H-1

KIRKEENG

COMMENTS ON FINAL DRAFT

I like the approach that you have taken. Not being any sort of an expert in physical modeling I appreciate attempting to put numbers on calibration/verification. Be warned, the evaluation team has not been able to do this because of so many factors that influence each model. The "numbers" we have put together don't really seem to provide the whole picture the modeler needs. The best we have thus far requires a consideration of several pieces of information that must be jointly analyzed by the modeler.

I tried to understand your methodology but did not spend a significant amount of time on the KA case study nor the data presented on each of the 30 models (graphs). I hope to still spend time looking at this portion of the report.

Cumulative frequency plot
Unsure about the computations in developing the cumulative frequency curve

Is the y axis the value of the parameter and the x-axis is how many times that values of that parameter occurred in that range? Y-axis is the parameter value and the frequency. The number of times that parameter values fell within each class relative to the total number of values defines the frequency of values for each interval. The cumulative frequency curve describes the cumulative number of events up to that parameter value.

The number of classes is this dividing the parameter up into equal portions and then the frequency is how many times the value fell into the class? yes, this is similar to what is done when analyzing hydrologic data.
I am not sure that I understand how these plots are developed. Is the data divided up into classes (ranges) and how many data values in the range is the frequency? Virtually the procedure used in analyzing hydrologic data, esp. for developing duration curves.

Is the overall or eventual concept to apply this measurement technique to every model study in the future? Present evaluation does not suggest using this approach, just parameter plots by range will be used/recommended. My preference is to continue the CF line of thinking with the aim of helping define some confidence bands that can be used in assessing how well the model agrees with the prototype. One problem with using the CF approach (and why it needs more work) is that the curves as presented don't give you any idea of where in the model reach you are for each of the parameter values -- you may have the same number of "large" values in model and prototype, but they could be in different places.

How were the ranges selected, throughout the whole model or just a certain study area? Was the same criteria used for selecting ranges for small scale vs large scale models? Range locations were selected somewhat arbitrarily, but the large models ant the 2 kate-aubrey micromodels tried to locate ranges where prototype survey ranges were taken -- prorotype ranges are locations where historical survey data have always been taken.

Can we derive from this some quantification of what values of each parameter are needed to call a model calibrated/verified? It is unlikely that such a quantification can be achieved because so many variables are involved. For instance, the amount of variability

in the prototype has an influence on how well the model can replicate the bathymetry -- a reach with less variability, a bedrock channel, should be "easier" to model than one that has bed elevations that rise and/or fall by 30 feet or more (with a typical channel depth of 25 feet) You can easily see the difficulty in coming up with any definitive guidelines. However, it may be possible to expand the CF technique to provide confidence bands that one could use in **interpreting** both the state of calibration AND MAYBE the changes resulting from alternative analysis. For this to happen, there has to be a way to link location in the reach with the actual parameter values. As stated in previous comments, this method looks good (in my personal opinion) but has a way to go.