Interactive comment on "Evaluating residual error approaches for post-processing monthly and seasonal streamflow forecasts" *by* Fitsum Woldemeskel et al.

F. Woldemeskel

fitsum.woldemeskel@bom.gov.au

Received and published: 16 June 2018

Response to Referee #2 General comment: This paper presents a comparison of three variants of a post-processing approach for long-range (here monthly to seasonal) streamflow forecasts in Australia. The paper is well-written and easy to read. The research is interesting for several reasons. First, the topic of long-range forecasting and especially how the skill of such forecasts can be improved is currently raising a lot of attention from hydrologists. There are many practical applications for which seasonal forecasts are required for decision-making. Second, Australia is a vast country which includes a broad range of hydro-climatic conditions, and the authors efficiently gath-

ered a data base of 300 catchments, to ensure that their findings are as generalizable as possible. The authors explicitly address the specific case of dry catchments and low-flow periods, which is of great practical interest in several areas on the planet. Indirectly, the research has implications for socio-economic issues, as the management of water-scarce catchments could benefit from better seasonal forecasts. The paper also fits well within the scope of HESS.

This paper is definitely suitable for publication in HESS. However, I do have a few specific comments and suggestions for the authors to improve it prior to publication. In my opinion, all comments are minor.

Author response: We thank the reviewer for this very positive assessment of our manuscript and for recognising its practical importance. We are pleased that reviewer found our manuscript suitable for publication in HESS. We provide response to the review comments as follows.

Specific comments/suggestions/questions:

Referee comment 1: Data assimilation/link with short-term forecasts

From the paper, I am not entirely sure how GR4J's state variables are managed (see also my next comment). From what I understand, there is no data assimilation at all. Perhaps this can be justified in the context of long-range forecasts as the effect of data assimilation would fade out quickly (probably before the one month horizon).

I was also wondering if there is a link (operationaly at BoM) between short- and longrange forecasts. Surely the hydrological model is the same, but what about the meteorological forecasts? Are the short- and long-ranges connected in some way? Surely, in operations, there must be a certain form of data assimilation for short lead times.

Considering the above interrogations, I would very much appreciate short comment regarding data assimilation in the paper.

Author response: The reviewer is correct that we have not used data assimilation to up-

date GR4J state variables. However, data assimilation of ocean observations has been incorporated in the climate model (POAMA2.0) from which the rainfall forecasts have been obtained. We agree with the reviewer that the benefit of data assimilation for seasonal forecasts is limited. However, Gibbs et al. (2018) showed that monthly streamflow forecasting could benefit from state updating when the issue of non-stationarity is also handled. This is something to be investigated further in the future. We will provide a short comment about this issue in the manuscript.

Regarding the link between short- and long range forecasts provided by the Bureau of Meteorology, the two have independent systems due to the different needs of the forecasting service stakeholders. Short-range forecasts require a daily update with a focus on timely delivery of forecasts to anticipate quickly evolving events. Long-range forecast are more connected to longer-term decision making, which requires monthly update and statistically reliable forecasts. In addition, the GCM inputs are sourced from different models: short-term streamflow forecasts use weather forecasting model with limited modelling of ocean dynamic, whereas long range streamflow forecasts use climate outlook model with strong coupling between ocean and atmospheric models. As of now, updating state variables through data assimilation is also not yet implemented in short-range forecasts, but there are plans to incorporate this in the future.

Overall, the streamflow forecast relies on data assimilation included in the climate model and of robust hydrologic modelling technique highlighted in the paper. It is also worth to mention that the Bureau of Meteorology prioritised investments in developing hydrologic modelling within a robust uncertainty framework, followed by streamflow and rainfall post-processing. In our view, the incremental benefits from data assimilation is likely to be less than these components.

Referee comment 2: Simulation and Forecast steps vs model calibration and warm-up

Section 2.2 and 2.3: I am slightly puzzled by that division into "simulation step", which includes model calibration, and "forecast step". Reading the description of the "simu-

lation step", one could thing that you re-calibrate the model several times, once before each forecasting step. Is that so? If so, why? You want the model parameters to be dynamic?

I would tend to think that what the steps would rather be (1) calibrate the model (40 times using the MCMC-based method you mention) once and for all, (2) simulate streamflow over the entire period and save the state variables (I assume no data assimilation) and (3) launch GR4J in "forecast mode", by fetching the appropriate state variables for a specific date and feeding the model with meteorological forecasts.

I would very much appreciate if you could clarify those issues in the paper. In particular, think that calibration should be separated from simulation.

Author response: The three steps the reviewer described are correct. However, there is an additional process in step 1, i.e., we estimate parameters in a moving 5 years leave-out cross-validation approach. This is done in order to validate forecasts with an observed data independent from the ones used for calibration. When we use moving 5 years leave-out cross-validation scheme, the parameters would be slightly different for each year.

We use data from 1980-2008 for cross-validation with a model warm-up period of 5 years (i.e. 1975-1979).

We will further clarify these points in the paper as well as make the distinction between calibration and simulation clearer as suggested by the reviewer.

Referee comment 3: Discussing the choice of model for dry catchments

Section 5, lines 535-536, you mention that "This finding can be attributed to the challenge of capturing key physical processes in modeling dry and ephemeral catchments (: : :)". In my opinion, this sentence leads to questioning whether of not GR4J is an appropriate model for very dry catchments. I know this model very well and I can appreciate its many qualities. GR4J works well for a very wide variety of hydro-climatic

conditions. In addition, I do understand the practicality of having only one (very simple) model for all catchments on the entire country. However, there is no soil per se in GR4J. It is a very simple conceptual model which cannot, for instance, model soil sealing phenomena for dry catchments. I don't see how this model could ever capture the physical processes, as mentioned in your sentence.

In my opinion, this issue (the choice of a very simple conceptual model) should be briefly discussed following lines 535-536.

Author response: We thank the reviewer for pointing out this issue. The reviewer is correct that the model structure of GR4J, in particular its simplifying assumptions, might be responsible for the relatively lower forecast skill in dry catchments as compared to wet. Having said that from our experience, uncertainty of rainfall forecast dwarfs the hydrologic uncertainty. Our intent with respect to hydrological modelling is to use a model that can perform as best as possible in different hydro-climate conditions without necessarily being complicated and non-parsimonious. Whilst using a single simple conceptual model is attractive for a practical operational system, there may be gains in exploring alternative model structures for difficult catchments (e.g. Clark et al., 2008; Fenicia et al., 2011). We intend to explore alternative model structures for difficult ephemeral catchments. We will elaborate on these issues on lines 535-536 as well as highlight some of these points further in the revised manuscript.

It is also worth mentioning that forecasting in dry catchments will remain an issue regardless of the hydrological model used due to the limited amount of information contained in streamflow records (high number of zero flow values) and high frequency of convective storms.

Referee comment 4: Citing papers from HESS Discussion

In my opinion, citing papers from HESS Discussion should be discouraged. After all, there is no real filtering of the papers before they can be published in discussion. The revision process takes place around the Discussion paper. To me, a paper that never

makes it to HESS (after the Discussion) should be considered as rejected, even though it remains publicly accessible on the web. You wouldn't cite a paper that was rejected from other "more traditional" journals for which the revision is not as public as for HESS. Of course you could argue that if a paper in Discussion receives excellent comments but never makes it to HESS, it could be a case where the authors purposefully decided not to spend time editing it according to the reviewer's comments and re-submitting it. In my opinion, this practice, if it exists, should not be encouraged. Again, it wouldn't be possible with the majority of other journals.

Therefore, I would very strongly recommend that you remove all references to HESS Discussion. Set (2006) should therefore not be cited.

The citation for Mendoza et al (2017) should be updated as it is now published. Same for Turner et al (2017). The titles have also changed in the published version.

Author response: We agree with the suggestion not to cite HESS Discussion papers. Therefore, we will remove or modify the above references as appropriate in the revised manuscript. We will also update the references as suggested.

Referee comment 5: Forecasts' value

Section 5.3 lines 584-587, you briefly touch on the issue of forecasts value. I personally don't think measures of skill could ever be linked to the socio-economic value of forecasts. Most studies focussing on forecast values in the current literature largely over-simplify the problem. For the issue of forecasts value to be tackles in a more realistic way, researchers from humanities and social sciences would inevitably have to be involved. Forecasts value involves complex issues related to human psychology, economic theory, communication, social studies, etc. See for instance Morss et al. (2010), Matte et al. (2017), Toon et al. (2017) and Solin et al (2018).

In my opinion, forecasts skill is a pre-requisite for forecast value but in no way a guaranty. I don't see how metrics related strictly to the skill of a forecast (as in comparing

the forecast to observation) could be a predictor of forecasts value on their own.

Morss et al (2017) Examining the use of weather forecasts in decision scenarios: results from a US survey with implications for uncertainty communication, METEOROLOGICAL APPLICATIONS, 17(2), 149-162

Matte et al (2017) Moving beyond the cost-loss ratio: economic assessment of streamflowforecasts for a risk-averse decision maker, HYDROLOGY AND EARTH SYSTEM SCIENCES, 21 (6), 2967-2986.

Toon et al (2017) Integrating Household Risk Mitigation Behavior in Flood Risk Analysis: An Agent-Based Model Approach, RISK ANALYSIS, 37 (10), 1977-1992

Solin et al (2018) Vulnerability assessment of households and its possible reflection in flood risk management: The case of the upper Myjava basin, Slovakia, INTERNATIONAL JOURNAL OF DISASTER RISK REDUCTION, 28, 640-652.

Author response: We thank the reviewer for these insights on the value of forecasts as well as for the suggestion of relevant literatures. We agree with the reviewer that forecast skill is a pre-requisite but not a guarantee of its value. From an operational point of view, having a way to link skill and value would be highly valuable. The Bureau actively woks with its stakeholders to provide evidence about this point by developing forecast application case studies (http://www.bom.gov.au/water/ssf/case_studies.shtml). A socio-economic study conducted by London Economics has also highlighted the value of seasonal forecasts, among other Bureau services (Duke et al. 2016). However, a link between skill and value is a very complex issue as mentioned by the reviewer. We will modify the text in line 584-587 to briefly elaborate on these aspects.

Referee comment 6: Typos/spelling/format/figures

- Page 10 line 255: I think the word " trial" should be replaced by "tried". - Page 13 equation 11: The CRPS is usually computed by averaging the values over a large sample of forecasts-observation groups. Therefore, I think it is important that equation

(11) be modified to be more explicit about this averaging. - Page 14 line 388: "lead to misleading" is a bit strange to read. I would advise rephrasing - Page 15 lines 413-414: there seem to be an awkward space between those two lines. Please verify. - Page 16 lines 443-444: Is "from in excess of 150%" the correct phrasing? Also, there is a typo in the parenthesis "(Figure 45i)". - Page 18 line 493: remove comma after "scheme" - Page 37, figure 8: please include the units for streamflow (y axis) on this figure. In addition, I am not entirely sure I understand the time step (x axis). Counting the points, I understand that the time step is one month, which would be coherent with the text, but not explicitly specified for this figure. In my opinion the x axis label could also be clearer. - Page 38 figure 9: An "S" is missing for the y axis label of the top row. It should be CRPSS and not CRPS.

Author response: We thank the reviewer for pointing out the above editorial corrections. We will incorporate these corrections in the revised manuscript.

References

Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gup ta, T. Wagener, and L. E. Hay.: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models,Water Resour. Res., 44, W00B02, doi:10.1029/2007WR006735, 2008.

Fenicia, F., D. Kavetski, and H. H. G. Savenije.: Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development, Water Resources Research, 47(11), W11510, 10.1029/2010wr010174, 2011.

Gibbs, M., McInerney, D., Humphrey, G., Thyer, M., Maier, H., Dandy, G., and Kavetski, D.: State updating and calibration period selection to improve dynamic monthly streamflow forecasts for an environmental flow management application. Hydrology and Earth System Sciences Discussions, 22(1), 871-887, 2018.

Duke, C., Godel, M., Koch, L., Suter, J., and Ladher, R.: A study of the economic

impact of the services provided by the Bureau of Meteorology, London Economics, United Kingdom, 2016.

Printer-friendly version

Discussion paper