

Hydrol. Earth Syst. Sci. Discuss., referee comment RC2
<https://doi.org/10.5194/hess-2021-605-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Review of “Large-sample assessment of spatial scaling effects of the distributed wflow_sbm hydrological model shows that finer spatial resolution does not necessarily lead to better streamflow estimation” by Aerts et al., 2021

Shervan Gharari (Referee)

Referee comment on "Large-sample assessment of varying spatial resolution on the streamflow estimates of the wflow_sbm hydrological model" by Jerom P. M. Aerts et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-605-RC2>, 2022

Review of “Large-sample assessment of spatial scaling effects of the distributed wflow_sbm hydrological model shows that finer spatial resolution does not necessarily lead to better streamflow estimation” by Aerts et al., 2021

The presented manuscript is trying to evaluate the added values of modelling at finer resolutions for the streamflow simulation in a large sample hydrology framework (CAMELS data set).

I enthusiastically accepted to review the manuscript as I was interested to see the developments from the Delft team regarding hydrological modelling and infrastructure for enhancing modelling capabilities. I am a bit disappointed I should say...

In the following, are my general and specific comments:

- The manuscript fails to advance modeling infrastructure and capabilities and hydrological understanding. I think the senior co-authors can do a better job in directing and balancing coding and science for younger generations.
- It was interesting to see the mentality behind the modeling in Delft. And I am a bit puzzled why the directions are the way they are! In land surface community, which the manuscript completely misses to cover, the recent tendency is toward vector-based implementation (away from the grid-based simulation, Gharari et al., 2020, HESS, for example, advantages of vector-based setup are explained in that paper in detail). This is also true for the routing models. Additionally, the routing models and land models are more and more decoupled which means the vector-based routing models such as

RAPID or mizuRoute can use runoff simulation at any (un)structured forms (grids, HRUs, GRUs, conforming or non-conforming subbasins). The land modeling community has spent a significant amount of time dealing with the upscaling of the DEM for grid-based setup. While the vector-based routing models can be simulated for any modeling resolution with underlying routing models and parameters remains identical (I mean really identical). Additionally, the grid setup results in excessive and unnecessary computational burden. If vector implementation is used, grids with similar soil, veg, and forcings are grouped, and computational costs will be significantly lower and the most optimal (explained in Gharari et al., 2020).

- I have a very bad feeling regarding the modeling resolutions the authors used! The resolutions are much finer than the actual forcing resolution used. Basically, what we see here is just the effect of forcing resampling at a finer resolution (and temperature lapsing). As the resolutions are smaller, I am afraid the actual precipitations are not very different across the scale of modeling, so the difference is coming from another source (routing perhaps?) or numerical implementation of the model.
- Following comments 2 and 3, are we just redoing a similar simulation when going to finer resolution here?
- 22 on the scale of KGE may not be that meaningful. The authors can design a simple experiment, perturb the model precipitation with a few percent, redo the simulations and see if the 0.22 gain in KGE is warranted. Or alternatively, they can use the EMDNA data set Tang et al 2021, ESSD. Personally, I am suspecting this 0.22 is well below the margin of the KGE envelop for the forcing uncertainty. I leave the streamflow uncertainty out here!
- Additionally, and following the previous comment, the margin of KGE improvement can be evaluated using methods presented by Clark et al., 2021 (cited by the authors). Why not just try?
- Back to the issue of routing with various resolutions. My understanding is that the routing parameters at various resolutions are different, although made consistently based on the work of Eilander et al., 2021. Then the setups are slightly different at various resolutions. Is that correct? If so, how the effect of this upscaling is seen on the calibrated parameters. And why line 326-328 is stated as it is!? This reinforces the use of vector-based routing in which the routing setup (network topology and its parameters) can be kept identical (even when modelling resolution or decisions are changed). Sorry if I misunderstood anything here.
- It seems that the authors have investigated lakes and reservoirs in setting up the model (I see hydrolakes). Are there any resolved lakes (on the river network) on the CAMELS data set? How significant are the lake areas within the subbasins?
- Why 454 CAMEL subbasins and not more?
- I missed how the calibration is done. Is this a single parameter calibration? I review the method section a few times but cannot comprehend it. Suggest clarifying. And sorry if I missed it.
- Back to the title, hydrology is partly about streamflow simulation. Simulation of other fluxes, states, and processes is also important. At least looking into the snow simulation for a few of the catchments might be helpful (like Figure 5 of Gharari et al., 2020 in which there are significant snow simulation differences with similar NSE for streamflow simulation). Also, the answer to the title is clear! Streamflow can be predicted with least complexity among other fluxes and states.
- Just an opinion, the authors could use an alternative model for their work. Any reasons why this model was chosen? I mean the model was not developed for small-scale hydrological applications.
- References are a bit haphazard. Before reading the manuscript, I could somehow guess much of the mentioned literature while a significant body of literature is missing. Land surface modelling community has done investigations on the effect of resolutions. There is only one citation in this work directed to those efforts (Melsen, et al., 2016). Suggest being more inclusive (which is actually very helpful for this manuscript). Is the research question relevant? Perhaps yes, but this has been discussed time and time again. I am

surprised to see no references from 80s/90s that looked into streamflow simulation and non-identifiability or equifinality of parameters and model simulation configurations (essentially this is what the authors are trying to demonstrate here, the sensitivity of streamflow simulation to various model setup configurations). The introduction talks about MPR and self-organization, while to what I see none of those principles were used. Sometimes references are given without reflection and completeness. For example, the choice with the deficiency of objective function. The author can mention many more relevant studies. For example, Gharari et al., 2013 is very relevant here and one of the papers mentioned here (lines 351 to 360) is a specific and simplified version of that work.

- Are we looking at the scaling or just a change in modelling resolution? Scaling is more about collective behaviours at a given scale (and hopefully understanding or explaining it). Change of resolution is only a model application at different configurations. Any explanation on that?
- And final comment, I am a bit concerned about using off the shelves material and packages. I think there is a line between being a user or a developer. Are you fully confident in all the bits and pieces of your workflow in setting up the model? for example ESMValTool. It seems to be a very good package however isn't it limiting in sense of forcing preparation? Why not use a more as more data agnostic package such as EASYMORE or more elaborated ESMF? Maybe my knowledge of EMSValTool is limited but the first look it doesn't seem to be designed for forcing preparation for models primarily. Another question, why move to Julia and not go to Fortran for speed? The speed is not comparable to python for example. Also, Fortran commands are more stable over time compared to Python packages which makes maintenances much easier (a pandas commands might change in future releases).

Overall, I think the manuscript needs a major-major revision. Hope my comments are useful and hope the authors reflect on my comments. Reflection is rather missing in today's hydrology I would say. Also, don't hesitate to contact me personally if any questions arise and sorry for the delay in review process.

With regards,

Shervan Gharari

20th February 2022, Saskatoon