Lecture 1.5

Analyzing Hydrologic Time-Series
a road map

Flood Frequency Analysis
Hydrologic Engineering Center
Beth Faber, PhD, PE

This lecture is a bit of a review and connect the dots of some material already covered, a preview of a topic that's coming up, and then some new topics that fit within the theme.

Goals

Review (or preview) some ways we might study a time-series, and what we can learn from it

1. What information we extract
2. What assumptions we make
3. What values or probabilities we can estimate
4. What type of analysis each requires

2

Topics

Assumption

↓

- Annual Extremes — independence
  - annual maximums or minimums, Bulletin 17B/C
  - Instantaneous flows or longer duration avg flow/volume
- Partial Duration (peaks over threshold) — independence
- Daily flows or stages — No independence - persistence
  - Duration Curves
  - Summary Hydrographs

3

These are the topics I'll cover. We divide the time period of the data, and whether we use all the data or isolate the extremes. These choices affect the assumptions we can make about the data set. Isolating annual extremes provides enough time between values that we can assume independence. Using all the data at a daily step, the values are persistent, and can't assume independence.

Topics

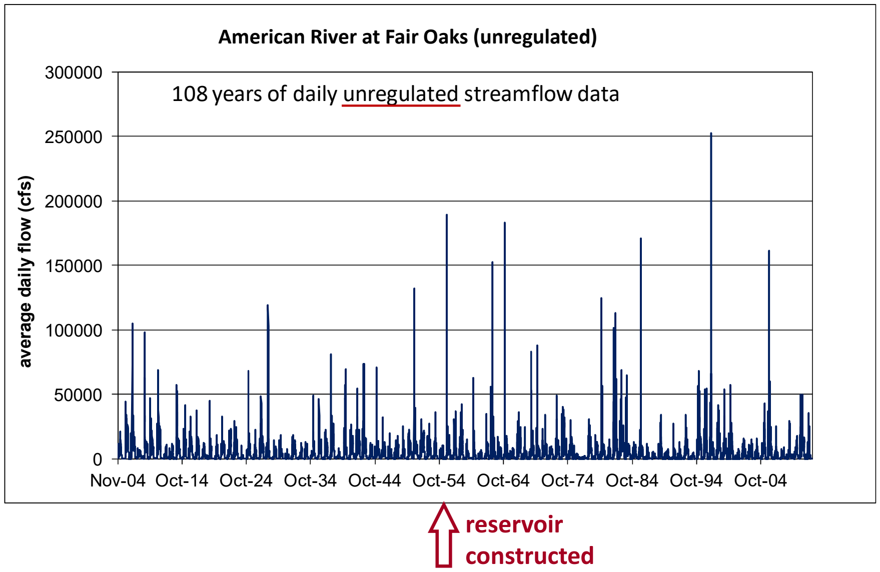
Assumption ↓

- Annual Extremes — independence
  - annual maximums or minimums, Bulletin 17B/C
  - Instantaneous flows or longer duration avg flow/volume
- Partial Duration (peaks over threshold) — independence
- Daily flows or stages — No independence - persistence
  - Duration Curves
  - Summary Hydrographs

4

These are the topics I'll cover. We divide the time period of the data, and whether we use all the data or isolate the extremes. These choices affect the assumptions we can make about the data set. Isolating annual extremes provides enough time between values that we can assume independence. Using all the data at a daily step, the values are persistent, and can't assume independence.

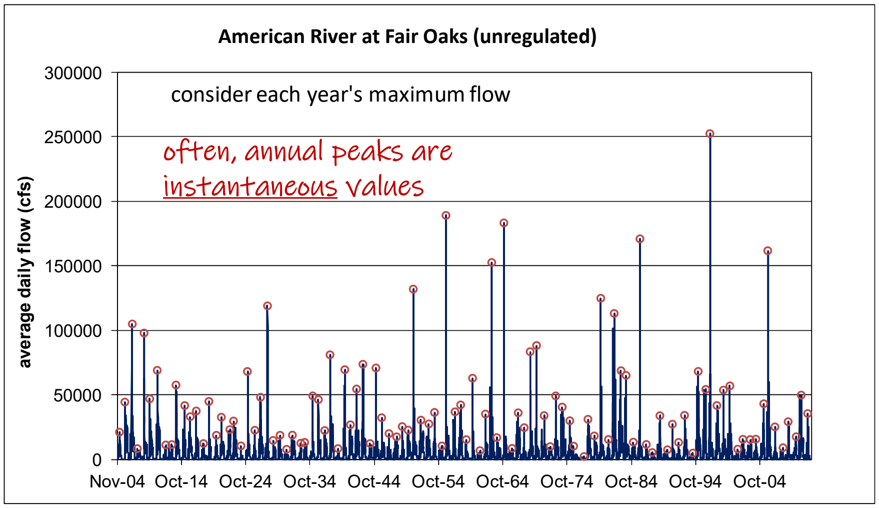
Streamflow data, daily flow record



5

Basically, this lecture is about the various ways we use streamflow time-series data to perform different probabilistic analyses. This is a 108 year daily time-series for the American River in California, and I'll use it to demonstrate many of the topics.

Streamflow data, find annual max

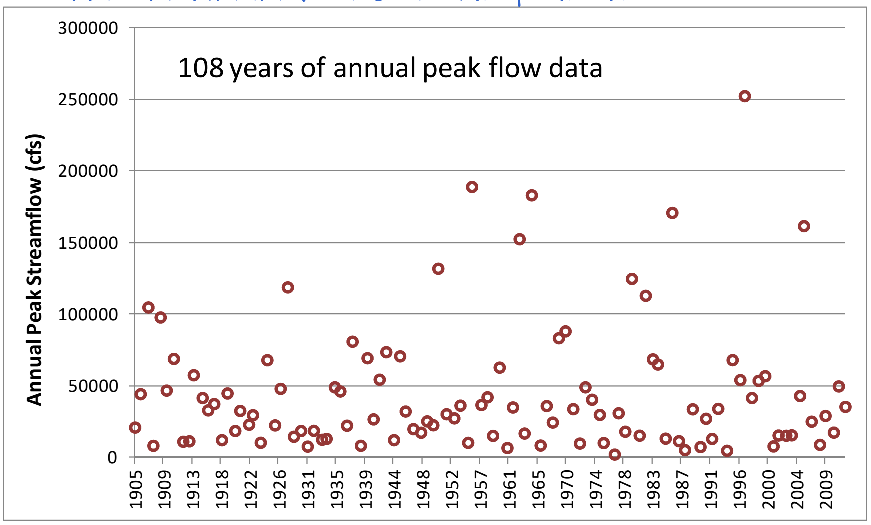


6

This is how we generate the annual maximum series – choosing the largest value of each year, based on some definition of "year," such as calendar (starting Jan 1) or water (starting Oct 1).

American River, ANNUAL MAX SERIES, AMS

annual maximum values are independent



7

The data set is now 108 values of annual maximum/peak flows.

Assumptions

- What we have: series of annual maximum flows
- Treat these flows as a random, representative sample from the flood population of interest
- *Generally, we assume the sample is* ***IID***
  - annual peak flows are random and independent
  - peak flows are identically distributed – homogeneous data set
  - sample is adequately representative of the population

    (estimate of the distribution improves with sample size)

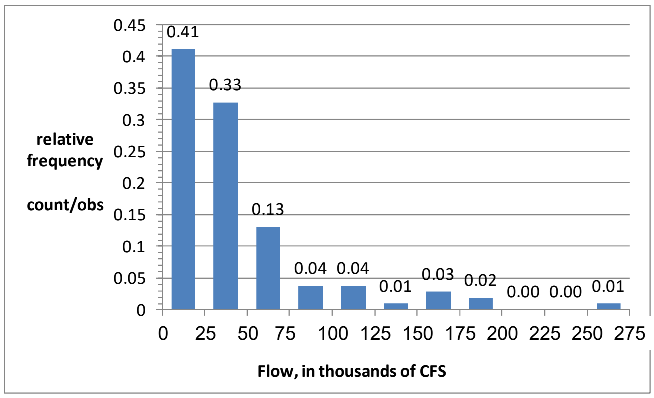
NOTE, we're assuming the data is homogeneous, stationary….

8

IID = independent and identically distributed. Not so much that they ARE identically distributed as that we assume it is reasonable to fit a single distribution to them. Except when we can't assume that….

Histogram

divide data range into bins, count number of observations in each bin (frequency), divide by total (relative frequency)



this shape is similar to a PDF, probability density function

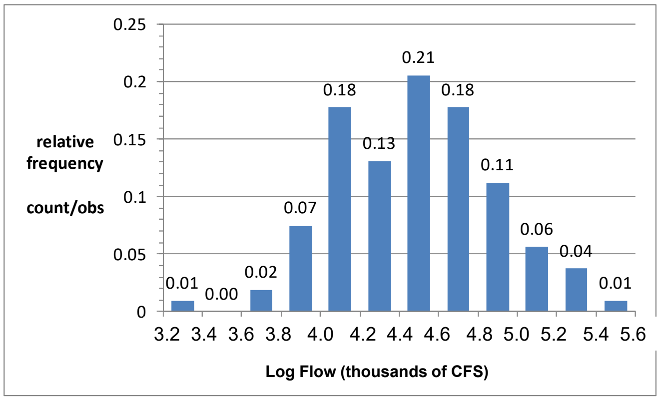
9

A histogram gives a good view of probability of the sample.

Histogram

divide data range into bins, count number of observations in each bin (frequency), divide by total (relative frequency)

switch to log(flow)



this shape is similar to a PDF, probability density function

10

Flow often is easier to work with after transforming to base 10 logarithm.

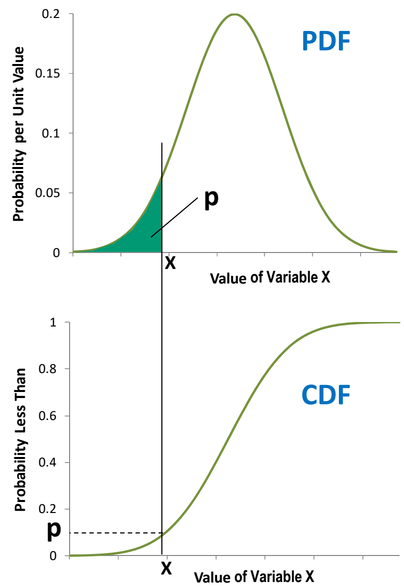
Definition of PDF and CDF

A Probability Density Function (PDF), f(x), defines the probability of occurrence for a continuous random variable.

area under curve = probability

The Cumulative Distribution Function (CDF), F(x) is the probability the random variable is less than some value

curve = probability



11

Empirical Cumulative Distribution

→ based on observation

Empirical or Graphical estimate: based on order statistics, using plotting position (estimated non-exceedance probability)

- Cumulative non-exceedance probability of observation magnitude is based on its position within the sample
- Each observation has incremental prob = 1/N

i.e., relative frequency

$$P[X \le x_i] = \frac{m_i}{N}$$

m_i = rank of ordered observation i
(m=1 as smallest value, N as largest)
N = total # of observations

estimate probability that will be smaller than observation x_i by the fraction of the observed sample that is smaller

12

Rank is where that sample member is in the line-up of all sample members ordered by magnitude. So, it's a way of counting the number of events less than or equal to.

Plotting position is the estimate of the cumulative prob of each observation, based on the sample size and the position of each observation within the sample that's been sorted by magnitude. So, effectively we're computing the relative frequency of being less than a given observation, based on how many values in the sample were less than that observation – defined by the rank.

Empirical Cumulative Distribution

Plotting Position:

General Formula

$$P = \frac{m + a}{N + b}$$

Weibull: a = 0.0, b = 1.0

Median: a = -0.3, b = 0.4

plotting position

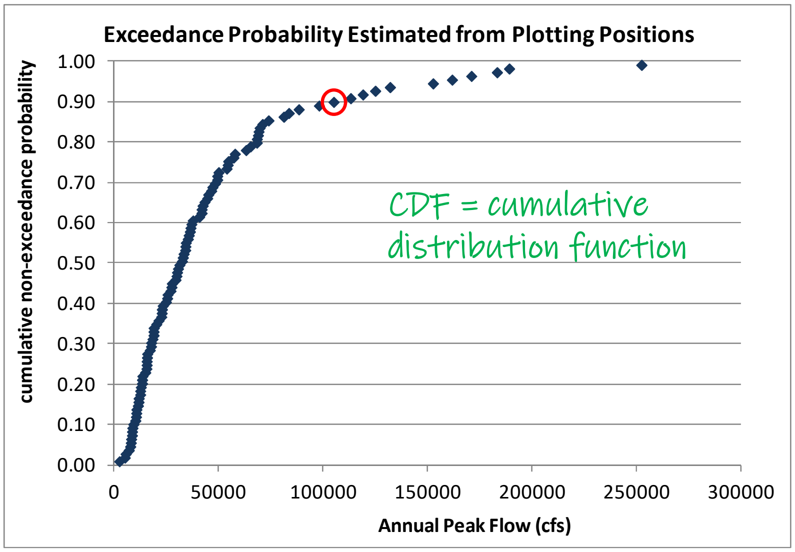
Year	Rank	Flow	Non-Excdnc Probability
1997	108	252431	0.99
1955	107	189073	0.98
1964	106	183242	0.97
1986	105	170960	0.96
2005	104	161731	0.95
1963	103	152614	0.94
1950	102	132000	0.94
1980	101	124915	0.93
1928	100	119000	0.92
1982	99	113126	0.91
1907	98	105000	0.90
1909	97	98000	0.89
1970	96	88316	0.88
1969	95	83526	0.87

Year	Rank	Flow	Non-Excdnc Probability
1955	14	10528	0.13
1975	13	10389	0.12
1972	12	10046	0.11
2008	11	9150	0.10
1966	10	8659	0.09
1939	9	8500	0.08
1907	8	8460	0.07
2001	7	8045	0.06
1931	6	7920	0.06
1990	5	7606	0.05
1961	4	6914	0.04
1988	3	5447	0.03
1994	2	5009	0.02
1977	1	2359	0.01

13

These are the largest 14 and smallest 14 values in the 108-year data set.

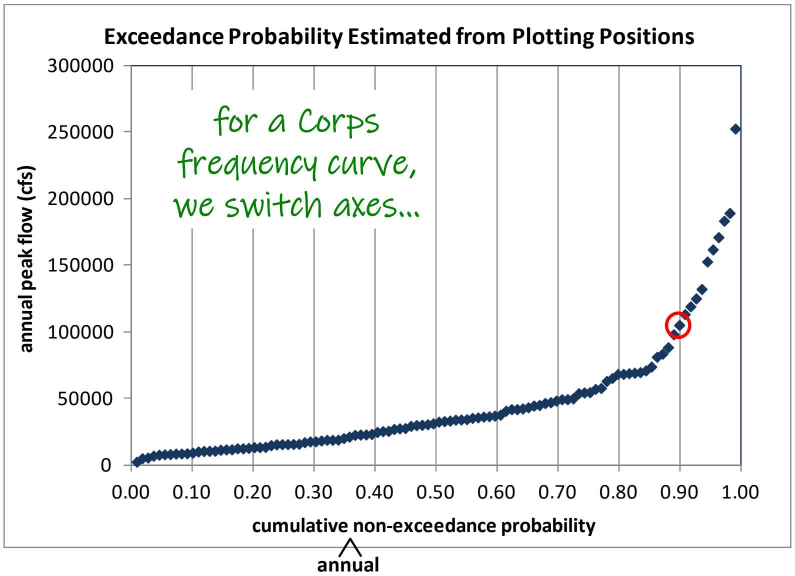
Empirical Cumulative Distribution



14

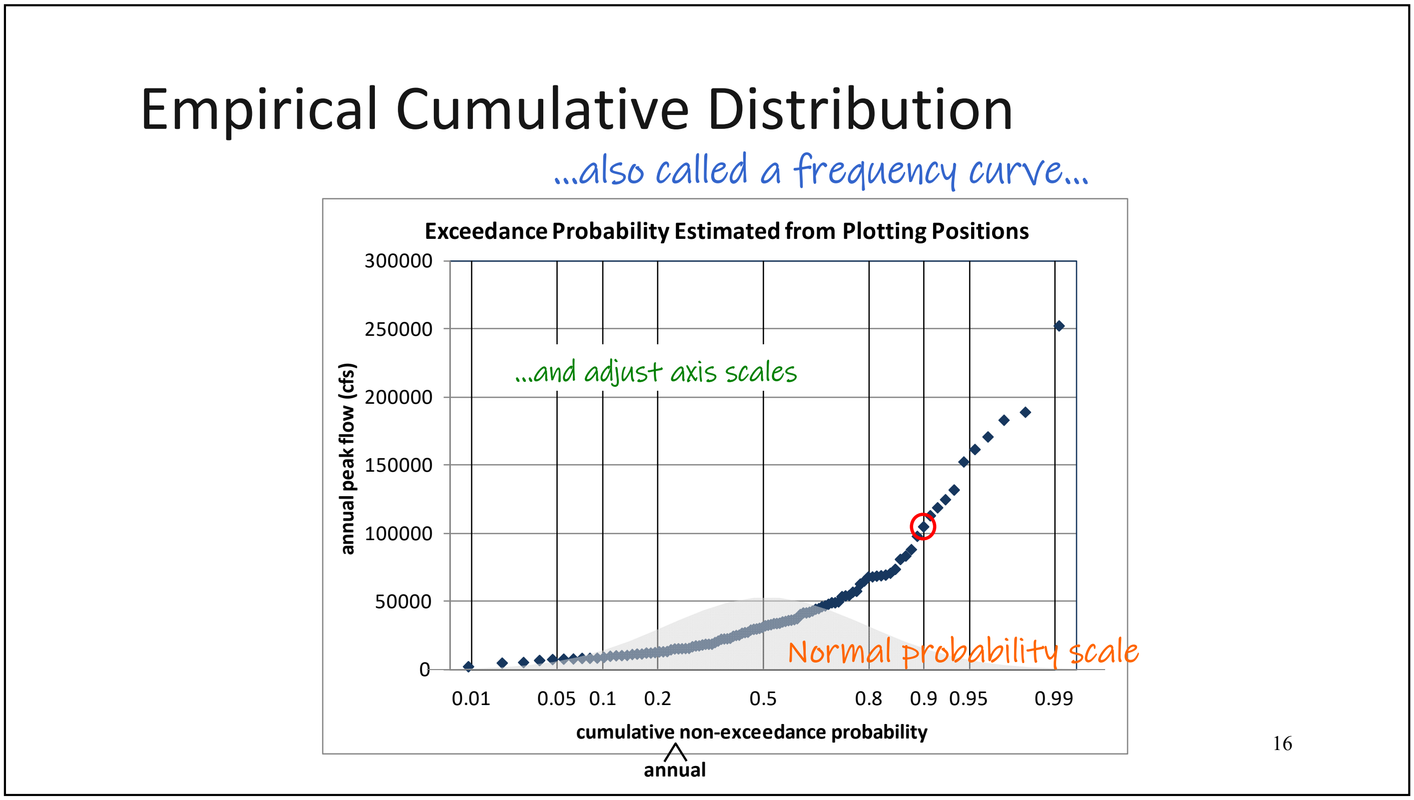
First, plot in the typical CDF form, with flow on horizontal and cumulative probability on vertical. But, now can plot every observation, rather than every bin.

Empirical Cumulative Distribution



15

Switch axes to probability on horizontal and probability on vertical. Note, annual probability because one value each year, and computing relative frequency based on years.



Switch probability axis from linear to Normal. Note, normal data would plot as a straight line on normal axis.

How to we create a normal axis?

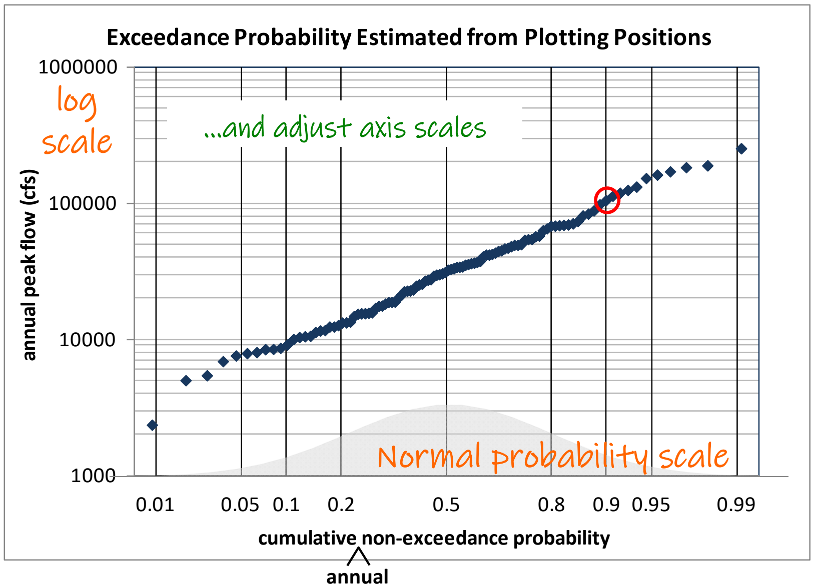
(1) the axis is linear in Zp, the standard normal deviate of probability p.

The axis is based on the normal PDF that you can see as a shadow along the horizontal, so

(2) we plot a given probability where the area under the PDF (from left to right) equals that probability.

Empirical Cumulative Distribution

...also called a frequency curve...



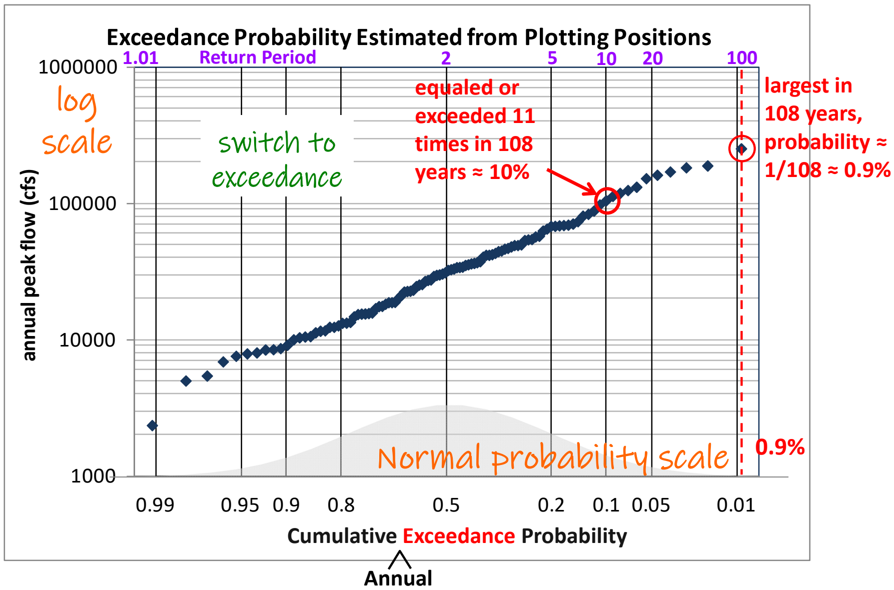
17

Switch vertical axis to log(flow).

Note, the data becomes close to a straight line, so log(flow) is close to Normal, and flow is close to logNormal in this case.

Empirical Cumulative Distribution

...also called a frequency curve...

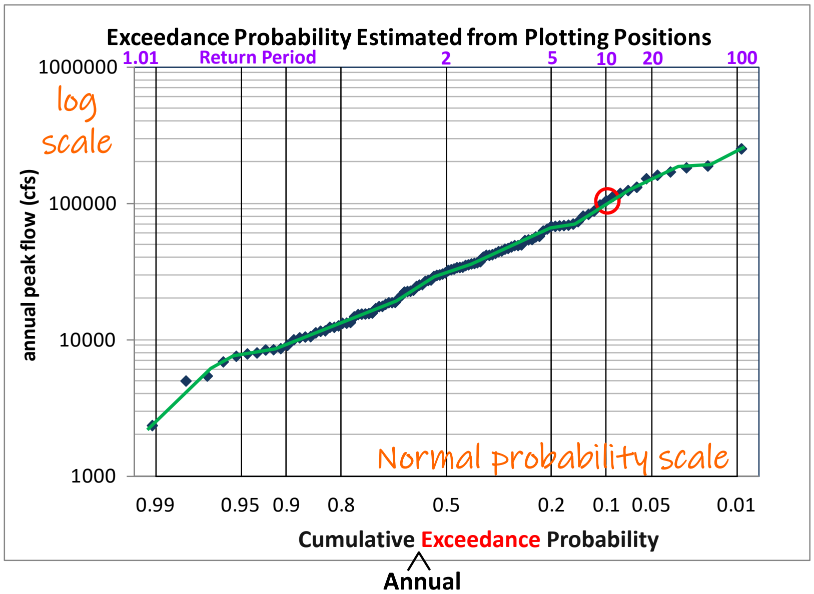


18

Switch from non-exceedance to exceedance probability. Note that return period = 1 / (exceedance prob).

Recall the interpretation of the plotting positions. The largest even is explicitly interpreted as equaled or exceeded once in 108 years, etc

Empirical Graphical Cumulative Distribution

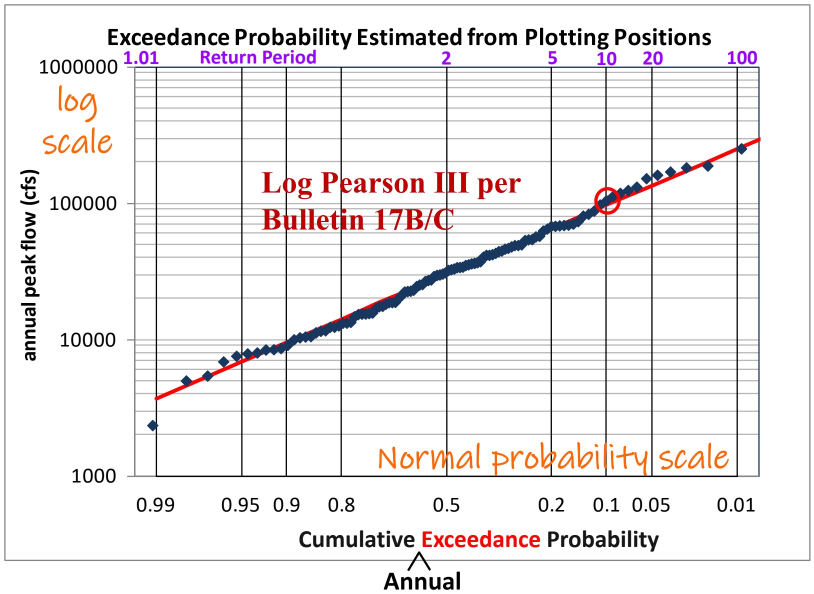


19

With no further assumptions, we have an empirical or graphical curve only.

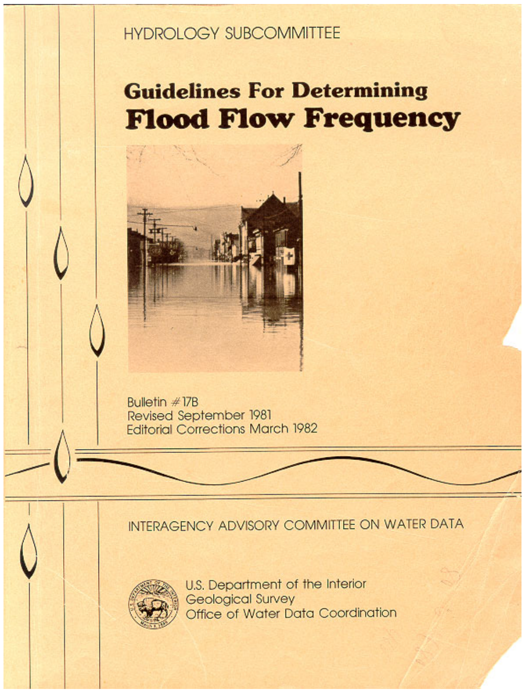
Analytical Distribution

for unregulated, instantaneous annual maximums



20

Or, we can assume an analytical distribution, and estimate its parameters from the 108 member sample of values. This is the log Pearson type 3 distribution recommended in Bulletin 17B/C for unregulated annual maximum values



Bulletin 17B

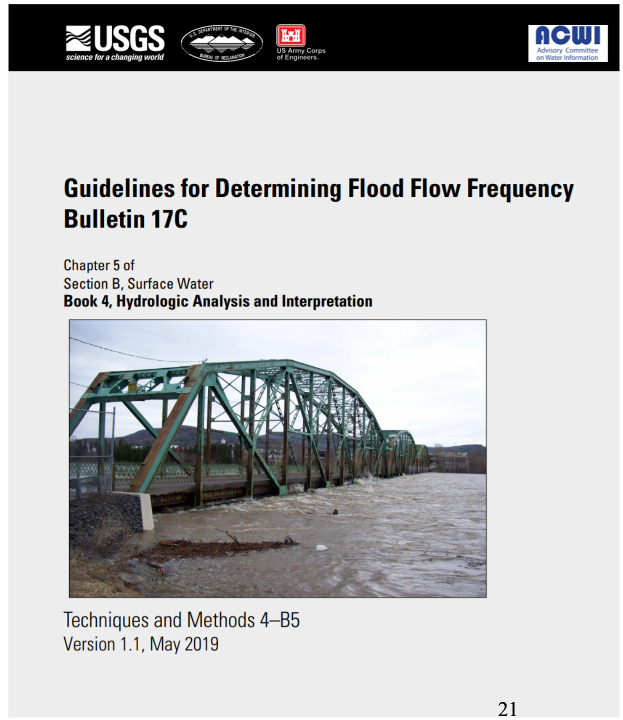
dated March 1982

←

Bulletin 17C

dated Feb 2018

→



Reference

B17B/C: Frequency Analysis of Annual Peak Flows

- Estimate **LogPearson III** distribution for unregulated annual peak flows
  - **Method of Moments** → *moments of sample used to estimate moments of distribution*
- Data challenges:
  - Missing Flows (Broken or Incomplete Record)
  - Low and High Outliers
  - Zero Flows
- Additional information to improve estimates:
  - Historical/Paleo Information
  - Regional Skews (use a *weighted* skew)
  - Allowing a longer record site to improve short record estimate

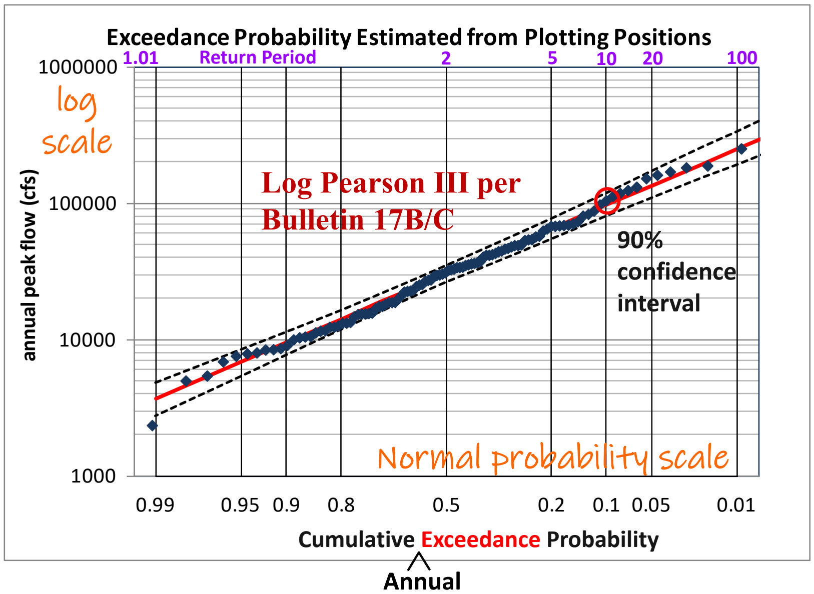
many of the methods for handling data challenges and using additional information are improved in Bulletin 17C

22

Fitting an LP3 to annual maximums is the primary guidance. But the details are in dealing with difficulties in the data set, and in bringing in additional data that can provide more information.

Analytical Distribution

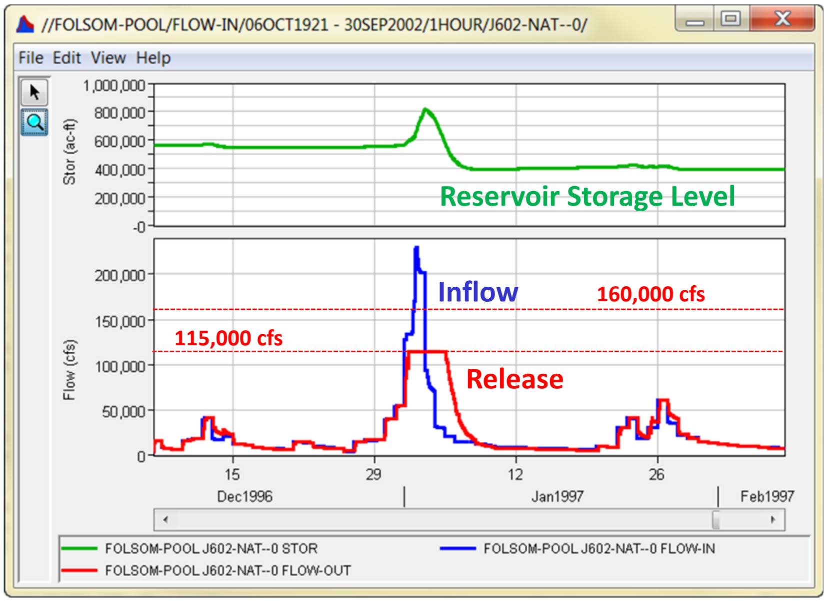
for unregulated, instantaneous annual maximums



23

The fit of the LogPearson type 3 distribution (LP3) also includes estimating the 90% confidence interval. The interval can be interpreted simply as meaning there is a 90% change the correct population frequency curve is within this range.

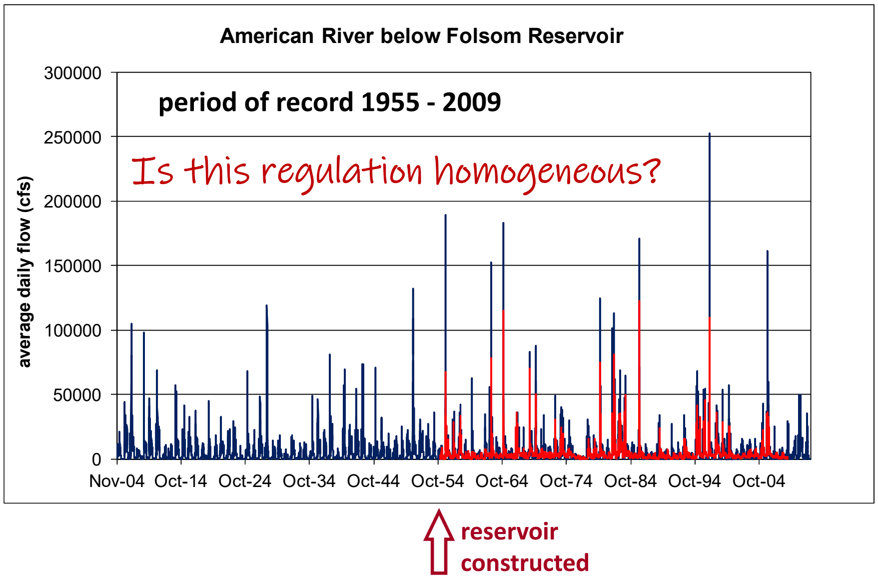
Also Interested in **Regulated** Flow



24

This figure shows the regulation of a flood event through a reservoir, where the reservoir stores all of the water volume in excess of the same channel flow or "objective release" and then releases that volume more slowly after the event.

Streamflow data, regulated

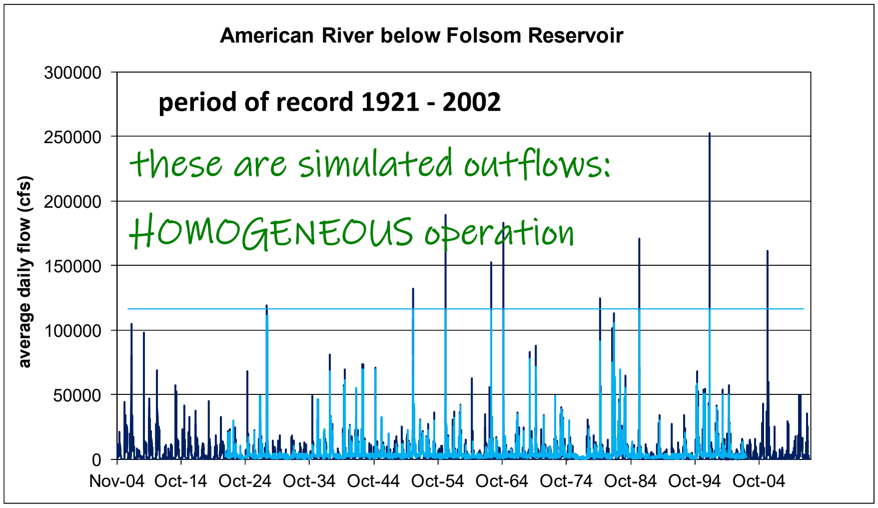


homogeneous = identically-distributed = stationary

25

The red line is the reservoir outflows, beginning when the reservoir was constructed. These actual releases may have been following different operation strategies over the last 60 years. We are less interested in what was done in the past as we are in using the past to study what could happen in the future. So, we need to see consistent operations.

Streamflow data, regulated



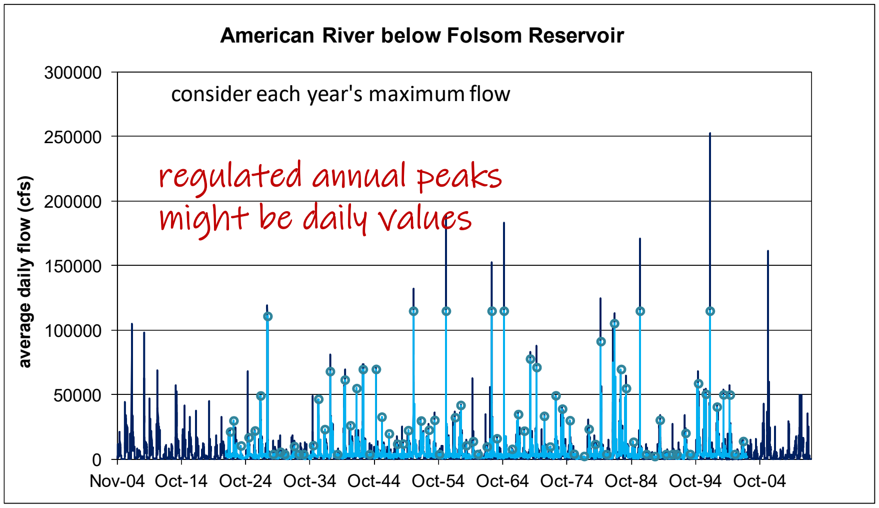
2 benefits:

homogeneous operation, and longer period

26

Here, we see the unregulated record simulated with the current operation strategy. So, the operation is homogeneous. Hopefully, the inflow record is, too!

Streamflow data, regulated



27

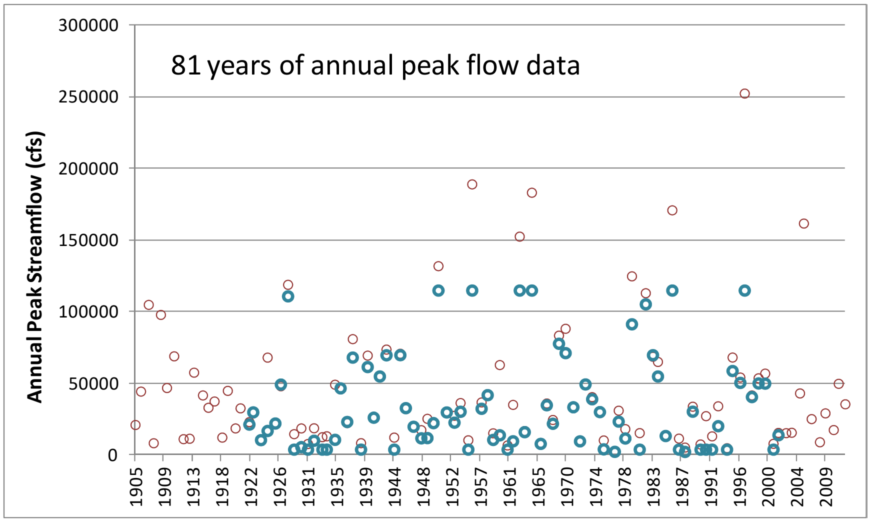
27

Create the annual maximum series (AMS) in the same way we do for unregulated flows.

Note, the simulated period is not the complete unregulated record.   So, the POR is a bit shorter.

Folsom Release, ANNUAL MAX

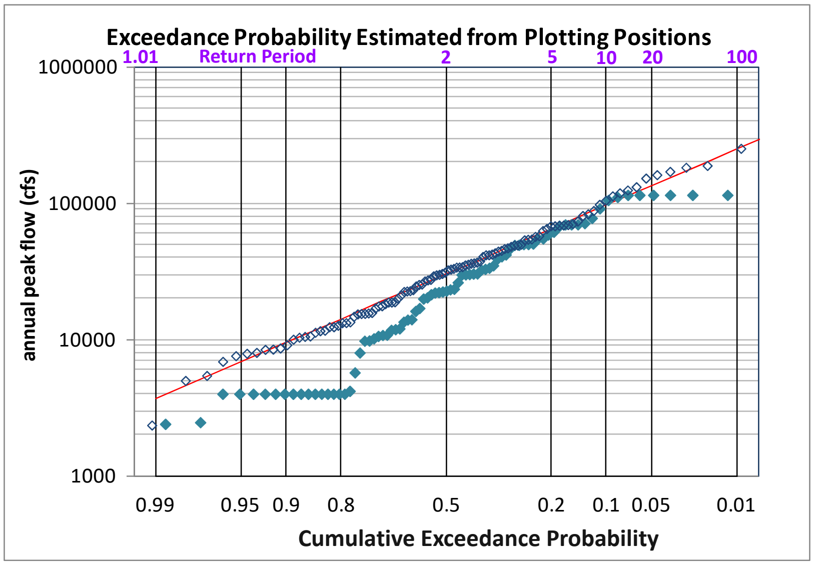
annual maximum values are independent



28

Now we have a record of regulated peaks.

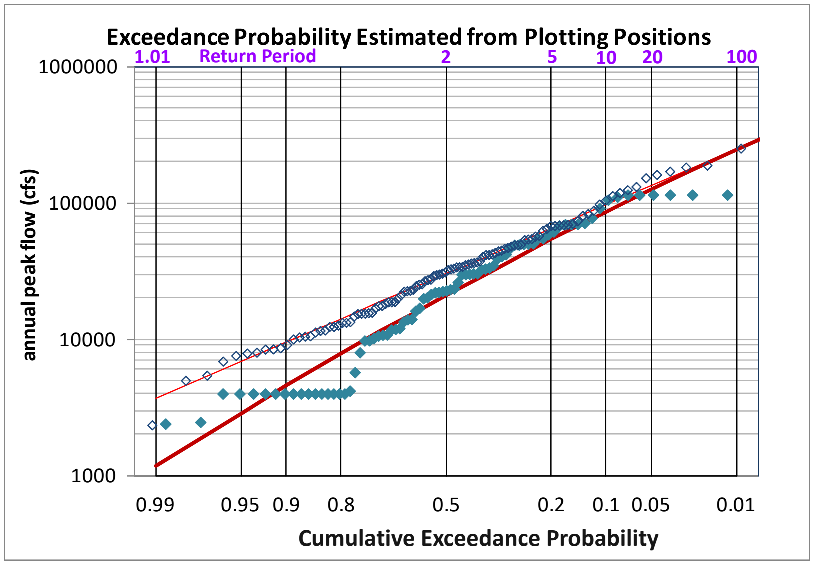
Plotted Annual Maximums



29

Hollow diamonds are the unregulated record. Solid diamonds are the regulated record, plotted in the same way. Note that there are relevant flow thresholds, such as channel capacity at the upper end, that are horizontal over a range of exceedance probability. These thresholds are important!

Does an Analytical Curve Make Sense?

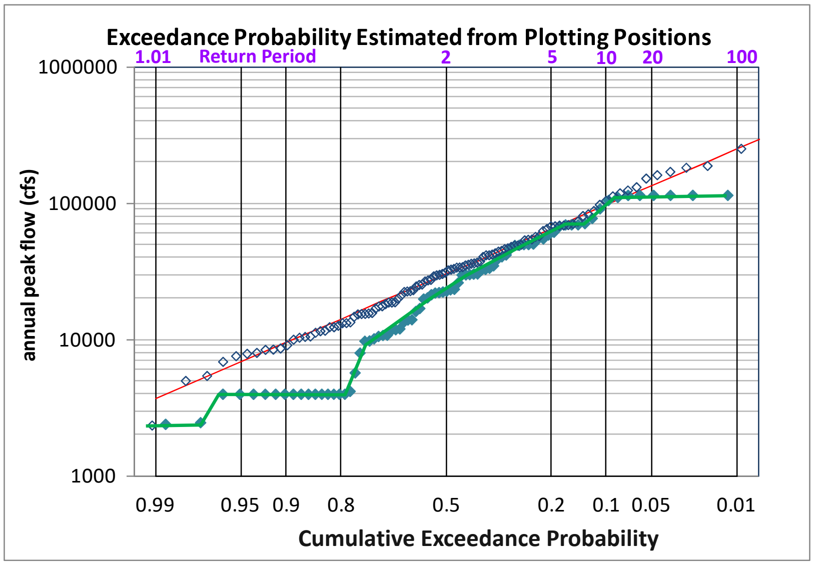


30

An analytical fit to this data would smooth through the important flat areas that represent flow thresholds. This is not what we want.

Sometimes we do want to smooth through sampling error. But, in this case, the flats spots are meaningful.

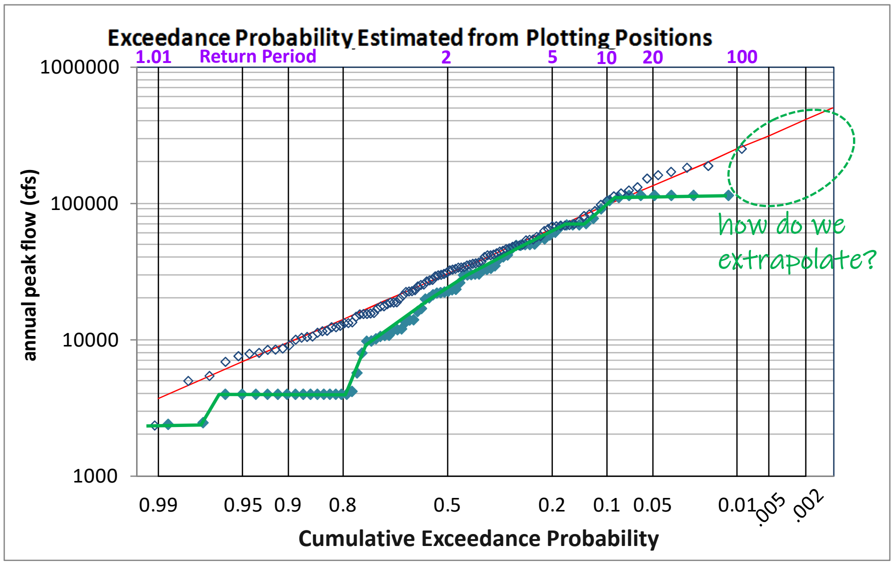
Graphical Curve is Better for this



31

It is better to do a graphical fit to this data, so we can capture the thresholds that are relevant, and smooth through the ones that are not.

Graphical Curve is Better for this



32

A problem with graphical frequency curves is that there is no basis for extrapolation beyond the range of the data (other than the clear upper bounds offered by the regulated frequency curve.) How should we extrapolate?

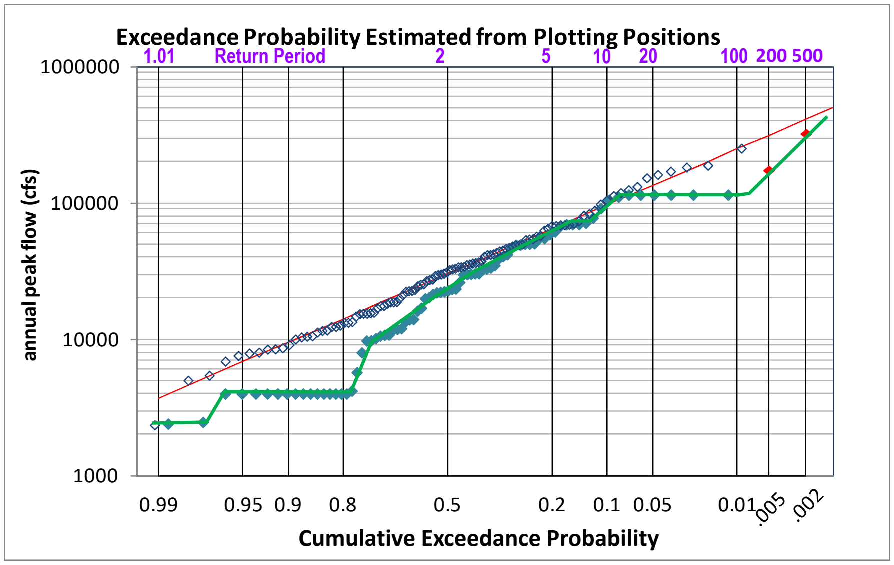
Extending a Regulated Curve

- To extend a regulated frequency curve, we can **construct synthetic events** with a specified frequency, e.g. 100-yr, 200-yr, etc
  - Estimate frequency curves for longer duration average flows or volumes *(3-day, 7-day, 30-day, etc)*
  - Construct hydrograph for a given frequency
  - Route through operation model to get peak outflow
  - ***NOTE: will assume likelihood of peak release is related to likelihood of peak inflow***

33

We extrapolate a graphical curve by estimating higher values (lower probability) regulated flows based on unregulated flood events of the desired probability.

Graphical Curve is Better for this



34

Using synthetic events, we can estimate points higher on the frequency curve, to fill in the gap.

Topics

Assumption ↓

- Annual Extremes — independence
  - annual maximums or minimums, Bulletin 17B/C
  - Instantaneous flows or longer duration avg flow/volume
- Partial Duration (peaks over threshold) — independence
- Daily flows or stages — No independence - persistence
  - Duration Curves
  - Summary Hydrographs

35

These are the topics I'll cover. We divide the time period of the data, and whether we use all the data or isolate the extremes. These choices affect the assumptions we can make about the data set. Isolating annual extremes provides enough time between values that we can assume independence. Using all the data at a daily step, the values are persistent, and can't assume independence.

Longer Durations of Flow/Volume

Volume Frequency Analysis

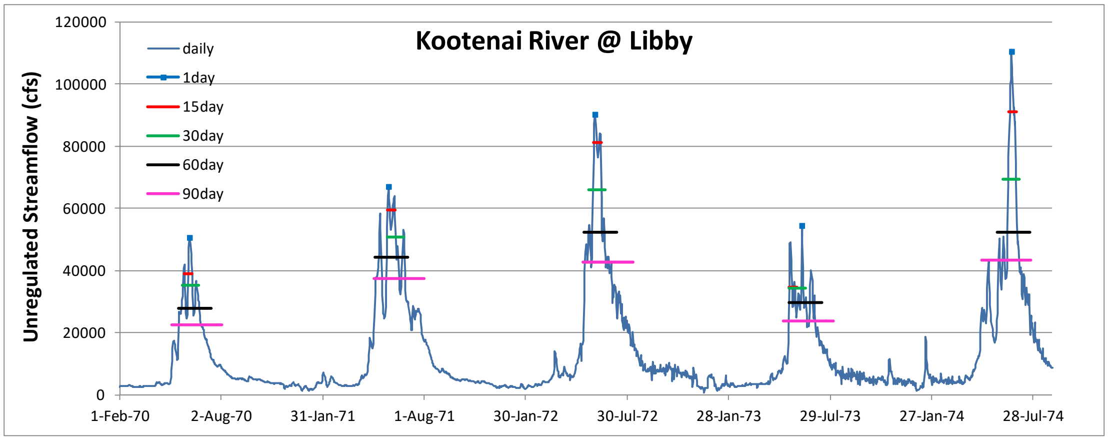
Create a family of frequency curves for a given site that specify average flow across various durations

1. Compute X-day moving average TS for various durations, X
   - X might be 1-day, 5-day, 30-day etc.
2. Extract annual maximums for each duration
3. Perform frequency analysis for each duration
   - Similar to Bulletin 17B/C procedure, but not instantaneous peak
   - Compute plotting positions and distribution parameters from each period of record sample

36

Why are we looking at longer durations? When reservoir are involved for providing flood storage, we're more interested in the volume of water that will arrive across some period of time than just the highest flow. So we want to estimate the likelihood of flow volumes.

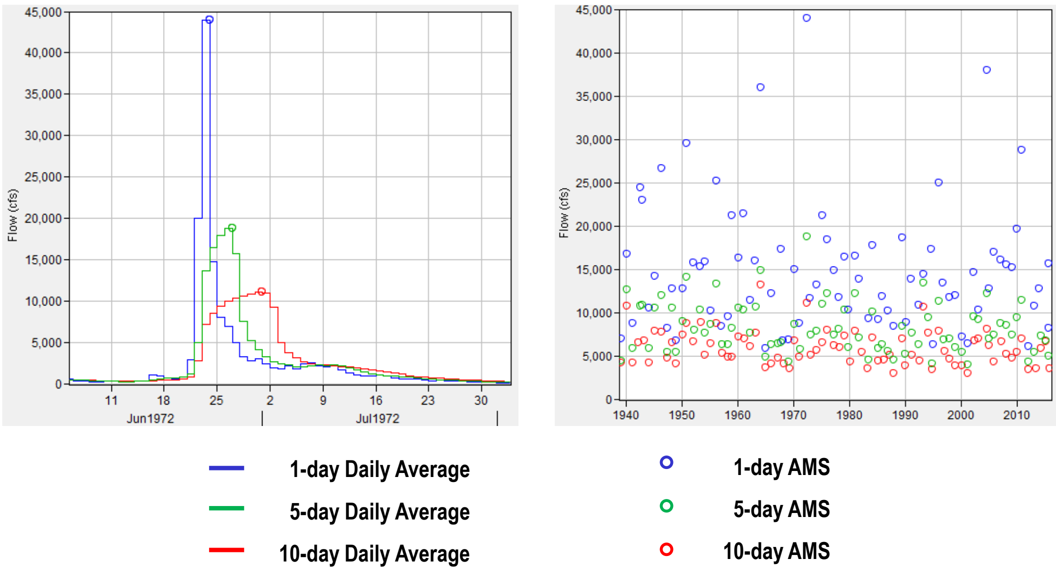
Longer Duration Max Average Flows



37

This is a location where the highest flows are snow-melt floods that last longer than rain floods. It lets us see the annual maximum average flows more easily.

Longer Durations Max Average Flows



38
38

These are rain floods, that are shorter – only a few days. On the left is the time-series of moving X-day averages for 1 event, with the maximum values noted. On the right is the record of annual maximums generated this way across the record.

Annual Maximums of Various Durations

Volume-Duration Data

Highest Mean Value for Duration, Average Daily FULL NATURAL FLOW in CFS

Year	1		3		7		15		30	
	Date	FULL NAT...	Date	FULL NAT...	Date	FULL NAT...	Date	FULL NAT...	Date	FULL NAT...
1945	02/02/1945	70900	02/04/1945	40733	02/07/1945	23661	02/15/1945	14983	05/18/1945	11098
1946	12/22/1945	32400	12/30/1945	25533	12/28/1945	22086	01/05/1946	18041	01/19/1946	11576
1947	02/13/1947	20100	02/14/1947	12410	02/18/1947	7654	03/24/1947	5867	04/08/1947	5297
1948	04/18/1948	17600	04/19/1948	15233	04/23/1948	[illegible]	[illegible]	[illegible]	[illegible]	10561
1949	03/03/1949	25500	03/05/1949	15647	04/28/1949	[illegible]	[illegible]	[illegible]	[illegible]	10483
1950	02/06/1950	22800	01/24/1950	20067	01/24/1950	[illegible]	[illegible]	[illegible]	[illegible]	10053
1951	11/21/1950	132000	11/21/1950	107500	11/24/1950	61757	12/02/1950	31688	12/17/1950	30557
1952	02/02/1952	30500	02/04/1952	20800	05/29/1952	19800	06/01/1952	19360	06/09/1952	18423
1953	04/28/1953	27600	04/29/1953	20867	04/30/1953	15571	05/07/1953	12411	05/21/1953	10350
1954	03/10/1954	36500	03/11/1954	26100	03/15/1954	16126	03/23/1954	10651	05/04/1954	9301
1955	05/09/1955	10528	05/23/1955	[illegible]	05/13/1955	9850	05/23/1955	8585	06/03/1955	7786
1956	12/23/1955	189073	12/24/1955	127450	12/27/1955	70984	01/02/1956	40608	01/17/1956	28224
1957	05/19/1957	36924	05/20/1957	23572	05/24/1957	15854	03/10/1957	13393	03/23/1957	9654
1958	04/03/1958	42302	04/03/1958	33635	04/05/1958	25279	04/04/1958	19005	06/01/1958	17154
1959	02/17/1959	15394	02/19/1959	12262	02/22/1959	8854	03/02/1959	5875	05/01/1959	4631
1960	02/08/1960	63014	02/10/1960	34802	02/14/1960	18198	02/16/1960	10095	04/05/1960	8010
1961	04/04/1961	6914	04/05/1961	6455	04/08/1961	5417	05/24/1961	4926	05/27/1961	4425
1962	02/10/1962	35216	02/16/1962	20717	02/16/1962	19671	02/23/1962	12694	05/10/1962	9814
1963	02/01/1963	152614	02/02/1963	93881	02/06/1963	49107	02/14/1963	26738	02/28/1963	15499
1964	11/15/1963	17002	01/22/1964	9561	05/19/1964	8886	05/25/1964	7987	05/27/1964	6583
1965	12/23/1964	183242	12/24/1964	140371	12/27/1964	87834	01/04/1965	50629	01/18/1965	33147
1966	04/02/1966	8659	04/03/1966	8414	04/08/1966	8169	04/13/1966	7964	04/28/1966	7110
1967	03/17/1967	36197	03/18/1967	29825	05/26/1967	22572	05/30/1967	19699	06/14/1967	15752
1968	02/21/1968	24697	02/22/1968	22341	02/26/1968	17942	03/03/1968	12405	03/17/1968	8709
1969	01/21/1969	83526	01/22/1969	71862	01/26/1969	49450	02/02/1969	32444	02/17/1969	21018
1970	01/22/1970	88316	01/23/1970	68756	01/23/1970	49561	01/28/1970	38627	02/08/1970	23501
1971	03/26/1971	34047	03/28/1971	25337	04/01/1971	16884	04/08/1971	12270	04/22/1971	9916
1972	03/04/1972	10046	03/06/1972	9585	03/10/1972	8824	03/18/1972	8471	03/25/1972	7823
1973	01/12/1973	49291	01/14/1973	31181	01/18/1973	25659	01/23/1973	16754	02/10/1973	11074
1974	01/17/1974	40631	01/19/1974	35520	01/21/1974	28113	01/27/1974	18103	01/25/1974	14662
1975	03/25/1975	30037	03/27/1975	20826	05/20/1975	15749	05/28/1975	14191	06/09/1975	13863

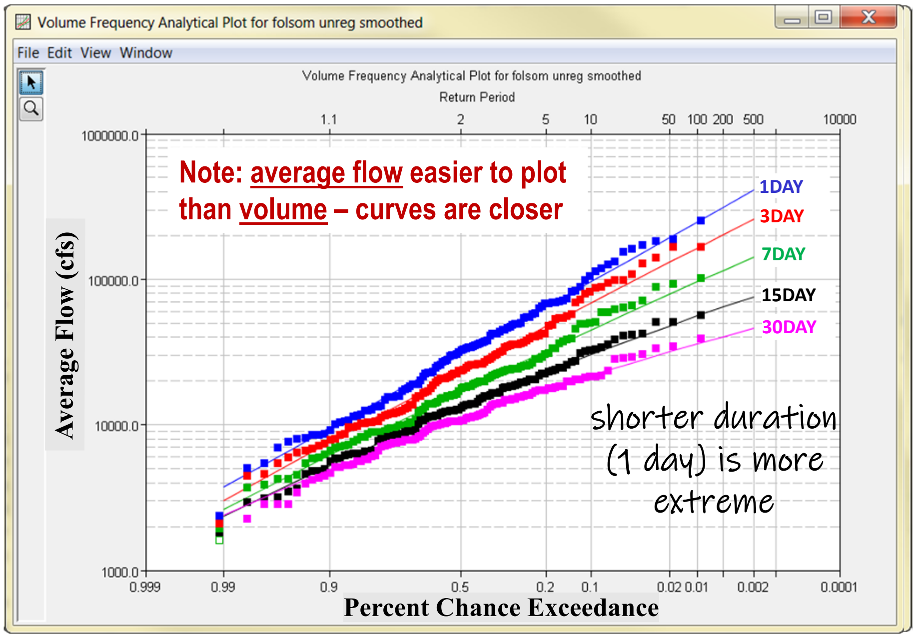
highest 3-day average

Extract Volume-Duration Data

39

SSP offers a summary of the annual maximum flows of every duration specified.

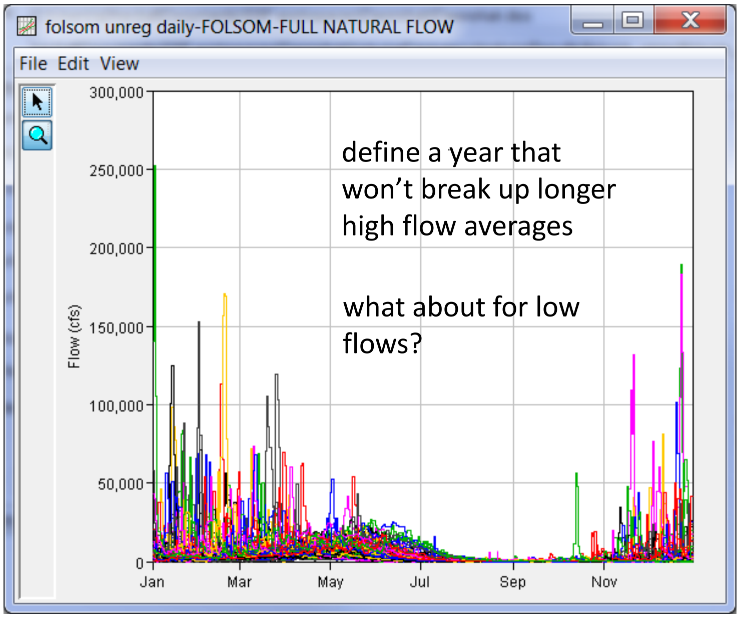
Max Flow at Folsom Volume-Freq Curves



40

This figure has both the sorted and plotted annual maximums, and fitted LP3 curves for each. Note, these are X-day average flows, because X-day volumes would be much higher for longer durations and not plot nicely. Showing average flows let's us show them all on one plot this way.

Note: it's important to define the year



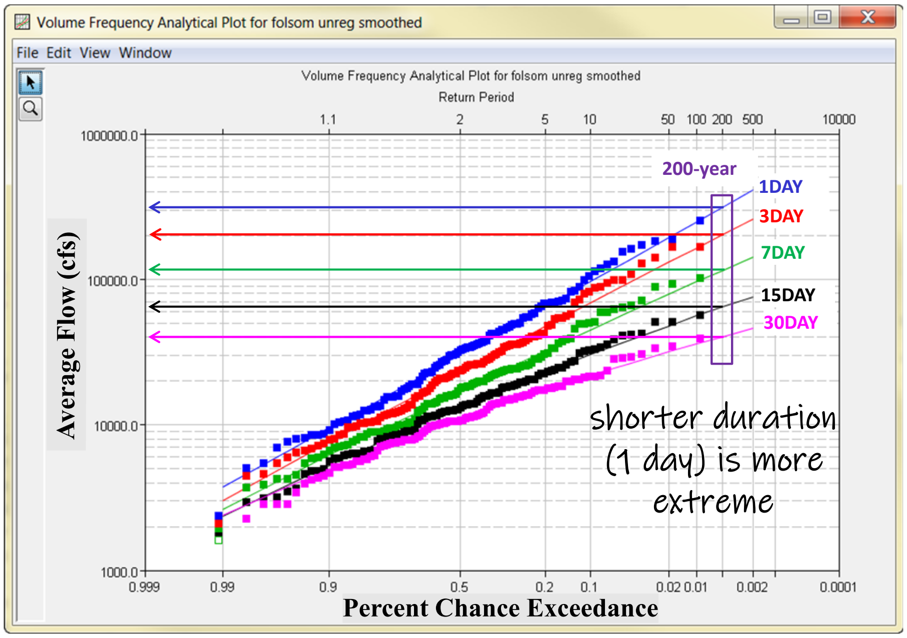
It is important to choose a start date that captures all flood events from a certain hydrologic regime. If high flows generally occur between November and May, then the year should not start between these months. This will minimize the possibility that the same flood event is used for consecutive years.

Synthetic Events

Create synthetic, low-frequency event hydrographs to model extremes, or other watershed conditions (e.g. regulated flow)

1. perform volume frequency analysis
   - produce the family of X-day average flow frequency curves
2. choose an historical event as a hydrograph "shape"
3. scale the hydrograph to match flow of chosen exceedance probabilities (e.g., the 1% chance event)
   - read the average flow for each duration for the chosen exceedance probabilities
   - scale the hydrograph up or down to match that avg flow

Max Flow at Folsom Volume-Freq Curves

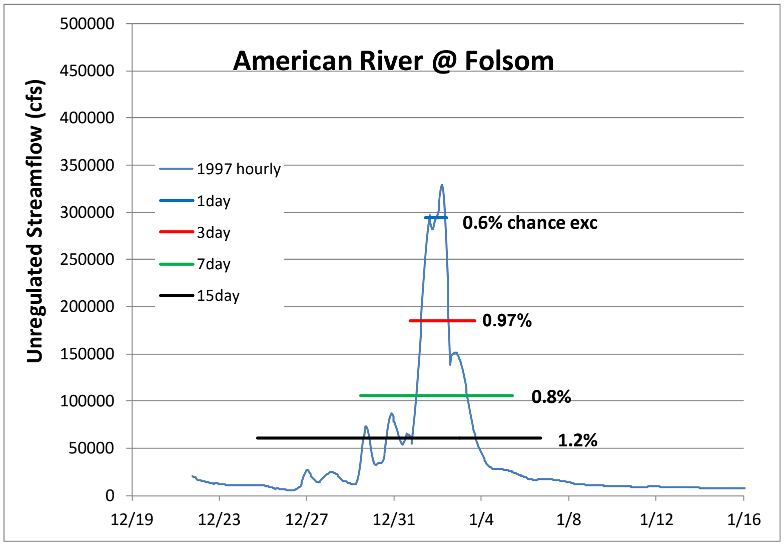


43

The average flow exceeded with a 0.005 (or, 0.5%) exceedance probability can be read for each curve. This value will allow us to create a synthetic event with that the desired estimated probability.

Rescaling Historical Event

1997 flood event

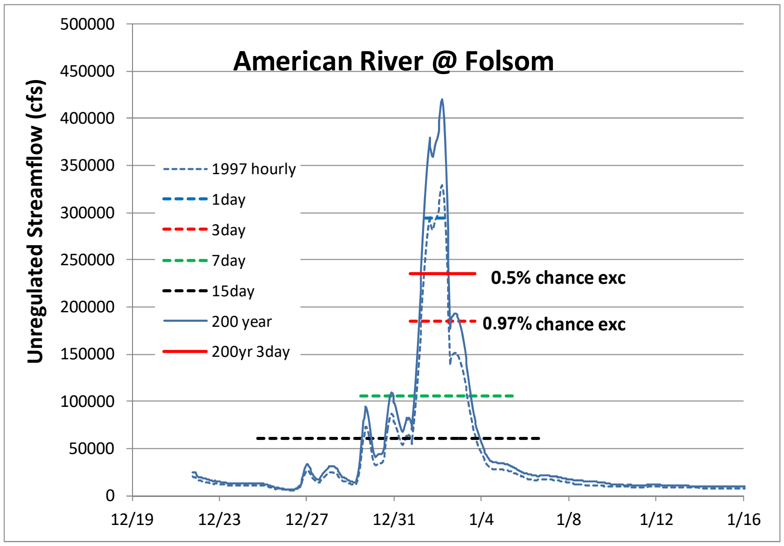


44

This is an historical event on the American River, from 1997. The exceedance probabilities of the maximum X-day average flows are noted.

Rescaling Historical Event

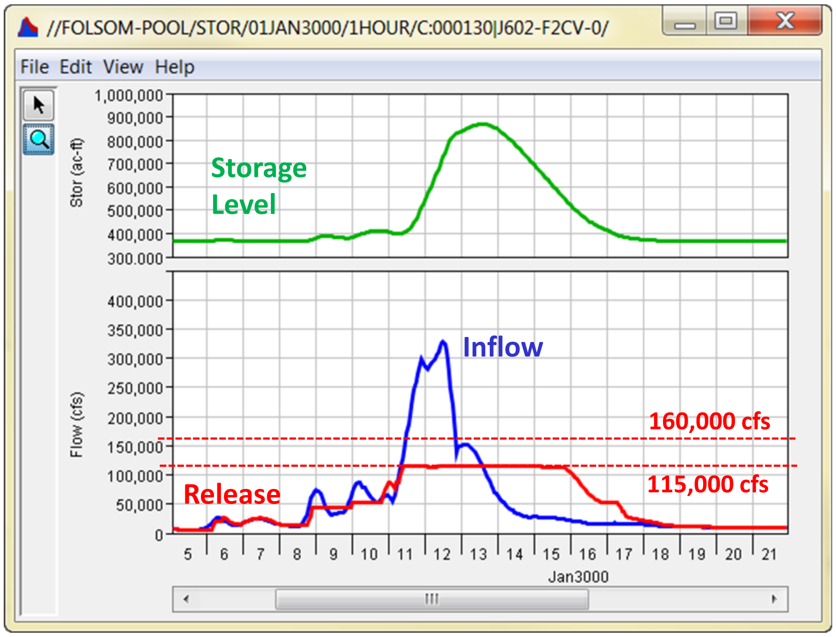
1997 flood event, scaled to 3-day 1-in-200%



45

This is the 1997 flood event scaled to the 0.5% (1 in 200 year) 3-day max average flow.

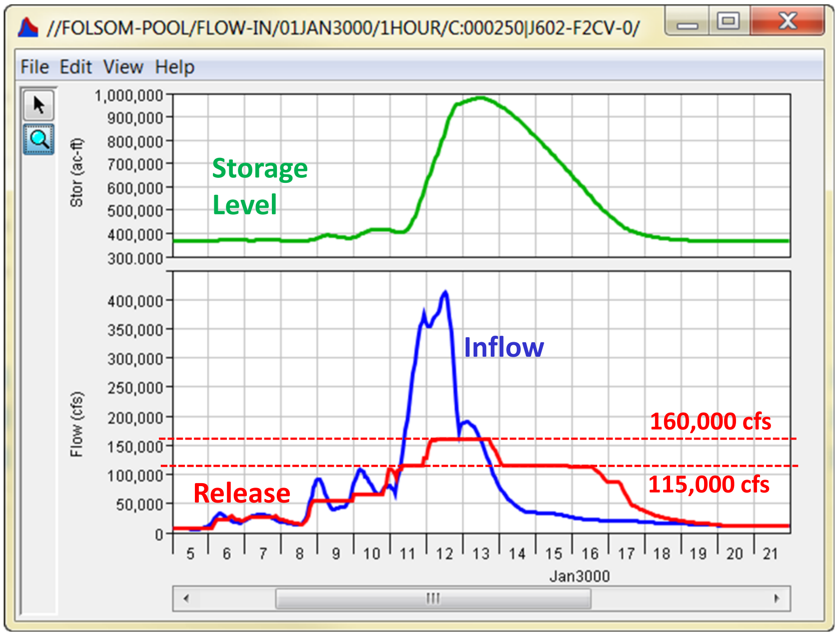
Modeled Operations for actual 1997



46

Simulating the 1997 event through the reservoir shows it not exceeding the lower objective release.

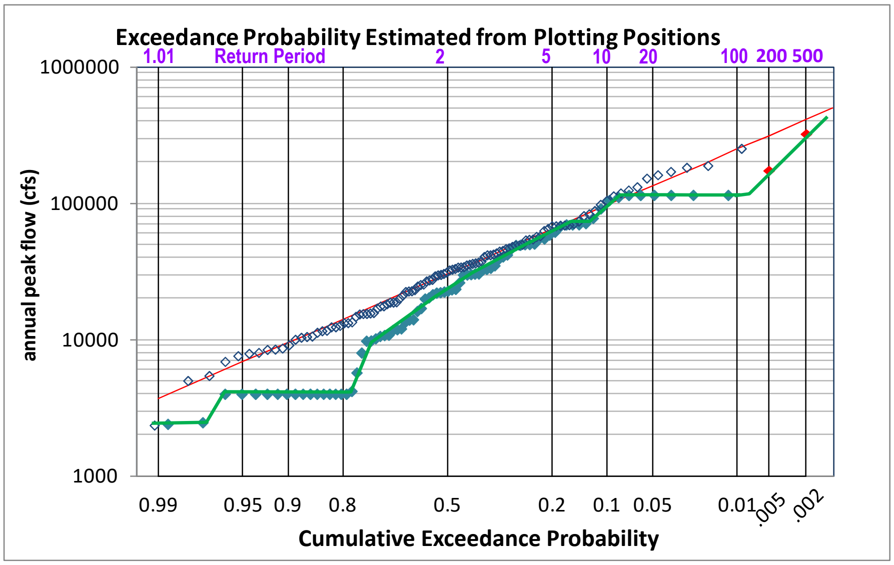
Modeled Operations for 200-yr 1997



47

Simulating the scaled up 0.5% version of the 1997 event through the reservoir shows it reaching the higher objective flow thresholds.

Graphical Curve is Better for this

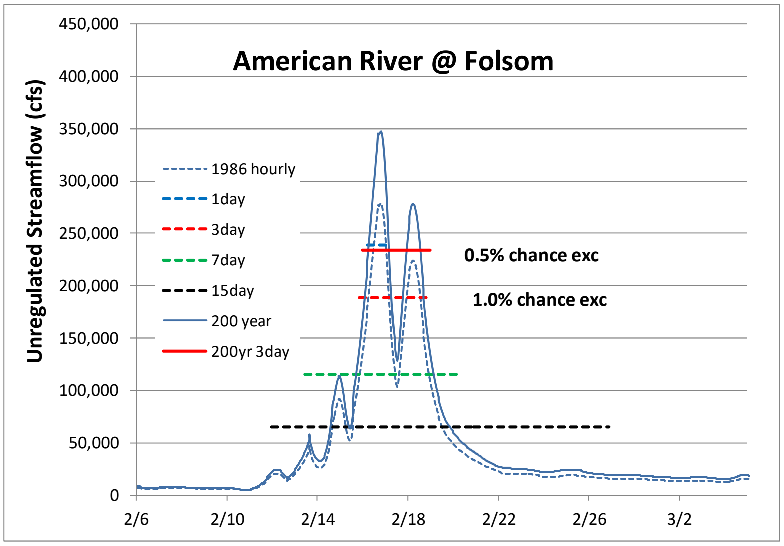


48

48

Rescaling Historical Event

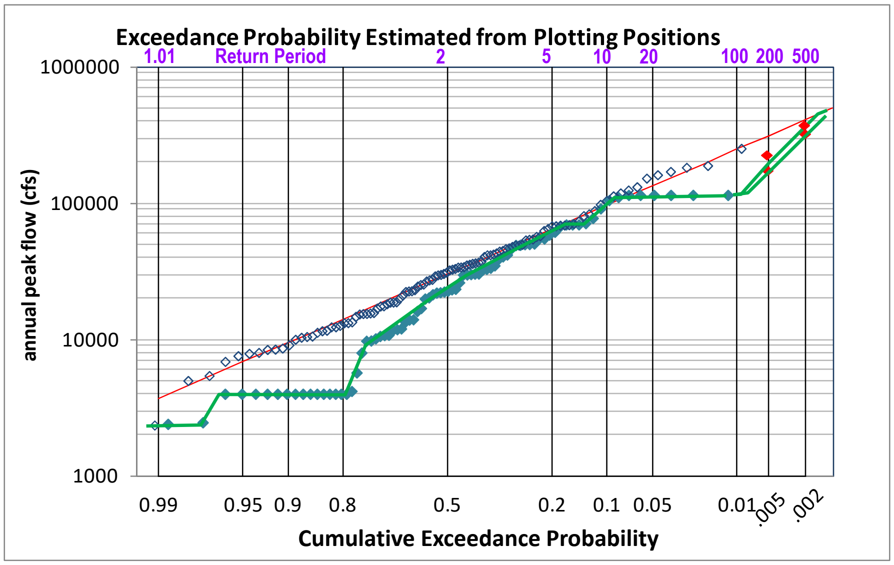
1986 flood event, scaled to 3-day 1-in-200%



49

We might also scale up additional historical flood events, if their shapes are different enough to demonstrate a relevant operational challenge.

Graphical Curve is Better for this



50

However, multiple estimates for the flow of a given exceedance probability must be resolved.

Perhaps we will assign a likelihood for each shape being the one experienced at a given magnitude. But, note that combining probabilities must be done horizontally, not vertically!

Volume Frequency Analysis

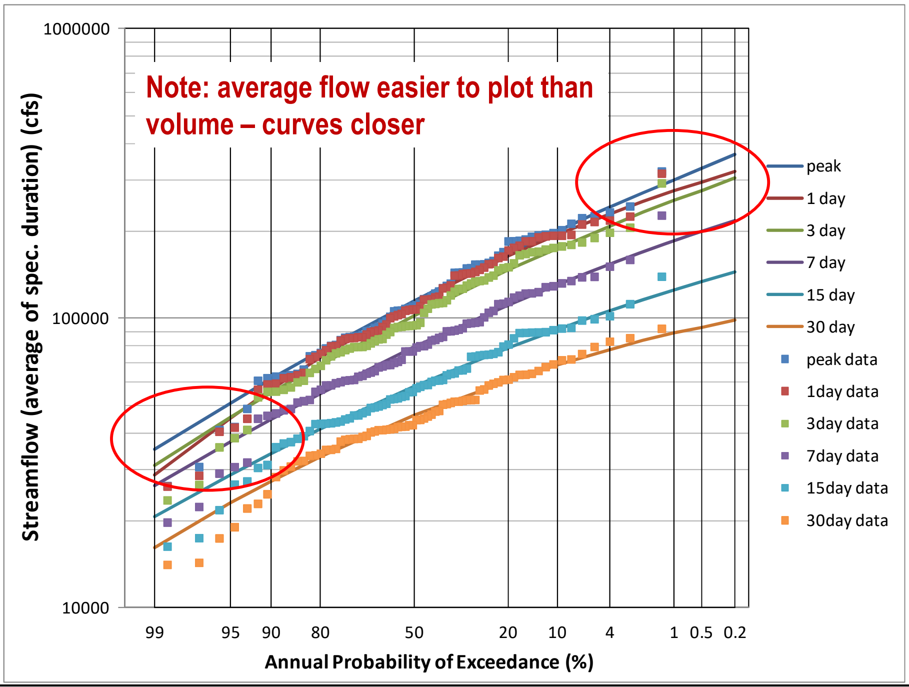
Sometimes there's an additional step needed in the volume frequency analysis…

3. Check curves for consistency
   - Curves should not cross
   - Smooth by plotting log moments vs duration, or versus each other

51

The curve have to make sense.

When volume-frequency curve require adjustment



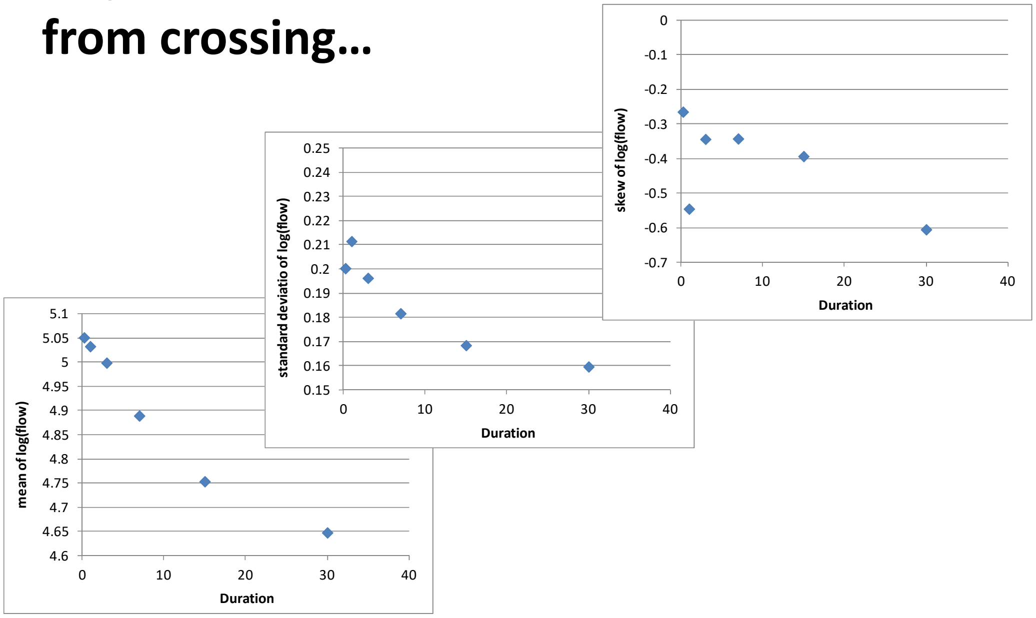
52

The sample for each duration is ranked and plotted separately, with an LP3 curve fitted to each. For high flow data, LP3 fit is generally successful.

The probability axis shows exceedance probability, and so extreme values plot to the right. Shorter durations are more extreme (in this case, higher) than longer durations.

When LP3 parameters are estimated for each separate data set (each duration column), the curves might not form a regular set, might cross. The parameters sometimes need to be adjusted…

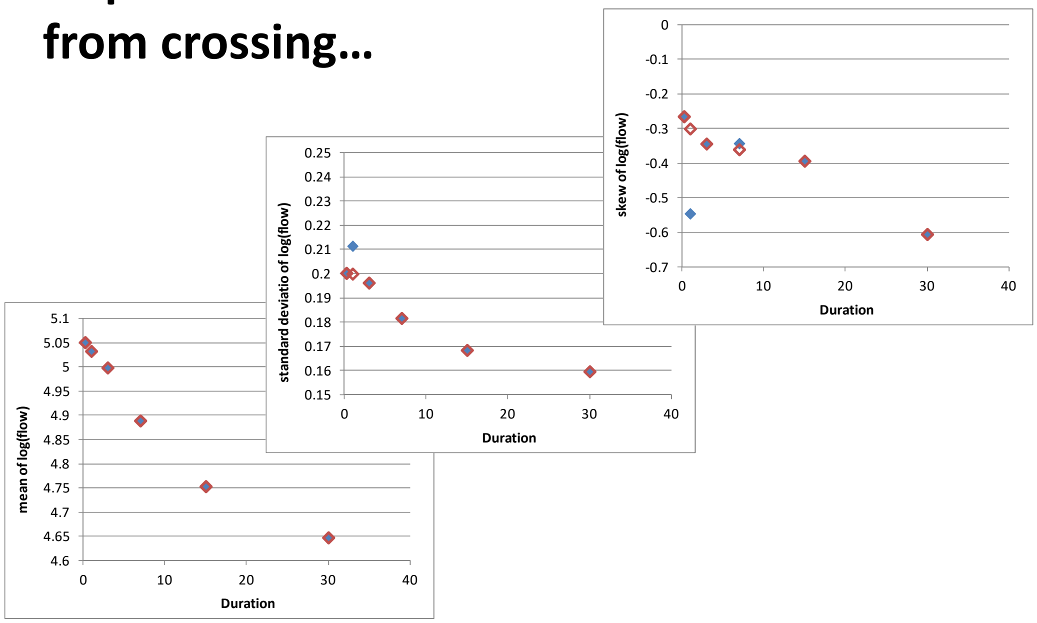
To prevent curves from crossing…



53

Here standard deviation is plotted against mean, and smoothed.

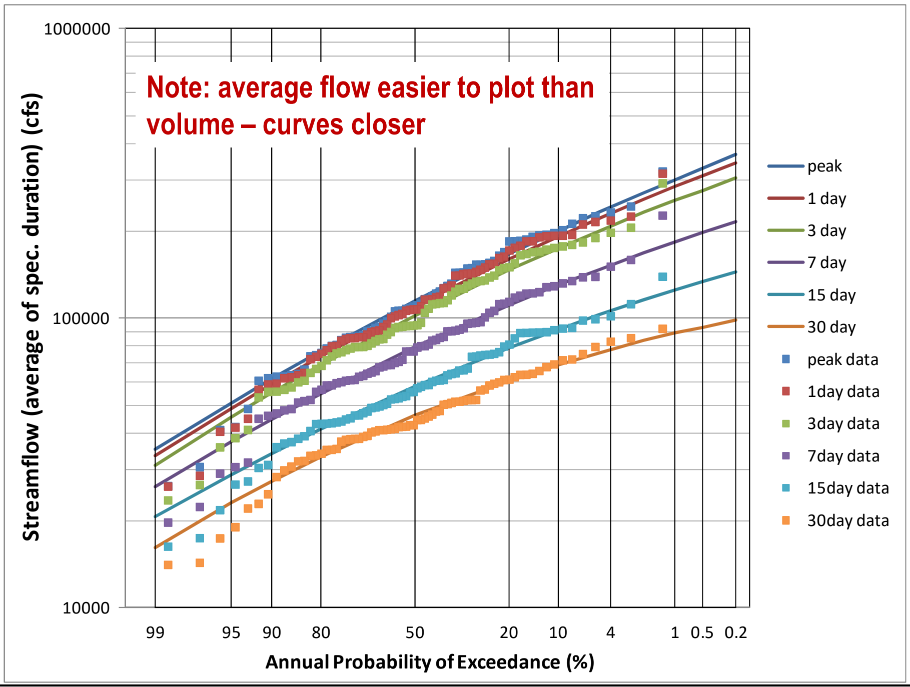
To prevent curves from crossing…



54

Here standard deviation is plotted against mean, and smoothed.

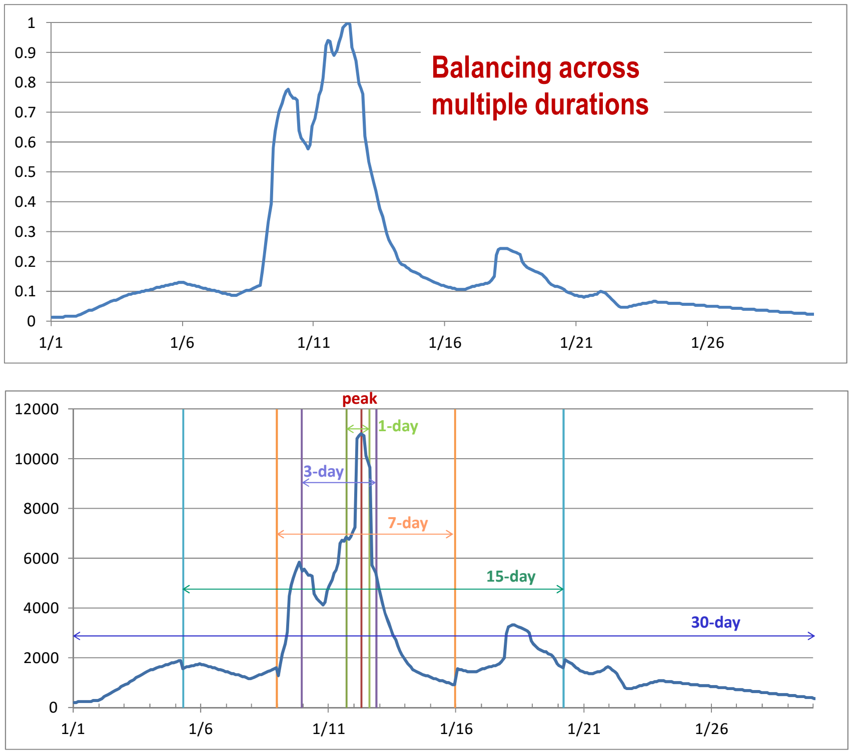
Resulting Volume Duration Frequency Curves



The sample for each duration is ranked and plotted separately, with an LP3 curve fitted to each. For high flow data, LP3 fit is generally successful.

The probability axis shows exceedance probability, and so extreme values plot to the right. Shorter durations are more extreme (in this case, higher) than longer durations.

When LP3 parameters are estimated for each separate data set (each duration column), the curves might not form a regular set, might cross. The parameters sometimes need to be adjusted…



56

Using the 1% chance exceedance values of each duration, create a "balanced" hydrograph that has those flow volumes. Could be scaled from an historical event. Note, this historical event might not be the best model, as the 1% values show a greater difference between peak and 1-day average max flow.

Low Flow Frequency Analysis

- Based on annual minimum flows
- Estimate the probability that X-day average flow will not exceed a given value

  example: 7-day 10-year low-flow

  ⇨ There is a 10% chance that the 7-day consecutive low-flow will be less than this value in any year

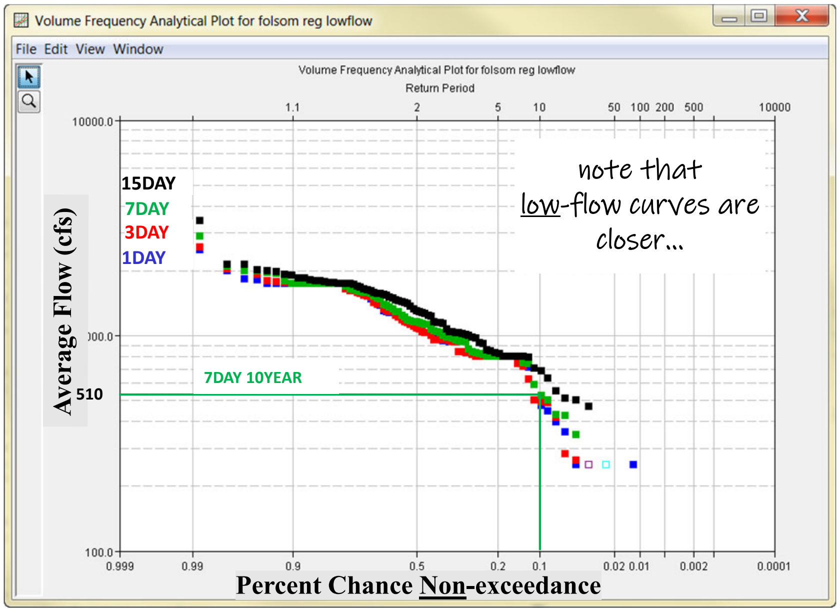
- As the duration increases to months, independence between events becomes less likely (7-day is common)

57

For low flows, we work with NON-exceedance probabilities.

A metric that's sometimes used in watershed restoration analysis and low flow regulation is ….

Regulated Low Flow Volume-Freq Curves

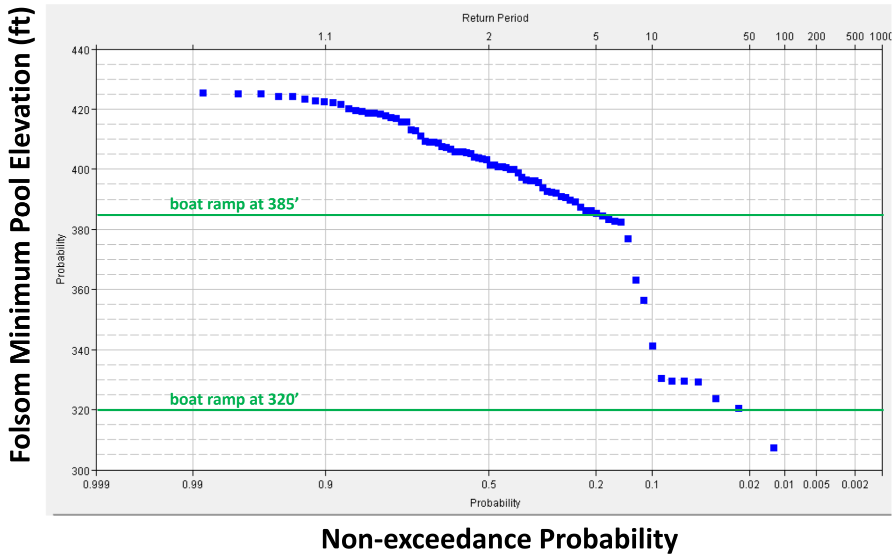


58

We’re interested in extremes. So, for annual minimums, we plot the NON-exceedance probability.

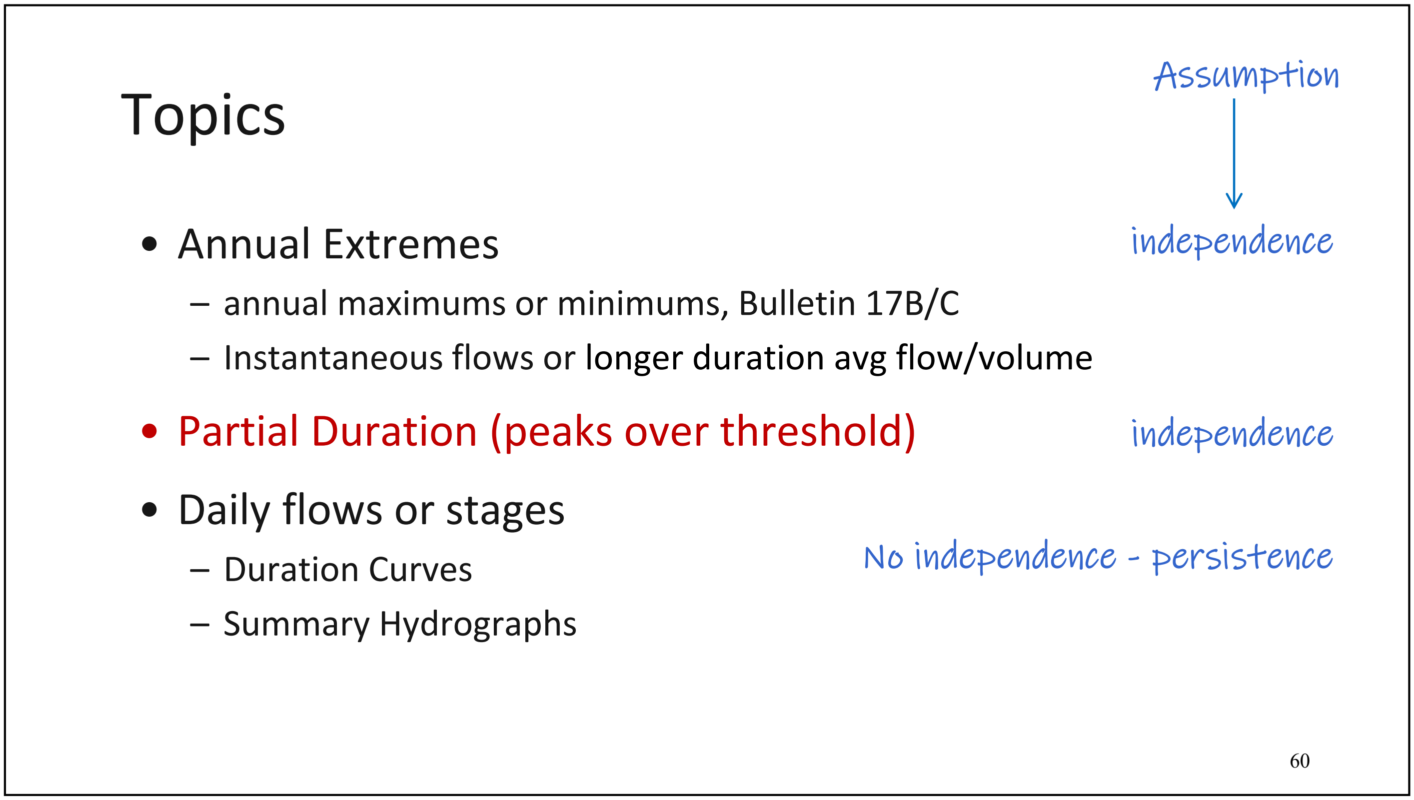
We usually keep the extremes on the right. So, zero non-exceedance probability is still on the right. The curves for different X-day average flows are closer, because low flows are not a peaky as high.

Folsom Minimum Pool Frequency



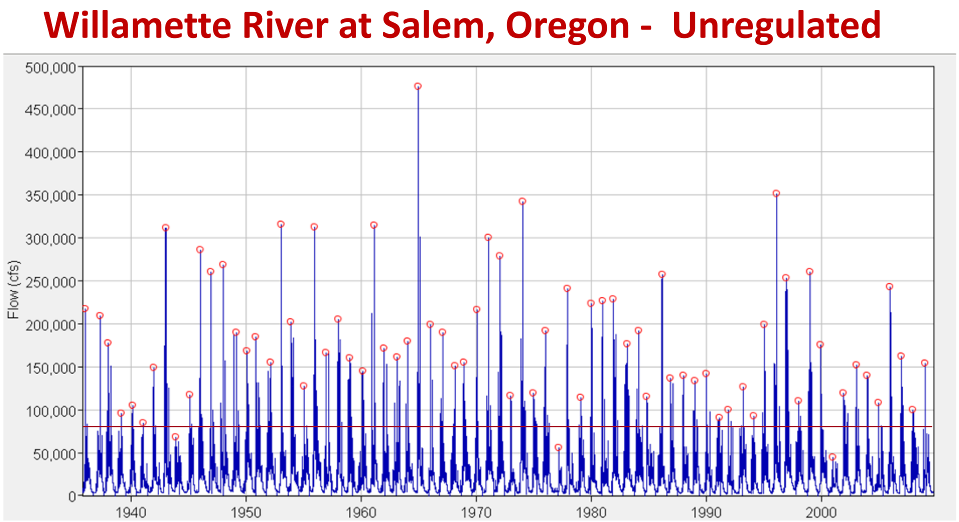
59

Also look at annual minimums (and maximums) for other variables. In this case, minimum reservoir pool.



These are the topics I'll cover. We divide the time period of the data, and whether we use all the data or isolate the extremes. These choices affect the assumptions we can make about the data set. Isolating annual extremes provides enough time between values that we can assume independence. Using all the data at a daily step, the values are persistent, and can't assume independence.

Annual Maximum Series (AMS)

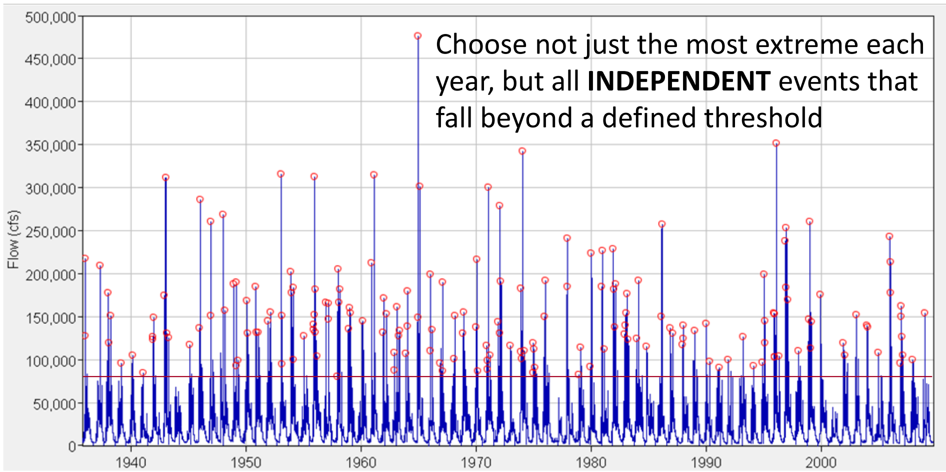


61

Annual maximum series is the largest value in each year.

Partial Duration Series (PDS) Peaks over Threshold

Willamette River at Salem, Oregon - Unregulated



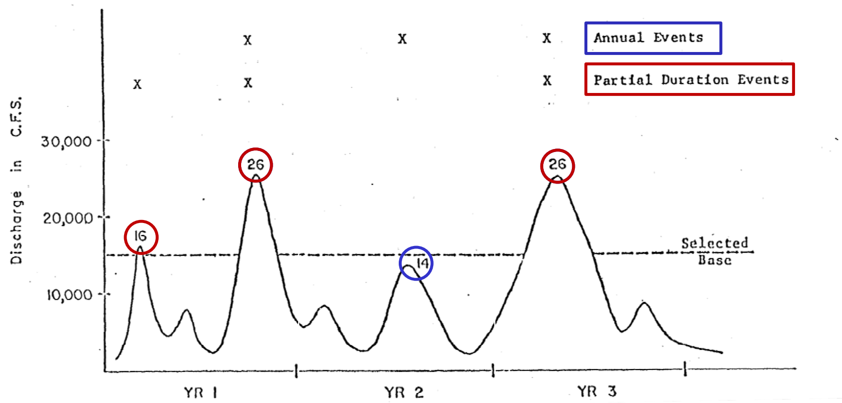
62

A partial duration series is all values above some defined threshold (also called peaks over threshold), no matter what year they occur in.

Partial Duration Series

All peak events above (or below) a particular threshold

Annual Series vs Partial Duration Series



63

Note, one annual event per years. But some years have no partial duration events, and some have several.

Empirical Partial Duration Series (PDS)

Calculating plotting positions:

If there are a total of **k** events in N years, then the plotting position for the smallest event (rank = k) is based on **k/N**

- If **k** > **N**, might be greater than 1.0   *can't interpret as probability...*
- The plotting position for this event is interpreted, not as exceedance probability, but as average of **k/N** events per year

Example Smallest event = 1000 k=50, N=25	Then the plotting position estimate states that on the average it is expected that 2.0 events per year will exceed 1000 cfs

64

It is possible to plot partial duration series empirically, with plotting positions. The exceedance probabilities are still interpreted as relative frequencies of equaling or exceeding a value, based on N years. The rank is interpreted the same way, as number of exceedances.

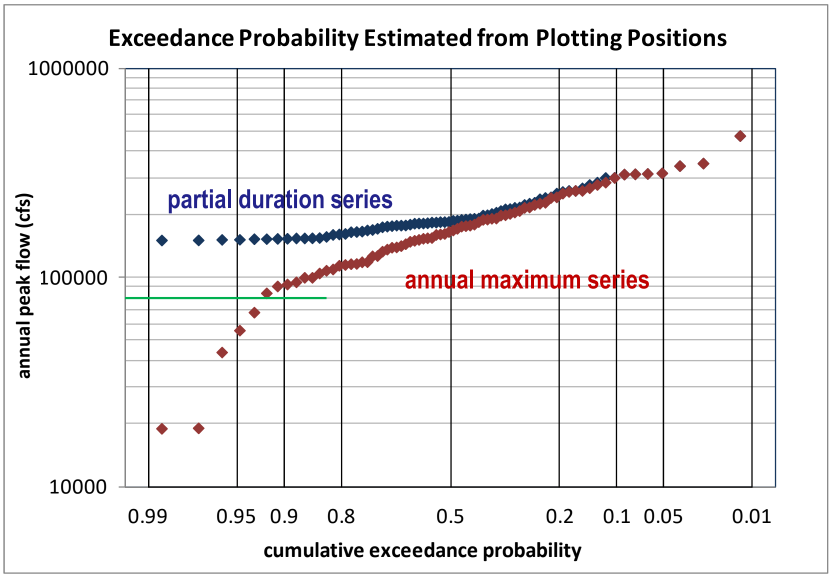
But, rank k might be higher than number-of-years N.

However, it is always possible to interpret as average number of exceedances per year, rather than likelihood of exceedance. This interpretation always makes sense.

Empirical Frequency Curves: Willamette at Salem partial duration vs annual max

can only plot the top N points on a probability axis

N=years



65

AMS and PDS curves are very similar at the top. But, they tend to differ quite a lot below the median, and often start to diverge below 10%.

We use PDS when we care about the lower end of the annual maximum frequency curve.

Analytical Analysis – PDS vs AMS Differences

- Analytical model for Partial Duration Series has two parts
  - Model for event **arrivals** (count per year)
  - Model for event **intensity** (magnitude over threshold)
- LP3 is not suitable for modeling flow magnitude
  - Ideal model is the Generalized Pareto distribution for intensity (and Poisson for arrival)
- Establishing independence of peaks is more challenging

66

Topics

Assumption ↓

- Annual Extremes — independence
  - annual maximums or minimums, Bulletin 17B/C
  - Instantaneous flows or longer duration avg flow/volume
- Partial Duration (peaks over threshold) — independence
- Daily flows or stages — No independence - persistence
  - Duration Curves
  - Summary Hydrographs

67

These are the topics I'll cover. We divide the time period of the data, and whether we use all the data or isolate the extremes. These choices affect the assumptions we can make about the data set. Isolating annual extremes provides enough time between values that we can assume independence. Using all the data at a daily step, the values are persistent, and can't assume independence.

Flow/Stage Duration Curve

Fraction of days exceeding a given value, referred to as percent of time exceeded

OR

not necessarily the annual extreme...

Estimate of the likelihood that a daily flow will exceed a particular level (refers to future, not soon...)

- Graphical display of stream characteristics
- Navigation, run-of-river hydropower, water supply

68

When using daily data, we compute percent of time exceeded.

Flow- or Stage-Duration Curves

Plot of Flow vs Percent of time exceeded

(see Maidment, 1992, pg. 8.27, 18.53, 27.40)

Computation

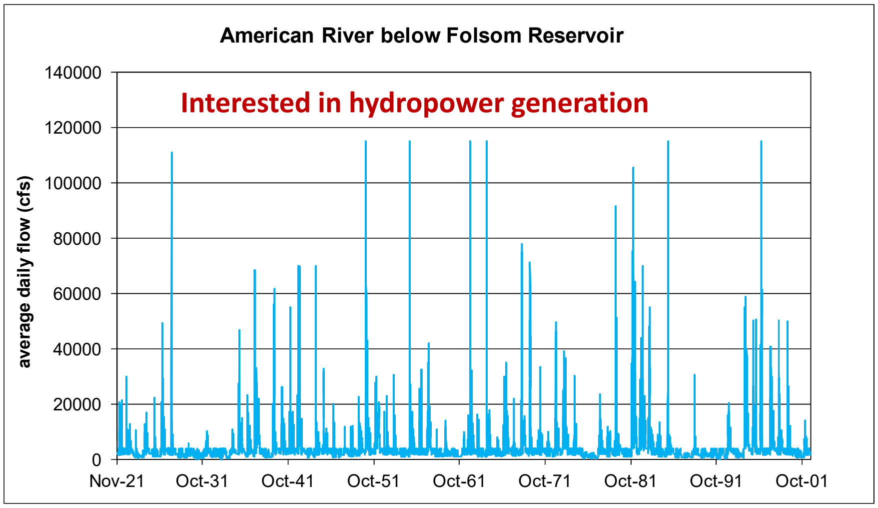
1. Rank all the daily flows for period of record from largest to smallest
2. Percent of time exceeded for each = $\dfrac{\text{rank}}{\text{total number of days in record}}$

69

Starting with a daily flow record…

Question: How much of the time can we expect future flows to be between 2,000 and 4,000 cfs?

How often were past flows between 2,000 and 4,000 cfs?



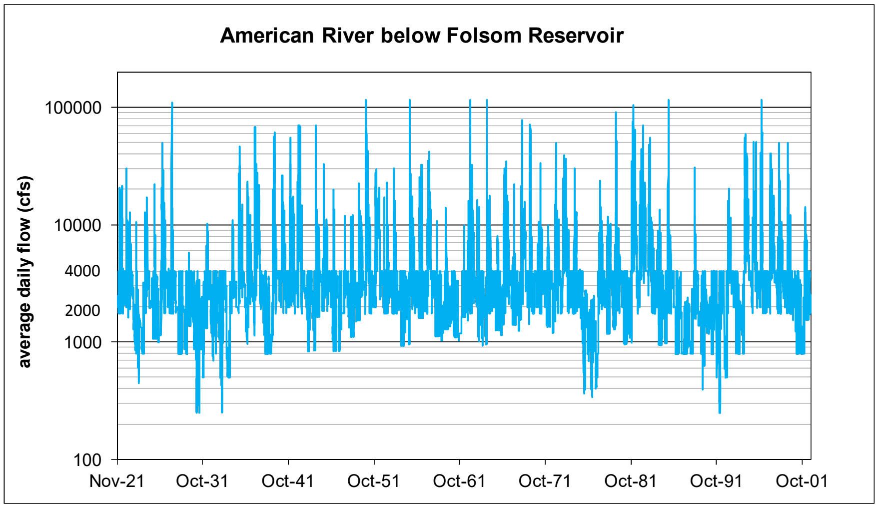
70

Here's the daily outflow record from the reservoir simulation we've been looking at.

Starting with a daily flow record…

Question: How much of the time can we expect future flows to be between 2,000 and 4,000 cfs?

How often were past flows between 2,000 and 4,000 cfs?



71

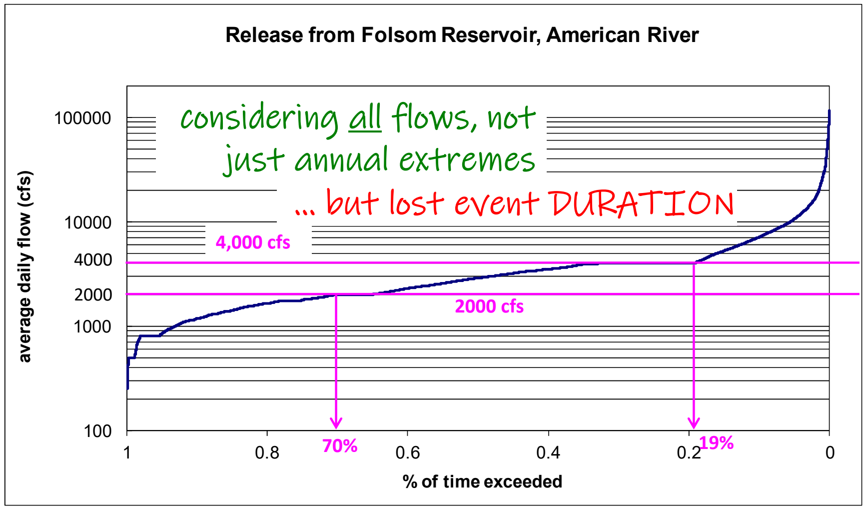
This is the daily outflow record on a log(flow) axis, so we can see the 2000 – 4000 cfs range more clearly.

Order the data from low to high, and compute relative frequency of exceedance

More than 4000 cfs is released 19% of the time, and

at least 2000 cfs is released 70% of the time,

so, 51% between 2000 and 4000 cfs



72

72

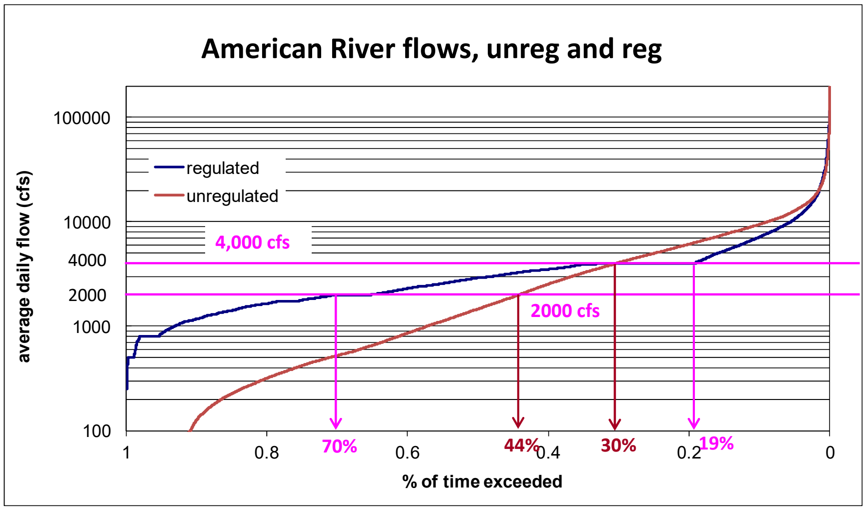
We can plot every value versus its percent of time exceeded.

Though this is called a "duration curve," we do not maintain information about the duration of time of each exceedance of a given flow. All days of exceedance are clumped together.

Comparing Unregulated to Regulated

Regulated flows are between 2000 and 4000 cfs 51% of the time,

unregulated flows are in the range 14% of the time.



73

73

The red curve uses the same process of generating a duration curve, but for unregulated flow. By comparing, it's clear that the reservoir allows release to be within the hydropower generation range much more often than a run-of-river hydropower plant would have.

Flow Duration Analysis Cautions

- Data points are not independent, so frequency does not approximate probability in the near term
  - Only approximates probability in future
  - An estimate of daily probability, not annual
- Analytical distributions do not fit data
- Need sufficient record length to estimate distribution tails

74

Topics

Assumption → independence

- Annual Extremes
  - annual maximums or minimums, Bulletin 17B/C
  - Instantaneous flows or longer duration avg flow/volume
- Partial Duration (peaks over threshold) — independence
- Daily flows or stages — No independence - persistence
  - Duration Curves
  - Summary Hydrographs

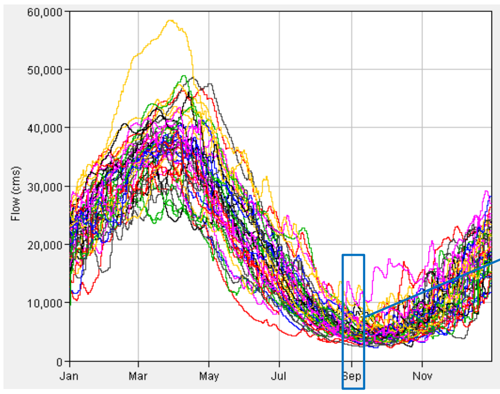
75

These are the topics I'll cover. We divide the time period of the data, and whether we use all the data or isolate the extremes. These choices affect the assumptions we can make about the data set. Isolating annual extremes provides enough time between values that we can assume independence. Using all the data at a daily step, the values are persistent, and can't assume independence.

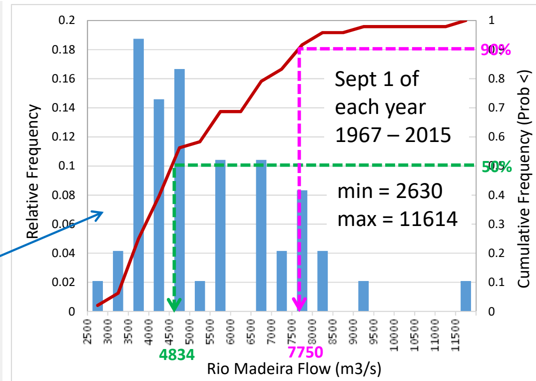
Summary "Hydrographs"

Rio Madeira @ Porto Velho, Brazil

1967 - 2005



Sept 1 only



76

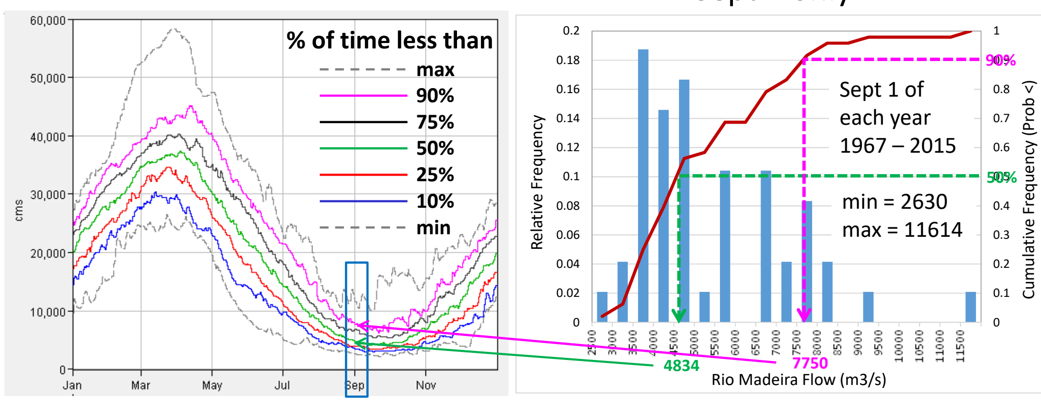
Summary hydrographs look at day of year, to give an overall summary of flow at a site across seasons. The method does a frequency analysis for each day of the years, and either plots all values, or plots summary values.

This image shows the frequency analysis for Sept 1 on the right, as a histogram and a cumulative histogram.

Summary "Hydrographs"

Rio Madeira @ Porto Velho

1967 - 2005 | Sept 1 only



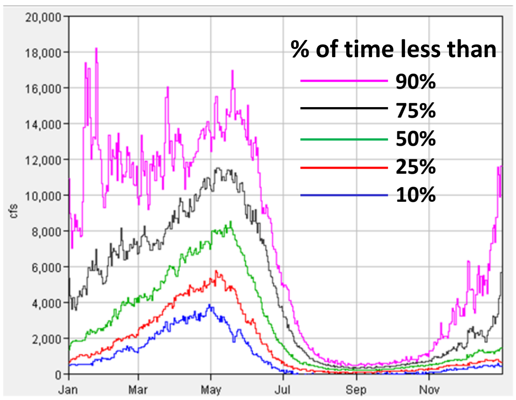
77

Given the frequency analysis for each day, the left plot can show percentiles, rather than all flow years. Note that the curves on the left are not really hydrographs. The daily values do not follow each other.

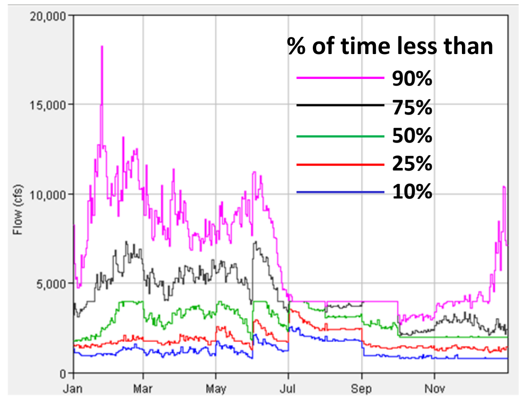
Summary "Hydrographs"

American River blw Folsom

Unregulated



Regulated

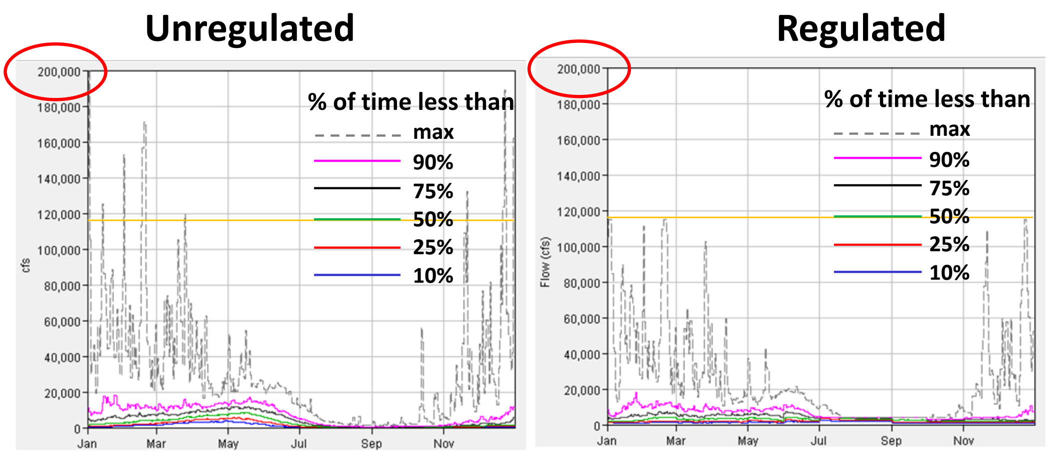


78

Summary "hydrographs" for American River, unregulated and regulated. Note wet winter and dry summer, and how regulated record shows lower high flows, and higher low flows.

Summary “Hydrographs”

American River blw Folsom



79

Summary “hydrographs” for American River, unregulated and regulated.

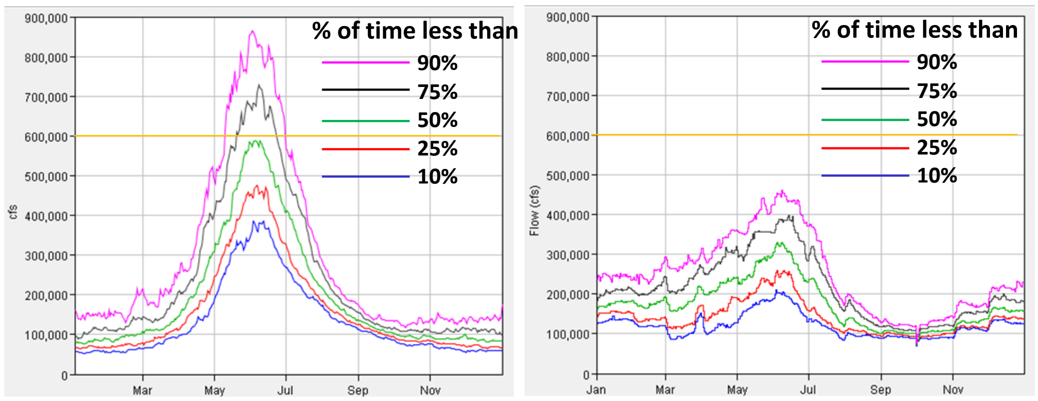
This version includes maximums for each day. Note, we see nearly all winter flood events, because they usually happened on different calendar dates. However, it makes it more clear that the max “hydrograph” doesn’t represent a flow time-series that can occur.

Seeing the maximum flows makes it clear that the reservoir is able to maintain each of the historical events to the channel capacity.

Summary “Hydrographs”

Columbia River at the Dalles

Unregulated **Regulated**



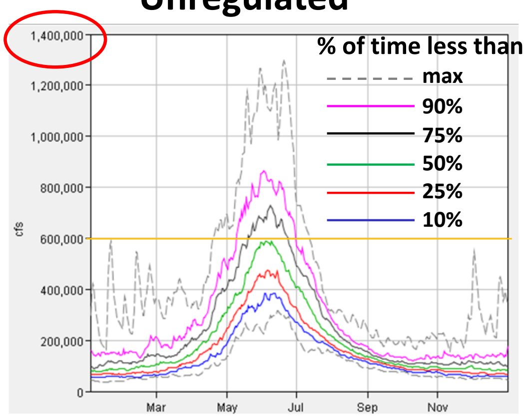
80

Summary “hydrographs” for Columbia River, unregulated and regulated. Note snowmelt in spring, and how regulated record shows a large decrease in the snowmelt flood flow.

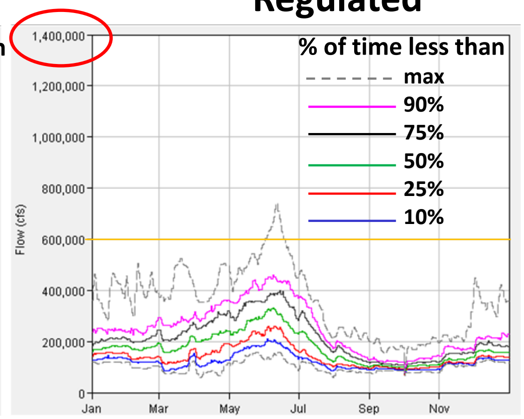
Summary "Hydrographs"

Columbia River at the Dalles

Unregulated



Regulated



81

Seeing the maximum flows makes it clear that the reservoir system is able to maintain nearly all of each of the historical events to the channel capacity.

Topics

Assumption

- Annual Extremes (Flow or Stage Frequency Curves)
  - annual maximums or minimums, Bulletin 17B/C — independence
  - Instantaneous flows or longer duration avg flow/volume
- Partial Duration (peaks over threshold) — independence
- Daily flows or stages — No independence - persistence
  - Duration Curves
  - Summary Hydrographs

82

References

- Corps of Engineers, 1993. Hydrologic Frequency Analysis, EM–1110-2-1415, Department of the Army, Washington D.C.
- Dracup, J.A., Lee K.S., and E.G. Paulson, Jr., 1980, On the Definition of Droughts, Water Resources Research, April, v16(2), p297-302.
- Hydrologic Engineering Center, 1985. Stochastic Analysis of Drought Phenomena, TD 25, p 140.
- Maidment, Davis R. (editor), 1992. Handbook of Hydrology, Mcgraw-Hill, Inc., New York.

83