OCT 2023

Central Valley Project Water Temperature Model Platform Review

Final Report

Delta Science Program

DELTA STEWARDSHIP COUNCIL

Table of Contents

List of Abbreviations

Abbreviation	Meaning	Agency
CVP	Central Valley Project	USBR
DSC	Delta Stewardship Council	CA
Review Panel	WTMP Peer Review Panel	DSC
RISE	Reclamation Information Sharing Environment	USBR
TCD	Temperature Control Device	USBR
USACE	U.S. Army Corps of Engineers	USACE
Reclamation	Bureau of Reclamation	USBR
WTMP	Water Temperature Modeling Platform	USBR

Panel Charge

Reclamation requests that the Review Panel identify both:

- Modeling elements that are appropriately represented and consistent with the project objectives; as well as
- Critical input and associated direction and recommendations to improve the model development and application.

Specific questions were developed to guide the Review Panel for the Mid-Term and Final Reviews.

1. Does the modeling design (e.g., model selection, framework) include the necessary processes and resolution (spatial and temporal) to represent the short-term and long-term temperature dynamics expected in the reservoir and river environments throughout the CVP project area?
2. Are the models adequate for describing water temperature during extreme hydrologic/storage conditions (e.g., droughts/low storage)?
3. Are unique features (i.e., selective withdrawal devices, thermal curtains, and submerged structures) adequately represented?
4. Are available data sufficient for the development of the selected models and intended uses?
   a. Where data gaps have been identified, are the assumptions and methodologies used to address them suitable?
5. Are testing methods (calibration and validation) adequate to demonstrate confidence in model performance for the historic period?

6. Does the modeling documentation include adequate information, assumptions, and detail to allow for transparency and replication of model results?
7. For Clear Creek and the Trinity, American, and Stanislaus River systems: A summary of items 1-6, regarding model development, calibration/validation, documentation, etc.
8. Are the model framework linkages adequate between models?
9. Are the models, in forecast mode, adequate for intended real-time and seasonal planning purposes (i.e., forecast period ranges from 3- to 5-days to six months into the future), based on performance measures, uncertainty, and the fidelity with which the models represent driving processes?
10. Do the model projections adequately account for the range of expected variability (e.g., hydrology and meteorology) from climate projections?
11. Are the metrics and methodology for describing and incorporating uncertainty in input data adequate and is model uncertainty described and quantified appropriately?
12. Are the modeling processes and approaches associated with model application appropriately documented?
13. What should be included in the models in the future to improve their accuracy, resolution, or other features?

Key Findings

The Water Temperature Management Platform (WTMP) developed by the Project Team within the Central Valley Project (CVP) provides important and commendable features, including (a) transparency along with open software, data, and metadata (Reclamation Information Sharing Environment,[1] RISE); (b) stakeholder engagement through open science; and (c) dissemination of models and data to build community capabilities both in-house and within stakeholder communities.

WTMP uses a systems framework with data flow through the modeling elements, allowing the automation of essential services. The WTMP Project Team developed a vision for the framework that accommodates running the systems at different spatial-temporal scales and for different purposes. This provides the ability for the analysis of model behavior at both element and system scales, as well as over both short- and long-term applications.

The Review Panel notes that the WTMP Project Team followed best practices for designing the model framework, as well as using a data management system (RISE) that offers the ability to proctor and serve data for user needs. The WTMP Project Team provides detailed, high-quality documentation of model development and implementation processes, as well as their performance relative to observed conditions. The WTMP Project Team also provides cross-model comparisons using CE-QUAL-W2 and ResSim that corroborate system performance.

The Review Panel would like to highlight the strong, effective, and efficient WTMP Project Team consisting of Reclamation and consultants, which includes an experienced and capable domain professional (Randi Field) who understands -- and

---

[1] https://data.usbr.gov

has extensive knowledge of -- daily operations, strategic vision, and stakeholder engagement. The Review Panel also finds that WTMP Technical Reports were appropriately and carefully revised based on suggestions and comments provided during the mid-term review.

The Review Panel finds that modeling elements are appropriately represented and consistent with project objectives. The Review Panel supports the modeling framework used to understand and predict water temperatures within CVP River-Reservoir Systems (i.e., Clear Creek, and the Shasta, Trinity, American, Stanislaus River systems). The integrated river-reservoir model, CE-QUAL-W2, and reservoir model, ResSim, are physically based models that have been widely used for similar systems and have been appropriately calibrated and evaluated. Each reservoir and river element are discretized and modeled using established scientific and engineering methods to incorporate both physical features as well as dynamic meteorological and hydrologic inputs.

A state-of-the-art Data Management System provides key information for both input boundary conditions as well as internal observations for model comparisons. The database provides key information needed to evaluate both real-time operations as well as longer-term planning. A data custodian curates data integrity to ensure that missing and inappropriate readings are identified and corrected prior to model simulation runs.

The WTMP framework is consistent with other model-development efforts. Critical inputs and associated system components are included. While model uncertainties have been rigorously examined, the Review Panel provides some modest recommendations. The Review Panel believes that the WTMP Project Team will be implementing these, and additional, improvements as they gain experience with model implementation. The WTMP framework provides a solid foundation upon which further improvements in characterizing model uncertainties can be identified based on comparison of model predictions with observations.

The WTMP scope and technical development effort effectively targets overall product improvement and credibility. These features should assure the Modeling Technical Community and stakeholders of the model's suitability and functionality.

Broad Recommendations

The Review Panel has several recommendations that may improve the transparency of the technical reports.

Overview of WTMP Documents

The documentation for the WTMP is extensive and was burdensome to understand during the review process due to the need to address the complexity of the CVP system. Consider providing a “capstone”, or summary report, that summarizes and synthesizes information for readers to answer key questions on the modeling capabilities, quality, and application in a single, readable document. In other words, address the Panel Charges directly and explicitly in a way that is accessible and useful to both technical and more casual users or stakeholders. Such a document could extensively reference the existing documentation and provide a much-needed roadmap to understand the scope of the project and build confidence in its capabilities for readers who are not skilled modelers or steeped in the details of CVP operations.

If a standalone capstone summary report is beyond the scope, available time, or available funding for the project, the Review Panel recommends adding an Executive Summary and documentation roadmap to the beginning of project documents that can better synthesize findings. A short overview document (or webpage), of no more than a few pages, should provide the overall project background and motivation. The overview summarizes the documentation and assists readers on what information is contained in each document. Fact sheets, or other broadly accessible documents or webpages, could provide a summary of WTMP capabilities and applications tailored to stakeholder and broader communities.

Technical Recommendations

Many of the provided reports (e.g., the Data Development Report) are vague on the difference between required ResSim and CE-QUAL-W2 boundary conditions, leading to ambiguous or incorrect statements throughout. For example, CE-QUAL-W2 requires dew point temperature as an input (not wet bulb temperature, relative humidity, or atmospheric pressure). This vagueness (and other occasional terminology errors) may hinder the evaluation and replication of model results for the different models.

Reporting goodness-of-fit statistics for reservoir stage is not a cogent metric as this is controlled by the modeler via use of the distributed tributary function in CE-QUAL-W2. Instead, it is the calibration target when closing the water balance for each reservoir. Instead, it is more useful to show the size of the distributed tributary. The Review Panel suggests instead reporting and discussing the size of the distributed tributary values applied to the various models relative to the size of the total inflows and outflows. The Review Panel recognizes that the distributed tributary is large at times and difficult to explain but notes that this is common across CE-QUAL-W2 modeling efforts and does not diminish model quality. Such a discussion provides insight into the structure of the model domain and helps identify errors and discrepancies within data sources. It also supports transparency and aligns with the idea of continuous model evaluation.

It is unclear whether replication is possible operationally or in “Forecast Mode”; the Review Panel suggests greater consideration be given to metadata or automated model archiving of individual runs, which can be used for replicability and transparency. This may be part of future work as the forecasting capabilities are further developed and refined.

Readability and Documentation

The Review Panel recommends that additional efforts be made to reduce redundancy throughout and across all documents, as well as the use of additional tables and figures to concisely deliver information. For example, descriptions of general approaches are repeated across basins or even individual models in many instances. This is unnecessary and can be confusing to a reader. The Review Panel suggests providing an introductory section on general approaches, which can then be referenced through the remainder of the documents.

- Provide greater referencing to other sections of report (or other reports) and to figures throughout all documentation to help readers navigate the volume of material.
- Provide a clearly labeled map or diagram of the entire project domain that can be referenced in the introduction and throughout each document. Ensure that this map includes all flow routes and clearly depicts the route of trans-basin transfers and locations referenced in the documents (e.g., Carr Powerhouse, Goodwin Dam).
- Similarly, ensure that all inflows and outflows are labeled on the provided maps of many reservoirs. Inflow and outflow locations are not intuitively clear from the topography or natural hydrology due to the complex engineered within the CVP system.
- Incorporate diagrams of dam outlets and other relevant structures (curtains) for all dams in relevant documents. The data development report has very nice diagrams that are immensely helpful to understand the logic behind flow routing and selective withdrawal in the models. The Review Panel

suggests making much greater use of these across documents, most notably, in the Development, Calibration, Validation, and Sensitivity Analysis report.

- Ensure that all figures have complete legends (preferable) or complete explanations in the accompanying caption. Many figures have incomplete labeling and cryptic colors or symbols that hampers their use and readability.
- Ensure a consistent set of units for each metric throughout all documents. The Review Panel understands that a mixed system of U.S. customary and metric units is commonly used by Reclamation and other stakeholders, but a single unit should be applied consistently throughout the document for individual parameters (i.e., temperature should be reported in either degrees Fahrenheit, degrees Celsius, or both throughout all documents)
- Ensure that all acronyms and abbreviations are defined upon first usage and in the front matter of each document. Reduce use of acronyms, abbreviations, and jargon throughout all documents, where possible.
- It would be useful to provide a brief overview of the regulatory framework, management requirements for Reclamation, and location of temperature management points on the rivers (and thus the reasons for the spatial extent of the modeling domain as background). This could be very brief but would help readers contextualize the scope and scale of the project.

Question 1. Does the modeling design (e.g., model selection, framework) include the necessary processes and resolution (spatial and temporal) to represent the short-term and long-term temperature dynamics expected in the reservoir and river environments throughout the CVP project area?

The modeling design, encompassing the model selection, framework, data management, and quality control, are well crafted to represent both short- and long-term temperature dynamics expected in both reservoir and river environments throughout the CVP project area. A detailed vision is presented for a framework that accommodates running the systems at different spatial-temporal scales and for different purposes. The WTMP Project Team follows best practices to develop the models and model framework. As evidenced by the excellent performance of the models compared with the observed data, the models and framework are robust, integrating the data flow, modeling, and automation. Furthermore, the presented framework is versatile, designed to operate across diverse spatial-temporal scales and tailored for various objectives. Detailed analyses have been conducted on the model's behavior at the elemental scale, ensuring adherence to industry best practices during setup. All previous concerns identified during the Mid-Term Review have been addressed, such as the minimal influence of hyporheic flows on temperature profiles, particularly at the modeled scales. Previous questions have also been addressed. For example, hyporheic flows have been demonstrated to have a small effect on the temperature profiles, especially given the scales being modeled.

Additional comments are provided in the following sections.

Wind Sheltering Coefficients

Constant wind sheltering coefficients are used in the Sacramento, Trinity, and American River CE-QUAL-W2 models. Using constant coefficients offers simplicity and helps ensure consistency in the simulations, both of which are useful for comparing alternative scenarios. However, this approach may not capture temporal variability due to factors like vegetation growth or seasonal changes on the magnitude and direction of the wind field. While it streamlines the modeling process, there is a risk of potential inaccuracies if the constant coefficient does not represent typical conditions throughout the year. It is crucial to weigh the benefits of simplicity against the potential for missing important episodic or seasonal events that could affect the modeled system. Some discussion is needed regarding the selection of constant wind sheltering coefficient values for the Sacramento-Trinity and American River CE-QUAL-W2 models. The WTMP team performed an assessment of how sensitive these models are to this coefficient, as presented during the August 2023 meeting. Please include this analysis in the documentation.

Technical Memorandum: Water Temperature Modeling Platform: Model Framework Application and Design (DRAFT)

Table 4 states that a benefit of general command-line models is that they can be run in an automated mode on a server. CE-QUAL-W2 and HEC-5Q are also command line models. HEC-RAS simulations can also be run from the command line.

Technical Memorandum: Water Temperature Modeling Platform: Model Selection (DRAFT)

The CE-QUAL-W2 entry in Table 3-1, in the URL column, please add a link to the CE-QUAL-W2 fact sheet[2]. This has been updated for Version 4.5 and provides the US Army Corps of Engineers (USACE) Engineer Research and Development Center (ERDC) and Portland State University (PSU) points of contact.

Table 3-5 lists CE-QUAL-W2 as a discrete model rather than a system model. CE-QUAL-W2 is often treated as a discrete model, but it can simulate multiple water bodies and branches. However, unlike HEC-ResSim, the current version of CE-QUAL-W2 does not compute reservoir releases (i.e., using a set of rules) or system operation of two or more reservoirs. In this context, HEC-ResSim is clearly a system model, while CE-QUAL-W2 is targeted at a smaller subset of water bodies and is typically treated as a discrete model, as it is for the WTMP. The introduction provides a brief definition, but interconnecting river reaches seems to be the defining feature, putting CE-QUAL-W2 in the same category as HEC-5Q and HEC-ResSim. A section should be added outlining the criteria that determines whether a model is considered a discrete or system model for the purposes of the WTMP.

---

[2] https://www.erdc.usace.army.mil/Media/Fact-Sheets/Fact-Sheet-Article-View/Article/554171/ce-qual-w2/

Question 2. Are the models adequate for describing water temperature during extreme hydrologic/storage conditions (e.g., droughts/low storage), as defined from both the historical record and future climate change projections?

The models appear suitable for modeling historical conditions that encompass the range of historical extremes. The Review Panel appreciates that the WTMP Project Team addressed concerns from the mid-year review to highlight different hydrologic conditions by water year types, such as wet and dry (for some models, but not all; suggest expanding this to include all). Additional summary discussion of performance broken out by year type could help modelers and stakeholders understand under what conditions uncertainty in predictions increases (or, conversely, provide additional confidence in results under extreme conditions). For example, a comparison of the range of stages under which each model was calibrated juxtaposed with the range of projected stage conditions under climate change. Similarly, some discussion on the range of meteorological conditions under which the models were calibrated would be useful to build confidence in the calibration.

Parameterizing boundary conditions based on regression relationships is a reasonable, well-documented approach and appears to produce high-quality modeling results as indicated by the model performance statistics. The Review Panel does not recommend changes for the current modeling effort but suggests some future consideration into the limits of regression approaches under extreme conditions. Similarly, the robustness and stationarity of regression relationships may become issues in the future. Efforts to establish targeted new data sources or implement a periodic revisiting of boundary condition estimation methods may be valuable in the future as the WTMP is implemented as a living framework (see discussion under Question 13)).

Without a detailed idea of the range of future climate change projections under consideration, it is difficult for the Review Panel to determine if the models are adequate to simulate climate change scenarios. There does not appear to be any reason to expect that the models would not perform well given their excellent calibration (assuming the boundary condition estimation methods remain valid), but explicit discussion of potential climate change evaluation would be useful in the documentation.

Question 3. Are unique features (i.e., selective withdrawal devices, thermal curtains, and submerged structures) adequately represented?

The WTMP Project Team selected an efficient modeling framework to simulate water temperature within the three river systems: Trinity-Sacramento, American, and Stanislaus. The modeling framework allows flexibility to change modeling elements to address processes at different temporal and spatial scales. The WTMP Project Team has also identified those unique features that require special attention within each modeling element. These features are unique to each system because the governing equations, which model hydraulics and heat transport (water temperature), may not capture them due to limitations of data, of spatial resolution and/or of modeling element dimensionality (one dimensional, 1D, or 2D instead of 3D). These unique features are (a) the temperature control device (TCD) at Shasta Dam, (b) thermal curtains in in Lewiston Lake and Whiskeytown Reservoir, and (c) tunnels in the Trinity-Sacramento system, (d) the selective water withdrawal shutters at Folsom Dam in the American River System and (e) the submerged dam in New Melones Lake in the Stanislaus River System.

The effects of unique features on flow hydraulics and water temperature depend on the model modeling element dimensionality (e.g., 1D vs 2D) and their temporal and spatial resolutions. Reservoirs within the modeling framework are simulated with HEC ResSim as a 1D vertical model and CE-QUAL-W2 as the 2D transverse-averaged model. Consequently, the inclusion of unique features is treated differently in ResSim and CE-QUAL-W2. The team properly evaluated their effects in both ResSim and CE-QUAL-W2 modeling elements and provided adequate model performance for each modeling element at each system.

The TCD presents several challenges, which include (a) leakages at the panels, (b) large panels that withdraw water from broad vertical bands with potentially different water temperatures, and (c) lateral locations of the TCD relative to the center of the dam. These characteristics cause 3D flow behaviors that neither 1D nor 2D approaches can model. Consequently, the most effective approach is to parameterize their behavior on water temperature releases based on reservoir water stage and TCD operation. As demonstrated, this is an effective approach that provides good performance when comparing predicted and measured water temperatures both in calibration and evaluation ("validation") stages in both ResSim and CE-QUAL-W2 modeling elements.

Similarly, the effect of the temperature shutters at Folsom Dam on water temperature at the outlet is parameterized considering their geometry and their operations by accounting for shutters position and leakage due to different shutter operational configurations. Performance of both ResSim and CE-QUAL-W2 to simulate water temperature releases from Folsom indicates that temperature shutters are well implemented in both modeling elements.

Temperature curtains have been implemented as a physical structure, floating skimmer weir, in CE-QUAL-W2, but their effects have been lumped into the entrainment algorithm or calibration of the withdrawal envelope for ResSim. The latter was done because ResSim 1D dimensionality does not allow it to add the structure as is done within CE-QUAL-W2. The good performance of the modeling elements (CE-QUAL-W2 and ResSim) for both calibration and evaluation ("validation") stages suggests that the features are well captured in both Lewiston and Whiskeytown Reservoirs.

New Melones Reservoir has a unique feature, i.e., the original Melones Dam that is now submerged within the reservoir. This feature affects the flow hydraulics near the New Melones Dam outlet. The submerged dam is properly characterized as an internal topographical feature in CE-QUAL-W2 and by a restriction on the withdrawal envelope in ResSim. These are adequate representations of this feature that is properly represented within the dimensionality of the selected modeling elements.

Other potential unique features have been identified and properly represented within the selected modeling elements.

Question 4. Are available data sufficient for the development of the selected models and intended uses?

a. Where data gaps have been identified, are the assumptions and methodologies used to address them suitable?

The WTMP Project Team should be commended for the detailed summary of their approach, the systematic development of data systems to support management of cold-water resources, and the attention to model data quality and relevance.

Here, we note a few items that could assist the developers in documenting and managing data resources.

The Review Panel suggests that the WTMP Project Team investigate additional methods for gap-filling to manage “Invalid” data resulting from gaps, spikes, steps (offsets), and trends due to equipment and operator failures. Another Review Panel recommendation is to calculate evaluation metrics (RMSE etc.) and regression equations using logarithmic discharge data. The metrics should include standard error of estimate, t-stat, and significance of regression estimated coefficients.

Because there is not an instantaneous correlation between flows in nearby river system, the Review Panel encourages the WTMP Project Team to consider using lagged regression models for Sulanharas Creek and other ungaged tributaries, such as the convolution sum:

$$Y(t) = a(o)\,X(t) + a(1)\,X(t\pm1) + a(2)\,X(t\pm2) + \ldots + a(m)\,X(t\pm m) = \sum a(i)\,X(t\pm i)$$

where a is the response function between X (e.g., Q_{sac}) and Y (e.g., Q_{sul}). Note that lags could be negative in some systems ($i < 0$), especially when estimating smaller, ungaged tributary inflows using gaged flows in larger rivers.

To provide consistency between systems, the Review Panel suggests that the WTMP Project Team indicate the date for each bathymetric map, including whether post-dam construction sedimentation is considered, and the technology used to generate the data.

The Review Panel encourages the WTMP Project Team to provide a table (or plots) of reservoir pools (e.g., surcharge, flood control, full, conservation, drought, inactive) with respect to elevations and/or volumes. This would include providing maximum cold-water pool volumes for each water (or calendar) year, and the anticipated volume required to meet downstream cold-water needs.

The WTMP Project Team might consider augmenting weather station data at a point with downscaled meteorological inputs over the watershed using assimilated weather data (e.g., NASA's LDAS[3]). Yet, this approach may not provide the temporal resolution required by WTMP models and may introduce excessive modeling efforts with limited utility.

The Review Panel also suggests (a) plotting discharge (as well as reservoir storage volumes) on logarithmic axes, (b) providing crossplots of observed vs modeled data in addition to time-series plots, (c) replacing "Average Annual Flow (cfs/mi^2)" and "Monthly Timing Factor" with "Water Yield", and (d) summarizing sites where additional data collection (e.g., channel cross sections, meteorologic, river and reservoir water temperatures, Whiskeytown Dam upper and lower outlet schedules) might assist with model calibration, as well as sites where current data collection may not be useful and could be discontinued.

[3] https://ldas.gsfc.nasa.gov/

Question 5. Are testing methods (calibration and validation) adequate to demonstrate confidence in model performance for the historic period?

In aggregate, the Review Panel responds positively to Question 5 in the Review Panel's charge. Model results for the historic period are generally good across the entire period and for all projects and compare favorably with other large-scale modeling efforts for stream temperature with which the Review Panel is familiar.

The WTMP is a complex modeling application and care needs to be taken to document and identify the model versions, model configurations, parameter sets, and model input data that were used to produce a particular model simulation or set of model simulations. That is particularly true for model simulations that are disseminated or that are the basis for operational decisions. Ideally this process should be automated rather than left to each individual model operator. The Review Panel therefore strongly recommends that metadata be developed that document these model configurations.

Discussion of model calibration and validation is expanded in the Final Report following review comments by the Review Panel during the midterm review. At the time, the Review Panel suggested moving away from the term "Validation", because it implies that a model is either good (if "Valid") or bad (if it fails to pass some a priori established performance criteria). Instead, the Review Panel encouraged the WTMP Project Team in the midterm evaluation to think more broadly about model evaluation to learn about the strengths and weaknesses of the model approach.

In the final report, the WTMP Project Team provided additional discussion of model calibration and validation. The WTMP Project Team retains the term "validation" because it is used widely within the Bureau and in discussion with its partners and stakeholders and is consistent with California Water and Environment Modeling Forum[4]. At the same time, the WTMP Project Team tested the sensitivity of the model to many parameters and is also developing workflows that continually evaluates model performance (e.g., through annual hindcasts). Detailed discussions of calibration and sensitivity results are provided in the updated documentation. These practices are well-aligned with the midterm review recommendations from the Review Panel.

Similarly, as suggested, the WTMP Project Team specified different performance criteria for different systems, e.g., separate performance criteria for large, medium, and small reservoirs and for river segments. They also provide different performance thresholds and metrics for different variables, in particular stage, flow, and water temperature. These changes are viewed positively by the Review Panel.

Plots and graphics are generally clear and informative, although their sheer number makes it at times difficult to find the desired information. The Review Panel suggests revising bar plots that are based on limited data (e.g., Figures 4-135 to 4-138 and 4-140 to 4-143) which use a bar to represent validation results. In the figures, the bar only represents two values (n=2); it is better to show the individual values.

Although it was discussed during the in-person presentations in September 2023, the project documentation provides little information about the calibration approach in the project documentation (*Technical Memorandum: Model Development, Calibration, Validation, and Sensitivity Analysis (DRAFT; August 2023)*. The Review Panel does not object to the use of a manual calibration approach, but

[4] https://cwemf.org/

at the very least this should be explained in the documentation. The term "Manual Calibration" is not mentioned anywhere in the relevant technical memorandum.

A manual calibration approach has disadvantages. Recalibrating the model, e.g., when model upgrades are made or when new data becomes available, can be labor-intensive and therefore expensive. At the same time, modelers gain important experience regarding model behavior and model performance when models are calibrated manually. Automated calibration approaches can be difficult to implement and at times have difficulty converging. They may also result in model parameter combinations that are difficult to interpret from a physical perspective. Here, we simply ask that the calibration approach be clearly documented, which can be brief. Because the model in its current form performs well, it is not clear that automated calibration procedures are a short-term priority.

Some information is already in the "*Model Framework Selection and Design*" document, which refers to itself as the *first document*, but that is not clear until the user starts reading this document. That document also refers to "*Technical Memorandum 6: Water Temperature Modeling Platform: Model Selection*", but that numbering is no longer present in the updated form of the document.

Question 6. Does the modeling documentation include adequate information, assumptions, and detail to allow for transparency and replication of model results?

The presentations by - and discussion with - the WTMP Project Team at the final model review symposium provided a thorough and convincing case for the quality and breadth of the modeling development, calibration, framework, and forecasting preparation. The organization of the presentations and following discussions was clear and focused and provided the Review Panel with the information needed to address the charge questions.

Similarly, the modeling documentation is highly detailed and thorough, but does not benefit from the same degree of focus and organization that the presentations did. In fact, the thoroughness is such that it becomes overwhelming for a reader to (1) find specific information related to the model assumptions, calibration, framework, forecasting methods, or (2) develop an overall understanding of the strengths, limitations, and overall capabilities of the project. This may hinder the ability of both the technical and stakeholder communities to appreciate the work done. It is an ironic limit on the transparency of the documentation. The project is impressively executed and has developed cutting-edge modeling capabilities – the documentation should reflect these achievements.

Comments specific to Technical Memorandum: Model Development, Calibration, Validation, and Sensitivity Analysis (DRAFT; August 2023):

The first few pages (4-1 to 4-8) of *Chapter 4: Model Calibration, Validation, and Sensitivity: Sacramento – Trinity Rivers System* contain information about the calibration, validation, and sensitivity approach that is more general than just the Sacramento – Trinity Rivers (e.g., description of metrics, thresholds, and overall approach). Similarly, general information about the sensitivity analysis is presented in the pages immediately following Figure 4-67.

This information should be moved to earlier in the document because it applies to the other sub-models and projects as well. It may be useful to collect all information about the calibration, sensitivity, and sensitivity into a single chapter immediately following Chapter 2, which can then be referenced by all the chapters describing the calibration, validation, and sensitivity experiments for the individual projects.

Question 7. For the Trinity River, Clear Creek, American River and Stanislaus River systems: a summary of Questions 1-6, above, regarding model development, calibration/validation, documentation, etc.

See responses for Questions 1 through 6, where the Review Panel has incorporated its responses for Clear Creek and the Trinity, American, and Stanislaus Rivers into its responses.

Question 8. Are the model framework linkages adequate between models?

The Review Panel finds that the modeling framework is suitable for separating the system into its key elements (e.g., rivers, tunnels, reservoirs), where each element is then simulated with a model informed by external forcing and/or outputs from the upstream modeling elements. This results in a cascading of modeling elements with information flowing from the upstream to the downstream elements. Consequently, linkages between models are used to ensure downstream models are properly informed.

The modular approach has the advantage of using models of different spatial and temporal resolutions and scales for selected elements, e.g., rivers, reservoirs. The CVP-WTMP has several elements that include reservoirs, rivers, lakes, tunnels, and canals. Rivers have been modeled using a one-dimensional (1D) longitudinal model, whereas reservoirs are modeled with a 1D vertical model or two-dimensional (2D) transversally averaged model depending on the simulation objectives.

This approach provides the flexibility required to address objectives with different temporal scales. For instance, ResSim is much faster than CE-QUAL-W2 to run and converge and would be a tool for long-term or for multi-scenarios analyses. However, the modular approach also presents challenges, because 1D and 2D models may require different input conditions. This is especially important when selecting different modeling elements, e.g., CE-QUAL-W2 and ResSim, within the same modeling framework. CVP-WTMP has developed a modeling framework that can accommodate swapping between modeling elements and ensure adequate information flow. This also ensures that the outputs from the upstream model have the required temperature resolution needed for the next element.

These linkages have been designed to automatically transfer inputs from one modeling element to the next.

Question 9: Are the models, in forecast mode, adequate for the intended real-time and seasonal planning purposes (i.e., forecast period ranges from 3- to 5-days to six months into the future), based on performance measures, uncertainty, and the fidelity with which the models represent physical processes?

In general terms, the Review Panel responds positively to Question 9 in its charge. It is more difficult for the Review Panel to comment on "*adequate for intended real-time and planning purposes*", because these purposes are external to the model technical evaluation. As presented and calibrated, the models should be adequate for the intended real-time and seasonal planning purposes (see responses to Questions 2, 3, and 5). In real-time and seasonal planning applications, the parameterizations and model parameters used to represent unique futures and the regression relationships used to parameterize boundary conditions should continue to capture the behavior of the system. The WTMP Project Team should be careful when interpreting results based on conditions that are sufficiently dissimilar from historic conditions that calibrations and parameterization may no longer hold (see response to Question 10). This challenge is not specific to the WTMP modeling effort but is a general challenge in the use of historically calibrated numerical models for climate change impact analysis.

Much of the actual forecast methodology and workflow remains to be developed and is separate from the development, implementation, and calibration of the WTMP. Consequently, the Review Panel limits its evaluation to the material made available as part of the final review. Documentation is provided for a suggested workflow for seasonal water temperature forecast simulations that relies on forecasted inflows and releases and meteorologic conditions from historic data or forecasts. An ensemble forecast can be developed by combining multiple initial conditions, inflow and release forecasts, and meteorological conditions. The monthly information that is used to create seasonal forecasts must be disaggregated into shorter timesteps (daily or shorter) to be used as input to the WTMP.

No information is provided for quantifying 3- to 5-day forecast uncertainties. While these errors are likely to be small (perhaps equal to model calibration errors), forecast uncertainty increases with lead time, with short-term accuracy having higher confidence than long-term forecasts. Mechanistic models are commonly superior to statistical time-series models (e.g., auto-regressive, moving average, neural network, extended Kalman filtering) for long-term forecasts and developing release strategies, while the autocorrelation structure between inputs and outputs may provide superior short-term forecasts.

A discussion of error sources that contribute to forecast uncertainty is provided in the *Technical Memorandum: Water Temperature Modeling Platform: Estimation of Uncertainty – Protocols (DRAFT)*. The content of this memorandum is less general than its title suggests, and it is not clear whether this information should stand as a separate memorandum. Chapter 2 provides general information about sources of uncertainty but is rather generic and could be part of the model development document. Chapter 3 discusses some of the uncertainty associated with forecasting using the WTMP but is limited in its discussion and its terminology is confusing. Most of the "*forecast process uncertainty*" in Chapter 3 does not directly address the uncertainty in the forecasted input and boundary conditions, but instead focuses on the uncertainty that originates from calibrated model error, selective withdrawal logic uncertainty, and disaggregation of coarse temporal data (uncertainty Types A-C). Even Type D uncertainty, which accounts for meteorology forecast uncertainty, is more strongly related to disaggregation of the monthly forecasts into daily values than to how well the monthly forecasts represent future months.

That is, the "*forecasting process uncertainty*" does address uncertainties associated with the process used to make forecasts but does not address the (likely) larger source of uncertainty associated with monthly inputs and boundary conditions, at least not as tested here. The WTMP Project Team is aware of the restrictive meaning of the term "*forecasting process uncertainty*", but it is likely to be confusing to any outside group (like the Review Panel or stakeholders). The true uncertainty is much larger than what is shown in the Technical Memorandum. Consider removing the term "forecast" from this uncertainty discussion.

Question 10. Is the proposed plan to manage the range of expected variability (e.g., hydrology, meteorology) from future climate projections adequate?

The Review Panel supports the proposed methodology for managing the uncertainty from future climate projections. The proposed model workflow provides the ability to account for expected (or known) climate variability by using external inputs that account for hydrologic and meteorologic variability. As noted in the presentation by the WTMP Project Team, this known uncertainty is based on (1) Emission Scenarios that (2) drive climate simulations, which are (3) spatially downscaled for (4) use by hydrologic models that (5) inform Operations Models that also consider (6) Bay-Delta Models.

In a larger sense, the possible universe of climate-change outcomes is unknown, so managing the range of expected variability resulting from future climate projections is a complex and multifaceted challenge. Climate change brings about a wide range of impacts, and managing variability involves a combination of mitigation and adaptation strategies. Managing the range of expected variability resulting from future climate projections requires a comprehensive and integrated approach. Collaboration, research, and adaptive management are key components of successful climate variability management.

The current model should be able to provide reasonable simulation if future climate forcing is within those historically included during model calibration. While predictions for conditions outside of the antecedent data envelope are possible, their accuracy is uncertain. Yet, predictions using physically based models are inherently more accurate than statistical forecast tools.

The WTMP Project Team has taken important and commendable steps to position the modeling framework along an adequate path. These steps include: (1) engagement with scientific community working on climate forecasts and downscaling, (2) engagement with stakeholders, (3) modeling and data sharing, (4) sharing documentation on modeling approach and performance, (5) system monitoring, (6) hinder approach for continuing model performance evaluation, and (7) building the system framework to operate within a Monte Carlo framework.

Question 11. Are the metrics and methodology for describing and incorporating uncertainty in input data adequate and is model uncertainty described and quantified appropriately?

The Review Panel agrees with the metrics and methodologies for describing and incorporating uncertainties. The WTMP Project Team used methods that are routinely used within the modeling profession. While model uncertainty is described in many parts of the documentation, the Review Panel believes that a more structured discussion that focuses on the multiple sources of forecast uncertainties be addressed. Models require two general types of information as input: physical features (e.g., reservoir bathymetric, stream morphology, dam dimensions) as well as internal and boundary conditions (e.g., tributary inflows, reservoir temperature profiles). Physical features are usually constant over time, while internal and boundary conditions are dynamic variables that are input as time series.

There are processes, however, that can change physical features; such a flood that scours a river channel or sedimentation that alters reservoir bathymetry. The Review Panel recommends that the WTMP Project Team monitor and/or develop plans to resurvey physical features at regular intervals (e.g., every decade) to update model features. New technology is likely to reduce the effort to update these parameters. Also, “as built” documentation that describes the final dimensions of structures should be used instead of the original design documents, in that modifications may have been made that are not reflected in the original design documents. A risk-informed analysis could be used to identify those features that most influence compliance with model performance objectives.

Based on information provided, there is little discussion of the uncertainty associated with physical features. A quality-assurance methodology that uses repetition (multiple measurements using a single technology), duplication (multiple measurements using different technologies), and redundancy (independent estimates using alternative approaches, such as tracers, mass balance calculations). Careful consideration of discrepancies between estimated features can be used to identify model uncertainties associated with physical features.

Uncertainties associated with time-series data could be better identified. A common problem associated with field data collection is specification of the timestamp. Switching between Standard and Daylight-Saving Time often causes uncertainty if it is not documented. Also, using UTC or UNIX time would minimize time uncertainties associated with data sensor calibration and reporting. Locations, elevations, and stage measurements are also problematic, in that multiple horizontal and vertical mapping schemes are in use. Adding metadata to each process could help to reduce these uncertainties. The WTMP Project Team could refer to the NSF-funded Water Data Markup Language (WaterML) developed by the Consortium of Universities for the Advancement of Hydrologic Sciences, Inc. (CUAHSI[5]).

Data integrity can be evaluated using observed correlations between variables, in that highly correlated variables should respond similarly. Outliers and erroneous data can be identified whenever the established relationship fails. These instances should be examined to evaluate the likely source of data inconsistency. Data gaps (i.e., missing data), steps, spikes, and trends can be examined using these tools.

---

[5] https://www.cuahsi.org

Appropriate interpolation strategies can take advantage of data correlations to better adjust these data. Also, routine equipment calibration should be used to estimate data measurement uncertainties.

Forecast uncertainties result from uncertainties in characterizing model features and boundary conditions. Ensemble averages can be used to evaluate the range of possible future responses. Yet a risk-informed method would focus on building forecasts for specific scenarios where the system could fail to meet regulatory objectives, such as extreme dry, warm conditions. Prioritizing data collection that minimizes uncertainties during these events would be more efficient than reducing uncertainties when risks are low. That is, performance-confirmation monitoring could be used to characterize uncertainties during critical periods.

Question 12. Are the modeling processes and approaches associated with model application appropriately documented?

The Review Panel commends the WTMP Project Team on the quality, comprehensiveness, and detail of the documentation. Given the extensive nature of the material, navigating specific details can prove challenging. The Review Panel has identified several areas that can be improved to ensure that technical and non-technical stakeholders can grasp, navigate, and apply the knowledge shared with greater efficacy, amplifying the utility of the WTMP.

Specific recommendations include:

Consolidating Information

Each document has evidently been crafted to stand on its own as much as possible. This naturally leads to considerable redundancies across documents. Instead, a dedicated section outlining general methodologies could be developed and then referenced in subsequent sections, making the documentation more streamlined. For each document, it would be also valuable to have an executive summary that provides a concise overview of the content. This helps users quickly understand the core objectives, methods, and findings without delving into intricate details. A fact sheet (1-3 pages) accompanying each document could offer stakeholders a snapshot of the model's critical features, applications, and benefits. This enhances transparency and provides an easy reference point. Finally, a separate document that provides a high-level overview of all the models, highlighting their broad similarities and differences would help identify relevant information to address stakeholder concerns, future modeling issues, and extension of the models and platform.

Improved Navigation

To aid readers, internal references to different report sections and visual aids should be increased. A clearly labeled project domain map, complete with flow routes and key reference points, can be indispensable.

Visual Enhancements

Using more tables, diagrams, and figures can simplify complex data representation. It's essential that diagrams of dam outlets and relevant structures are integrated where needed. Every visual must have comprehensive legends or explanatory captions for clarity.

Metadata and Provenance

Introducing a standardized metadata format allows for effective tracking of the provenance of model runs. This should encapsulate details like model version, data sources, and configuration settings, ensuring clarity and repeatability in the modeling process.

Consistency of Terminology and Units

It is imperative to establish a common lexicon. Ensuring that terminology is used consistently and holds the same interpretation for all stakeholders is essential for clarity and collaboration. Consistency in terminology and units is also vital. While the hybrid usage of U.S. customary and metric units is acknowledged, it is recommended to adopt one unit type for each parameter throughout. Finally, all acronyms and abbreviations must be defined at their initial use and listed in the beginning of each document. Their usage should be minimized for enhanced clarity.

Glossary of Terms

A dedicated section or document that defines all terms and acronyms used would be beneficial, especially for non-technical stakeholders. This can serve as a quick reference guide and enhance understanding.

Context

A brief introduction to the regulatory framework, Reclamation's management requirements, and temperature management points would offer readers a clearer perspective of the project's significance and context.

Future Pathways Document

After the completion of the WTMP, a forward-looking document that explores future avenues would be beneficial. This document should delve into potential implementations like forecasting, climate change assessments, and how the WTMP can be leveraged for these future challenges. It provides stakeholders with a vision and roadmap for the platform's evolution.

Question 13. What should be included in the models in the future to improve their accuracy, resolution, or other features?

The Review Panel has determined that the models adopted and developed by the WTMP Project Team are state-of-the-art for the designed application, especially when applied to the chosen spatial and temporal resolutions. No substantive changes of the WTMP models are recommended. However, improved accuracy could be achieved by securing meteorological and hydrological input values aligned with the models' time scale, thereby avoiding the need for disaggregation. While the models have been demonstrated to perform well, future improvement endeavors should focus on refining and extending the input data, workflows, and operating procedures. For example, forecast workflows need to be further developed to include restart procedures and updating model state.

One item to consider is the parameterization of the TCD in Lake Shasta and the shutter in Folsom Lake. Given that the model behavior is susceptible to temporal shifts influenced by structural processes and alterations in reservoir hydraulics, it is important to routinely monitor the temperature profile at the reservoir outlet. These measurements would help in maintaining and improving the parameterization, ensuring the models remain as precise and relevant as possible.