Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-387-RC1, 2017

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Verification of ECMWF System4 for seasonal hydrological forecasting in a northern climate" *by* Rachel Bazile et al.

K. Förster (Referee)

kristian.foerster@uibk.ac.at

Received and published: 21 July 2017

General comments

In this research article, the authors present a detailed study on the predictive skill of hydrological ensemble forecasts in 10 watersheds in Québec, Canada. Different methods are employed which are subject to different degrees of complexity. Among these methods, a simple application of historical streamflow data is seen as benchmark for more complex approaches. The second method, the ESP approach is based on historical meteorological data and accounts for initial conditions in each forecast. The initialization using known system states is also relevant in the third approach, a dynamical seasonal forecast method, in which meteorological forcing is obtained from

bias-corrected climate model forecasts (ECMWF's System4). Given a lead time of 1 month, the dynamical approach provides improved skill in terms of Continuous Ranked Probability Score (CRPS), while for longer lead times the predictive skill is similar to the corresponding ensemble forecasts using ESP. For the period of snowmelt in spring, the CRPS is lowest (best) in the case of ESP and the dynamical approach. In some watersheds, however, the first method which provides forecasts using historical streamflow data performs best. This comparison highlights the fact that the predictability is low in some watersheds. The study is interesting, the results are promising and the paper fits very well into both the special issue on "Sub-seasonal to seasonal hydrological forecasting" in particular and HESS in general. The methodology is comprehensively presented and the results are discussed in a balanced way. Related work and relevant references are mentioned and acknowledged. Especially the assessment of added value provided by each increased level of complexity (using streamflow data only -> ESP -> dynamical forecasts) is very useful. Another important point is that the paper presents a specific case study in which operational forecasts have already been issued and new methods are going to be implemented. This might be relevant for other forecasting centers. However, in my opinion, the paper needs a few minor revisions and technical corrections. It's my impression that the section on reliability seems to be detached to a certain degree given that the findings from this analysis are not really considered in the summary. Moreover, the paper would benefit from some additional explanations that might improve comprehensibility. Please find my suggestions below.

Specific comments

Page 1, line 9: the abbreviation "corr-DSP" is not explained in this context and might be omitted here

Page 1, line 9: Would it make sense to point out that "Simulated streamflow computed using observed meteorological data is used as benchmark."?

Page 2, lines 12-28: In this section, historical streamflow prediction (HSP) and ex-

tended ensemble streamflow prediction (ESP) are presented. In my opinion, some additional explanations might be helpful in this context. You could explain that using HSP is in general possible without using a hydrological model, even though, in particular, you involve the output of a model in your specific case study. ESP, in contrast, does require a hydrological model in order to improve forecasts through explicitly incorporating initial states in the forecasts. The relevance of using hydrological models, as already pointed out, might be helpful in the process of understanding the different methods you apply.

Page 3, line 2: I am not sure whether “questioning” is the appropriate verb in this context. As far as I know it would make sense if you have reason to doubt the usefulness. Instead, using “assessing” might be a better option.

Page 3, line 14: Please add appropriate references of the DEMETER project and also explain the project’s acronym.

Page 5, Table 1: Please add mean temperature and mean streamflow if easily available. As mean precipitation is indicated, averages of temperature and streamflow might gain insight into the climate characteristics.

Page 5, line 13-14: Does this mean that short-term forecasts are extended by the ensembles generated using ESP? Please consider rephrasing.

Page 6, lines 27-28: Do the Nash-Sutcliffe values are computed using daily time series?

Page 7, line 18: Forecasts are also computed using one day time steps?

Page 8, line 6: Please indicate why the number of forecasts amounts to 420. 35 years x 3 months x 4 seasons?

Page 8, line 23: Is the term ”confidence interval” really correct in this context? As far as I understand, say we consider a reliable forecast of a specific event, a probability of 95% should at best also refer to 95 out of 100 occasions in the observed dataset.

Please also define the terms “nominal” and “effective”.

Page 10, lines 13-14: This phrase is hard to understand. Please consider rephrasing.

Page 12, line 13: “evolution” might be more appropriate that “maturation” in this context.

Page 13, line 1: corr-DSP forecasts

Page 14, line 1: Here, you state that Fig. 6 presents a reliability diagram while in the figure’s caption it is labeled as PIT diagram. This is a little bit misleading and might cause confusions even if the type of information is similar to a certain degree. Please confirm or specify the type of figure more detailed.

Page 14, line 11: Here, you state that Fig. 7 presents a PIT histogram while in the figure’s caption it is labeled as rank histogram. Is this in line with your explanations in Sect. 4?

Page 14, lines 11-12: Further explanations might improve comprehensibility (e.g., by stating that an equal distribution indicate accurate ensemble forecasts).

Page 14, Figure 6: In my opinion, labeling each row of the diagram by stating the watershed’s numbers might be more intuitive (see, e.g., Figure 2). This is also relevant in the case of Figures 7, 8, 9, and 10.

Page 15, line 16: The bias correction is applied for each month. Single events at time scales smaller than one month might by be subject to biases different to the monthly values.

Page 16, line 11: By the way, the term dispersion is often used throughout the manuscript if the variability is overestimated (or underestimated). Variability might be more appropriate as mentioned in line 6 on page 15.

Page 17, line 11: Do you mean corr-DSP when discussing the results of ensemble meteorological forecasts?

Page 17, line 12: Is it possible to prove if the skill is significant or not significant from your analyses? The term significant should be proved by providing statistical measures.

Page 18, line 5-6: Please explain in brief why corr-DSP is less reliable. Is this finding relevant for winter or all seasons? Maybe you can refer to the reliability diagram?

Technical corrections

Page 1, line 6: Please add the CRPS (abbreviation) here as it is mentioned later without explanation (cf. line 14)

Page 2, line 32: I would suggest using the singular form of precipitation

Page 8, line 20: distributions (plural)

Page 10, Figure 2: Please add the dimension of the precipitation bias in the color bar.

Page 15, line 15: remain

Page 17, line 11: “is predictable” instead of “are predictable”

Page 18, line 6: originates

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-387, 2017.