**Hydrology and Earth System Sciences Discussions**

# *Interactive comment on* "Combining satellite radar altimetry, SAR surface soil moisture and GRACE total storage changes for model calibration and validation in a large ungauged catchment" *by* C. Milzow et al.

**Anonymous Referee #2**

Journal: Hydrology and Earth System Sciences (HESS)

Title: Combining satellite radar altimetry, SAR surface soil moisture and GRACE total storage changes for model calibration and validation in a large ungauged catchment

Author(s): C. Milzow, P. E. Krogh, and P. Bauer-Gottwein MS No.: hess-2010-347 MS Type: Research Article

Suggestion: MAJOR REVISIONS

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

This paper discusses the calibration and (mainly) validation of SWAT simulations over a large catchment in South-Africa. The model is forced with different precipitation data sets and the output is compared against SAR-based soil moisture, radar altimetry-based water stages and GRACE water storage.

GENERAL:

The strength of the paper lays in the study of several remotely sensed data products over a relatively limited monitored area, where remote sensing data are of crucial importance. However, the main weakness of the paper is the lack of 'truth' and the consequent hypothetical character of all results. The paper should be made more quantitative, instead of qualitative in order to meet publication standards. The following comments could help to improve the relevance of this research. Also, the structure of the paper is unconventional: a lot of data description is included in the introduction. Consequently, the introduction lacks a clear overview of applications of the used data or studies over this catchment in the past, and the introduction is very weak in indicating what is learnt from this research. Finally, the paper can be greatly condensed by adopting a more concise writing style.

At this moment, the paper is too weak and only very few (if any) lessons can be learnt from it; therefore, a major revision is suggested.

SPECIFIC:

a) + Abstract: "SMM data are compared. . . They cannot be used. . .but. . . bias identification in the precipitation data" + 4.3: Pre-processing of the SMM data + Manual calibration + overall findings

- I don't see where/how in the paper this bias identification is quantitatively made.

- The precipitation data are all at different resolutions in space (and time). No spatial aggregation/downscaling is discussed, so a lot of "bias" is scale dependent and may not be present when scales are taken care of in a more sophisticated way. As of now,

Interactive
Comment

there is no way to be sure of which precip product is biased and which not. Also, it would be good to have an estimate of the average HRU size to see how the resolution of the precipitation compares against that of the model resolution.

- The approximate size of the HRU is also important to compare simulations against the SMM data. The latter are 1km resolution data: are they simply all averaged over the catchment, and SMM data as well in the comparisons or are smaller scale features being looked at?

- How is the model exactly calibrated? Pure visually? At this time in research, I think a multi-objective *automated* calibration, using the multiple data sources and maybe using different measures of goodness of fit should be used: it would greatly enhance the relevance of this study and is needed to get this work published, but will require a lot of work (hence the major revision). There is published work out that describes thorough SWAT calibration studies, so it is perfectly possible to select a few sensitive parameters and perform automatic calibration, using a split sample of the available RS data. (note that without having an 'unbiased' model structure and parameters, it will also be too hypothetical to point out which forcings are biased.)

- You may want to step away from looking into the bias, if multiple products are used and if the model shows a different level of climatology: each product or model has its own 'climatology'. You don't want one product (at a specific scale) to exactly resemble another one (at another scale). One way to deal with this would be to look into anomalies. Since you have a long time series, it is worth it to try this. This would also circumvent having to deal with the above comments on bias identification.

In short: it is necessary to better identify bias at the correct scales for the different products (RS and model), as well as its sources, or step away from the bias issue and look into anomalies.

b) All results and the discussion are very much limited to visual comparison of time series graphs. It is really necessary to introduce a more quantitative discussion by

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

presenting measures like RMSE, correlation, regression slope between observed and simulated (or different simulated curves).

DETAILS:

- Title (and abstract/conclusions. . .) "Combining. . .": there is no combination or merging of data in this study, except maybe in the "manual calibration", but that procedure is not really technically sound, and hence not worth it mentioning in the title. It is rather a "Comparing of. . . versus model simulations".

- Abstract: acronym SAR first time in full

- p.9125, l.11: "The purpose of the model developed in this study. . .assess the impact of agricultural development on runoff. . .": If that is the purpose, why isn't that in any way discussed in the paper? No case is made on testing anything concerning agricultural development. Which input data exactly reflect that agricultural development, how is it used and do you see a difference if you use this dynamic input versus using a static input on land cover and other agricultural aspects. (same goes for p.9128, l.26: agricultural intensification. . .: which data are exactly referred to, what is input in the model?)

- p.9126, l.13: catchment outflow is thus measured? In that case, your catchment is not ungauged.

- p. 9129, l.9: acronym SCS - first time full

- p. 9129, l.15: Hargreave's formula: explain which data *are* used in this formula, rather than telling which data are not used (and not available).

- p.9130: l.6-10: we don't need that background info, just tell us you used 7 subbasins, because you had no data to justify more. Are parameters calibrated for each of those 7 basins individually? How? Please indicate the final model resolution.

- p.9135: please indicate which measure of goodness of fit is used for calibration,

which parameters are optimized, at which resolution, etc. . .; also, I would suggest an automated calibration.

- p. 9136: a sensitivity study is done. NEXT, one would expect to see a discussion of the calibration of the most sensitive parameters. Please re-organize and list the values of your calibrated parameters.

- Section 5: this whole section does not look too good, because the results tell that the model is not performing so well. You could use part of your 'validation' data for calibration and use the other part for validation to strengthen the paper.

- Figure 2: It is confusing to see the word "Annual" with precipitation/runoff as well as just 2 months (Aug-Jul and Nov-Oct). Please make sure you have the exact same overlap of months in both.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 9123, 2010.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper