

Relative effects of statistical preprocessing and postprocessing on a regional hydrological ensemble prediction system

Sanjib Sharma, Ridwan Siddique, Seann Reed, Peter Ahnert, Pablo Mendoza, Alfonso Mejia

Response to the reviewers' comments

We thank you very much for your thorough review of manuscript hess-2017-514. Below we provide a point-by-point response to each of the comments. The reviewers' comments are shown in blue font and our response follows immediately after that.

Comment from Reviewer #2: 1) I was intrigued by reading the statement “postprocessing alone performs similar, in terms of the relative mean error, skill, and reliability, to the more involved scenario that includes both preprocessing and postprocessing” in the Abstract (page 1, lines 24-25). This is one of the major conclusions of the work. However, further reading reveals that the results do not fully support this conclusive statement, for the following reasons:

- i) Figures 5 and 6 show appreciable performance gains of S6 over S5 for 5 cases out of 8. One can see that S6 outperforms S5 in terms of forecast lead times by 12 hours to 3 days.
- ii) The closeness of the results for the other cases (i.e., (e) and (f)) between S5 and S6 can be explained by the closeness of the raw GEFS and preprocessed GEFS precipitation, as shown in Figure 3.
- iii) The verification appears to be only conducted for large observed events without considering large forecast events, which can generate false-alarms. In short, I find this conclusion is inaccurate and can be misleading.

Response to reviewer #2: We agree with the reviewer. We have now modified the revised manuscript to indicate that the scenario involving both preprocessing and postprocessing, i.e. S6, consistently outperforms the other scenarios. However, we also indicate that in some cases the differences between S5 (only postprocessing) and S6 are not as significant. We believe, as the reviewer suggested, that this statement and conclusion is more consistent with the overall results that are presented in the manuscript.

In regards to the reviewer's point iii), we have now revised Figures 3-7 in the new version of the manuscript by computing all the verification metrics over the entire verification period (please also see our response to reviewer 1 regarding this issue). The revised figures show more clearly that S6 is consistently better than the other scenario. Qualitatively, the revised and original figures are overall similar. But some difference do emerge, as indicated in our response to reviewer 1, particularly between the warm and cool season. We have now revised the manuscript to note and discuss these differences.

2) How are the GEFS precipitation and temperature downscaled to force the HR-RDHM? A description should be provided.

Response to reviewer #2: As suggested by the reviewer, we added the following text in the revised manuscript: “The GEFSRv2 data are bilinearly interpolated onto the 4 x 4 km² grid cell resolution of the HL-RDHM model.”

3) Page 1, Line 12: Do you mean “Is comprised of “by “is comprised by”?

Response to reviewer #2: This modification was incorporated into the revised manuscript.

4) Page 1, Line 28: In “The intersection of climate variability and change, increased exposure from expanding urbanization, and sea level rise are increasing”, what do you mean by “The intersection of climate variability and change”?

Response to reviewer #2: We meant by this statement that climate variability and climate change, which act together, alongside expanding urbanization and sea level rise are making flood prediction more challenging. We now revised the manuscript to say only “climate change” as we think this makes the sentence more clear and easier to read.

5) Page 2, Line 6: In “for research purposes, meet specific regional needs, and/or real-time forecasting applications”, do you mean “to meet . . .”?

Response to reviewer #2: We have corrected the text following the reviewer’s suggestion.

6) Page 3, line 21: Shouldn’t it be U.S. Middle Atlantic region?

Response to reviewer #2: We incorporated this modification into the revised manuscript.

7) Page 5, line 16: “Also, HCLR has been shown to outperform other widely used preprocessors (Yang et al., 2017)”. Should be more specific here since the paper only compares the HCLR and BMA.

Response to reviewer #2: Following the reviewer’s comment, we made our statement more specific; it now reads as follows: “Also, HCLR has been shown to outperform other widely used preprocessors, such as Bayesian Model Averaging”.

8) Page 6, line 31: “6-houlry” is a typo.

Response to reviewer #2: Thanks for catching this. We incorporated this modification into the revised manuscript.

9) Page 7, line 23: “QR has similar skill performance in streamflow and normal space”. This sentence is not clear to me. Do you mean that QR has similar skill performance in the streamflow space as well as normal space?

Response to reviewer #2: We rephrased this sentence to incorporate the reviewer’s comment. The revised sentence reads as follows: “QR is applied here in streamflow space, since it has been shown that, in hydrological forecasting applications, QR has similar skill performance in streamflow space or normal space”.

10) Page 8, line 15: How many events result from this threshold? Is the sampled climatological probability distribution derived from the observed data? If so, will your conclusions still hold if events corresponding to forecasts with large magnitudes and high probabilities also included in the verification?

Response to reviewer #2: The reviewer makes a good point. As mentioned in one of our previous response, we have now modified the manuscript by computing the metrics (Figs. 3-7) over the entire verification period. Overall, our conclusions did not change based on the revised figures. As noted before, we do see now some seasonal differences (mainly, the performance of scenarios S4-S6 is more similar to each other in the warm season than it was before in the original manuscript) and the ability of S6 to outperform the other scenarios is more clear now.

11) Page 10, line 33: “QR displays better reliability than ARX (1,1) across lead times, basins, and seasons”. By what measure(s)?

Response to reviewer #2: This sentence was slightly modified in the revised manuscript to add clarity and address the reviewer’s comment. The new sentence reads: “We also computed reliability diagrams for the two postprocessors (plots not shown) and found that QR displays better reliability than ARX(1,1) across lead times, basins, and seasons.” The figures are not shown simply to keep the length of the manuscript and number of figures manageable.

12) Page 11, line 36: “reinforcing the fact that preprocessing may have little effect on the flood forecasts”. See the General Comments.

Response to reviewer #2: We agree with the reviewer on this comment and modified this sentence to read as follows: “However, the gain in skill between S3 and S4 is generally small, particularly at the short lead times, reinforcing the fact that preprocessing alone may have little effect on the flood forecasts.”