Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-564-RC1, 2017

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Mapping (dis)agreement in hydrologic projections" *by* Lieke Melsen et al.

J. Kiesel (Referee)

jenskiesel@gmx.de

Received and published: 11 November 2017

The manuscript "Mapping (dis)agreement in hydrologic projections" by Melsen et al. shows an extensive analysis of selected uncertainties inherited in climate change projections. The authors consider a cascade of uncertainties: three different hydrologic models; different model parameterizations that pass a performance threshold; five different climate models; 605 catchments across the contiguous US. They analyse these combinations on their direction of change of future mean discharge and timing of discharge. The results show significant disagreement in the direction of change for each uncertainty component.

The paper is well structured, well written, the results are well presented and the discussion and conclusion cover the major points. I therefore suggest minor revision.

The major comment I have is related to a more detailed discussion of the behavioural

model runs. Please find detailed comments regarding this and other issues below.

METHODOLOGY

p.2 l.26: I suggest to mention the information about which RCP you used in chapter 2.2

p.3 Fig 1d is not clear to me: why two bars in the upper positive change section? What is the black line with the two points and how is it created? Why is 'Frequency' written below the box?

p.4 l.1ff: Model description: I miss one sentence for explaining the model concept about runoff components (surface, lateral, groundwater) which are important for discharge timing

p.4 l.3: moisture

p.5 l.1: what are the "100 base runs with average parameter values" - how were these defined? e.g. nothing about this is mentioned in the parameter tables (Annexe)

p.5 l.4: I miss one or two sentences about the other input data. I know it is publicly available in the CAMELS dataset. But I think it is important to know the very basics here: Topography, land use, soil/geology and if catchment management (irrigation, damming) is considered and included in the dataset and in models. If catchment management is not included, could that be a reason for the non-behavioural catchments? I.e. the central US is subject to the highest ratio of agriculture and this distribution seems to fit well to the non-behavioural spots (https://www.agcensus.usda.gov/Publications/2007/Online_Highlights/Ag_Atlas_Maps/Farms/M085.php). If relevant, this could be an additional point for the discussion (see my comment at p.8 l.18ff).

p.5 l.8: For future studies, consider using the KGE': "For the variability ratio c we used CVs/CVo instead of rs/ro, which was proposed in the original version of the KGE-statistic (Gupta et al., 2009). This ensures that the bias and variability ratios are not cross-correlated, which otherwise may occur when e.g. the precipitation inputs are

biased." Kling et al. 2012, JOH.

p.5 l.16 and p.6 l.2: Please comment on why the two time periods differ in length. When comparing aggregated/average metrics the length of the comparison period is important since the longer the time series, the less influenced are the metrics by singular annual extremes.

p.5 l.18: I suggest to explain the choice of the rcp briefly in a sentence.

p.5 l.20: what is meant by "member" (i.e. what distinguishes the different 'members' of each GCM family - regional climate model, version, resolution, year, ...)?

p.6 l.6-7: I suggest to define what you mean by "ensemble mean change". E.g. I think something along these lines is clearer: "The ensemble mean change was then determined as the mean change over all behavioural parameter sets of each GCM-, hydrological model-, and catchment combination".

p.6 l.12: I assume number of "representative sample of parameter sets" is defined through the behavioural runs. For the other two uncertainty sources we know the number (three for the hydro models, five for the GCMs) - but for the chosen parameter sets you do not show them. However, I think it matters how many runs in each catchment are used to produce all the subsequent results. Could you show three additional maps of the CONUS (could also go to the Appendix) where the color of each catchment dot indicates the number of behavioural runs for each hydrologic model?

p.6 l.18-26: I really like this part of the analysis, but the paragraph is difficult to understand without having seen the results and I suggest to begin the paragraph with an explanation, e.g. something along the lines: "It is assumed that catchment characteristics can influence the agreement between hydrological models and GCM. To assess the influence on the hydrological model agreement, we divided all basins into three categories: ... "

RESULTS

p.8 l.7-8: Is this result not better suited for the section 3.2.1?

p.8 l.18ff: I think this is an important paragraph which is valid for the other sources of uncertainty as well. I.e. if you end up with model runs that generally depict the processes better, you may end up with less disagreement for the other uncertainty causes as well. So, an interesting hypothesis to test would be, if most of your mapped disagreement is caused by parameter sets at the lower end of the KGE and if high KGEs lead to higher agreement (though I am interested in this, this is just a side note, no need to do this within this paper).

However, I think the paragraph fits better to the discussion and I suggest to add:

- that it could be possible that improved process depiction in your models could reduce disagreement related to other uncertainty sources as well

- numerous studies (e.g. Pool et al. 2017 HESS 21) have found that looking at certain metrics without having used them in the optimization (in your case: selection of behavioural runs) can cause inadequate depiction of those metrics - so the actual selection of the objective function may influence (dis-)agreement

- a short statement if you can rule out that the non-behavioural results could be due to the selection of the parameters and ranges which may be more suitable for conditions significantly different from the catchments that are non-behavioural

p.8 l.29: Suggest to change to "... hydrologic models (dis)agree on the sign of..."

p.8 l.30: "in the north-east the models..."

p.11 l.6: suggest to replace "...was able to capture current..." through "...was non-behavioural (Figure 4e)"

p.11. l.29: a lower aridity? Figure 5.d. suggests higher aridity? how can aridity be both high for disagreement and non-behavioural catchments? also at Figure 7.d: the significance triangle for mean delta P should point down and be hollow or? Seems

like I have difficulties understanding the rose plots. If the plots are correct, I require more explanation how they need to be interpreted (e.g. already in the methods with an example rose plot).

p.15 l.10: This chapter is a very good summary of the uncertainties. But I miss that you explain how the combined uncertainty is produced in the methods. Did I miss something?

DISCUSSION

p.18 l.1: Ehret et al. 2012 (HESS Opinions) is also a suitable reference to this statement.

CONCLUSION

I suggest to add a few sentences which hydrologic model, which GCM and which combination led to the highest (dis)agreement. I know that the information is scattered throughout the results, but I would have liked to see this information summarized in the conclusions. I envision these sentences as a very concise summary of the whole Appendix.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-564, 2017.