Geophysical Research Abstracts
Vol. 18, EGU2016-9865, 2016
EGU General Assembly 2016

A new fit-for-purpose model testing framework: Decision Crash Tests

Bryan Tolson and James Craig
University of Waterloo, Civil & Environmental Engineering, Canada

Decision-makers in water resources are often burdened with selecting appropriate multi-million dollar strategies to mitigate the impacts of climate or land use change. Unfortunately, the suitability of existing hydrologic simulation models to accurately inform decision-making is in doubt because the testing procedures used to evaluate model utility (i.e., model validation) are insufficient. For example, many authors have identified that a good standard framework for model testing called the Klemes Crash Tests (KCTs), which are the classic model validation procedures from Klemeš (1986) that Andréassian et al. (2009) rename as KCTs, have yet to become common practice in hydrology. Furthermore, Andréassian et al. (2009) claim that the progression of hydrological science requires widespread use of KCT and the development of new crash tests.

Existing simulation (not forecasting) model testing procedures such as KCTs look backwards (checking for consistency between simulations and past observations) rather than forwards (explicitly assessing if the model is likely to support future decisions). We propose a fundamentally different, forward-looking, decision-oriented hydrologic model testing framework based upon the concept of fit-for-purpose model testing that we call *Decision Crash Tests* or DCTs. Key DCT elements are i) the model purpose (i.e., decision the model is meant to support) must be identified so that model outputs can be mapped to management decisions ii) the framework evaluates not just the selected hydrologic model but the entire suite of model-building decisions associated with model discretization, calibration etc. The framework is constructed to directly and quantitatively evaluate model suitability.

The DCT framework is applied to a model building case study on the Grand River in Ontario, Canada. A hypothetical binary decision scenario is analysed (upgrade or not upgrade the existing flood control structure) under two different sets of model building decisions. In one case, we show the set of model building decisions has a low probability to correctly support the upgrade decision. In the other case, we show evidence suggesting another set of model building decisions has a high probability to correctly support the decision.

The proposed DCT framework focuses on what model users typically care about: the management decision in question. The DCT framework will often be very strict and will produce easy to interpret results enabling clear unsuitability determinations. In the past, hydrologic modelling progress has necessarily meant new models and model building methods. Continued progress in hydrologic modelling requires finding clear evidence to motivate researchers to disregard unproductive models and methods and the DCT framework is built to produce this kind of evidence.

References:

Andréassian, V., C. Perrin, L. Berthet, N. Le Moine, J. Lerat, C. Loumagne, L. Oudin, T. Mathevet, M.-H. Ramos, and A. Valéry (2009), Crash tests for a standardized evaluation of hydrological models. *Hydrology and Earth System Sciences*, 13, 1757–1764.

Klemeš, V. (1986), Operational testing of hydrological simulation models. *Hydrological Sciences Journal*, 31 (1), 13–24.