

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

Anonymous Referee #3

Received and published: 15 February 2017

The authors proposed a methodology to give insight in the performance of ensemble streamflow forecasting systems in three streamflow categories (low, medium and high) and related runoff generating processes from lead times of 1 day to 10 day with a case study in a mountainous river catchment of less than 1000 sq km in Poland. The quantitative precipitation forecasts and temperature forecasts extracted from the European Centre for Medium-Range Weather Forecasts (ECMWF) are averaged with catchment as input of a lumped hydrological (HBV) to generate ensemble streamflow. Several intensively used verification measures (CRPS, CRPSS, Rank histogram, Reliability diagram and ROC) are selected to evaluate the ensemble forecasts. Additionally, the pre-processing, post-processing and updating of model initial states are adopted to

C1

improve the behavior of the system.

Generally speaking, the study gave an interesting investigation on the assessment of hydrological ensemble prediction system on different runoff processes including snowmelt, short-rain flood and so on, and a further analysis was made on the uncertainty source of these varied hydrometeorological conditions. There I suggest accept this manuscript after a moderate revision.

There are a few issues list below that the authors should address: 1) The logic in Paragraph 2 and 3 of Section 1 needs to be perfect. Some irrelevant statements can be removed, eg. SOME CONTENTS from Line 10 to Line 15 in Page 2 about EFAS are unnecessary to some degree. 2) Lines 18-20 Page 6: A further explanation is expected why the training period is defined from 2011-2013 while the years previous to 2011 is used to validation. 3) In Section 3.2, it is not necessary to introduce all the evaluation scores in details, for the CRPS, CRPSS, Reliability diagram and ROC can be regarded as "industry standards" in ensemble forecasting, so simply citing the relevant references. 4) In Section 4.1.2, it is confusing that since the QM pre-processing brings improvement to the precipitation and temperature forecasts, why the conclusion is that the strategy 0 results in the best CRPS. 5) The figures about rank histograms and reliability diagrams are missing or not shown intentionally? 6) The catchment area is less than 1000km² and the data used are daily. For flood forecasting in such catchment area, is it daily data too coarse? Perhaps 3h or 6h subdaily data are more useful for flood forecasting in such area. Please make it an elaborate story. 7) For flood forecasting, flood peak, volume and peak time are all important. Can these be analyzed in the study? 8) Page 9: It is not very clear how the errors are contributed in Section 3.3. Why can CRPSSim/CRPSSmeans represent the error contribution? Please add more details.

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

C2