Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-584-RC2, 2017

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Performance of ensemble streamflow forecasts under varied hydrometeorological conditions" *by* H.-J. F. Benninga et al.

Anonymous Referee #2

Received and published: 15 February 2017

General comments:

This paper summarizes the application of the widely used HBV hydrologic model to streamflow forecasting in a Polish mountain river. The project uses ECMRWF ensemble weather forecasts to drive the streamflow model, and explores both pre- and post-processing of the ensembles for bias correction. Useful results are obtained, and the study has significant potential. I recommend that the paper is accepted pending major revisions.

Detailed comments:

1. The paper repeatedly refers to HBV as a spatially lumped model. This isn't just

terminology, as around lines 20-25 of page 15, the manuscript seems to imply that the model assumes a snowpack to be present (or absent) across the entire model domain. There are a few versions of HBV, but it's normally viewed as semi-distributed, using (at a minimum) elevation bands.

2. The manuscript makes a good point on lines 29-30 of page 1 about socio-economic development increasing the impacts of extreme hydrometeorological events. It also probably bears mentioning that climate changes, both natural and anthropogenic, may further exacerbate these impacts. See Perkins, Pagano, and Garen, "Innovative operational seasonal water supply forecasting technologies," Journal of Soil and Water Conservation, 2009; and Fleming, "Demand modulation of water scarcity sensitivities to secular climatic variation: theoretical insights from a computational maquette," Hydrological Sciences Journal, 2016.

3. Terms could stand to a little better defined. For example, most flood and water supply forecasters who I know would regard "short-term" forecasts as having lead times of 0-10 days, and "long-term" forecasts as having lead times of weeks to months. So what the authors refer to here as "medium-term" would be referred to as "short-term" by many if not most others working in the field. And no effort is made here to distinguish medium-term from short-term hydrologic forecasting. More broadly, some of the wording throughout the manuscript would benefit from a re-think for better clarity and precision.

4. Why is only meteorological forecast uncertainty incorporated into the ensemble model? It's commonplace in the research literature for forecast models to include both meteorological uncertainty (NWP ensemble) and hydrologic model parameterization uncertainty (ensemble of hydrologic parameter values). This work is starting to make its way into operational practice too. Providing some justification for this choice might be a good idea.

5. The description of the model implementation isn't quite adequate. What was the

calibration-testing split, and what were the model performances during both phases? And it's stated that the objective function selected for calibration is "Y", which apparently combines the Nash-Sutcliffe efficiency with a volumetric error measure. Objective function selection is a key step in model calibration, and more information needs to be provided, starting with an explicit mathematical definition for "Y".

6. The updating of initial states was performed here for the slow-runoff and fast-runoff reservoirs. That's interesting and useful, but why was SWE not selected as the object of this data assimilation exercise? It seems like it would be a more rewarding, and certainly more conventional, choice in this northern continental European mountain catchment.

7, The literature review of ensemble hydrologic forecasting, pre- and post-processing for bias corrections, and data assimilation and model updating, is a good start but seems a little light. Citing more work would provide valuable context to the paper. A reasonable place to start might be recent work by Dominique Bourdin at the University of British Columbia and Hamid Moradkhani at Portland State University.

8. Some of the specific conclusions seem a little surprising. That's great, but it also means they'd benefit from additional discussion. In particular, the paper concludes in section 4.2.1 that the quality of the forecasts at lead times of less than 3 days is dominated by hydrologic initial conditions, and the weather forecasts become the dominant source of predictive skill after that. This would be a reasonable conclusion for a large or flat basin, but for a small, steep mountain river it seems a little surprising – these are typically flashy systems that respond to rain or snowmelt inputs within a day or so. Indeed, a few pages later near the end of section 5, the paper states that "in the hydrological model the lag time between a rainfall event and the streamflow peak is set to 1 day." It also seems that conclusions like this, which attempt to attribute predictive skill (and therefore also predictive error) to various different sources, might be difficult to make convincingly without using a more statically sophisticated and exhaustive data assimilation procedure, incorporating ensembles of hydrologic models and/or model

parameters, etc.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-584, 2016.