Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts

Louise Crochemore, Maria-Helena Ramos, and Florian Pappenberger

The authors are presenting a generally well-conceived study focussed on possible improvement of seasonal streamflow forecasts by applying bias correction to the forcing precipitation forecasts. I found the manuscript well written, with appealing and adequate artwork and well discussed results. I found that the whole manuscript well supports the results the authors are enumerating. The scores the chose are adequate and the combination of scores allows drawing conclusions that are not blended by the scores that show maximal improvement with respect to the defined references.

It is not clear which PET forcing is used in forecasts mode. If the authors use the observation-based reanalysis of PET in combination with the precipitation forecasts, then the appeal of this study would be quite reduced, in my "operational-minded" opinion.

Furthermore, the discussion and conclusions section is not adequately considering previous studies.

Issues to be addressed (Page(s) – Line(s)):

General comments:

While the Introduction is well balanced and gives useful insight on previous work on the topic and also references supporting the envisaged methodology, I found that the final paragraphs should possibly include more information on the novelty of the present manuscript. Also in the methodological section some more referencing is needed. See minor comments for this.

4 – 4-15: We learn here about the meteorological forcing. It is clear to me how you use precipitation, but as a forecast and as SAFRAN product. Concerning Potential Evapotranspiration (PET), only SAFRAN is declared. I'd like you to declare which PET is used in retrospective forecasts forced by the ECMWF products. If it is from ECMWF, you should state why you are not post-processing it. If you use SAFRAN, you should be able to assess how much uncertainty are you neglecting by using the best observed estimates of PET instead of using a forecasted value (which you need to do as soon as you will deploy the system in real-time). In our experience, for basins not affected by snow-melt, the post-processing of relative-humidity data (an important proxy the evaporation demand by the atmosphere) helps improving the estimation of hydrological droughts (Jörg-Hess et al., 2015).

4-25: I just reviewed another paper on seasonal forecasting where authors did not show any score concerning their calibration/validation and I amended it. Same here. I am happy with a table as supplementary material.

24 – 18: I like this evaluation very much, just, I miss some quantification supporting the description based on visual inspection you are giving. Be pragmatic.

25 -3: The discussion section is here quickly merged with the conclusions. The only link to current literature one is expecting here merely consists in a enumeration of possible post-processing of the forecasts with currently available methods. Here some more effort has to be shown to make also this section a valuable part of the manuscript.

26 – 2: You address here the issue of implementation in operational systems. Again, declare how you deal the PET, and then re-evaluate the potential for real-time operations.

Minor comments (Page(s) – Line(s)):

2 – 11: I guess here you should give one or two references for the statistical models, too. Eg. Some approaches relating winter snowpack to summer-flows (e.g.: Godsey et al., 2014; Jenicek et al., 2016).

5 – 4: Please support the “one-year-leave-out cross-validation method” with a reference.

6 -2: Please support “Precipitation and streamflow forecasts are evaluated with deterministic and probabilistic scores commonly used in ensemble forecasting” with a reference, e.g. Brown et al EVS paper.

8-15: Nice idea to use the ensemble of past-streamflow observations as a reference. If you would “sort-out” some past years by means on analogues techniques you might get a very challenging set of members for your ensemble forecast. Have you tried this?

8-22: Another interesting feature here. This definition of gain is very elucidative. Can you maybe elaborate on pro and contra of this kind of “gain” definition with respect to scores based on cost-loss considerations?. Why choosing such a large gap of day between the classes? Have you tried to make a 30-day moving window? Or a 15-days moving window?

9 – 2 & 9 -19: Both in Figure 2 & 4 CRPSS is showing increasing skill at weeks 5 and 9. We are also used to “struggle” in interpreting such cycles. Do you have some ideas on your particular case here?

11 – Figure 4: How would look like this figure if you use the “ensemble based on past streamflow” as a reference?

13 – Figure 6: Right margin is cropped. Additionally, the “too wet”=red is not really intuitive.

13 – 2: “the 2-month” or the “month-2” ? If you mean the one for the second month of the forecast I would find more adequate to use “month-2”.

17 – Figure 8 (and later 9): I like such Figures because the train my brain. Tell me if I am reading it wrong:

> If a look at a certain score in a certain season than for a particular bias correction method a percentage of the basins is showing improvement in lead time. Of this percentage a distinct portion shows improvement of let’s say 60 to 90 days.
>
> So largest improvement is in the PIT-Skill in summer and Winter for the EDMD methods.

Right?

22 – 8: This would be the only heading with a question mark. Maybe replace this with a sentence

23 – Figure 15: is there any special reason (beside readability) for having different scales in the three panels?

Final considerations:

I find this manuscript is a very solid communication for the growing community dealing with seasonal forecasting in hydrology. It uses a strong set of data and robust statistics and comes to valuable conclusions. I indicated some weakness that let me recommend to the editors to ask for moderate revisions for this manuscript.

Best regards

Massimiliano Zappa

Birmensdorf, 23. March 2016

References:

Brown, James D., et al. "The Ensemble Verification System (EVS): A software tool for verifying ensemble forecasts of hydrometeorological and hydrologic variables at discrete locations." Environmental Modelling & Software 25.7 (2010): 854-872.

Godsey, S. E., Kirchner, J. W., and Tague, C. L.: Effects of changes in winter snowpacks on summer low flows: case studies in the Sierra Nevada, California, USA, Hydrol. Process., 28, 5048–5064, doi:10.1002/hyp.9943, 2014.

Jenicek, M., Seibert, J., Zappa, M., Staudinger, M., and Jonas, T.: Importance of maximum snow accumulation for summer low flows in humid catchments, Hydrol. Earth Syst. Sci., 20, 859-874, doi:10.5194/hess-20-859-2016, 2016.

Joerg-Hess S, Kempf SB, Fundel F, Zappa M. 2015. The benefit of climatological and calibrated reforecast data for simulating hydrological droughts in Switzerland. Met. Apps. 22: 444–458. doi:10.1002/met.1474