



Evaluating prospective hydrological model improvements with consideration of data and model uncertainty

James Craig (1), Nicholas Sgro (2), and Bryan Tolson (3)

(1) University of Waterloo, Civil and Environmental Engineering, Waterloo, ON, Canada (jrcraig@uwaterloo.ca), (2) University of Waterloo, Civil and Environmental Engineering, Waterloo, ON, Canada (nasgro@uwaterloo.ca), (3) University of Waterloo, Civil and Environmental Engineering, Waterloo, ON, Canada (batolson@uwaterloo.ca)

New algorithms for simulating hydrological processes are regularly proposed in the hydrological literature. These algorithms are often promoted as being more physically-based or better at capturing hydrologic phenomenon seen in the field. However, the tests used to evaluate the effectiveness of these algorithms are typically no more than history matching – an improved model hydrograph is (often inappropriately) interpreted as an improved model. Here, a simple and more stringent method is proposed for comparing two model algorithms in terms of their ability to provide distinguishably different validation results under the impact of uncertainty in observation data and forcings. A key output of the test is whether results from two model configurations are fundamentally differentiable. This test can be used both to support improved algorithm development, but also to aid in hypothesis testing about watershed functioning or to support model selection. As may be expected, our ability to identify the preferred hydrologic algorithm is significantly diminished when model/data uncertainty is incorporated into the evaluation process. The information content of the data and compensatory parameter effects play a key role in our ability to distinguish one model algorithm from another, and the results suggest that simpler models justified by the available data may have more utility than complex physically-based models which can fit the data at the cost of poor validation performance. They also suggest that finding the “best” model structure is (unsurprisingly) dependent upon both the quality and information content of the available observation data.