Dear Reviewer,

Many thanks for taking the time and reviewing the submitted manuscript. We also thank you for your kind words, and even more so your critical yet constructive evaluation.

Below, we address your specific comments (in blue) and will outline how we will improve a re-submitted manuscript, especially in the introduction and discussion sections.

What spatial extent constitutes a global hydrological model? Is the continental scale as given in the present study can be treated global? This question arises because of the statement that few studies have attempted hyper-resolution modelling over CONUS, but they do not have global coverage. In a strict sense, why can't this study be termed as a continental scale application?

Many thanks for this comment. A global hydrological model covers the entire terrestrial surface except for, at least in most instances, areas below/above +60/-60 degrees North. In doing so, it typically draws upon input datasets with global extent, a uniform parameterization, and is not calibrated regionally or locally, to facilitate comparability across the Globe. You rightly point out that our application is not 'global' as it covers only the European continent. However, our application merely employs global input datasets, a global parameter set (as it's been derived from coarser model versions), and is not calibrated regionally or locally, which is key in deriving potentially globally-transferrable findings from model evaluation. More generally speaking, one could say that a global hydrological model allows continental scale applications everywhere around the globe, but a continental model not for global applications (even though the numerical scheme could probably be fed with global data and be run if the model does not depend on locally calibrated parameter values).

In the revised version of the manuscript, **we will make it clearer that we our aim is to evaluate a global hydrological model in a continental scale application** as the 1 km version of PCR-GLOBWB is still very much under development and therefore its evaluation profits greatly from being applied to data-rich areas now (see also your other remark below). On the longer term, the model will of course be applied and evaluated globally.

The major issue in hyper-resolution modelling is modelling the physical processes happening at smaller scales. When developing a hyper-resolution model over large spatial extent, the physical processes to be considered would vary from region to region. How to account for the spatial variation in physical processes in the model? Can a generic model be applied over the entire continent/globe without accounting for region specific physical processes? Or how to develop model which can consider automatically, the various hydrological processes appropriate for a region within the model domain?

We kindly thank the reviewer for the critical comment. Indeed, representing physical processes happening at smaller scales is one of the “grand challenge” (Wood et al., 2011; Beven and Cloke, 2012). While for coarser spatial resolutions, it was deemed acceptable to subsume these processes in one or more parameters, this approach reaches its limits at hyper-resolution. To what extent it is possible to ‘blindly’ apply a single process representation with varying parameters at hyper-resolution (and how to move forward from there) is in fact one key questions of the submitted manuscript (see page 3, lines 5-6). While we already reflect on the answer to this question and their implications in the current version of manuscript various times (e.g.: page 1, line 24; page 15, line 36), **we will sharpen the focus on a revised version**.

Please note that it is a key feature of global hydrological models to use one uniform way of describing or parameterizing physical processes world-wide. Balancing generality and specificness determine hereby the overall accuracy of the model. The range of available literature on global hydrological modelling (see Bierkens (2015) for the latest state-of-the-art review) strongly indicates that such an approach is scientifically sound and that its results are of societal importance. In a revised version of the manuscript, **we will re-formulate the relevant sections such that the interplay between continental/regional specificity and global modelling approach becomes more evident**.

One reason that is often mentioned as an advantage of hyper-resolution modelling is the ability to simulate hydrological processes over data scarce regions. If data scarcity prevents us from developing a detailed model over a particular catchment, then how can be confident about the processes simulated by a global model over such data scarce regions? Further, in hydrology, studies are there to demonstrate the transfer the information obtained over a data-rich region to a data scarce region with similar characteristics. How global hyper-resolution modelling will add value to the existing methods in understanding the processes over data scarce regions?

This is an interesting question. We would argue that it is not the advantage of hyper-resolution modelling to be able to simulate hydrological processes over data-scarce areas. It is rather the advantage of global hydrological models in doing so, regardless the spatial resolution chosen. Nevertheless, we agree that global hydrological models should not be derived in data-scarce areas and then applied to data-rich areas, but vice versa. This is also why we opted for Europe, a data-rich region, as initial test case for the 1 km evaluation and starting point for model development efforts on the global scale. Once a global model is developed, its transferability to data-scarce areas can be tested. Based on the 10 km model, overall transferability of model skill is expected to be good (see Fig. 2a and Fig. 8a in Sutanudjaja et al. (2018) for a comparison of discharge and terrestrial water storage with observations). As such, the here presented study does not intend to improve modeling of data-scarce areas, but pave the way for improved representation of hydrological processes “everywhere and locally relevant” via hyper-resolution (Bierkens et al., 2015).

As the topic of knowledge transfer from data-rich (what we do in the submitted manuscript) to data-scarce areas (what we will do in future research) is of importance, as you rightly indicate, we will put extra emphasis in discussing this topic in revised version of the manuscript.

On a similar note, is it possible to develop a nested model structure that is followed in numerical weather modelling? i.e., develop coarse resolution model over large region and the outputs of this would act as boundary conditions of a nested model over a smaller spatial extent but at a much finer spatial resolution.

Thank you for your remark and great suggestion. Such nested model structures are not very common in global hydrological modelling, most likely because producing bespoke local high-detail models is a bit against the 'philosophy' of global hydrological modelling. What is instead done in instances where finer output is required than model resolution allows, is downscaling. There are, however, a few examples where nested modelling approaches are used, but then with the aim to include physical processes that are not represented by a GHM, such as detailed two-dimensional floodplain routing (Hoch et al., 2019) or coastal boundaries (Eilander et al., 2022). This is not to say that applying nested models is not viable, but it would require common efforts of global hydrological modelers and experts at the regional/catchment scale. Another option would be to use flexible meshes which can be refined where needed. Surely, both are avenues which hydrological studies should explore in more detail to advance the societal impact of hydrological models. At the moment, unfortunately, this will be only feasible for bespoke case studies and cannot be automated for any area of interest – which is, again, against the global transferability and comparability mindset of the global hydrological modelling approach.

Can the authors throw some light on the improvements to be made on the numerical aspect of the models? i.e., how to improve the efficiency of the models through novel and recent numerical schemes? This might save time during model runs.

We thank the reviewer for pointing this out. As we have mentioned in the manuscript already, we applied the same numerical scheme as used in the 10 km model (Sutanudjaja et al., 2018). Our study is hence a 'blind-test' how good this scheme would work when directly applied to a much finer spatial resolution. While the overall ability for hydrological simulations is provided, the current scheme is not suitable to execute lateral processes between cells efficiently, which has consequences for the model's ability to simulate groundwater and river discharge (even though our validation shows very satisfying performance over Europe), as we also point out in the manuscript (page 5, line 20) already. Possible ways forward are mentioned too (page 15, last paragraph; page 16, first paragraph and line 6), but **we will ensure that the role and limitations of the numerical scheme is discussed more clearly** in the revised manuscript. Also, **we will expand this section** and include additional relevant literature such as the LUE framework (de Jong et al., 2022), Distributed memory parallel modeling (Verkaik et al., 2021), or the option for running models on GPUs (Shaw et al., 2021) or XPUs in general.

Why can't the 1K model be validated on a grid-by-grid basis using the available high quality in situ observations even for a smaller time period? For example, soil moisture and ET can be validated using in-situ datasets. Further, for soil moisture, comparison can be made against SMAP data that are available at relatively higher spatial resolutions than the ESA-CCI data? Similarly, for ET too, comparison can be made against available high-resolution products such as MODIS 16 ET, PML-V2 product (Zhang et al., 2019) etc. This can be useful to test if the model is really performing in a hyper-resolution manner. At present, I feel the model evaluation is not rigorous enough.

Many thanks for bringing this up. You are indeed right that a second layer of evaluations is needed, namely against high-resolution satellite or in-situ observations. Research is already on-going or planned for various catchments where such data is available. As the submitted manuscript is intended to serve as a baseline study of where the 1 km PCR-GLOBWB model stands and what current challenges are, we did not want to include this second layer as it would overload the manuscript.

Moreover, there is a practical limitation of using such high-resolution data (e.g. 500 m data) in the research framework of this study. As we evaluate three model resolution (1 km, 10 km, and 50 km) with observational data at 25 km and 300 km, a spatial aggregation level was needed which allows for a fair yet thorough comparison. Hence, we used water provinces over which we averaged results. The added value of very detailed observational data would thus be diluted by spatial averaging. Because of this, we do not claim but only hypothesize that the resolution of observational data impacts the outcome of the evaluation. **In the revised manuscript, we will lay out better the motivation and limitations of the use of water provinces** in our study.

Due to our main goal of having a very robust baseline study, we furthermore opted for observational data which has a long record rather than having highest spatial resolution available. We think that employing data from 2015 onwards (as in the case of SMAP data) does not include enough variability to allow for answering our research questions to our fullest satisfaction. For follow-up studies, especially those focusing on hyper-resolution data only and not on a wide range of spatial resolutions, this dataset can become key in the validation strategy. **In the revised manuscript, we will better highlight the need to compare against observations with sufficiently fine spatial resolution** in future research activities.

References

Beven, K. J. and Cloke, H. L.: Comment on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water" by Eric F. Wood et al., Water Resour. Res., 48, https://doi.org/10.1029/2011WR010982, 2012.

Bierkens, M. F. P.: Global hydrology 2015: State, trends, and directions: Global Hydrology 2015, Water Resour. Res., 51, 4923–4947, https://doi.org/10.1002/2015WR017173, 2015.

Bierkens, M. F. P., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., de Roo, A., Döll, P., Drost, N., Famiglietti, J. S., Flörke, M., Gochis, D. J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R. M., Reager, J. T., Samaniego, L., Sudicky, E., Sutanudjaja, E. H., van de Giesen, N., Winsemius, H., and Wood, E. F.: Hyper-resolution global hydrological modelling: what is next?: “Everywhere and locally relevant,” Hydrol. Process., 29, 310–320, https://doi.org/10.1002/hyp.10391, 2015.

Eilander, D., Couasnon, A., Leijnse, T., Ikeuchi, H., Yamazaki, D., Muis, S., Dullaart, J., Winsemius, H. C., and Ward, P. J.: A globally-applicable framework for compound flood hazard modeling, Hydrological Hazards, https://doi.org/10.5194/egusphere-2022-149, 2022.

Hoch, J. M., Eilander, D., Ikeuchi, H., Baart, F., and Winsemius, H. C.: Evaluating the impact of model complexity on flood wave propagation and inundation extent with a hydrologic–hydrodynamic model coupling framework, Nat. Hazards Earth Syst. Sci., 19, 1723–1735, https://doi.org/10.5194/nhess-19-1723-2019, 2019.

de Jong, K., Panja, D., Karssenberg, D., and van Kreveld, M.: Scalability and composability of flow accumulation algorithms based on asynchronous many-tasks, Comput. Geosci., 162, 105083, https://doi.org/10.1016/j.cageo.2022.105083, 2022.

Shaw, J., Kesserwani, G., Neal, J., Bates, P., and Sharifian, M. K.: LISFLOOD-FP 8.0: the new discontinuous Galerkin shallow-water solver for multi-core CPUs and GPUs, Geosci. Model Dev., 14, 3577–3602, https://doi.org/10.5194/gmd-14-3577-2021, 2021.

Sutanudjaja, E. H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H. C., Drost, N., van der Ent, R. J., de Graaf, I. E. M., Hoch, J. M., de Jong, K., Karssenberg, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M. W., Vannametee, E., Wisser, D., and Bierkens, M. F. P.: PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model, Geosci. Model Dev., 11, 2429–2453, https://doi.org/10.5194/gmd-11-2429-2018, 2018.

Verkaik, J., Hughes, J. D., van Walsum, P. E. V., Oude Essink, G. H. P., Lin, H. X., and Bierkens, M. F. P.: Distributed memory parallel groundwater modeling for the Netherlands Hydrological Instrument, Environ. Model. Softw., 143, 105092, https://doi.org/10.1016/j.envsoft.2021.105092, 2021.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, Water Resour. Res., 47, https://doi.org/10.1029/2010WR010090, 2011.