Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-534-RC4, 2019

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Data assimilation for continuous global assessment of severe conditions over terrestrial surfaces" *by* Clément Albergel et al.

Anonymous Referee #4

Received and published: 7 December 2019

The paper primarily discusses the verification of output from the global LDAS-Monda land surface data assimilation system for 2010-2018 at 0.25-degree spatial resolution. Satellite observations of surface soil moisture (SSM) from ASCAT and leaf area index (LAI) from GEOV1 are assimilated into the ISBA land surface model. The verification assesses the skill of the assimilation estimates against the assimilated observations and against independent in situ, model and satellite datasets. The skill of the assimilation estimates is compared to that of a model-only "open-loop" simulation (without data assimilation). Moreover, higher-resolution simulations with and without data assimaltion along with 4-day and 8-day forecasts of land surface conditions are presented for 2017-2018 over two regional domains, and the output was assessed in the context

of extreme land surface conditions. The authors find that the assimilation generally improves over the open-loop simulation, particularly for the more slowly-varying LAI estimates.

The manuscript is obviously of interest to HESS readers. I'm confident that the data assimilation system, the simulations, and the verification results are technically sound. However, I found the presentation of the material to be somewhat lacking. The paper is a bit tedious to read in that a number of results are discussed at length where the assimilation and the open-loop are essentially the same (e.g., for snow variables) or where the improvements from the assimilation are probably not meaningful, at least not in the sense of statistical significance. Perhaps worse still, there is no quantitative information about errors in the assimilation estimates, e.g, confidence intervals or other suitable uncertainty estimates. In my view, the discussion also over-emphasizes the comparison of the assimilation estimates against the assimilated observations, which is not independent verification (as the authors point out clearly in one part of the paper but mostly sweep under the carpet in another part). Finally, I was rather disappointed, given the illustrious author list, with the fairly careless pre-submission proof-reading.

I think the paper can be an important contribution and can eventually be suitable for publication in HESS, but at this point I recommend MAJOR revisions with consideration of the comments below.

Major comments:

1) Six of the thirteen figures that show results (i.e., not counting "data & methods" figures 1 and 2) are about evaluating the skill of the assimilation estimates *exclusively* against the assimilated observations (Figs 3 and 9-13). Comparisons against the assimilated observations are also included in Figs 4-6 (along with other variables) and Figs 14-15 (along with forecast estimates of SSM and LAI). While I agree that it is important to verify that the assimilation system works as intended, the authors over-emphasize the comparison against the assimilated observations.

2) The two figures about snow (Figs 7 & 8) could be simplified considerably because there is no meaningful difference between the assimilation estimates and the open-loop estimates, which is a rather trivial result (as the authors discuss).

3) There are no graphics in the main text (only in the supplement) about the validation of the results against *independent* in situ measurements (section 3.1.2). This independent validation should be reflected more prominently in the main paper.

4) The claim about "improvement" of the assimilation estimates vs. the open-loop estimates from the independent validation against in situ soil moisture estimates in section 3.1.2 (∼line 460) is on shaky footing. For none of the networks listed in Table S3 is there a difference of more then 0.02 in the R values between the assimilation and the open loop. In some cases, the 0.02 difference is negative (ie., degradation). For most networks the R difference is 0 or 0.01, that is, there really isn't a meaningful change. Here, and also for at least the other in-situ based results, it is imperative that the authors provide some estimates of whether the differences are meaningful (e.g., by including statistical confidence intervals), and then honestly discuss the results. The claim in line 460 about significant improvements at some sites may be true, but given the network-average neutral results there must then also be sites with a significant degradation, which is not mentioned in the paper.

5) The editing of the paper is rather careless. There are many small mistakes, and the organization of the text is lacking. Examples include the following:

a) The Introduction lacks a clear statement of the paper's objectives. The text in Lines 107-121 simply states what will be presented (with lots of references and details). It's hard to tell what the objectives might be.

b) There are several instances in the Results section of text that belongs in the Methods section, incl: Lines 384-387 - IMS snow cover product description Lines 405-409 - Fluxnet description Lines 440-447 - ISMN description

c) Section 3.2.2 is a *single* paragraph that stretches over nearly two pages. Really? There are several other paragraphs of excessive length.

d) Graphics:

Figure 1a: Use different color for zero values and no-data value. (currently, both are white, making it unclear whether there are data in, e.g., the western US, or whether those are screened, perhaps because of topography.

Figure 3: The label of the colorbar should read "RMSD of LAI [m2 m-2]", not just "LAI [m2 m-2]"

Figure 5: Units are missing for RMSD panels. (This is particularly important because this information is needed to judge whether the differences are in fact meaningful.)

Figure 6: Three panels only have a single tick & tick label on the y-axis. At least two are required to interpret the axis scale.

Figure 7: The color choices should be made consistent with Fig 4.

Figure 9: I could not find out what the thin cyan lines depict.

Figures 10+11: add "LAI" to plot title of c) and d); add "SSM" to plot title of g) and h)

Figure S2: NSE should vary from -infinity to 1. The colorbar is from -20 to 20, and darker blue values would clearly be greater than 1. Either the colorbar is wrong or the values show something other than NSE.

Table S3: The column headings on the 2nd page of the table still include French words.

6) In section 3.2.2, the authors no longer make it clear that the verification is against the assimilated datasets. While verification of forecast data against the assimilated dataset can be viewed as independent validation because the verification data have not (yet) been assimilated, there is an important distinction here between SSM and LAI. For SSM, the assimilation is done after rescaling (cdf-matching), which removes bias. For

LAI, however, the assimilation uses the raw LAI observations (I think). That is, the assimilation removes bias in the modeled LAI (w.r.t. the observed LAI). This technical difference between SSM and LAI assimilation, combined with the longer memory of LAI compared to SSM, should contribute to the results in section 3.1.2. Put differently, the LAI results of section 3.1.2 are not likely to hold if an independent LAI dataset had been used for validation that is itself biased against the assimilated LAI observations. (Different LAI datasets may not be as biased against each other as typical satellite SSM datasets, but there are considerable biases between LAI products.)

7) Figure 3c suggests that the change in GPP is negligible, at least in the zonal mean sense although Figure 4f suggests that GPP does change in terms of RMSD. Given the considerable change in the (zonal mean) LAI (Fig 3a), I would have expected a lot more change in the mean GPP. I suspect that the disconnect between the LAI and GPP changes is rooted in how these variables are connected in ISBA and how exactly the assimilation system goes about updating LAI. This rather counter-intuitive result requires clarification in the paper.

8) Fig 5h: The changes in EVAP are with +/- 0.02 (mm/d???). If my guess about the units is correct, this would amount to only a few mm per year, which is well within the uncertainty of in situ measurements. That is, the EVAP changes are not likely to be meaningful in a practical sense. This should be discussed more explicitly.

Minor comments:

9) Line 167: typo "bale" –> "able"

10) Line 209: "fifth generation of European reanalyses produced by ECMWF" I recommend phrasing this differently to avoid the misunderstanding that the reanalyses are just for the European domain. E.g.,: "fifth generation of global reanalyses produced by ECWMF"

11) Lines 293-295: How did you address the heterogeneity within the 0.25-deg grid

cells during spin-up? It is not obvious that the short spin-up period from April 2016 suffices for properly spinning up grid cells with strong heterogeneity at the sub-0.25-degree scale.

12) Line 379: Do you mean a decrease in RMSD or a decrease in skill?

13) Line 412: If I'm reading this correctly RMSD decreases while both bias and ubRMSD increase. This is quite counter-intuitive and requires a rather odd distribution of the metrics across the sites or networks included in the average. In any case, since bias and ubRMSD get worse, I do not think that the statement about "a small advantage of the analysis over the open-loop" is justified.

14) Line 429: "NSE values below -2 were discarded" requires a justification, otherwise it reads like cherry-picking.

15) Line 535: "the analysis is of better quality" Given the numbers, I see at best "slightly better quality"

16) Line 592: "surface (0-1 cm)" In section 3.2.2 the discussion was about the "(1-4 cm)" layer. Which is it?

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-534, 2019.