

## Comment on hess-2021-589

Anonymous Referee #1

---

Referee comment on "Seamless streamflow forecasting at daily to monthly scales: MuTHRE lets you have your cake and eat it too" by David McInerney et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-589-RC1>, 2022

---

I found this manuscript to be of high quality, interesting and very well written. In particular, I find the figures very clear and of high quality. I have a few minor comments that I would like to see addressed, but this should not require a lot of time.

1) When I first started reading the manuscript, including the title, I was very confused by what you refer to when you talk about a "model". To me, and I think for most hydrologists, our first reflex is to think about the hydrological model (GR4J in your case). It later became clear that you were most of the time referring to the post-processing method, but I was still bothered throughout the paper by the fact that it could be clearer. For one thing, I would suggest changing the title to indicate that you are interested by comparing two post-processing methods. Second, I would avoid the use of the QPP acronym and instead refer to the "post-processing method" or "post-processing model". I recognize that it would be a bit longer, but in my opinion it would be clearer. I'm one of those persons who don't like acronyms too much... Section 2 should be named "Post-processing methods" or "Post-processing models", and maybe the titles of the subsections could be adjusted accordingly. Finally, I would suggest mentioning GR4J explicitly in section 2.1 and referring to section 3.2 for greater details. In fact, I was initially a bit bothered by the fact that GR4J was described in the case study and not in the "Forecasting models" section, because to me, GR4J is a model and it can be used jointly with meteorological forecasts to produce streamflow forecasts. I think I understand what you wanted to do here: you wanted to avoid drawing too much attention to GR4J because it is not the central element of your study. So in the end, I don't mind too much GR4J being described in section 3.2, but I really think it would help clarify/distinguish things if you could mention it in section 2.1 too.

2) This is linked to point 1. When I read the abstract and intro, I had a bit of difficulty understanding what you meant by "designed" for a single time step", because I had, again, a hydrological model in mind. So I was thinking that, for instance, if GR4J runs at a daily time step, there is nothing stopping you from aggregating it at the monthly time step. So I wasn't completely sure where you were going with that, until I read section 2. It then became very clear that the monthly post-processing method is completely inapplicable at the daily time step and I understood what you meant in the abstract and intro. I think it

would be better it could be clearer right from the start, and I think that making it more explicit that you refer to post-processing methods would really help.

3) Page 1 line 20: Would it be more accurate to refer to ACCESS-S as an atmospheric model? When I think of a climate model, I think about the ones used for climate change study, that really model the climate over long periods and for which the fluctuations at smaller time steps have no real significance. I know they are basically very similar in their structures, but if I understand correctly, they are not run in the same way at all

4) I was really thankful for Figure 1. Section 2 was a bit heavy, but this figure is really helpfull to understand how everything is put together. Great work. Very small point: is it possible that the accronym REM is not defined in the text? Maybe I missed it? I understand it means Residual Error Model but can you please make sure it is defined in the text when first mentionned? Just for clarity.

5) Very very small point: Line 111, why chose the subscript "foc"? Why not "fcst" if it is forecasted? I think it would be clearer?

6) Page 5 line 141: do you mean each individual member or is it really each forecast? I understand you reduce (or "collapse") the forecasts to deterministic ones, so I guess it is really "each forecast", but I'm not sure. Can you maybe clarify? I was a bit lost.

7) Page 6 line 167: I like that you use the word "collapse" when refering to the transition from ensemble to deterministic. I think I will adopt this terminology myself in the future.

8) Section 2.4 is very helpful. Very clear. Thank you.

9) Page 10 line 258: I would suggest referring to Schepen et al (2018) as a pre-processing method instead of a post-processing method, especially in the context of your study, since you use the result from this method as an input.

10) I like section 3.4.1. The choice of performance indicators is very well justified and they are clearly appropriate.

11) I was surprised that Figure 5 b was not discussed more, especially in relations to the findings in the next page (statistical significance for the different metrics. When looking at Figure 5b, I can't help but think that the monthly post-processing method is not doing a very good job, and this is a bit surprising given that it is trained for monthly forecasts.

First, there is an underestimation of streamflow between 2000-2001, and then (2006-2008) a pretty large overestimation. Why is that? Any idea? Then later we learn that, according to the different performance metrics, the differences in performance are actually not often statistically significant. I was surprised by that, because of that figure. I would really appreciate a bit more discussion/explanation of what is happening in 2000-2001 and 2006-2008.

12) Line 388: I suggest removing the word "actually" to avoid repetition with the previous sentence