Hydrology and
Earth System
Sciences

Open Access

EGU

Discussions

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 #1

Received and published: 30 August 2016

This manuscript presents results of comparing estimated terrestrial water storage (TWS) from four hydrological models with GRACE derived TWS in 31 hydrological basins. Four metrics were used in evaluating model performances. Components of TWS as well as actual and potential ET were examined in selected basins to show the impact of model physics on estimated TWS.

The results and discussions are generally well presented and justified. But I think the paper can be further improved in a few areas. For instance, the fact that three of the four models do not model groundwater, which may contribute significantly to TWS changes, is not explicitly mentioned and discussed in the paper. In addition, the four metrics used in evaluation may be good for summarizing the differences but they do

not necessarily reflect the actual discrepancies between modeled and GRACE derived TWS. For instance, the amplitude and phase differences may not be important if TWS exhibits strong inter-annual variability.

Additional comments: Page 3, data set. Please emphasize the fact that three of the four models do not simulate groundwater and discuss its potential impacts on model estimated TWS in the result section. Also, do these models account for anthropogenic impacts such as groundwater abstraction? If not, how would this affect the comparison with TWS from GRACE which does detect changes associated with groundwater withdrawals?

Page 4, Line 8: I understand why you removed the trend but the ability to predict trend is also an important part of the models. Can you provide a scatter plot comparing trends from the models and those from GRACE in the 31 basins?

Page 5 Line 10, should the second "GRACE" be TWS?

Page 6. Line 20, I don't think it is appropriate to compare GRACE errors with the RMSE since the former represents instrument and post-processing errors and has nothing to do with how well models perform. In addition, basin-scale GRACE errors are smaller than the gridded errors which are spatially correlated (http://grace.jpl.nasa.gov/data/get-data/monthly-mass-grids-land/). Did you consider spatial correlation of errors (in both modeled and GRACE TWS) when calculating basin-scale RMSEs between the model and GRACE? Either way, I think it only makes sense to compare RMSEs among the models. The statistics in Table 2 shows the models generally did not performed well in the tropical climate. Why is that? Does it have something to do with runoff estimates as ET is energy limited in this type of climate? You don't necessarily need to collect in situ stream flow data, but some discussions and plots on runoff may be needed to explain this result.

Page 7. Line 5 to 12. As you pointed out that AET does not have a significant impact on TWS in humid areas, then what is the purpose of including three basins from that

climate in Fig. 7? I think including more basins from drier climate is more useful here.

Page 7. Line 13-22, I didn't learn anything from this paragraph and it can be removed. As you correctly pointed out that AET may be significantly different from PET which does not help much in explaining the result. Again, I think presenting runoff estimates is more useful.

Figs. 2-5: It would be very helpful if you provide time series of TWS for a basin(s) with the largest deviation from GRACE in either of these metrics. For instance, it's hard to visualize how significant a 45 degree difference in phase is. In addition, two of these metrics measure differences in seasonality which may not mean much when the interannual variability of TWS is much stronger. So providing actual TWS time series along with some discussions will be helpful for readers to understand the usefulness and limitation of these metrics.

Fig. 7. I think including runoff instead of PET is more appropriate here. Also, please try to use the same y-axis range for all plots which makes it to compare the magnitude of TWS and ET.

Fig. 8, "Subsurface water" should be soil moisture + groundwater storage for WGHM and soil moisture for all other models. Again, please use the same range for all y-axis if possible. In the caption, snow water content should be snow water equivalent.