Hydrology and Earth System Sciences Discussions

Interactive comment on "*HESS Opinions* "Crash tests for a standardized evaluation of hydrological models"" *by* V. Andréassian et al.

Anonymous Referee #1

Received and published: 16 June 2009

I enjoyed the paper, especially because it is well written, it has an extensive list of references, and it well describes some of the problems associated with model evaluation. For this reasons, I recommend that it will be published after some small revisions. However, while the paper well states the problems, I do not think it indicates a convincing solution.

First of all, as the authors mention, I am a believer that uniqueness of place reflects into uniqueness of models. As a result, I am not in favor of Klemes' four level testing scheme. From this it follows that I do not think that a model should be submitted to a wide range of natural conditions. Also, I do not agree with the analogy that as cars should be able to drive on every road, also models should be able to cope with a variety

of hydrologic conditions. The fact that cars are built to satisfy many needs makes them often over-accessorized and unnecessarily expensive.

As Kirchner (2006, your paper) writes, the fact that models may stand a chance of failing the test to which they are submitted is an advantage, as model failure gives a chance of learning new aspects of catchment behavior. A model that is built to pass all tests will be unnecessarily over-parameterized for a specific application.

However, this does not mean that research should focus on the idiosyncrasies of individual catchments (McDonnell et al., 2007). I think, instead, that model comparison and catchment classification is a promising way of learning something new in the field of hydrology.

Overall, in my opinion, there are model weaknesses that are desirable and others that are undesirable. Crash tests should be performed for those weaknesses that are not desirable, such as bugs in the code, numerical errors, wrong units conversion, and so on. The fact that a model may fail when confronted with additional data, or when applied to a different catchment, is in my opinion an advantage, as it is part of the process of scientific investigation (e.g. Fenicia et al., 2008).

Although I have a different opinion on some aspects, I respect the ideas of the authors, and I consider this a paper that deserves publication, maybe after some reinforcement of the last part, where suggestions for possible crash tests are given.

McDonnell, J. J., et al. (2007), Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, Water Resour. Res., 43, W07301, doi:10.1029/2006WR005467.

Fenicia, F., J. J. McDonnell, and H. H. G. Savenije (2008), Learning from model improvement: On the contribution of complementary data to process understanding, Water Resour. Res., 44, W06419, doi:10.1029/2007WR006386

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 3669, 2009.