Hydrology and Earth System Sciences Discussions

Interactive comment on "Performance and reliability of multimodel hydrological ensemble simulations based on seventeen lumped models and a thousand catchments" *by* J. A. Velázquez et al.

M. Zappa (Referee)

massimiliano.zappa@wsl.ch

Received and published: 8 September 2010

In recent times several strategies have been proposed in order to apply ensemble techniques in catchment hydrology. Several realization have been proposed with hydrological models being driven by an ensemble input meteorology. The setup presented here by Velázquez et al. is by far the most interesting experiment with "Multimodel Ensemble" that I'm aware of. It is of course a mostly relevant contribution to HESS.

C2160

The robustness of the paper is guaranteed by both the large number of models (17) and of catchments (>1000). Also the 10 years of available data are adequate for sound calibration and verification of the experiment. It would be nice if the authors would spend some words on the climatology of these 10 years as compared to long-term time series.

The authors are able to point to different aspects of using and evaluating "Mulitmodel ensembles" and are able to give enough information to well support their interpretations.

The selected catchments include basins with different characteristics and located in different climate regions. I wonder if all basins are headwaters (with areas between 10 and over 32000 km2) or if the biggest basins are combining the results from sub-basins. The referred work of LeMoine et al. (WRR, 2007) does not answer this doubt either. LeMoine et al. used less basins and the biggest one is smaller than 10000 km2.

So here splits my comment:

a) If sub-basins are used, then I wish the authors would indicate if routing has been adopted to combine small headwaters to large river basins and that they spend some words on the way they dealt calibration of nested sub-basins.

b) If all basins are headwaters, then how do the 17 different model deal with the routing? Do some of the parameters (e.g. storage coefficients) indirectly assume values that allow accounting for internal flow times for the larger river basins?

I assume that most of the used lumped models have been originally developed for mesoscale catchments with areas not exceeding 200 to 3000 km2). I ask the authors to explain their strategy since it is a real challenge to combine nested basins for obtaining ensemble hydrographs for large river basins (e.g. Jaun et al., 2008).

For this reason I also support the requirement of the other reviewer asking some indications on the dependence of the model results to the catchment size.

The choice of the score is always difficult. The selection presented here is adequate to obtain the information needed for evaluating the experiments.

Figure 12: The authors state that their findings "supports the finding of Viney et al. (2009) that the best ensembles are not necessarily those containing the best individual models, but it seems that the inclusion of some good models is essential". My experience with ensemble verification is that it is important, that the selected ensemble members present an adequate spread. An example of rank histograms with catchments sub-sets would allow to verify if the spread of the subset ist higher than the one of C0, C1 and C2.

Finally I can only recommend this paper for publication in HESS.

Literature: Jaun, S., Ahrens, B., Walser, A., Ewen, T., and Schär, C.: A probabilistic view on the August 2005 floods in the upper Rhine catchment, Nat. Hazards Earth Syst. Sci., 8, 281-291, doi:10.5194/nhess-8-281-2008, 2008.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 4023, 2010.