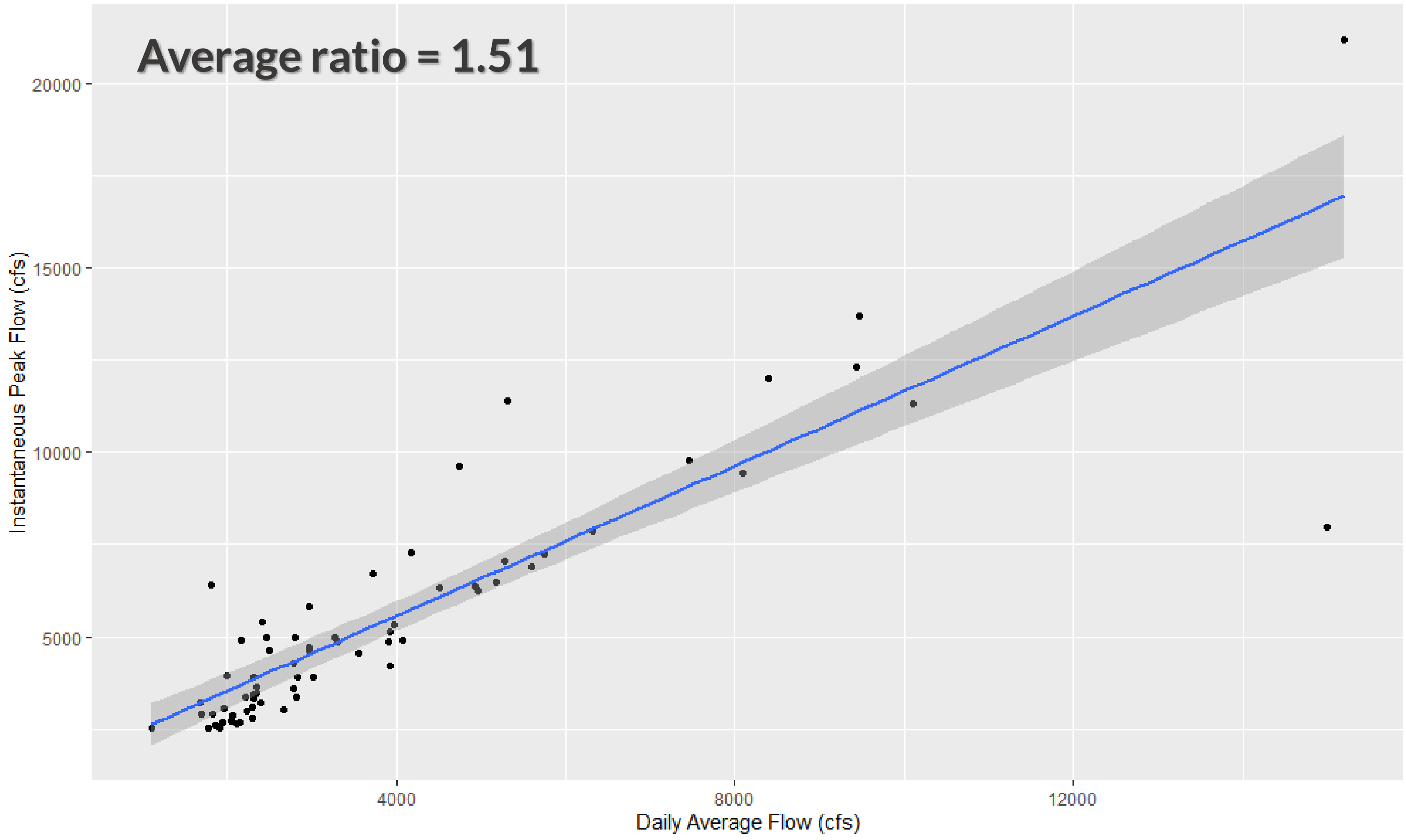
PDS with Instantaneous Data

- Shorter record, but
- No conversion



US Army Corps of Engineers®

Approximating Peaks



US Army Corps of Engineers®

Partial Duration Strengths

- Sample sizes often larger than AMS
- More samples of floods
- Automatic rejection of small annual maxima
- Lower error in estimating rare quantiles in many circumstances
- Samples are easier to sub-divide
  - Good for mixed population analysis

Mixed Population

- Split the PDS by flood-causing mechanisms
- Bigger sample size for each part compared to splitting AMS

US Army Corps
of Engineers®

Seasonal Analysis Example (Apple River)

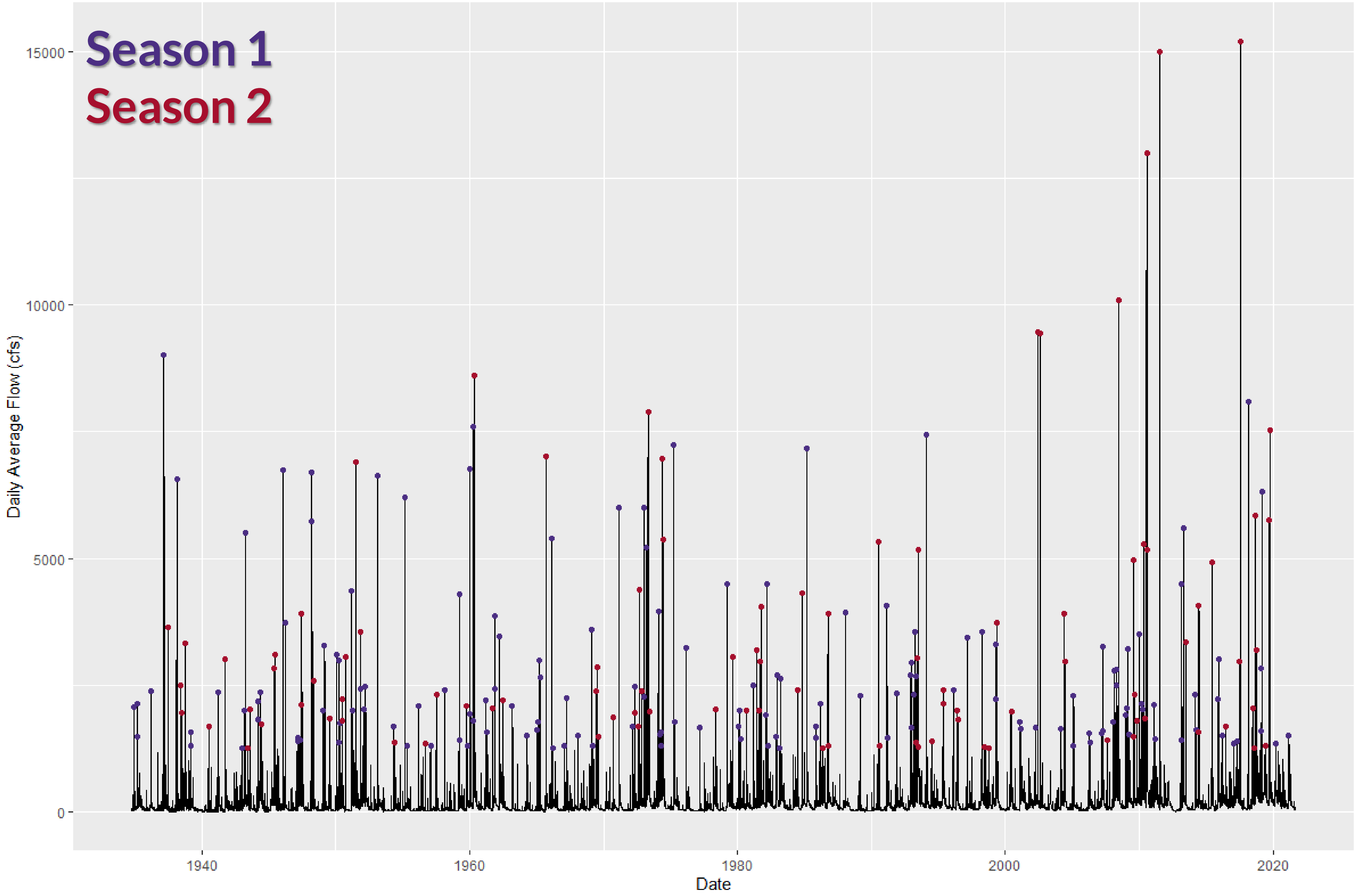
Season 1

- Nov – Apr
- 1.75 events/year
- Mean event = 2,800 cfs
- Snow/synoptic meteorology

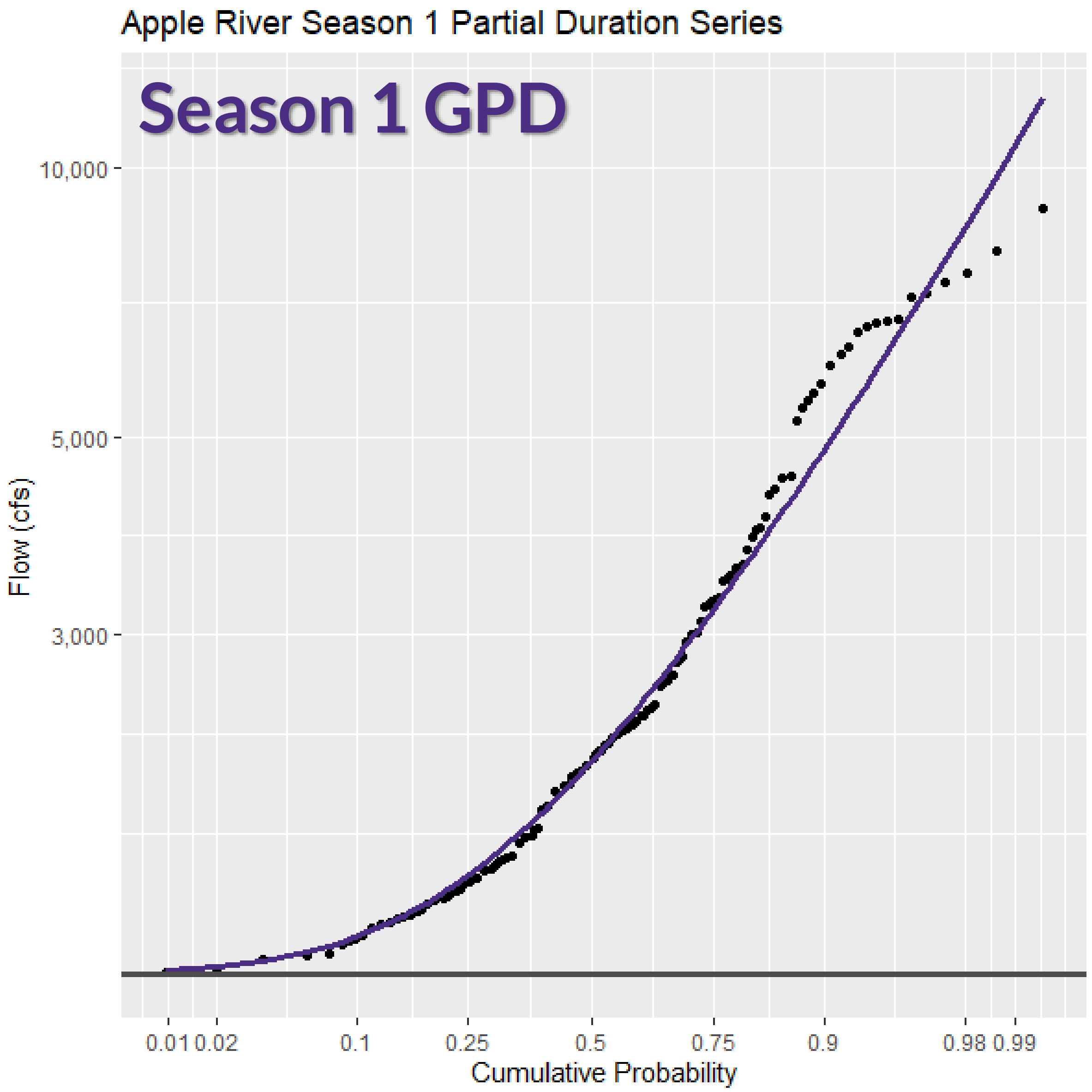
Season 2

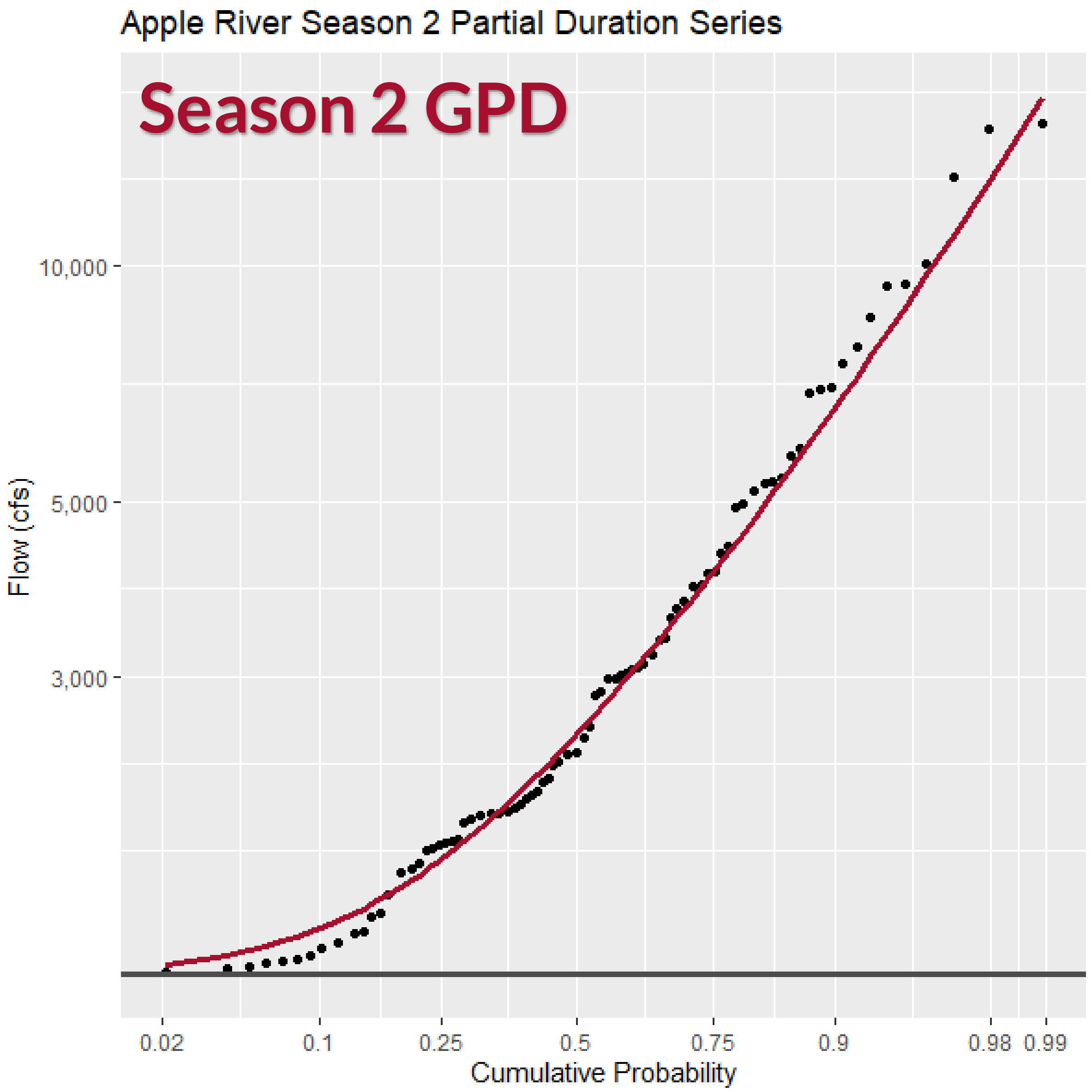
- May – Oct
- 1.10 events/year
- Mean event = 3,500 cfs
- Convective meteorology

US Army Corps of Engineers®



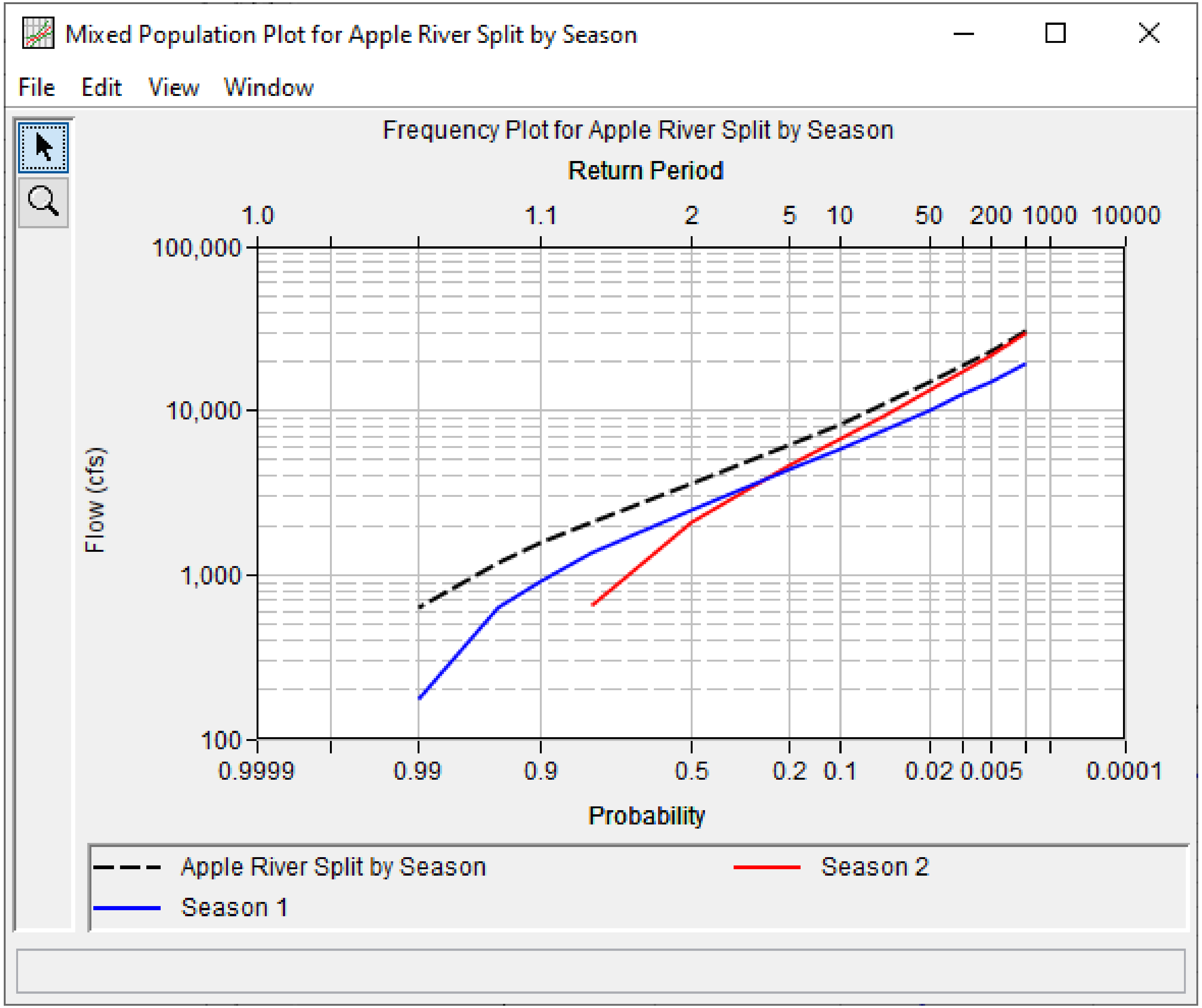
US Army Corps of Engineers®



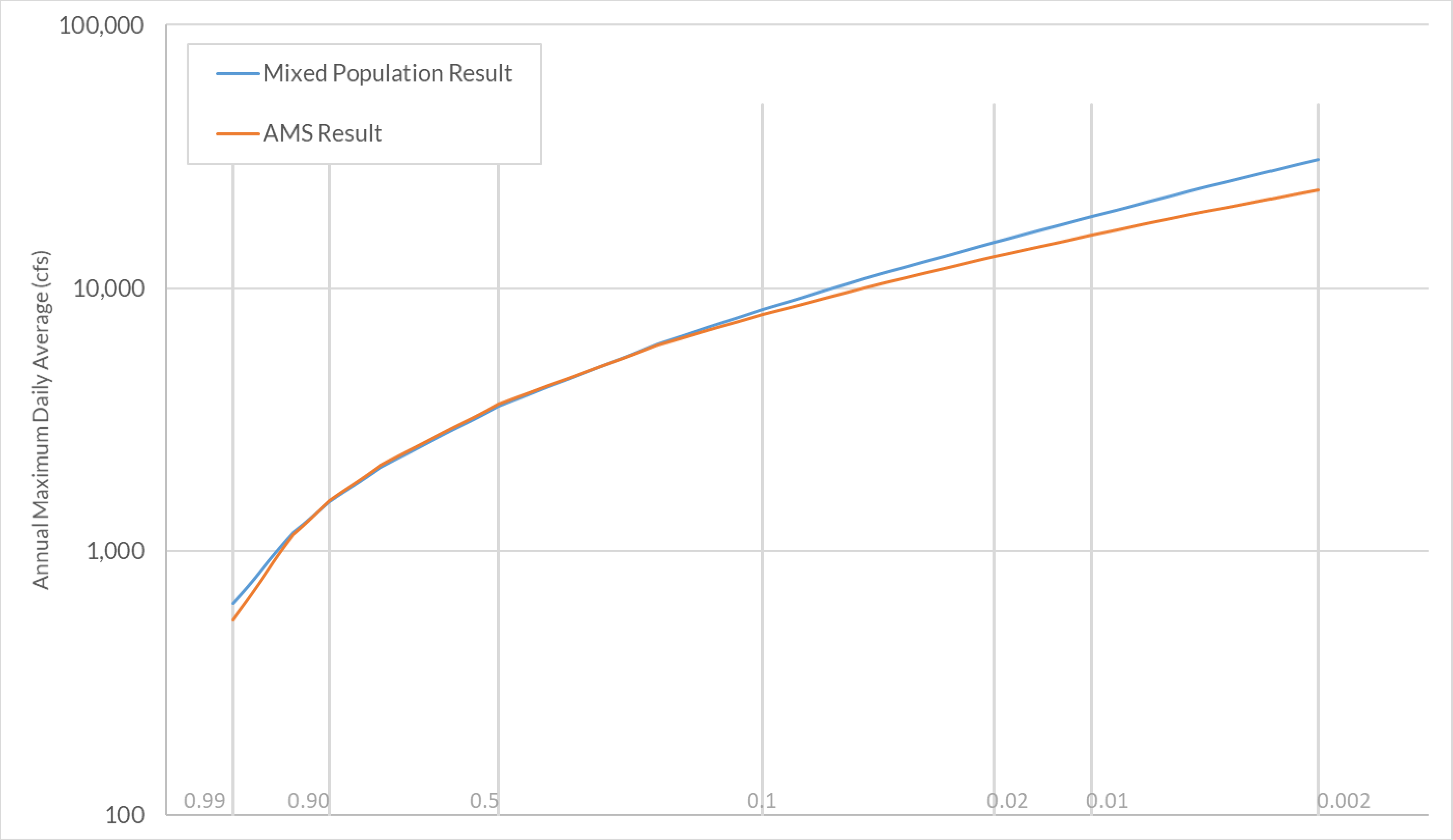


US Army Corps
of Engineers®

- Analytical annualization
- Compute mixed curve from equivalent GEV



US Army Corps of Engineers®



US Army Corps of Engineers®

Questions?

US Army Corps
of Engineers®