A Tree-Ring Based Assessment of Synchronous Extreme Streamflow Episodes in the Upper Colorado & Salt-Verde-Tonto River Basins

Final Report

July 2005

A Collaborative Project between The University of Arizona's Laboratory of Tree-Ring Research & The Salt River Project

Katherine K. Hirschboeck & David M. Meko
Laboratory of Tree-Ring Research
The University of Arizona

With the assistance from: Kiyomi Morino, Ela Czyzowska & Jennifer Lee
Laboratory of Tree-Ring Research, The University of Arizona
&
The Salt River Project / Water Resource Operations

ACKNOWLEDGEMENTS
This project was been funded by The Salt River Project
in collaboration with the City of Phoenix and the U.S. Bureau of Reclamation

The Laboratory of Tree-Ring Research
THE UNIVERSITY OF ARIZONA.

SRP
Delivering more than power.

A Collaborative Project Involving
The Laboratory of Tree-Ring Research (LTRR) & The Salt River Project (SRP)

A TREE-RING BASED HYDROCLIMATIC ASSESSMENT OF SYNCHRONOUS EXTREME STREAMFLOW EPISODES IN THE UPPER COLORADO AND SALT-VERDE RIVER BASINS

Katherine K. Hirschboeck & David M. Meko
Laboratory of Tree-Ring Research, The University of Arizona

FINAL REPORT [1]

July 13, 2005

ABSTRACT

Tree-ring reconstructions of total annual (water year) streamflow for gages in the Upper Colorado River Basin and Salt-Verde River Basin were computed and analyzed for the period 1521-1964. These reconstructed flow series were used to identify years of extreme low flow (L) and high flow (H) discharge in each basin, based on 0.25 and 0.75 quantile thresholds, respectively. Synchronous extreme events in the same direction in both basins (LL and HH events) were much more frequent than LH or HL events, which turned out to be extremely rare occurrences. Extreme synchronous low flow (LL) and high flow (HH) events tended to cluster in time. The longest period of consecutive LL years in the record was 3 years. In terms of multi-year extremes, a scenario of 2 extreme years occurring anywhere within a 3-yr or 4-yr moving window was the most common. The overall conclusion based on the long-term record is that severe droughts and low flow conditions in one basin are unlikely to be offset by abundant streamflow in the other basin.

INTRODUCTION

Due to the recent severe and geographically extensive drought in western United States, many questions have emerged about the relationship between the water supplies of the Upper Colorado River Basin and the Salt-Verde River Basin. Are the two river systems relatively independent of each other due to different climate regimes? Is annual streamflow variability in the Salt-Verde Basin independent of annual streamflow variability in the Colorado Basin? Is drought in the Salt-Verde system unlikely to be accompanied by drought in the Colorado Basin, such that the latter's water supply can serve as a buffer for the former? How frequently do extreme events, such as the recent drought, occur in *both* basins at the same time?

The purpose of this project was to use long term records of tree-ring reconstructions of annual streamflow to analyze variations of low flow and high flow extremes in the water supply of the Upper Colorado and Salt-Verde River basins over the past several hundred years. The central question guiding the research was: **How frequently have extreme droughts or high flows occurred in both basins simultaneously in the past?** Specifically the project goals were: to determine how streamflow extremes in each basin have co-varied over time, to assess the hydrometeorological and hydroclimatological causes of this past co-variation, to provide probabilistic estimates of the likelihood of various scenarios of synchronous low-flow and high-flow extremes, and to devise an assessment tool for implementing the project's results into operational water supply decision-making.

---

[1] This report summarizes the results of work completed by the University of Arizona's Laboratory of Tree-Ring Research (LTRR) at the request of the Salt River Project (SRP). The project was funded by SRP in late 2003 and the bulk of the research took place in 2004. The project results also include a companion website containing the appendices, relevant data, and related links. It can be found at the following URL: http://fpcluster.ccit.arizona.edu/khirschboeck/srp.htm

The project generated a large volume of data, documentation, statistics, and analytical results presented in both tables and figures. The bulk of this information is included in the appendices which can be found on the companion Project Website.

An overview of the project methodology is shown at right. The work accomplished and the project findings will be presented in this report via a question/answer format that should be broadly applicable for use in a variety of settings such communicating with water managers, water resource scientists, water consumers, legislators and the general public.

OVERVIEW OF PROJECT METHODOLOGY:

(1) Use existing tree-ring data to refine and re-calibrate previous tree-ring reconstructions of streamflow; produce new reconstructions and compare with observed record

(2) Develop procedure to identify extreme streamflow episodes: drought (low flow, L) and "flood" (high flow, H) in each basin

(3) Define extreme streamflow scenarios: LL, HH, LH, HL in both the observed & reconstructed records

(4) Determine likely weather & climate causes of scenarios

(5) Develop climate-linked, probability-based assessment tool to transfer information into useful operational decision-making format

1 -- STREAMFLOW AND TREE-RING DATA USED IN THE PROJECT

The main conclusions of this project are based on long-term time series of annual streamflow values for the Upper Colorado and Salt-Verde-Tonto River Basins reconstructed from networks of annual tree ring widths. Available tree-ring chronologies, gaged streamflow data, and existing streamflow reconstructions were gathered, assessed and re-computed to produce updated and consistently derived streamflow reconstructions for key river locations in the basin.

(1-a) *What river basins and streamflow gages were reconstructed?*

Table 1 lists the 8 rivers / gages for which streamflow reconstructions were completed. **Appendix 1** provides additional information on the gages used, the data sources (e.g., for natural flow values), and the observed water-year average flows for these river systems. **Figure 1** shows streamflow gaging sites, tree-ring sites, and climatic stations that were used in the project.

Table 1 -- Rivers reconstructed and statistics for observed flows (cfs)

			Statistics[4]			
Code[1]	River/gage[2]	Coverage[3]	Mean	Std dev.	cv	skew
A.	Colorado at Lees Ferry	1906-1995	20813	5940	0.29	0.05
B.	Salt + Verde + Tonto	1914-2002	1673	1215	0.73	1.55
C.	Gila at head of Safford Valley	1915-2002	500	444	0.89	2.00
D.	Green at Green River, UT	1906-1995	7502	2224	0.30	0.38
E.	Colorado near Cisco, UT	1906-1995	9517	2703	0.28	0.21
F.	San Juan near Bluff, UT	1906-1995	3033	1202	0.40	0.32
G.	Salt + Tonto	1914-2002	1044	807	0.77	1.56
H.	Verde	1914-2002	629	431	0.68	1.50

[1] Letter code for gage (used throughout report)
[2] Name of river or gage (see **Appendix 1** for details)
[3] Start and end year of water-year totals used for statistics
[4] Mean, standard deviation, coefficient of variation and skewness of the water-year average flows in cubic feet per second (cfs)

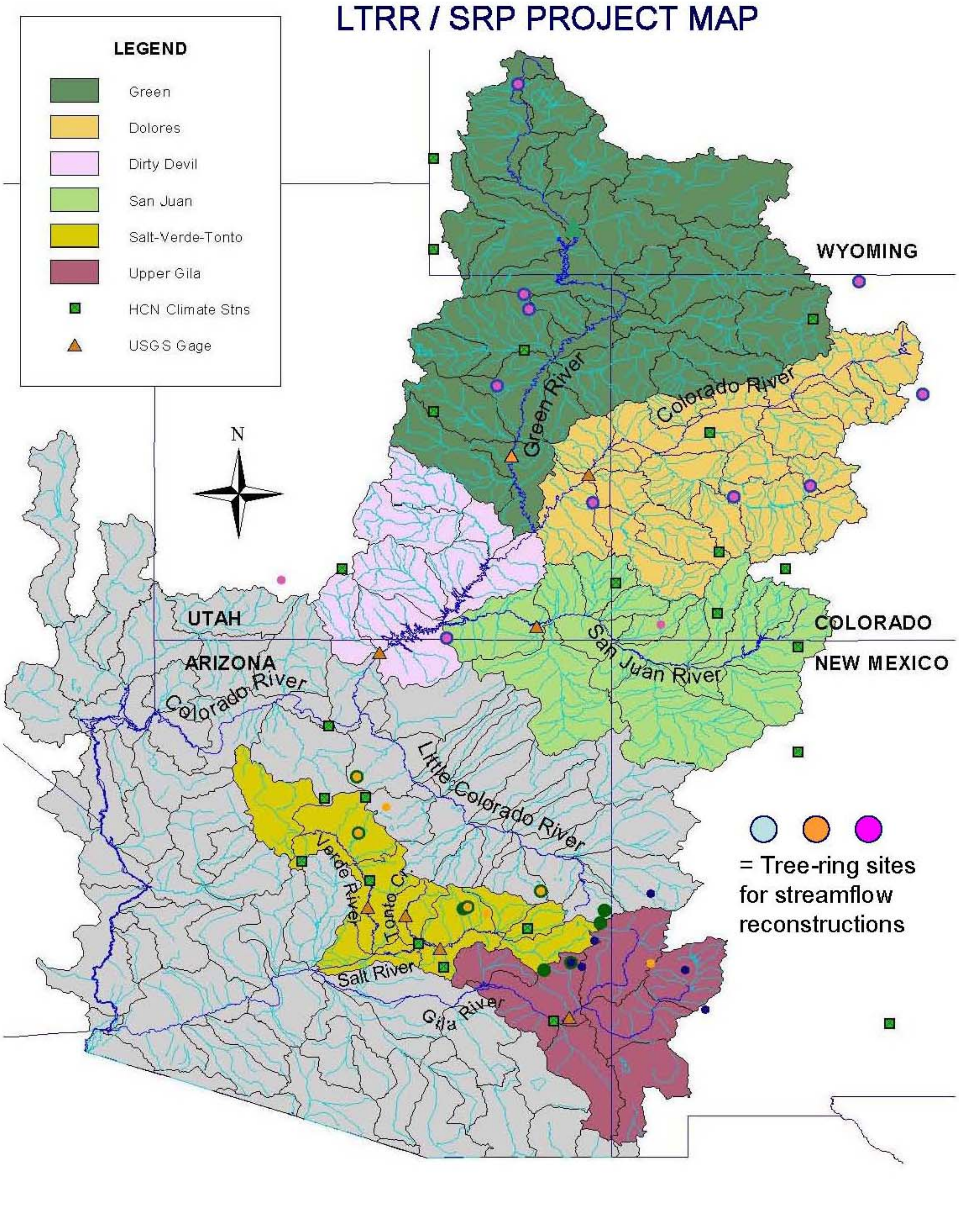


Figure 1 -- Project map showing watersheds, tree-ring, streamflow, and climate stations

(1-b) *What information obtained from tree rings is used to link tree growth with streamflow?*

The basic tree-ring data for this study consisted of "site chronologies," which are dimensionless time series representing the annual tree-ring width variations from many trees (e.g., 10-20) at a location (tree-ring site) (Fritts 1976). A site chronology (see **Figure 2**) can be regarded as the proportion of normal growth in each year such that values above 1.0 represent higher than normal growth and values below 1.0 represent lower than normal growth. The minimum value that can occur theoretically is zero, or no growth in a given year.

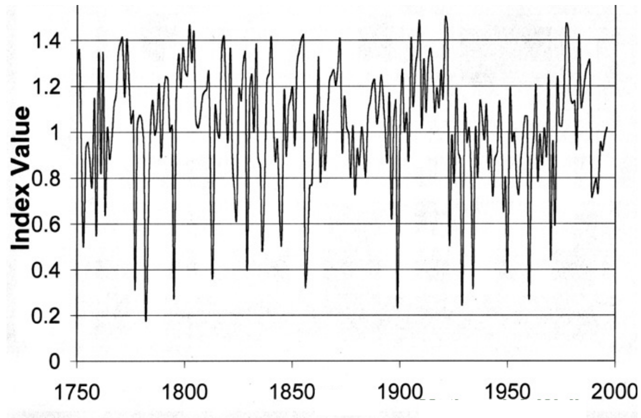


Figure 2 -- Example of a tree-ring site chronology with standardized index values representing ring widths

(1-c) *Why should there be a relationship between streamflow and tree growth in the arid and semi-arid western United States?*

Tree growth can be limited if there is not sufficient moisture in the soil. This can arise from either a lack of precipitation, or excessive evapotranspiration in warm, dry climates-- or both. Regionally, precipitation minus evaporation determines the amount of runoff available in rivers and streams. Hence both trees and streams can respond to weather and climate patterns that produce sustained drought condition. In the project study areas, the snow that falls in the winter *prior* to a tree's growing season is reflected in the tree's growth in spring and summer via soil moisture storage.

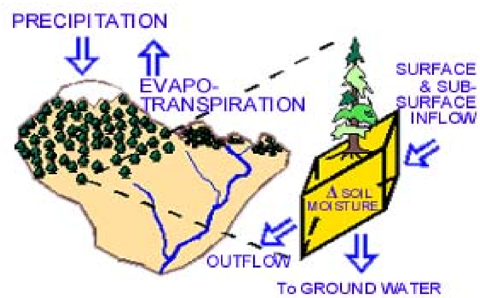


Figure 3 -- Why tree growth and streamflow variations can be correlated (see Meko et al. 1995)

(1-d) *What's the benefit of using tree-rings to study past streamflow variation?*

Gage records for the Salt-Verde River system and the Upper Colorado watershed are some of the longest in the United States, yet they represent only a short segment of these rivers' long-term natural variability. Droughts and floods can have devastating impacts on people, but because they are rare events, even 100 years of instrumental streamflow records may not be long enough to capture the range of possible extreme streamflow episodes. Climate-sensitive tree-ring records can extend the record of climate and hydrologic variability back many centuries to provide important information unavailable in the gaged record. In one of the first statistical streamflow reconstructions from tree rings, Stockton (1975), Stockton and Jacoby (1976) demonstrated how unusually wet the early 20th century was in the Colorado River in comparison to the long-term mean. Water supply planning decisions made on the basis of unusually wet or dry episodes of a river's history may not properly represent the true nature of the supply.

(1-e) *How were the tree-ring sites selected for the analysis?*

The study used only pre-existing tree-ring chronologies. No field work or new chronology development was included due to the project's limited timeframe. In selecting the tree-ring data, sites inside the basins or within a buffer of about 200 km of the basin boundaries were considered a-priori as potentially useful for the analysis. An additional constraint was imposed that the tree-ring data at the site cover at least the period from the mid-1660s through 1961. The end year was dictated by the collection dates of available chronologies and our preference to have a deeply replicated site network. The start year was selected such that even the shortest chronologies would sample the well-known "Pueblo" drought that occurred in the late 1660s (Smith and Stockton 1981).

Figure 4 shows the locations of sites in the tree-ring networks used in the reconstructions for the Colorado at Lees Ferry (Fig 4a) and the Salt-Verde-Tonto (Fig 4b). The tree-ring network goes back to 1279 in the Upper Colorado Basin and 1199 in the Salt-Verde-Tonto Basin. However, the best reconstruction model for the Colorado at Lees Ferry emerged from a subset of well-replicated tree-ring sites whose records started in 1521 (sites indicated by □ symbol in Figure 4a.). Note that there is minimal overlap of sites in this subset of the Upper Colorado network with that of the Salt-Verde-Tonto network, thereby avoiding the problem of built-in correlations between the two basins due to the use of

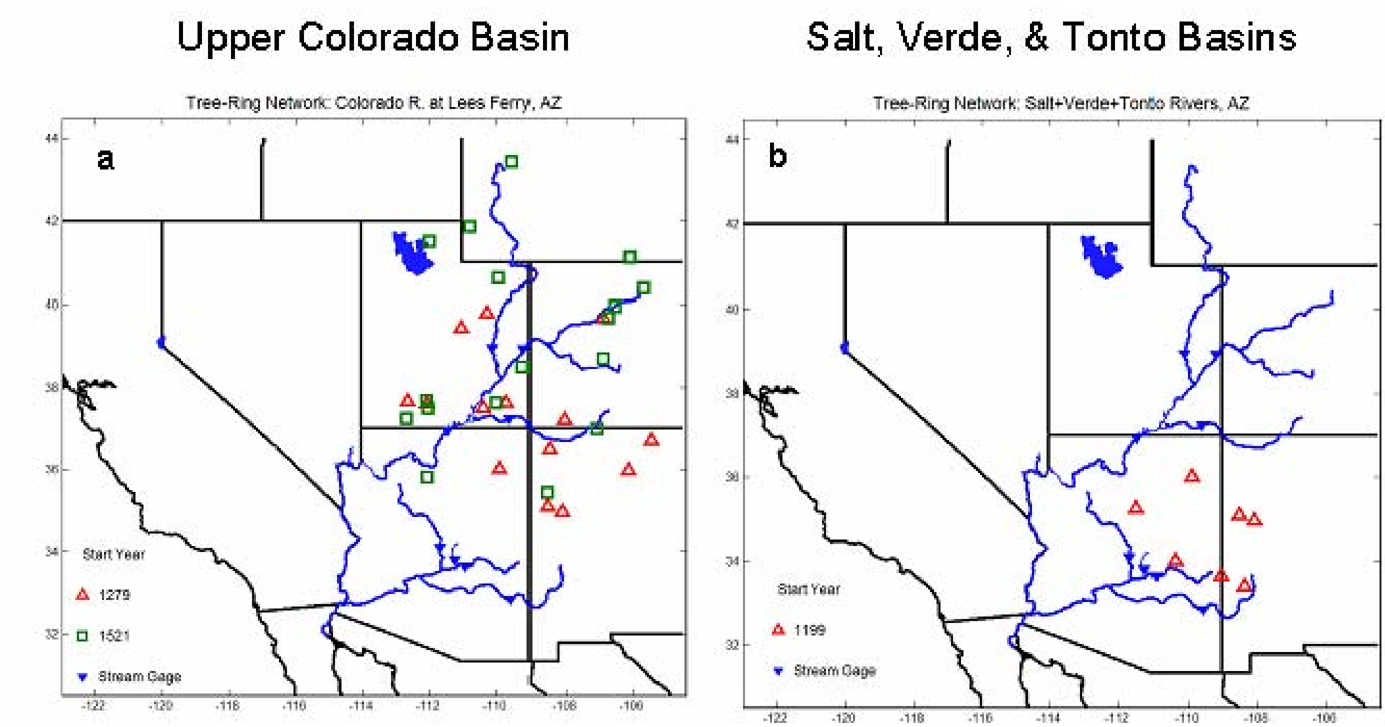


Figure 4 Locations of tree-ring sites used in the streamflow reconstructions.

duplicate chronologies. The post-1521 sites in the Upper Colorado were last collected in 1964. Because of this, the Colorado streamflow reconstruction ends in 1964 and the period 1521-1964 was designated as the common period for comparison between the two basins in our subsequent analysis of synchronous extreme flow episodes in the reconstructed time series.

(1-f) *<u>What were the different sources of the tree-ring data?</u>*

Tree-ring data were obtained from four sources: (1) The International Tree-ring Databank (ITRDB) which is the major repository of tree-ring data in North America; (2) a North American tree-ring network used in various studies of spatial patterns of drought (Meko et al. 1993; Cook et al. 1999; Cook et al. 2004); (3) a network of tree-ring data used recently for 1000-year reconstruction of precipitation and drought index for Arizona (Ni et al. 2002); and (4) a network of chronologies from living trees augmented by archaeological samples and previously used in a 1500-year reconstruction of annual streamflow of various rivers in Arizona (Graybill et al., in press).

There is considerable redundancy among these four networks from which data were drawn. Where chronologies were available from multiple sources, priority was given first to the Ni et al. (2002) dataset, as those chronologies had been specially processed to maximize retention of multi-decadal climate information and reduce possible effects of diminished sample size (number of trees) in the earlier years of the chronology.

“Standard” chronologies, defined as dimensionless tree-ring indices not adjusted for removal of autocorrelation (Cook and Kairiukstis 1990) were preferred for the study, but so called “residual” chronologies were also acceptable, as the reconstruction method is fairly insensitive to pretreatment for autocorrelation. The complete network for this study includes 119 chronologies: 82 from the ITRDB, 11 from Ni et al (2002), 23 from Meko et al. (1993), and three from Graybill et al. (in press). A complete listing of chronologies and sources is included in **Appendix 2** . Site locations are mapped in **Figures 1 and 4**. The network includes nine different species. Time coverage varies by chronology. The most recent chronology includes growth year 2000, while the longest extends to B.C. However, no tree-ring data before A.D. 1199 were used in the reconstructions to avoid basing conclusions on poorly replicated tree-ring data.

2 -- RECONSTRUCTING STREAMFLOW

The streamflow reconstruction procedure used in this study had four stages (I- IV, see **Figure 5**). The procedure is a modification of a reconstruction method previously developed for reconstruction of precipitation from time-varying subsets of tree-ring indices (Meko 1997).

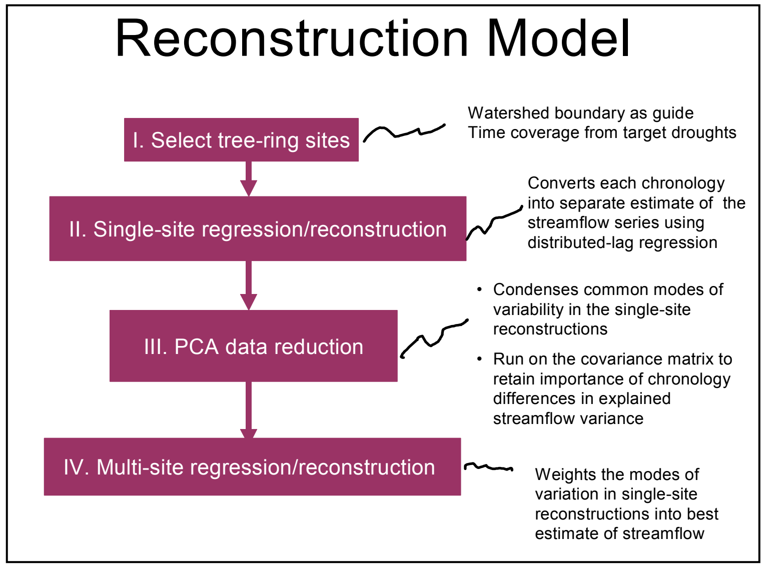


Figure 5 – Stages in the streamflow reconstruction method

(2-a) *What guided the design of the streamflow reconstruction model?*

One objective was to obtain reconstructions that take advantage of the extensive replication provided by the dense network of available tree-ring chronologies. Another goal was to make the model robust, i.e., free of uncertainties stemming from possible unknown factors in the past, or non-climatic disturbances at individual sites.

(2-b) *What are the steps of a streamflow reconstruction procedure?*

Stage I -- Building the predictor pool The first stage identified the initial pool of predictor tree-ring chronologies, i.e. the set of chronologies potentially useful for reconstruction for a specific gage. This initial pool essentially comprised those chronologies from the full 119-site network (see above) that were within the watershed above the gage or roughly a 100-mile buffer around the watershed. Depending on the river, several subsets of predictor pools with different time coverage were assembled.

Stage II -- Screening and filtering The second stage eliminated from the pool those chronologies statistically unrelated in a bivariate sense to streamflow, and filtered the remaining chronologies into individual estimates of the streamflow record. The screening and filtering was accomplished by multiple linear regression of the flow at time t on a tree-ring chronology at lags -3 to +1 years relative to the year of streamflow. We refer to these regression models as the "single site regresson models".

The predictand for a single-site regression model is the observed annual flow, either in original units or log-transformed (depending on an analysis of regression residuals). The pool of potential predictors is the chronology and its lagged values. A forward-stepwise regression on the full overlap of tree-ring data and streamflow was first run to identify the order of entry of predictors.[2] After estimating the full-calibration-period model and recording the order of entry of predictors, a split-sample calibration-validation procedure was done to check that the identified model was stable over time. The first and second halves of the full calibration period were designated A and B, respectively. The model was calibrated stepwise

[2] The criterion for entry was that the F-value for entry of a predictor has a maximum p-value of 0.05. At each step the F-value for removal was also checked, and the maximum p-value a predictor can have without being removed from the equations was set at 0.10.

on A and validated on B, with statistics of calibration and validation accuracy recorded at each step. The process was then reversed – calibrating on B and validation on A. The final step for the stepwise procedure was identified as the step beyond which an additional predictor led to a decrease in modeling accuracy as measured by the root-mean-square error of validation in either of the split-sample regressions. The stepwise regression model was then re-calibrated on the full calibration period with stepwise entry truncated at the appropriate step indicated by the split-sample modeling

Substitution of the long-term tree-ring indices into the suite of resulting regression equations yielded a set of "single-site" reconstructions of streamflow. The single-site regression served as a screening tool to eliminate from further consideration any chronologies without a statistically significant correlation with the gaged streamflow record. The weakest single-site reconstructions (no significant relationship between the lagged tree-ring index and flow) were eliminated from further consideration.

Stage III -- Data reduction The third stage of the reconstruction procedure used principle components analysis (PCA) to reduce the single-site reconstructions, which contain much redundancy, into a new set of orthogonal variables describing independent dominant modes of streamflow-related tree-ring variation[3]. This principal components transformation was done for different sub-periods of tree-ring data coverage corresponding to target droughts. The most inclusive subset of chronologies covered the period back to the mid-1660s, and was aimed at allowing assessment of the well-documented "Pueblo drought" that impacted Arizona around1670 (Smith and Stockton 1981). An intermediate subset of chronologies targeted the period to A.D. 1570, based on the severe late-1500s drought identified by several researchers (e.g., Schulman 1956; Stockton and Jacoby 1976; Stahle et al. 2000). The longest subset extended to the late A.D. 1100s in an attempt to capture the "Great Drought" believed to have impacted the Anazazi civilization (Schulman 1956). Only some of the basins used in this study have tree-ring coverage for this earliest period.

Stage IV -- Multi-site Regression / Reconstruction The fourth stage of the reconstruction procedure was multiple linear regression (MLR) of streamflow on the scores of the principal components of the single-site reconstructions. The estimated regression coefficients weighted the principal-component scores into final estimates of flow. The predictand for the MLR was the observed annual flow series and the predictors were the principal component (PC) scores of the single-site reconstructions. Initial runs of the model were later adjusted to produce the final streamflow reconstructions.[4-5] Complete details of the reconstruction modeling for each gage (A through H, see Table 1) are provided in **APPENDIX 3**.

---

[3] The principal components analysis for this data-reduction step was run on the covariance matrix, rather than the correlation matrix, of the single-site reconstructions because it was desirable to retain the information inherent to the differences in variances of those reconstructions: the variance of a single site reconstruction is proportional to the variance of flow explained by the chronology and its lagged values.

[4] In our initial runs, only the most important PCs, as measured by the eigenvalues from the PCA, were included as potential predictors in a stepwise forward model. To be included in the pool, the eigenvalue was required to be larger than the mean eigenvalue of all PCs. This criterion is analogous to an "eigenvalue of 1" selection criterion when the PCA is run on the correlation matrix (Mardia et al. 1980). Experience showed that this rule was generally too conservative, and was later relaxed to allow any PC explaining more than 5% of the variance of the single-site reconstructions to be included. (To avoid confusion, however, we kept the "mean eigenvalue" threshold for the Salt+Verde+Tonto model because the reconstructions by the two methods differed only slightly and much further analysis had already been done on the earlier version). Regardless of which rule was used to initially truncate the PCs, the number of predictors in the pool was not allowed to exceed 25 percent of the number of observations for calibrating the model. The 25% rule is a conservative application of a guideline suggested by Haan (2002).

[5] The predictors were entered stepwise (same *p*-enter and *p*-remove thresholds as single-site models) and the model was cross-validated at each step by a leave-*m*-out procedure, where *m* was set to one plus the maximum number of positive and negative lags used in any of the single-site regression models. This modification of leave-1-out cross-validation of Michaelsen 1987 was implemented to retain the independence of the calibration and validation subsets (see Meko 1997). The cross-validation was used in conjuction with a stopping rule that entry of predictors be terminated whenever cross-validation error as measured by the root-mean-square error of validation started to increase.

3 – THE COLORADO AND SALT-VERDE-TONTO STREAMFLOW RECONSTRUCTIONS

Figure 6 compares the water year streamflow reconstructions of the Upper Colorado at Lees Ferry and the Salt-Verde-Tonto Basins with corresponding annual observed (gaged) values in the 20th century.

(3-a) *How good are the tree-ring based reconstructions of flow when compared to the gaged record?*

The reconstructions track the timing of high and low flow years quite well and do a good job of capturing the magnitudes of the flow, especially in the dry years. The reconstruction models explain 77.4 % of the variance for the Colorado at Lees Ferry (r = 0.88) and 57.8 % of the variance for the Salt-Verde-Tonto (r = 0.76). The magnitudes of extreme wet years are reconstructed more accurately in the large Colorado basin than in the Salt-Verde-Tonto basin, in part because of the latter smaller basin's "flashier" streamflow. The multi-site calibration and cross-validation statistics for each basin's reconstruction can be found in **APPENDIX 3. APPENDIX 4** provides a detailed assessment of other models that have been used to reconstruct Colorado and Salt-Verde-Tonto streamflow time series.

(3-b) *Are the gaged and reconstructed records comparable in terms of mean?*

Note in **Figure 6** that in both basins, the mean of the observed period is higher than the longterm reconstructed mean. This indicates that, in general, the 20th century has been wetter in both basins than in previous centuries. (The recent drought years of the late 1990s and early 2000's were not included in this analysis, however, and their inclusion would lower the observed-record mean.)

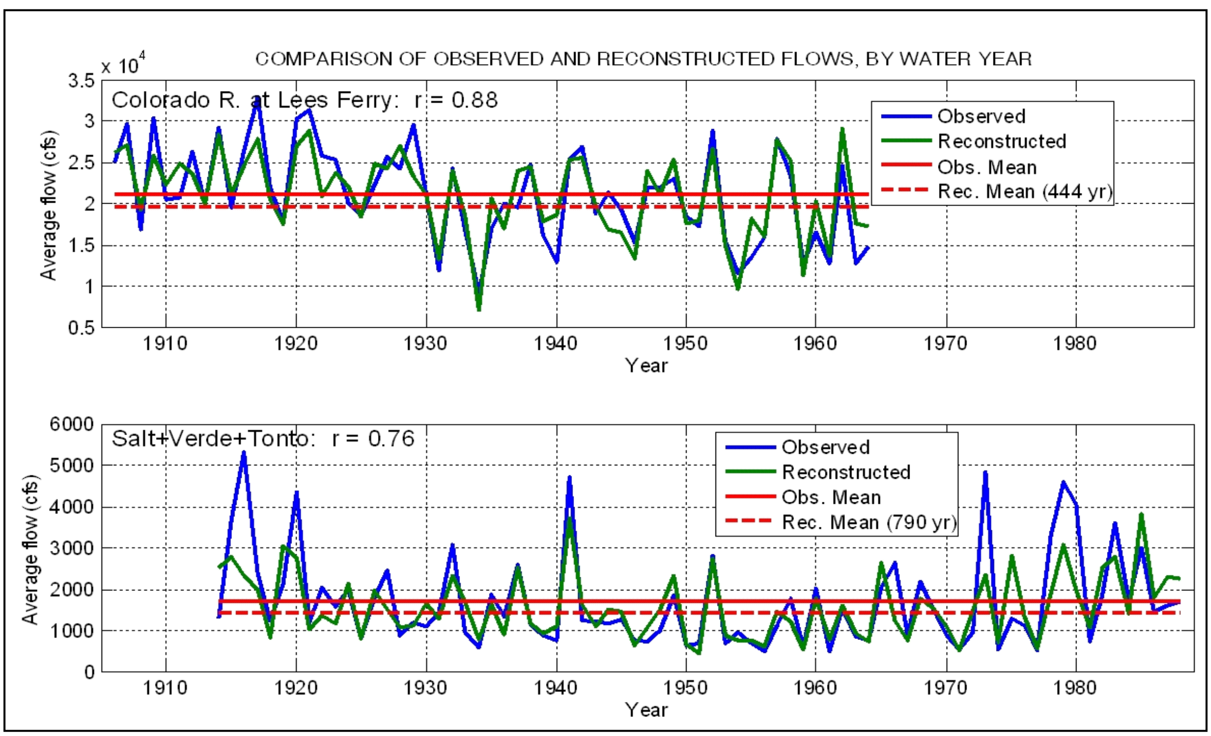


Figure 6 – Comparison of observed and reconstructed flows in the Upper Colorado and Salt-Verde-Tonto Basins

(3-c) *What do the plots of the entire reconstructed streamflow records look like?*

Figures 7a and 7b (next page) show the long-term reconstructed time series plots for averaged annual water year streamflow (in cubic feet per second) for each basin, with error bars shown in red. The 50% confidence interval indicated by the red bars is based on the root-mean-square error from the cross-validation procedure (Stage II). When comparing the two basins it is important to note the large difference in scaling of vertical axes seen in Figures 7a and 7b. In order to plot and compare high and low observed and reconstructed streamflow episodes for both basins on the same axis, the discharge was converted to percent of normal flow, and is plotted as such in several of the subsequent graphs.

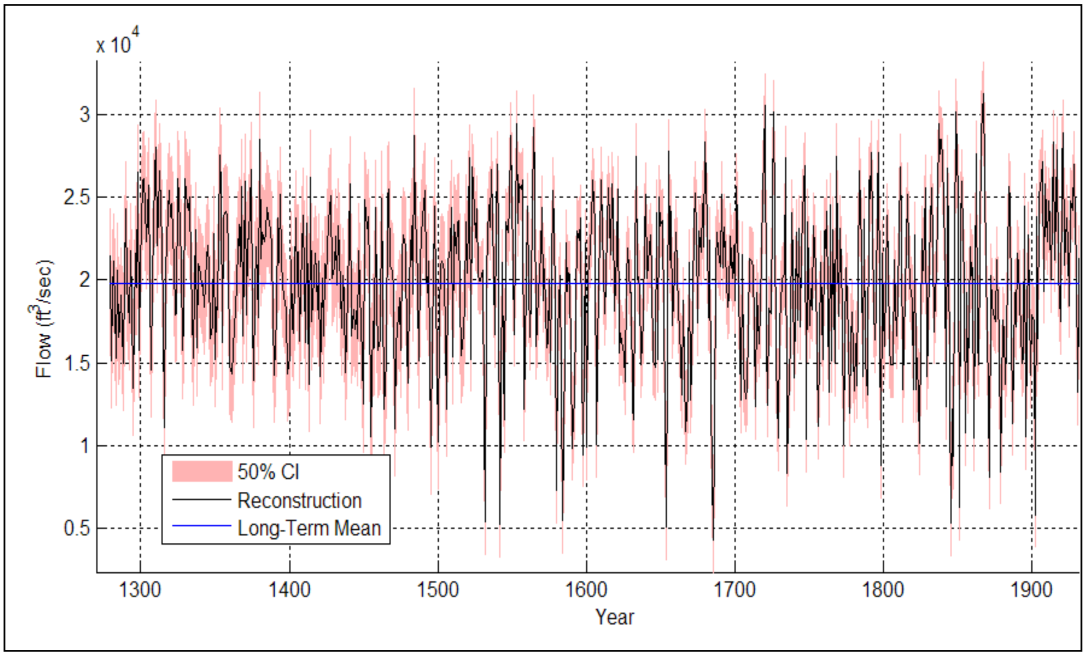


Figure 7a -- Reconstructed annual water year flows, Colorado River at Lees Ferry.

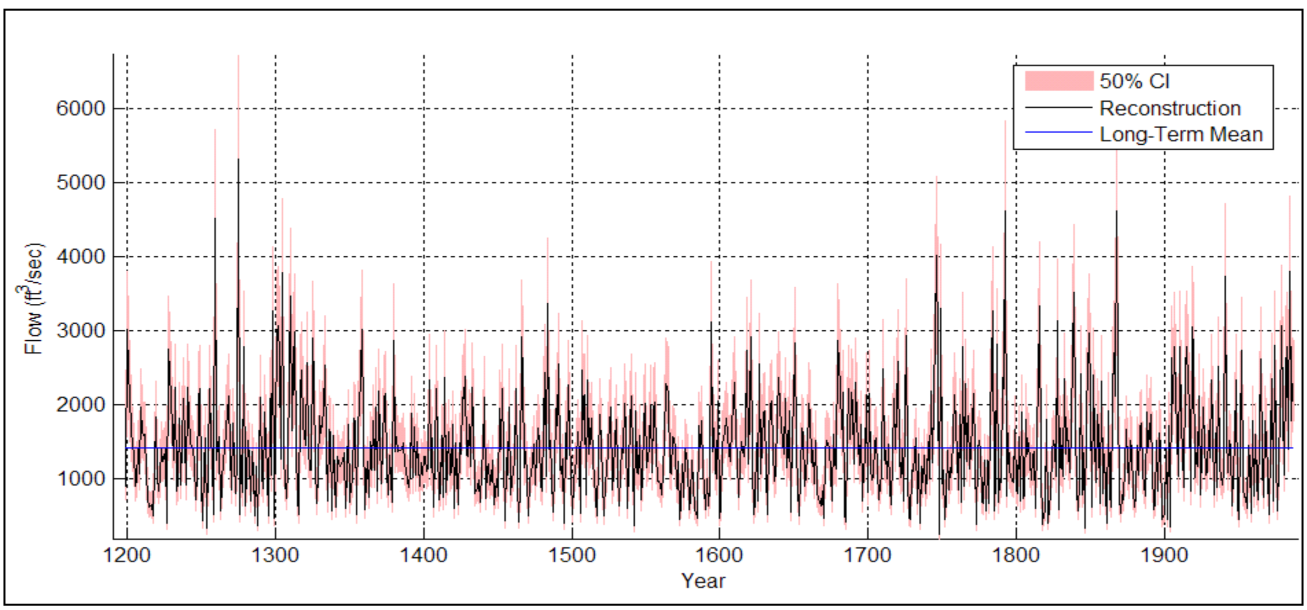


Figure 7b -- Reconstructed annual flows, Salt+Verde+Tonto Rivers, Arizona

4 – THRESHOLD IDENTIFICATION OF INDIVIDUAL AND SYNCHRONOUS EXTREME STREAMFLOW EPISODES IN BASINS

(4-a) *<u>What criteria were used to define extreme streamflow episodes in each basin?</u>*

Through discussions with SRP Water Resource Operations specialists, a threshold-based procedure was developed to identify the extreme low and high flow years in each basin. The thresholds used are quantiles of water-year annual discharge (compared to the mean). **High flow** years **(H)** are those with flow > 0.75 quantile and **Low flow** years **(L)** are those with flow < 0.25 quantile. At a later stage of the analysis, thresholds of > 0.90 and < 0.10 were also examined in order to assess the *most* extreme years. **Table 2** displays the quantile thresholds used in each basin along with corresponding streamflow units for each basin.

Table 2 – Thresholds for determining low (L) & high (H) flow extremes in each basin.

THRESHOLDS OF OBSERVED & RECONSTRUCTED FLOW FOR HH / LL ANALYSIS

BASIN		QUANTILE	% MEAN	CFS	Thousands of Acre-Feet
UCRB @ Lees Ferry	L	0.25	77.4 83.3	15,910 16,326	11,526 11,827
	H	0.75	119.9 119.6	24,649 23,422	17,857 16,968
S + V+ T (Salt + Verde + Tonto)	L	0.25	49.5 63.1	835 887	605 642
	H	0.75	123.1 126.1	2,077 1,772	1,505 1,283

	Observed	Reconstructed
Base period for quantiles: water years	1914-2001	1521-1964
Means: A. Colorado R at Lees Ferry:	20,564 cfs	19,589 cfs
B. Salt, Verde and Tonto:	1,687 cfs	1,405 cfs

Key episodes of extreme low flow and high flow in the basins were identified in both the gaged and reconstructed streamflow records. Upper .25 and Lower .75 thresholds were selected and applied to water year mean discharge in both the observed and reconstructed records. **Figure 8** shows the plotted threshold levels for each basin. Note that the thresholds differ somewhat between the observed and reconstructed series due to differences in the mean and standard deviation of the two series.

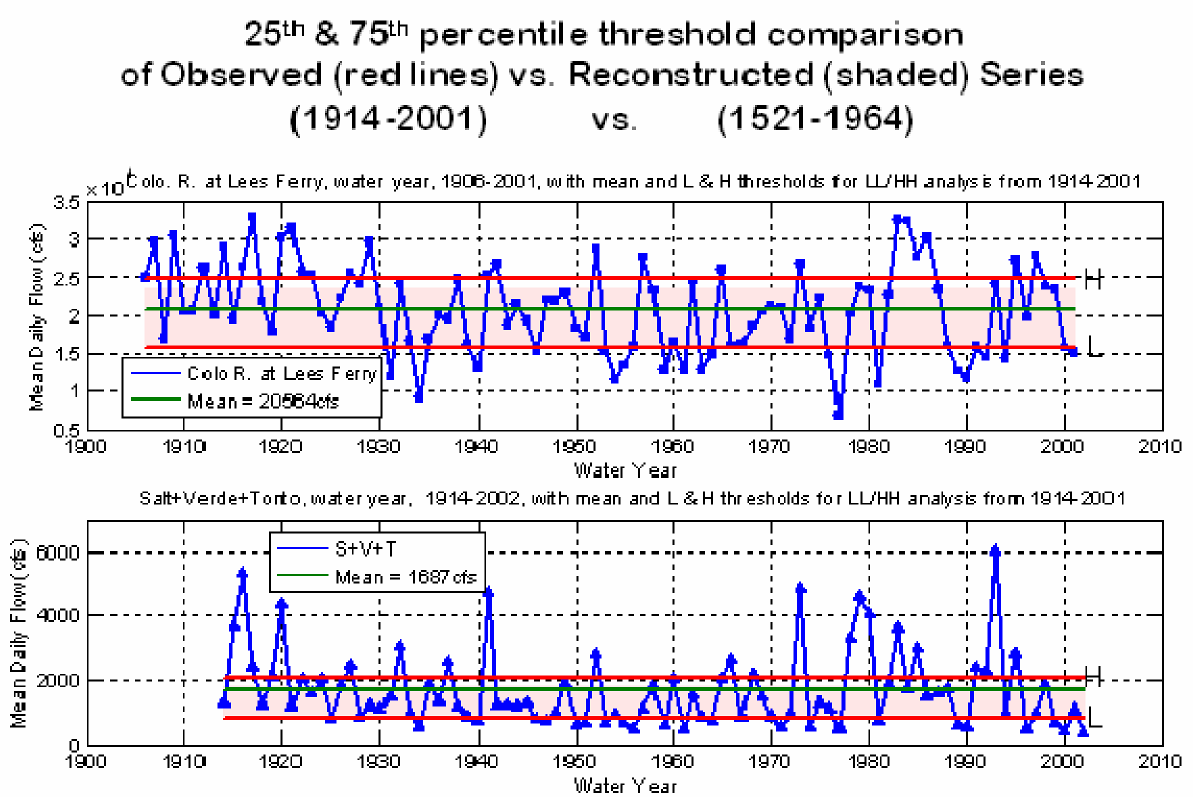


Figure 8 – Comparison plot of quantile thresholds for observed & reconstructed time series

(4-b) *What types of synchronous extreme streamflow scenarios have occurred in both the observed and reconstructed record?*

Four synchronous scenarios are possible when comparing the Upper Colorado and the Salt-Verde-Tonto River systems and are indicated by a two-letter code. (For consistency, the first letter will always refer to the Upper Colorado and the second to the Salt-Verde-Tonto):

When the two basins are responding in the same direction:

HH = High flow (H) in the Colorado at the same time as high flow (H) in the Salt-Verde-Tonto
LL = Low flow (L) in the Colorado at the same time as low flow (L) in the Salt-Verde-Tonto

When the two basins are responding in the opposite direction:

HL = High flow (H) in the Colorado at the same time as low flow (L) in the Salt-Verde-Tonto
LH = Low flow (L) in the Colorado at the same time as high flow (H) in the Salt-Verde-Tonto

(4-c) *What were the results?*

The graphical results of the quantile threshold analysis -- including the number of **HH**, **LL**, **HL** and **LH** events -- are displayed in **Figures 9 and 10** for the observed record and **Figures 11 and 12** for the reconstructed record. In both the observed and reconstructed records, **HH** and **LL** events were much more frequent than **HL** and **LH** events, especially in the long, 444-year reconstructed time series. In fact, no **HL** events at all occurred in the reconstructed record, and only 2 **LH** events occurred. In the observed record, only 3 **HL** events and no **LH** events occurred. In order to examine some **LH**-like scenarios in the observed record, the Colorado threshold was relaxed to < 0.50, yielding **LH** events. Due in part to the quantile method, the number of **LL** events tends to be counterbalanced by the number of **HH** events, but overall – in both the observed and reconstructed records – **LL** events are more frequent occurrences than **HH** events. A year-by-year listing of the **HH**, **LL**, **HL** and **LH** events, color coded by each year's streamflow threshold (>0.90, > 0.75, < 0.25, < 0.10) is presented in **APPENDIX 5**.

(4-d) *Were these results expected?*

A widely held working hypothesis is that the Upper Colorado River Basin (UCRB) can serve as a buffer to compensate for extreme low flow in the Salt-Verde-Tonto Basin during drought periods. One reason this is plausible is because of the size of the UCRB basin and the overall magnitude of its discharge. It has also been assumed that streamflows in the two river systems are relatively independent of each other due to a difference in the climatic regimes influencing each basin (e.g. R. Sedlacek and R. Siegal, unpublished report).

While the above presumptions sound logical, our analysis did not substantiate them. We found that the simple correlation between the annual flow values of the two 444-year reconstructed time series was r = .599 (r^2 = 0.359), which is a statistically significant correlation, given the large sample size. Far more compelling with respect to extreme events, however, was the overwhelming dominance of **HH** and **LL** events in comparison to extremes where the basins responded in the opposite direction.

(4-e) *What are the implications of these results?*

These results suggest that annual streamflow variability in the Salt-Verde Basin – especially extreme streamflow – is *not* independent of annual streamflow variability in the Colorado Basin. Therefore it is not unlikely that severe drought in one basin will be accompanied by severe drought in the other basin, such as has occurred during the western drought scenario of the late 1990s and early 2000's. It should be noted that the immense water supply of the large Upper Colorado basin may allow it to continue to serve as a buffer even when it is in a low flow scenario because a little flow can go a long way in a smaller watersheds such as the Salt-Verde-Tonto, however ever-increasing demand on the water supply of the UCRB by many stakeholders is likely to reduce the size of this buffer in the future.

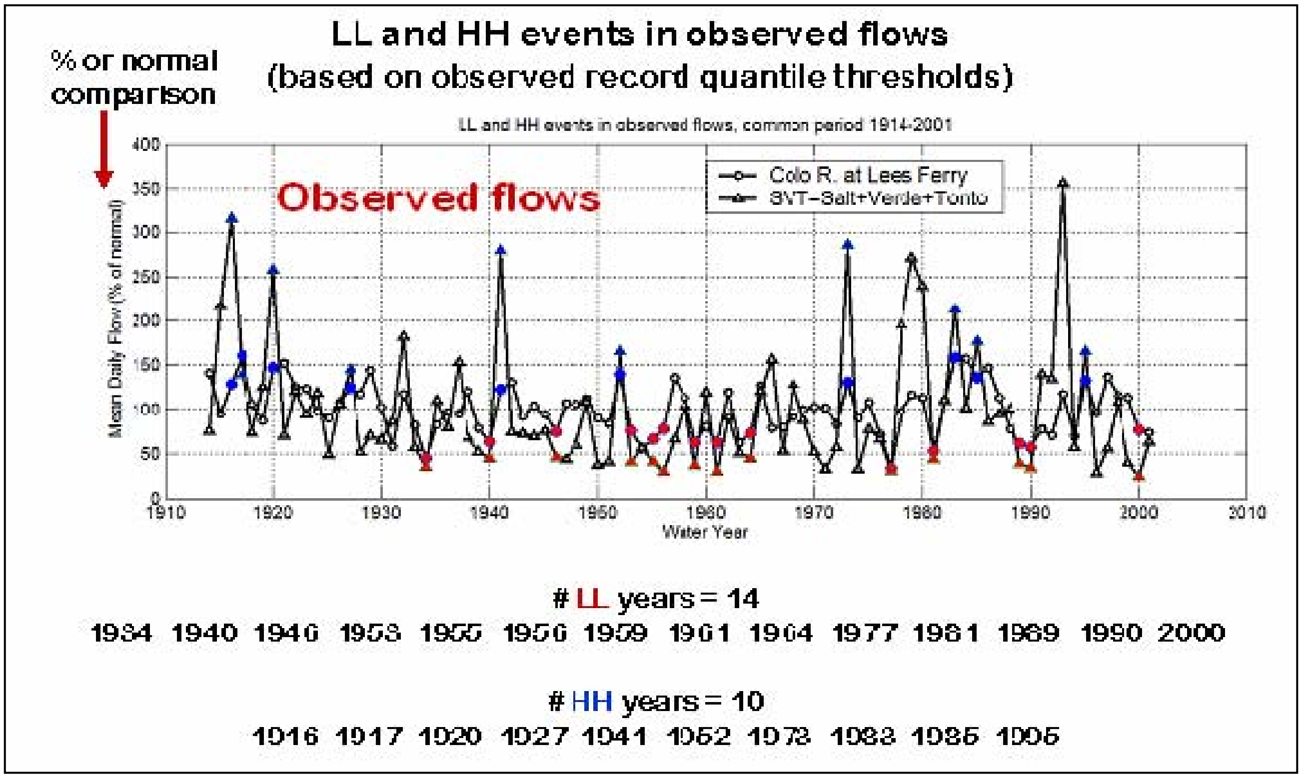


Figure 9 – 20th century **HH** and **LL** water years based on observed-record quantile thresholds

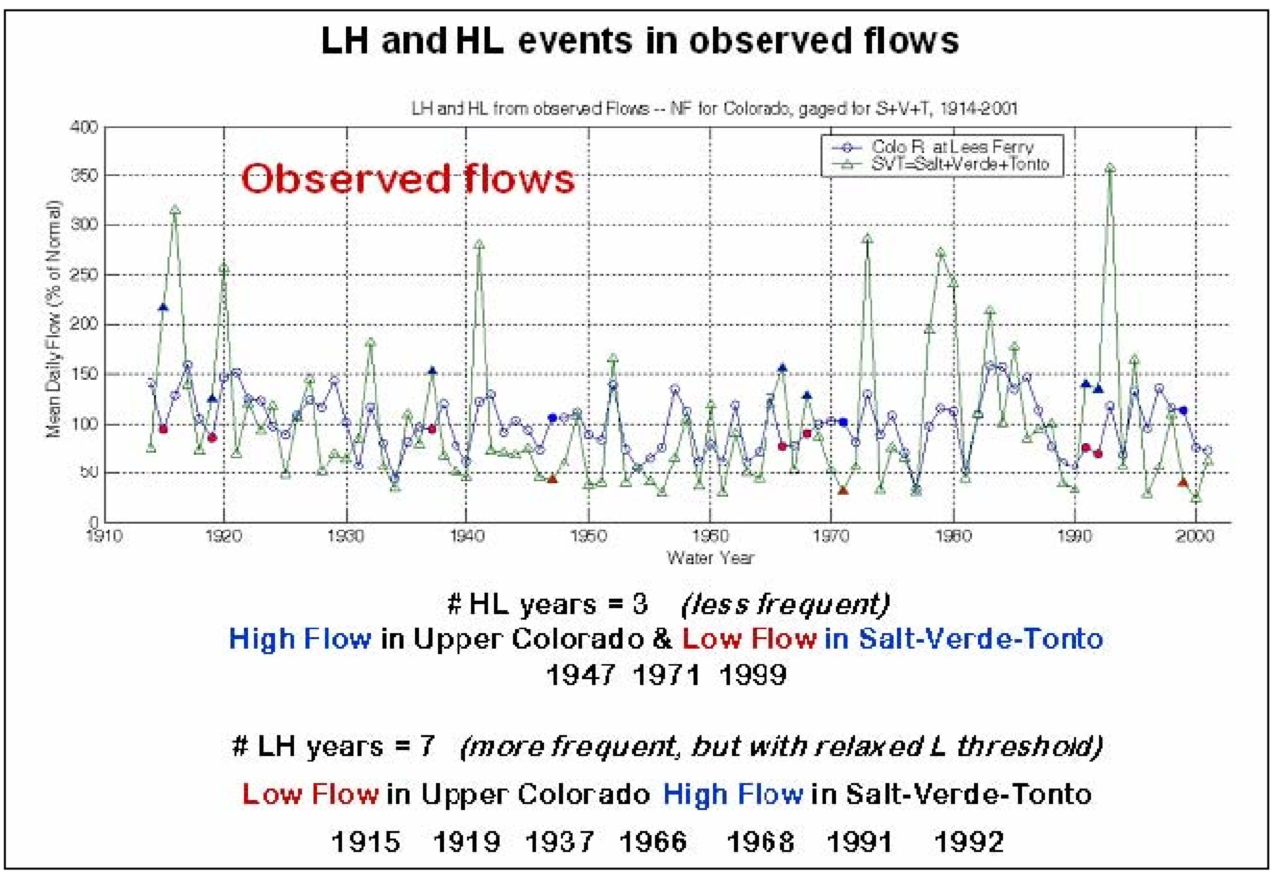


Figure 10 – 20th century **HL** and **LH** water years based on observed-record quantile thresholds
[NOTE: ***LH*** *years were identified only by relaxing the Colorado* ***L*** *threshold to 0.50]*

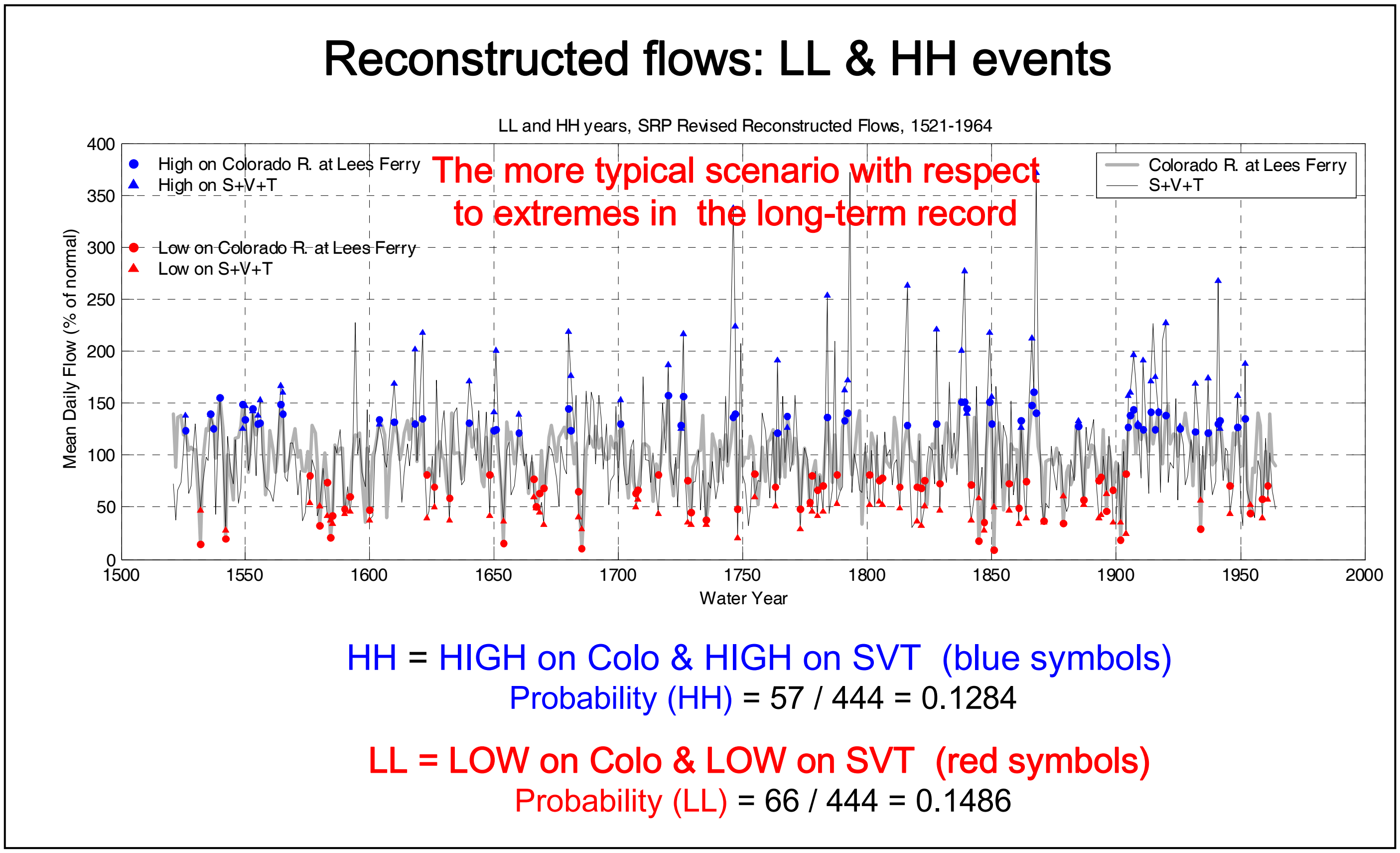


Figure 11 – **HH** and **LL** water years based on reconstructed-record quantiles from 1521 - 1964

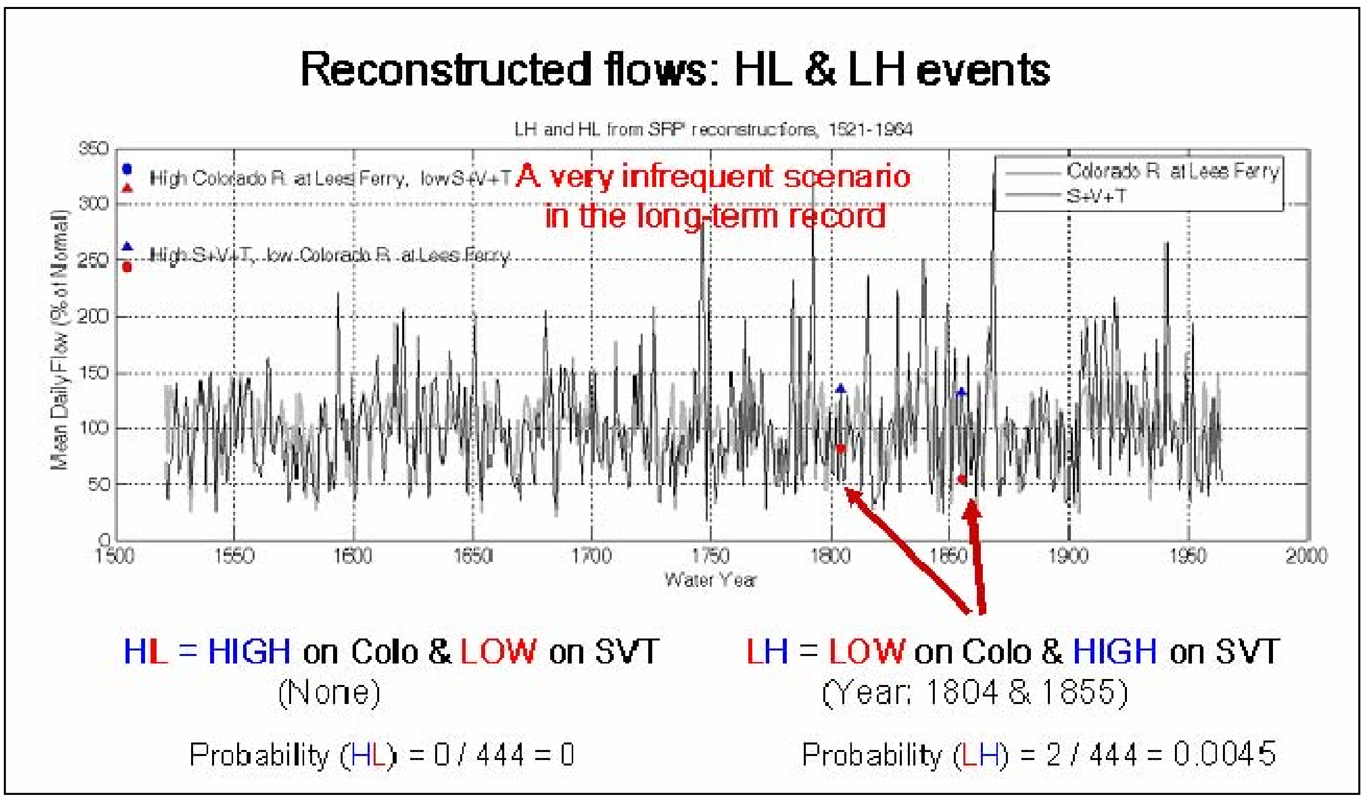


Figure 12 – **HL** and **LH** water years based on reconstructed-record quantiles from 1521 - 1964

(4-f) *Is the observed frequency of occurrence of **HH** and **LL** years greater than can be expected by chance?*

A hypergeometric test (see box) indicates that the probability of the observed numbers of joint **HH** and **LL** years is less than one in a billion if the extreme streamflow in the two basins were indeed independent of one another!

Hypergeometric Test

Using the 0.25 and 0.75 quantiles of reconstructed flow for the full 444-year overlap of the SVT and Lees Ferry reconstructions as thresholds for dry years and wet years, we designate by definition 444/4 = 111 low-flow years and 111 high-flow years in each basin. A count indicates that in 66 years flows were low in both rivers (66 LL years) and in 57 years flows were high on both rivers (57 HH years).

The hypergeometric distribution can be used to show that 66 LL years or 57 HH years is much greater than expected if flows in the two rivers were independent of one another.

The hypergeometric distribution gives the probability of drawing up to *x* of a possible *k* items in *n* drawings without replacement from a group of *m* objects.

Assume the "items" are defined as low-flow years in the Lees Ferry reconstruction, and that the "drawings" are samples from the Lees Ferry reconstruction in years with low flow on the SVT reconstruction. The problem for the **LL** case is set up as follows:

k=111 is the number of possible items (low-flow years in Lees Ferry record)
n=111 is the number of drawings, or years of low flow in the SVT record
m=444 is the number of objects, the total number of years in the Lees Ferry record
x=65 is one less than the number of number of successful drawings, or the number of drawings that results in one of the k items (in other words, an **LL** year)

The cumulative distribution function of the hypergeometric distribution gives the probability of x or fewer successful drawings. If this probability is given by p^*, then

$$p = 1 - p^*$$

is the probability of more than x successful drawings. For the example above, p is the probability of getting 66 or more **LL** years if the occurrence of low-flow years on the two rivers is indeed unrelated. The estimated chance probabilities for 66 or more **LL** years and 57 or more **HH** years are listed below

	k	*n*	*m*	*x*	*p**	*p*
LL	111	111	444	65	>0.99999999999999	<1E-12
HH	111	111	444	56	0.9999999999989624	1.037E-12

In summary, the probability of the observed numbers of **LL** and **HH** years is less than one in a billion if the two flow record were indeed independent of one another.

(4-g) *What's the probability of different lengths of LL and HH sequences ?*

To address this question, probabilities of **HH** and **LL** events were empirically derived based on the number of their occurrences in the 444-year reconstructed times series. In addition, a moving-window procedure (see **Figure 13**) was used to determine the number of **HH** and **LL** events in a moving *n-year* window, both sequentially and non-sequentially.

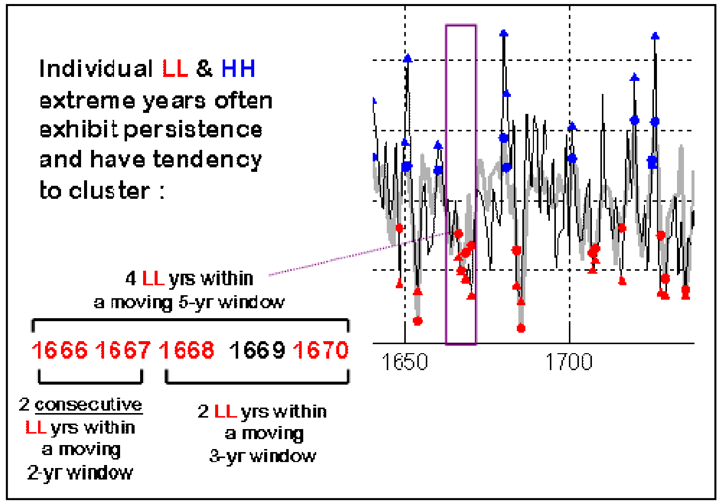


Figure 13 –Moving window probability count procedure

The results displayed in **Table 3** indicate that there is > 10% chance that a single year could be an extreme **LL** year or an extreme **HH** year. There is also a 10% chance that 2 **LL** years (or **HH** years) could occur in any given 4-year period and a > 5 % chance that 2 years out of a moving 3-year window will be extreme **LL** years (or **HH** years).

Table 3 -- Probability Counts of LL & HH Event Scenarios

Most probable	Over the period 1521-1964	LL # events/ # possible (probability)	HH # events / # possible (probability)
***	Individual 1-yr events	66 / 444 (0.149)	57 / 444 (0.128)
	Consecutive Sequences		
*	2 consecutive years *(within a moving 2-yr window)*	10 / 443 (0.023)	14 / 443 (0.032)
	3 consecutive years *(within a moving 3-yr window)*	1 / 442 (0.002)	3 / 442 (0.007)
	Clustered Sequences		
**	2yrs *(within a moving 3-yr window)*	22 / 442 (0.050)	29 / 442 (0.066)
***	2 yrs *(within a moving 4-yr window)*	45 / 441 (0.102)	47 / 441 (0.107)
	3 yrs *(within a moving 4-yr window)*	5 / 441 (0.011)	9 / 441 (0.020)
*	3 yrs *(within a moving 5-yr window)*	13 / 440 (0.030)	16 / 440 (0.036)
	4 yrs *(within a moving 5-yr window)*	1 / 440 (0.002)	0 / 440
	4 yrs *(within a moving 6-yr window)*	1 / 439 (0.002)	0 / 439
	5 yrs *(within a moving 6-yr window)*	0 / 439	0 / 439

*** = probability > 10% ** = probability > 5% * = probability > 2%

Figure 14 illustrates another perspective on the tendency for **HH** and **LL** events and episodes to cluster, using a smoothed time series approach. It demonstrates the cumulative effect that short-term clusters of extreme years can have over longer time periods.

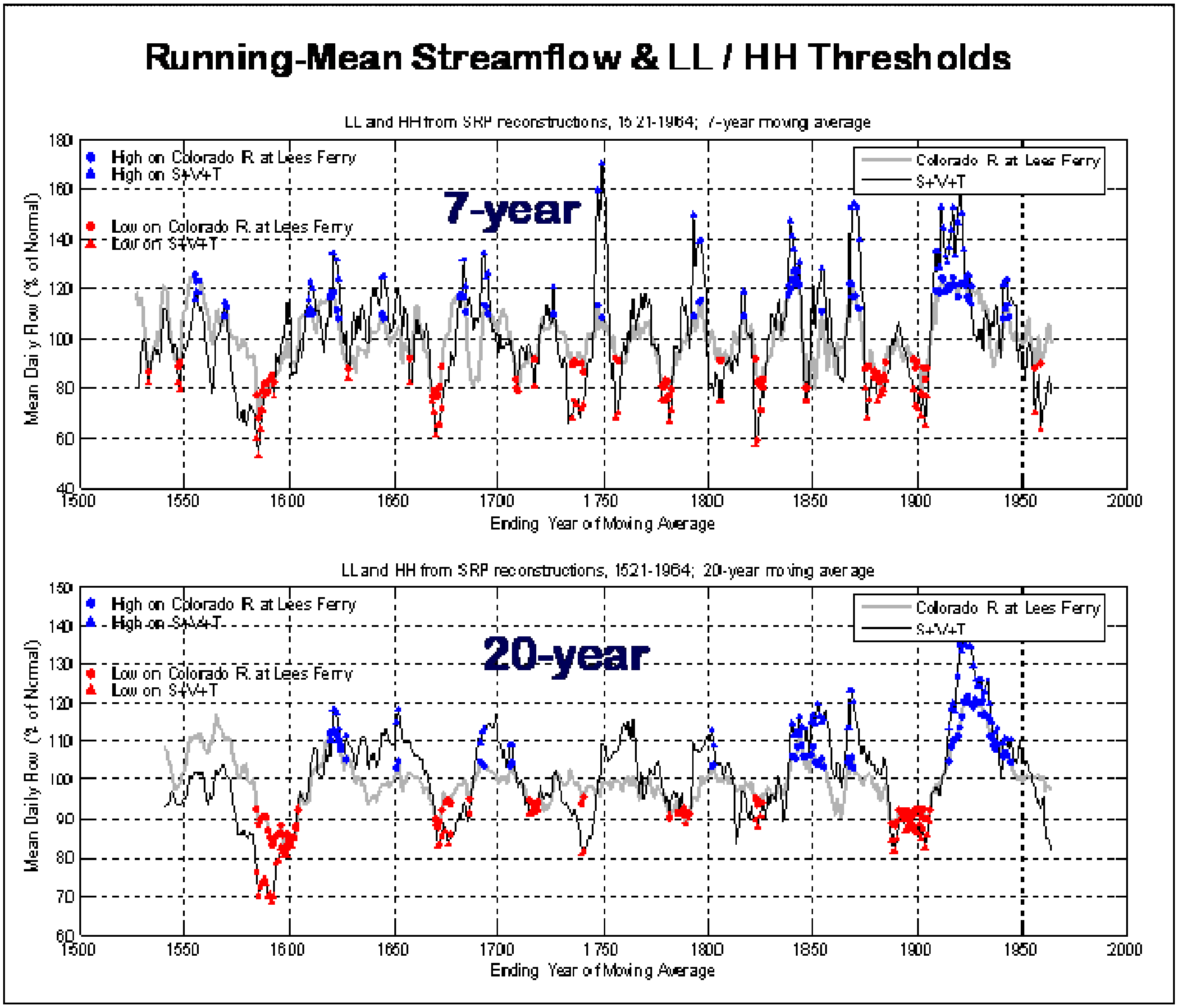


Figure 14 – Smoothed time series: thresholds based on 7-year and 20-year running means

(4-h) *What are the implications of this tendency for extreme years to occur in sequences or clusters?*

It is possible that if the number of wet extreme years is about equal to the number of dry extreme years, that the two extremes could "cancel each other" on a year-to-year basis such that there would be little long-term stress on water supply operations. However, because of the clustering tendency of extreme events, it is more probable that episodes of sustained drought or sustained high flow will persist, placing more of a burden on water systems management and operations. In other words, reservoir storage will buffer water supplies in the Upper Colorado River Basin and Salt-Verde-Tonto region during single year droughts, but supplies are increasingly strained as droughts extend over multi-year periods.

(4-i) *How else can the existence of a linkage in multi-year drought occurrence in the two basins be supported statistically?*

Monte Carlo Study of Simultaneously Low Moving Average Flows

To summarize multi-year simultaneous drought, the reconstructed flow series for the Colorado River at Lees Ferry and the Salt-Verde-Tonto (SVT) were analyzed as moving averages, or running means, using the 0.25 quantile of running means for the base period 1521-1964 as a relative drought threshold. A count of running means below the respective threshold on both rivers represents the frequency of joint drought for any given length of moving average. The single year analysis, which yielded 66 **LL** events and 57 **HH** events, can be thought of in this sense as a one-year moving average. We have generalized this method of tabulating drought (**LL** events) to moving averages of length 1, 2, 3, …, 20 years.

It is useful to compare counts of **LL** events to the expected number of events given no relationship between the flow series for the Colorado and the SVT, but a theoretical expected frequency is not straightforward to calculate because once the time series has been smoothed into a moving average the independence of successive observations is compromised. We therefore took the following Monte Carlo approach:

1. Convert the annual reconstructed flows, 1521-1964, to moving averages and count the number of **LL** events
2. Repeat the count for the SVT against 1000 simulated reconstructed flow series for the Colorado at Lees Ferry
3. Compare the observed count of joint low-flow events to the cumulative empirical distribution of counts from the Monte Carlo simulation.

For the Monte Carlo simulations, the 1521-1964 reconstructed flow series for Lees Ferry was modeled as a first-order autoregressive (AR(1)) process. The model is given by

$$y_t = 0.208 y_{t-1} + e_t$$

where y_t is the flow in year *t* expressed as a departure from the 1521-1964 mean, and e_t is a random sample a standard normal distribution. A total of 1000 simulations were generated in this way, and each was subsequently scaled to the same mean and variance as the reconstructed flow series.

The results of the Monte Carlo analysis are summarized in **Table 4** (next page)**,** which lists for each length of moving average the number of **LL** events in the reconstructed flows and threshold number of **LL** events with empirical non-exceedance probability 0.50, 0.95, 0.99 and 0.999 by chance. For example, for the 7-year moving average, the Lees Ferry reconstruction and SVT were simultaneously below their drought thresholds in 62 years, which is far more than expected by chance: the Monte Carlo tally indicates the probability is $p = 0.999$ of fewer than 50 **LL** events if the flows on the Colorado and the SVT are unrelated. The expected number of chance events, given by the $p = 0.50$ probability point is only 27. In summary, none of the observed **LL** counts (column labeled N_3) has a greater than 0.001 probability of occurring by chance alone. Since there were only 1000 simulations done for the analysis, it is also true *no* simulation had a count of **LL** events as large as for the reconstructed series themselves.

The results strongly support a linkage in multi-year drought occurrence in the two basins.

Table 4. Monte Carlo summary of frequency of jointly low reconstructed moving-average flows on the Colorado River at Lees Ferry and the Salt-Verde-Tonto Rivers.

	Threshold[2] Flow (cfs)		Events[3]			Probability Point[4]			
M[1]	Lees F	SVT	N_1	N_2	N_3	.50	.95	.99	.999
1	16326	887	444	111	66	28	35	38	40
2	16943	1061	443	111	58	28	36	38	41
3	17139	1108	442	110	65	27	36	41	47
4	17514	1163	441	110	62	27	36	41	49
5	17597	1174	440	110	62	28	37	41	47
6	17814	1188	439	110	66	27	38	42	49
7	17954	1197	438	109	62	27	38	42	50
8	18115	1220	437	109	65	27	39	44	49
9	18135	1240	436	109	67	27	40	46	49
10	18255	1252	435	109	67	27	40	47	52
11	18306	1251	434	108	66	26	40	47	51
12	18417	1260	433	108	64	27	41	48	51
13	18427	1264	432	108	65	27	42	47	54
14	18590	1275	431	108	71	27	43	48	56
15	18594	1280	430	107	64	26	43	49	55
16	18630	1279	429	107	66	26	44	50	54
17	18652	1284	428	107	71	26	45	50	55
18	18658	1304	427	107	66	27	44	50	54
19	18768	1303	426	106	66	26	44	49	55
20	18775	1308	425	106	63	26	44	51	57

[1] Length of moving average (yr)

[2] Threshold for low-flow event, defined as 0.25 quantiles of the moving-average annual reconstructed flows, 1521-1964

[3] Number of moving averages in series (N_1), number of moving averages below threshold in each series(N_2), and number of simultaneous low flows (N_3)

[4] Threshold number of joint droughts with specified probability of not being exceeded, based on 1000 Monte Carlo simulations

(4-j) *How have low and high streamflow episodes in each basin individually varied from century-to-century? How have the numbers of* ***LL*** *and* ***HH*** *years varied from century-to-century?*

Figure 15 (and **APPENDIX 5**) display a century-by-century record of extreme streamflow events in the Upper Colorado and the Salt-Verde-Tonto River basins. The 19th century experienced more **LL** years than any other in the record (22 events) and the 18th and 20th centuries (prior to 1965) experienced the most **HH** years (14 events).

Figures 16 and 17 display century-by-century timelines of extreme **LL** and **HH** events showing the individual-year discharge values associated with the extreme events and their tendency to cluster. **Figures 18 and 19** show similar plots, but for sub-basins of the Upper Colorado (Green, Dolores and San Juan Rivers) and the Verde, Salt + Tonto, and Gila Basins in Arizona. The sub-basin plots provide more detailed spatial information on where the extreme streamflow events were concentrated in each watershed.

Figure 15 – **LL** and **HH** water years with quantile time series

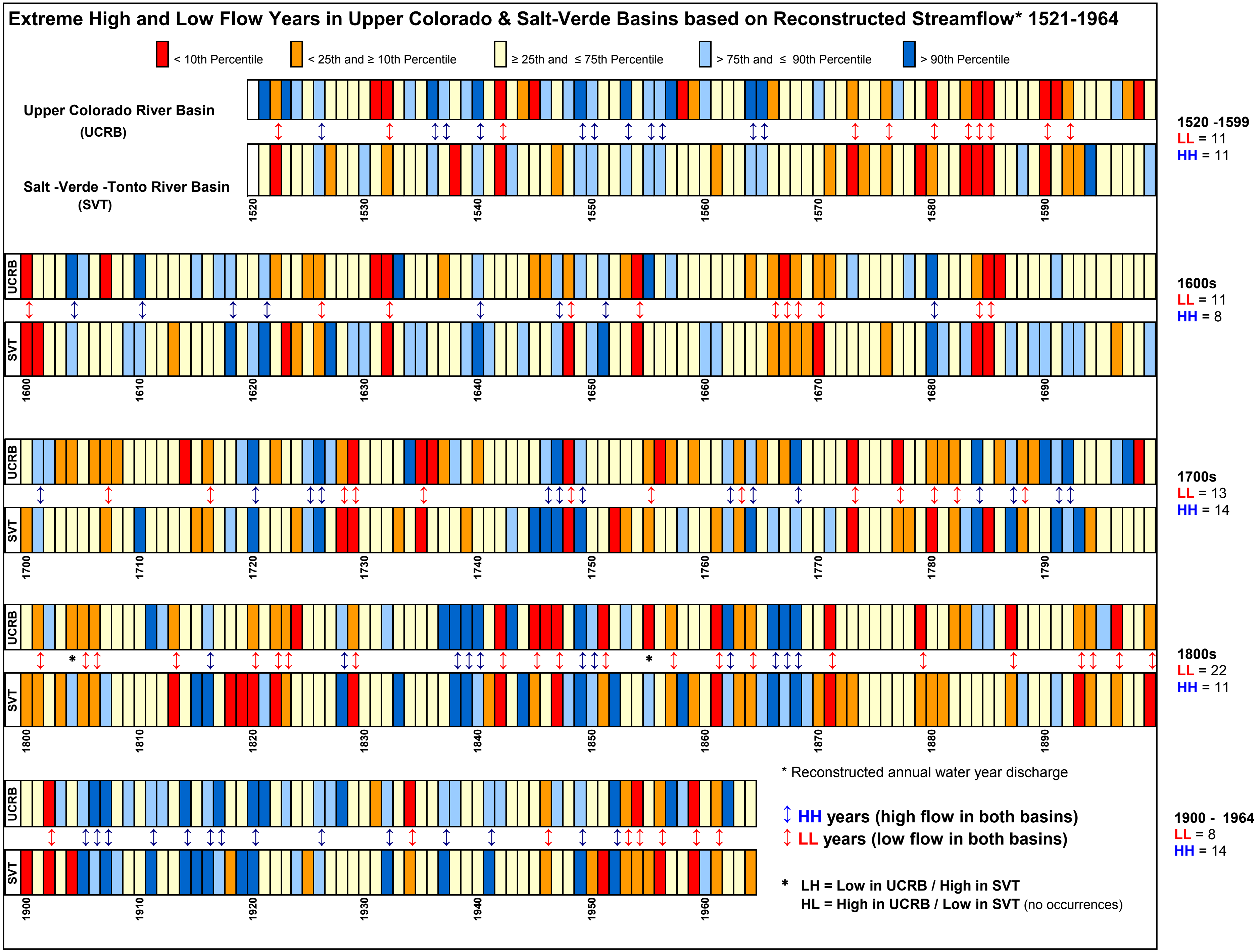


Figure 16 – Low discharge events in each basin by century, showing tendency to cluster

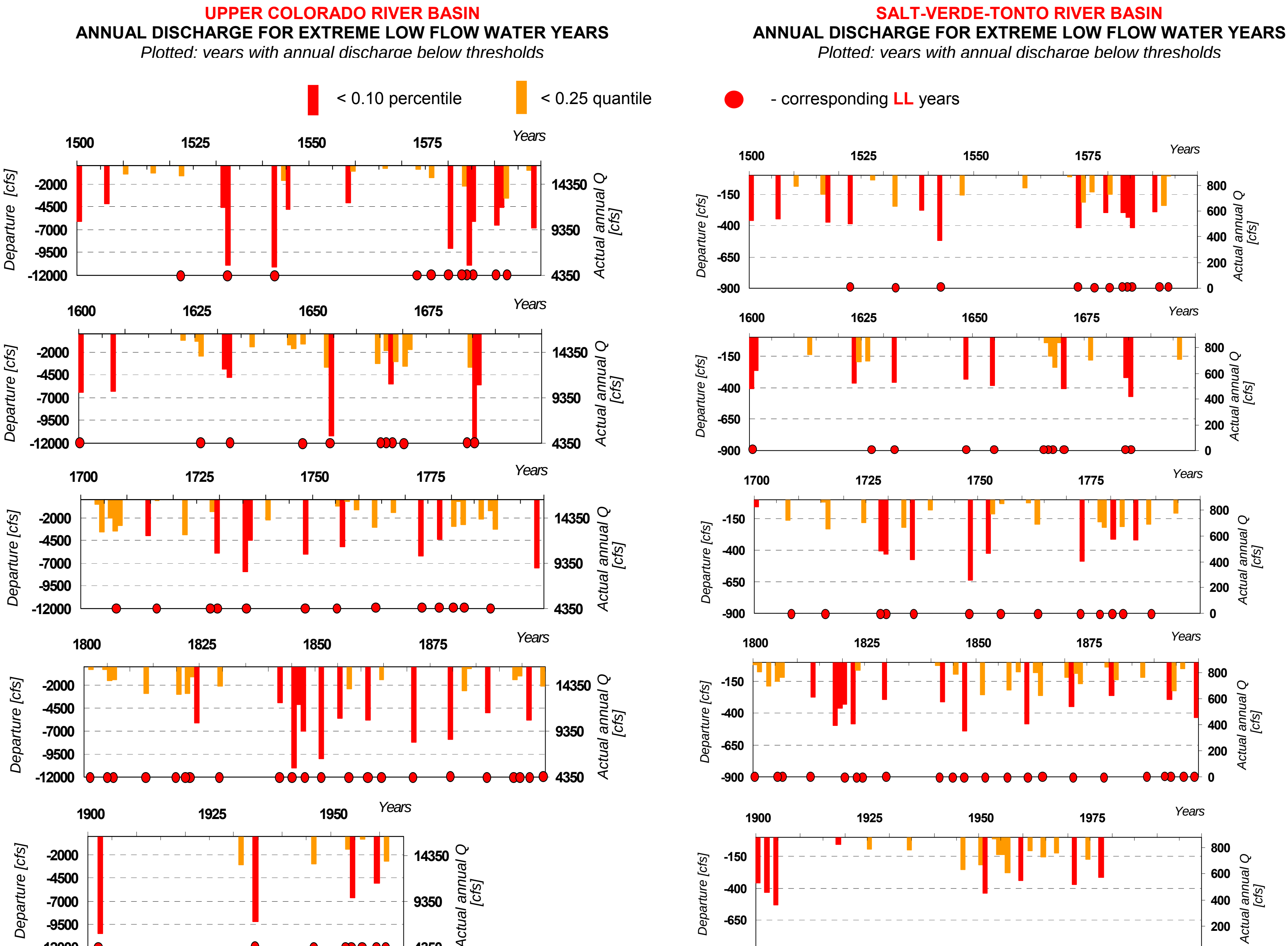


Figure 17 – High discharge events in each basin by century, showing tendency to cluster

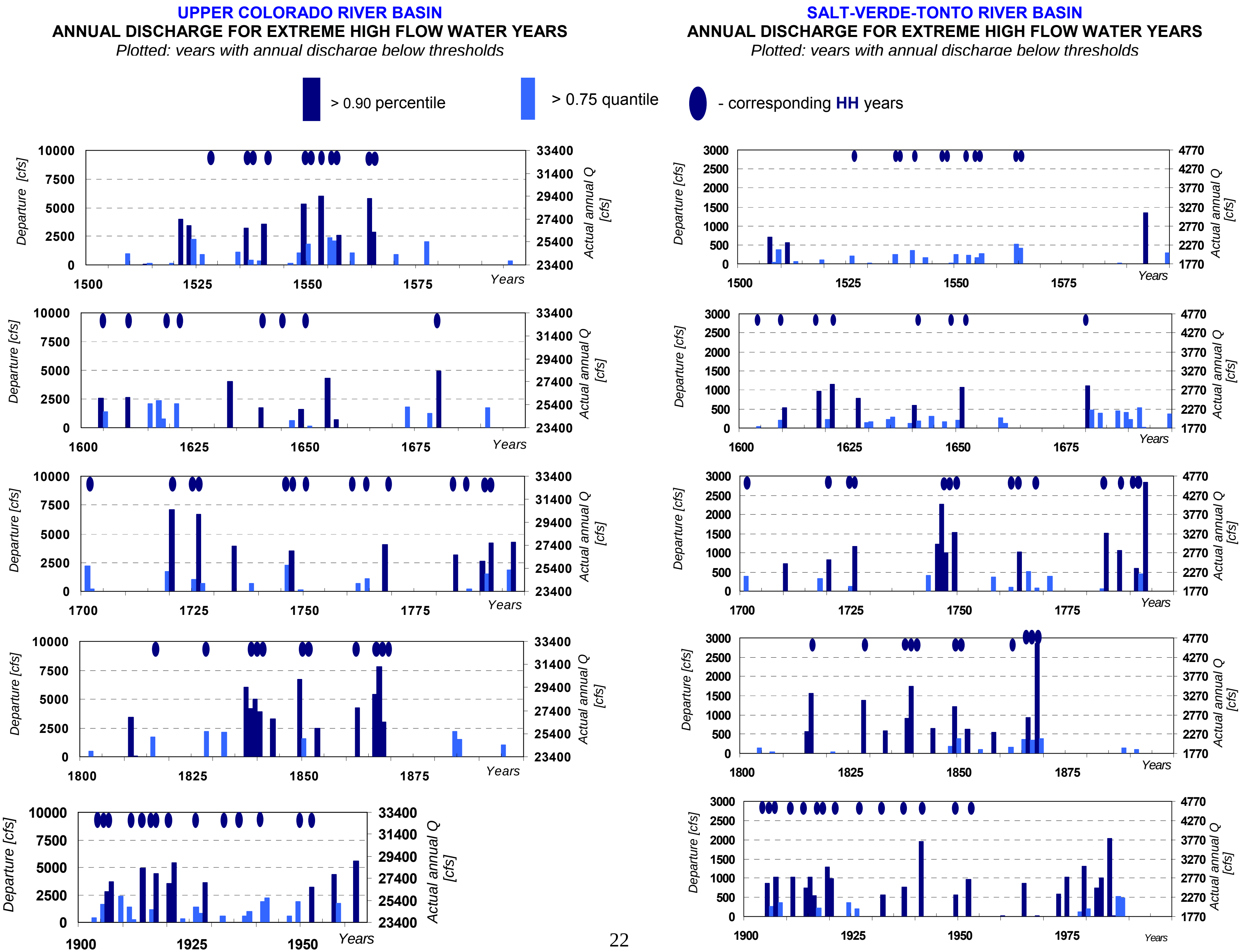


Figure 18 – Sub-basin comparison of low discharge events by century, showing tendency to cluster

Plotted: annual discharge of years below threshold in individual watersheds

< 0.10 percentile   < 0.25 quantile

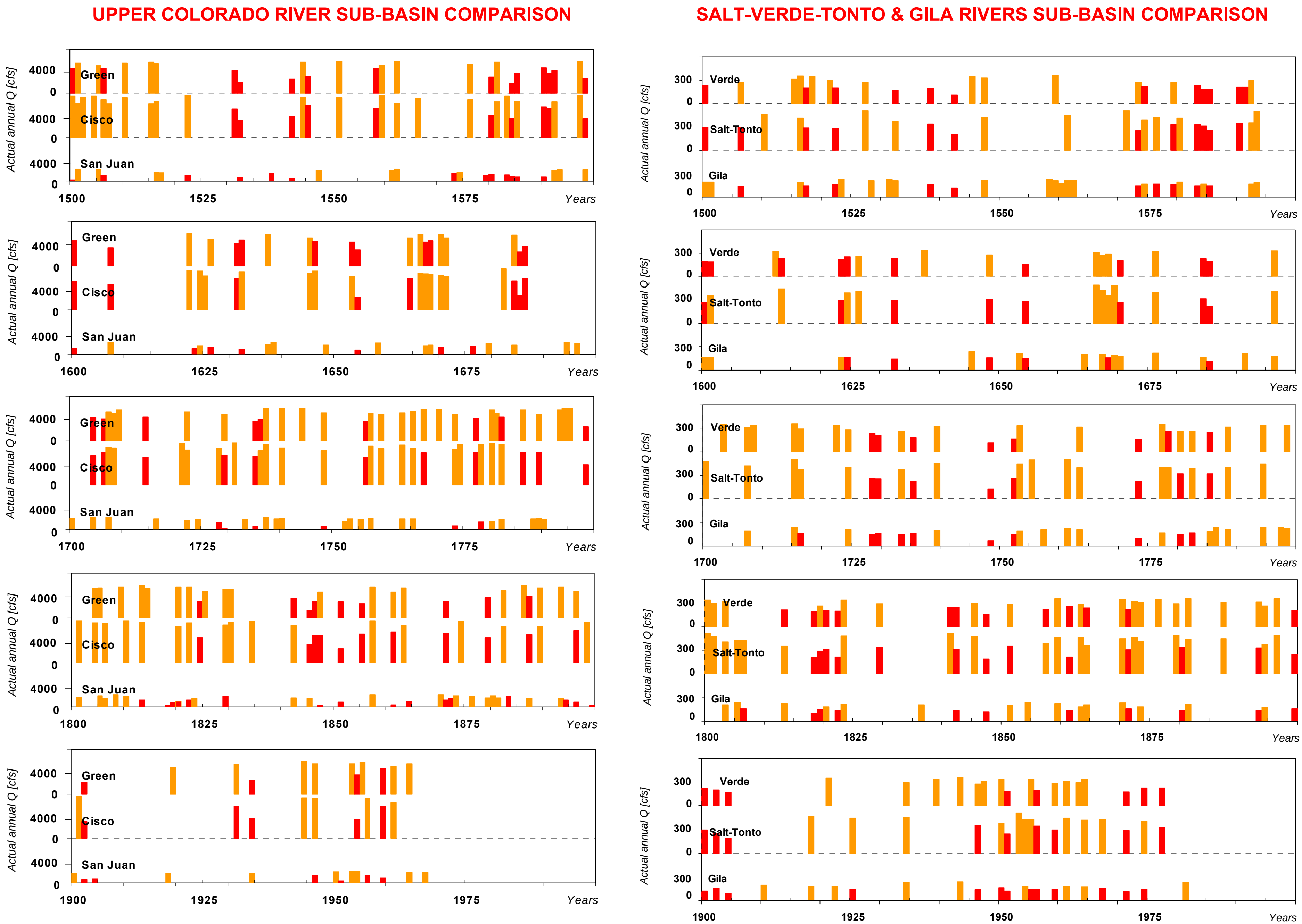
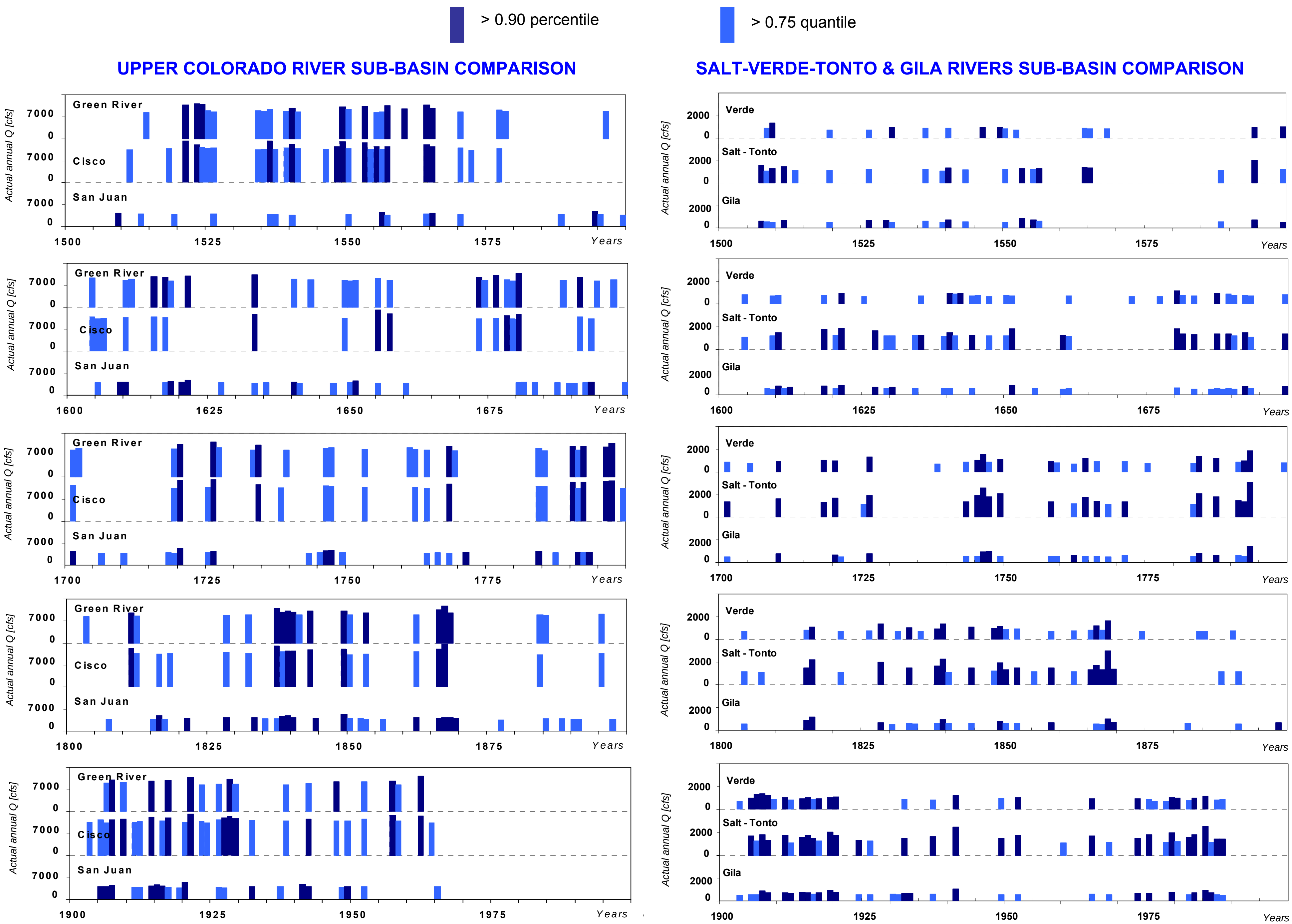


Figure 19 – Sub-basin comparison of high discharge events by century, showing tendency to cluster

Plotted: annual discharge of years above threshold in individual watersheds



5 – CLIMATIC CONTEXT OF SCENARIOS LEADING TO INDIVIDUAL AND SYNCHRONOUS EXTREME EVENTS

(5-a) *Does the precipitation record substantiate the project findings of synchronous extreme events in the two basins?*

The instrumental record of precipitation supports the finding to a certain degree. **Figure 20** shows percentiles of 20th century annual precipitation in the Upper Colorado and Salt-Verde-Tonto Basins, based on divisional precipitation data. There are differences between the two basins such that not all of the driest and wettest years are shared indicating a difference in storm track location, etc. However, many of the *most* extreme dry (< 0.10) years are regional events. A contingency table of joint occurrence of precipitation extremes (**Table 5**) strongly suggests that when one basin is extremely wet, so is the other and vice versa for dry years.

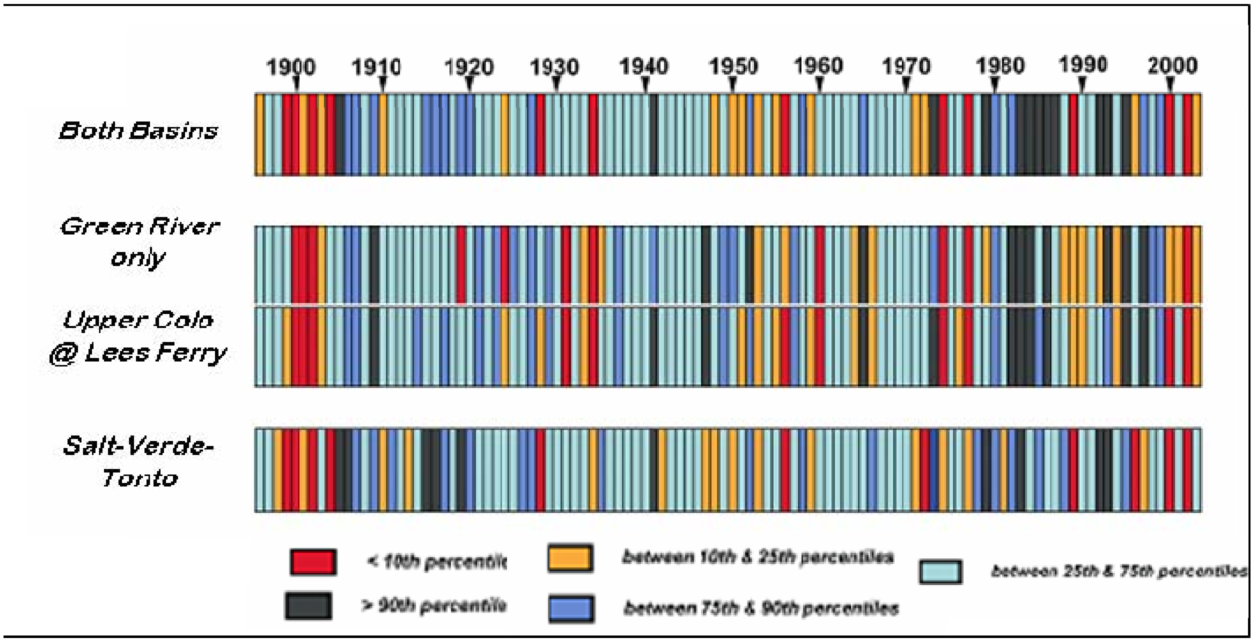


Figure 20 – Climatic context: 20th century time series of divisional precipitation

Table 5 – Joint occurrence of quantile-based divisional precipitation

JOINT OCCURRENCE OF PRECIPITATION EXTREMES IN UPPER COLORADO & SALT-VERDE-TONTO

1 – driest years, < 10th percentile;
2 – dry years, between 10th & 25th percentiles;
3 – average years, between 25th & 75th percentiles;
4 – wet years, between 75th & 90th percentiles,
5 – wettest years, > 90th percentile

(Based on divisional data roughly coinciding with each basin)

Shaded boxes = basins in sync

C-at-LF-an-avr	SVT-an-avr 1	2	3	4	5	sum	
1	5	4	2	0	0	11	108
2	3	2	10	1	0	16	
3	3	8	30	7	6	54	
4	0	0	8	6	2	16	
5	0	2	3	3	3	11	
sum	11	16	53	17	11	46	
	108						

(5-b) *<u>What atmospheric circulation patterns lead to the **LL**, **HH**, **LH** **and** **HL** scenarios?</u>*

Seasonal composites of upper level (500 mb geopotential height) circulation anomalies for each type of extreme event scenario (based on the observed record) are shown in **Figure 21a and 22b**. The characteristic circulation pattern for **LL** events is higher-than-normal upper level pressure over the west in early winter (Oct -Dec) and over the North Pacific ocean storm track region in mid- to late winter (Jan - Mar). The inverse of this pattern leads to **HH** events. **LH** and **HL** scenarios (Figure 21b) arise when the Pacific storm track appears to shift to an anomalous poleward (**HL**) or equatorward (**LH**) location.

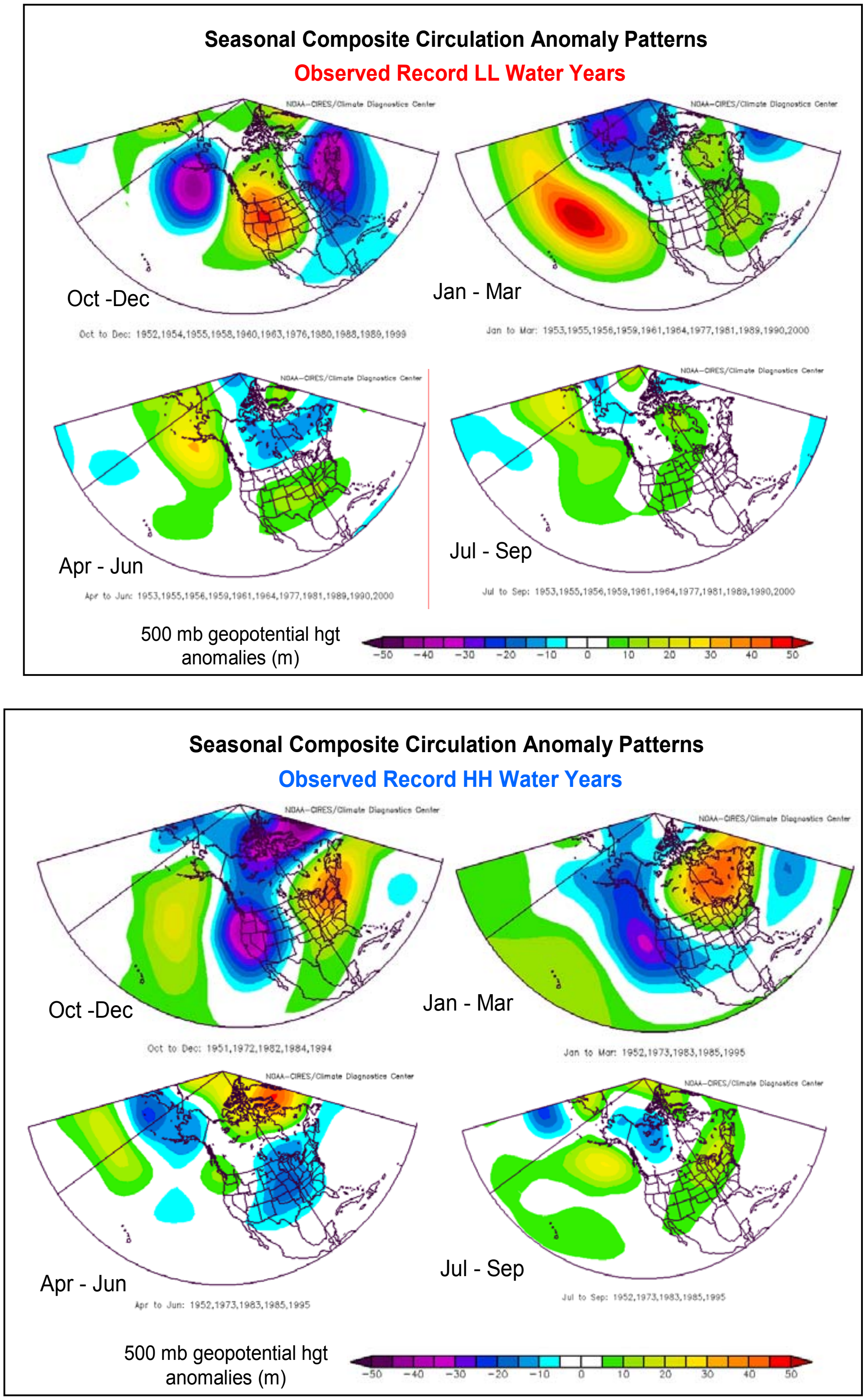
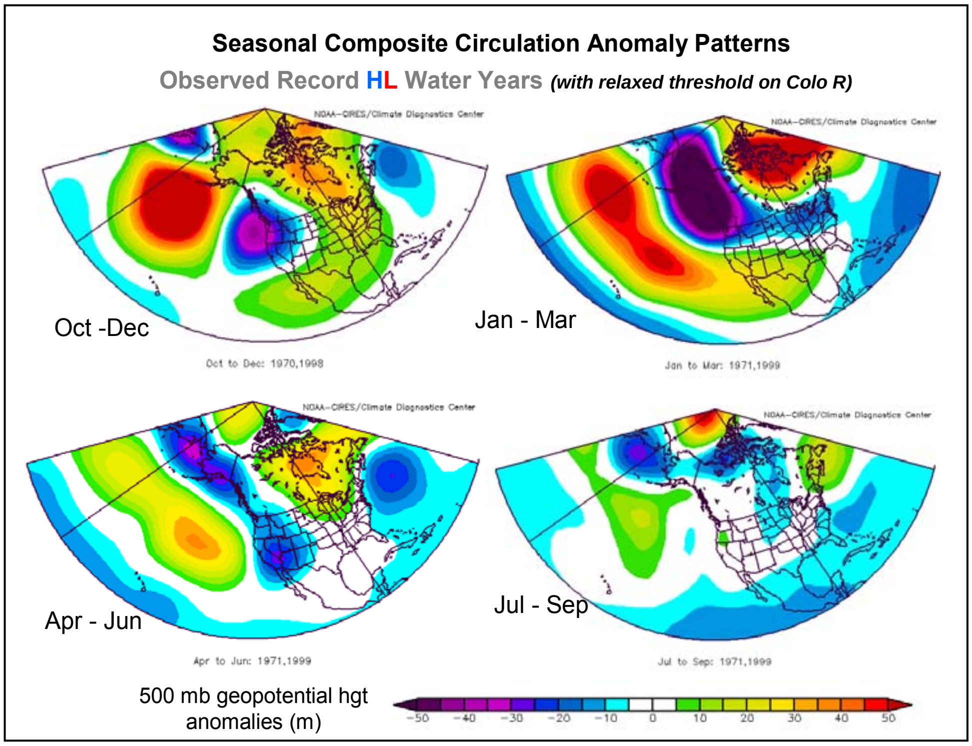


Figure 21a – Composite circulation anomaly patterns for LL and HH years



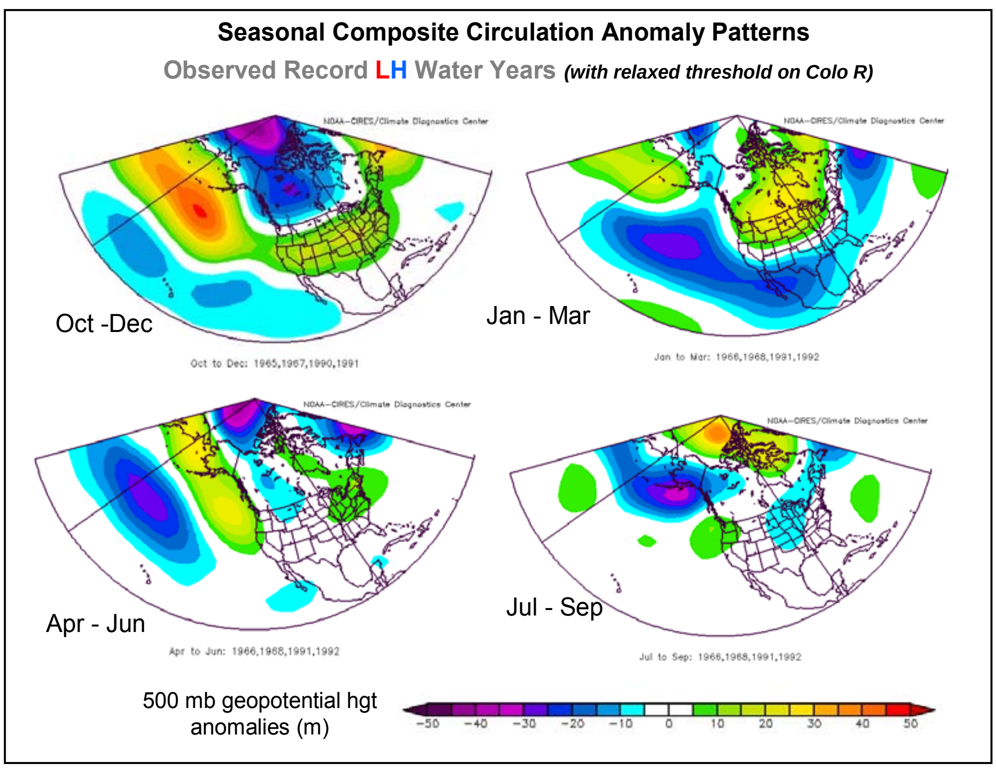


Figure 21b – Composite circulation anomaly patterns for HL and LH years

(5-c) *Are there any linkages of these anomaly patterns to climate-scale driving mechanisms such as sea surface temperature anomalies, El Niño, La Niña, etc.?*

Preliminary examination of El Niño, La Niña influences and ocean indices such as the Pacific Decadal Oscillation (PDO), and the Atlantic Multidecadal Oscillation (AMO) suggest linkage to some – but not all **LL** years in the observed record (**Figure 22**). A University of Arizona M.S. thesis is in preparation to examine this in more detail. A cross-spectral analysis was completed to examine the possibility of low-frequency variability and possible driving mechanisms (see **APPENDIX 6**).

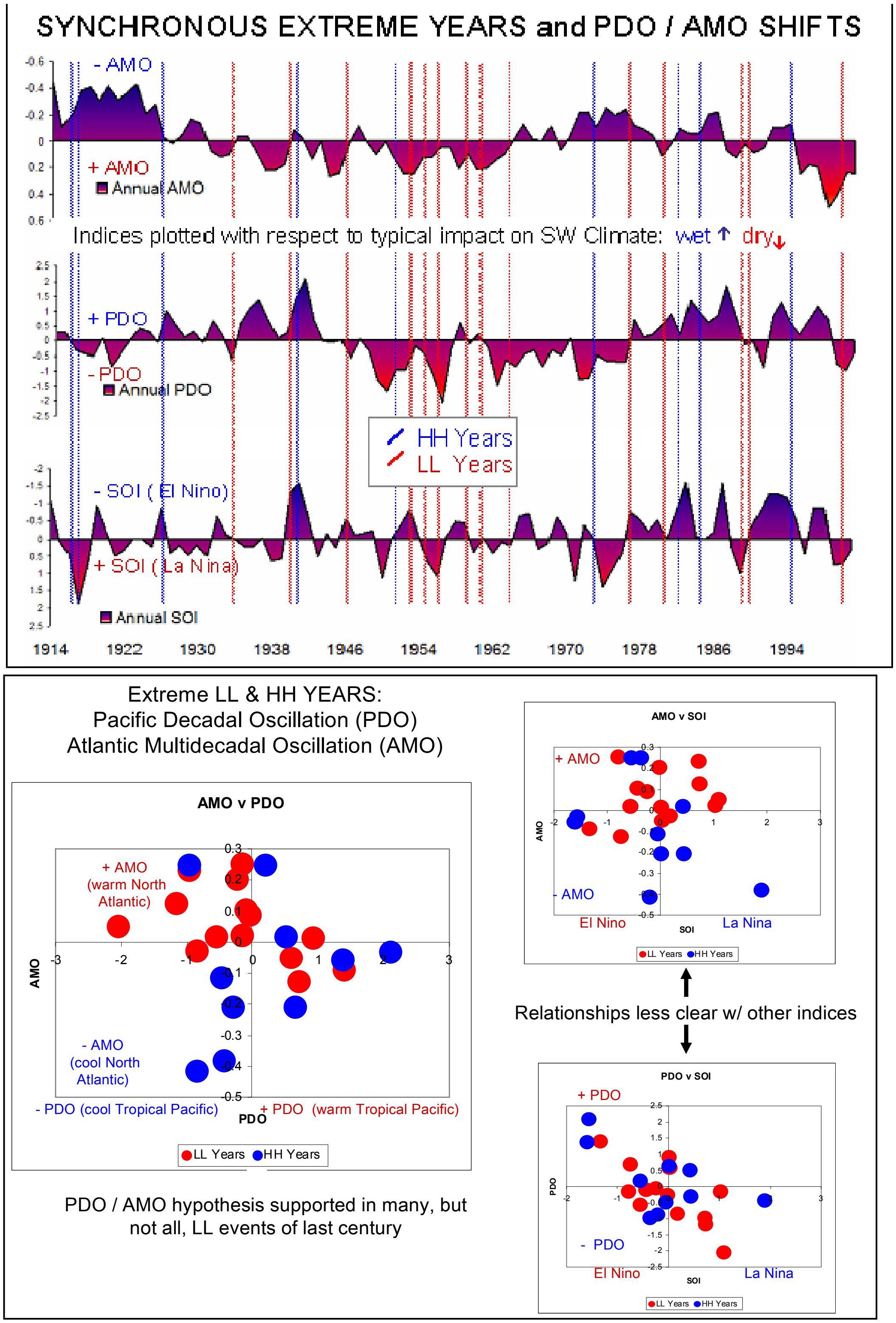


Figure 22 – Extreme years and PDO / AMO / ENSO

(5-d) *Is the recent drought unprecedented in severity?*

Extended periods of below-normal streamflow and low reservoir levels in the late 1990s and early 2000s clearly mark the most recent decade or so as one of the driest in the period of observed streamflow records in the Salt-Verde watershed. The Salt+Verde+Tonto reconstruction, A.D. 1199-1988, allows us to put this recent period into an 800-year context.

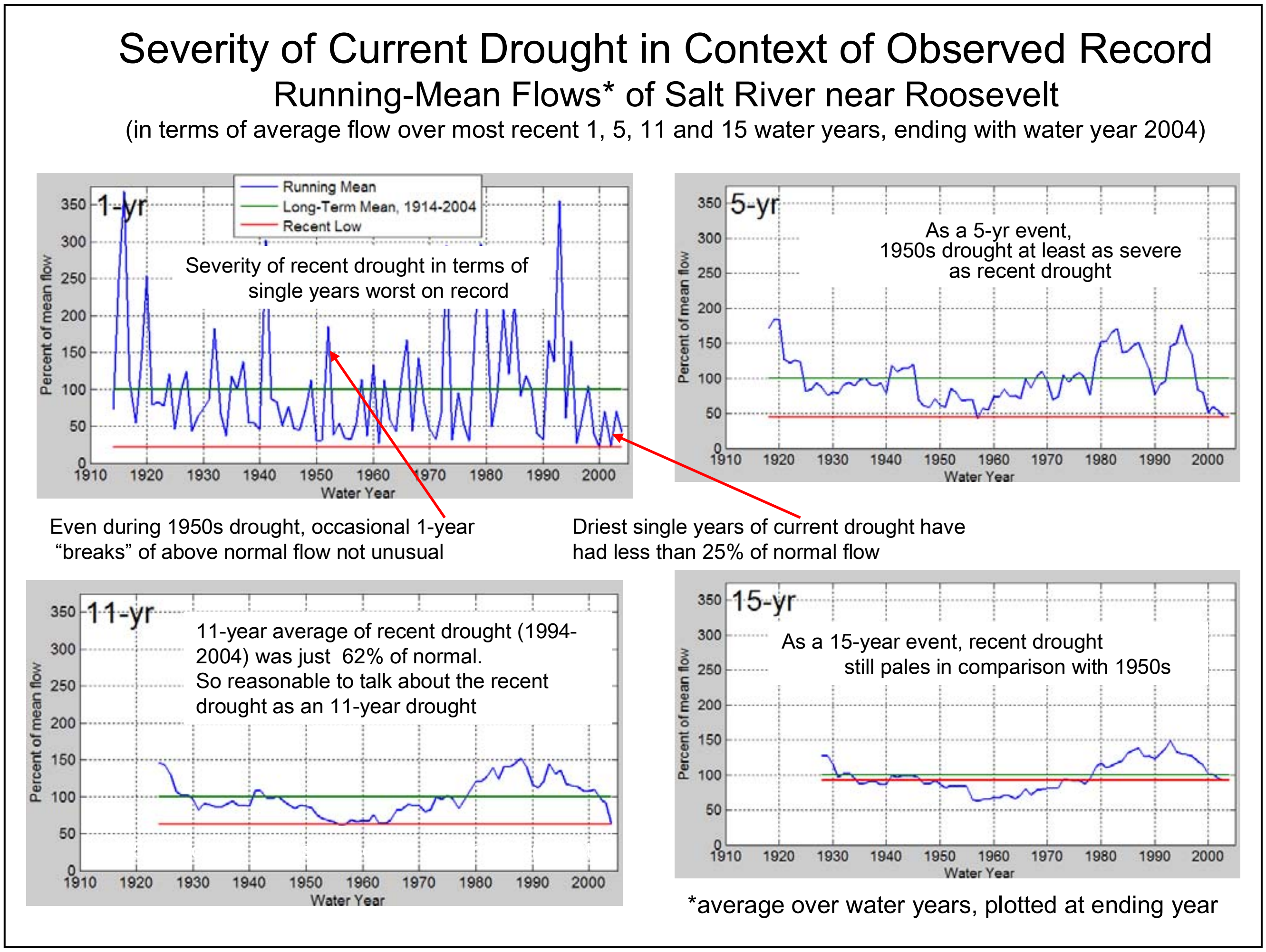
First, it is useful to put the recent drought quantitatively in historical perspective using the observed flow record itself. Data for the gage on the Salt River near Roosevelt, Arizona, cover water years 1914-2004. The annual (water-year) totals for that gage clearly mark the recent drought as exceptional. The annual flows, as well as running means of length 5, 11 and 15 years are plotted in **Figure 23** as percentages of the 1914-2004 mean. The lowest point of each curve in the recent drought is marked with a horizontal line extended back over the length of the full gaged record. This line serves as a baseline for comparing severity of past droughts and the current drought.

The annual flows began an extended dive in water-year 1994 that culminated in single-year flows for 2000 and 2002 lower than any previously experienced in the observed record (**Figure 23a**, top left). The 5-year, 11-year and 15-year running means reached their recent low points in the periods ending with water year 2004. As a 5-year running mean, the recent drought is about as severe as the lowest-flow period in the 1950s. The same is true of the 11-year running mean, suggesting that the period commencing with the decline in water year 1994 and continuing through water year 2004 ranks with the driest conditions in the entire gaged record. When the running mean is extended to 15 years (**Figure 23a**, lower right), the recent drought no longer ranks among the most severe; the reason for this is the wet sequence of years in the early 1990s enters the moving average. But up to an averaging period of 11 years, it appears the recent drought is at least comparable in severity to any earlier drought in the gaged record.

The tree-ring reconstruction for the Salt+Verde+Tonto (SVT) ends in 1988, and so does not cover the recent drought. Nevertheless that reconstruction does sample the 1950s drought, and because the 1950s drought was characterized by flow departure of roughly the same magnitude as the recent drought, we can use the lowest reconstructed flows of the 1950s to indirectly evaluate the relative severity of the recent drought in the context of the reconstruction to A.D. 1199.

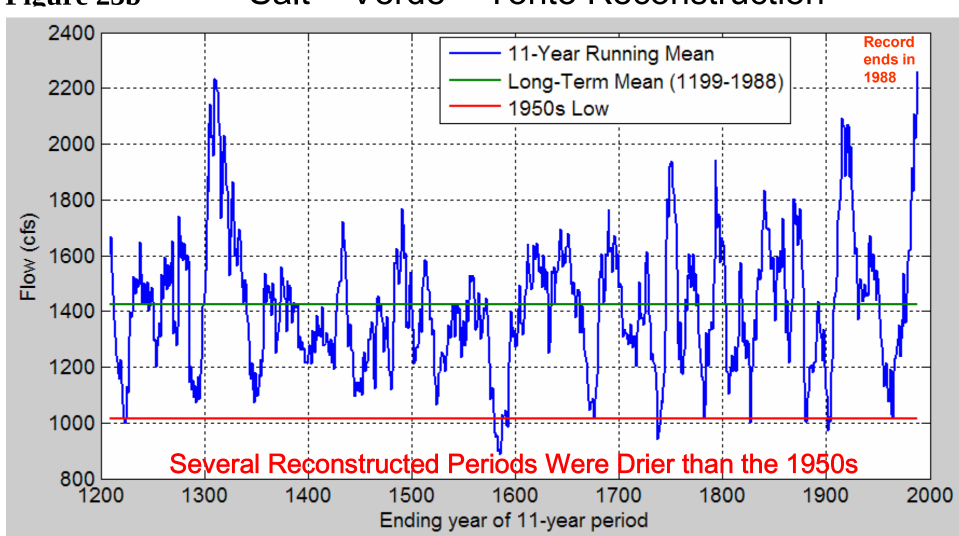
A plot of 11-year running means of the SVT reconstruction with the baseline marked as the low point in the 1950s suggests that the current drought was exceeded in severity several times in the past 800 years (**Figure 23b)**. Eight distinct periods before the start of the gaged record show lower 11-year mean flow than the lowest reconstructed value of the 1950s. The most severe of the tree-ring droughts was in the late 1500s, during the well-documented “mega-drought” of North America (Stahle et al. 2000), when 11-year average flow is reconstructed about 100 cfs below the lowest flows of the 1950s. The recent drought, while severe, does not appear to be unprecedented when viewed in a multi-century context.

Figure 23a – The recent drought in context



Severity of Current Drought in Context of Reconstructed Record:

Figure 23b Salt + Verde + Tonto Reconstruction



-- Current drought was about as severe as 1950s in terms of flows averaged over 11 years

-- 8 other droughts were as severe, according the tree-ring record

-- Late 1500s megadrought was much more severe

6 -- OVERALL PROJECT CONCLUSIONS [6]

- Synchronous extreme events in the same direction (LL and HH events) were much more frequent than LH or HL events.
- Extreme synchronous low flow (LL) and high flow (HH) events tended to cluster in time.
- The longest period of consecutive LL years in the record was 3 years.
- In terms of multi-year extremes, a scenario of 2 extreme yrs within a 4-yr or 3-yr moving window was the most common.
- BOTTOM LINE: Severe droughts and low flow conditions in one basin are unlikely to be offset by abundant streamflow in the other basin.

7 - REFERENCES

Bloomfield P. (2000) *Fourier analysis of time series: an introduction, second edition*. John Wiley & Sons, Inc., New York. 261 p.

Cook E. R. and Kairiukstis L. A., editors (1990) *Methods of dendrochronology: applications in the environmental sciences*. Kluwer Academic Publishers, Boston. 394 p.

Cook E. R., Meko D. M., Stahle D. W. and Cleaveland M. K. (1999) Drought reconstructions for the continental United States. *J. of Climate* **12**, 1145-1162.

Cook E. R., Meko D. M. and Stockton C. W. (1997) A new assessment of possible solar and lunar forcing of the bi-decadal drought rhythm in the western United States. *Journal of Climate* **10**, 1343-1356.

Cook E. R., Woodhouse C., Eakin C. M., Meko D. M. and Stahle D. W. (2004) Long-term aridity changes in the western United States. *Science* **306**, 1015-1018.

Fritts H. C. (1976) *Tree Rings and Climate*. Academic Press, London. 567 pp.

Graybill D. A., Gregory D. A., Funkhouser G. S. and Nials F. L. (in press) Long-term streamflow reconstructions, river channel morphology, and aboriginal irrigation systems along the Salt and Gila Rivers. Prepared for inclusion in: Environmental Change and Human Adaptation in the Ancient Southwest, edited by Jeffrey S. Dean and David E. Doyel.

Mardia K., Kent J. and Bibby J. (1979) *Multivariate Analysis*. Academic Press. 518.

Meko D. M., Cook E. R., Stahle D. W., Stockton C. W. and Hughes M. K. (1993) Spatial patterns of tree-growth anomalies in the United States and southeastern Canada. *J. of Climate* **6**, 1773-1786.

Meko D. M. and Woodhouse C. A. (2005) Tree-ring footprint of joint hydrologic drought in Sacramento and Upper Colorado River Basins, western USA. *Journal of Hydrology* **308**, 196-213.

Mitchell J. M., Jr., Stockton C. W. and Meko D. M. (1979) Evidence of a 22-year rhythym of drought in the western United States related to the Hale Solar Cycle since the 17th Century. B.M. McCormac and T.A. Seliga, eds., Solar-Terrestrial Influences on Weather and Climate, p. 125-143, D. Reidel Publishing Company, Dordrecht, Holland.

Ni F., Cavazos T., Hughes M. K., Comrie A. C. and Funkhouser G. (2002) Cool-season precipitation in the southwestern USA since A.D. 1000: comparison of linear and nonlinear techniques for reconstruction. *International Journal of Climatology* **22**, 1645-1662.

Sedlacek, R. and Siegal, R. (unpublished report) Probability of simultaneous droughts for the Salt and Colorado river systems. (on file with SRP)

Smith L. P. and Stockton C. W. (1981) Reconstructed streamflow for the Salt and Verde Rivers from tree-ring data. *Water Resources Bulletin* **17**(6), 939-947.

Stahle D. W., Cook E. R., Cleaveland M. K., Therrell M. D., Meko D. M., Grissino-Mayer H. D., Watson E. and Luckman B. H. (2000) Tree-ring data document 16th century megadrought over North America. *EOS Transactions* **81**(12), 121-125.

Stockton, C.W., 1975, Long term streamflow records reconstructed from tree rings, University of Arizona Press, Tucson, Arizona, 111 pp.

Stockton, C.W., and Jacoby, G.C., 1976, Long-term surface-water supply and streamflow trends in the Upper Colorado River Basin, Lake Powell Research Project Bulletin No. 18: National Science Foundation, 70 pp.

[6] Note: The final streamflow reconstructions for each gage can be found in **Appendix 7**.