

## ***Interactive comment on “Validation of Terrestrial Water Storage Variations as Simulated by Different Global Numerical Models with GRACE Satellite Observations” by Liangjing Zhang et al.***

**Anonymous Referee #2**

Received and published: 25 September 2016

The manuscript gives an excellent overview on the current performance of four global hydrological models validated by using the latest GRACE Release. 31 river basins worldwide covering different climate zones are assessed and deeper insights into a few basins in arid and snow dominated zones are given. The study addresses current scientific issues and goes further than previous works. It is well structured and the results are presented in a clear and comprehensible way. I think that the study will contribute to the improvement of hydrological models.

In one point the paper could still be improved: especially in Section 3.2 and in Section 4 I miss some interpretation of the findings. Is it possible to discuss a few reasons for the different model behaviors?

C1

I have a few minor comments and questions which do not include the issues already discussed by the first reviewer.

Specific comments

Page 1, Line 7: What is the meaning of ‘different spatial characteristics’ of the individual storage compartments? Do you mean that e.g. groundwater is simulated using a different number of layers?

Page 2, Line 20: I think that an accuracy of 1 cm equivalent water height is a very optimistic estimate for areas as small as 100,000 km<sup>2</sup>.

Page 2, Line 24: You say, that there are more than 13 years of GRACE data available, but Figure 7 indicates that you use only 10 years of data instead of the full record. Why do you not use the whole time series of GRACE data? Is the data from the models missing?

Page 3, Data sets: For WGHM you explicitly mention the water storage compartments. Please add information about the storage compartments for the other models (to some extent you already did this in your response to the first reviewer).

Page 4, Line 27: Are the local re-scaling factors introduced for each grid cell?

Page 4, TWS Estimates from GRACE: Did you also remove the trend from the GRACE time series for the same period as for the models?

Page 5, Evaluation metrics: The relative annual amplitude differences and the phase differences are interesting metrics for river basins with a strong annual cycle that follow approximately a sine curve. Can you discuss the meaning of these metrics for river basins where interannual signals dominate (also with respect to Fig. 2 and Fig. 3)?

Page 5, Line 16: Why did you estimate the trend? You subtracted it in the preprocessing step.

Page 5, Global evaluation: This paragraph gives an excellent overview on the current

C2

performance of the four models on a global scale. Is it possible to add some interpretation of the results, i.e. can you explain the results by the model structure or the parametrization? E.g. why do the models in general have an earlier seasonal storage maximum than GRACE, why does LSDM have smaller phase differences than the other models, what is the problem of the models in basins with small explained variance, . . . ?

Page 6, Line 24: Did you mean: 'most basins have low SNR values'?

Page 7, Line 8: Fig.7 shows Amazon, Zaire, Mekong, and Niger instead of Chari, Indus, Murray, and Niger. Is this intended?

Page 7, Line 9: I do not think that the performances of the models at those four basins are quite consistent with each other. In fact, JSBACH performs quite differently.

Page 9, Line 7-14: This is a good summary of the performance of the four models. Can you provide any reasons for the strengths and the deficiencies of the individual models?

Fig. 6: When the annual signal is removed the boxes become much smaller for all models except for LSDM. Do you have an explanation for this behavior?

#### Technical comments

Page 2, Line 12-16: Did you want to say that instead of using observations of precipitation (P), evapotranspiration (E), and runoff for closing the water budget equation, one can also use water vapor content and moisture flux convergence from atmospheric reanalysis data and river discharge? Is runoff and river discharge the same in both cases? Furthermore, P and E can also be taken from atmospheric reanalysis (Rodell et al. (2011) Estimating evapotranspiration using an observation based terrestrial water budget. *Hydrological Processes*, 25:4082–4092.) Maybe you would like to reformulate the sentence.

Page 9, Line 24-27: The structure of the sentence is strange. Probably you should delete either 'In future' or 'in our next step'.

C3

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-330, 2016.

C4