Hydrol. Earth Syst. Sci. Discuss., 8, C2315–C2318, 2011
www.hydrol-earth-syst-sci-discuss.net/8/C2315/2011/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Evaluating uncertainty estimates in hydrologic models: borrowing measures from the forecast verification community" *by* K. J. Franz and T. S. Hogue

Anonymous Referee #3

Received and published: 21 June 2011

The paper tackles a very important point for the scientific community involved in hydrologic predictions. In fact, the recent development of probabilistic prediction methods has not been accompanied by an adequate advance in techniques for the evaluation/verification of probabilistic forecasts. Some of the existing verification techniques have been developed in the field of meteorology, where probabilistic forecast were first introduced; other tools comes from the econometric and statistical fields where risk is associated with convincing economic reason. Some of these verification techniques are presented in this manuscript in the honorable attempt of promoting a wider diffusion and use of such tools in hydrology.

The paper uses a model for deterministic prediction of river flow (SAC-SMA) and three approaches (GLUE, W-GLUE and SCEM) to identify parameter sets for the generation of ensemble predictions. These ensembles are considered representative of the distribution characterizing the ensemble forecast and they are used to present, test and comment on a set of verification tools presented in Section 2.4. The main comments to the paper are the following:

1) Ensemble predictions are operationally appreciated as indicative of the probability distribution of future (predicted) values, however they do not represent the full forecast probability distribution over all possible values. As stated by the authors, ensembles aim at modeling the parameter uncertainty, but I disagree on that "the approach outlined here is readily transferable to evaluation of uncertainty from all potential sources of error" [page 3089, lines 1-2]. How it would be possible to model the uncertainty associated for example to "model structure, input (forcing) data and validation data" with the approach presented? Moreover, the ensemble approach is not the only form of probabilistic forecasts; I suggest mentioning the existence of different probabilistic methods in hydrology, for example the important contribution by Krzysztofowicz (2001, The case for probabilistic forecasting in hydrology, J. Hydrol., 249, 2–9).

2) The verification tools presented in the paper (Section 2.4) are various and addressed at evaluating different characteristics of the ensemble prediction, most of them intended at evaluating its statistical correctness. In the hydrologic context, however, a rather important aspect is the operational value of the probabilistic forecast, that has to be accounted for when weighting the probability associated to an extremely high (or low) flow. As briefly mentioned by the Authors as well: large confidence bounds might correctly include the data, but are not useful by an operational point of view. The concept of operational value and consequent evaluation of probabilistic forecast is detailed for example in Laio and Tamea (2007, "Verification tools for probabilistic forecasts of continuous hydrological variables", HESS, 11, 1267-1277), which propose to evaluate the probabilistic method through i) the expected costs associated to the predicted distri-

bution and ii) a cost/loss function modeling the risk severity. This kind of evaluation criteria, borrowed from meteorologists and developed for continuous (hydrologic) variables, is relevant for this paper and should be considered.

3) Tools for evaluating the statistical correctness of probabilistic forecast miss at least one major point well outlined in the forecast verification literature in the field of Econometrics. Christoffersen (1998, Evaluating interval forecast, Int. Econ. Rev., 39, 841–862), for example, brings the attention to the importance of independence in the sequence of inclusion/exclusion of data with respect to interval forecasts or confidence bounds. In other words, events inside (or outside) the intervals should not come clustered together, and the sequence must be random. Christoffersen's test for conditional coverage might be too strict for hydrologic variables with long memory and strong autocorrelation, such as streamflow, but the statistical correctness of interval forecast requires a test of independence. Such point should be discussed in the paper.

4) On a different chord, a remark to the paper is related to the organization and effectiveness of information conveyed. The Result section is too long and overly detailed: 8 HESSD pages and 9 Figures (not considering Figure 1), with a total of 53 panels, challenge even the most interested reader. As a consequence, important contents cannot be distinguished from minor details/differences and the key messages are missed. In my opinion, the Result section should be drastically reduced and only illustrative cases reported; figures don't have to cover all cases but only the most informative ones which are functional to the key messages the Authors want to convey. Only with a strong reorganization, the paper will provide the information in a more effective and readable way.

5) A final, but not less important point is that in the Discussion session is not clear if the probabilistic forecast methods or the verification tools are being evaluated. The main goal of the paper, drawn from title and introduction, seems to be the presentation and interpretation of the forecast verification measures, while the first half Discussion focuses on the performance of GLUE, W-GLUE and SCEM, creating confusion about

the ultimate objective. This impression is also strengthened by the detailed presentation of the three probabilistic frameworks in Section 2, and probably also by the several details given in the Result section. In the paper reorganization, some attention could be devoted to focus and emphasize the real goal, avoiding misunderstandings.

Concluding, my suggestion for the paper is the publication in HESS only after major revisions.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 3085, 2011.