

complex model, such as CALSIM II. The model is clearly too complex, and not complex enough. The root of this difficulty is that when such a model is constructed, it is not clear what level of detail is needed, so the model must be made sufficiently complex to ensure it is complex enough. And the complexity needed to address some issues will remain in the model when it is used to address other less complex issues, or the same issues at less complex locations. One approach to addressing this issue is to develop different linkable modules of CALSIM II having different complexities. In this way the level of detail can be varied to be consistent the application or study at hand, and level of sophistication and resources available to the user.

Other weaknesses model users would like addressed include:

- The model provides limited and inadequate coverage of non CVP or SWP water and of the California water system south of the Delta.
- The model assumes that facilities, land-use, water supply contracts and regulatory requirements are constant over this period, representing a fixed level of development rather than one that varies in response to hydrologic conditions or changes over time.
- Groundwater has only limited representation in CALSIM II.
- Groundwater resources are assumed infinite, i.e., there is no upper limit to groundwater pumping.
- The linear programming model considers only the current month, and hence CALSIM II operating rules are required to determine annual water allocations, to establish reservoir carryover storage targets, and to trigger transfers from north of Delta to south of Delta storage.
- Better quality control is needed both for the model and its current version and the input data. Procedures for model calibration and verification are also needed. Currently many users are not sure of the accuracy of the results. A sensitivity and uncertainty prediction capability and analysis is needed.
- Need improved ways of altering the model's geographic scope and resolution and its temporal resolution to better meet the needs of various analyses and studies.
- Need to improve the model's comparative as well as absolute (or predictive) capabilities.
- CALSIM II needs better capabilities for analyzing economic, water quality, and groundwater issues.
- Need improved documentation explaining how the model works, its assumptions, its limitations, and its applicability to various planning and management issues.
- DWR and USBR have not provided a centralized source of support for CALSIM II. More training for CALSIM II is needed. There is a need for more people who can run CALSIM II. There is a need for a well-publicized user group. A more extensive users' guide is needed.
- Improved capabilities are needed for real-time operations especially during droughts, gaming involving stakeholders during a simulation run, handling of evapotranspiration and agriculture demand changes over time, water transfers, Delta storage, carryover contract rights, refuge water demands and more up to date representation of Feather River, Stanislaus River, Upper American River, San Joaquin River and Yuba River operations.

- Need an improved graphical user interface to facilitate input of model data, setting of model constraints and weights, operating the model, and displaying and post analysis of model results.
- Need to be able to change the model time period durations for improved accuracy of model results.

6. Limitations, Uncertainties, and Impediments

6.1 Absolute Values or Comparative Results

Modelers sometimes make a distinction between the use of a model for *absolute* versus *comparative* analyses. In an absolute analysis one runs the model once to predict an outcome. In a comparative analysis, one runs the model twice, once as a baseline and the other with some specific change, in order to assess change in outcome due to the given change in model input configuration. The suggestion is that, while the model might not generate a highly reliable absolute prediction because of errors in model specification and/or estimation, nevertheless it might produce a reasonably reliable estimate of the relative change in outcome. The panel is somewhat skeptical of this notion because it relies on the assumption that the model errors which render an absolute forecast unreliable are sufficiently independent of, or orthogonal to, the change being modeled that they do not similarly affect the forecast of change in outcome; they mostly cancel out. This feature of the model is something that would need to be documented rather than merely assumed.

In our opinion CALSIM II has not yet been calibrated or validated for making absolute predictions values. Yet it is apparent that there has been a distinct need by model users for absolute predictions. In the absence of alternatives, users are adopting CALSIM II results as the best absolute prediction available and they are likely to continue to do so. We recommend that model developers recognize the requirement for CALSIM II to provide absolute predictions. To satisfy this new purpose, additional calibration of the model will be required to ensure that the output it produces is fit for this purpose. Regardless of how possible it is to match the model closely with observed behavior, statistics on the accuracy of the calibration run should be supplied to users to enable them to gauge the likely errors involved with using the model output.

6.2 Sensitivity and Uncertainty Analyses

Sensitivity analyses would be useful to identify which parameters and input data have major impacts on decisions and system performance criteria of concern. Uncertainty analyses would help users of the model understand better the risks of various decisions and the confidence they can have in various predictions.

6.3 Graphical User Interface

input parameters to the model. This was done to a limited extent in the CALSIM II validation run with three regulatory periods modeled related to decisions made by the State Water Resources Control Board. It is also legitimate to incorporate growth in demand especially if that growth is described in a manner that is consistent with the way that demand is specified in the production run. Demand north of the Delta was specified in the validation run by inputting the historical crop areas.

A Calibration/Validation report should be very useful in demonstrating the accuracy of the model. However there are a number of elements in the CALSIM II validation run and the validation report which reduce that confidence including:

- State Water Project (SWP) demands south of the Delta were set at historical deliveries in years with no restriction and at the contractor's request level in restricted years. Neither of these pieces of information is available to a production run which calculates demand based on crop areas. Therefore the validation run does not provide reliable information on how well the model can represent these demands.
- The validation run omitted Article 21 deliveries. Although this omission will not affect the delivery of 'Table A' volumes south of the Delta, it will affect flow in the Delta and Delta water quality. Also, in the example model run presented in the paper by Draper A.J. et al (2003) which was supplied as part of the review, changes to Article 21 deliveries constituted the largest impact resulting from a change to the allowable pumping capacity at Banks between March and December. This suggests that the modeling of these demands is important.
- The DWR (2003) report produces estimates of SWP and Central Valley Project (CVP) deliveries south of the Delta but then adjusts them for changes in storage before presenting comparisons of those results with observed deliveries. This process merely checks that the model is preserving a water balance and does not present a legitimate validation of model deliveries.
- The report provides statistics on long term average deliveries and flows but no statistics on the fit for individual years. Additional analysis of the output would assist stakeholders to assess whether the estimate of water supply reliability and in particular the modeled volumes of water available in the most restricted years are accurate.
- In some instances, such as the examination of water quality in the Delta, the ability to accurately model monthly flows and deliveries will be important. The validation report contains no information that would enable the ability to model monthly flows to be assessed.
- A key model output is the water quality in the Delta. It would assist the validation of the model if a comparison of parameters such as the location of the X2 boundary was provided.

The users of CALSIM should recognize that models are a summary of what one believes to be true and important about a system. Validation is then an exercise to test how good that summary and understanding really is.

Appendix I contains brief descriptions of calibration modeling in the Murray-Darling Basin in Australia and in the State of Texas.