

Review of Hamitouche et al. revision

Main points:

The manuscript is much improved, and the authors mostly addressed my 3 main comments. I suggest Minor Revisions.

First, adding the HD results as a baseline improves the paper. The poor HD performance shows a real need for model development, and justifies the study contribution and why it fits in a Model Development journal. It would be good to highlight this more in the manuscript. In the results, I suggest including a figure from one of the locations in the main manuscript (i.e., add one location from Figure S4 to the manuscript).

Second, in their response, the authors make a compelling argument for keeping the WRF-Hydro calibration results as a separate section (rather than integrated as I had asked for). However, I didn't see the author's points reflected in the manuscript, which might help the reader see this more clearly. (I am referring to the points in their response, copied/pasted below; I suggest integrating some of these comments in the manuscript later in my line-by-line comments):

- 1. Unlike WRF-Hydro, CaMa-Flood and Noah-MP do not currently support basin-specific or spatially variable parameter calibration, and Noah-MP (within ENEA-REG) is not two-way coupled with CaMa-Flood. To ensure fairness in the corecomparison between the two routing models, we restricted the main evaluation to non-calibrated configurations.*
- 2. WRF-Hydro calibration was facilitated thanks to the availability of the PyWRFHydroCalib package, whereas no equivalent tool exists for CaMa-Flood in our framework.*
- 3. Calibration is computationally expensive, and we considered it more appropriate to present the results as an "add-on" analysis, illustrating the potential benefits of calibration where resources allow, rather than as a standard step in the intercomparison.*

For these reasons, we structured the calibration results as a separate subsection, highlighting their added value while keeping the baseline comparison consistent and unbiased. We have revised the section headings to clarify this rationale and improve the readability of the calibration analysis.

Third, the paper is improved in clarity, organization, and flow, and I only offer a few suggestions in my line-by-line comments below.

Introduction

Line 106. I would remove "interpreted as a baseline reference". Now that you have added the HD results (which could also be interpreted as a baseline reference) I think this is confusing.

Line 111-112 – replace "with a particular focus on WRF-Hydro" with "focusing on WRF-Hydro",

2. Materials and methods.

123. Suggest replacing sentence starting with “Considering”, by saying: “The ENEA-REG model, developed with the MED-CORDEX framework, (Anav et al., 2021)...”

Line 128. Are these the same 10 rivers you mention in line 118? Can you reorganize this paragraph to better connect which rivers are the largest versus which are in ENEA-REG versus the rivers that you actually validate in this study? This would help the flow.

Line 153. It might be useful here to organize the paragraph to introduce the reader systematically to your subsequent experiments for simulating daily discharge. Maybe draw upon your response points (see above). You could also note you are moving from more simple to more complex. You start with simple, CaMaFlood – which only requires runoff and doesn’t support parameter calibration. Then you do more complex – WRF-Hydro where the inputs are the set of atmospheric variables, but you do this using default parameters (uncalibrated) so it’s a fair comparison. Then, you add complexity by calibrating the WRF-Hydro model parameters.

Line 200. Seems you need to add something for the reader, either earlier or here or in the next paragraph, referring to your response points (see above) for why you added this section as a stand alone.

Line 285. Thanks for adding the HD results, which clearly show how the models you are presenting are a clear improvement. If you have space in the ms, I would suggest putting one of the locations from Figure S4 in the manuscript to more clearly make the point. If you need to swap out a figure to make space, I suggest putting Figure 7, the scatterplot of lowflow and high flow bias, into the Supplemental.

Conclusions

Line 541. Suggest explicitly saying that both models evaluated in this paper showed improvements over the HD baseline.