Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-378-AC1, 2020

Hydrology and Earth System Sciences
Discussions

Interactive comment on "HESS Opinions: Improving the evaluation of groundwater representation in continental to global scale models" *by* Tom Gleeson et al.

Tom Gleeson et al.

tgleeson@uvic.ca

Received and published: 17 November 2020

RC1 - Keith Beven (Referee) received and published 1 September 2020:

Is evaluation of a global model the right approach?

I understand the requirement to have improved groundwater modelling capability in earth system science but I am really not sure that this paper is advocating a correct strategy for achieving that. It starts from the viewpoint of a community of global groundwater modellers that there are global groundwater models available that need to be evaluated (with a view to improvement). As such it completely ignores the experi-

ence of what are here called the regional hydrogeologists in implementing operational groundwater models (with all their difficulties of conceptual models of the geology, spatial heterogeneities in transmissivities, fracturing, disconnections between layers and local confinement, patterns of (sometimes unlicensed) abstractions, etc etc). There is, for example, no mention of the Danish National Water Resources Model that has tried to do this at a national scale (and even then run into scale, conceptualisation, and parameterisation problems). So, in that groundwater is framed by local geology, which can vary at below the global groundwater model grid scale, it would seem to be much more productive if the approach to the global problem was to provide a portal to make use of that regional information much more directly than the portal for model evaluation suggested here.

We deeply appreciate Prof. Beven's thoughtful commentary and suggestions - these provided many opportunities to reflect again on past, present and future large-scale modeling that includes groundwater parameterizations. Prof. Beven is making the point that no global model will have the same level of detail that local models have. That is true, however, the solution is not that we therefore need to reproduce all the local information in the global model. We do not need to reproduce the world at hyper-resolution to address some of the questions we want to answer at the large scale. The question is rather what aspects of the local system have to be reproduced and what scale/resolution is required to do so. We agree that the experience of regional hydrogeologists in implementing operational groundwater models is very important and was under-represented in the manuscript. A number of the coauthors have directly developed (e.g. Fogg, Cuthbert, Bierkens, Zipper, Gleeson, Bresciani) or evaluated regional groundwater models (Hill), or are close project collaborators of people who do. So even though we did not emphasize this experience proportionally, we bring this important perspective to the problem of evaluating large-scale models as well.

In Section 2 (Current Model Evaluation Practices) we had a final paragraph highlighting the evaluation of regional-scale groundwater models, as well as a list of five ways that

evaluation of describing how regional-scale and large-scale model evaluation practices are different. In order to provide greater emphasis on regional models we moved this paragraph on regional scale groundwater model evaluation forward to the beginning of Section 2 and expanded on it as the new section '2.1 Learning from regional-scale modeling'.We also clarified the important differences between large-scale and regional scale models in the new Section 2.2 Adding subsections to Section 2 also helps with the overall flow of this section. We agree that the Danish National Water Resources Model is a good example of such an effort as are numerous regional models by the USGS starting with the Regional Aquifer System Assessment (RASA) program in the 1970's. We feel it is beyond scope to mention or describe individual models or modeling programs, but rather more important to point towards efforts to synthesize regional models and learnings about these models. With this in mind we now referenced the Rossman & Zlotnik (2014) review in Section 2, which is a useful synthesis of regional models across the western United States. To our best knowledge there is no other similar synthesis of regional models for other regions of the world. As a side note, we have referenced Troldborg et al. (2007) and Refsgaard et al. (2007) which are both related to but not directly about the Danish National Water Resources Model.

Finally, we note that we fully agree that a portal to make use of that regional information would be highly useful. This is exactly what we meant by including "regional-scale models that meet the standards described above and could facilitate inter-scale comparison (Section 3.2)" in our list of desired ingredients in our envisioned 'Groundwater Modeling Data Portal". (it is interesting that "geology" as such does not appear in the text – only in authors affiliations – we have to infer it from "conceptualisations").

This is a useful and interesting observation and reflection, and Prof. Beven is correct that the crucial role of 'geology' on groundwater systems is somewhat obfuscated under 'conceptualizations' in the Opinion manuscript. As noted by the author affiliations, many of the coauthors identify as, and have foundational training in geology or earth science. In addition, we have been involved in various efforts such as developing the

global permeability maps based on a compilation of geology maps and data for large-scale groundwater analysis and modeling (Gleeson et al. 2011; Fan et al. 2015; Fan et al. 2020).

An earlier version of this manuscript entitled "Groundwater representation in continental to global hydrologic models: a call for open and holistic evaluation, conceptualization and classification" (also available on as Version 1 of this manuscript on EarthArXiv) had an entire section entitled "SYSTEM CONCEPTUALIZATION" that we are removed from the HESS Opinion submission to focus the manuscript on model evaluation; we plan a follow-up manuscript on the role of different geological environments and different approaches to conceptualizing the key geological elements of these geological environments. In sum, we fully agree that geology is crucial to groundwater analysis and modeling at all scales, we feel it is out of scope of this manuscript to focus specifically on geologic differences in groundwater systems in different regions or more generally.

To acknowledge the critical role of geology, we added a sentence about this as a possible future direction in large-scale models that represent groundwater in Section 1. We primarily refer to Condon et al. (in prep) which contains much more detail, as described above. There will, of course, be gaps in the global coverage where there are important groundwater bodies but where no regional or local models are available and data is poor. Certainly in those situations we would need to resort to expert elicitation in creating a suitable model to make the coverage more complete. But that is a different problem. Because bringing in regional hydrogeologist expertise to evaluate the global modelling (as suggested) would seem to be doomed to failure. The global grid scale and variation in parameters is too crude. We agree that expert elicitation will be challenging because of many differences such as the technical ones mentioned here (grid scale and parameter variation) but also perceived or real differences in perspectives, priorities and backgrounds between regional-scale and large-scale modelers. But we think that this challenge could be potentially worth it but for the pure scientific advances but also the potential new synergies between these two communities. As described

above, many of us move between and have strong connections in both communities, so we can be the bridges to help make these connections. Scale and commensurability remain crucial scientific challenges so these efforts can hopefully contribute to these challenges.

This type of direct comparison using expert elicitation as Prof. Beven is suggesting here would be useless. Such an expert elicitation exercise has to be done in the context of the purpose of the global model. What is the model for and what local properties have to be preserved to achieve this. Clearly we would not propose to elicit a hillslope hydrologist regarding her specific knowledge of a particular hillslope to evaluate a catchment scale model (the only answer could be that the model has to be rejected). However, we might reasonably ask her to suggest what variability we should expect across all hillslopes in a region (or something like that).

Therefore in Section 3.2 we added this text: "Expert elicitation also has a number of challenges including: 1) formalizing this knowledge in such a way that it is still usable by third parties that did not attend the expert workshop itself; and 2) perceived or real differences in perspectives, priorities and backgrounds between regional-scale and large-scale modelers.

The generation of recharge rates and evapotranspiration rates when the water table is near the surface just cannot be properly represented when the grid scale cannot reflect the local variations in topography, but there is nothing here about evaluating such fluxes (and getting such boundary conditions right is surely rather important.... Or is that being left to the global land surface modellers rather than the global groundwater modellers?).

The scale dependency of recharge rates and evapotranspiration is important and many large-scale models are too coarse to correctly resolve groundwater-surface water interaction and the interaction with evaporation unless subgrid-processes are incorporated.

Some advancements strongly depend on increasing the resolution of global models

as extensively discussed Fan et al. (2019, WRR). We are not disputing this at all. However, we still need to find a hierarchical strategy to evaluate whether global models are good enough for their purpose, or whether they lack key behaviours that need to be preserved. These could be included either through increased resolutions or through statistical sub-grid representations.

The land surface models that calculate diffuse and concentrated groundwater recharge have sub-grid parameterizations to account for un-resolved fractions of saturated soils (TOPMODEL, ARNO-scheme) and the two-way coupled version of our global groundwater model accounts for groundwater discharge through small streams, seeps and springs by a subgrid parameterization. Furthermore, resolutions are increasing (look at PARFLOW across the U.S and Europe; we have recently been running 1 km transient 2 layer MODFLOW 6 for the globe for 50 years within 1 day), so these issues will become less important in the near future.

It is also important to note the formerly clear boundaries between land surface modellers and groundwater modellers are dissolving so this classification is less useful now. Each model has different strengths, capacities and approaches to solving for water fluxes and stores below, at and above the land surface.

Finally we note the recharge rates and evapotranspiration are mentioned in Section 2.3 and included in Table 1. Since these are standard fluxes for evaluation of many large-scale models, it does not seem necessary to describe them in detail in Section 2.3. The paper recognises the issues of commensurability, but has no suggestion for how to take that into account (except for the use of "signatures" but it is then not explicit about how that might actually work).

This was a useful critique - we did spend much more time recognizing the issues of commensurability than directly taking it into account. We note that in Section 3.2 we describe how inter-scale model comparison could use variables routinely simulated in regional-scale models such as baseflow or recharge to evaluate large-scale models.

“In this way, the output fluxes and intermediate spatial scale of regional models provide a bridge across the “river of incommensurability” between highly location-specific data such as well observations and the coarse resolution of large-scale models.”

In our minds, a number of the future priorities outlined in Section 3.1 can partially address or at least improve upon some of our current commensurability challenges (Section 2) - but this did not make it from our minds to the manuscript! For example, using data that is more consistent with the scale modelled grid resolution will reduce the commensurability challenges. We therefore added this text to Section 3.1: “Using data (such as baseflow, land subsidence, or the spatial distribution of perennial, intermittent, and ephemeral streams) that is more consistent with the scale modelled grid resolution will hopefully reduce the commensurability challenges.” And some of the suggestions for “evaluation” seem to me to be rather circular (see comments on manuscript).

In the attached we have also responded to each of the detailed manuscript suggestions as comments in the pdf.

And then there is karst. This is rather important in some parts of the world. It does get just a passing mention in the text (through the paper of Hartmann et al., 2017) but there is no discussion of how this might fit into a global model based on PDE continuum approaches.

We agree that karst is crucial to consider in groundwater modeling and one of the co-authors (Andreas Hartmann) is an emerging leader on new approaches to karst in large-scale models. Here too, there is significant overlap with Condon et al. (in prep., so we feel that it is not appropriate to add significant text about model development to this manuscript that is focused on model evaluation. Therefore, we added a reference to karst in Section 1 and in this reply we give a more fulsome discussion of the role of karst in large-scale modeling: 1. When regional balances are considered, equivalent porous medium approaches that do not explicitly consider karst have shown to perform not too bad (Scanlon at el, 2003 JoH), which may justify disregarding karst het-

erogeneities/discontinuities to some extent. 2. A new paper using WoKaS database (Hartmann et al., 2020) shows that the delay of the recharge signal through karstic groundwater dynamics has to be considered at the seasonal scale while Hartmann et al. (2017) showed that recharge is a very sensitive variable that large scale models need to account for. 3. For recharge and for groundwater, karstic discontinuities are important to groundwater quality (not the focus of this manuscript). Global water quality data sets, when finally compiled, will help to evaluate how much the lack of fast flow representation affects large-scale model simulations (Hartmann, 2016). 4. Regional applications of karst specialized groundwater models (MODFLOW-CFP, Reimann and Hill, 2009, among others) that provide combined PDE applications for laminar Darcy flow in the karst matrix and turbulent flow in a discrete conduit network may provide some directions for future improvements.

The paper is the outcome of a workshop on global groundwater modelling but for all the expertise available it seems to me to be wrong about how to approach the problem (perhaps because the expertise of operational groundwater modellers was not that well represented). Evaluation of the type of global groundwater models being suggested is not really the issue. For all sorts of reasons we can expect that they will be too crude and too approximate and will not make best use of local information where that is available (even to the point of local rejection). Global groundwater is an aggregation of regional and local groundwater systems with all their different geological and other characteristics. If the problem of using that regional and local information directly is only computational, then ways could be found of simulating the responses more efficiently (not necessarily using a coarser grid, but, for example, perhaps using machine learning). Where a regional or local model cannot be used because of bureaucratic reasons, then it will be necessary to construct a simulator in the same way as for a data-sparse area, but again without data there can be no evaluation (the expert elicitation will already have been used in the construction). I would suggest, therefore, that this paper needs more, and deeper thought, and should not necessarily have a starting point of here is global groundwater model how do we evaluate it, but rather here

are all the important aquifers worth representing, how should their response be best represented (which might of course be locally/regionally a PDE continuum model – or not)?

We agree with Prof. Beven's thoughtful suggestion of focusing on regional differences. In Section 3 we had this text: "Considering regional differences in model evaluation suggests that global models could in the future consider a patchwork approach of different conceptual models, governing equations, boundary conditions etc. in different regions. Although beyond the scope of this manuscript, we consider this an important future research avenue."

We expand on this idea of a patchwork approach by adding this text to Section 2: "Given the strengths of regional models, a potential alternative to development of large-scale groundwater models would be combining or aggregating multiple regional models in a patchwork approach (as in Zell and Sanford, 2020) to provide global coverage. This would have the advantage of better respecting regional differences but potentially create additional challenges because the regional models would have different conceptual models, governing equations, boundary conditions etc. in different regions. Some challenges of this patchwork approach include 1) the required collaboration of a large number of experts from all over the world over a long period of time; 2) regional groundwater flow models alone are not sufficient, they need to be integrated into a hydrological model so that groundwater-soil water and the surface water-groundwater interactions can be simulated; 3) the extent of regional aquifers does not necessarily coincide with the extent of river basins; and 4) the bias of regional groundwater models towards important aquifers which as described above, underlie only a portion of the world's land mass or population and may bias estimates of fluxes such as surface water-groundwater exchange or evapotranspiration. Given these challenges, we argue that a patchwork approach of integrating multiple regional models is a compelling idea but likely insufficient to achieve the purposes of large-scale groundwater modeling described in Section 1. Although this nascent idea of aggregating regional models is

beyond the scope of this manuscript, we consider this an important future research avenue, and encourage further exploration and improvement of regional-scale model integration from the groundwater modeling community.."

We hope this text distills the possibilities, but also the challenge of this approach. There are many more comments on the manuscript. Please also note the supplement to this comment:https://hess.copernicus.org/preprints/hess-2020-378/hess-2020-378-RC1-C3 In the attached we have also responded to each of the detailed manuscript suggestions as comments in the pdf.

Please also note the supplement to this comment:
https://hess.copernicus.org/preprints/hess-2020-378/hess-2020-378-AC1-supplement.pdf

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-378, 2020.