

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2022-173-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2022-173

Anonymous Referee #1

Referee comment on "Continental-scale evaluation of a fully distributed coupled land surface and groundwater model, ParFlow-CLM (v3.6.0), over Europe" by Bibi S. Naz et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2022-173-RC1>, 2022

The model evaluation paper of Naz et al. describes a version of the established ParFlow-CLM model applied over Europe and evaluated its hydrological components.

ParFlow-CLM is an established modeling tool, and a publication of a model evaluation paper that builds a foundation for future scientific use is certainly something I would like to support. Unfortunately, in the current stage, the manuscript does not deliver on this goal and seems to purposely hide model shortcomings. In the current version, I can only suggest significant revisions.

I would like to focus on two aspects that are currently flawed. Firstly, the paper's motivation could be much clearer from the beginning. In the light of the many publications that already exist on ParFlow and CLM, what is the added value of this model evaluation paper? What is the model's purpose within the range of continental and global models? What questions can it help to answer? Outlining this much more clearly from the beginning will be helpful for the scientific community in making this publication a helpful reference for future research.

Secondly, I cannot accept the current evaluation of the groundwater component. The authors use groundwater in the title and motivate the model's usefulness with the argument of an active groundwater component but provide a not convincing evaluation. I do not expect the model to be able to perfectly represent the water table. Still, I think we can only progress if we are open about our models' shortcomings and clearly communicate uncertainties. Poor model performance is not a reason for not publishing something as long as there is a proper discussion on the causes. Currently, the paper is not doing that and uses oversimplified evaluation methods to obfuscate the actual model behavior. Furthermore, existing literature and models are omitted as well.

Additional notes:

* While I know that it is difficult to find a repository to host a large amount of data, I employ the authors to think about if selected model outputs could be made available in the spirit of OpenScience principles!

* Is it really necessary to use the overcomplicated PF-CLM-EU3km as a name? Why not stick with ParFlow-CLM in the paper? If it is a very different model, why is that not the name used in the title?

L. 1: How are these large-scale models useful for water resource management? I see how they are helpful for large-scale policy and fostering scientific understanding but are they really useful for management? Please also define what high-resolution means in brackets - people have very different interpretations about that, and it is changing fast.

3: How is the coarse spatial resolution linked to the lateral fluxes and groundwater components - isn't that mixing up things? What small scale processes specifically?

4: what does more complex refer to? Complex in what regard?

11: what is PF-CLM-EU3km? It has not been introduced; quantify good agreement

17: this is the first-time heterogeneities are mentioned. Is it implied that this is a result of the higher spatial resolution? This should be explained

Fig. 1 c) WTD in log scale without indicating what red is. Is that deeper than 100 m? How deep is it? Why is the WTD so deep near larger rivers? Why so shallow in mountainous regions? What is the reasoning here why this is plausible? Is it plausible in the light of the performance of other large-scale models?

415: I get the problem of inconsistent WTD elevation data. Still, this should be solvable for at least some regions in Europe. I feel that the authors feared that the model performance would be judged too harshly. Whatever the reason, the solution shown here is not acceptable. Furthermore, you can't simply select only the cells that simulate WTD < 10!! This is the range almost all models do a good job. This is not advancing our science. This is far from ok.

Please show how much the model deviates from observations. You motivate your paper with the statement that representation of groundwater is essential and then skip a proper evaluation of your model.

I suspect it will not perform perfectly - no large-scale model currently can, and you are providing some reasonable answers by referring to Gleeson et al., 2021, which is good but not enough. Please provide a more extensive discussion on how the performance differs from other existing research.

417: ?? = Fig. 7