Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC2
https://doi.org/10.5194/hess-2021-589-RC2, 2022

Comment on hess-2021-589

Anonymous Referee #2

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-RC2, 2022

Summary

In this paper, the authors introduce a Multi-Temporal Hydrological Residual Error (MuTHRE) model that enables the production of seamless streamflow forecasts (e.g., daily, weekly, fortnightly, monthly) within the range of 1-30 days. The approach is described and compared against a non-seamless streamflow post-processing (QPP) model implemented by the Australian Bureau of Meteorology's Dynamic Forecasting System. The comparison is performed in 11 Australian catchments in terms of several forecast attributes, and the authors conclude that the MuTHRE model is not only capable of providing good performance for daily streamflow forecasts and cumulative volumes, but also similar performance to that obtained with the non-seamless QPP model for monthly flows.

Overall, this is an interesting manuscript that contributes with encouraging results on the use of seamless streamflow forecasting frameworks. The motivation is clearly stated and the results are nicely presented. There is, nevertheless, a lot of room for improving explanations of the model formulation, streamflow forecast generation, and verification, so that any reader could reproduce the results presented here. There are other minor comments and editorial suggestions that may also help the authors to improve the quality of their manuscript.

Main comment:

1. Model description (section 2): I found this section very hard to understand. I think the manuscript would greatly benefit from re-organizing the material and improving definitions and descriptions. For example:

- It seems that the two approaches compared here follow the same general model structure (equations 1 and 2). Is that what you mean with "both QPP models"? Can you please be more explicit? Also, Q_t is described as a "probability model for streamflow" (equation 1), and then as a "residual error model" (L103, equation 2) when it is, in reality, the sum of deterministic model output and a residual error term. I wonder if you actually need equation (1) in this description.
- I think it would be better to have the information presented in L191-214 (differences between MuTHRE and monthly QPP model) right after section 2.1. The authors should consider separating Figure 1 (which is very nice) into two figures: one for model structure (which could include model equations for more direct comparisons between model structures), and another figure for model calibration and forecasting.
- The meaning of z should be included after presenting equation (2) (perhaps in line 107).
- Since X_t is also used to describe state variables in the hydrology literature (especially in data assimilation books/papers), I think u_t would be more appropriate for meteorological forcings (e.g., Liu and Gupta 2007). Additionally, in L125 you describe s_t as a time-varying scaling factor, while the same variable is used to describe hydrological model states in L100.
- L113: I presume that the raw streamflow forecasts do not account for uncertainty in hydrologic model parameters. Can you please clarify?
- L135: Is the ensemble size still N_{foc} after adding the residual term?
- L141: What do you mean by "individual raw forecast"? Each ensemble member produced with the ensemble of rainfall forecasts?
- L148: how is m* determined?
- Since the paper should be self-contained, additional information on the calibration procedures referred to in L162 and L190 should be provided (what are the calibration period, objective functions, and optimization algorithms?). A couple of sentences should suffice.
- L167: "and then collapsed to a deterministic forecast by taking the median". Is this current operational practice?
- L111, L112, L135, L168, L169, and elsewhere: is "replicate" the same as "ensemble member"?

Additional minor comments

2. L33-35: It makes more sense to me to describe common practice before referring to the need for seamless forecasts. Also, it would be worth highlighting that non-seamless forecasting efforts have been (and are being) conducted in South America (e.g., Souza Filho and Lall 2003; Mendoza et al. 2014), Europe (e.g., Ionita et al. 2008; Hidalgo-Muñoz et al. 2015), Asia (e.g., Pal et al. 2013) and everywhere else around the world, with appropriate citations.

3. L37: "This is the focus of our study". This reads out of place here. I recommend deleting this sentence or moving it toward the end of the introduction.

4. Figure 2: How many values are contained in each boxplot? One per basin? Since you have only 11 catchments, I think it would be better to show one line per basin. Further, it

would be informative for readers to have a table with the name of the station, basin-averaged elevation, area, mean annual runoff, mean annual precipitation, mean annual temperature, annual runoff ratio, and aridity index.

5. L273: Are you working with calendar years or water years? Are daily forecasts produced each day in year j with MuTHRE, or only at the beginning of each month? What is the final ensemble size of your forecasts?

6. L274-275: The problem of hydrologic memory in Australian catchments and its implications for cross-validation has been previously documented (e.g., Robertson et al. 2013; Pokhrel et al. 2013). I recommend the authors read and cite these papers here. The following blog article is also relevant: https://hepex.inrae.fr/how-good-is-my-forecasting-method-some-thoughts-on-forecast-evaluation-using-cross-validation-based-on-australian-experiences/

7. L292: Perhaps it would be better to replace the word "uncertainty" with "spread". Also, it would be informative to state that sharpness is a forecast attribute only (i.e., it does not depend on the observations).

8. L296: Since CRPS measures the difference between forecast and observation CDFs, it would be better to refer to "probability forecast errors" instead of "combined performance".

9. Figure 4: How are confidence limits generated? Do you compute the metric merging forecasts from all basins? Please clarify these points in the manuscript.

10. L368-370: You mention that reliability results are similar, although the boxplots look different. I recommend applying a statistical test to determine whether the distributions of these metrics are significantly different.

11. L370 and elsewhere: "practically significant" or "significant". Are the authors referring to a statistically significant result? If not, I suggest re-wording or deleting the word 'significant'.

12. Figure 6: I think you should say "overall monthly performance" in the caption, and perhaps remind readers here what "overall" means. Are you grouping the results of all basins? In the left panels, how many points are contained in each boxplot? In the center and right panels, how are the confidence limits computed?

13. L421: Shall we expect persistence in rainfall, given the chaotic nature of the atmosphere?

14. L426: I encourage the authors to replace the last sentence of this paragraph (which reads a lot like "propaganda") with a more quantitative statement regarding the performance of MuTHRE.

15. L467: “This was not feasible in this study”. If you cannot provide an explanation on why was not feasible, I suggest deleting this sentence.

16. L479: “High-quality forecasts”. Note that the quality depends a lot on the forecast attributes you are analyzing. I think it would be good to provide a brief discussion (maybe in section 5) about tradeoffs between the metrics included here (e.g., how your forecast system can improve reliability at the cost of losing sharpness), and what makes a forecast “good” or “high-quality”.

Suggested edits

17. L51: ‘Hydro-electric’ -> ‘hydropower’.

18. L70: ‘drop in’ -> ‘loss of’.

19. L312: ‘which’ -> ‘who’.

20. L377: ‘in 1 month (September)’ -> ‘in September’.

21. L380: delete ‘similar/better performance in all months, with practically’.

22. L382: delete ‘similar/better performance in 19 out of 22 years, with practical’.

23. L451: ‘Simplifies’ -> ‘A simplified’.

24. L473: ‘to forecasts’ -> ‘compared to forecasts’.

References

Hidalgo-Muñoz, J. M., S. R. Gámiz-Fortis, Y. Castro-Díez, D. Argüeso, and M. J. Esteban-Parra, 2015: Long-range seasonal streamflow forecasting over the Iberian Peninsula using large-scale atmospheric and oceanic information. *Water Resour. Res.*, **51**, 3543–3567, doi:10.1002/2014WR016826.

Ionita, M., G. Lohmann, and N. Rimbu, 2008: Prediction of Spring Elbe Discharge Based on Stable Teleconnections with Winter Global Temperature and Precipitation. *J. Clim.*, **21**, 6215–6226, doi:10.1175/2008JCLI2248.1. http://journals.ametsoc.org/doi/abs/10.1175/2008JCLI2248.1 (Accessed February 18, 2015).

Liu, Y., and H. V. Gupta, 2007: Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resour. Res.*, **43**, W07401, doi:10.1029/2006WR005756.

Mendoza, P. A., B. Rajagopalan, M. P. Clark, G. Cortés, and J. McPhee, 2014: A robust multimodel framework for ensemble seasonal hydroclimatic forecasts. *Water Resour. Res.*, **50**, 6030–6052, doi:10.1002/2014WR015426.

Pal, I., U. Lall, a. W. Robertson, M. a. Cane, and R. Bansal, 2013: Predictability of Western Himalayan river flow: melt seasonal inflow into Bhakra Reservoir in northern India. *Hydrol. Earth Syst. Sci.*, **17**, 2131–2146, doi:10.5194/hess-17-2131-2013. http://www.hydrol-earth-syst-sci.net/17/2131/2013/.

Pokhrel, P., Q. J. Wang, and D. E. Robertson, 2013: The value of model averaging and dynamical climate model predictions for improving statistical seasonal streamflow forecasts over Australia. *Water Resour. Res.*, **49**, 6671–6687, doi:10.1002/wrcr.20449.

Robertson, D. E., P. Pokhrel, and Q. J. Wang, 2013: Improving statistical forecasts of seasonal streamflows using hydrological model output. *Hydrol. Earth Syst. Sci.*, **17**, 579–593, doi:10.5194/hess-17-579-2013.

Souza Filho, F. A., and U. Lall, 2003: Seasonal to interannual ensemble streamflow forecasts for Ceara, Brazil: Applications of a multivariate, semiparametric algorithm.

Water Resour. Res., **39**, 1307, doi:10.1029/2002WR001373.