

Gerald Johns' Memo

Jan de Leeuw

October 23, 2005

These are brief comments on a Gerald Johns' memo to the Monterey Plus EIR Committee dd 90/22/05 re CALSIM II.

The argumentation in the memo is somewhat defensive and not always convincing. It is certainly true that the use of a model "to simulate actions in the system is valid and recognized practice". In fact, there is no alternative, because there is no physical copy of the system which can be used for experimentation. This does not imply, however, that the use of CALSIM II is "an acceptable and valuable analysis method". Such a conclusion certainly does not follow from the fact the models in general are useful, and asserting from the start that CALSIM II is valuable obviously begs the question if it is indeed valuable. The argument that DWR and other agencies have "consistently used CALSIM II" is not convincing either, unless it can be shown that this use has indeed led to better or more reliable predictions of supply than any of the available alternatives.

The memo does little more than summarizing the support materials that are available, without emphasizing that some of this material is quite critical of CALSIM II. The fact that CALSIM II is a cooperation between federal and state agencies may be important to federal and state bureaucrats, but it cannot be used as an argument in favor of the model. Similarly, the argument that using an optimization engine is what many other modeling

systems do is not a cogent argument, because it relies solely on authority and not on superior performance compared to other approaches. Intuitively, optimization engines do not seem to be very natural tools to model complex systems, because they assume all operators in the system are optimizers in possession of perfect information. Also, by its very nature, if you maximize deliveries to purveyors, then of course you overestimate actual deliveries.

The memo adds the by now standard DWR disclaimer that “Traditional model calibration and verification process is difficult to apply to a planning model, such as CALSIM II, that simulates operations and water supplies at a fixed level of development”. To me this has always seemed in appropriate testimonium paupertatis on the one hand, and a unconvincing attempt to have your cake and eat it too. Either CALSIM II is a model that describes the operations of SWP/CVP and is used to make predictions about future water supply to the purveyors, or it isn't. If it is, then it surely must be able to withstand traditional calibration and verification processes, because these are the only ways to assess its quality. If it is not such a model, then it clearly should not be used as such.

As I have said elsewhere, the main problems with CALSIM II is that it is only tested on a very limited set of realistic scenarios to assess its sensitivity, that it does not take the essentially stochastic nature of model inputs into account, that it optimizes deliveries instead of modeling actual deliveries, and that it has weak links between optimizations for successive time points. The actual implementation of CALSIM II has many additional problems. The documentation is virtually non-existent, the course of the program is not available, and the actual manipulations by DWR and its consultants to produce results from the program are shrouded with a mysterious layer of unassailable expertise.

In fact, at a recent CALSIM meeting in Sacramento, someone from the audience asked the chair of the peer review committee if anyone on the committee had ever tried to actually run the CALSIM II program. Somewhat surprisingly, the answer was no. For actual

CALSIM II results the peer review relied on runs provided by DWR and its consultants. It is clear that such a strategy implies that what we studying is not just the actual computer program, but it is whatever goes on when the experts are producing their results, including possible non-documented tweaks and iterative attempts to “improve” the results of the runs. The CALSIM II provides much opportunity for such tweaking, for instance by relaxing “hard” to “soft” constraints.

I want to make this point perfectly clear, admittedly exaggerating (although I don’t know by how much). There is no guarantee that the results in the SWP Delivery Reliability Report actually come from CALSIM II runs. The whole process of arriving at these results is not sufficiently transparent for others to be able to reproduce the results, and whenever that is the case a true peer review and true scientific scrutiny are not possible.