An Assessment of Potential Severe Droughts in the Colorado River Basin[1]

Homa Salehabadi, David G. Tarboton, Bradley Udall, Kevin G. Wheeler, John C. Schmidt

Department of Civil and Environmental Engineering (Salehabadi, Tarboton), Utah State University, Logan, UT, USA; Colorado Water Institute (Udall), Colorado State University, Fort Collins, Colorado, USA; Environmental Change Institute (Wheeler), University of Oxford, Oxford, UK; Department of Watershed Sciences (Schmidt), Utah State University, Logan, UT, USA. (Correspondence to Salehabadi: homa.salehabadi@gmail.com).

Research Impact Statement: Plausible severe drought scenarios developed by resampling from historical and tree-ring reconstructed streamflow indicate a need to rethink Colorado River management to face future droughts.

ABSTRACT: Much has been learned about Colorado River hydrology since the severe sustained drought study in 1995. We summarize our updated understanding of plausible future drought conditions by considering historical flows, tree-ring reconstructions, and climate change. We focus on natural streamflow at Lees Ferry, the primary metric used to quantify the runoff in the Colorado River Basin. We identify drought periods using historical records and tree-ring reconstructed streamflow at Lees Ferry, which we then use to characterize potential future droughts. Resampling from past drought periods generates plausible future conditions to consider during planning. We produced three drought scenarios, each comprising 100 streamflow sequences to be used as input to systems operation and management models. We used analysis of the duration-severity and cumulative deficit relative to the mean natural flow to evaluate droughts and drought simulations and show that the current millennium drought that started in 2000 has an average flow far less than the historical record. However, the flows reconstructed from tree rings or future flows projected from climate models indicate that even more severe droughts are possible. When used as input to the Colorado River Simulation System the drought scenarios developed indicate considerable periods when Lake Powell falls below its hydropower

[1] This is the Final submitted version of the following article. Please refer to the published article for the final version of the manuscript.
Salehabadi, H., D. G. Tarboton, B. Udall, K. G. Wheeler and J. C. Schmidt, (2022), "An Assessment of Potential Severe Droughts in the Colorado River Basin," *JAWRA Journal of the American Water Resources Association*, https://doi.org/10.1111/1752-1688.13061.

penstocks, indicating a need to rethink management and operation of these reservoirs during these critical conditions.

(KEYWORDS: Colorado River Basin; Drought; Streamflow; Stochastic Models; Dendrochronology; CMIP; CRSS.)

INTRODUCTION

The Colorado River (Schmidt et al., 2022) is a critical source of water for the southwestern United States and northwestern Mexico (Figure 1). Earth's warming climate is expected to cause a decline in runoff in the Colorado River Basin, a highly utilized basin where current demand exceeds supplies (Lukas et al., 2020). Long-range planning of the water supply provided by the Colorado River requires an assessment of the impact of potentially extreme future droughts within a changing climate. Any model addressing Colorado River management is ultimately driven by assumptions about the watershed's future hydrology, even though the precise characteristics of that future are unknown. Past data, whether historically observed or reconstructed from tree rings, contain information useful in planning for the future. However, hydroclimate conditions that occur in the future will not precisely match those of the past due to the randomness of nature and ongoing anthropogenic climate change.

The Index Sequential Method, or ISM (Ouarda et al., 1997; Payton, 2020), has been used in several past Colorado River planning studies to generate flow sequences that are input to the Colorado River Simulation System or CRSS. CRSS is Reclamation's primary long-term planning model implemented in RiverWare for the Colorado River Basin (Zagona et al., 2001; USBR, 2012; Payton et al., 2020). The ISM develops a set of future streamflow sequences based on a historical record, with each future sequence replicating a portion of that historical record. Each sequence begins at each year in the record and flows for that sequence are assumed to match the historical period following that starting date. When the end of the historical record is reached, the remaining years for the sequence are determined by inserting or wrapping years from the beginning of the historical record to follow the end of the historical record. By starting with each year in the recorded past, the number of sequences produced is exactly the same as the number of years in the record. In ISM, the magnitudes of the future flows precisely match historical flows, and the year-to-year sequence of flows is unchanged, except for the years around the "wrap". This gives rise to the primary limitation of ISM, that it only allows analysis of past events precisely as they occurred in the past, and therefore does not provide a variety of "statistically plausible" sequences (Prairie et al., 2006). In other words, ISM cannot simulate longer or more intense droughts or wet periods than those that occurred within the historical record. Salehabadi et al. (2020) illustrate in a table (Table 14) and visually (Figure 19) how ISM works.

Stochastic hydrology has evolved as a field to generate sequences, often of streamflow, that are different from, but statistically equivalent to past records. These streamflow sequences serve as diverse inputs to systems planning and operations models, to test their resilience for what may occur in the future (Bras and Rodriguez-Iturbe, 1985; Loucks et al., 2005). In the Colorado River Basin, stochastic methods have been used, over time, to overcome limitations of ISM and provide novel hydrologic scenarios in which longer and more severe sequences of droughts or wet periods can be considered (Tarboton, 1994; 1995; Tarboton et al., 1998; Prairie et al., 2006; 2007; 2008; Barnett and Pierce, 2008).

Today, there is an increasing consensus that gaged flows alone, only a few of which date back to the late 1800s, do not adequately represent the potential range of natural variability that has occurred in the past few centuries and that might occur in the future. Thus, there is a need for

Paleo methods (tree-ring hydrology and other surrogate approaches) that provide estimates of past flows prior to the start of gaging. This was recognized early on by LaRue (1925) who developed Colorado River flow estimates prior to gaged records through use of the historical record of the Great Salt Lake levels. Later, Schulman (1946) successfully applied tree-ring science for the first time to reconstruct streamflow in the Colorado River Basin. Since then, many efforts have been conducted to improve the Colorado River streamflow reconstructions (Stockton and Jacoby, 1976; Michaelsen et al., 1990; Hidalgo et al., 2000; Woodhouse et al., 2006; Meko et al., 2007; 2017). The recent tree-ring reconstructions of the Colorado River flow provide robust information about past hydrology and a more complete picture of the range of variability experienced in the Colorado River prior to what is recorded in the gaged records. These reconstructions reveal the occurrence of a number of megadroughts estimated to have been more severe than any droughts during the gaged period. It is plausible that events similar to these megadroughts estimated from tree-ring hydrology could occur in the future, because they have happened in the past. Meko et al. (2012) showed that events similar to the severe and sustained mid-1100s drought (originally estimated by Meko et al., 2007) may have a frequency of occurrence of once every 400-600 years. The probability of such a drought occurring in the future may be more likely because of the effects of a warming climate. Udall and Overpeck (2017) found that megadroughts in the medieval period, which caused flow reduction of -16%, would, if they were to recur in a warmer future climate, result in even greater flow reduction to -21.5% and -34.5% under a 1°C and 3°C future warming, respectively. Such studies indicate the importance of considering tree-ring reconstructions in future planning.

How climate related changes will affect the Colorado River Basin runoff is another prominent question that should be considered in future planning. There is a consensus among many studies that the future runoff of the Colorado River Basin will decline as it warms, although shorter periods of high flows are also likely to occur (McCabe and Wolock, 2007; Vano et al., 2014; Woodhouse et al., 2016; Udall and Overpeck, 2017; McCabe et al., 2017; Milly and Dunne, 2020; Lukas et al., 2020). Almost all approaches in these studies begin with using climate models (known as General Circulation Models or GCMs) driven by a scenario based on the future emission of greenhouse gases into the atmosphere. Although GCMs are currently good tools for exploring future climate changes on continental and greater scales, these models still have weaknesses in representing some climate features over smaller areas such as the Colorado River Basin. Of particular concern is that GCMs do not adequately capture the frequency of drought and pluvial events that occurred in the past and may underestimate the risk of future megadroughts (Ault et al., 2012; Ault et al., 2013). Udall and Overpeck (2017) reported that half of the CMIP5 models and one-quarter of CMIP3 models used in the projections of U.S. Bureau of Reclamation (hereafter Reclamation) for the Colorado River Basin (USBR, 2014) are unable to simulate the 21st century drought. On the other hand, individual future years projected from Reclamation's 2012 Basin study using the Christensen et al. (2004) methods do include rare annual flows as large as 45 maf/yr, 80% higher than the highest historical year (USBR, 2012). Despite these uncertainties and limitations, which augurs for caution in their direct use for future planning, GCMs do have significant value for decision making. Their strengths include their ability to project future warming with a high level of certainty, and warming is the main driver of shifts in hydroclimate toward lower spring snowpacks, earlier snowmelt, lower annual runoff volumes, and increasing water demand (Lukas et al., 2020).

The so-called Law of the River (MacDonnell et al., 1995; MacDonnell, 2021a; 2021b), has evolved to govern management of the Colorado River and is the subject of much writing and analysis (Megdal, 2022; Kuhn and Fleck, 2019). Broadly, within the U.S., the Colorado River Basin is divided into upper and lower basins at Lee Ferry. Upper basin water is distributed on a percentage basis, while water deliveries to lower basin states and Mexico have a set amount. Currently, there are three elements in the Law of the River used to govern the Colorado River in response to increasing droughts and shortages; the 2007 *Colorado River Interim Guidelines for Lower Basin Shortages and Coordinated Operations of Lake Powell and Lake Mead* (USDOI, 2007), the 2019 *Drought Contingency Plan* (DCP, 2019) agreement among the seven Basin States in the United States, and *Minute 323* of the 1944 Bi-national Water Treaty between the United States and Mexico that have been used to manage water allocations during the current persistent period of drought that started in 2000 (Minute 323, 2017; Treaty, 1944). These agreements will expire in 2026, which, combined with the drought still persisting, is motivating much work towards informing their renegotiation (Wang et al., 2020; Milly and Dunne, 2020; Lukas and Payton, 2020; Salehabadi et al., 2020; USBR, 2020b; Woodhouse et al., 2021; Wheeler et al., 2021; Williams et al., 2022).

The objective of this study was to quantitatively derive a set of drought scenarios that characterize plausible future Colorado River droughts that should be considered in planning. We used historical and tree-ring reconstructed natural flow at Lees Ferry. We developed an analysis methodology to characterize and visualize drought severity based on calculating flow means for a particular duration and cumulative deficits relative to the historical mean flow. From this analysis we identified three past severe sustained droughts: the period between 2000 and 2018, with mean flow of 12.44 million acre feet/year (maf/yr), that we refer to as the millennium or current, ongoing drought; the period between 1953 and 1977, with mean flow of 12.89 maf/yr, referred to as the mid-20th century drought; and the period between 1576 and 1600 from the tree-ring record, with mean flow estimate of 11.76 maf/yr, referred to as the paleo tree-ring drought.

Planning for the future is inherently uncertain, and made more so due to climate change and the resulting lack of stationarity (Milly et al., 2008). Nevertheless, past flows are important in the development of scenarios for future planning. Here we have developed streamflow scenarios that are a resampling of past drought scenarios but for which precise future probabilities are difficult to determine or are lacking in meaning due to climate change and non-stationarity. These scenarios may thus be regarded as having level three uncertainty in the sense of van Dorsser and Walker's uncertainty levels for future planning and policy setting (Walker et al., 2013; van Dorsser et al., 2018). They represent situations in which one is able to enumerate multiple plausible alternatives without being able to quantify precisely how probable they are (van Dorsser et al., 2018). Such an approach was suggested for the Colorado River by Wang et al. (2020). We implemented a resampling scheme that assumed that years of low runoff that occurred in the worst of past droughts might occur again in the future but that the sequence in which these years of low runoff occur in the future might differ from what occurred in the past. Thus, we simulated possible future droughts by randomly selecting flows from the records of the three severe past droughts identified above. Each grouping of randomly-assigned sequences of low-flow years drawn from one of these past droughts is referred to as a scenario. Multiple (100) sequences were simulated for each scenario.

To use the annual drought scenario sequences developed in CRSS, we applied annual block disaggregation, further described in the methodology section, to produce monthly flows at each of the 29 nodes where flow enters CRSS. These data were used in the April 2020 version of CRSS initialized with the projected January 1, 2021 reservoir conditions and the current interpretation of the Law of the River within CRSS to evaluate Colorado River System performance during the next 20 years. The results indicate considerable periods with Lake Powell at a level below its hydropower penstocks (14 to 34 percent of the time across the drought scenarios), a situation that would be critical as power could no longer be generated and the ability to control outflows is diminished. This indicates a need for rethinking the management paradigms for operation of these reservoirs in the face of future droughts.

The paper is organized as follows: It starts with a description of the data we used including historical natural flows, tree-ring reconstructions of streamflow, and flow projections under climate change. Then, the methodology section describes a novel duration-severity analysis visualization approach used to identify drought scenarios, the drought scenario approach used for streamflow simulation, its implementation, the method used to identify performant climate projections, and the CRSS model. Results include the severe droughts identified first in past flows, and then compared to climate projections. Then we show the severity of droughts in the scenarios we developed and the outcomes from using these as input to CRSS. The discussion, which puts the results in context and reviews limitations and assumptions, is followed by conclusions on the plausibility and importance of the drought scenarios we developed.

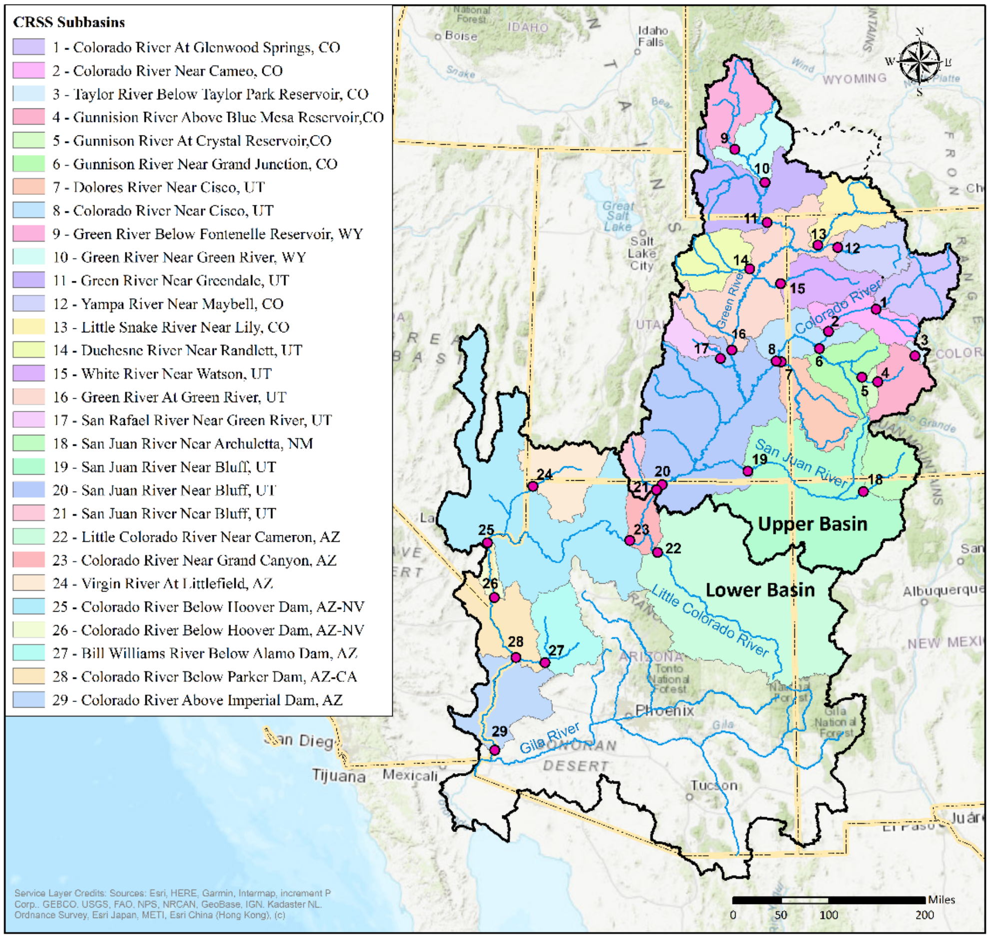


Figure 1. Colorado River Basin with 29 inflow nodes used in the CRSS model and the local-watersheds of each node. Note that the local-watershed of each node only includes the area downstream from any node further upstream (Mapped using ArcGIS from ESRI).

DATA

Natural Flows

Natural flow is the flow that would have occurred in the absence of human activities such as trans-basin diversions, consumptive irrigated agriculture, municipal and industrial uses, and reservoir evaporation. The agencies involved with the Colorado River have, over time, used a variety of methods to estimate natural flow (Prairie and Callejo, 2005; Colorado Water Conservation Board, 2012; Upper Colorado River Commission, 2018). Depending on calculation methods and the extent to which upstream consumptive losses are accounted for, the natural flow estimated by different agencies may differ. In this study, we used Reclamation (USBR, 2020a) estimates of the natural flow at 29 sites in the Colorado River Basin. These data are published and updated as the Colorado River Basin Natural Flow and Salt Data (USBR, 2020a). This database includes the estimated natural flow originating within each local-watershed (referred to as intervening flow in the Reclamation data) as well as the flow coming from further upstream. These data are regularly revised due to source data updates, and the most recent update as of May 2022 has data to the end of the 2019 water year, with provisional estimates for water years 2020 and 2021 (USBR, 2021). Prairie and Callejo (2005) described methods and assumptions used to calculate Reclamation's natural flows since 1971 from the gaged records.

Tree-Ring Reconstructions of Streamflow

There are multiple tree-ring reconstructions available that estimate the streamflow of the Colorado River at Lees Ferry. We examined the latest versions of tree-ring reconstructions of Lees Ferry natural flow (Woodhouse et al., 2006; Meko et al., 2007; 2017) to select the reconstruction to use in this study. Tree-ring reconstructions differ, due to different statistical methods, different calibration periods, and different natural flow estimates used for calibration. There have also been questions raised about bias in the variability, and Robeson et al. (2020) developed a bias-correction approach that was applied to the Meko et al. (2007) reconstruction. We applied this approach to the Meko et al. (2017) reconstruction and evaluated both original and bias-corrected reconstructions using statistical metrics such as Coefficient of Determination (R^2), Nash-Sutcliffe Efficiency (NSE), and Root Mean Square Error (RMSE) to quantify differences and the goodness of fit, between observations and tree-ring reconstruction model results. A further measure of similarity was also obtained from the quantile-quantile relationships between the distribution of the observed and reconstructed values. On the basis of these comparisons we found the model labeled as most skillful in the Meko et al. (2017)study (or M17-SK), covering the period 1416-2015, to be generally best in our evaluation and selected it for use in this study to quantify the severity of droughts back through this time (Figure 2). Readers are referred to Salehabadi et al. (2020) for a detailed analysis of the comparisons that we did between multiple tree-ring reconstructions and natural flow at Lees Ferry.

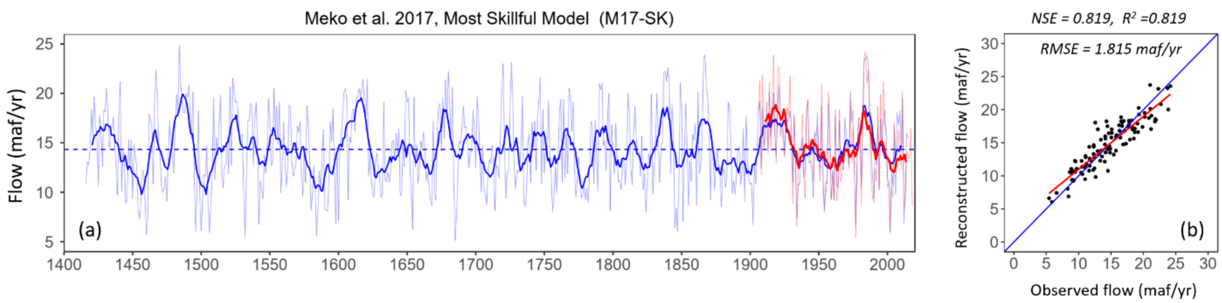


Figure 2. Tree-ring reconstructed flow of the Colorado River at Lees Ferry estimated by the most skillful model of Meko et al. (2017). (a) Annual time series of the reconstruction (light blue line) and 10-year moving average (blue line), along with annual time series of the observed natural flow (light red line) and 10-year moving average (red line). (b) Relationship between observed and reconstructed flow of the Colorado River at Lees Ferry (R^2, Nash Sutcliffe Efficiency (NSE), and RMSE).

Streamflow projections under Climate Change

While the primary focus of this study was on the development of drought scenarios from historical and tree-ring reconstructed streamflow, we used climate model projections of streamflow to place this work in the context of climate change. We used Reclamation's flow projection dataset in which climate model projections from the Coupled Model Intercomparison Project 5 (CMIP5) were linked to the Variable Infiltration Capacity (VIC) hydrology model to compute natural streamflow at Lees Ferry from 1950 to 2099 (USBR, 2014). There are 97 hydrology projections in the CMIP5-VIC simulations available from Reclamation for a range of potential greenhouse gas emission scenarios referred to as Representative Concentration

Pathways or RCPs, specifically RCPs 2.6, 4.5, 6.0, and 8.5. In these simulations, output from the climate models is used as input to VIC to estimate streamflow. VIC is a semi-distributed physically based land surface model that has been used for sensitivity studies and climate change projections of hydrology in the Colorado River Basin (Wood et al., 2020). Given that CMIP models project significant future temperature rises, it is important to assess the temperature sensitivity of the models used to estimate projected streamflow from climate model outputs. Vano et al. (2012; 2014) examined the sensitivity of eight hydrologic models in the Colorado River Basin. Sensitivities ranged from -3% to -14% per degree of temperature change, with those reported for VIC in the middle to the high end of the range (-6% to -10%). Hoerling et al. (2019), on the other hand, suggested lower temperature sensitivities (-2.5% per 1°C). A more recent study by Milly and Dunne (2020) estimated a decrease of -7.8 % to -12.2% (mean -9.3%) per 1°C of warming. Overall, this is still an active research area and the interpretations of VIC deduced streamflow should be interpreted in the context of this uncertainty.

METHODOLOGY

Duration-Severity Analysis

We used the duration-severity approach to analyze the flow data with the aim of identifying drought scenarios based on severity and duration. In this approach, the severity was quantified in terms of the mean flow for a particular duration. To effectively present the results of this analysis, we developed duration-severity plots in which each mean flow for a particular duration (number of years) was depicted as a single point (Figure 3). Given a streamflow record such as the annual natural flow at Lees Ferry (Figure 3-a), we can look for durations with the lowest multi-year average flow (orange points in Figure 3-b). Accordingly, Figure 3-b shows that the one year with lowest flow was 1977 (point A), there was a 5-year period from 2000 to 2004 with mean flow of 9.47 maf/yr (point C), and the minimum 12-year mean flow was 12.13 maf/yr from 1953 to 1964 (point E). We can also compute and depict the second lowest averages of each duration (blue points in Figure 3-b). This shows that, for example, the year with the second lowest flow is 2002 (point B), and the second lowest 5-year mean was from 1988 to 2002 (point D). The periods corresponding to each of these points are indicated in Figure 3-a. Continuing this procedure for the entire flow series leads to a complete duration-severity plot as presented in the results (Figure 5 and on below).

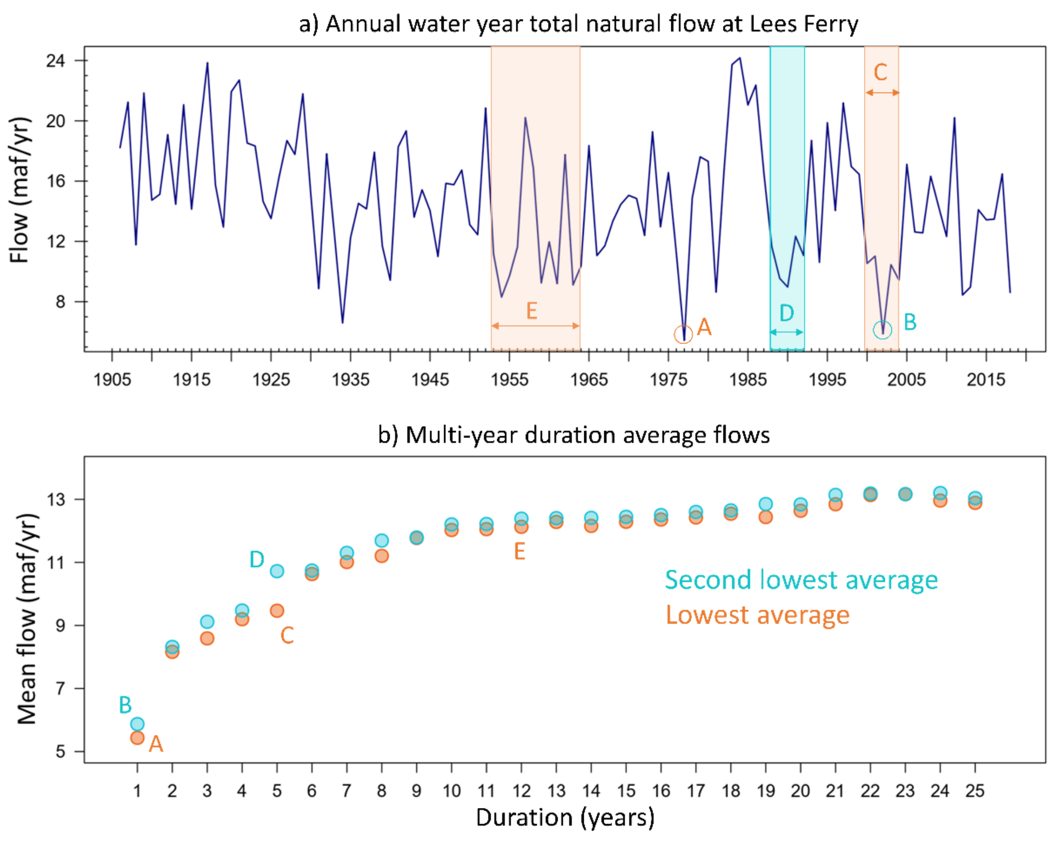


Figure 3. An example to show the procedure of performing the duration-severity analysis. (a) annual natural flow at Lees Ferry, (b) durations with the lowest and second lowest multi-year flow average. Capital letters represent the corresponding durations in the upper and lower figures.

We also examined the cumulative magnitude of departure from average conditions for different durations. The cumulative departure from average conditions, or “cumulative deficit”, of each n-year duration represents the cumulative deficit, in acre feet, during those n years relative to the long-term average natural flow of 14.76 maf/yr for the 1906 to 2018 period. Our use of the long-term average for 1906-2018 as a reference, was selected, because it is a value that is commonly understood by water-supply managers. Cumulative deficit analysis enables interpretation in terms of what the total deficit over each duration is, relative to the mean.

Duration-severity analyses were used to not only present the tradeoff between drought severity and duration in the flow data, but also to provide an approximate measure of the probability of the recurrence of droughts with different severity under the assumption that past records serve as a basis for probability estimation, while noting that such assumptions do not hold for a changing climate. These analyses helped compare between alternative drought scenarios and also compare these scenarios with either historical or paleo reconstructed natural flows in a way that useful information can be discerned at a glance using these plots.

Streamflow Simulation

Using multiple sequences of streamflow allows consideration of, and planning for, variability in the future hydrology, and is needed to provide a diversity of inputs against which alternative management strategies may be evaluated. We used drought scenario resampling of flows at Lees Ferry to provide plausible simulations for future drought scenarios. We resampled the data that comprise the identified drought scenarios to provide a range of sequences in which the years of low runoff of a drought might occur. In this drought scenario resampling approach, we constructed 100 traces each 42 years long, so as to provide data to use in CRSS to project up to 42 years, spanning, in our case, a 2019 to 2060 mid-century modeling time frame. The 100 traces were obtained using random sampling with replacement drawing only from years in a drought scenario to provide a very stringent stress test on the system, but which, by drawing upon yearly flow values that have all occurred in the past, remains grounded in reality.

An immediate concern with resampling with replacement is that it does not preserve correlation. The annual year-to-year lag 1 correlation of flows at Lees Ferry is 0.24, which being based on L=113 years of data exceeds the threshold $1.96/\sqrt{L} = 0.184$ for statistical significance with 95% confidence. However, resampling that is restricted to be from a drought scenario does result in persistence of the drought scenario, which is the basic property that correlation quantifies. Recognizing the many stochastic hydrology approaches that have been developed to preserve correlation, among other statistics, we evaluated a number of models in addition to random resampling with replacement, and have included our findings as supplementary online supporting information (section 1). This information reports, for each of the three drought scenarios considered, as well as for the entire historical and entire tree-ring records, results from block resampling with lengths of 1, 2 and 5 years, and an auto-regressive order 1 (AR 1) model with correlation selected from the respective full historical record (i.e. observed flow correlation for observed drought scenarios and tree-ring correlation for tree-ring scenarios). A block length of 1 is uncorrelated resampling. The supplementary material findings indicate that uncorrelated resampling is most defensible among the methods evaluated for the purposes of scenario-based resampling.

Note that this uncorrelated scenario-based resampling approach does not select flows from the full record nor strive to reproduce statistics from the full record. Rather it is intended as a way to produce and focus on plausible difficult situations to come that should be considered in planning. This perspective is thus more one of multiple plausible alternatives in support of robust and resilient decision making than one of precise probabilistic risk quantification (van Dorsser et al., 2018). The sequences produced have the statistics of the drought period selected. By being resampled from past data, these are plausible scenarios in the context of a multiple plausible alternatives for planning. Note also that our drought scenarios do include years of above average flow, reflecting that in the past high flow years have occurred within a severe sustained low flow period. The scenarios derived from resampling will also have these occasional high flow years reflecting that short periods with high flow do not necessarily take the system out of a critical condition due to persistent drought.

We used duration-severity analysis to identify the minimum mean flow for each duration, and by considering these across all traces in a scenario determined percentiles of duration-severity variability. We superimposed these on duration-severity and cumulative deficit plots from observed and tree-ring reconstruction records to provide a measure of the historical probability and indicate the plausibility of these scenarios.

To use our drought scenarios as inputs to CRSS we used a nonparametric resampling approach referred to as "water year block disaggregation" to disaggregate annual flow at Lees Ferry into monthly flow at each of the 29 CRSS natural inflow sites. This approach has its roots in other block bootstrap approaches that have been applied in hydrology (Vogel and Shallcross, 1995; Srinivas and Srinivasan, 2005; 2006). Nowak et al. (2010) used a similar approach to disaggregate annual to daily flows. For the historical droughts, given a Lees Ferry flow resampled from a drought scenario, a block of flows from the historical record corresponding to the year of the Lees Ferry flow is used to provide monthly flows across the 29 sites (Figure 4). For tree-ring droughts, there is no corresponding year in the historical record. To address this, for each resampled water year, the nearest observed water year of natural flow at Lees Ferry to the tree-ring reconstructed flow was chosen as the "parallel" year. The corresponding blocks to these parallel years were selected to do the temporal and spatial disaggregation. To preserve the consistency between the generated flows, the ratio of flow between Lees Ferry tree-ring water year and the nearest historical water year was used to adjust the entire block of 348 values. This is a non-parametric approach, in that no distributional or correlation model assumptions are needed.

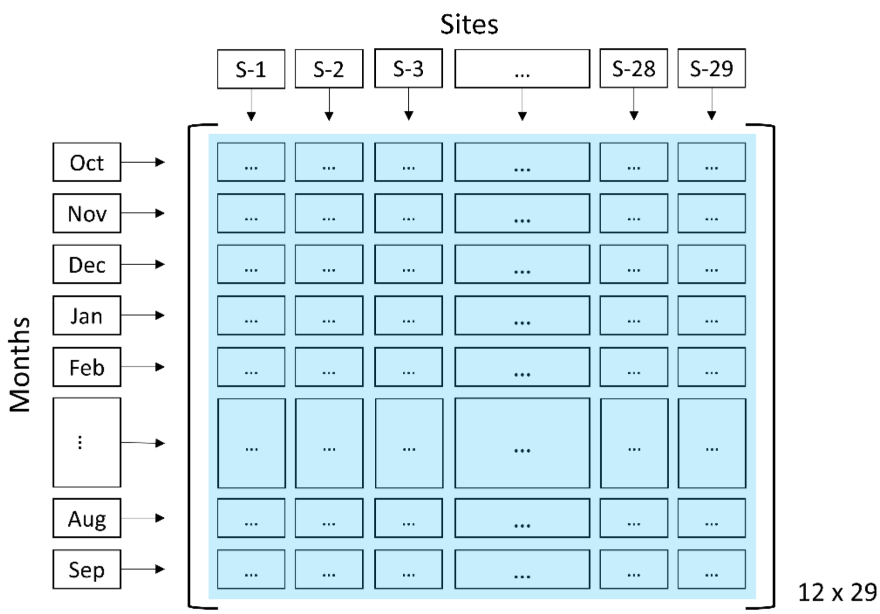


Figure 4. Water year block disaggregation method with each block comprising 348 flow values (12 months × 29 sites)

Implementation Algorithm

Let $X = x_1, x_2, \ldots, x_n$ denote the full record of historical annual flows of the reference site (Lees Ferry), $D=d_1, d_2, \ldots, d_k$ denote the annual flows of one of the past droughts identified using the duration-severity analysis, which may be historical or tree-ring reconstructed, and s denote the number of sites for spatial disaggregation. The steps to generate an ensemble of monthly flows across s sites within the basin are:

1- One of the annual flows of D is selected based on random resampling with replacement (d_i).

2- The year with the equal or nearest flow to d_i is selected as the parallel year (x_j) from the full historical annual flow record.

3- An adjustment ratio is calculated by dividing the parallel year annual flow (x_j) by the resampled year annual flow (d_i), as $C = d_i/x_j$. If the resampled year is available in the historical record (i.e., $d_i = x_j$), the adjustment ratio is 1.

4- A block of monthly flows of s sites from the historical record corresponding to the parallel year (x_j) is selected as matrix B with 12 × s dimensions (12 months and s sites). Flow values within this matrix are multiplied by the adjustment ratio C to provide monthly flows across s sites.

5- Steps 1 to 4 are repeated for each simulation year to generate an ensemble of monthly flows across s sites.

By construction, this approach preserves spatial and temporal correlations within water years, because it is a resampling of the data. Across water years, correlations may not be exactly preserved because of the year-to-year transitions. Additionally, this approach does not introduce spatial and within year temporal variability into the simulations.

Identifying drought performant streamflow projections

Previous studies have used a variety of methods to evaluate and screen future climate projections for the western U.S. (Brekke et al., 2008; Pierce et al., 2009; Mote et al., 2011; Ahmadalipour et al., 2015; Rupp et al., 2017). Prior work has noted that GCMs may not adequately capture the frequency of drought (Ault et al., 2012; 2013). Their internal variability may be different from climate system observations. To address this limitation we evaluated how well each of the

CMIP5-VIC streamflow projections represents past droughts. While it is recognized that climate models cannot be expected to reproduce individual extreme events (Nasrollahi et al., 2015), a measure of how good they are at modeling droughts is how well the projections reproduce the statistical distribution of historical droughts quantified using the duration-severity analysis method presented here. If a CMIP5-VIC model produces droughts during the historical period that differ significantly, in a statistical sense, from what was observed, the internal variability of that model may be different from that of the actual climate system and the ability of that model to project future droughts may be questioned. There were 97 CMIP5-VIC streamflow projections that started in 1950 and ran until 2099. These projections, in general, strive to replicate radiative forcing for the historical period from 1950 to about 2011 (IPCC, 2014) then project forward using representative concentration pathways or RCPs for the future (up to 2099). Thus, statistically these models should represent droughts in the historical period that they simulate. We used duration-severity analysis comparisons to select 10 model projections that best reproduced the severity of the observed record of droughts during the overlapping historical period (1950-2018) based on streamflow at Lees Ferry. Duration-severity analysis, introduced above, is a statistical approach that quantifies drought and its persistence over a range of durations. In this analysis we compared the lowest multiyear mean flow versus duration from historical and CMIP-derived streamflow at Lees Ferry and computed the mean square difference across drought durations ranging from 1 to 25 years. The 10 projections selected were those for which this mean square difference was smallest. The models selected thus statistically reproduce historical droughts across a range of durations over the common historical data period, when external radiative forcing that is a driving input to CMIP models is consistent with history, and thus may be regarded as having greater credibility for projecting future drought statistics.

Colorado River Simulation System (CRSS)

As the primary long term planning model for the Colorado River Basin, CRSS incorporates key components of the Colorado River's channel network such as the river's main stem, major tributaries, intervening flows between gages, diversions, and reservoirs, along with the operational rules such as lower basin drought contingency plan and upper basin drought response operation presently implemented as the Law of the River (Wheeler et al., 2019; Zagona et al., 2001; USBR, 2010). The model identifies 20 inflow nodes in the Upper Colorado River Basin, where natural flow is considered to enter the drainage network (i.e. Colorado River at Lees Ferry), and nine inflow nodes in the Lower Colorado River Basin (i.e. downstream from Lees Ferry). Collectively, these nodes divide the Colorado River Basin into 29 distinct local watersheds defined by the area that drains directly to each node excluding area that drains to a node further upstream (Figure 1).

RESULTS

Severe Droughts in the Colorado River Basin

We analyzed the annual historical and tree-ring-reconstructed (M17-SK) natural flows of the Colorado River at Lees Ferry in order to quantify the severity of past droughts using duration-severity and cumulative deficit analyses (Figure 5 to Figure 7). The severity of the 21st century drought was notable in the historical record (Figure 5 and Figure 6) where the first or second most severe droughts for all durations up to 19 years started during the 19 years since 2000. In these plots, red dots represent the averaging periods that start post-2000 associated with the 21st century drought. Comparison with the other data on these plots demonstrates that the post-2000 average flow is far less than that of the entire 1906-2018 period. Note also that there were only 19 years of data post 2000, thus the red dots ended beyond the duration of 19. However, the lowest 20- and 21-year durations, begin in 1999 and 1998 respectively, and, while not red dots in our labeling, are comprised predominantly of 21st century flows. The current 21st century drought was thus the worst in the historical record when considering flows averaged over longer than the 19 years of post-2000 data we used. We refer to the 21st century drought as the “millennium drought”. The characteristics of the millennium drought scenario were computed from the historical record for the 2000-2018 period as mean flow of 12.44 maf/yr and a cumulative deficit of 44.08 maf. Three years of additional data, two of them provisional natural flow estimates from Reclamation indicate that the millennium drought has persisted into 2021 (and perhaps 2022) with a 1906-2021 average of 14.67 maf/yr, 2000-2021 average of 12.3 maf/yr, and a cumulative deficit of 54 maf.

Examination of Figure 5 reveals that the most severe 25-year drought occurred between 1953-1977. During this sustained drought, the average flow at Lees Ferry was 12.89 maf/yr, which was 87% of the long-term average from 1906 to 2018. For purposes of defining scenarios, we defined the “mid-20th century drought” scenario as the 25-year drought that occurred between 1953-1977.

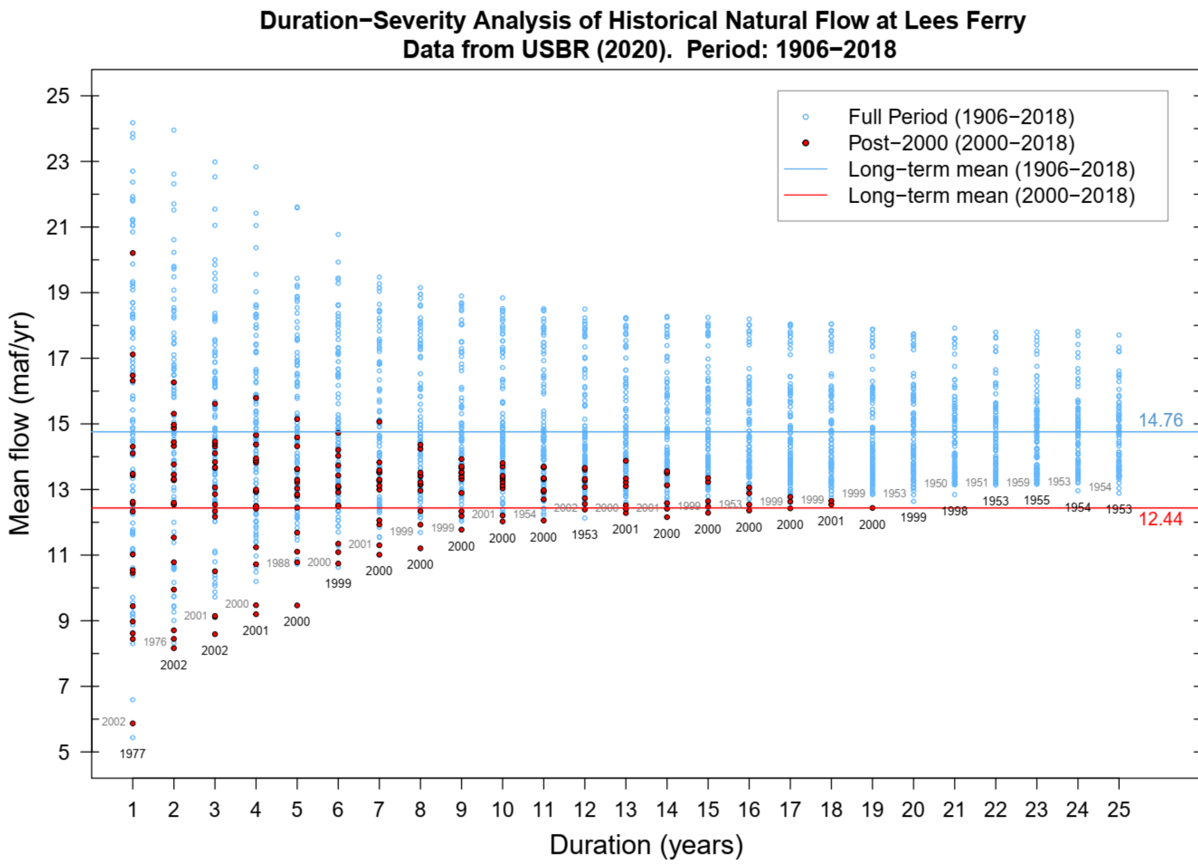


Figure 5. Duration-severity analysis of the historical natural flow of the Colorado River at Lees Ferry. Each dot represents water year mean annual flow averaged over the duration. There is a dot for each duration (including overlaps) within the record. Dot labels give the start year of the lowest (black number) and second lowest (gray number) duration average. The spread of the dots for each duration characterizes how mean flow may vary for different durations.

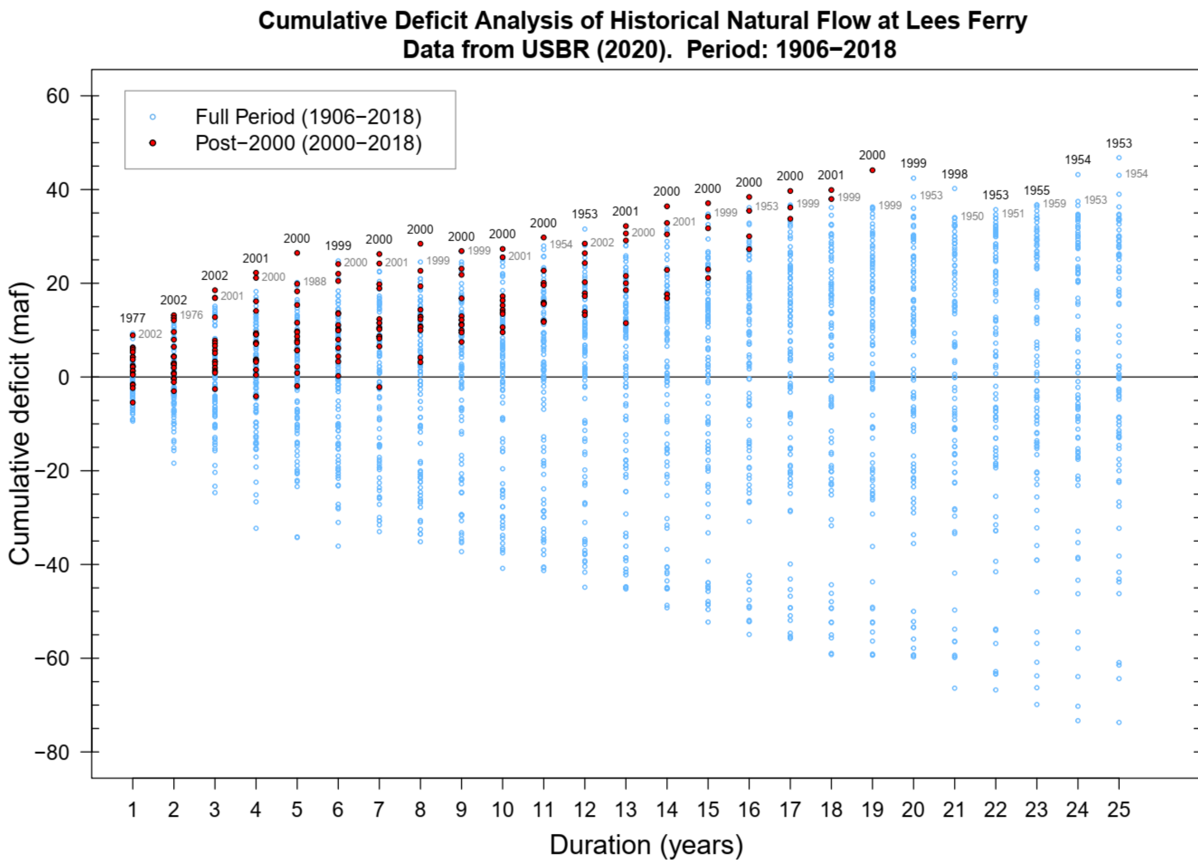


Figure 6. Cumulative deficit analysis of the historical natural flow of the Colorado River at Lees Ferry. Each dot represents water year mean annual flow deficit with respect to the 1906-2018 average of 14.76 maf/yr aggregated over the duration on the x-axis. There is a dot for each duration (including overlaps) within the record. Dot labels give the start year of the highest (black number) and second highest (gray number) cumulative deficit for each duration. The spread of the dots for each duration characterizes how cumulative deficit may vary for different durations.

In the M17-SK reconstructed flow based on tree-ring estimation methods, the most severe and sustained (25-year) drought occurred during the 1576-1600 period (Figure 7). A cumulative deficit plot similar to Figure 6 for the tree-ring record is included as Figure S-18 in the supplementary material. During 1576-1600 the average tree-ring reconstructed flow at Lees Ferry (11.76 maf/yr) was 82% of the reconstruction long-term average. We defined this drought as the “paleo tree-ring drought” scenario. Therefore, overall, three drought scenarios were defined (Table 1).

Note that while we identified the 1576-1600 period and used it in our study, examination of Figure 2 (10-year moving average) shows multiple severe droughts only slightly less severe than this period. This suggests that extreme droughts in the Colorado River Basin occur naturally, and, at multi-century time scales not infrequently.

It is worth noting here that these scenarios identified as 19 and 25 year periods of sustained low flows do include years of above average flow. In particular the millennium drought includes one year (2011) with natural flow 20 maf/yr and four years with flow greater that 15 maf/yr.

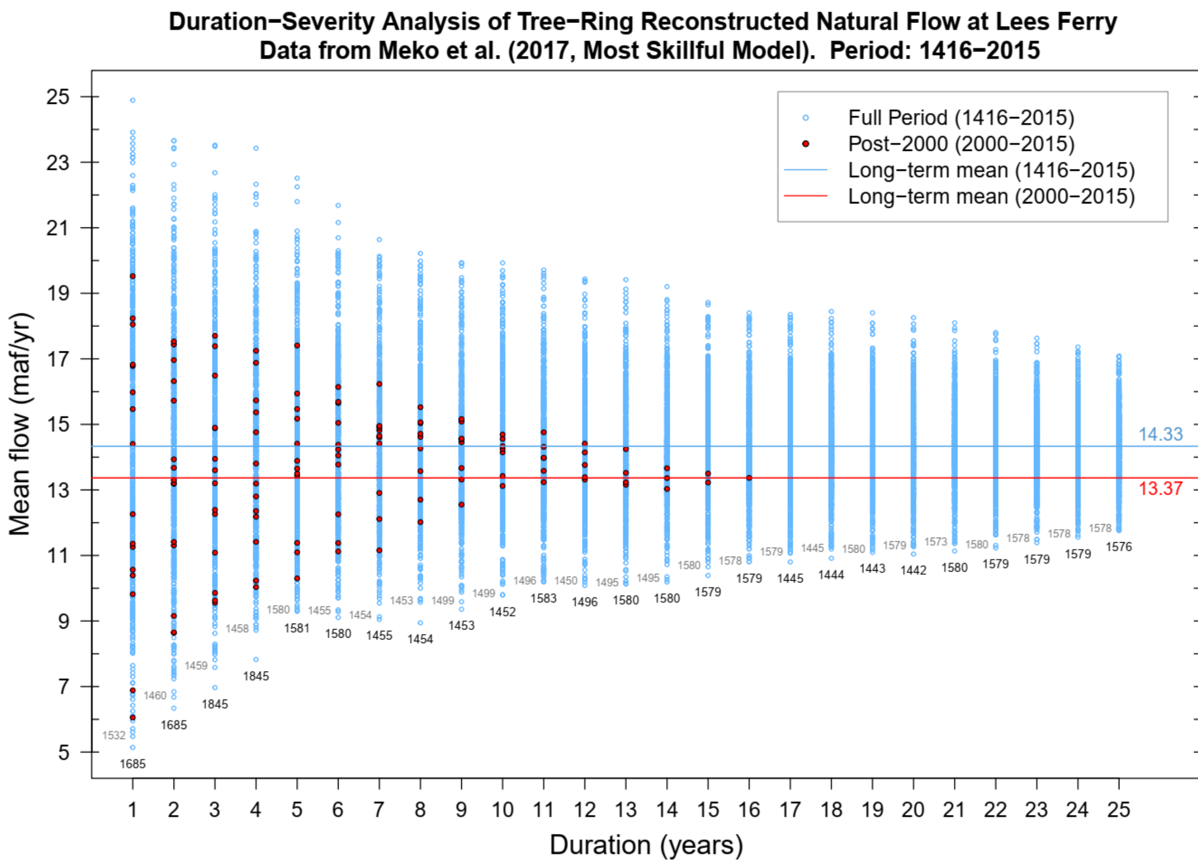


Figure 7. Duration-severity analysis of the tree-ring reconstructed flow (Meko et al., 2017) of the Colorado River at Lees Ferry. See Figure 5 caption for further details.

Table 1. Identified drought scenarios based on natural and tree-ring reconstructed flows

Scenario	Flow data	Period	Duration (years)	Mean flow (maf/yr)	Cumulative deficit (maf)
Millennium drought	Observed natural flow (USBR, 2020a)	2000-2018	19	12.44	44.08
Mid-20th century drought	Observed natural flow (USBR, 2020a)	1953-1977	25	12.89	46.75
Paleo tree-ring drought	Tree ring reconstructed flow (Meko et al., 2017)	1576-1600	25	11.76	75

Severity of Droughts in Climate Change-Informed Streamflow Projections

The minimum duration-severity for each of the 97 CMIP5-VIC projections compared to the minimum duration-severity for observed and tree-ring reconstructed flows indicates climate projections that produce droughts more severe, less severe and similar to historical droughts over the 1950 to 2018 period for which the past climate and hydrology are known (Figure 8). In this figure the minimum mean flow corresponding to each averaging duration within the 69-year CMIP5-VIC run was plotted for all 97 models (gray lines). The corresponding minimum mean flows for the observed and tree-ring reconstructed natural flows for the same period were also plotted (red and blue lines). The 10 performant projections with smallest mean square error between the minimum duration-severity values of the observed natural flow and CMIP5-VIC projected flows were identified and colored light blue in Figure 8. Note that the minimum duration-severity values were compared during the overlapping period (1950-2018) of the observed natural flow and CMIP5-VIC flow projections.

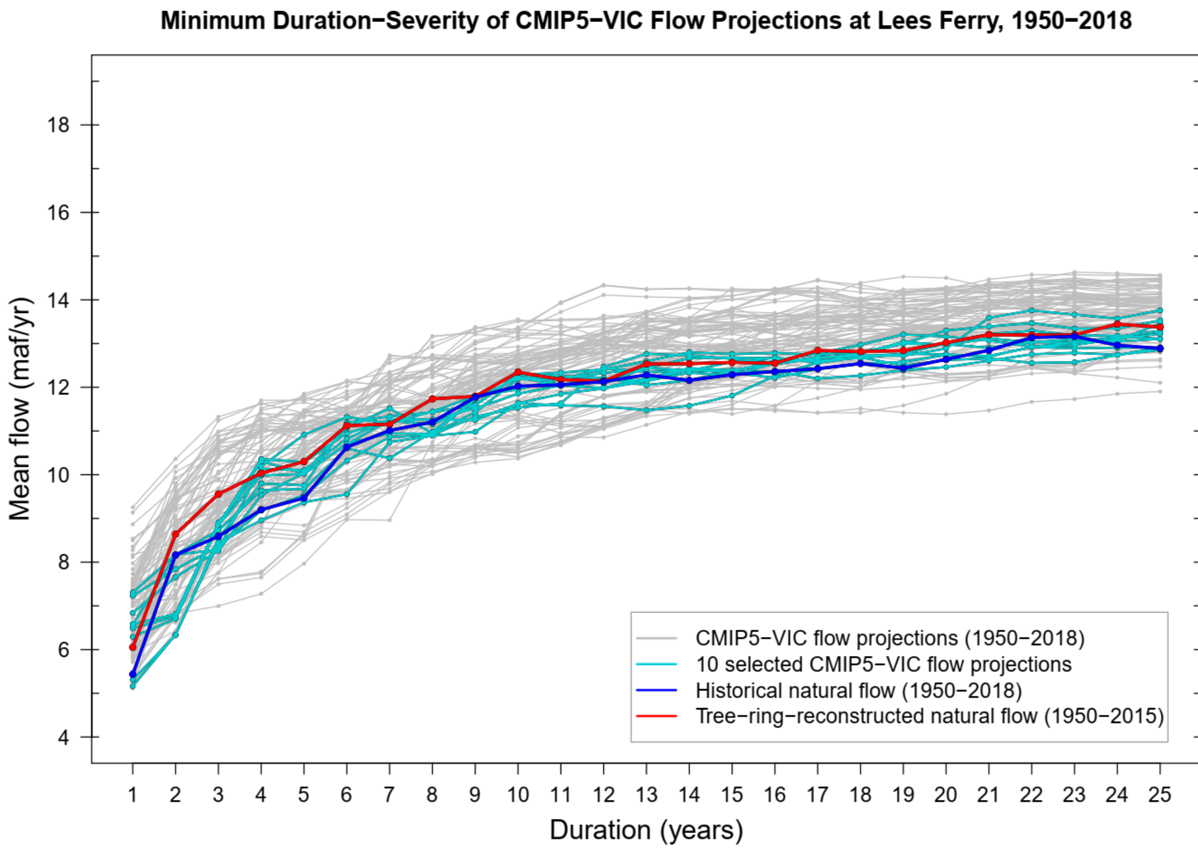


Figure 8. Minimum duration-severities of the CMIP5-VIC projected streamflow at Lees Ferry during 1950-2018 compared to the historical duration-severities and 10 selected simulations that best represent Lees Ferry sustained droughts.

Potential extreme droughts for the full set of CMIP5-VIC projections, and 10 selected projections were examined for the period 2020-2099 (Figure 9). Very extreme droughts with flows as low as 9 maf/yr for durations of 25 years are projected in the future for some of the CMIP5-VIC projections. When we limit analysis to only those 10 selected CMIP5-VIC projections that do a reasonable job of representing the historical droughts (Figure 9), the duration-severity values for six of the 10 models show drier conditions in the future than the past represented by observed natural flow for most durations across the one- to 25-year range. Note that Figure 9 also depicts the duration-severity lower bound for the tree-ring record.

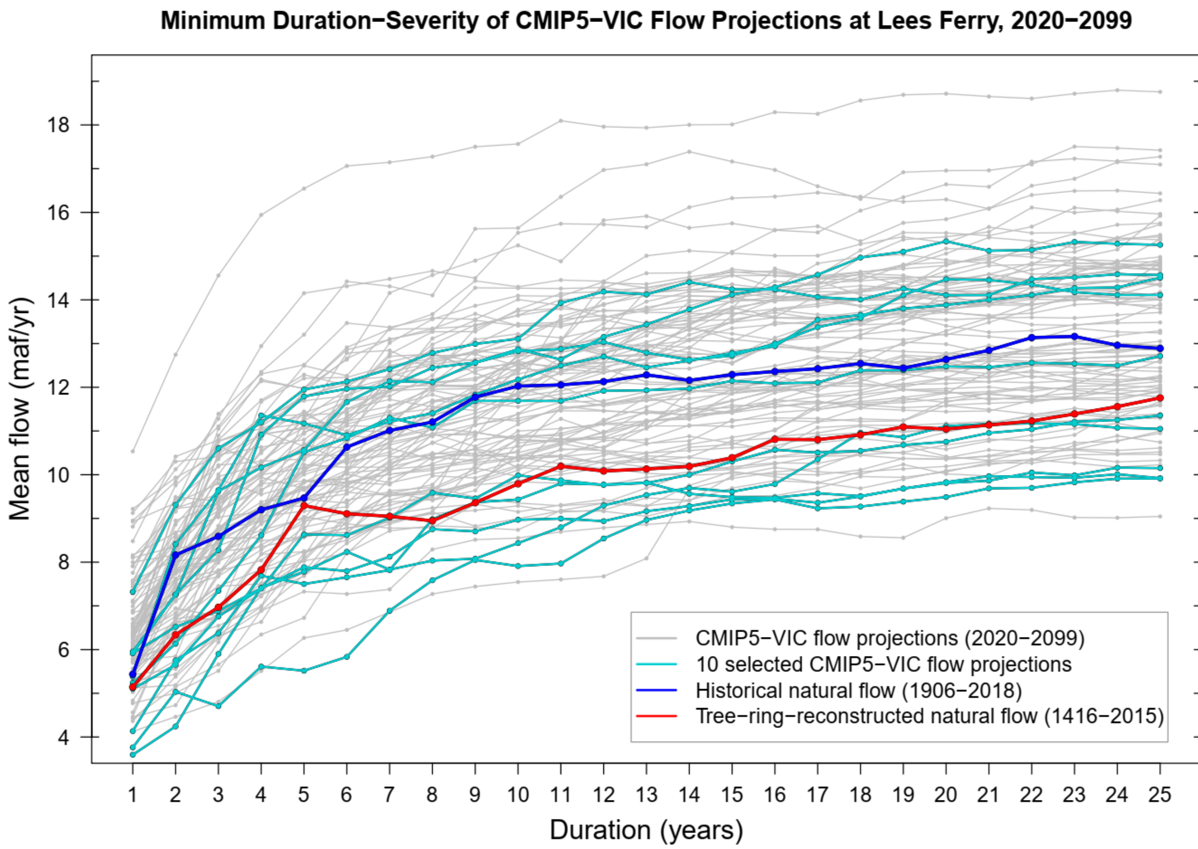


Figure 9. Minimum duration-severities of the CMIP5-VIC projected streamflow at Lees Ferry over 2020-2099 compared to the historical and tree-ring duration-severities and 10 selected projections that best represent Lees Ferry sustained droughts

Streamflow Simulation

We used the methods described above to generate 100 streamflow traces for each of the three drought scenarios (millennium, mid-20th century, paleo tree-ring). Mean and standard deviation of monthly and annual streamflow simulations are included in the supplementary material (Figure S-19 to Figure S-24). Each trace comprised 42 years of monthly streamflow for 29 sites in the Colorado River Basin. To compare the severity of the drought scenarios, we considered the lowest duration-severity at Lees Ferry from each of the 100 traces. The lowest duration-severity for a drought scenario shows the variability of the minimum of the mean flow for different durations based on each of the 100 simulated flow traces. To place the severity of the drought scenarios in a historical context, the range of 10^{th} to 90^{th} percentiles of the lowest duration-severity of each scenario (blue, yellow, and orange areas in Figure 10) was positioned on the duration-severity plots of the observed and tree-ring reconstructed flows (dots in Figure 10). This shows where the range of extreme cases for each of the 100 traces for each scenario falls, with respect to past flows. For example, in the millennium drought scenario (yellow area in Figure 10), five-year flow durations have 10^{th} percentile and 90^{th} percentile mean flows of 8.78 and 10.73 maf/yr respectively. The recent five-year drought we observed in the early 21^{st} century with mean flow of 9.47 maf/yr over 2000-2004 (shown earlier in Figure 5) falls near the middle of this range. In the paleo tree-ring drought scenario, the simulated flows represent even more severe droughts over various durations than those of the millennium drought scenario while the simulated flows in mid-20^{th} century drought scenario represent similar or slightly less severe droughts. Overall, the 10^{th} to 90^{th} percentile ranges of drought scenarios are typically consistent with what has previously occurred. These scenarios are thus consistent with the idea that if it has happened in the past, it can happen again and should be planned for.

We also overlaid the 10^{th} to 90^{th} percentile range of minimum duration-severity of 10 selected CMIP5-VIC projections on the duration-severity plots of the past flows (gray area in Figure 10). Note that with just 10 CMIP5-VIC projections this is the range of the 8 middle models from those shown in blue in Figure 9. The spread of climate projection traces is wider than from our resampled scenarios, and they do extend somewhat lower in the duration-severity plot space than the proposed drought scenarios. CMIP5-VIC streamflow projections are thus indicating possibly more severe conditions than the resampled scenarios.

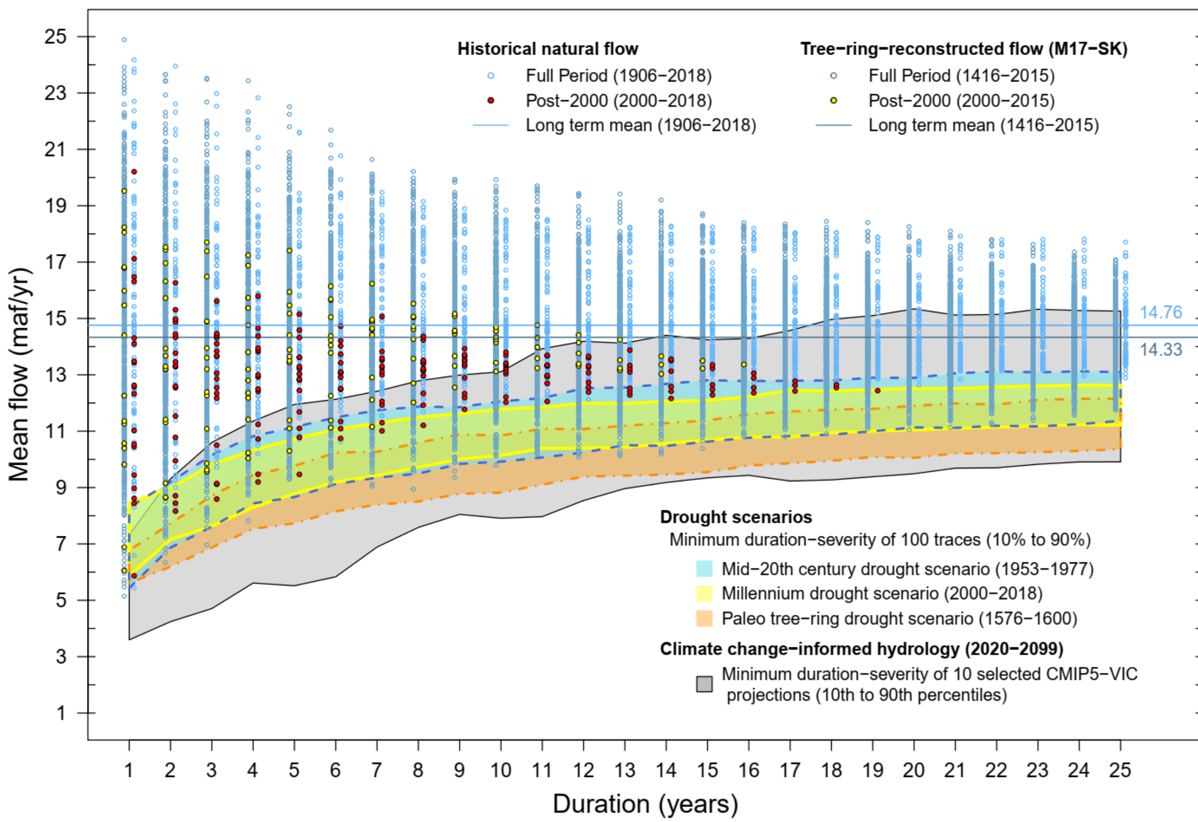


Figure 10. Drought scenario duration-severity set in the context of historical and tree-ring reconstructed mean streamflow versus duration. The 10th percentile to 90th percentile range of lowest drought scenario duration-severity and 10th to 90th percentile range of 10 selected CMIP5-VIC projections juxtaposed over full record of historical and tree ring duration-severities. Darker blue dots from tree-ring record set to the left of historical record dots to keep apart.

The above duration-severity analysis was recast as cumulative deficit relative to a reference flow over the duration. In this case, the highest cumulative deficit range for a drought scenario illustrates the variability of the maximum of the cumulative deficit (relative to the mean flow of 14.76 maf/yr from 1906 to 2018) for different durations based on each of the 100 simulated flow traces (Figure 11). The position of scenario extreme deficits relative to historical and tree-ring reconstruction deficits can be used to quantify the historical probability of the recurrence of droughts with this severity, noting that precise probability estimates under the non-stationarities associated with climate change become infeasible. This estimation of probability, which assumes stationarity, nevertheless provides some measure of quantification as an indication them being plausible alternatives for consideration in future planning. In the millennium drought scenario, the historical probability of simulated droughts with different durations are low in terms of observed natural flow and are a little higher in terms of tree-ring reconstructions. Cumulative deficit figures for the mid-20th century drought and paleo tree-ring drought similar to Figure 11 are included in the supplementary material (Figure S-25 and Figure S-26). In the mid-20th century drought scenario, the simulated droughts are less severe and as a result their estimated probabilities are slightly higher than those in the millennium drought scenario. In the paleo tree-ring drought scenario, the cumulative deficit of simulated flow traces is higher than the other two drought scenarios, leading to estimated probabilities that are lower and are zero for many of the levels calculated. These low stationarity-based historical probabilities indicate that the droughts produced by the paleo tree-ring scenario are worse than droughts in the historical or tree-ring records.

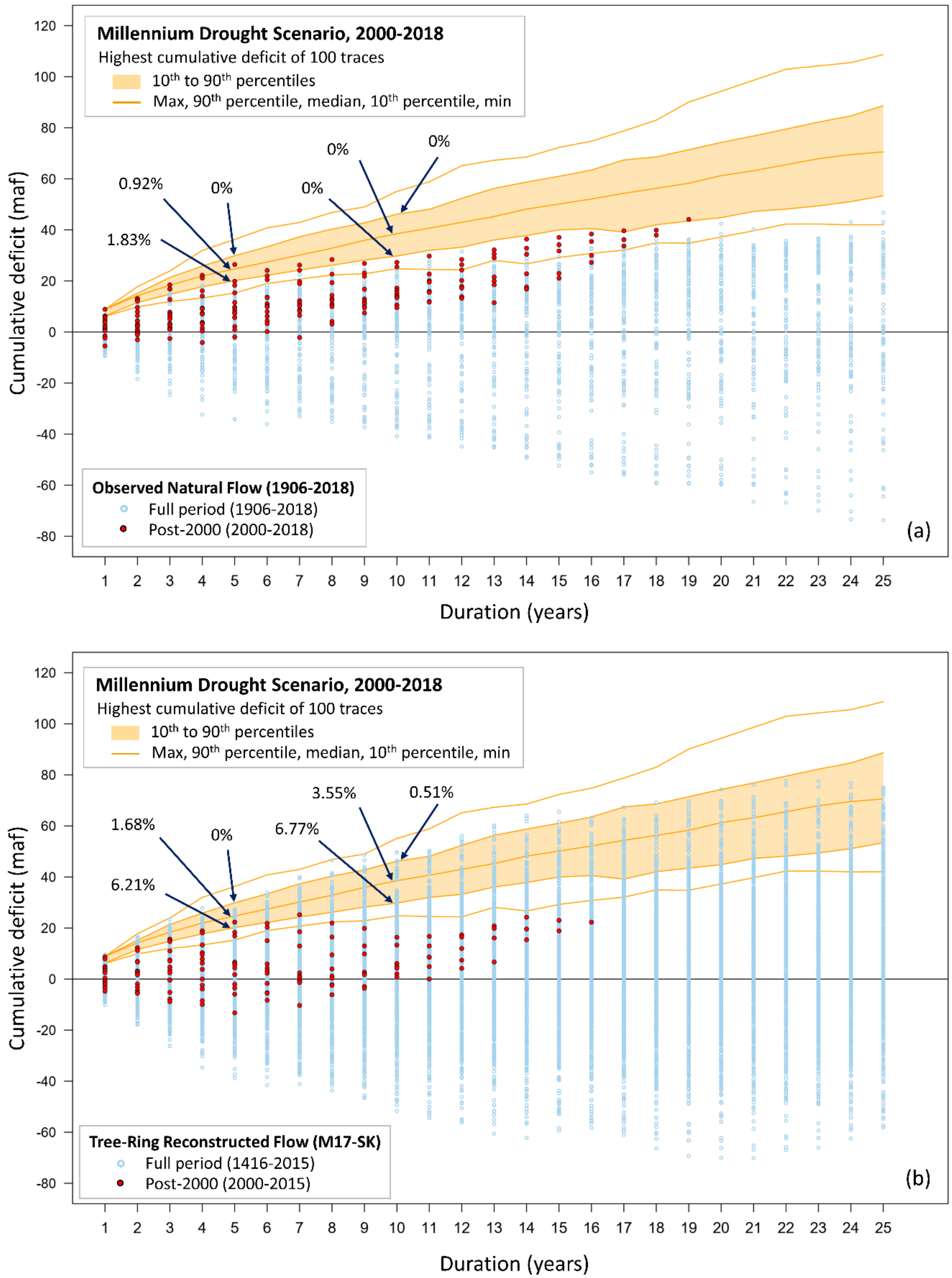


Figure 11. Highest millennium drought cumulative deficit compared with a) historical, and b) tree-ring reconstructed natural streamflow cumulative deficit.

Impacts of Various Hydrologic Scenarios on Lake Powell Levels

Each of the 100 traces from each drought scenario was disaggregated to provide inflows to CRSS at the 29 inflow points. The hydrologic scenarios were analyzed using the April 2020 version of CRSS representing the projected initial reservoir conditions for December 2020 and the current interpretation of the Law of the River as represented in the model. These assumptions include the 2007 Interim Shortage Guideline operation rules and the 2019 lower basin drought contingency plan and upper basin drought response operation. The exceedance probability of Lake Powell pool elevations for the drought scenarios was compared with the CRSS runs that use the Full and Stress Test hydrology scenarios from Reclamation. The Full hydrology, which is Reclamation's highest mean flow scenario, is an ensemble of 113 traces generated by applying ISM to the full period of the observed record (i.e. 1906-2018). Stress test hydrology, on the other hand, is a drier scenario from Reclamation including 31 traces generated by using ISM for a subset of observed flow from 1988 to 2018. Figure 12 demonstrates that the effect of the resampled scenarios we have developed persisting over a 20-year duration result in conditions more problematic than the current Reclamation stress test. A considerable fraction of the scenarios and sequences indicate that the reservoir elevation would fall below minimum power pool (19-41% of time), and even in some cases below the penstock intake levels (14-34% of time).

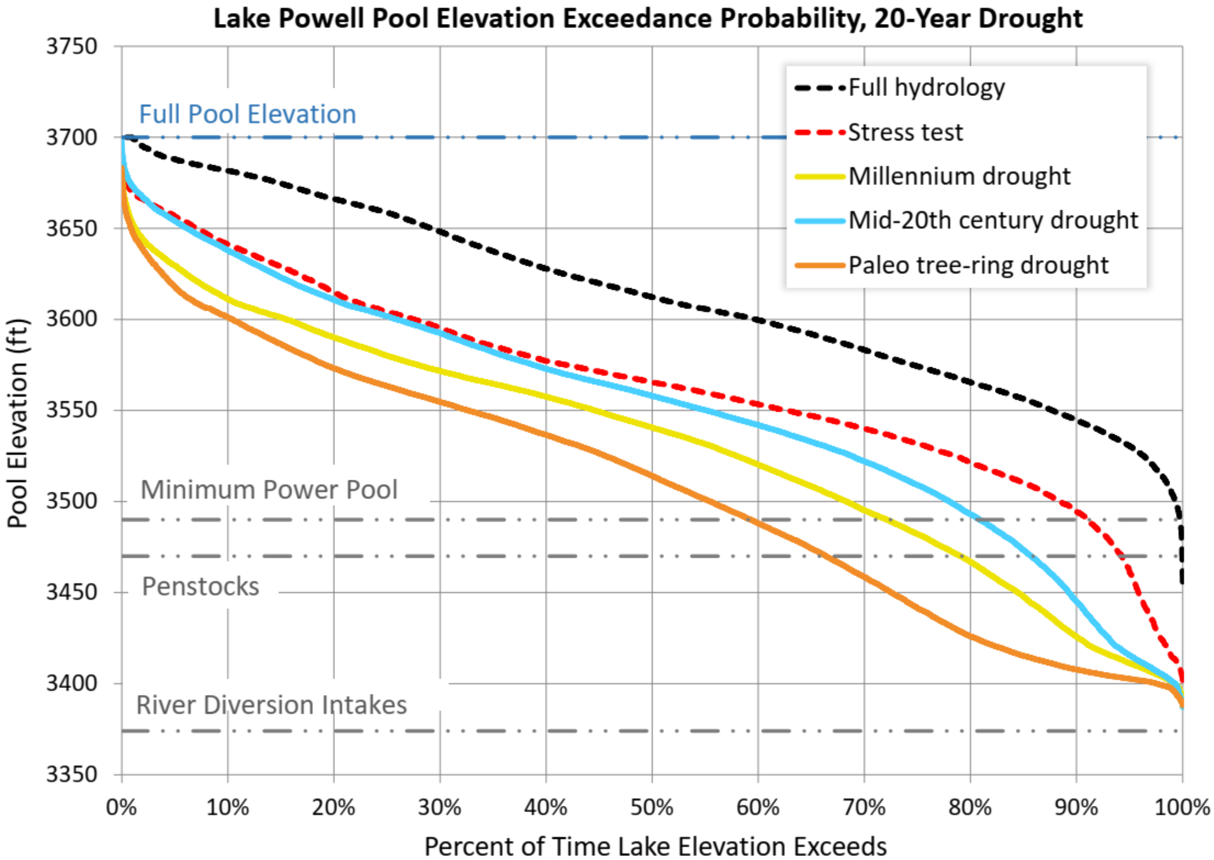


Figure 12. Lake Powell pool elevation in response to each of the drought scenarios over a 20-year period

DISCUSSION

Through analysis of historical streamflow data, this paper has documented that the Colorado River suffers from periodic severe and sustained drought and that water management planning should take this into account. The most recent 22 years (2000-2021) are the driest in the historical record (1906-2021). However, this period is not without precedent in paleo reconstructions using tree rings and also not as severe as some climate projections. Conversely, the first 24 years of the commonly used historical record (1906-1929), known as the Early 20th century pluvial period, are the wettest on record. It is interesting that the cumulative deficit analysis (Figure 6) appears skewed with surplus values (at the bottom) larger than deficit values at the top for longer durations. This reflects the fact that the early 20th century pluvial was a greater departure from the mean, but in the opposite direction, than even the present millennium drought, yet another indicator of how unusual the 1906-1929 pluvial period was. The tree ring cumulative deficit analysis (supplementary material Figure S-18) does not have this skewness, reflecting perhaps a level of balance between pluvials and droughts over the longer record.

Tree-ring reconstructions of streamflow serve to extend the observable record and provide evidence of multiple severe sustained droughts in the past. These data show that extreme droughts in the Colorado River Basin occur naturally and, at multi-century time scales frequently, further underscoring the need for planning scenarios where they are considered. The tree rings thus provide an ample source of believable potential low future flows from which to sample.

In much prior work, hydrologic scenarios have been developed using the index sequential method, which is a recycling of flows for a selected period. ISM introduces limited and, in our view, insufficient variability. In this study, we used resampling with replacement from three extreme past drought scenarios identified, combined with water year block disaggregation to obtain monthly streamflow at each of the 29 CRSS natural inflow sites to be used as hydrology inputs to CRSS and potentially other management models. These provide greater variety and serve to comprehensively stress the system for planning purposes. We note that resampling with replacement does not preserve annual correlation, which at least for lag 1 is statistically significant in the Colorado River. This is a limitation, justifiable in our judgment, given that resampling from a period selected as a drought scenario does introduce a degree of persistence in the scenarios developed. We provide further evaluation of this in online supplementary information. We note that while the water year block disaggregation does reproduce monthly and cross site correlation, it does not preserve correlation across water years. The water year break

September-October is, in the western U.S., a point where streamflow resets from one year to the next. Spring runoff peaks have typically waned by July, and August and September are typically months of low flows. The first new winter snow that builds up for the next seasons flow is just starting around September and October. For these reasons, September to October correlations are typically small.

Another drawback of the water year block disaggregation approach is that the July to November monsoon season spans the water year transition, and resampling of past flows may not reflect monsoon changes associated with climate change (Grantz et al., 2007). Given that the Colorado system has a high degree of storage, and short-term (less than a year) regional droughts are less of a concern for our focus on Colorado River basin-wide drought planning, these limitations are not addressed here.

One element of the variability represented by these scenarios is that sequences do include above average flow years, as a consequence of the past droughts they are based on including above average flow years. These occur at different and variable points in each sequence and reflect the fact that within persistent droughts there can be short periods with high flow that do not necessarily take the system out of a critical condition.

Climate change is an important factor and source of uncertainty that affects future streamflow. Climate change is occurring on top of the natural occurrence of droughts. Although GCM precipitation projections are highly variable in the Colorado River Basin, there is a consensus among most of the climate studies that the future runoff of the Colorado River Basin will decline as it warms and that the future might include megadroughts that are even worse than the drought scenarios quantified in this study. Temperature has increased across the Colorado River Basin, a trend that is virtually certain to continue and a factor that will likely reduce flows further. The information from CMIP5-VIC streamflow projections that we juxtaposed over our analyses of historical data and simulated scenarios frame our simulations in a climate change context. The simulations based on the paleo tree-ring drought are more severe than the most severe historical drought but plausible and perhaps not even that extreme when juxtaposed and considered in the context of climate change.

As noted earlier GCMs may be limited in their ability to model drought frequency (Ault et al., 2012; 2013) and this limitation has to be considered in evaluating the CMIP5-VIC duration severity projections (Figure 9). Our approach to selecting performant climate projections based on evaluating how well each of the projections represents the statistical distribution of historical droughts was an effort to address this. Among the CMIP5-VIC scenarios that did represent historical droughts well, five of 10 models project future droughts worse than the tree-ring record, indicating that it is not unreasonable to generate scenarios for planning purposes from the most severe tree-ring drought. It is notable that one performant CMIP5-VIC model projects a 25-year duration mean close to 10 maf/yr (Figure 9).

We do note that there are other approaches that could have been taken to select performant climate-derived streamflow projections, such as evaluation of large-scale weather pattern reproduction. There may also be limitations associated with the use of bias corrected and spatially downscaled model outputs parsed through a hydrologic model, not being discerning of a climate model that does not well reproduce historical large-scale climate patterns of temperature and precipitation. Evaluating such approaches was beyond the scope of what we could do, noting that while we have drawn conclusions based on the statistically performant models we selected, we do present results for all the CMIP5-VIC simulations (Figure 9). It is interesting to note that, while examining soil moisture droughts estimated from tree-ring reconstructions, Williams et al.

(2020) found that, due to climate change, the millennium drought (which would otherwise have been a moderate drought) had become comparable to mega droughts in the tree-ring soil moisture record.

The field of water resources planning is grappling with approaches to address non-stationarity due to climate change and other factors (Milly et al., 2008). The scenarios developed here are offered for use by the water management community as inputs to CRSS and other planning tools, as severe but realistic possible future flow conditions in the sense of being plausible but for which precise future probabilities, or risk, is difficult to determine and somewhat lacking in meaning due to non-stationarity. While, it is possible to estimate risk based on assumptions of stationarity that no longer hold, we have avoided attempting to too precisely assess risk for these scenarios, recognizing them as having level 3 uncertainty (Walker et al., 2013; van Dorsser et al., 2018). They represent plausible alternatives for which precise probabilities are not evaluated. Plausibility comes from the fact that the simulated flows across the basin in any one year are resampled from what has occurred in the past. Another consideration for these scenarios and their risk is that efforts to quantify risk may lead to highly uncertain low probability values that result in them being ignored in planning. We feel that this paper has established their plausibility and that ignoring them is unwise. Rather, our perspective is that robust and resilient decision-making demands consideration of rare but plausible events such as in the scenarios developed, and that was our goal, namely providing severe streamflow inputs to help prepare for a highly uncertain future.

Further support for plausibility has been provided by juxtaposing the duration-severity results over past duration-severity data and minimum mean flow duration-severity from climate projections. We found that given what is projected in terms of climate change (Figure 10), the severe droughts resampled from past flows are within the ranges of, and thus consistent with, climate projections.

While quantifying precise risk is uncertain, we did nevertheless use juxtaposition of our drought scenario cumulative deficits in comparison to the observed full historical and tree-ring natural streamflow data to estimate probabilities. Such probability calculations assume stationarity, which is why we refer to them as historical probabilities. In the case of the millennium drought scenario we note that the simulated flows are resampled from the 2000-2018 period of the historical natural flow, which has been experienced by the current generation of water managers and users. Stakeholders should consider these drought scenarios to evaluate paradigms for water allocation under these circumstances as part of drought planning.

The evaluation of CRSS simulated Lake Powel elevations under current Law of the River operations and demand allocations indicates considerable periods where elevations would be at an unsustainable and unacceptable level below hydropower penstock intakes. This would be catastrophic both for water supply, power generation and the ecosystem downstream from Lake Powell, and motivates the need to develop alternative management paradigms to account for the possibility of the inflow scenarios developed here. A wide variety of alternative management paradigms to stabilize the Colorado River using the hydrology scenarios developed here have been investigated by Wheeler et al. (2021) and (2022).

CONCLUSIONS

Climate warming has already been shown to reduce runoff in the Colorado River Basin, a highly utilized basin where current demand exceeds supplies. Future warming is projected to cause additional significant losses. This will occur on top of periodic severe and sustained droughts, which have occurred in the past, and have high likelihood of occurring in a warming climate. To better manage the Colorado River System in this uncertain future and mitigate vulnerabilities, water managers need to evaluate the system behavior under defensible worst-case possible scenarios.

This paper has examined available information on the hydrology of the Colorado River Basin and constructed plausible drought scenarios that address the multiple factors and uncertainties involved and are intended to be used as a basis for testing alternative operation and management paradigms.

Three past periods of severe and sustained droughts in the Colorado River Basin were identified using the average of streamflow, and the cumulative deficit relative to the mean flow, over varying durations within historical and tree-ring reconstructions of natural streamflow at Lees Ferry. These were identified using a novel visualization approach that depicts the mean flow, or cumulative deficit relative to the long term mean, for a range of durations with each sequence of each duration plotted as a point. This plotting technique illustrates the interplay between duration and the distribution of mean flow depicting drought severity. This visualization approach is general and can be applied across any streamflow scenario to help assess differences across critical flow sequences for basins beyond just the Colorado River Basin.

The millennium drought from 2000-2018 is characterized by a water year average flow of 12.44 maf/yr, significantly below the 1906-2018 mean of 14.76 maf/yr. This leads to a 19-year cumulative deficit of 44 maf. Extending this to 2021 including provisional natural flow estimates for the later years, this drought has a water year average flow of 12.3 maf/yr and 22 year cumulative deficit of 54 maf. The mid-20th century drought from 1953-1977 has a water year average flow of 12.89 maf/yr. These are both droughts in the historical record whose potential recurrence should be planned for. Tree-ring reconstructions of streamflow serve to extend the observable record and provide evidence of multiple severe sustained droughts in the past. The paleo tree-ring drought from 1576-1600 had an average flow of 11.76 maf/yr, notably lower than the historical droughts, and is representative of extreme droughts that occur naturally within the Colorado River Basin.

We used resampling of the flows at Lees Ferry for the drought scenarios we identified to provide an ensemble of 100 plausible annual streamflow traces. We used a nonparametric resampling approach referred to as water year block disaggregation to split the simulated annual flow at Lees Ferry into monthly flow at each of the 29 CRSS natural inflow sites. The scenarios we present, while extreme, are grounded in past streamflow records and the maxim that if it has occurred in the past, it may occur again. These scenarios therefore serve as plausible stress tests for future hydrology of the Colorado Basin. The plausibility of these scenarios was established by comparison to past historical and tree-ring reconstructed natural streamflow and climate projections selected to be performant in their reproduction of the duration and severity of past droughts.

The 100 sequences from each scenario that we developed are available for use in CRSS and other planning tools as severe but realistic possible future flow conditions that should be planned for. Minimum duration-severity and maximum cumulative deficit relative to the 1906-2018 natural flow mean were used to quantify drought severity for each of the scenarios developed. The 10 to 90 percentile range of five-year cumulative deficits is from 20.13 to 29.88 maf for the millennium drought scenario and represents a significant but plausible deficit of flow to plan for. The recent 2000-2004 drought falls in the middle of this range. When used as input to the currently configured CRSS, the scenarios developed indicate considerable periods with Lake Powell at a level below its hydropower penstocks, a critically low level for operation of the reservoir and supply of water to the lower basin, indicating the need for rethinking the management paradigms for operation of these reservoirs in the face of future droughts.

DATA AVAILABILITY

The data used and scenarios developed in this study are publicly available in HydroShare https://www.hydroshare.org/resource/ca2e152c9fca4b2aa7c3294a388c522d/ (Salehabadi and Tarboton, 2022c). Scripts used to produce duration severity and cumulative deficit plots are in HydroShare https://www.hydroshare.org/resource/bbe8dffacb07458783b2e6924aa615bb/ (Salehabadi and Tarboton, 2022b). A HydroShare collection holding these resources is http://www.hydroshare.org/resource/6d351874f16947609eab585a81c3c60d (Salehabadi and Tarboton, 2022a).

SUPPORTING INFORMATION

Additional supporting information may be found online under the Supporting Information tab for this article: Evaluation of other stochastic models (multi-year block disaggregation, and AR1) designed to preserve correlation, a cumulative deficit analysis for tree-ring-reconstructed flow, statistics (mean and standard deviations) of monthly and annual streamflow simulations, and highest cumulative deficit plots of mid-20th century and paleo tree-ring drought scenarios are included as supplementary online supporting information.

ACKNOWLEDGMENTS

We are grateful to the Catena Foundation and the Walton Family Foundation for their financial support for this project. Guidance, advice, feedback and insightful conversation greatly improved the quality of this effort – thanks to Chris Harris, Don Ostler, Sara Larsen, Amy Haas, Jim Prairie, Eric Kuhn, and Sara Goeking for their thoughtful contributions. The future of the Colorado River team worked hard to refine and direct these ideas: Jian Wang, Bethany Neilson, David Rosenberg, Christy Leonard, Lindsey Bruckerhoff, and Bryce Mihalevich. Thanks to Lael Gilbert for editing support. This paper has also benefited from the comments of anonymous JAWRA reviewers and associate editor Flavio Lehner whom we thank for their ideas and suggestions.

LITERATURE CITED

Ahmadalipour, A., A. Rana, H. Moradkhani and A. Sharma, (2015), "Multi-criteria evaluation of CMIP5 GCMs for climate change impact analysis," *Theoretical and Applied Climatology*, 128(1): 71-87, https://doi.org/10.1007/s00704-015-1695-4.

Ault, T. R., J. E. Cole, J. T. Overpeck, G. T. Pederson, S. St. George, B. Otto-Bliesner, C. A. Woodhouse and C. Deser, (2013), "The continuum of hydroclimate variability in western North America during the last millennium," *Journal of Climate*, 26(16): 5863-5878, https://doi.org/10.1175/JCLI-D-11-00732.1.

Ault, T. R., J. E. Cole and S. St. George, (2012), "The amplitude of decadal to multidecadal variability in precipitation simulated by state-of-the-art climate models," *Geophysical Research Letters*, 39(21), https://doi.org/10.1029/2012GL053424.

Barnett, T. P. and D. W. Pierce, (2008), "When will Lake Mead go dry?," *Water Resour. Res.*, 44: W03201, http://dx.doi.org/10.1029/2007WR006704

Bras, R. L. and I. Rodriguez-Iturbe, (1985), *Random Functions and Hydrology*, Addison-Wesley, Reading, MA, 559 p.

Brekke, L. D., M. D. Dettinger, E. P. Maurer and M. Anderson, (2008), "Significance of model credibility in estimating climate projection distributions for regional hydroclimatological risk assessments," *Climatic Change*, 89(3): 371-394, https://doi.org/10.1007/s10584-007-9388-3.

Christensen, N. S., A. W. Wood, N. Voisin, D. P. Lettenmaier and R. N. Palmer, (2004), "The effects of climate change on the hydrology and water resources of the Colorado River basin," *Climatic change*, 62(1-3): 337-363, https://doi.org/10.1023/B:CLIM.0000013684.13621.1f.

Colorado Water Conservation Board, (2012), "Colorado River Availability Study Phase I," Colorado Water Conservation Board, https://dnrweblink.state.co.us/cwcb/0/doc/158319/Electronic.aspx?searchid=78f0eafa-0b8f-4d8a-9ff3-faf67cc82f52.

DCP, (2019), H.R. 2030 - Colorado River Drought Contingency Plan Authorization Act. https://www.congress.gov/bill/116th-congress/house-bill/2030 - https://www.usbr.gov/dcp/finaldocs.html

Grantz, K., B. Rajagopalan, M. Clark and E. Zagona, (2007), "Seasonal Shifts in the North American Monsoon," *Journal of Climate*, 20(9): 1923-1935, https://doi.org/10.1175/JCLI4091.1.

Hidalgo, H. G., T. C. Piechota and J. A. Dracup, (2000), "Alternative principal components regression procedures for dendrohydrologic reconstructions," *Water Resources Research*, 36(11): 3241-3249, https://doi.org/10.1029/2000WR900097.

Hoerling, M., J. Barsugli, B. Livneh, J. Eischeid, X. Quan and A. Badger, (2019), "Causes for the Century-Long Decline in Colorado River Flow," *Journal of Climate*, 32(23): 8181-8203, https://doi.org/10.1175/JCLI-D-19-0207.1.

IPCC, (2014), "Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change," [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)], IPCC, Geneva, Switzerland, 151 p., https://www.ipcc.ch/report/ar5/syr/.

Kuhn, E. and J. Fleck, (2019), *Science Be Dammed: How Ignoring Inconvenient Science Drained the Colorado River*, University of Arizona Press, 288 p, https://uapress.arizona.edu/book/science-be-dammed.

LaRue, E. C., (1925), *Water power and flood control of Colorado River below Green River, Utah*, USGS Water Supply Paper No. 556. , US Government Printing Office, Washington, D.C.

Loucks, D. P., E. van Beek, J. R. Stedinger, J. P. M. Dijkman and M. T. Villars, (2005), *Water Resources Systems Planning and Management: An Introduction to Methods, Models and Applications*, UNESCO, Paris, 676 p, http://hdl.handle.net/1813/2804

Lukas, J., E. Gutmann, B. Harding and F. Lehner, (2020), "Climate Change-Informed Hydrology," Chap. 11 in: *Colorado River Basin Climate and Hydrology: State of the Science*, edited by J. Lukas and E. Payton, Western Water Assessment, University of Colorado Boulder, p.384-449, https://doi.org/10.25810/3hcv-w477.

Lukas, J. and E. Payton, (2020), *Colorado River Basin Climate and Hydrology: State of the Science*, Western Water Assessment, University of Colorado Boulder, https://doi.org/10.25810/3hcv-w477.

MacDonnell, L., (2021a), "Colorado River Basin. Waters and Water Rights, Lexis-Nexus, CORB-1," *SSRN*, https://ssrn.com/abstract=3780342.

MacDonnell, L., (2021b), *The Law of the Colorado River: Coping with Severe Sustained Drought, Part II*, Preprint submitted to JAWRA, https://www.researchgate.net/publication/350328112, https://ssrn.com/abstract=3811024.

MacDonnell, L. J., D. H. Getches and W. C. Hugenberg Jr, (1995), "The Law of the Colorado River: Coping with Severe Sustained Drought," *Water Resources Bulletin*, 31(5): 825-836, https://doi.org/10.1111/j.1752-1688.1995.tb03404.x.

McCabe, G. J. and D. M. Wolock, (2007), "Warming may create substantial water supply shortages in the Colorado River basin," *Geophysical Research Letters*, 34(22), https://doi.org/10.1029/2007GL031764.

McCabe, G. J., D. M. Wolock, G. T. Pederson, C. A. Woodhouse and S. McAfee, (2017), "Evidence that recent warming is reducing upper Colorado River flows," *Earth Interactions*, 21(10): 1-14, https://doi.org/10.1175/EI-D-17-0007.1.

Megdal, S. B., (2022), "Water Policy Options as Arizona Adapts to a Drier Colorado River: A Perspective," *Renewable Resources Journal*, 37(3): 2-10, https://rnrf.org/wp-content/uploads/2022/05/RRJV37N3.pdf.

Meko, D., C. Woodhouse and K. Morino, (2012), "Dendrochronology and links to streamflow," *Journal of Hydrology*, 412: 200-209.

Meko, D. M., C. A. Woodhouse, C. A. Baisan, T. Knight, J. J. Lukas, M. K. Hughes and M. W. Salzer, (2007), "Medieval drought in the upper Colorado River Basin," *Geophysical Research Letters*, 34(10), https://doi.org/10.1029/2007GL029988.

Meko, D. M., C. A. Woodhouse and E. R. Bigio, (2017), "Southern California Tree-Ring Study," Final Report to California Department of Water Resources, Agreement 4600011071, University of Arizona, https://cwoodhouse.faculty.arizona.edu/content/california-department-water-resources-studies.

Michaelsen, J., A. Loaiciga, L. Haston and S. Garver, (1990), "Estimating Drought Probabilities in California Using Tree Rings. California Department of Water Resour Report B-57105," Univ of California, Santa Barbara, CA.

Milly, P. C. and K. A. Dunne, (2020), "Colorado River flow dwindles as warming-driven loss of reflective snow energizes evaporation," *Science*, 367(6483): 1252-1255, https://doi.org/10.1126/science.aay9187.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier and R. J. Stouffer, (2008), "Stationarity Is Dead: Whither Water Management?," *Science*, 319(5863): 573-574, http://doi.org/10.1126/science.1151915.

Minute 323, (2017), "Minute No. 323, Extension of Cooperative Measures and Adoption of a Binational Water Scarcity Contingency Plan in the Colorado River Basin dated September 21, 2017," International Boundary and Water Commission, Available online at: https://www.ibwc.gov/Files/Minutes/Min323.pdf.

Mote, P., L. Brekke, P. B. Duffy and E. Maurer, (2011), "Guidelines for constructing climate scenarios," *Eos, Transactions American Geophysical Union*, 92(31): 257-258, https://doi.org/10.1029/2011EO310001.

Nasrollahi, N., A. AghaKouchak, L. Cheng, L. Damberg, T. J. Phillips, C. Miao, K. Hsu and S. Sorooshian, (2015), "How well do CMIP5 climate simulations replicate historical trends and patterns of meteorological droughts?," *Water Resources Research*, 51(4): 2847-2864, https://doi.org/10.1002/2014WR016318.

Nowak, K., J. Prairie, B. Rajagopalan and U. Lall, (2010), "A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow," *Water Resources Research*, 46(8), https://doi.org/10.1029/2009WR008530.

Ouarda, T. B. M. J., J. W. Labadie and D. G. Fontane, (1997), "Indexed Sequential Hydrologic Modeling for Hydropower Capacity Estimation," *JAWRA Journal of the American Water*

Resources Association, 33(6): 1337-1349, https://doi.org/10.1111/j.1752-1688.1997.tb03557.x.

Payton, E., (2020), "Historical Hydrology," Chap. 9 in: *Colorado River Basin Climate and Hydrology: State of the Science*, edited by J. Lukas and E. Payton, Western Water Assessment, University of Colorado Boulder, p.337-360, https://doi.org/10.25810/3hcv-w477.

Payton, E., R. Smith, C. Jerla and J. Prairie, (2020), "Primary Planning Tools," Chap. 3 in: *Colorado River Basin Climate and Hydrology: State of the Science*, edited by J. Lukas and E. Payton, Western Water Assessment, University of Colorado Boulder, p.82-111, https://doi.org/10.25810/3hcv-w477.

Pierce, D. W., T. P. Barnett, B. D. Santer and P. J. Gleckler, (2009), "Selecting global climate models for regional climate change studies," *Proceedings of the National Academy of Sciences*, 106(21): 8441, https://doi.org/10.1073/pnas.0900094106.

Prairie, J. and R. Callejo, (2005), "Natural Flow and Salt Computation Methods, Calendar Years 1971-1995," U.S. Department of the Interior, Bureau of Reclamation, Salt Lake City, Utah, https://digitalcommons.usu.edu/govdocs/135/.

Prairie, J., K. Nowak, B. Rajagopalan, U. Lall and T. Fulp, (2008), "A stochastic nonparametric approach for streamflow generation combining observational and paleoreconstructed data," *Water Resour. Res.*, 44(6): W06423, http://dx.doi.org/10.1029/2007WR006684

Prairie, J., B. Rajagopalan, U. Lall and T. Fulp, (2007), "A stochastic nonparametric technique for space-time disaggregation of streamflows," *Water Resour. Res.*, 43(3): W03432, http://dx.doi.org/10.1029/2005WR004721

Prairie, J. R., B. Rajagopalan, T. J. Fulp and E. A. Zagona, (2006), "Modified K-NN Model for Stochastic Streamflow Simulation," *Journal of Hydrologic Engineering*, 11(4): 371-378, http://dx.doi.org/10.1061/(ASCE)1084-0699(2006)11:4(371).

Robeson, S. M., J. T. Maxwell and D. L. Ficklin, (2020), "Bias correction of paleoclimatic reconstructions: A new look at 1200+ years of Upper Colorado River flow," *Geophysical Research Letters*, 47: e2019GL086689, https://doi.org/10.1029/2019GL086689.

Rupp, D. E., J. T. Abatzoglou and P. W. Mote, (2017), "Projections of 21st century climate of the Columbia River Basin," *Climate Dynamics*, 49(5): 1783-1799, https://doi.org/10.1007/s00382-016-3418-7.

Salehabadi, H. and D. G. Tarboton, (2022a), "Data Collection to Supplement the Future Hydrology of the Colorado River Basin Study," HydroShare, https://doi.org/10.4211/hs.6d351874f16947609eab585a81c3c60d.

Salehabadi, H. and D. G. Tarboton, (2022b), "Duration-Severity and Cumulative Deficit Analyses for Colorado River Streamflow at Lees Ferry," HydroShare, https://doi.org/10.4211/hs.bbe8dffacb07458783b2e6924aa615bb.

Salehabadi, H. and D. G. Tarboton, (2022c), "Hydrology scenarios that characterize plausible future drought conditions in the Colorado River Basin," HydroShare, https://doi.org/10.4211/hs.ca2e152c9fca4b2aa7c3294a388c522d.

Salehabadi, H., D. G. Tarboton, E. Kuhn, B. Udall, K. G. Wheeler, D. E. Rosenberg, S. A. Goeking and J. C. Schmidt, (2020), "The Future Hydrology of the Colorado River Basin," White Paper 4, Future of the Colorado River Project, Center for Colorado River Studies, Utah State University, 71 p., https://qcnr.usu.edu/coloradoriver/files/WhitePaper4.pdf.

Schmidt, J. C., L. Bruckerhoff, H. Salehabadi and J. Wang, (2022), "The Colorado River," in: *Large Rivers: Geomorphology and Management*, edited by A. Gupta, John Wiley & Sons, Ltd, p.253-319, https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119412632.ch10.

Schulman, E., (1946), *Tree-ring hydrology of the Colorado River basin*, Vol 16, University of Arizona, Tucson, Arizona.

Srinivas, V. V. and K. Srinivasan, (2005), "Hybrid moving block bootstrap for stochastic simulation of multi-site multi-season streamflows," *Journal of Hydrology*, 302(1): 307-330, https://doi.org/10.1016/j.jhydrol.2004.07.011.

Srinivas, V. V. and K. Srinivasan, (2006), "Hybrid matched-block bootstrap for stochastic simulation of multiseason streamflows," *Journal of Hydrology*, 329(1): 1-15, https://doi.org/10.1016/j.jhydrol.2006.01.023.

Stockton, C. W. and G. C. Jacoby, (1976), "Long-Term Surface-Water Supply and Streamflow Trends in the Upper Colorado River Basin Based on Tree-Ring Analyses," *Lake Powell Research Project Bulletin*, No. 18, National Science Foundation.

Tarboton, D. G., (1994), "The Source Hydrology of Severe Sustained Drought in the Southwestern United States," *Journal of Hydrology*, 161: 31-69, http://dx.doi.org/10.1016/0022-1694(94)90120-1.

Tarboton, D. G., (1995), "Hydrologic Scenarios for Severe Sustained Drought in the Southwestern United States," *Water Resources Bulletin*, 31(5): 803-813, https://doi.org/10.1111/j.1752-1688.1995.tb03402.x.

Tarboton, D. G., A. Sharma and U. Lall, (1998), "Disaggregation Procedures for Stochastic Hydrology based on Nonparametric Density Estimation," *Water Resources Research*, 34(1): 107-119, http://dx.doi.org/10.1029/97WR02429

Treaty, (1944), Treaty Between the United States of America and Mexico, signed February 3, 1944. Available online at: https://www.ibwc.gov/Files/1944Treaty.pdf

Udall, B. and J. Overpeck, (2017), "The twenty-first century Colorado River hot drought and implications for the future," *Water Resources Research*, 53(3): 2404-2418, https://doi.org/10.1002/2016WR019638.

Upper Colorado River Commission, (2018), "Seventieth Annual Report of the Upper Colorado River Commission," http://www.ucrcommission.com/RepDoc/UCRCAnnualReports/70_UCRC_Annual_Report.pdf.

USBR, (2010), "Colorado River Modeling Work Group Charter," https://www.usbr.gov/lc/region/programs/climateresearch/Charter_ModelingWorkGroup.pdf.
USBR, (2012), "Colorado River Basin Water Supply and Demand Study," U.S. Bureau of Reclamation, Available online at: https://www.usbr.gov/lc/region/programs/crbstudy/finalreport/studyrpt.html.
USBR, (2014), "Downscaled CMIP3 and CMIP5 climate and hydrology projections: Release of hydrology projections, comparison with preceding information, and summary of user needs," US Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, Colorado, 110 p., https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/techmemo/BCSD5HydrologyMemo.pdf.
USBR, (2020a), "Colorado River Basin Natural Flow and Salt Data", Colorado River Basin Natural Flow Database. Last updated January 10, 2020, https://www.usbr.gov/lc/region/g4000/NaturalFlow/current.html.
USBR, (2020b), "Review of the Colorado River Interim Guidelines for Lower Basin Shortages and Coordinated Operations for Lake Powell and Lake Mead, Upper and Lower Colorado Basin Regions," U.S. Department of the Interior, Bureau of Reclamation, December 2020, https://www.usbr.gov/ColoradoRiverBasin/documents/7.D.Review_FinalReport_12-18-2020.pdf.
USBR, (2021), "Colorado River Basin Natural Flow and Salt Data", Colorado River Basin Natural Flow Database. Last updated April 20, 2021, https://www.usbr.gov/lc/region/g4000/NaturalFlow/current.html.
USDOI, (2007), "Colorado River Interim Guidelines for Lower Basin Shortages and the Coordinated Operations for Lake Powell and Lake Mead - Final Environmental Impact Statement," https://www.usbr.gov/lc/region/programs/strategies/FEIS/index.html.
van Dorsser, C., W. E. Walker, P. Taneja and V. A. W. J. Marchau, (2018), "Improving the link between the futures field and policymaking," *Futures*, 104: 75-84, https://doi.org/10.1016/j.futures.2018.05.004.
Vano, J. A., T. Das and D. P. Lettenmaier, (2012), "Hydrologic sensitivities of Colorado River runoff to changes in precipitation and temperature," *Journal of Hydrometeorology*, 13(3): 932-949, https://doi.org/10.1175/JHM-D-11-069.1.
Vano, J. A., B. Udall, D. R. Cayan, J. T. Overpeck, L. D. Brekke, T. Das, H. C. Hartmann, H. G. Hidalgo, M. Hoerling and G. J. McCabe, (2014), "Understanding uncertainties in future Colorado River streamflow," *Bulletin of the American Meteorological Society*, 95(1): 59-78, https://doi.org/10.1175/BAMS-D-12-00228.1.
Vogel, R. M. and A. L. Shallcross, (1995), "The Moving Blocks Bootstrap Versus Parametric Time Series Models," *Water Resour. Res.*, 32(6): 1875–1882.

Walker, W. E., R. J. Lempert and J. H. Kwakkel, (2013), "Deep Uncertainty," in: *Encyclopedia of Operations Research and Management Science*, edited by S. I. Gass and M. C. Fu, Springer US, Boston, MA, p.395-402, https://doi.org/10.1007/978-1-4419-1153-7_1140.

Wang, J., D. E. Rosenberg, K. G. Wheeler and J. C. Schmidt, (2020), "Managing the Colorado River for an Uncertain Future," White Paper 3, Future of the Colorado River Project, Center for Colorado River Studies, Utah State University, 30 p., https://qcnr.usu.edu/coloradoriver/files/CCRS_White_Paper_3.pdf.

Wheeler, K. G., E. Kuhn, L. Bruckerhoff, B. Udall, J. Wang, L. Gilbert, S. A. Goeking, A. Kasprak, B. Mihalevich, B. Neilson, H. Salehabadi and J. C. Schmidt, (2021), "Alternative Management Paradigms for the Future of the Colorado and Green Rivers," White Paper 6, Future of the Colorado River Project, Center for Colorado River Studies, Utah State University, 91 p., https://qcnr.usu.edu/coloradoriver/files/WhitePaper6.pdf.

Wheeler, K. G., D. E. Rosenberg and J. C. Schmidt, (2019), "Water Resource Modeling of the Colorado River: Present and Future Strategies," White Paper 2, Future of the Colorado River Project, Center for Colorado River Studies, Utah State University, 47 p., https://qcnr.usu.edu/coloradoriver/files/WhitePaper2.pdf.

Wheeler, K. G., B. Udall, J. Wang, E. Kuhn, H. Salehabadi and J. C. Schmidt, (2022), "What will it take to stabilize the Colorado River?," *Science*, 377(6604): 373-375, https://doi.org/10.1126/science.abo4452.

Williams, A. P., B. I. Cook and J. E. Smerdon, (2022), "Rapid intensification of the emerging southwestern North American megadrought in 2020–2021," *Nature Climate Change*, 12(3): 232-234, https://doi.org/10.1038/s41558-022-01290-z.

Williams, A. P., E. R. Cook, J. E. Smerdon, B. I. Cook, J. T. Abatzoglou, K. Bolles, S. H. Baek, A. M. Badger and B. Livneh, (2020), "Large contribution from anthropogenic warming to an emerging North American megadrought," *Science*, 368(6488): 314, https://doi.org/10.1126/science.aaz9600.

Wood, A., L. Woelders and J. Lukas, (2020), "Hydrologic Models," Chap. 6 in: *Colorado River Basin Climate and Hydrology: State of the Science*, edited by J. Lukas and E. Payton, Western Water Assessment, University of Colorado Boulder, p.221-252, https://doi.org/10.25810/3hcv-w477.

Woodhouse, C. A., S. T. Gray and D. M. Meko, (2006), "Updated streamflow reconstructions for the Upper Colorado River Basin," *Water Resources Research*, 42(5), https://doi.org/10.1029/2005WR004455.

Woodhouse, C. A., G. T. Pederson, K. Morino, S. A. McAfee and G. J. McCabe, (2016), "Increasing influence of air temperature on upper Colorado River streamflow," *Geophysical Research Letters*, 43(5): 2174-2181, https://doi.org/10.1002/2015GL067613.

Woodhouse, C. A., R. M. Smith, S. A. McAfee, G. T. Pederson, G. J. McCabe, W. P. Miller and A. Csank, (2021), "Upper Colorado River Basin 20th century droughts under 21st century warming: Plausible scenarios for the future," *Climate Services*, 21: 100206, https://doi.org/10.1016/j.cliser.2020.100206.

Zagona, E. A., T. J. Fulp, R. Shane, T. Magee and H. M. Goranflo, (2001), "Riverware: A generalized tool for complex reservoir system modeling," *JAWRA Journal of the American Water Resources Association*, 37(4): 913-929, https://doi.org/10.1111/j.1752-1688.2001.tb05522.x.

SUPPORTING INFORMATION
for
"An Assessment of Potential Severe Droughts in the Colorado River Basin"

Homa Salehabadi, David G. Tarboton, Bradley Udall, Kevin G. Wheeler, John C. Schmidt

Contents

1- Evaluation of Methods for Streamflow Simulation

The random resampling with replacement method used for stochastic generation of streamflow drought scenarios in the paper does not preserve lag 1 correlation. Given many stochastic hydrology approaches developed to preserve correlation, among other statistics, in streamflow simulation, this supplement presents results from evaluation of a number of these approaches indicating that uncorrelated resampling is most defensible among the methods evaluated.

The lag 1 correlation of 1906-2018 natural annual flow at Lees Ferry is 0.24. Being based on L=113 years of data the threshold for statistical significance with 95% confidence is $1.96/\sqrt{L}$ = 0.184. Thus, the lag 1 correlation of 0.24 is statistically significant.

The following methods were evaluated

- 1-year block resampling
- 2-year block resampling
- 5-year block resampling
- Auto-regressive order 1 (AR1) model with correlation from past flows

The 1-year block resampling approach is equivalent to random resampling with replacement used in the paper. The 2- and 5-year block resampling methods are designed to capture some year-to-year dependence in a non-parametric way, while the AR1 model is a parametric approach to reproduce correlation in simulations.

For each of the three drought scenarios considered, as well as for the entire historic and entire tree-ring records we generated 100 traces using each of these methods. We set the mean and variance of the AR1 model to the mean and variance of the respective scenarios with correlation selected from the respective full historic record (i.e. observed flow correlation for observed drought scenarios and tree-ring correlation for tree ring scenarios). We compared lag 1 to 5 correlation from all these simulations with the lag 1 to 5 correlation computed from the full historical streamflow, full tree-ring record, and the scenario being used (Figure S-1 to Figure S-5). We also computed the cumulative deficit versus duration for each of these scenarios and compared to historical and tree-ring cumulative deficit (Figure S-6 to Figure S-17), as was done in the paper for the uncorrelated block length 1 resampling. We found that the lag 1 correlation (and also higher lag correlations) computed within the drought scenarios was essentially zero for the Millennium and mid-20th century droughts and had a value of 0.27 for the paleo drought. Given the negligible correlation for the two historic streamflow droughts, the use of block lengths of 2 and 5 did not reproduce any notable correlation in the sequences simulated. For the tree-ring drought, the 5-year block resampling did reproduce the tree-ring drought scenario correlation. The AR1 models in all cases reproduced the full historical record correlation as expected. We did also note that multi-year block resampling within a shortish (19-25 year) scenario does sometimes introduce a bias due to under-sampling of edge values. We found that cumulative deficit versus duration was not different across these methods to any degree that would alter our conclusions. The use of greater block lengths did change the spread of cumulative deficit versus duration, increasing it for the millennium and paleo droughts and reducing it for the mid-20th century drought. This we attribute to the specific pattern of low flows in each of these scenarios that multi-year block resampling preserves. The median cumulative deficit versus duration pattern did not change significantly. The AR1 model also had non significantly changed median cumulative deficit versus duration, but even greater spread due to the higher correlation that it reproduced. While this spread would be meaningful for simulations based on the full record, we feel it is invalid to consider with simulations that are effectively offset to the mean of the drought scenario already, that within it had negligible correlation. It effectively double counts the effect of a drought, because of the persistence due to lag 1 correlation being combined with the persistence of simulations centered on a drought scenario mean. We thus feel that uncorrelated resampling is more defensible than multi-year block resampling (which introduces mean biases) and AR 1 simulation that has this double-counting effect.

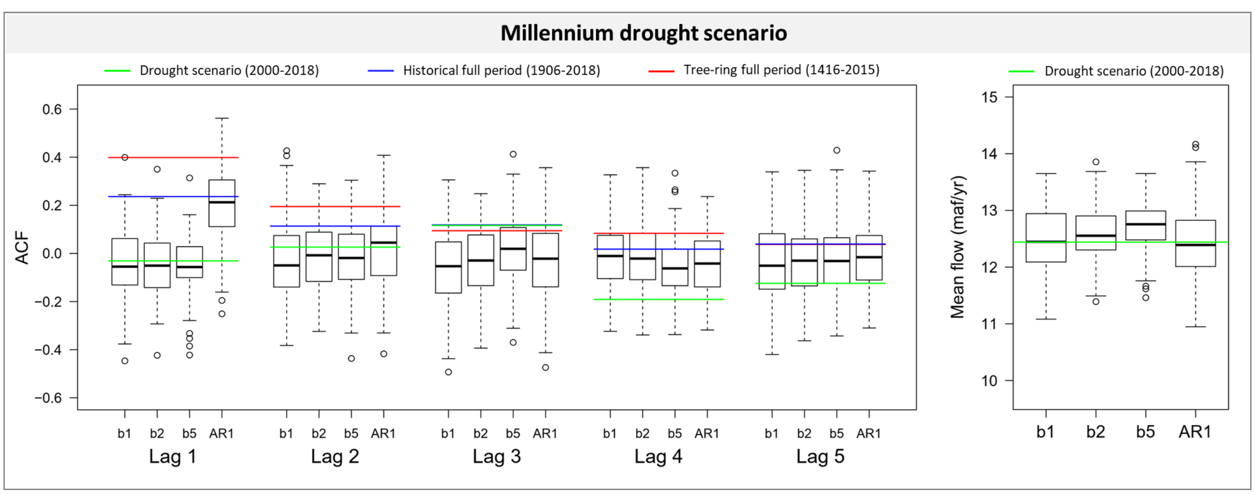


Figure S-1. Millennium drought scenario simulations using 1-yr block resampling (b1), 2-yr block resampling (b2), 5-yr block resampling (b5), and AR1. Left plot: Lag 1 to 5 correlation, and Right plot: Mean of simulations.

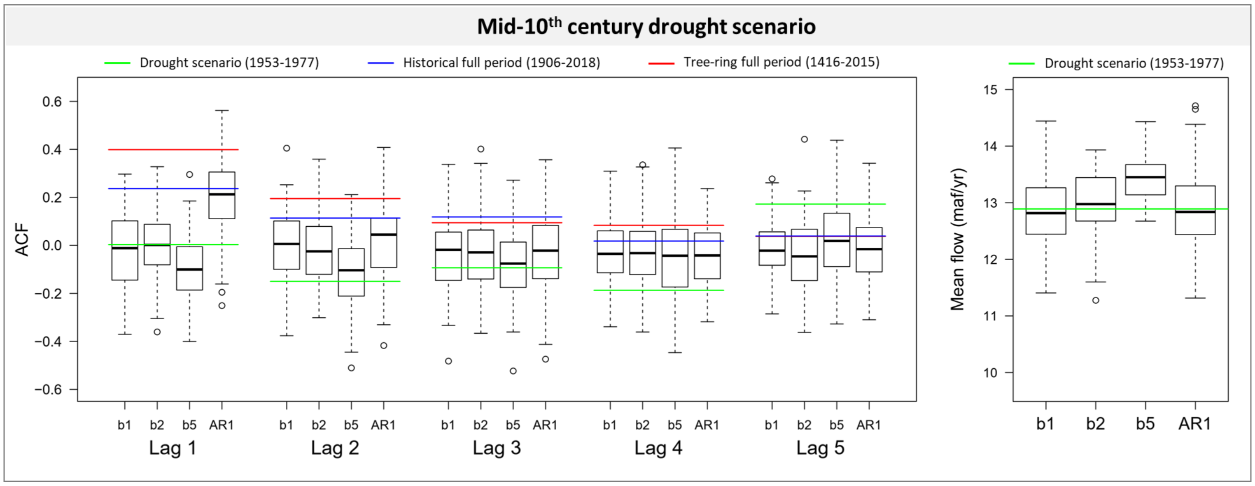


Figure S-2. Mid-20th century drought scenario simulations using 1-yr block resampling (b1), 2-yr block resampling (b2), 5-yr block resampling (b5), and AR1. Left plot: Lag 1 to 5 correlation, and Right plot: Mean of simulations.

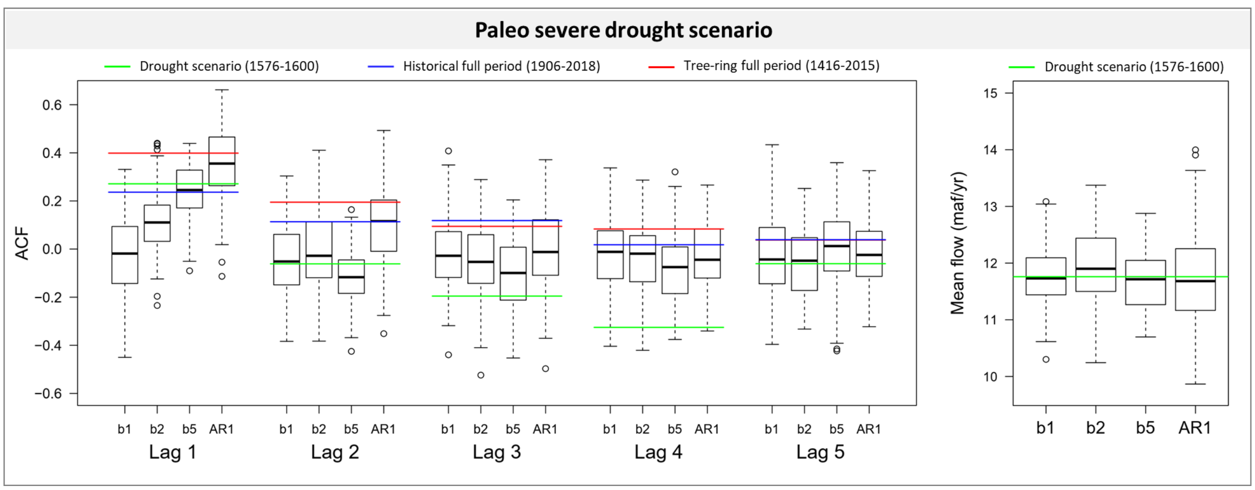


Figure S-3. Paleo severe drought scenario simulations using 1-yr block resampling (b1), 2-yr block resampling (b2), 5-yr block resampling (b5), and AR1. Left plot: Lag 1 to 5 correlation, and Right plot: Mean of simulations.

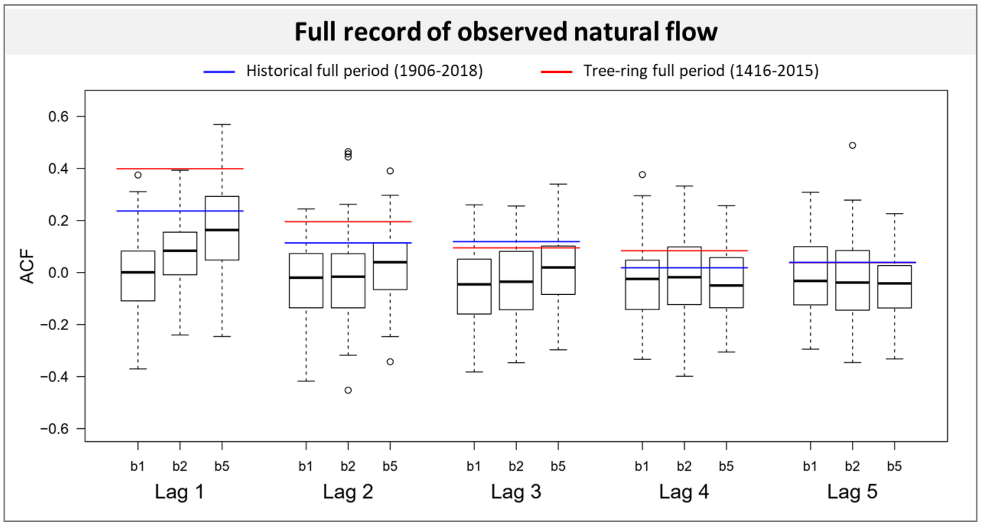


Figure S-4. Simulations from the full record of the observed natural flow using 1-yr block resampling (b1), 2-yr block resampling (b2), and 5-yr block resampling (b5). Left plot: Lag 1 to 5 correlation, and Right plot: Mean of simulations.

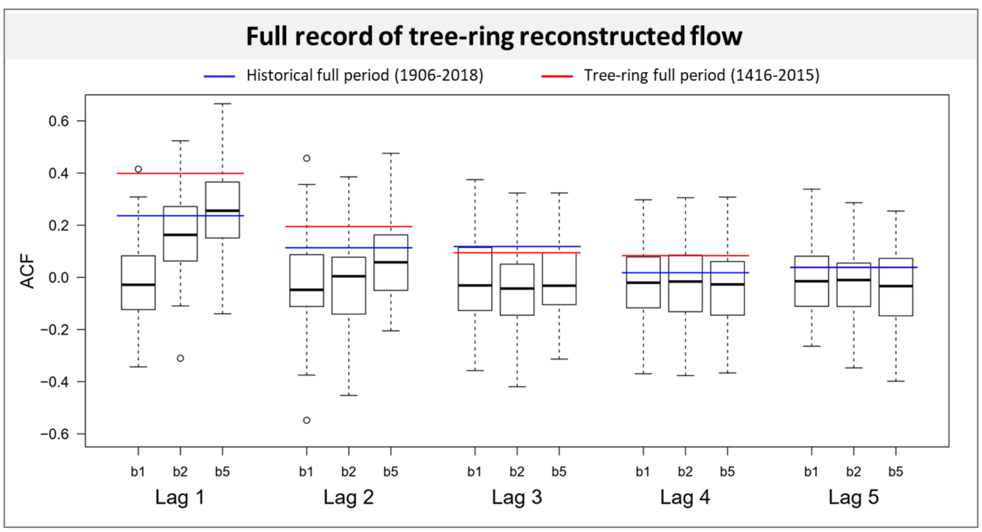


Figure S-5. Simulations from the full record of the tree-ring reconstructed flow using 1-yr block resampling (b1), 2-yr block resampling (b2), and 5-yr block resampling (b5). Left plot: Lag 1 to 5 correlation, and Right plot: Mean of simulations.

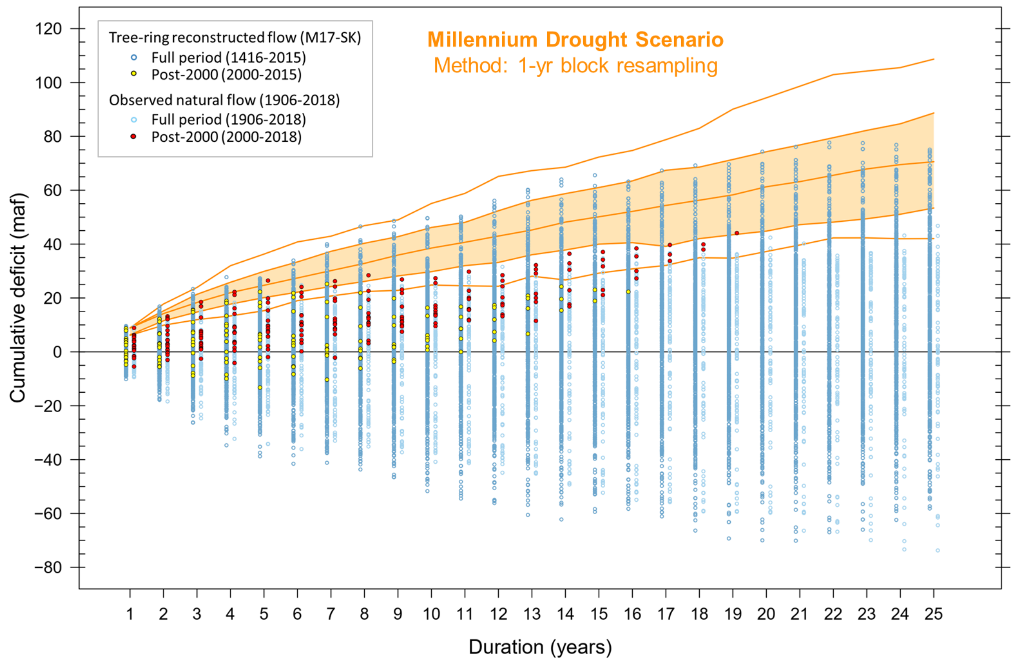


Figure S-6. Cumulative deficit versus duration for the flows simulated from the **millennium drought** scenario using **1-yr block resampling** (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

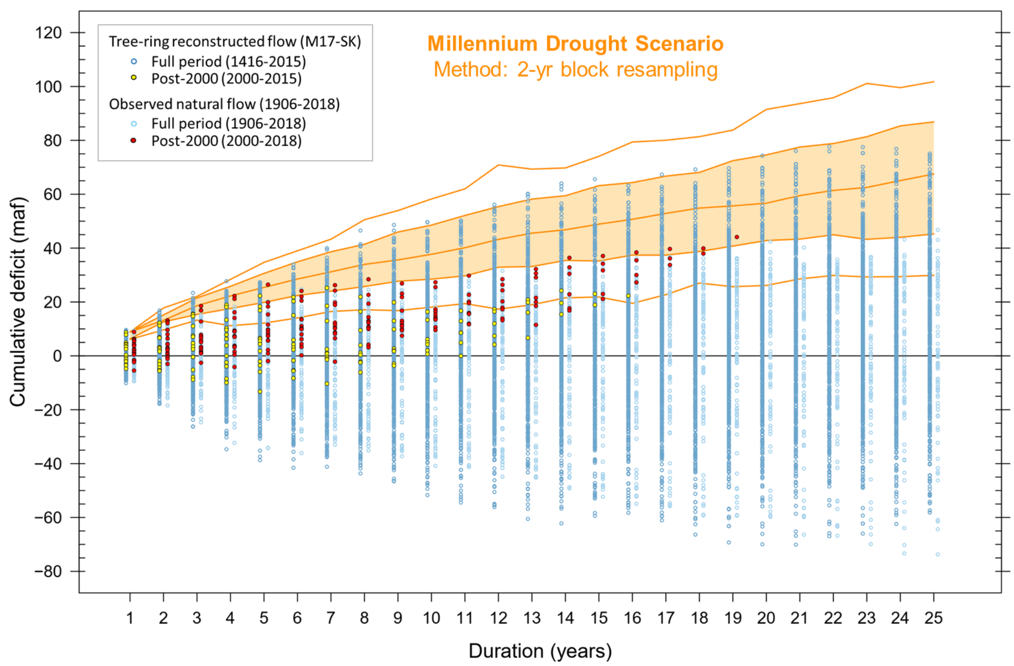


Figure S-7. Cumulative deficit versus duration for the flows simulated from the **millennium drought** scenario using **2-yr block resampling** (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

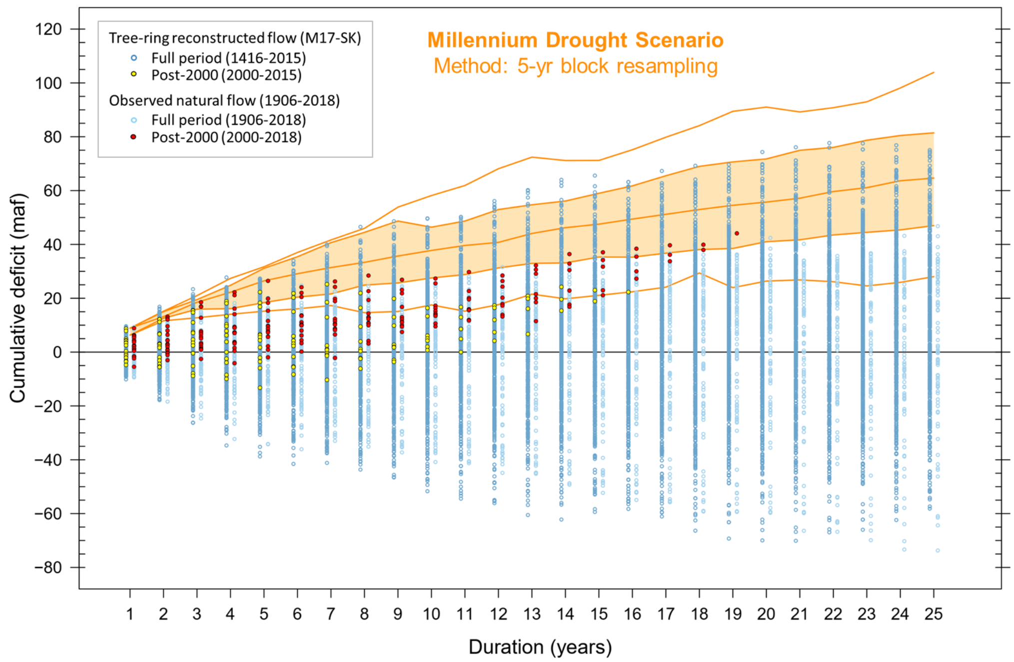


Figure S-8. Cumulative deficit versus duration for the flows simulated from the **millennium drought** scenario using **5-yr block resampling** (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

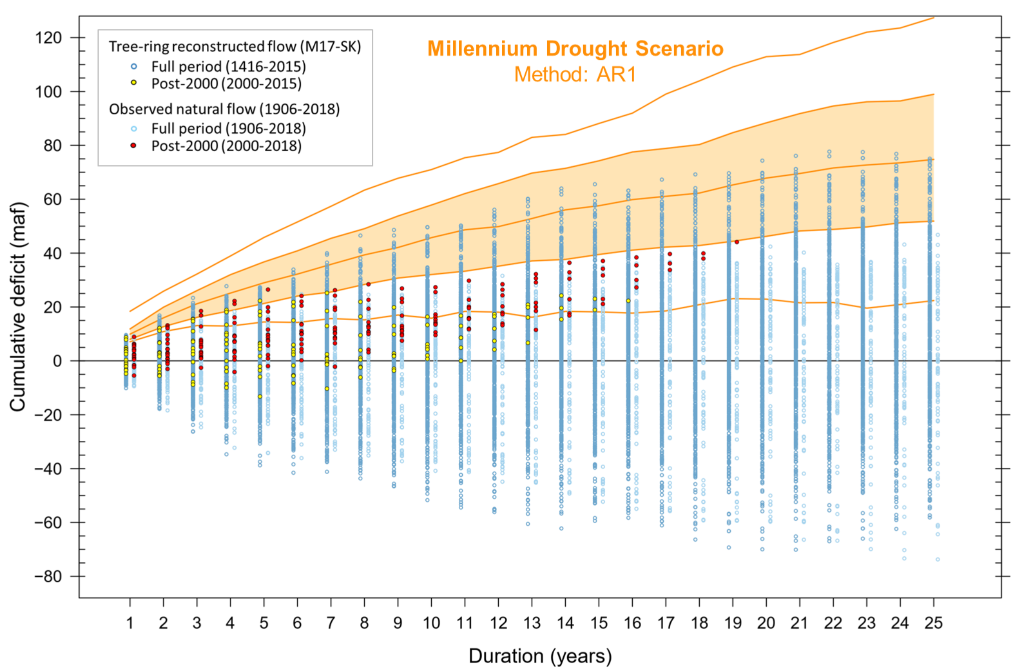


Figure S-9. Cumulative deficit versus duration for the flows simulated from the **millennium drought** scenario using **AR1** model (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

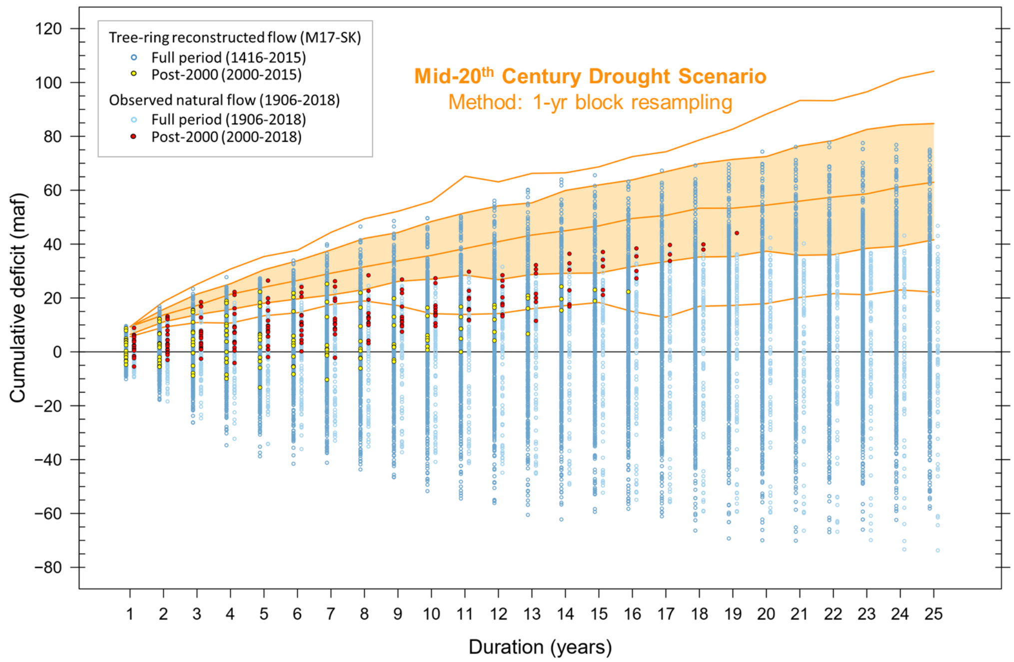


Figure S-10. Cumulative deficit versus duration for the flows simulated from the **mid-20th century drought** scenario using **1-yr block resampling** (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

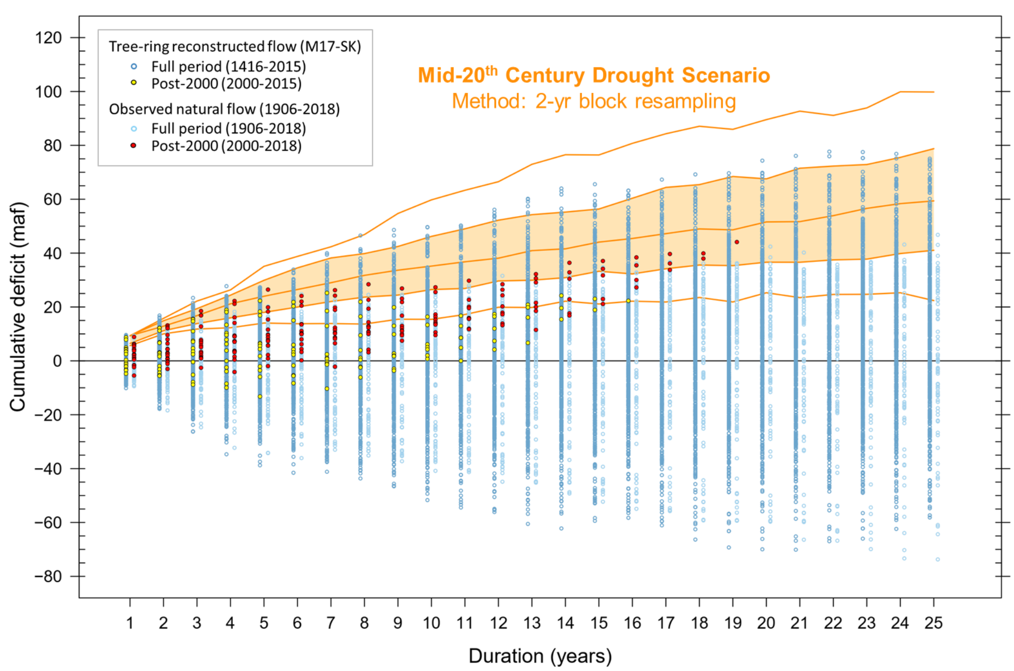


Figure S-11. Cumulative deficit versus duration for the flows simulated from the **mid-20th century drought** scenario using **2-yr block resampling** (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

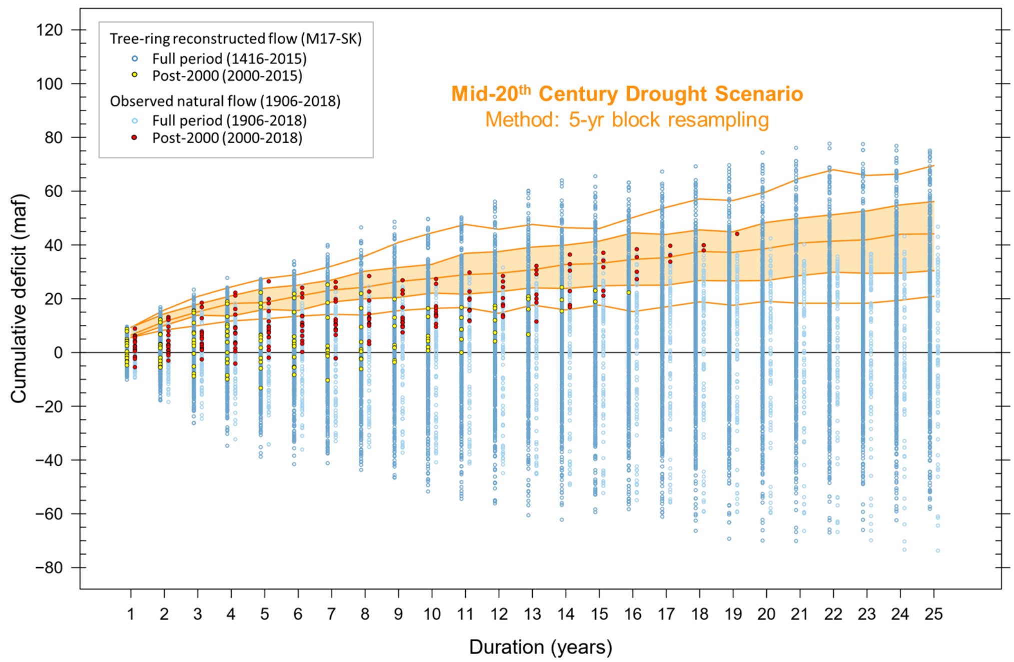


Figure S-12. Cumulative deficit versus duration for the flows simulated from the **mid-20th century drought** scenario using **5-yr block resampling** (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

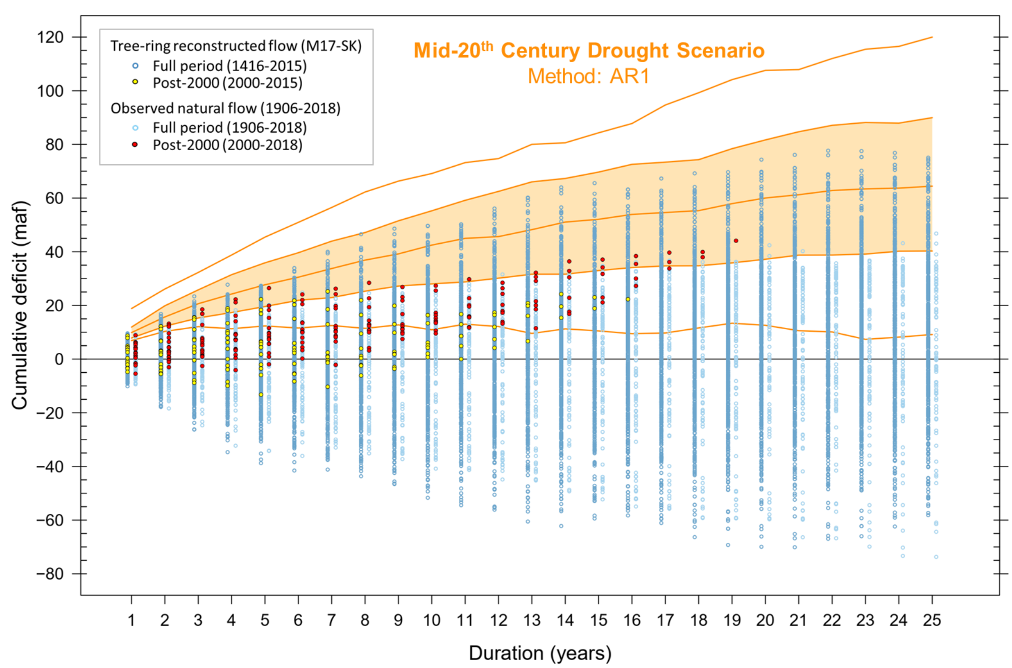


Figure S-13. Cumulative deficit versus duration for the flows simulated from the **mid-20th century drought** scenario using **AR1** model (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

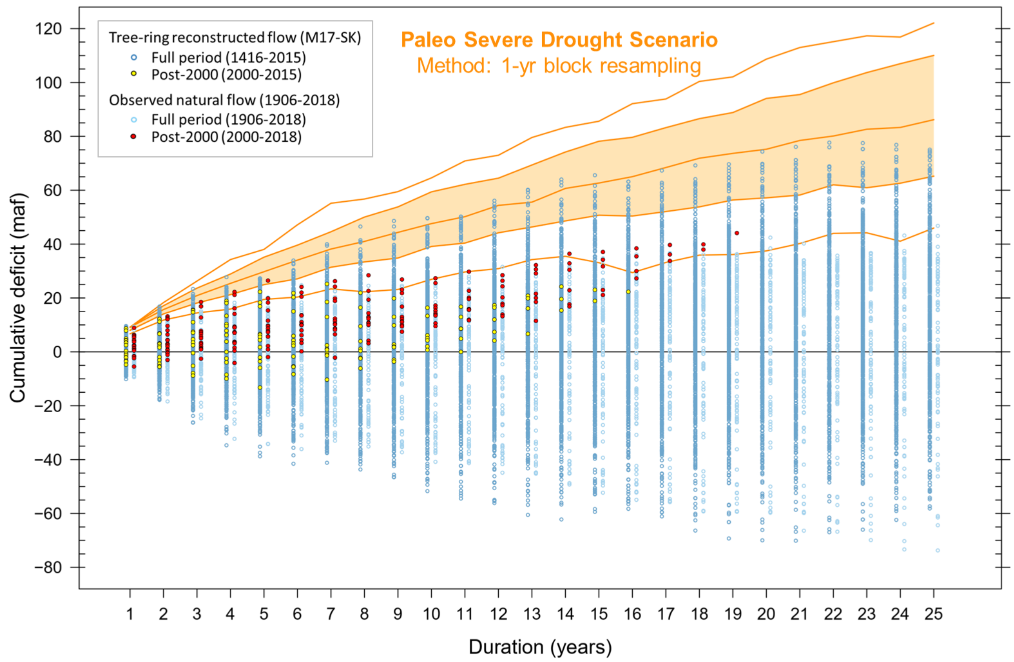


Figure S-14. Cumulative deficit versus duration for the flows simulated from the **paleo severe drought** scenario using **1-yr block resampling** (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

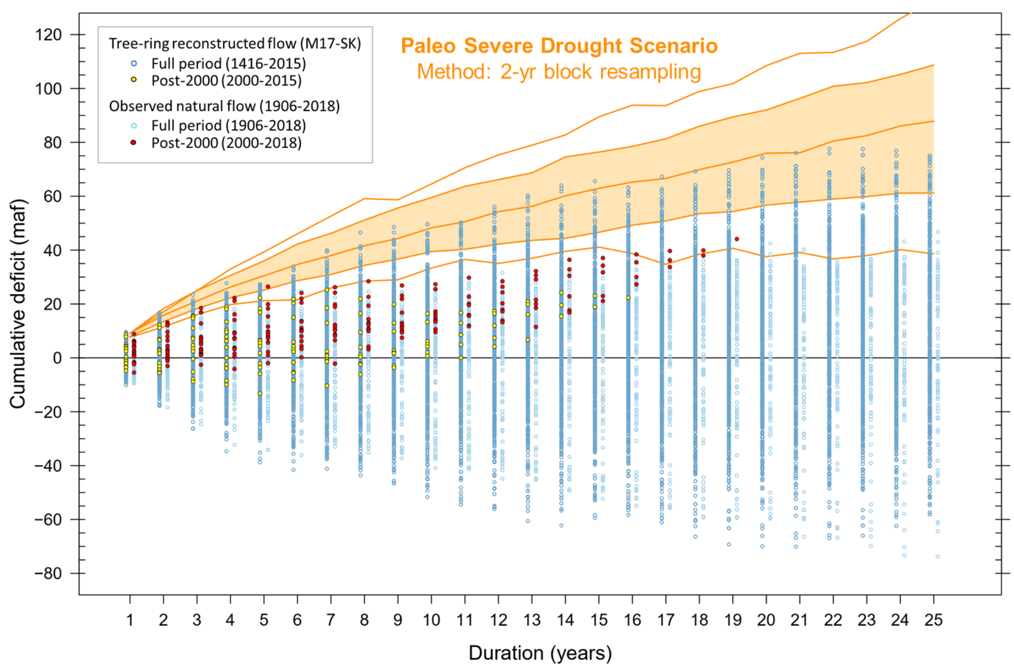


Figure S-15. Cumulative deficit versus duration for the flows simulated from the **paleo severe drought** scenario using **2-yr block resampling** (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

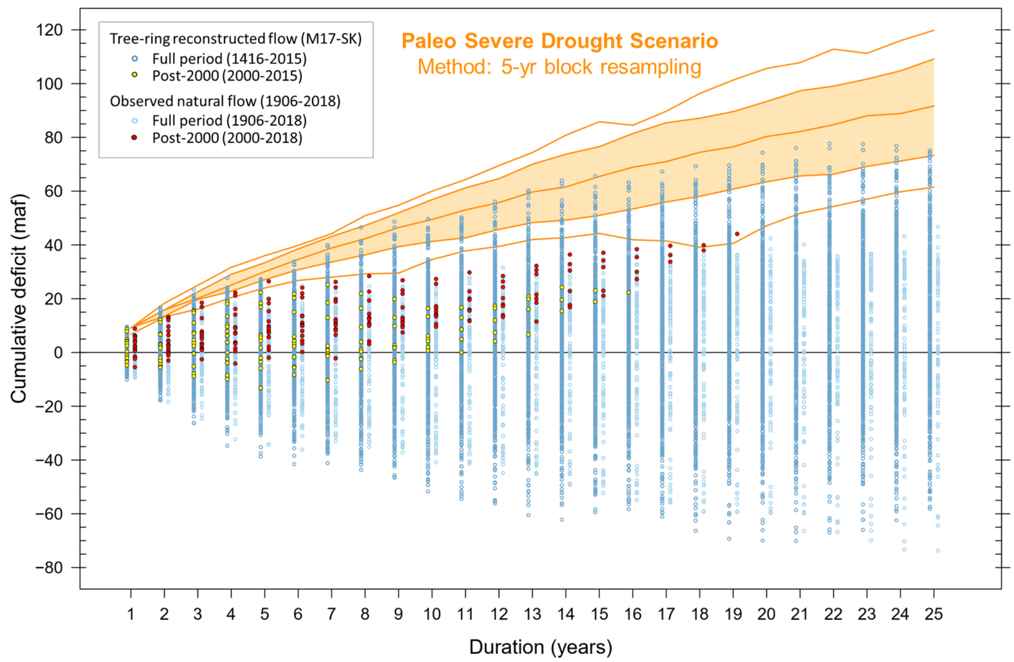


Figure S-16. Cumulative deficit versus duration for the flows simulated from the **paleo severe drought** scenario using **5-yr block resampling** (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

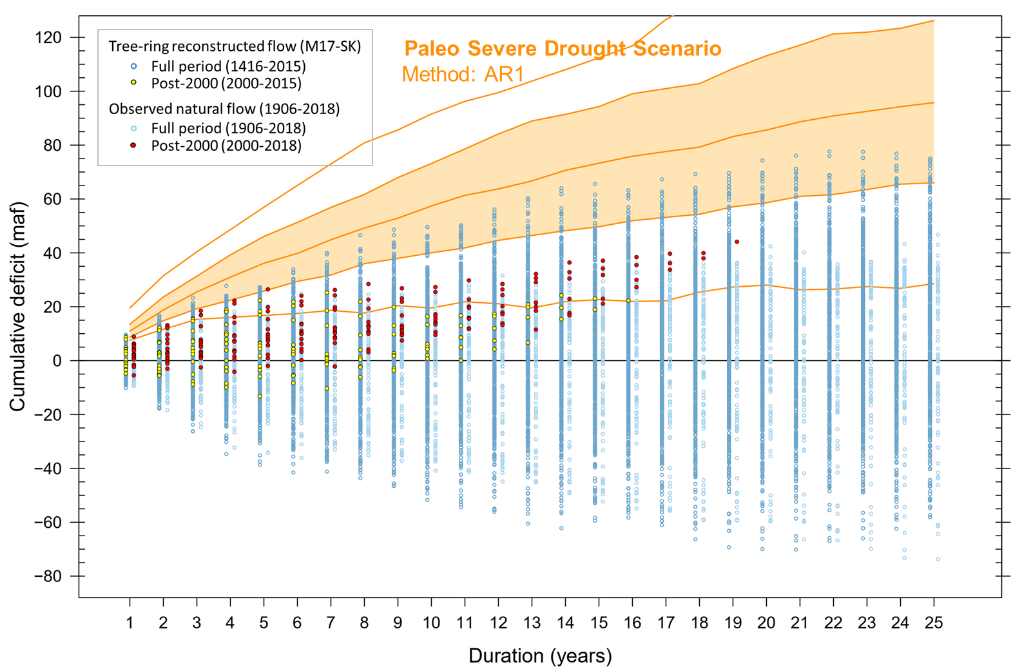


Figure S-17. Cumulative deficit versus duration for the flows simulated from the **paleo severe drought** scenario using **AR1** (the orange lines and area) in the context of observed and tree-ring reconstructed flows (the dots). The orange lines show minimum, 10th percentile, median, 90th percentile, and maximum ranges of the simulation. The orange area highlights the 10th to 90th percentile.

2- Cumulative Deficit Analysis for Tree-Ring Reconstruction

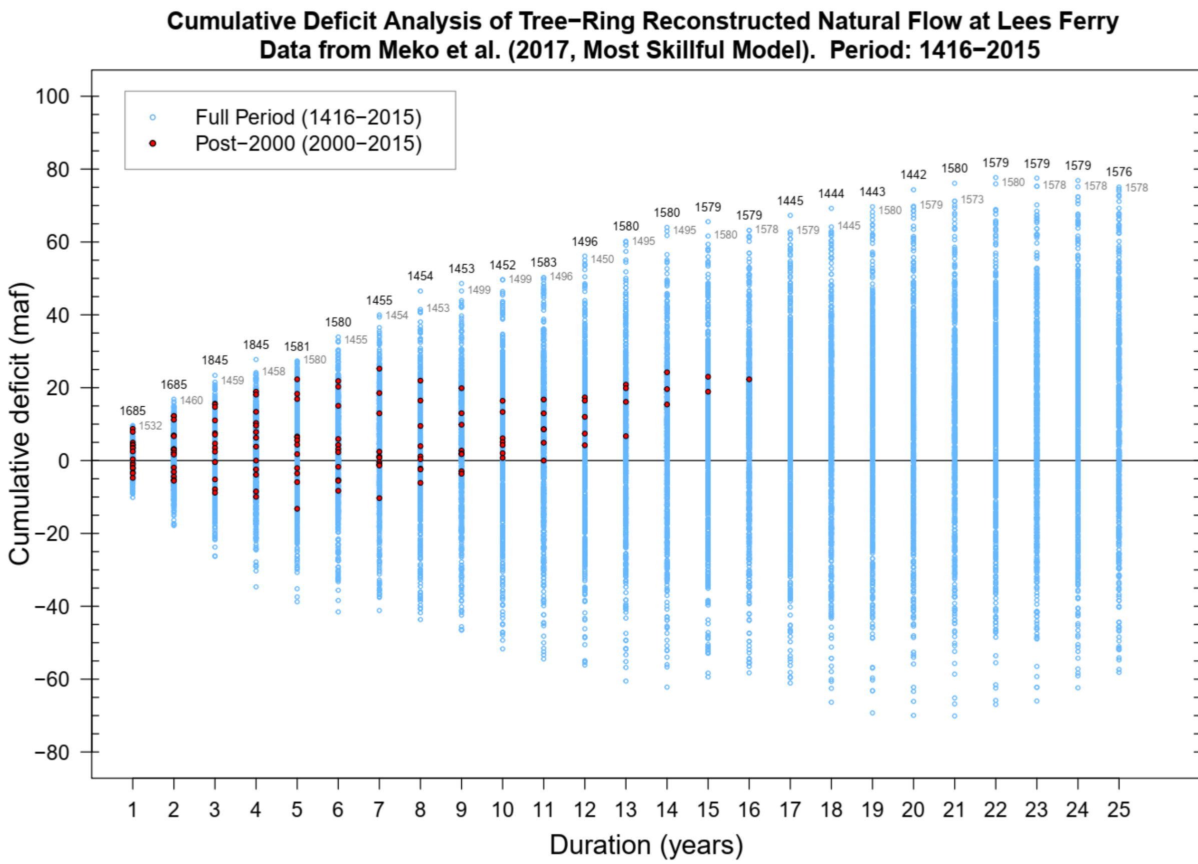


Figure S-18. Cumulative deficit analysis of the tree-ring reconstructed flow (Meko et al., 2017) of the Colorado River at Lees Ferry. Each dot represents water year mean annual flow deficit with respect to the 1906-2018 average of 14.76 maf/yr aggregated over the duration on the x-axis. There is a dot for each duration (including overlaps) within the record. Dot labels give the start year of the highest (black number) and second highest (gray number) cumulative deficit for each duration. The spread of the dots for each duration characterizes how cumulative deficit may vary for different durations.

3- Mean and Standard Deviation of Monthly and Annual Streamflow Simulations

This section presents additional statistics for each of the three drought scenarios (millennium, mid-20th century, and paleo tree-ring drought scenarios). Each scenario is comprised of 100 traces with 42 years of monthly streamflow for 29 sites in the Colorado River Basin disaggregated from annual flows using block disaggregation. The ranges of mean and standard deviation of monthly and annual streamflow simulations at Lees Ferry are shown in Figure S-19 to Figure S-24 below.

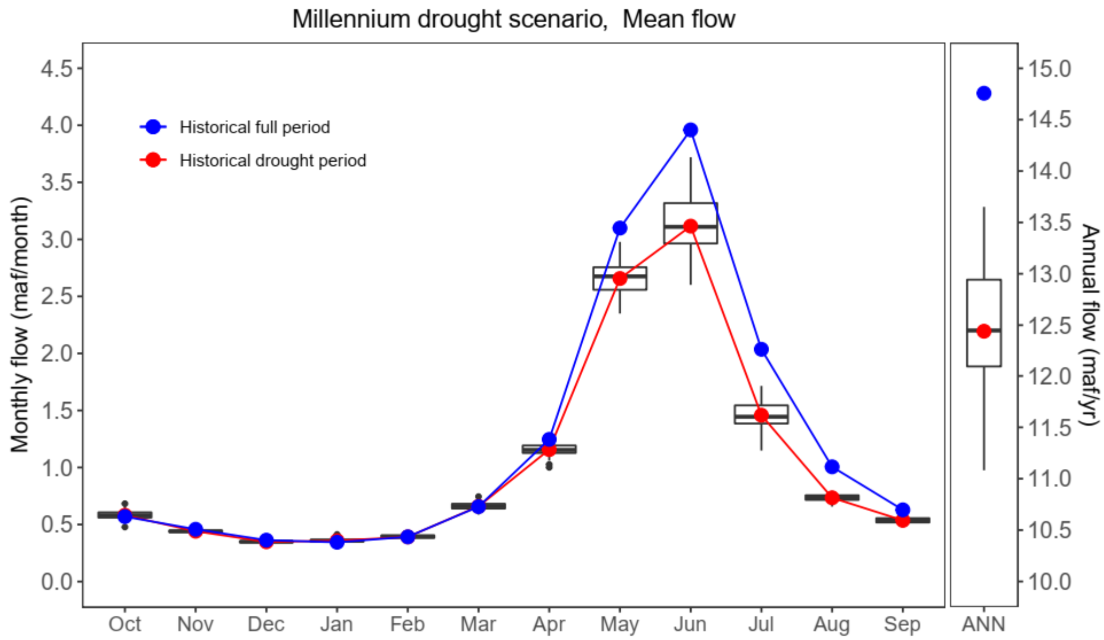


Figure S-19. Comparison of the monthly and annual mean flows of the millennium drought Lees Ferry flow simulations with the mean flows of the historical full period (1906-2018) and historical drought period (2000-2018)

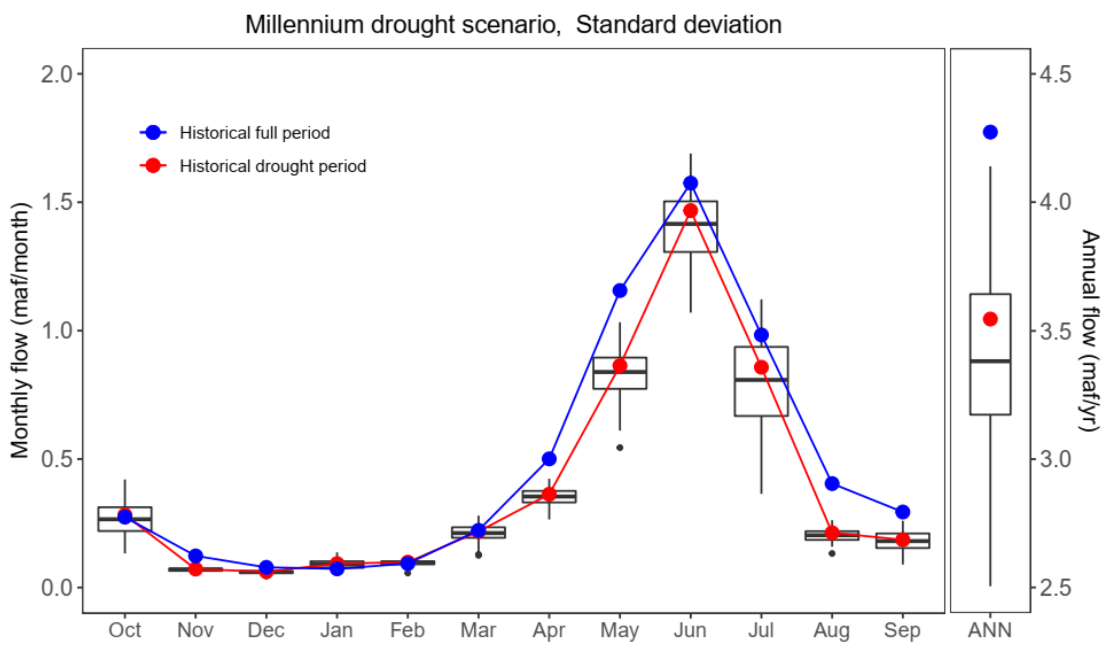


Figure S-20. Comparison of monthly and annual standard deviations of the millennium drought Lees Ferry flow simulations with the standard deviations of the historical full period (1906-2018) and historical drought period (2000-2018)

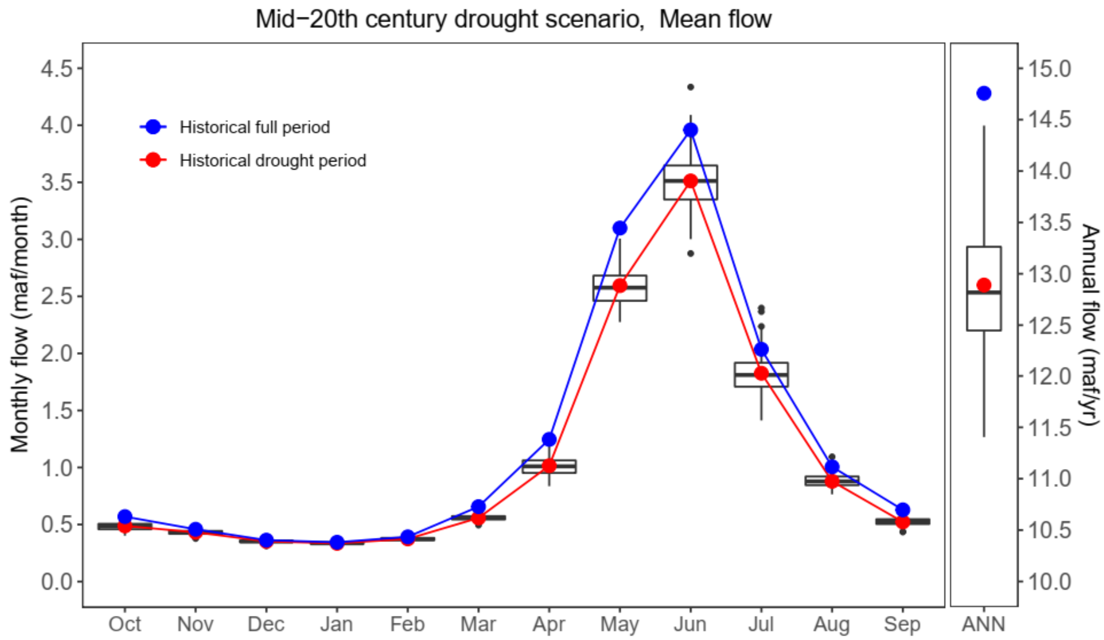


Figure S-21. Comparison of the monthly and annual mean flows of the mid-20th century drought Lees Ferry flow simulations with the mean flows of the historical full period (1906-2018) and historical drought period (1953-1977)

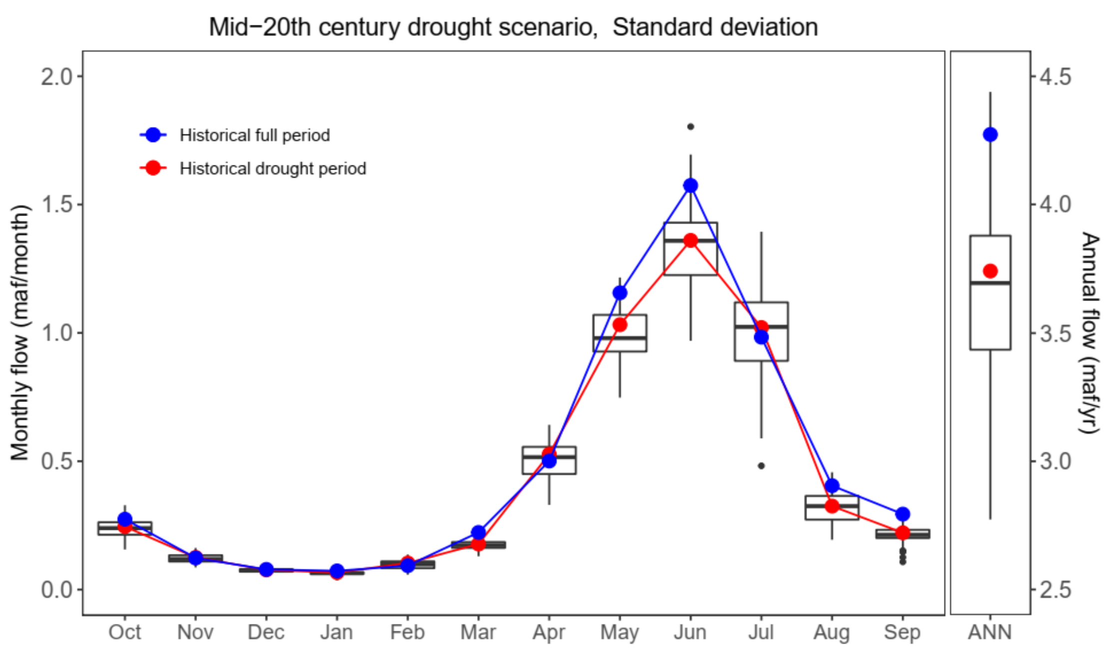


Figure S-22. Comparison of the monthly and annual standard deviations of the mid-20th century drought Lees Ferry flow simulations with the standard deviations of the historical full period (1906-2018) and historical drought period (1953-1977)

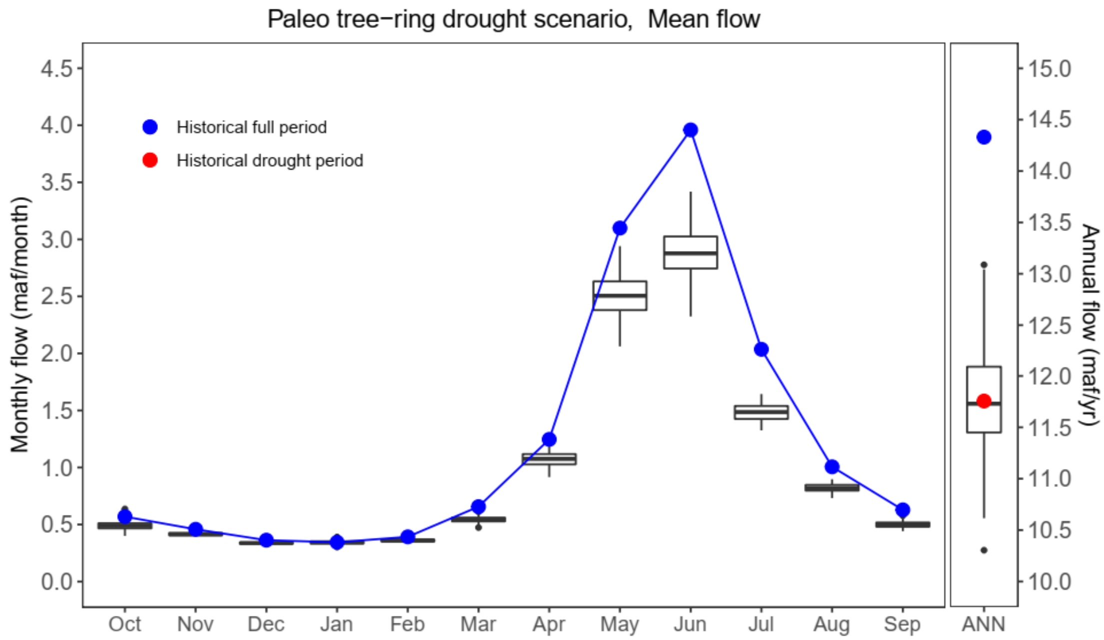


Figure S-23. Comparison of the monthly and annual mean flows of the paleo tree-ring drought flow simulations at Lees Ferry with the mean flows of the historical full period (1906-2018) and historical drought period (1576-1600).

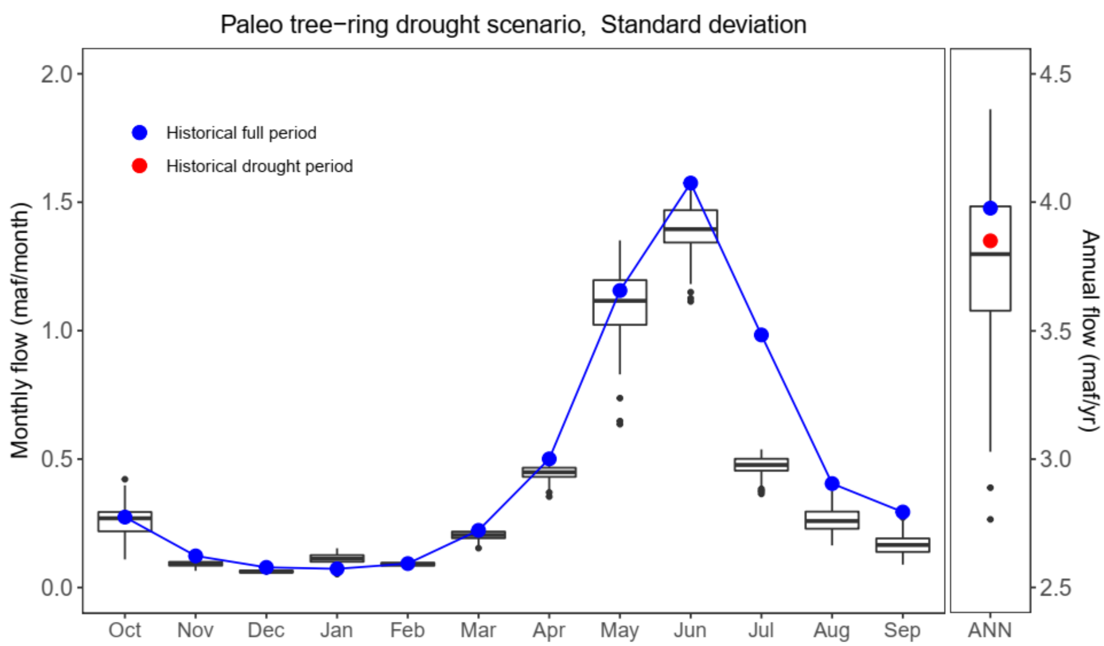
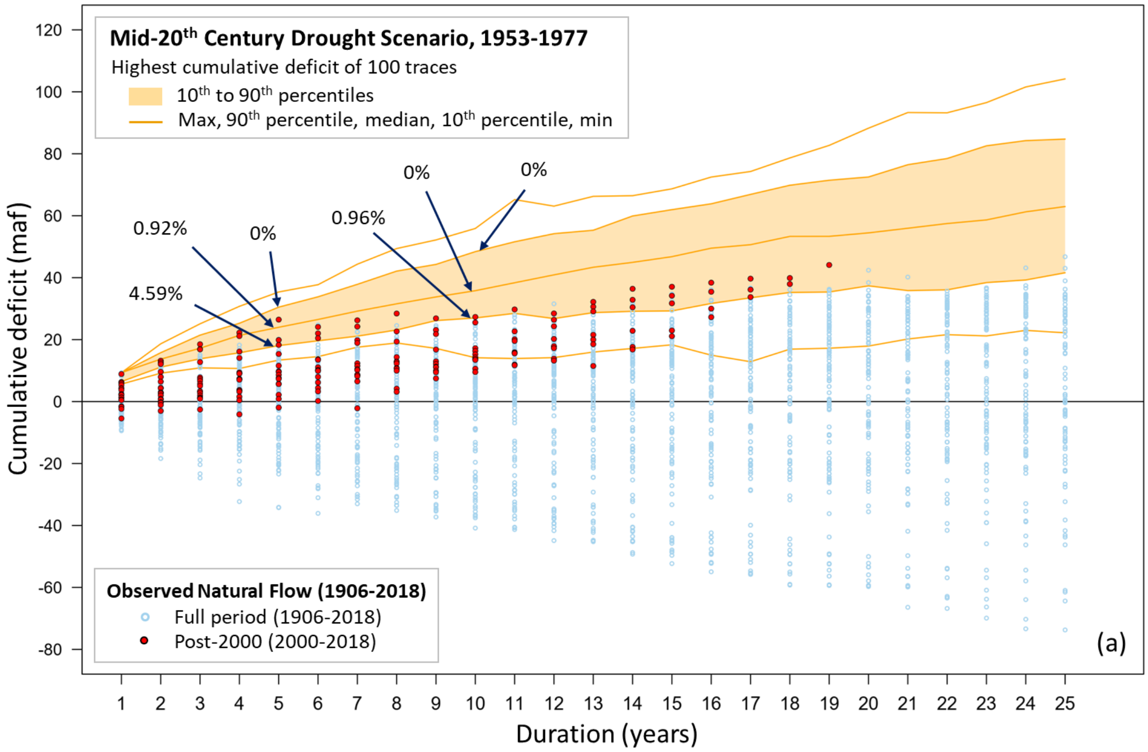


Figure S-24. Comparison of the monthly and annual standard deviations of the paleo tree-ring drought flow simulations at Lees Ferry with the standard deviations of the historical full period (1906-2018) and historical drought period (1576-1600).

4- Highest Cumulative Deficit Results for Mid-20th Century and Paleo Tree-Ring Drought Scenarios

The paper presented the highest cumulative deficit range for the Millennium drought scenario depicting the variability of the maximum of the cumulative deficit (relative to the mean flow of 14.76 maf/yr from 1906 to 2018) for different durations based on each of the 100 simulated flow traces (paper Figure 11). Here similar results are given for the mid-20th century and paleo tree-ring drought scenarios (Figure S-25 and Figure S-26).



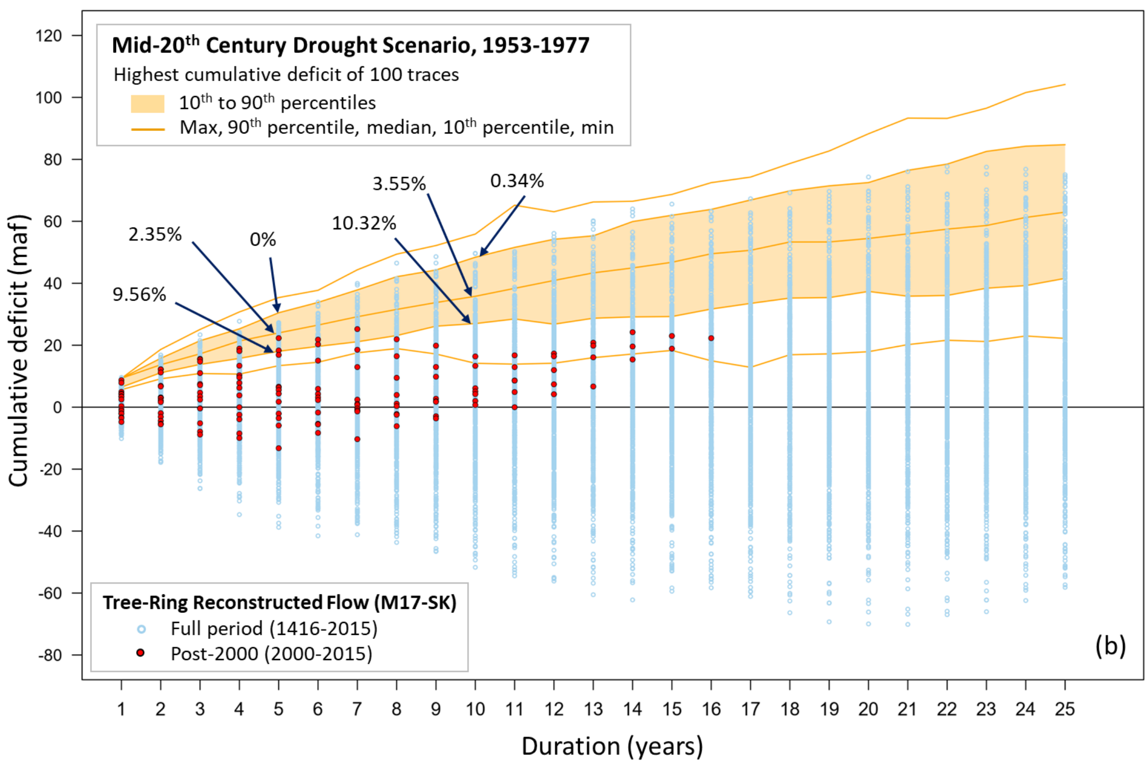


Figure S-25. Highest mid-20th century drought cumulative deficit compared with a) historical, and b) tree-ring reconstructed natural streamflow cumulative deficit.

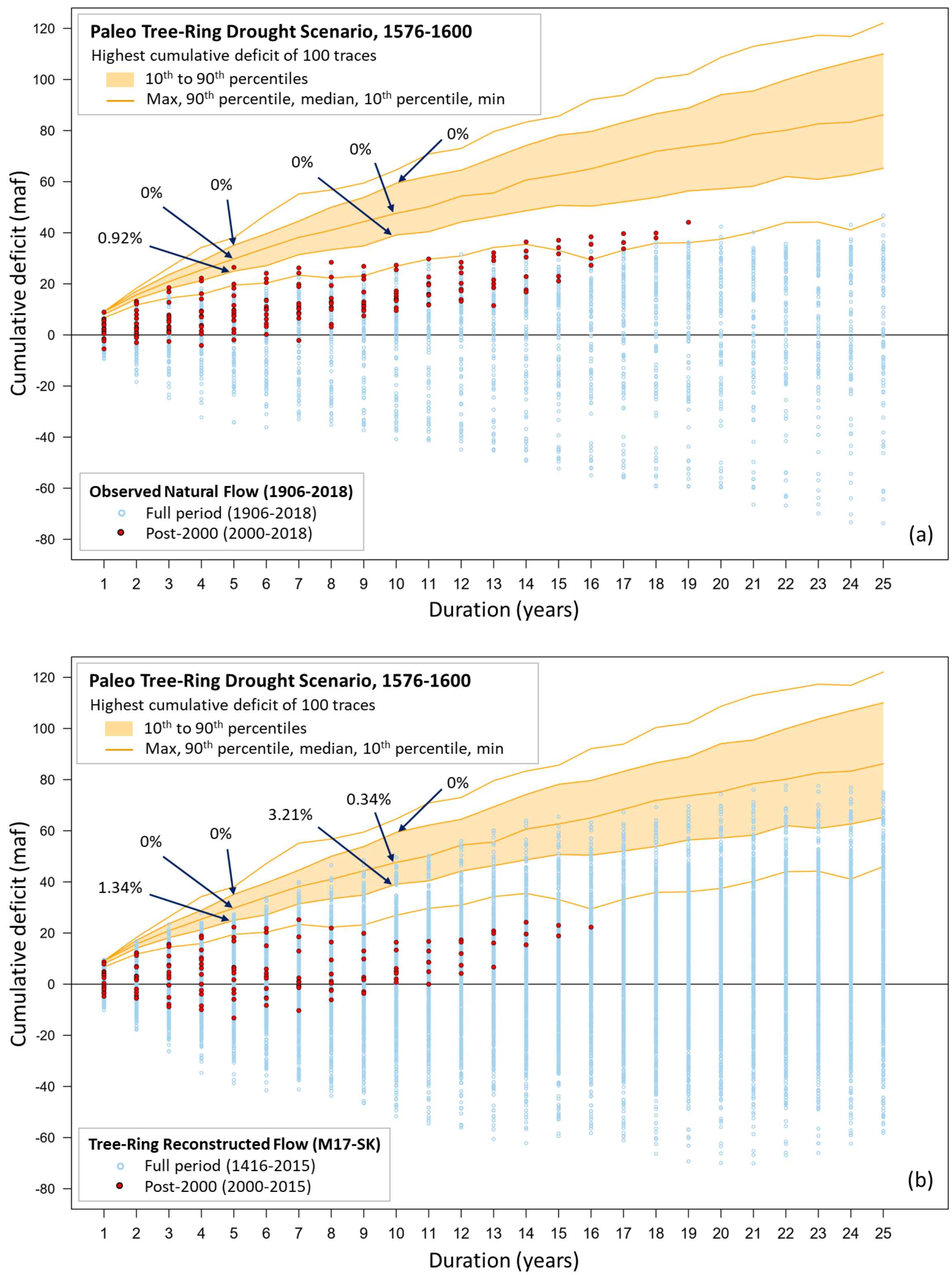


Figure S-26. Highest paleo tree-ring drought cumulative deficit compared with a) historical, and b) tree-ring reconstructed natural streamflow cumulative deficit.