Hydrology and
Earth System
Sciences
Open Access
Discussions

Interactive comment on "Benchmarking hydrological models for low-flow simulation and forecasting on French catchments" *by* P. Nicolle et al.

Anonymous Referee #2

Received and published: 6 January 2014

The manuscript reports on a model intercomparison project (MIP) on low flow simulation and forecasting in a number of rivers in France. It is well written, and methods are generally clear, with some exceptions (see below). The study is certainly relevant and could fill a niche as no such experiments have yet focused on low flows. Reading a paper with this topic I expected more specific conclusions on what needs to be done to better simulate low flow thoguh – this information, i.e. the relation to of the results back to the actual model differences with respect to low flow modelling are still a bit weak. In order for the paper to make a convincing and useful contribution to wider hydrological sciences, therefore, it needs to be more focussed, preferably shifted from a wide

reporting of the (provocatively put: rather boring and not directly transferrable) very detailed results on all individual models and score measures to e.g. the (commendably: really interesting and useful) questions already now addressed in the discussion section. The manuscript would then have the potential for a much wider impact.

Major comments:

1) As said above, the interesting aspects are currently hidden in the discussion section. On the other hand a proper discussion relating the results to the literature is missing. The analysis done is largely sufficient for example to make these discussion questions into the main objectives of the manuscript and thus make it visibly more useful to a wider community than the French participants of this specific MIP.

2) The frequently repeated conclusion that "model performance depends more on the catchment than on the choice of model", which is presented as the main outcome of the results as they are presented now, needs to be better supported. The analysis doesn't prove this consistently e.g. by testing the degree of systematic ranking of models versus the degree of systematic ranking of catchments with the same approach.

3) Related to the previous point is the highlighted but unsupported not-found link to catchment characteristics. For example: p13995 lines 18/19 What is the basis of this statement "satisfactory streamflow simulation seems to depend more on catchment characteristics than on the model"? It contradicts the next sentence and the sentence in the abstract that "all attempts to relate model efficiency to catchment characteristics remained inconclusive". Two of the three predictors tested are not catchment characteristics, but streamflow characteristics. Only drainage density is a catchment characteristic. In any case what are the hypotheses that these three should be influencing model performance? If influences are tested statistically there should be a hypothesis. Besides characteristics, also processes not represented by the models will hopefully influence the performance. What these might be could also be the basis for such tests and lead to more transferrable results. Here the approach needs to be clearer and

more focussed.

4) One methodological aspect the functioning of which and the relevance for the results needs to be better explained and elucidated are the post-correction methods and the assimilation methods. As these are related to model application rather than to the models themselves they can distort results. It would have been useful to test these somehow separately. But at least their effect needs to be discussed in detail with reference to literature (see comment on lacking discussion in general).

5) The approach of constructing scenarios appears not well suited to look at the two very severe events in forecasting mode, as this results in most 'spaghettis' above the actual low flow. The value of this sub-analysis should be reconsidered or it should be assessed more in terms of the models themselves. What do those that can model these extreme events well have that others don't?

Specific and technical comments:

6) The abbreviations and variable names of the skill scores are very variable with some being words, some three or four letter abbreviations, some one letter variables. This makes the manuscript difficult to read. Perhaps they could be homogenized in their presentation or less could be selected (see comment above on "too detailed")

7) Catchment yield (Table 2), Runoff Yield p. 13994, line 4. Choose one term. "Runoff Ratio" may be the more common term, anyway.

8) Last sentence of the abstract was unclear to me without having read the paper.

9) Last sentence of 1.1 is a bit disconnected and surprising and would require some further information of why this is mentioned at all.

10) p. 13983 line 8ff This sound a bit complicated. Why can it not be called ensemble low flow forecasting (similar to ensemble flood forecasting)?

11) 1.3 The list of models or forecasting systems in France is of little use to the paper

if no information on the specific experiences or strengths and weaknesses made with them are provided; i.e. what is it about them that informed the present study.

12) Tables 4 and 5 are really just lists and could be included in the text instead.

13) The results text on the maps mostly refers to the name of the river/station. If the text is written like this, Fig. 1, which the reader can use to find the locations, should show the names. Looking up the number in the table, then the location in Fig.1 and then read the results again is an unreasonable demand on the reader. Alternatively, the text could refer to the numbers in Fig. 1., but names may be preferable for readers with some knowledge on French hydrography.

14) Figure 10 is unreadable in the version published by HESSD and needs to be improved. I cannot distinguish the line colors in the plot and the lines do not correspond to the thickness in the legend, which would be correct.