Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-389-RC1, 2017

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Long-term ensemble forecast of snowmelt inflow into the Cheboksary reservoir under the differently constructed weather scenarios" *by* Alexander Gelfan et al.

Anonymous Referee #1

Received and published: 14 August 2017

The research presented in this manuscript compares two methods for producing long-term hydrological forecasts: Extendred Streamflow Prediction and forecasts in which the metheorological scenarios are based on a weather generator. In both cases, the hydrological model is ECOMAG, a semi-distributed process-based model. The methods are applied on an interesting case study, which is a reservoir that is fed by a very large snow dominated catchment. Unfortunately, in my opinion the manuscript would require many improvements regarding both content and presentation before being published in HESS. I would personally advise that it be rejected in its present form, even if I think that the general objective of implementing a long-term hydrological forecasting system for the Cheboksary reservoir is relevant and interesting.

Printer-friendly version

Discussion paper

In the following, I will explain my opinion in greater detail. I really appreciate the effort of the authors and in no way I wish to discourage them. On the contrary, I sincerely hope that this discussion can be useful for them.

MAJOR COMMENTS

1. The original contribution is not clear/significant enough

Extended streamflow prediction (ESP) was proposed by Day in 1985 and is very well-known. In consequence, this is not an original element of the research presented herein. The specific weather generator used in the research was previously developed in another study, so that also does not represent a strong original content. The same consideration applies to the hydrological model ECOMAG. While the authors clearly state the problematic in their introduction (finding long-term predictability for snow-dominated catchments, in particular large multipurpose reservoirs), their contribution is not highlighted.

The three types of long-term forecasting systems currently operational on the VKRC system, described on page 2 lines 8-17, are never compared to the ESP and WG-based forecasts proposed by the authors. In my opinion, this is very unfortunate. As mentioned above, ESP is not a new technique. However, it would be interesting to compare it to a variety of other techniques, more or less sophisticated. I also think it would help the reader to appreciate your contribution. In fact, many of the results presented in the manuscript do not appear to me as a clear improvement over climatology. For instance, in Table 5, the skill scores obtained by the WG-based forecasts are all below 0.5. While I do agree that this represents an improvement over climatology, it is not a large one. I don't have any problem with this (a slight improvement over climatology), but it would probably be much more convincing to see (also) the improvement relative to at least one of the current operational methods mentioned on page 3. Figures 13 and 14 also support my comment: the forecasts presented on Figure 14 appear only slightly different from the climatology (please also see comment 2 about climatology) presented on Figure 13. This is especially true for the forecasts issued on March 1st (Figure 14 left) compared to Figure 13.

To sum up my first major comment: I think the authors should clarify and emphasize their original contribution, which is not clear for me at the moment. One way to do so, besides phrasing it more clearly, is to include a comparison with at least one of the 3 operational forecasting systems described on page 2.

2. Some methodological/conceptual elements need clarification

- Page 3 line 30: I disagree with the formulation: "(. . .) incorporating a stochastic weather generator (WG) that will allow for reproduction of a hydrological system response to a large variety of possible weather conditions (. . .)". I think you might want to say that "(. . .) incorporating a stochastic weather generator (WG) that will allow for a large variety of possible weather conditions that can then be provided to the hydrological model (. . .)". In fact, the WG itself doesn't reproduce the hydrological system's response. It produces meteorological scenarios. Then, hopefully, if the hydrological model is truly a faithful and complete representation of the hydrological system (which in many cases could be discussed), passing those weather scenarios into it will indeed provide a simulation of the basin's hydrological response.

- Page 6, line 11: Is ECOMAG really taking daily precipitation intensities as inputs? As in mm/hour? All the models I know rather use total daily precipitation. Although it is true that mm/day can be seen as an intensity (since it is a quantity over time), it seems a bit unusual to me.

- Page 8, line 29: How many months is "several"? Is it at least one full year?

- Page 9, Figure 4: On which basis did you chose to generate 1000 members from the WG while there are 50 in the ESP system? My guess is that you wanted to be sure to include a very wide variety of scenarios. While I agree that it could be interesting, (1) most operational agencies that issue meteorological forecasts from dynamic models produce 50 ensemble members or less and (2) the number of members influences metrics such as the Continuous Ranked Probability Score (or similarly the Ranked Probability Score). See for instance Ferro et al. (2008) for a demonstration. I suggest either setting the WG to issue the same number of ensemble members as EPS or at least justifying the choice of 1000 members and discussing the impact of ensemble size on performance assessment metrics. Page 14 line 17-18: According to Murphy (1973), "Hedging is said to occur whenever a forecaster's forecast r does not correspond to his judgement p (...) ". I don't understand how you associate your results to hedging. Hedging, by definition, arise from human intervention. Since your research does not involve human forecasters, I don't think hedging is the appropriate term here. Perhaps you want to refer to a systematic over forecasting bias in the forecasting system? In my opinion, this overforecasting is to be expected if the historical database includes many years with "higher than usual" precipitations. EPS (and WG) are very much dependent on the sample of data you have.

-Page 15 line 1: What do you mean by "forecast by chance"? Please define.

-Page 15 line 10: "(. . .) comparing forecasts to climatology." I suspect you mean "streamflow climatology"? If so, this should be explicitly mentioned in the text everywhere applicable.

-Page 18 line 12: When you write "confidence bands", do you mean "confidence intervals"? If so, please provide the level of confidence and if not, please define what you mean by "confidence bands".

-In section 4.4: are the results for ESP or WG? Globally, the explanations in this section (page 20) are difficult to follow. In my opinion, a schematic representation of the methodology would be helpful. And since this portion is, I think, more methodological, it should be moved to section 3.3

References:

Ferro C.A.T., Richardson D.S. and Weigel A.P. (2008) On the effect of ensemble size on the discrete and continuous ranked probability scores, Meteorological Applications, 15: 19-24.

Murphy A.H. (1973) Hedging and skill scores for probability forecasts, Journal of Applied Meteorology, 12, 215-223.

3. The analysis and discussion of the results is too shallow

Section 4 of the manuscript is labelled "Results and discussion". I was therefore expecting results to be discussed (rather than simply presented) in this section. However, I find this is not the case for many figures and tables. Specifically: Table 2 (what do does values mean?), Table 3 (the "first" Table 3 on page 13) and Figure 8. I think than Table 5 could also be discussed more, since, as I mention above, the improvement over climatology is still modest. However, without any other basis of comparison (such as the current operational forecasting system), it is hard to put the results in perspective. This is also related to Figures 13 and 14, which are just presented but not discussed. Those figures show that the improvement over climatology is very modest and hence, it is difficult to appreciate the authors' contribution. Similarly, Figure 15 should be analyzed more deeply (i.e. the explanations behind the results, not simply describing the figure). In the conclusion (page 24 line 19-20), the authors implicitly mention a comparison with "the deterministic forecasts of inflow into reservoir that are used in common practice in Russia (. . .)", but this comparison is not explicitly shown in the manuscript.

Another thing that struck me is that the authors are not discussing the performance of their systems in terms of the relative importance of resolution and reliability. For instance on page 15 line 2, it is mentioned that the forecasts are "capable of detecting the occurrence of rare extreme events (. . .)" This is an indication that points toward forecast reliability, but what about resolution? If the forecasts are very widely dispersed, they will likely include any events but with very low power of discrimination. This should be studied and could help to improve the discussion.

4. There are numerous spelling, orthographic and typographic errors throughout the manuscript

I will provide some specific examples in my specific comments. However, there are many such errors and I don't think it is my duty to list them exhaustively. English is not my native language either and I perfectly understand how challenging it can be to write a scientific paper in another language. From experience, I would recommend the authors to have their manuscript thoroughly proof-read by a native English speaker before resubmitting. There are also many typographic errors that should be corrected, mostly problems with parenthesis in references.

SPECIFIC COMMENTS

1. All figures except the first one and Figure 8 need reworking:

- Figure 2: the resolution of the right hand side figure is very poor. All the small grey pixels should be removed.

- Figure 3: I don't think this figure brings much information to the manuscript. It has no legend, and I think most readers are familiar with the requirements of a distributed, process oriented model. I suggest removing this figure.

- Figure 4: The three small figures in the center of each middle box (representing plots of time series) are much too small and of poor resolution. I suggest either modifying them to make them readable, or removing them.

- Figure 5: The x-axis is labeled 'years' while the text says "daily inflow discharge". The label of the axis should reflect what is plotted on the figure. Labeling it "years" means that you would plot yearly values, not daily.

- Figure 6: The legend is missing and the y-axis for 3 of the 4 panels need to be completed ("Simulated inflow volume, kmˆ3" rather than just "Simulated")

- Figure 7: Why is the Taylor diagram elliptical? Should it not be more spherical (a portion of a circle)?

- Figure 9: The axes should be labeled (titles) ! Since all panels will all likely have the same axis titles, I would suggest writing axis titles only once for each: the x axis at the bottom of the figure, centered and the y axis completely on the left, also centered.

- Figures 10-11: The text on the figures (labels, ticks, etc.) is so small, it is absolutely impossible to read anything. It should be made readable, both by increasing character sizes and figure resolution. In addition, the labeling "1", "2", etc under each panel is quite unusual. I advice labeling sub-figures (a), (b), . . . above each panel, as it is usually done.

- Figure 12: It is also difficult to read, although not as much as figures 10-11. The resolution of the figure could be substantially improved. Again, the labeling of the panels should be placed above, not below, each panel.

- Figures 13-14: Same thing: difficult to read. The legend is missing for Figure 14. Figure 15: Labels for panels (a, b, . . .) are again misplaced. The axes ticks are very difficult to read (size and resolution).

3. Table 3 on page 13 (the "first" Table 3): the units are all missing in the first column (W, QMax, Nq and Nmax).

4. There are two tables labeled "Table 3"

5. The Taylor diagrams should be explained briefly in the methodology. At the moment, all other performance assessment tools are at least mentioned in the methodology except this one.

6. English errors and typos

(This is not an exhaustive list)

- Page 3: line 9 instead of "in (Gelfan and Motovilov 2009)", it should be "in Gelfan and Motovilov (2009)". Similarly, at line 20, remove parenthesis around 2017 in "Arnal et al. (2017)". There are many similar errors with parenthesis around references in the manuscript.

- Page 2 line 31: Change "(. . .) allows forecaster to provide user (. . .)" either to "(. . .) allows the forecaster to provide the user (. . .)" or to "(. . .) allows forecasters to provide users (. . .)"

- Page 2 line 33-34: Change "Recent studies illustrating ability of the ensemble (. . .)" to "Recent studies illustrating the ability of the ensemble(. . .)"

- Page 3, line 11: remove the "the" from "Water Problems Institute of the Russian Academy of the Sciences".

- Page 3, line 23: Change "(. . .) but for possible weather condition (. . .)" to "(. . .) but also for possible weather conditions (. . .)"

- Page 4, line 15: Replace "Also, analysis of (. . .)" by "Also, an analyse of (. . .)".

- Page 8, line 22: Replace "(. . .) leads to increase of the model robustness. List of the (. . .)" by "(. . .) leads to an increase of the model's robustness. A list of the (. . .)"

- W, Max, Nq and Nmax are sometimes in italics, sometimes not. Sometimes, the "max" in "Nmax"is in subscript ans sometimes not. Sometimes with a capital "M" and sometimes not. This needs to be uniformed according to the HESS's guidelines.

- Page 10 line 11: remove "into" in the sentence "(. . .) in which the observation fell into (. . .)"

- Page 12 line 1: Replace "Magnitude of the used metrics and error estimation has led to an assumption that the model is suitable to act as a core component of (. . .)" by "The magnitude of the performance assessment metrics and error estimations lead to the conclusion that the model is suitable as a core component of (. . .)"

- Page 12 line 12: Replace "(. . .) tested through its ability" by "(. . .) tested through their ability (. . .)"

-Page 13 line 8: Replace "For the verification purposes (. . .)" by "For the purpose of verification (. . .)"

- Page 13 line 8 and several other instances: replace "ensemble of the forecasted" by "ensemble forecasts"

- Page 14 line 3: I absolutely don't understand what you mean with this sentence "(. . .) show RMSE values around 0.55-0.65 fraction of the standard deviation of the corresponding observed characteristics." Please rephrase. In particular, which "observed characteristics" are you referring to?

- Page 15 in general (for instance line 16) and in many other places: "forecasts" should be plural. Unless you examine one single forecast for a particular time step.

- Page 17 line 10: What do you mean by "efficiency"? To me the RPSS is a measure of forecasts overall quality.

-Page 17: The mathematical variables in the text, such as Fm(j) and Om(j) should be in italics and with subscripts, as in equation (2).

- Page 20 line 25: Please correct the reference for "Report. . ., 2002"

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-389, 2017.