Review of 'Large-scale sensitivities of groundwater and surface water to groundwater withdrawal' by Bierkens et al.

for Hydrology and Earth System Sciences

Summary

Bierkens et al. present a novel analytical approach to assess the impact of groundwater withdrawals on the disconnection of streamflow and groundwater, that can be applied globally. The approach further facilitates the calculation of critical pumping rates that are needed to maintain the surface-groundwater connection. Results include an assessment of model parameters and qualitiative comparisons to other studies.

Recommendation

I am a new reviewer to this submission, which has already undergone a first round of major reviews. Thus, my comments focus on how well the authors have addressed the reviewer comments - but I am adding a few comments that I deem important and may require a few (minor) adjustments.

In general, I think that the authors addressed most (if not all) major comments from the reviewers. Major changes include, e.g., a reformulation of the manuscript focus (from sustainability to the disconnection of streamflow-groundwater interactions), additional validation and comparison exercise using other data sets (originating outside the group), additional supplementary materials were added, among other minor revisions to address the reviewer's concerns. There is just one overarching concern, raised by all reviewers, that could be addressed in a slightly better way. I would thus recommend that the authors revise their response to this concern in another minor revision.

Issues with validation / intercomparisons

All reviewers requested additional validation/comparisons to other studies. In response, the authors added (i) a comparison to other input data sets (i.e., critical transition times from Cuthbert et al), and (ii) qualitative intercomparisons to other global (de Graaf et al.), continental (Condon and Maxwell), and regional/local (Wen and Chen) studies. The latter includes a local validation with streamflow and groundwater head trends for the Republican River basin in the US.

While the authors went through a lot of effort employing these other data sets for the validation and intercomparison - which is appreciated -, I believe this is the only important part that may need clarification. While I understand that these intercomparisons are difficult, especially since most of these variables cannot be measured, I think a bit more effort in explaining the differences and the expected shortcomings of the approach is needed.

This concerns, for example, the differences in critical transition times in Fig. 5 derived from the two data sets (Sutanudjaja et al.; and Cuthbert et al.). The differences here are between a factor 100 and a factor 1000. I am not an expert on this, but it is indeed "quite striking" (l. 399). The text does not really explain what the impact of these differences on the further results in the manuscript is (and if these are used at all). The explanation (l. 399-404) is not very convicing to me (but again, I am not an expert).

Further, the intercomparisons to other, e.g. numerical models such as ParFlow-CLM, remain rather qualitative. I would be okay with that if it was intuitive, but the text barely touches upon the reasons for differences. Instead, the comparisons are lacking behind. E.g., l. 546f: "Figure S12 shows again that the analytical approach yields larger depletion estimates than ParFlow, but the results are more similar than with de global model of De Graaf et al (2019)." ... but likely for different reasons, no? Also, for the comparison to de Graaf, the differences appear (maybe only) larger outside of the US.

Along these lines, personally, I think that a few more notes related to the shortcomings of the approach and its expected performance may be required at the end / in the summary. This is especially, because applications of this analytical approach with other data sets / models are encouraged. So I think the reader needs to know (i) why differences to the aforementioned studies appear, and (ii) where the model is expected to perform well and where not. I think this concern is in line with concerns from all reviewers (esp. #1 and #3), who criticize some of the assumptions - even though they are explicitly mentioned throughout the text.

Some more minor notes

- Please check the consistency of units throughout the text (e.g. use SI unit "d" instead of "day" and unify "yr", "y" and "year"
- l. 657-670: this should maybe be moved to the methods
- l. 104-105: check sentence; "these transitions do not occur" and "is that ... is that"
- l. 255: It should be "h_s(infinity)" I think
- l. 283: check sentence; "that are of interest to show" or remove "of"
- l. 360: and?
- l. 374: delete one "the"
- Fig. 6+7: could the authors add what is considered 'negligible' ?
- l. 545: "de" --> "the" (a little bit more Dunglish ;))
- Table S1: small v instead of large V for consistency with main manuscript?

Concluding remark

In general, I am very appreciative of the work and fully support the notion towards large scale hydrology. In particular, I welcome the broad applicability of the approach, e.g. at various scales and with other models/observations - and concur with the authors that this approach may be useful for (i) bridging the time gap until global numerical approaches are ready, but (ii) also for benchmarking, intercomparisons and uncertainty analyses.