Earth Syst. Sci. Data Discuss.,
doi:10.5194/essd-2016-55-RC1, 2017

Open Access Earth System Science Data Discussions

Interactive comment on "A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset" *by* Jaap Schellekens et al.

Anonymous Referee #1

Received and published: 14 January 2017

General Comments: This paper describes the eartH2Observe Tier-1 dataset and presents a benchmark for the key water cycle variables simulated by a suit of land surface models (LSMs) and global hydrological models (GHMs). The authors provide an overview of current state-of-the-art models and analysis framework along with tools that enables benchmarking repeatable as new improvements are made to the models and forcing datasets. I'm impressed and excited about the open access and completeness of the datasets and results of the paper. I think that the authors did great job designing the framework for identifying the model consistency/inconsistency via the use of common forcing and the SNR analysis, including both GHMs and LSMs and investigating the uncertainty in the precipitation forcing, and verifying with the benchmark

dataset. The paper contains valuable findings to the modeling community, in which the strength of ensemble mean over any single model is demonstrated in some variables and the areas of importance for further model dataset development are identified.

Specific Comments: -Page3, Line8: I was not sure if the second question "is the current modeling protocol with one forcing dataset and the selected output variables sufficient for evaluation of (global) water resources" was answered -"Continental water budget" is referring to water budget over land? I thought of individual continents (i.e. mean over North America, etc) but just global budget was presented. -Page13, Line 25: "the spread in ET is large" and that the model estimates are higher than the reference datasets are indeed concerning points. Only ORCHIDEE and WaterGAP3 include irrigation or water-use currently, but incorporating irrigation in other models will likely increase ET even more. -Is there a reason why you didn't use the snow cover from GLOBSNOW-2 and used IMS instead? -Table4: It states that the difference in model mean ET (and products?) are due to different periods used for the comparison. Do they match over the common overlapping period, 2003-2011? Additional information on spread of the three ET products can be helpful as a first cut uncertainty estimate, given that quality of ET validation datasets is difficult to assess. -Page15, Line10: "Although this may seem to be a large mismatch. . .." I don't see how this makes it more comparable. Could you elaborate? -Page19, Line 3 and top panel of Figure9: is the precipitation increase after 1997 evident in the reanalysis observation based datasets as well or has it been evaluated elsewhere? -Just curious, what does eartH2Observe stand for? Is there Tier-2 dataset (is it going to be the next round with error estimation and higher resolution etc.)?

Technical Corrections: -Page1, Line11: "at" -> remove -Page3, Line8: "modelling" -> typo -Page22, Line 25: "bets" -> typo -Figure 3 needs description on the line, box, and error bars. -Table 9 and Table 10 seem to be identical. I doubt that it is true since the global summary Table 6 shows different values for snow cover.