Interactive comment on "State Updating and Calibration Period Selection to Improve Dynamic Monthly Streamflow Forecasts for a Wetland Management Application" *by* Matthew S. Gibbs et al.

Anonymous Referee #3

This manuscript presents an original study focused on two important aspect of monthly streamflow forecasting: state updating and the selection of an appropriate calibration period. The watershed on which the methods are implemented and tested is a very interesting case study. This semi-arid watershed is extensively impacted by human intervention, as it includes diversions and "Drain M", so runoff water can be directed either toward the north to benefit wetlands or toward the south to maintain fish ecosystems and comply with other constraints. Adequate management of this watershed appears very difficult and it is evident from the authors' description of the situation that monthly hydrological forecasts are essential. The authors show that selecting a calibration period during which the conditions are close to the expected forecast conditions can improve forecasts performance substantially.

In my opinion, this is a well-written paper (very clear!) that brings interesting novel knowledge in the field of ensemble monthly streamflow forecasting. I also find it completely appropriate for publication in HESS, especially since the case study raises issues regarding water management and tradeoffs between ecosystem services and human activities.

I appreciate that the authors included a short discussion about the uncertainty related to the gauging measurements (page 8 first paragraph). I also like the idea of adding the "split" parameter to GR4J in the calibration process.

The conclusions drawn by the authors regarding the tradeoff between the length of the calibration database versus its "representativeness" is interesting. I appreciate that they recognize that said conclusions might be limited to their specific case study and that further investigation for a wider range of hydro-climatic regimes would be needed. I only have very few minor comments and suggestions that I think could improve the paper. I strongly recommend that it be published in HESS.

Thank you for the constructive comments that will improve the manuscript.

Minor comments:

1. Why did you consider only the median of the hydrological model predictions rather than the whole ensemble?

This question kept bugging me all along while I was reading. You have access to meteorological ensemble forecasts (page 6 line 30 to page 7 line 5). They are expected to account for the uncertainty related to the meteorological conditions (or at least a part of this uncertainty). Why then did you not keep all the scenarios after passing everything through the hydrological model? Is it to make it more comparable to the case where the hydrological model is forced by (deterministic) observations? If so, I personally don't see why this would be necessary. And then after, you use a statistical method to dress the median back to an ensemble. Is it because you found out that hydrological ensemble forecasts built only from meteorological ensembles were underdispersed and would have needed post-processing?

In my opinion, those choices (i.e. using the median, thus ignoring the other ensemble members, and then dressing the median into an ensemble) really need further explanations/justifications.

The focus of this work was to investigate the impact of the state updating and calibration periods on streamflow predictive uncertainty, and a current best practice post-processing approach was adopted to estimate this predictive uncertainty. It is common practice to use deterministic inputs, often based on summarising an ensemble of forecasts using summary metrics (e.g. mean or median) as the input to post-processing methods (e.g. Lerat et al., 2015; Matte et al., 2017; Schepen et al., 2017; Wani et al., 2017). As such, is was considered beyond the scope of this work to also consider improving methods for applying post-processing models.

However, it is acknowledged that that by only considering the median streamflow simulated across the forecast rainfall ensemble that some information is lost. The paper will be modified to include the references above to outline that the current practice for applying post-processor error models has been adopted in Section 3.5, and highlight that further work is needed to more fully utilise the information from ensemble forecasts when developing post processing models.

2. There are a couple of (very minor) elements that could be clearer

- Page 5 line 30: please add a reference for the Ramsar list. I didn't know this list before reading the manuscript so I looked it up on the web. I think that a reference would be helpful to be sure that other readers like me know what you are talking about.

The reference Matthews (1993) will be added.

- Page 9 line 21: What is a "burn-in" period? At first I thought it was a synonym of "warm-up period" in the context of the DREAM algorithm, but I am really not sure.

Effectively, the two approaches are the same. The term "warm-up" typically used in hydrological modelling, and "burn-in" typically used in Markov Chain Monte Carlo modelling. The commonly used terminology is proposed to be maintained, with a qualifier added in the DREAM description for clarification:

A maximum of 25,000 evaluations were used, including a burn-in period *(similar to the warm-up period used in hydrological modelling)* where the initial samples were discarded…

3. There are a few typos in the manuscript and I am unsure of the spelling for 1-2 words:

- Page 3 line 27-28: parenthesis typo, replace "(McInerney et al, 2017)" by "McInerney et al. (2017)"

The typo will be corrected.

- Page 11 equation (5): Something seems wrong with the curly brace

The curly brace is commonly used notation for conditional expressions, in this case to avoid dividing by 0 when λ=0. No change is proposed.

- Page 11 line 1: I think that the sentence "(…) are available, for example, the ensemble Kalman filter (…)" should be split, as in: "(…) are available. For example, the ensemble Kalman filter (…)"

It is agreed that the suggested change is clearer and will be adopted.

- Page 11 line 12: day "n" and month "t" should be in italics.

The change will be adopted.

- Page 12 line 17: I think that "straightforward" should be written in one word.

The change will be adopted.

- Page 13 equation 12: A summation seems to be missing at the denominator. Also, I don't think there should be a "t" index for the average streamflow, since by definition it is independent from time (it is the average of the time series).

The summation typo will be corrected. It is agreed that the index t for the average is also incorrect. Thank you for picking these up.

- Page 14 line 6 (and several other places in the manuscript): Why do you write "observ***ed*** rainfall" but not "forecast***ed*** rainfall" I am not a native English speaker so perhaps I am completely wrong, but I could not help but finding this strange.

This is a strange nuance of the English language. Both "forecast" and "forecasted" can be used as past tense for the verb forecast. It was considered that "forecast" is more commonly used in the literature and is proposed to be maintained throughout.

- Page 14 line 16: Is there a "to" missing in "The changes due adopting (…)"?

The typo will be corrected.

References

Lerat, J., Pickett-Heaps, C., Shin, D., Zhou, S., Feikema, P., Khan, U., Laugesen, R., Tuteja, N., Kuczera, G. T., M, and Kavetski, D.: Dynamic streamflow forecasts within an uncertainty framework for 100 catchments in Australia, 36th Hydrology and Water Resources Symposium: The art and science of water, Barton, ACT, 1396-1403, 2015.

Matte, S., Boucher, M. A., Boucher, V., and Fortier Filion, T. C.: Moving beyond the cost–loss ratio: economic assessment of streamflow forecasts for a risk-averse decision maker, Hydrol. Earth Syst. Sci., 21, 2967-2986, 2017.

Matthews, G. V. T.: The Ramsar Convention on Wetlands: its History and Development, Ramsar Convention Bureau, Gland, Switzerland, 1993.

Schepen, A., Zhao, T., Wang, Q. J., and Robertson, D. E.: A new method for post-processing daily sub-seasonal to seasonal rainfall forecasts from GCMs and evaluation for 12 Australian catchments, Hydrol. Earth Syst. Sci. Discuss., 2017, 1-27, 2017.

Wani, O., Beckers, J. V. L., Weerts, A. H., and Solomatine, D. P.: Residual uncertainty estimation using instance-based learning with applications to hydrologic forecasting, Hydrol. Earth Syst. Sci., 21, 4021-4036, 2017.