

Interactive comment on “Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration” by H. Müller Schmied et al.

Anonymous Referee #2

Received and published: 3 April 2014

This paper explores the sensitivity of a number of hydrological indicators to five perturbations of the WaterGAP GHM; changes in 1) climate (using different climate forcing datasets for 1971–2000), 2) land-cover (using different homogenised land-cover datasets), 3) structure (not different model parameters but an exploration of the effect of implementing new parameterisation schemes), 4) human water use, and 5) whether the model is calibrated or not.

It's nice to see a paper that explores uncertainties within a macro-scale hydrological

C796

model. Papers like this are quite common within the catchment-scale hydrological modelling community so it's good to see the practice transfer to larger-scale models because there is a need to understand the uncertainties that are inherent within individual global hydrological models (GHMs) and not only across different GHMs. To this end, I consider the article to be a valuable contribution but I would like the authors to address my comments below.

General comments _____

1) It could be argued that there is little value in the NoCal-STANDARD comparisons that appear throughout the paper. This is because all GHMs to some extent will be calibrated, whether it is through a comprehensive basin-level calibration with a runoff correction factor (e.g. WaterGAP), or simply by tuning model parameters until there is a good fit (e.g. MacPDM). No model will ever be used without some level of tuning having taken place during the model development stage. Thus the NoCAL-STANDARD maps that show large differences (e.g. Fig 7 and Fig 9) is hardly surprising and offers little value to the paper. It could also be argued that the inclusion of this comparison skews the conclusions of the paper, because it would be considered infeasible and inappropriate to run a model that had not been tuned or calibrated at any point during its construction. To put it another way, it's a bit like comparing the STANDARD model with a model where the LAI parameter for each vegetation type is increased by 200%, to explore the effect of vegetation parameterisation. While this would tell us something about how hydrological response is sensitive to LAI, the leaf sizes would be unrealistic in some cases, so a comparison of the two simulations would be nonsensical. In the same way, here the authors are indirectly suggesting that it would be feasible to run an un-calibrated model, when of course it is not. I suggest that the authors either miss out the NoCal comparisons from a revised version of the manuscript or that they provide a rigorous justification for why they make the NoCal comparisons. If the authors decide to keep the NoCal comparisons, then can they also please explain clearly and concisely, which parameters are calibrated in the STANDARD version – i.e. is it only γ , CFA

C797

and CFS that are perturbed during calibration, or are other parameters related to, for instance, interception capacity, rooting depth, and field capacity etc. perturbed too?

2) It's good to see that the authors acknowledge on line 11 of page 1589 that they neglect parameter uncertainty. It is not unreasonable to assume that the uncertainty that might have arisen from running a perturbed parameter ensemble with the GHM would have been as large as the greatest source of uncertainty identified by the authors and this should be mentioned in the Discussion. This is often the case with climate modelling experiments that have run PPEs (e.g. see Rowlands et al. 2012 and Collins et al. 2006). To some extent, recent multi-GHM climate change impact assessments (Prudhomme et al. 2013; Davie et al. 2013; Hagemann et al. 2013) provide a glimpse of the range of uncertainty that can arise from different GHM structures, although strictly, GHM structure is not explored systematically in those cases.

3) This is related to 2) above. It is worth discussing that one of the reasons the simulations appeared less sensitive to land-cover than the other modifications, is because the parameters associated with each land-cover class were kept constant. The differences would probably have been much larger if land-cover-associated parameters were modified (e.g. LAI, interception capacity, surface roughness etc.), especially those related to PET.

Specific comments —————

1) The abstract includes a lot of information and it is rather long. I suggest reducing the amount of text included in the abstract. Moreover, lines 15-25 of the abstract could be shortened and it could be stated more concisely the order (increasing or decreasing) in which each uncertainty impacts on one chosen hydrological indicator (e.g. Q, or AET).

2) Lines 6-23, page 1592. Can the authors please clarify whether for both CLIMATE datasets they forced WaterGAP with monthly data, or daily data (disaggregated from monthly means).

C798

3) Line 10, page 1592, should "e.g. wind undercatch" actually be "e.g. precipitation undercatch"?

4) Line 6, page 1593, please write fully what IGBP stands for.

5) Figure 2 – it would be useful to include a third map that shows areas where there is different land-use according to the two sources of information. Grid cells where there is a difference could be shaded in a single colour. While this would be quite a simple map, it would make it easier to observe where the differences are.

References ————— Collins M et al. (2006) Towards quantifying uncertainty in transient climate change. *Climate Dynamics* 27: 127-147.

Davie JCS et al. (2013) Comparing projections of future changes in runoff and water resources from hydrological and ecosystem models in ISI-MIP. *Earth System Dynamics* 4: 359-374.

Hagemann S et al. (2013) Climate change impact on available water resources obtained using multiple global climate and hydrology models. *Earth System Dynamics* 4: 129:144.

Prudhomme C et al. (2013) Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proceedings of the National Academy of Sciences* 111: 3262-3267.

Rowlands DJ et al. (2012) Broad range of 2050 warming from an observationally constrained large climate model ensemble. *Nature Geoscience* 5: 256-260.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 11, 1583, 2014.

C799