September 2023

Independent Peer Review of the State Water Project – Delivery Capability Report, Part 2

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 "Risk-Informed Future Climate Scenario Development for the State Water Project Delivery Capability Report," which was prepared by the California Department of Water Resources for the State Water Project - Delivery Capability Report (DCR) for 2023, performs a thorough analysis of many of the risks that the State Water Project (SWP) and the Central Valley Project (CVP) may face in the next 20 years due to changes in local temperature, precipitation, and sea-level rise that fall outside of an envelope of observed historical natural variability. Temperature, precipitation, and sea-level rise could differ significantly from what has been observed in the last 100 years, so analytical approaches to assessing SWP and CVP risks based solely on historical observations are likely to underestimate risks to those projects. The report bases its analysis on a survey of downscaled climate model projections using the Localized Constructed Analogs Version 2 (LOCA2) (Pierce et al, 2023) downscaling routine that has been applied to a set of Coupled Model Intercomparison Project - Phase 6 (CMIP6) models (Eyring et al, 2016; O'Neill et al, 2016), which are then sampled and used, in conjunction with probabilistic estimates of sea-level rise, to drive CalSim3 simulations to perform system stress testing for the SWP and CVP for a 30-year period centered around 2043. The choice of 2043 was meant to represent 20 years into the future from this 2023 report.

This report is a companion report which is meant to be complementary to another part of the 2023 Delivery Capability Report that has been written by the California Department of Water Resources entitled "Evaluation and Adjustment of Historical Hydroclimate Data." Both this report and the companion report are the subject of Independent Peer Review.

Overall, there is a significant amount of carefully-considered work that is summarized in this report. This work involves a large number of steps to support the development of a probabilistic approach to assessing the system stresses that the SWP and CVP will be facing in the next 20 years, subject to a large number of constraints, chief amongst which is uncertainty in the surface air temperature and precipitation fields that the State of California (particularly northern and central California) will face over this time period. Still, there is widespread recognition of, and peer-reviewed support for, the concern that the surface air temperature and precipitation fields of the next several decades in California will differ fundamentally from those that have occurred in the last 100 years or even those for which there is historical proxy information. For these reasons, a comprehensive evaluation of what climate model projections specifically mean for the likelihood that there will be different water resource levels available to the SWP and CVP is very much warranted.

At the same time, climate models are blunt tools and provide information that is too far removed from the actionable information that is required for regional, state, and local planning purposes. Climate models provide projections of hydroclimate variables at coarse spatial resolution (~100 km) and sometimes only at monthly temporal resolution, and they exhibit large biases in the historical simulations relative to the observational record due to their sampling of different modes of natural variability in the Earth system that impact hydroclimate (e.g. El Nino Southern Oscillation, Pacific Decadal Oscillation, etc.) and representation errors that may have minimal impact on global model performance but can have major impacts on regional, state, and local model performance. The use of downscaling solutions, especially ensembles of downscaling solutions that capture secular trends in surface air temperature and precipitation, while necessarily retaining realism because the downscaling approach incorporates historically-observed spatial patterns, is well-taken.

Given the mismatch between the information that climate models and downscaling solutions can provide and the information that planners need, plus inherent uncertainty in California hydroclimate projections, this report seeks to manage these challenges in order to develop stress tests for the SWP and CVP, which provide some quantitative (and certainly qualitative) information about those systems' performances in the coming decades. The assessment of system stress due to secular changes in California and Western US-wide hydroclimate is not, and likely will not be for some time, a monolithic analysis. The development of climate model projections, downscaling techniques, tools to assess SWP and CVP system performance, and analysis of these different components are all disciplines in and of themselves and, while they are not being developed completely independently, they are largely separate undertakings.

There are inherently significant analysis challenges associated with developing stress tests for assessing SWP and CVP system performance under a set of different, plausible hydroclimate states. These include: (1) the CalSim3 operational code is a requisite tool for evaluating the implications of hydroclimate for the SWP and CVP, but it is not designed to incorporate secular change in hydroclimate, and

(2) CalSim3 can only be reasonably run a small number of times, so running CalSim3 across an entire ensemble of climate model simulations is infeasible. Another, relatively minor inherent analytical challenge is that sea-level rise is a boundary condition for CalSim3 and there are ranges of projected sea-level rise that marginally complicate the assessment of SWP and CVP system performance where sea-level rise is below a threshold of about 1-2 feet in the California Delta. When sea-level rise goes above a certain threshold, though, salt water intrusion would impact the findings of this report, which this report should note. It appears that a 1-2 foot sea-level rise in the California Delta would be unlikely in 2043, given current projections (Kopp et al, 2014).

To manage many of these inherent analytical challenges, this report rightly employs a multi-pronged approach that includes both bottom-up and top-down approaches to assessing SWP and CVP vulnerabilities and resilience in the next 20 years, and generally describes the rationale for the approaches taken in ways that are defensible.

In spite of the thorough analysis presented here, there is a point that does need to be raised where future analyses would be helpful: the assessment of hydroclimate risk presented here relies on the use of statistically-downscaled climate model products. An independent assessment of the uncertainty introduced from a statistical downscaling solution itself would be helpful to determine if the sole use of statistically-downscaled solutions in this report materially impacts the results. There are longstanding academic discussions and debates about the proper value(s) and role(s) of statistical and dynamical downscaling (e.g., Wood et al, 2004, Wilby et al, 2004; Giorgi et al, 2009; Maraun et al, 2010, Barsugli et al, 2013) and future Delivery Capability Reports do not need to be entrained in those per se, but there are key differences between the statistical and dynamical downscaling projections for the same parent climate model that conceivably could impact the findings of this report. As shown in Figures 3 and 4 of Walton et al, (2020), the patterns of temperature and precipitation change in statistical and dynamical downscaling solutions of future surface air temperature and precipitation for the same model are very different between the low, middle, and peak elevations in California's Sierra Nevada mountains. The dynamical downscaling solution is not necessarily superior to the statistical downscaling solution, but the dynamical solution, which is based on the Weather Research and Forecasting (WRF) model produces very different dynamical and thermodynamical changes in complex terrain, with the WRF model's changes exhibiting patterns that have robust dynamical and thermodynamical features (e.g., Rhoades et al, 2022).

There are inherent challenges to adapting the analysis in the report to incorporate dynamical downscaling climate models: first and foremost, it is unlikely that there will be ensembles of dynamically downscaled climate model results to analyze because they are so much more computationally expensive to produce. In lieu of having ensembles of dynamically downscaled climate model outputs, a more straightforward analysis would involve comparing statistical and dynamical downscaled climate model outputs for the same model to determine if there are consistent spatial patterns to the differences in the two downscaling approaches, and then perform a small number of runs of CalSim3 to establish the sensitivity of that model to the spatial pattern differences between the two downscaling approaches. For example, the findings in Walton et al (2020) suggest that statistically downscaled climate models disagree with dynamically downscaled climate models, where the dynamical approach shows snow increasing at high elevations and decreasing at lower elevations. If CalSim3 produces very different results (i.e., Figures 5-13 to 5-17) if snow has different elevational distributions from the past, then those differences are important for future DCR analysis and are an important caveat for this DCR.

Another point that does need to be raised pertains to the level of sea-level rise at which problematic conditions arise in CalSim3 and in general for the SWP and CVP. At a certain level of sea-level rise, saltwater intrusion becomes very problematic for water use planning, so this report should state that and use Kopp et al (2014) tables to estimate the likelihood that such sea-level rise will occur in the next 20 years, but note that the DCR is not relevant for SWP and CVP system performance above that threshold.

Specific Comments

There are also a number of specific points that should be addressed. They are listed here:

<u>Page 2-1, lines 17-18:</u>
The point regarding uncertainty in climate change impacts needs to be made cautiously. It can easily be construed or misunderstood as a statement equivalent to the future being uncertain, which then easily can be conflated to justify arguments that future projections, analysis, and planning are futile exercises. To be clear, they are not. The hydroclimatic and hydrometeorological future of California

is indeed uncertain, but not completely so. Rather, some remarkably strong statements about the hydroclimatological and hydrometeorological future of California: first and foremost, the likelihood of stationarity in hydroclimate in California throughout the coming decades of the 21st Century relative to the historical record is extraordinarily low. Second, there is very little uncertainty in the sign of changes in temperature and snowpack (Siirila-Woodburn et al, 2021). There are other changes too for which little future uncertainty exists, but these clear cases are listed because there are clear connections to physical processes that are well-understood now and are well-understood to exert change on California's future hydroclimate (such as the impact of increased greenhouse gasses on temperature, the impact of rising temperature on snow, precipitation amount/phase, etc.). Because of these quite likely conditions in the coming decades, assessments of a likely range of climate change impacts are needed to ensure the highest likelihood of state water resource plans being able to meet as many of the diverse stakeholder needs as possible under different, plausible, realistic hydroclimate states. The language in the report should be modified to reflect that there are aspects of California's future hydroclimate that are uncertain, and those that are, in fact, quite clear, and the combination of those certain and uncertain changes in aspects of California's future hydroclimate form the basis for resilience and vulnerability assessments, such as the one being conducted in this report.

<u>Page 2-1, lines 18-19:</u>
There is much literature on the challenges of developing realistic climate change projections with a single climate model realization and the likelihood of risk underestimation therein (e.g., Murphy et al, 2004; Bengtsson et al, 2006; Easterling and Wehner, 2009; Pierce et al, 2009; Kay et al, 2015, Steinschneider et al, 2015; O'Reilly et al, 2021; Charn et al, 2022). The report should consider adding these citations.

<u>Page 2-1, lines 27-28:</u>
Citations are needed to support the statement that multiple future scenarios allow for more robust planning. The report should consider citations of Barsugli et al, 2013; O'Neill et al, 2016; Moss et al, 2017; Deser et al, 2020; and Feldman et al, 2021, which, among many others, support the use of multiple future scenarios and ensemble members.

<u>Page 2-1, line 27-28:</u>
Here, the point about natural variability as a complicating factor for delivery

capability projections needs to be made. However, the discussion of natural variability in California's hydroclimate needs to be made judiciously in this report. The following is suggested: “The Earth system exhibits many different modes of natural variability due to the oscillating nature of its atmospheric and oceanic circulation. These modes can create apparent warming and cooling periods, sometimes 1-2 decades in length. Unfortunately, natural variability creates an envelope of uncertainty for looking at long-term trends because it is difficult to assess where the Earth actually is in terms of the current cycles of natural variability (like on the ascending or descending part of a wave). Therefore, developing an unbiased range of hydroclimate projections requires sampling natural variability, which is akin to sampling the ascending and descending parts of a wave to be able to isolate, identify, and quantify long-term trends that are occurring independent of Earth system natural variability.”

Page 2-2, line 3:
Change "climate change impact analysis" to "climate change impacts analysis".

Page 4-1, after lines 7-13:
This review reiterates the point made above regarding the overemphasis of uncertainty, which should be coupled with statements regarding the very high likelihood of hydroclimate state transitions relative to the historical record (e.g., Siirila-Woodburn et al, 2021).

Page 5-2, Line 3:
Figure 5-2 is meant to be a schematic, but some of the graphics in it are confusing. What is shown in the “Operations and Planning Models” figures in the left and right columns? What is shown in the “Conditional System Performance Projections” figure? What is shown in the “System performance prediction”? All of these smaller figures within the schematic need to be clearly-labeled, not pixelated.

Page 5-5, Line 3:
Figure 5-2 is pixelated and should be displayed or recreated at higher resolution. Also, the meaning of the dots in the upper-left contour plot on this figure are unclear. Presumably, these dots are outputs from the weather generator, but the color of the dots is not made clear. The colorbar label is not legible. Finally, the meaning of the red dots on the upper-left contour plot is not clear.

Page 5-6, Line 36 to Page 5-7, Line 1:
The phrasing is awkward and potentially misleading. A suggested change is the following: “Between one and 10 initial condition variants were available for each of

the three different future scenarios (SSPs). Initial condition variants represent an approach to estimate the envelope of natural variability by running multiple instances of a single climate model simulation but with infinitesimally small changes in initial conditions. Over the time-scales of the initiation of the climate model simulations, these differences produce model simulations that have identical representations of Earth system processes but the different ensemble members sample different phases of Earth System natural variability as realized by the climate model."

<u>Page 5-7, Line 3:</u>
Please consider adding a URL at which LOCA2 data can be obtained, such as https://cirrus.ucsd.edu/~pierce/LOCA2/.

<u>Page 5-18, Line 4:</u>
The slopes and intercepts of the major and minor axes of the ellipse should be listed, and confidence intervals on the estimates of those slopes should be added.

<u>Page 5-20, Line 4:</u>
The sentence "In theory, the only differences between these variants should be initial conditions and underlying climate stochasticity" is misleading and incorrect. The only difference between initial condition variants is their initial conditions. They are produced from an identical code-base otherwise, and they are deterministic codes (so the same code run with the same initial conditions will produce the same results to within the computational precision of the machine on which the code is run, which is typically double-precision). Please consider changing this sentence to the following: "The only difference in these variants is the conditions with which they were initialized. Specifically, these variants were created by perturbing their initial conditions by a value just larger than the computing precision of the specific computing system(s) on which they were run (typically ~10^{-14}) and then the variants are spun-up for hundreds or even thousands of years with the boundary conditions that the Earth system experienced (Kay et al, 2015; Deser et al, 2020). The result of this process is that initial condition variants exhibit differences that sample the space of natural variability as simulated by the climate model. This natural variability in the Earth system creates many features in short-term time-series that can appear to be trends, since the different physical mechanisms that contribute to natural variability produce modes (akin to sinusoidal waves) and it is difficult to know the exact phasing of the Earth system in these modes (akin to not knowing where on the wave you are until after it has passed) at any given time. Ultimately,

natural variability presents an analysis challenge with which to contend, but it only masks long-term hydroclimate trends associated with the changing amount of energy available to the Earth system due to anthropogenic influences."

Page 5-20, Line 8 and Line 20:
Caution should be exercised in the use of the word "stochasticity." It has a specific denotative meaning in statistics and a range of connotative meanings outside of the field of statistics. Consider other phrasing related to the range of potential responses of the Earth system to identical forcings, and the wide amount of literature on Earth system climate feedback analysis that supports such phrasing.

Page 5-20, Lines 25:
According to Pierce et al (2023), LOCA2 uses ensemble bias-correction, which means that each of initial condition ensemble members from a given model will exhibit biases in their historical simulations relative to the observational record, but the ensemble of all initial condition ensemble members is not biased in their historical simulations relative to the observational record. This approach to bias-correction in LOCA2 produces a range of values and trends that this report has shown. The primary purpose of ensemble bias-correction is to produce realistic short-term (daily, multi-day) extremes in temperature and precipitation, since fixed bias-correction techniques produce unphysical (capped) extremes in temperature and precipitation and the capping of extremes is inconsistent with Earth system dynamics and thermodynamics. The ensemble bias-correction does capture the linear impacts in Earth system natural variability. The information on ensemble bias correction in LOCA2 should be included, perhaps in a less verbose way, in this report.

Page 5-21, Lines 17-21:
Caution must be exercised in attempting to isolate the signal of climate change from natural variability. Indeed, the decadal climate prediction literature is replete with an exploration of this very approach. There are fundamental limitations in the ability of data analysis to do this (e.g., Weatherhead et al, 1998). The text should cite the Weatherhead reference or similar references related to using models to isolate and detect climate change signals (e.g., Leroy et al, 2007; Feldman et al, 2013). These and many other references show that there are fundamental limitations to isolating climate change signals from natural variability due to data record length, errors in models, and uncertainty in the structure and phasing of Earth system natural variability.

Page 5-21, Line 4:
The essential take-away of Figure 5-6 is that uncertainty in precipitation estimates grows with projection time. This should be stated explicitly.

Page 5-22, Line 13:
The slopes and intercepts of the major and minor axes of both ellipses should be listed and the confidence intervals of the estimates of those slopes should be added. Statistically significant differences between these slopes would indicate that the variant averaging procedure is destroying important information about the likelihood of future covariate changes in temperature and precipitation in California.

Page 5-23, Line 8:
Describe the details of what is being plotted in the box-whisker diagrams in Figure 5-8. What do the whiskers represent? 5-95 percentile? What about the box? What about the line and the 'X'?

Page 5-23, Line 18:
The comparison of precipitation change from 2040-2050 against 1981-2010 is ill-posed given natural variability in the Earth system. Please revise this figure so the change calculation between the historical and future hydroclimates is made between the future period over a longer period of time, preferably 30+ years.

Page 5-25, Line 6:
It appears that the word "bivariant" should be replaced with "bivariate"

Page 5-25, Line 7:
A justification is needed for the random sample size of 10,000. Presumably, it is meant to be a very large number of samples, but no information was provided as to whether estimates derived from this sample are sensitive to the sample size. Ideally, the sample size is large enough so that estimates are insensitive to the sample size, but information is needed on why 10,000, and why not 1,000 or 100,000, for example.

Page 5-25, Line 13-14:
Change "the considered socioeconomic and radiative concentration pathways" to "the widely-utilized shared socioeconomic pathways that exhibit different radiative forcing values at the end of the 21st Century."

Page 5-29, Line 3:
Move the y-axis label "Change in Temperature" so it is to the left of the figure

panels. Also include historical average response metrics in the capture of each figure panel so that the reader can understand the importance of the changes in response metric values that are shown in each of the contour plots.

Page 5-33, Line 4:
What do the colors of each dot mean?

Page 5-36, Line 4:
The explanation of red and gray dots needs to also be included in the caption or a legend for Figure 5-16, and not just in the main text.

Page 5-38, Line 3:
The explanation of red and gray dots needs to also be included in the caption or a legend for Figure 5-17, and not just in the main text.

Page 5-41, Line 23:
Reconsider the use of the word "daily weather trace" in favor of "realization of a set of simulated daily weather events". Otherwise, please define what is meant by "trace."

Page 5-42, Line 3:
The phrasing "Generally, a scaling rate of 7 percent per °C is expected" is problematic. The average precipitation definitely does ***not*** scale following the 7% per °C of the Clausius-Clapeyron relationship. Global and even regional precipitation has been found by numerous authors in observations, models and theory to scale at ~2% per °C (e.g., Boer, 1993; Betts, 1998; Trenberth, 1998; Allen and Ingram, 2002; Held and Soden, 2006; Boos, 2012; Li et al, 2013). It is true that extreme precipitation can exhibit Clausius-Clapeyron scaling (Allan et al, 2010) or scaling even above the Clausius-Clapeyron thermodynamic relationship (Lenderink et al, 2017; Martinkova and Kysely, 2020). However, if the stress test described here is to be more realistic, it should be performed with a 2% per °C scaling, unless single, or multi-day extreme precipitation represents the greatest source of stress to the system, in which case the bounding of 0%, 7%, and 14% is acceptable.

Page 5-42, Line 34:
As before, please consider a different terminology instead of "trace," or define "trace" formally.

Page 5-50, Line 15:
Please indicate the sea-level rise value above which there would be medium, large, and very large changes to operations.

<u>Page 6-2, Line 4:</u>

Please provide peer-reviewed citations on VIC's relative insensitivity to wind-speed.

Charge Question #1

Is the procedure developed by DWR appropriately documented? Is there anything missing from the documentation?

Response to Charge Question #1

In general, the procedure developed by DWR is appropriately and thoroughly documented, and much of the justification for all of the parts of the procedure has been provided. There are a few minor points that are missing, however, and those should be addressed in a revision to this report. They are listed below:

1. A description of the changes in LOCA2 downscaling relative to LOCA v1 is needed. In particular, there are changes that are described in Pierce et al (2023) and need to be included after the paragraph that ends on Page 5-6, Line 31. The changes to LOCA2 that need to be highlighted in this report include (1) that LOCA2 provided downscaling of CMIP6 models instead of CMIP5 models, (2) that the LOCA2 effort produced downscaling of many initial-condition ensemble members per model per emissions scenario rather than a single ensemble member per model as was done with CMIP5, (3) that the use of ensembles of downscaled solutions instead of a single solution has already been shown to be important for realistic sampling of regional and local hydroclimate change factors across the CMIP6 ensemble (Longmate et al, 2023), (3) that bias correction techniques in LOCA2 ensure that the ensemble of LOCA2 results for a given model and emissions scenario is unbiased in its historical simulations while retaining the ability to capture natural variability, and (4) that LOCA2 ensured that temperature and precipitation extremes that are under- or unconstrained by the historical record are unbounded and represented by a Generalized Extreme Value distribution with parameters estimated at each grid box from historical observations.
2. The climate uncertainty sampling described in Section 5.2.5 focuses on “levels of concern” with a reference to Francois et al, (in prep). So long as this reference is unavailable to the reviewers and the larger scientific community, the methods described in Francois need to be included in appendix material for this report.

3. A discussion of the historical performance of CalSim3 in terms of risk assessment for the SWP and CVP is needed. Has that model always produced actionable information for end-users? Has it had to be corrected recently in order to account for providing information that isn't always actionable for end-users? Does CalSim3 require structural modification to be implemented with climate-change-scaled temperature and precipitation inputs?
4. A more detailed discussion is needed of the sea-level rise at which the findings of this report would need to be revisited due to major issues arising from saltwater intrusion.
5. There are a number of minor clarifications that were noted in the specific comments.

Charge Question #2

Does the procedure apply rational and defensible evidence for the steps taken and techniques used to capture the probability of projected changes related to climate and sea level rise? Why or why not?

Response to Charge Question #2

In general, the procedures do apply rational and defensible evidence to support the steps taken and techniques used to capture the probability of projected hydroclimate and sea-level rise changes in the next twenty years.

There are many details to the procedure and most of them are rational and defensible, but there are a few details which do need to be addressed.

First, the efforts to use a broad range of LOCA2 data are commendable, but the report needs to recognize that the use of LOCA2 here represents an initial assessment of the importance of including Earth system natural variability in the course of establishing the vulnerability of the SWP and CVP systems to changing hydroclimate and sea-level rise. The actual Earth system natural variability is much higher than what was presented for two reasons. First, the initial condition variants produce lower assessments of natural variability than an ensemble that also includes perturbed physics ensembles (Palmer, 2000).

Second, statistical downscaling produces generally stationary downscaling solutions which underestimate the range of changes that may occur at the local level for a given change in large-scale circulation because they have inherent assumptions of stationarity (even if newer statistical downscaling approaches attempt to address where their solution space is grossly insufficient and unphysical at very high quantiles by using GEV functional forms to produce unbounded projections of extrema). The report needs to address the relative importance of downscaling method accuracy against developing a realistic range of climate model projections.

Third, more justification is needed for rejecting the solutions from those CMIP6 models that exhibit high climate sensitivity (the "hot models" as discussed in Hausfather et al, (2022)). The report needs to openly discuss when it is justified to discard climate model projections due to their poor historical or future performance. There are some clear examples where gross structural deficiencies in a model are problematic, but likely many others where caution must be exercised so as to avoid underestimating actual risks in an unknown future. The gross structural deficiencies could include the non-conservation of energy on a global scale, the lack of Clausius-Clapeyron water vapor scaling with temperature in the atmospheric boundary layer, unrealistic changes in ocean dynamics, and other clear violations of atmospheric physical processes that are known to be invariant, regardless of future climate change.

Fourth, more information is needed on the "Levels of Concern" to understand what quasi-probabilities mean relative to true probabilities. This information is being written up in Francois et al, but it needs to be provided in the review and also needs to be vetted by peer review.

Finally, the report authors should be commended for their efforts to validate their procedures against independent analyses to assess the reasonableness of this report's procedures. The comparisons of trends in Figures 5-9 and 5-10 provided important validation by indicating that the lengthy chain of analysis performed on LOCA2 data did not inadvertently introduce artifacts in trend analysis.

Charge Question #3

Do the new scenarios provide enhanced information for water users about *potential future conditions and system reliability risks? If not, why?*

Response to Charge Question #3

In general, the analysis presented in this report and the different "Levels of Concern" do provide enhanced information for water users about potential future conditions and system reliability risks as an initial part of the discussion with water users.

Figure 5-15, in particular, succinctly summarizes the findings of the likelihood and associated magnitude of reduced water resources for the SWP and CVP. It indicates that there is an extremely low likelihood that, in 20 years, there will be approximately the same level of water resources as there are now. At a high-level, this figure should be considered for a summary slide/document/handout (though the labels need to be simplified for such presentations).

The report also provides an assessment of more stressing sets of conditions for which water-users can plan, and walks those users through those, although the water-users may legitimately wonder what "levels of concern" and associated "quasi-probabilities" mean.

All this being said, the report does need to have an expanded set of statements in Chapter 6 on the limitations of what was done.

The sole use of statistical downscaling solutions and the sole use of initial condition variants (instead of both initial condition and perturbed physics variants) mean that Figure 5-5 through 5-17 are subject to revision, at least in that the range of results will grow. All of the ellipses will grow in size when multiple, plausible downscaling solutions are used along with perturbed physics variants, as will the size of the scatter plots, distances between 50th and 95th percentiles, and the height of the cumulative distribution function. It is very important to include these limitations, because a revised analysis for a future DCR based on CMIP7 downscaled solutions would very likely have larger ellipses, larger sizes of scatterplots, larger distances between the 50th and 95th percentiles, and larger heights of cumulative distribution functions than this report, and it will be important to highlight that this current report was aware of, and sought to mitigate as much as possible, limitations to this DCR's analysis.

Another limitation which needs to be mentioned pertains to sea-level rise. Chapter 6 needs to state that when sea-level rise increases above a certain threshold (DWR likely can provide an estimate of that threshold), saltwater intrusion into the California Delta would be significant enough to require a wholesale revision to this DCR. Again, this will be important in case there are major changes to sea-level rise which are currently considered low probability but may be nonetheless realized.

Charge Question #4:

Is this procedure an improvement over other previously used approaches to climate scenario selection/development? Why or why not?

Response to Charge Question #4:

The procedures described in this report are a welcome improvement to previous efforts that used limited climate model selections and scenarios to assess the risk to the SWP and CVP. As mentioned in the General and Specific Comments in this Review Letter, SWP and CVP system risk assessment due to nonstationary hydroclimate and sea-level rise requires a probabilistic approach which must consider Earth system natural variability, secular changes in temperature and precipitation, and rising sea-levels. The analysis presented throughout the report took that probabilistic approach and sought to estimate levels of risk and their likelihood, which should support a straightforward set of information for water-users to consider and around which to plan.

Charge Question #5:

Are there specific investigations or improvements that should be undertaken in future updates of this approach or use of this procedure to develop additional scenarios at time periods further into the future?

Response to Charge Question #5:

There are a number of specific investigations that are recommended for future updates to these approaches, including in preparation for an ever-more-ambitious, voluminous and likely non-convergent set of hydroclimate simulations that will be produced for California as part of future climate assessments, including for CMIP7.

The fundamental question that should motivate future specific investigations is this: is a probabilistic assessment of SWP and CVP in the next 20 years strongly dependent on choices of model democracy vs meritocracy, on the choice(s) of downscaling technique(s), and the use of initial condition and/or perturbed physics variants?

There are really three parts to that question which can be answered systematically. For the first part of the question, the importance of model democracy vs meritocracy to the probabilistic assessment shown here, especially with respect to disregarding or regarding high climate-sensitivity models, can likely be addressed with existing LOCA2 data.

The choice of downscaling technique is slightly more involved, but as Walton et al (2020) showed, there are potentially very different precipitation patterns predicted at the local level in California depending on which downscaling technique is used. While ensembles of different dynamical downscaling solutions do not currently exist, there are some very large ensembles of dynamically downscaled solutions (Rahimi et al, 2022; Rahimi et al, 2023) and hybrid downscaling approaches (Gutmann et al, 2016), which can explore, at least qualitatively, when and where Earth system variability and trends are being under-sampled. The approaches taken in this report do assume that the range of natural variability is adequately sampled by the analysis presented in this report. However, the range of natural variability is almost certainly still under-sampled. This is because there is a lot more diversity in climate model simulations than is contained in the LOCA2-CMIP6 dataset. Ultimately, analyses are needed that explore if CalSim3 is sensitive to the differences between statistical and dynamical solutions for California. Sensitivity analyses produced by simply scaling the LOCA2 output by precipitation and temperature values in those areas where LOCA2 and WRF disagree would be sufficient for gauging this risk.

Finally, the importance, or lack thereof, of including both initial condition and perturbed physics variants as part of the analyses in Delivery Capability Reports needs to be established. Indeed, California's hydroclimate appears to be very sensitive to the state of the ocean and atmosphere (e.g., in water year 2023, California kept on getting storms even though you would not necessarily expect it given the ENSO state). This is where the perturbed physics variants will be important to ensure that the range of the details of the ocean and atmosphere that matter for California are included in an assessment of natural variability, and that single sets of model realizations do not lead to an underestimation of risk.

References

Allan, R.P., et al. (2010) Current changes in tropical precipitation, *Environmental Research Letters*, 5, 025205, doi: 10.1088/1748-9326/5/2/025205.

Allen, M. R., and W. J. Ingram, (2002) Constraints on future changes in the hydrological cycle. *Nature*, 419 , 224–228, https://doi.org/10.1038/nature01092.

Barsugli, J. J., et al. (2013), The Practitioner's Dilemma: How to Assess the Credibility of Downscaled Climate Projections, *Eos Trans. AGU*, 94(46), 424, doi: https://doi.org/10.1002/2013EO460005.

Bengtsson, L., Hodges, K.I., Roeckner, E. et al. (2006), On the natural variability of the pre-industrial European climate. *Clim. Dyn.*, 27, 743–760, doi: https://doi.org/10.1007/s00382-006-0168-y.

Betts, A. K., (1998) Climate–convection feedbacks: Some further issues. *Climatic Change*, 39 , 35–38, doi: 10.1023/A:1005323805826.

Boer, G. J., (1993) Climate change and the regulation of the surface moisture and energy budgets. *Climate Dyn.*, 8 , 225–239, doi: https://doi.org/10.1007/BF00198617.

Boos, W. R., (2012) Thermodynamic Scaling of the Hydrological Cycle of the Last Glacial Maximum. *J. Climate*, 25, 992–1006, doi: https://doi.org/10.1175/JCLI-D-11-00010.1.

Charn, A. B., O'Brien, T. A., Risser, M. D., Longmate, J. M., & Feldman, D. R. (2022). Sign of observed California temperature trends depends on data set homogenization: Implications for weighting and downscaling. *Geophysical Research Letters*, 49, e2022GL099186. doi: https://doi.org/10.1029/2022GL099186.

Deser, C., Lehner, F., Rodgers, K.B. et al. (2020), Insights from Earth system model initial-condition large ensembles and future prospects. *Nat. Clim. Chang.* 10, 277–286, doi: https://doi.org/10.1038/s41558-020-0731-2.

Easterling, D. R., and Wehner, M. F. (2009), Is the climate warming or cooling? *Geophys. Res. Lett.*, 36, L08706, doi: 10.1029/2009GL037810.

Eyring, V., et al. (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, 9(5), 1937-1958, doi: https://doi.org/10.5194/gmd-9-1937-2016.

Feldman, D. R., D. M. Coleman, and W. D. Collins, (2013) On the Usage of Spectral and Broadband Satellite Instrument Measurements to Differentiate Climate Models with Different Cloud Feedback Strengths. *J. Climate*, 26, 6561–6574, doi: https://doi.org/10.1175/JCLI-D-12-00378.1.

Feldman, D.R., Tadic, J., Arnold, W., Schwarz, A., (2021) Establishing a Range of Extreme Precipitation Estimates in California for Planning in the Face of Climate Change, *Journal of Water Research Planning and Management*, 147(9), doi: https://doi.org/10.1061/(ASCE)WR.1943-5452.0001410.

Giorgi, F., C. Jones, and G.R. Asrar, (2009) Addressing climate information needs at the regional level: the CORDEX framework. WMO Bull., 58, 175-183.

Gutmann, E., I. Barstad, M. Clark, J. Arnold, and R. Rasmussen, (2016) The Intermediate Complexity Atmospheric Research Model (ICAR). *J. Hydrometeor.*, 17, 957–973, https://doi.org/10.1175/JHM-D-15-0155.1.

Hausfather, Zeke; Marvel, Kate; Schmidt, Gavin A; Nielsen-Gammon, John W; Zelinka, Mark. (2022) Climate simulations: recognize the 'hot model' problem, *Nature*, 605, 26-29, doi: https://doi.org/10.1038/d41586-022-01192-2.

Held, I. M., and B. J. Soden, (2006) Robust Responses of the Hydrological Cycle to Global Warming. *J. Climate*, 19, 5686–5699, doi: https://doi.org/10.1175/JCLI3990.1.

Kay, J. E., et al, (2015), The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. *Bull. Amer. Meteor. Soc.*, 96, 1333–1349, doi: https://doi.org/10.1175/BAMS-D-13-00255.1.

Kopp, R.E., Horton, R.M., Little, C.M., Mitrovica, J.X., Oppenheimer, M., Rasmussen, D.J., Strauss, B.H. and Tebaldi, C. (2014), Probabilistic 21st and 22nd century sea-level projections at a global network of tide-gauge sites. *Earth's Future*, 2: 383-406. doi: https://doi.org/10.1002/2014EF000239.

Lenderink, G., R. Barbero, J. M. Loriaux, and H. J. Fowler, (2017) Super-Clausius–Clapeyron Scaling of Extreme Hourly Convective Precipitation and Its Relation to Large-Scale Atmospheric Conditions. *J. Climate*, 30, 6037–6052, doi: https://doi.org/10.1175/JCLI-D-16-0808.1.

Leroy, S. S., J. G. Anderson, and G. Ohring, (2008) Climate Signal Detection Times and Constraints on Climate Benchmark Accuracy Requirements. *J. Climate*, 21, 841–846, doi: https://doi.org/10.1175/2007JCLI1946.1.

Li, G., Harrison, S. P., Bartlein, P. J., Izumi, K., and Colin Prentice, I. (2013), Precipitation scaling with temperature in warm and cold climates: An analysis of CMIP5 simulations, *Geophys. Res. Lett.*, 40, 4018–4024, doi: 10.1002/grl.50730.

Longmate, J.M., Risser, M.D. & Feldman, D.R. (2023) Prioritizing the selection of CMIP6 model ensemble members for downscaling projections of CONUS temperature and precipitation. *Clim. Dyn.*, doi: https://doi.org/10.1007/s00382-023-06846-z.

Maraun, D., et al. (2010), Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, 48, RG3003, doi: 10.1029/2009RG000314.

Martinkova, M., and J. Kysely, (2020) Overview of Observed Clausius-Clapeyron Scaling of Extreme Precipitation in Midlatitudes, *Atmosphere*, 11(8), 786, doi: https://doi.org/10.3390/atmos11080786.

Moss, R. H. et al, (2017), Nonstationary Weather Patterns and Extreme Events Informing Design and Planning for Long-Lived Infrastructure, Tech. rep., Workshop Report on Nonstationarity, ESTCP Project RC-201591

Murphy, J., Sexton, D., Barnett, D. et al. (2004), Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, 430, 768–772. doi: https://doi.org/10.1038/nature02771.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M. (2016) The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, *Geosci. Model Dev.*, 9, 3461–3482, doi: https://doi.org/10.5194/gmd-9-3461-2016.

O'Reilly, C.H., Befort, D.J., Weisheimer, A. et al. (2021) Projections of northern hemisphere extratropical climate underestimate internal variability and associated uncertainty. *Commun Earth Environ*, 2, 194, doi: https://doi.org/10.1038/s43247-021-00268-7 .

Palmer, T.N. (2000) Predicting uncertainty in forecasts of weather and climate, *Reports on Progress in Physics*,63, 71, doi: 10.1088/0034-4885/63/2/201

Pierce, D.W., T. P. Barnette, B. D. Santer, P. J. Gleckler (2009) Selecting global climate models for regional climate change studies, *Proceedings of the National Academy of Sciences of the USA*, 106(21) 8441-8446, doi: https://doi.org/10.1073/pnas.0900094106

Pierce, D. W., D. R. Cayan, and B. L. Thrasher, (2014) Statistical Downscaling Using Localized Constructed Analogs (LOCA). *J. Hydrometeor.*, 15, 2558–2585, doi: https://doi.org/10.1175/JHM-D-14-0082.1.

Pierce, D. W., D. R. Cayan, D. R. Feldman, and M. D. Risser, (2023) Future Increases in North American Extreme Precipitation in CMIP6 Downscaled with LOCA. *J. Hydrometeor.*, 24, 951–975, doi: https://doi.org/10.1175/JHM-D-22-0194.1.

Rahimi, S., Krantz, W., Lin, Y.-H., Bass, B., Goldenson, N., Hall, A., et al. (2022). Evaluation of a reanalysis-driven configuration of WRF4 over the western United States from 1980 to 2020. *Journal of Geophysical Research: Atmospheres*, 127, e2021JD035699. doi: https://doi.org/10.1029/2021JD035699.

Rahimi, S., Huang, L., Goldenson, N., Risser, M., Feldman, D. R., Lebo, Z. J., Norris, J., Dennis, E., Thackeray, C., and Hall, A. (2023), Mean-state biases in CMIP6 GCMs are a predictor of precipitation biases in dynamical downscaling, *Geophysical Research Letters*, Submitted.

Rhoades, A.M., Hatchett, B.J., Risser, M.D. et al. (2022) Asymmetric emergence of low-to-no snow in the midlatitudes of the American Cordillera. *Nat. Clim. Chang.*, 12, 1151–1159. doi: https://doi.org/10.1038/s41558-022-01518-y.

Steinschneider, S., McCrary, R., Mearns, L. O., and Brown, C. (2015), The effects of climate model similarity on probabilistic climate projections and the implications for local, risk-based adaptation planning. *Geophys. Res. Lett.*, 42, 5014–5044. doi: https://doi.org/10.1002/2015GL064529.

Siirila-Woodburn, E.R., Rhoades, A.M., Hatchett, B.J. et al. A low-to-no snow future and its impacts on water resources in the western United States. *Nat Rev Earth Environ*, 2, 800–819 (2021). doi: https://doi.org/10.1038/s43017-021-00219-y.

Trenberth, K. E., (1998) Atmospheric moisture residence times and cycling: Implications for rainfall rates with climate change. *Climatic Change*, 39 , 667–694, doi: https://doi.org/10.1023/A:1005319109110.

Walton, D., Berg, N., Pierce, D., Maurer, E., Hall, A., Lin, Y.-H., et al. (2020). Understanding differences in California climate projections produced by dynamical and statistical downscaling. *Journal of Geophysical Research: Atmospheres*, 125, e2020JD032812. doi: https://doi.org/10.1029/2020JD032812.

Weatherhead, E. C., et al, (1998) Factors affecting the detection of trends: Statistical considerations and applications to environmental data. *J. Geophys. Res.*, 103 (D14), 17,149–17,161, doi: https://doi.org/10.1029/98JD00995.

Wilby, R.L. et al. (2004) Guidelines for use of climate scenarios developed from statistical downscaling methods. IPCC Doc., 27 pp. Available at www.ipcc-data.org/guidelines/dgm_no2_v1_09_2004.pdf

Wood, A.W., Leung, L.R., Sridhar, V. et al. (2004) Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs. *Climatic Change*, 62, 189–216, doi: https://doi.org/10.1023/B:CLIM.0000013685.99609.9e.