Reply to Trine J. Hegdahl's comments

We thank Trine Hegdahl for her comments on our manuscript, which will help to improve its overall quality. Below we give our replies (in blue) to her comments (in black) on how we intend to modify the paper to account for her suggestions and recommendations.

This is a well written paper. For a large sample, 121 catchments in France, different combinations of model structures, calibration strategies, and spatial frameworks are tested. The amount of data, results and analysis is substantial. I believe the topic is of interest to the hydrological community, and that the paper contributes to the research and practical application of modelling approaches to improve streamflow simulations.

Thank you for this positive feedback.

L245: The split-sample test was used, and calibration applied to the two separate time periods. Did you then use two different parameter sets and two different evaluation periods for each catchment?

Exactly, the procedure is as follow:
1) Calibration between 1999 and 2008 / Evaluation between 2009 and 2018
2) Calibration between 2009 and 2018 / Evaluation between 1999 and 2008

When we evaluate our simulations, we average our composite criterion obtained on these two untrained periods. We will make this point clearer in the revised manuscript.

L251: The use of streamflow transformation is set to +0.5, +0.1 and -0.5. Could you please explain what this means and why these numbers were selected? Experience, other?

Our team has investigated this issue of the impact of flow transformations on calibration results in several studies (e.g. Oudin *et al.*, 2006; Pushpalatha *et al.*, 2012). More recently, in the study by Thirel et al. (2023), we showed that it is difficult to represent a wide range of streamflow with a single flow target and that the use of streamflow transformations for hydrological models calibration leads to better represent specific ranges of the hydrograph. Here we selected three transformations which target high flows (+0.5), low flows (-0.5) and one more generalist (+0.1), according to this study. We will refer to these articles in our manuscript to support these statements.

Figure 3: I find this figure very informative and helpful. Return to it several times.

Thank you for the comment. We will add a few references to this figure in the text to help the reader.

Figure 6: Am I right that positive values are lumped better than semi-distributed? It was a bit difficult to grasp from the y-axis information. Could that be stated in the subtitle or perhaps put "median (KGE_L – KGE_{SD})" as label on the y-axis?

Positive values indicate that semi-distributed is better than lumped. Indeed, the figure can be difficult to understand. We will improve it.

Figure 10: It was difficult to see the thin grey lines. Especially for the printed version. Is it possible to make this a bit more visible.

Noted, we propose the following modification:

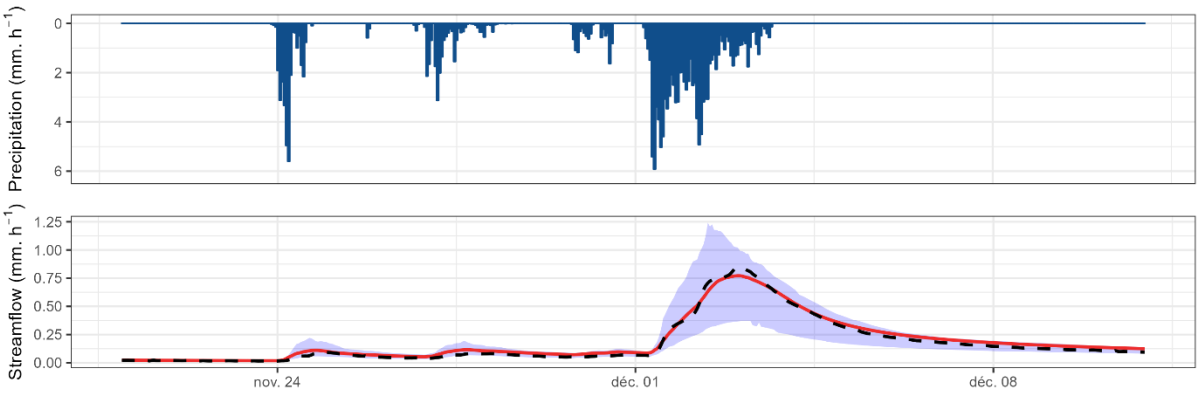

Figure 19: Include the total number of catchments in the figure text. (Why not 121?) It would be nice to include the -0.01 line in the plot, referred to in L515.

The 121 catchments were not all used because we limited these tests to semi-distributed configurations with one and only one upstream sub-catchment. We will clarify this in the figure cation and include the proposed threshold in the figure.

L552: I totally agree that it would be very interesting to see how this approach would apply to catchment with a steeper gradient/elevation and snow.

We hope that our paper will stimulate similar initiatives in other hydroclimatic and physical conditions.

References

Oudin, L., Andréassian, V., Mathevet, T., Perrin, C. & Michel, C. (2006) Dynamic Averaging of Rainfall-Runoff Model Simulations from Complementary Model Parameterizations. *Water Resources Research* **42**. doi:10.1029/2005WR004636

Pushpalatha, R., Perrin, C., Moine, N. L. & Andréassian, V. (2012) A review of efficiency criteria suitable for evaluating low-flow simulations. *Journal of Hydrology* **420–421**, 171–182. doi:10.1016/j.jhydrol.2011.11.055

Thirel, G., Santos, L., Delaigue, O. & Perrin, C. (2023) On the use of streamflow transformations for hydrological model calibration. *EGUsphere* 1–26. Copernicus GmbH. doi:10.5194/egusphere-2023-775