Hydrol. Earth Syst. Sci. Discuss., 6, C894–C897, 2009
www.hydrol-earth-syst-sci-discuss.net/6/C894/2009/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Evaluation of a probabilistic hydrometeorological forecast system" *by* S. Jaun and B. Ahrens

S. Jaun

simon.jaun@wsl.ch

Received and published: 25 May 2009

We would like to thank the anonymous Referee #3 for his positive and constructive comments. Here we want to address the issues raised by referee #3 and comment on possible modifications to the paper.

1. If the goal of the verification is to analyze the hydrological forecast quality based on weather forecasts, it will give some misleading results, if you not take into consideration the time of concentration of the catchment. For example in Fig. 6 the box-plot for the Rhine catchment (C23) for leadtime 1 will probably not show any effect of the forecasted precipitation, but will only indicate how well the hydrological model is able to reproduce the routing. I don't know how long the routing effects will last, but for a catchment of this size, I could imagine that the forecasted precipitation will influence

the outflow of the whole catchment after 1 to 3 days. That is also one reason, why the ensemble spread of C23 is that small. This is important also for all kinds of skill score measurements integrated over different catchments with different time of concentrations.

The goal of this evaluation study is to analyze the entire forecast system. By merging the results from the different catchments, we try to estimate the general performance of HEPS with regard to two reference forecasts (HDET and HART) without restrictions due to climatological/topographical conditions at the catchment scale. The box-plot for C23 in Fig. 6 does show the effect of the forecasted precipitation for all leadtimes. Spread for leadtime 1 is of course relatively small, but the ensemble members only differ in the values of the meteorological variables used as input. In the case of a forecasted large scale event (or a local event in the northern part of Switzerland), the contributing catchment area for C23 grows quickly and shows up in the (small) ensemble spread. On the other hand, a forecast of a local event in an alpine catchment will not be reflected in the ensemble spread of C23 (for short leadtimes due to the concentration time issue, for longer leadtimes due to averaging). However, this is the real forecast situation and as we treat all forecasts the same way, none of them should benefit. Nevertheless your comment regarding the concentration time is absolutely justified and the fact that ensemble spread/error can additionally be influenced by the concentration time of a catchment should have been mentioned in the discussion of Fig. 6. We will include this in the revised version of the paper. Using HREF as runoff reference eliminates errors emerging from the hydrological model itself, therefore the results are not dependent of the routing abilities of the hydrological model. Anonymous referee #2 expressed similar concerns (that the combination with larger catchments could push the scores). But ruling out the 6 largest catchments does only result in a minor reduction of skill scores and leads to the same conclusions. We actually thought about estimating the contribution of each of the downstream subcatchments separately, but decided against it, as we would possibly need an additional routing model and measurement errors show a higher impact.

HESSD

6, C894–C897, 2009

2. Maybe it would be worth to construct the spread of the artificial ensemble (HART) in a different way without using the ensemble median. Instead of the application of the linear correlation between the median and the sorted ensemble members, a quantile regression could be applied directly by using the reference discharge as depended variable.

The intention of constructing HART was to mimic a deterministic forecast with some sort of uncertainty information. Of course, the possibilities to construct such uncertainty information based on climatological informations are manifold, and the quantile regression is generally more robust than a linear regression. But using the reference discharge as base of HART would additionally eliminate lead-time dependent errors/biases. Also, it is not possible to construct the ensemble in a forecast situation, as the reference simulation/observation is not yet available. We deliberately chose the HEPS ensemble median as base for HART to assure that the two ensembles only differ in terms of the given uncertainty information. Using HDET (or HREF) would result in an advantage (disadvantage) for HEPS (cf. deterministic evaluation of the HEPS median). One could of course base the quantile regression on the ensemble median, but we found the idea of using climatological ensemble model information more appealing, an information which would be lost using the quantile regression. This approach also shows that the knowledge of model output statistics (regarding the link between forecast value and spread/range) is not sufficient to meet the performance of the ensemble forecast system.

3. In order to compare the deterministic and the probabilistic forecast quality the operational value of the continuous forecast (Laio and Tamea, 2008) could be calculated. Even when the cost-loss function in this methodology is maybe over-simplified, it is an appropriate way to compare different forecast systems (including deterministic ones) and taking all (continuous) data into account without restrictions (breaking the data into categories).

Evaluating the operational value of the continuous forecast is another interesting ap-

HESSD

6, C894–C897, 2009

proach for this kind of problem. Nevertheless we chose a two-step approach using the median ensemble forecast and a forecast with artificial spread. We think that this concept is easier to communicate. We could have used the CRPS to avoid the use of categories (as suggested by anonymous referee #1) but opted against it, as it is common practice in operational hydrological forecasting to work with several warning-levels.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 1843, 2009.