Discussion of “Crash tests for hydrological models” by V. Andreassian et al.

Referee Comment by John C. Schaake

Recognizing that all hydrological models are inherently limited approximations of real-world catchments the authors explore the familiar question of how should we test models to understand their strengths, limitations and opportunities to improve them. A key hypothesis is that models ideally should be tested with large data from a large number of basins from a wide range of physical and climate regimes.

I enjoyed reading this paper because it presents a nice summary of different rationales for how to evaluate models and approaches to understanding model limitations. I completely agree with their final conclusions that: (i) it would be valuable to set up a large international data set for testing models and (ii) further progress in hydrologic modeling will come in part from intercomparisons based on large data sets.

I’d like to suggest that an important question to consider is “what can hydrologists do to make our models most useful to society”? I’m sure there are many ways to approach this question, including: make better models. One approach is from the perspective that models are often used to get information (e.g. predictions) about the state of a hydrologic system of interest. But how much do we really know about how well any of our models can meet this need for information – before we apply a model or if we apply a model to an “ungaged” location?

Looking at models from the perspective of how much information they provide is a way of recognizing that models are not perfect and that statistical as well as physical principles may be of value to get the most information to meet a particular need. When we began to develop the MOPEX project we wanted to improve our ability to make a priori estimates of hydrological model parameters so that we might make better predictions for ungaged areas or that we might improve model parameter estimation by having better a priori constraints. I remember one study where I tried to correlate calibrated values of model parameters with physical and climatological variables for many of the MOPEX basins in the United States. I also tried to see if I could make an a priori prediction of measures of model error of the calibrated model. Surprisingly, I found I could predict measures of the model error much better than I could predict any of the calibrated model parameters.

Many hydrologic systems of practical interest are influenced not only by natural hydrologic processes but by upstream human influence, including unknown influences, both natural and human. Moreover, our measurement systems are less than perfect too. Therefore, including statistically as well as physically based tools in our tool box should help us to get the most value for society from our models. But we need data from more than one country to do that.

One of the ways we might use a large, diverse international data set is to test how well we can predict various measures of the information content of our models.