JULY 2023

Independent Peer Review of the State Water Project – Delivery Capability Report

An individual letter review for the Delta Science Program

Prepared by

Dr. Daniel Feldman – Lawrence Berkeley National Laboratory

Delta Science Program

DELTA STEWARDSHIP COUNCIL

Table of Contents

General Comments

The report entitled "Evaluation and Adjustment of Historical Hydroclimate Data: Improving Representation of Current Hydroclimatic Conditions in Key California Watershed," which was prepared by the California Department of Water Resources for the State Water Project - Delivery Capability Report for 2023, performs a thorough analysis of historical hydroclimatic data across the State of California for the watersheds that are critical for water resources for a very diverse set of needs across the State Water Project (SWP) and Central Valley Project (SVP).

The rationale for this thorough analysis is that the assumptions of stationary statistics for hydroclimate data may not be tenable (see Milly et al, (2008) for a high-level discussion), and developing science-driven plans that are based on computer-model projections instead of previous observations requires very careful planning. The first step to this planning is to ensure that the use of historical data is well-positioned to be compared against a range of future projections. This document seeks to take that step. It does so by determining the extent to which historical observations of California's hydroclimate already contain discernible impacts from climate change that would materially impact planning for how future changes in California's hydroclimate impact the SWP, CVP, and other major efforts by the California Department of Water Resources (DWR) to support the diverse water resource interests across California.

The document undertakes a large number of analyses of hydroclimate data and uses these to develop a case for producing a focused set of findings related to the need to make adjustments to the historical time-series of a metric built from historical observations of water resources from major rivers that feed the SWP and CVP called Full Natural Flow (FNF).

The results of the analyses and most of the logic presented in the case to make historical adjustments to FNF are ultimately tenable, given that the analysis convincingly shows that the standard deviation and coefficient of variation in FNF has changed in recent decades for interannual flows and the seasonal timing of FNF has changed for sub-seasonal flows. Planning for a future with different hydroclimates would be more robust if it starts with a baseline of FNF data that is adjusted such that recent changes in FNF standard deviation and coefficient of variation on interannual timescales, and FNF seasonal timing on subseasonal timescales, are removed.

That being said, there is a major comment that should be addressed before the document is released. First and foremost, the document would benefit from some reframing and reorganization, because the logic and rationale for the analyses, and the ultimate reasons for undertaking historical adjustments are not made clear, so readers and users of the report may not be convinced that adjustments to historical data are needed.

The authors of the document should recognize that the burden of proof is on them to convince a skeptical audience that adjustments to historical observations set the foundation for better preparation for climate resilience. Historical observations are the primary and ultimate source of information about California's hydrometeorology and hydroclimatology, so strong justifications are needed for making any adjustments to well-established observations and observational metrics. The findings of this document do support the making of such adjustments, but the document is not organized now to actually make the case for those adjustments.

A reorganization of the document would help make that case by approaching the problem more holistically. A suggested way of doing this is to start the document building off of the background information that is already there, then discussing the role of hydroclimate data and providing a background on the advantages and disadvantages of adjusting historical observationally-based data. The provision of simple, even toy examples, of cases where historical data should be adjusted would be very helpful for readers who would and should be skeptical of adjusting data.

Then, the document should address the primary driving factors and what the goals are that would motivate and justify such adjustments. Then, the document would greatly benefit from a section that frames all the analysis that is performed: why are certain datasets analyzed in the ways that they are. This would be followed by specific descriptions of the methods and why they would or would not be appropriate for adjustments.

Together, such reframing and reorganization of the document will be helpful to justify to a skeptical audience why it would possibly make sense at all to adjust observations in a rational way in order to develop resilience to climate change impacts on water resources for the future.

Second, the analysis presented in the document should recognize, discuss, and consider focusing on temperature and precipitation datasets that are homogenized (e.g., Easterling et al., 1996 and Menne and Williams, 2009). Recent work (e.g., Charn et al, 2022) showed that trends in temperature can very much be impacted by the choice of whether or not temperature products are homogenized.

Third, there are a large number of clarifying comments that should be addressed to improve the document and they are listed below:

Specific Comments

Page 1, first and second paragraphs: Describe what you mean by "baseline."

Page 1, second paragraph after the phrase “now considered non-stationary”: Cite Milly et al, 2008 (Stationarity is Dead: Whither Water Management) as a driver for this analysis.

Page 1, last paragraph: The end of this paragraph is better posed as a question: is there sufficient evidence, between the historical observations and the consensus of climate model projections of the impacts of climate change on California hydrology to warrant a modification of the datasets utilized for setting a historical baseline for the SWP and CVP from historical datasets alone to adjusted historical datasets that take into account the nonstationarity of California’s hydroclimate during the period of observations?

Page 3: Change “The precipitation falling as rain …” to “Generally speaking, the precipitation falling as rain …” to account for the occasional mid-winter heatwave that leads to snowmelt.

Page 5, second paragraph: Clarify that the San Joaquin Valley watersheds are fed by the southern Sierra and that some of the watersheds of the Sacramento are fed by the Sierra, but some are fed by different mountain systems.

Page 5, third paragraph: Change from “more affected by snow” to “more snow-dominated”. Also, a brief statement about the differences in precipitation and evapotranspiration between the northern and southern Sierra Nevada may be warranted here.

Page 5, last paragraph: "These three basins are important surface water supply sources for the SWP and CVP, respectively." This sentence needs to be amended, since there are three basins and two projects being mentioned, so the use of the word "respectively" does not provide one-to-one assignment of basins to supply sources for the SWP and CVP.

Figure 2 caption: reiterate the California Data Exchange Center station IDs.

Page 7: What do you mean by "un-modified historical precipitation"? It suggests that historical precipitation values were used, but there was some consideration to adjust these values based on temperature? If that is the case, then replace "un-modified" with "actual".

Page 8: The approach for calculating daily max temperature, daily temperature range, and daily minimum temperature is consistent with many others, including what is done with LOCA and LOCA2. It may be worthwhile to include this point.

Figure 3: Please change the color of the lines to contrast more: purple and blue are difficult to differentiate. What is meant in the title of Figure 3 by “Key Reservoirs”? All of the reservoirs listed above or something else? What years are being covered here? Is this an average, or a single water year? There does need to be information on which water years are being plotted, and if more than one, shading to include a range of results so that the reader can see if there are significant differences between the hydrographs. Also, the sub-title of “Sum of All Inflows modified VIC Bias Correction Period (70-03)” is esoteric and likely meaningless to most readers. Please clarify or remove. Also, please include water year totals.

Page 8: An explanation for why temperature detrending changes the timing, and to a lesser extent, the amount of the VIC-calculated hydrograph in Figure 3 is needed. It would be expected that detrending would lead to cooler temperatures as inputs into VIC, because the nominally upward-sloping temperatures would be removed at least for part of the record. Cooler temperatures would tend to lead to later runoff and lower peak winter runoff, all else being equal (which is a big assumption). The result in Figure 3 is non-intuitive, anyways, and requires a high-level explanation of why temperature detrending leads to the change in modeled hydrograph shown.

Page 9, last paragraph: Please define the acronym DSC VA, in case this differs from DWR VA. Also, a statement needs to be made about whether a -1% difference in the annual sum of flows is significant for the CWP or SVP. How does this compare to the skill of the VIC simulation (which could be roughly estimated by comparing the VIC simulation with and without calibration)?

Figure 4: What is meant by “11 Rim Inflows” in the figure subtitle? As with Figure 3, what is meant by “Key Reservoirs”? The ones listed in this report?

Page 10, Figure 4 caption: What is meant by “the hydrologic model”? SAC-SMA? Also, water year totals are needed in the table below Figure 4. Finally, there does need to be information on which water years are being plotted, and if more than one, shading to include a range of results so that the reader can see if there are significant differences between the hydrographs.

Page 10: As with Figure 3, an explanation for why the detrended temperature forcing for the hydrologic model produced enhanced wintertime flow and reduced springtime flow is needed. Also, are the differences in the hydrographs between Figure 3 and 4 reflective of the differences in hydrologic models (presumably VIC vs SAC-SMA)?

Page 11: The Section “Analysis of Detrending Studies Conclusion” needs more information in it. It needs to explicitly state the differences in hydrographs associated with different detrending methods, whether the differences are significant between detrended and non-detrended temperature inputs for a given study or for all of the studies, and why this led to the conclusion to use FNF data. Presumably the use of FNF data introduces less uncertainty than the use of hydrologic models with and without temperature detrending. However, the impacts of temperature on hydrology are implicit.

Page 12: The challenges listed for using hydrologic models require citation support, either in the form of a report, a peer-reviewed publication, or even a personal communication.

Page 12: The challenge that “evapotranspiration” poses to the use of hydrologic models requires clarity. Is this report claiming that “Differences in evaporation simulation in hydrologic models appears to be a key driver of differences in runoff timing and amount”?

Page 12: The challenge presented that using hydrologic models “was beyond the scope of this effort” is misplaced: it is not necessarily the case that one or any hydrologic model will be able to capture both the observed FNF data along with temperature and precipitation sensitivities because any hydrologic model will require extensive bias-correction to produce realistic FNF data, and that can produce uncertainty in the sensitivity of specific water resources to temperature and precipitation.

Page 12: The justification that because streamflow is an aggregate measure of climatological changes from hydrology, that FNF data should be used does not make sense. A better justification is that the FNF data is strongly grounded in timevarying observations from measurements with stream gages, discharge volumes, reservoir heights, etc.

Page 12: The justification that FNF data should be used because “streamflow allows for more simplified statistical manipulations of historical data to represent current conditions” does not make sense. A better justification is that FNF data requires very little, if any bias correction, and bias correction complicates any statistical analysis.

Page 12: Regarding the working group statement, the document needs to make clear, or at least have a clear reference to, the 63 CalSim2 watersheds and the 201 CalSim3 rim watersheds.

Page 13: The phrasing of “Unambiguous enough” juxtaposed with the discussion of the significance value of the Mann-Kendall test at $p<0.05$ is confusing. Is this paragraph trying to convey that analysis of variables where the significance threshold of $p<0.05$ is achieved are “unambiguous enough”?

Page 13: More information and discussion is needed in the section titled “Basin Averaged Temperature and Precipitation,” especially regarding temperature. One minor point: do the topographic adjustments also include temperature? If so, how was that done? Spatial interpolation of temperature in complex terrain should account for temperature lapse rates (which are curiously but persistently small, as shown in Minder et al, JGR, DOI:10.1029/2009JD01349). A more major point is that there are major potential differences between temperature datasets depending on which is used. A recent paper by Charn et al. (2022) showed that historical trends in temperature in California can be of different signs (!!) depending on which dataset is used, and this is largely reflective of whether or not the temperature datasets have been homogenized. Homogenization may or may not be warranted for this analysis (likely it is), but a discussion is needed of how historical temperature datasets do have artifacts and those do need to be accounted for.

Figure 5: Use thicker lines so that any differences between FNF data is more easily seen by the reader.

Page 15: Do the paleo streamflow reconstruction time-series used cover the watersheds of the Sacramento and San Joaquin Valleys that are analyzed here?

Page 15: Statements about the independent validation of paleo streamflow data would be helpful. This could take the form of simply citing community-wide validation efforts (e.g., PAGES Hydro2k Consortium 2017) or discussing them in some detail so that the reader has more information on how robust these datasets are (or are not) and how they should be properly used.

Figure 10: Move most of the description of the figure on the last paragraph of page 19 to the caption for Figure 10. The caption should read something like "Reconstructed annual flows from the four rivers of the Sacramento watershed (faint grey), 30-year rolling mean annual flow (red line), and +/-1 30-year standard deviation from the mean flow from 1500–2011 (black solid line). A linear smoothing of the 30-year rolling mean annual flow +/-1 30-year standard deviation from the mean flow (blue lines). The dashed orange line highlights the maximum and minimum historical levels of variability prior to 1950." Then, include a description of Figure 10 results either before or after the figure. This description would include information about the envelope between 1500 and 1950, and the distinct increase in interannual variability after 1950.

Page 21: When analyzing runoff efficiency, does the calculation of a double mass analysis look at cumulative precipitation and runoff for a given water year or some other period of time? This period of time over which the cumulative observations are gathered needs to be established. A choice of cumulative analysis for time intervals longer or shorter than a water year would require justification.

Page 22: Typo in first line, should read "...support an explicit adjustment of runoff to account for this shift."

Page 26: The finding of statistically-significant increasing standard deviation and coefficient of variation in FNF also has some level of theoretical support as an emergent climate-change phenomenon that is expected to be comparably detectible to changes in temperature (see Pendergrass et al. 2017 for details and consider citing that paper).

Page 26, Paragraph that begins with "Historical Runoff Adjustments": The phrase "without imparting additional unwanted trends or modifications" requires revision. This phrase is too vague as it is written. A suggested change is to focus on historical runoff adjustments that preserve, or at least do not substantially corrupt estimates of the unforced modes of variability in California's hydroclimate as they impact runoff.

Page 30, last line: did you mean "Table 7"?

Page 31, Table 7: It is difficult to understand the logic that uses the results from the metrics listed in Table 7 to reach the result that RC-YTM is the superior method for adjusting the FNF data.

Page 31, typo: Should read "A description of the RC-YTM method …"

Page 38: It is unclear how to evaluate the findings on additional adjustments to CalSim inputs (Temperature, Precipitation, Demand, WYT) since this area is incomplete. Suggestions regarding temperature, precipitation, and demand adjustments involve using well-established detrending techniques. One well-established approach is described in Wu et al. (2007).

Charge Question #1

Is this method an improvement over the use of unadjusted historical data (i.e., an assumption that the historical time series is stationary) for representing current conditions? Why or why not?

Response to Charge Question #1

The methods presented here are heavily-grounded in historical observations and look at a wide variety of datasets and analytic techniques. Ultimately, the methods seek to use these datasets to address the problematic assumption of stationarity in historical time-series of data that inform water management in California. By looking for significant changes in the historical data of first and second-order statistics of key variables related to water resources in California, the methods first determine whether issues of stationarity, or lack thereof, apply to the problem at hand. They find that the standard deviation and coefficient of variation for a metric of discharge from major rivers that serve California's State Water Project and Central Valley Project have been changing. Therefore, to create a baseline of data to plan for climate change impacts of the future, it is appropriate to adjust the historical data in such a way that significant changes in river discharge statistics due to climate change are not present in the baseline historical data.

Without such adjustments, some of the impacts of climate change may be counted incorrectly. Specifically, analyses of future projections of water resources in California that are interested in changes in the multi-year standard deviation and coefficient of variation in metrics of discharge from major rivers will be problematic without historical baseline data that is adjusted to be as stationary as possible. To be blunt, a comparison between projections and unadjusted historical data would actually underestimate the impact of future climate change on metrics of interest where climate change has already occurred because unadjusted data would show a smaller change than adjusted data. The primary reason that this smaller change is unrealistic is that climate change impacts are already happening, so any differences in metrics of interest between historical data and projections need to account for any climate change impacts that have already occurred.

The very detailed analyses of the different adjustment methods were generally convincing in that the adjustments at least were able to reduce the standard deviation and coefficient of variation in the historical metrics of river discharge. That being said, the rationale of why methods were used for the adjustment of historical data was unclear other than ultimate performance relative to (well-posed) criteria established for such adjustments (namely that the adjustment reflects recent, significant trends without corrupting or biasing historical data in unintended ways). The result is that the adjustment methodology is entirely phenomenological.

Charge Question #2

How well does the new method account for statistically significant trends to represent a quasi-stationary current climate while avoiding bias or trends that are artifacts?

Response to Charge Question #2

The new method does account for statistically significant trends to represent the quasi-stationarity of the current climate. The new method also recognizes that there is some degree of stationarity in the current climate, but some level of non-stationarity as well. The important aspects of the analysis presented here is that they seek to address the non-stationary components in addition to the stationary ones.

While the time-series of adjusted and unadjusted FNF data do appear to show differences that are consistent with making a more consistent level of interannual variability in FNF in the adjusted data, more information would be helpful on whether the use of the RC-YTM method does actually remove trends in the standard deviation or coefficient of variation of the FNF data on interannual time-scales, and if the RC-YTM method removes trends in seasonal FNF timing and amount on sub-annual time-scales. Explicit explanations of why the RC-YTM method is chosen to remove trends in unadjusted FNF statistics would be most helpful.

Charge Question #3

What specific investigations or improvements should be considered in future updates of this dataset?

Response to Charge Question #3

First, an explanation of why the RC-YTM method is able to remove trends in the standard deviation and coefficient of variation in FNF data would be very helpful. Second, an inclusion of figures that show that trends in standard deviation and coefficient of variation in adjusted FNF data are indistinguishable from zero would also be very helpful.

Finally, the analysis of the temperature and precipitation datasets should be updated to include records that have been homogenized to remove any artificial trends or lack of trends associated with changing instrumentation, changing conditions (e.g., land use) near the measurements, other known and/or diagnosed measurement artifacts, and any other known deficiencies in the observations.

Charge Question #4

How frequently should DWR consider updating this dataset?

Response to Charge Question #4:

The dataset developed here uses methods that are minimally impacted by river discharge, precipitation, temperature, evapotranspiration, or other methods in a given water year. The use of a 30-year window for low-pass filtering of metrics of river discharge is tenable, especially for identifying trends and ultimately detrending data. At the same time, the use of such a window suggests that updates to the dataset can be quite infrequent: no more than every 5 years, and probably closer to every 10 years.

However, there could be events that arise in the near-future that do warrant a revisitation of this dataset. Specifically, the analyses presented in the document that focus on runoff efficiency may need to be revisited on a 3-5 year basis. If the runoff efficiency changes for the watersheds presented here, there will need to be a much more detailed discussion about the role and value of historically-adjusted data for planning for a California with hydrologic function that is fundamentally dissimilar from the past.

Charge Question #5

The draft Climate Adjusted Historical Hydrology dataset presented for review is adjusted to a 1992-2021 climate condition. This period is entirely retrospective. With a goal of more accurately simulating the range of hydrologic variability under current climate conditions, what are the pros and cons of taking a more prospective approach in future iterations by, for example, including modeling of potential future conditions to capture a 30-year climate period centered on the current year rather than concluding with the current year?

Response to Charge Question #5:

The primary benefit of considering future conditions as part of a baseline for analysis of water delivery capabilities is that there may be latent hydrologic variability that is not contained within the last one hundred years, but which may be impactful to water delivery capabilities. There are many spatial and temporal modes of variability in temperature, precipitation, surface, and subsurface hydrology throughout California and the northeastern Pacific Ocean which are responsible for the observed hydrologic variability. Paleo reconstructions of streamflow attempt to capture some of the variability that has occurred in California, and they paint a striking picture of precipitation and drought extremes, not just from storms but for multiple years, that simply have not occurred in the last 100 years. Analyses that ignore the proxy evidence of the much larger hydrologic variability that has occurred before 100 years ago will not be helpful for planning for extended precipitation and drought extremes in California. Therefore, analyses that do take into account more hydrologic variability, which may be captured by near-term future projections, will enable planning that is more cognizant of the range of hydrologic variability that can occur in California, regardless of whether that variability is or is not forced by climate change.

There are disadvantages to considering future conditions as part of a baseline for analysis of water delivery capabilities, however. First and foremost, the skill with which climate models are able to produce probabilistic multiyear predictions of precipitation and temperature in California is low. It is driven by dominant oceanatmosphere modes of variability in the climate system such as the El Niňo Southern Oscillation (ENSO), the Pacific Decadal Oscillation (PDO), and the Atlantic Multi-Decadal Oscillation (AMO), and the phases of those modes cannot be predicted with certainty on multiyear time-scales. Therefore, the direct use of climate model projections on decadal time-scales for California hydroclimate planning is very challenging.

Because of the low skill in hydrometeorological predictions at multiyear time-scales, the question of what information to use for establishing a baseline immediately arises. Does it make sense to include dry years, wet years, or a mixture of the two? Which, if any, climate model projections for the next few years should be used? The answers to those questions are not clear. Reasonable arguments can be made for including only dry years, or only wet years, or a mixture of the two in the future data that support the construction of a baseline here. Additionally, reasonable arguments can be made to support the use of all climate model projections for the next few years to construct a baseline instead of sampling from periods outside of the last 100 years, as can arguments be made to use a handful of climate model projections of only those models that reproduce California's observed hydroclimate in their historical simulations, as can arguments be made to use just a single, highly-performing climate model projection. The point is that there is unresolvable uncertainty in what data should be used as a baseline without a reframing of the goals of what is desired from including data outside of the last 100 years as part of the baseline.

To that end, one potentially defensible strategy for creating a baseline with data outside of the last 100 years is to include estimates of metrics of river discharge (FNF), drawn from paleo reconstructions, that include the water years that would be most problematic for managing the SWP and CVP. This would essentially introduce a very realistic (since it actually occurred), worst-case scenario for baseline analyses. If such worst-case scenario baseline planning is too conservative (i.e., planning for much more extreme events outside of the last 100 years), then including some of the more problematic water years that have occurred outside of the last 100 years, but which did not occur in the last 100 years, would be warranted.

Alternatively, analyses of the decadal climate predictions from the Coupled Model Intercomparison Project - Phase 6 (CMIP6), especially from the Decadal Climate Prediction Project (DCPP) could be performed but would again require a clear set of goals regarding what is desired from including such data. If, as indicated above, the goal is to establish a baseline that does not underestimate actual hydroclimate variability in California, then using those DCPP results for the water years that are most problematic for the SWP and CVP should be considered. Model sub-selection and/or weighting would need to be considered and formally addressed for any such activity that uses near-term climate model projections, and there are many schools of thought for appropriate ways of sub-selection and weighting, including some schools that think that no sub-selection or weighting should be performed.

References

Easterling DR, Peterson TR, Karl TR (1996) On the development and use of homogenized climate datasets. *Journal of Climate* 9(6): 1429-1434. https://doi.org/10.1175/1520-0442(1996)009<1429:OTDAUO>2.0.CO;2

Menne MJ and Williams, Jr. CN (2009) Homogenization of temperature series via pairwise comparisons. *Journal of Climate* 22(7): 1700-1717. https://doi.org/10.1175/2008JCLI2263.1

Milly, P.C.D. et al (2008) Stationarity is Dead: Whither Water Management? Science, 319(5863), 573-574. https://doi.org/10.1126/science.1151915

Charn AB, O'Brien TA, Risser MD, Longmate JM, Feldman DR (2022) Sign of observed California temperature trends depends on data set homogenization: Implications for weighting and downscaling. *Geophysical Research Letters* 49(15): e2022GL099186. https://doi.org/10.1029/2022GL099186

PAGES Hydro2k Consortium (2017) Comparing proxy and model estimates of hydroclimate variability and change over the Common Era. *Climate of the Past*, 13(12): 1851–1900. https://doi.org/10.5194/cp-13-1851-2017.

Wu Z, Huang NE, Long SR, Peng C-K (2007) On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proceedings of the National Academies of Sciences* 104(38): 14889-14894. https://doi.org/10.1073/pnas.0701020104.

Pendergrass AG, Knutti R, Lehner F, Deser C, Sanderson BM (2017) Precipitation variability increases in a warmer climate. *Scientific Reports* 7:17966. https://doi.org/10.1038/s41598-017-17966-y