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

Anonymous Referee #3

This paper demonstrates the usefulness of applying the COSMO-LEPS in combination with a hydrological forecasting system. It is well written, explaining the methodology thoroughly and showing most appealing results. Just a few comments regarding the verification: 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. 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. 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).