

Interactive comment on “Evaluating the surface energy partitioning in ERA5” by Brecht Martens et al.

Anonymous Referee #1

Received and published: 11 March 2020

The paper investigates the skill of the land surface energy partitioning in the ERA5 reanalysis. For reference, the ERA5 skill is compared to that of its predecessor, ERA-Interim (ERA-I). Skill is determined in several different ways: (1) directly vs flux tower measurements, (2) vs fluxes from water and energy balance estimates, (3) by driving the GLEAM land surface scheme with reanalysis data and validating the output vs. flux tower measurements, and, finally, (4) by driving the CLASS4GL boundary layer model with reanalysis data and validating the output vs. balloon observations.

The authors find that ERA5 land surface energy partitioning is generally improved over that of ERA-I. In particular, the overestimation of the latent heat flux in ERA-I is reduced (but not eliminated) in ERA5.

The paper is of interest to GMD readers and makes an important contribution by doc-
C1

umenting the quality of ERA5 land surface estimates. By and large, the writing and graphics are clear and concise, and the conclusions drawn from the study are supported by the results. I recommend eventual publication of the paper in GMD provided the authors address the comments below. It would be particularly helpful to include other reanalysis data in the comparison.

Major comments (no particular order):

1) The title and introduction do not make it sufficiently clear that the turbulent fluxes investigated here are for the land only. I suggest changing the title to: "Evaluating the *land* surface energy partitioning in ERA5" and occasionally replacing the expression "surface [...] fluxes" with "land surface [...] fluxes", e.g., P3/L15, P4/L3, and probably a few more places.

2) P4/L1: Aren't there advances in land data assimilation from ERA-I to ERA5? And do these not matter for the quality of the land surface turbulent flux estimates? Land data assimilation in ERA-I and ERA5 and its impact on land surface estimates should be included briefly in the Introduction and further discussed in the Results and Discussion section.

3) Section 2 is a somewhat odd mix of "methods" and "data". E.g., the FLUXNET 2015 section 2.2 includes discussion of how the climatology is derived for the computation of the standardised anomalies, but this would also apply to the other datasets (incl. ERA-I and ERA5). The entire section needs to be reorganized and be more clearly separated into "Data" and "Methods".

4) The violin plots are great visual tools, but I assume their construction involves is some fitting of the distributions. These details should be in the "Methods" section.

5) There are gaps in the literature discussion. E.g., Draper et al (2018) is a highly relevant assessment of reanalysis estimates of land surface energy flux estimates, incl ERA-I. Draper, C. S., R. H. Reichle, and R. D. Koster (2018), Assessment of

MERRA-2 Land Surface Energy Flux Estimates, Journal of Climate, 31, 671-691, doi:10.1175/JCLI-D-17-0121.1.

6) Related to 5), the present paper only investigates ERA-I and ERA5. The paper would be considerably more relevant to readers if these ECMWF products were assessed along with at least one or two other, major reanalysis products (such as MERRA-2).

7) If I understood this correctly, the GLEAM and CLASS4GL analyses work as follows: (i) use ERA-I and, separately, ERA5 to force GLEAM (or CLASS4GL), then (ii) evaluate the results against tower (or balloon) measurements. This approach depends on the assumption that GLEAM and CLASS4GL are very good models, or at least that they do not have compensating errors. If, say, errors in GLEAM were to compensate for errors in ERA-I, forcing GLEAM with a better reanalysis isn't necessarily going to deliver better GLEAM outputs. Similarly for CLASS4GL. This is a major caveat that needs to be discussed prominently in the paper.

8) Figures A.1-A.4 are oddly placed. Either they need to be put into a proper Appendix, with discussion in the Appendix as well, or they need to be placed in a separate "Supplementary Information" document. As assembled, Figures A.1-A.4 are simply out of order.

9) Figures 2 and 4 are a confusing mix of dimensional metrics (MD for raw fluxes) and unit-less metrics (MD for raw Bowen ratio, and MAD and R for standardised anomalies). These figures also lack basic information about a dimensional 2nd-order metric such as MAD or RMSE for dimensional (raw or anomaly) variables. The readers are going to want to know typical MAD or RMSE values for fluxes in units of W/m², incl. and/or excl. the seasonal cycle.

I suggest revisiting the assembly of Figures 2, 4, A.1-A.4. The paper would be much easier to follow and more informative if, say, one figure includes only (dimensional) metrics computed from raw time series and another figure includes only (non-dimensional) metrics from standardised anomaly time series.

C3

10) P10/L15-16: "Figure 4b shows that ERA5 is better at capturing..." Doesn't this invalidate the conclusion drawn from Fig 4a, that is, the evaluation of ERA-I and ERA5 through GLEAM? See also comment 7) above.

11) P11/L3-8: One aspect that may come into play here is that GLEAM+ERA5 is an off-line (land-only) modeling system that does not permit feedback whereas ERA5 is a coupled land-atmosphere modeling system. This may be related to a finding by Draper et al (2018):

"Finally, the SH results for MERRA-Land are troubling. While MERRA-Land did have the desired reduction in the LH biases compared to MERRA (to 1 W/m² in the global land annual average), it also had a compensating, and much larger, increase in the SH bias (up to 15W m² in the global land average)" [beginning of p689 of Draper et al. 2018]

See also comment 5) above about the need for a better integration of the results of the present paper into the literature context.

12) P11/L26-P12/ L2: There is no discussion of Fig 9! In this paragraph, insert explicit references to Figs 7, 8, and 9 in the relevant place within the paragraph. E.g., reference Fig 7 in P11/L28, reference Fig 8 in P11/L33. This reveals that Fig 9 has not been discussed.

13) In some cases the results are overstated.

E.g., P12/L22: "The improvements are less clear..." suggests that there are some (hard-to-see) improvements, when in fact the results are neutral at best.

P14/L3-4: The statement here is not consistent with the results of Fig 4 that show that ERA-I estimates of the sensible heat flux and Bowen ratio are better than those of ERA5.

14) The tower validation results should come with some measure of statistical significance or error bars. Are the improvements, that is, the small shifts in the distributions

C4

of the metrics as shown in the violin plots, meaningful?

Minor comments:

a) P2/L13: Reichle et al. 2017 is a reference primarily for MERRA-2 land surface estimates. MERRA-2 which is a full atmospheric reanalysis, similar to ERA-I and ERA5. In this place, however, the authors are here referring to land-only reanalyses, such as ERA-Interim/Land and MERRA-Land. For the latter, Reichle et al. (2011) is a better reference. Note that there is **not** a land-only reanalysis associated with MERRA-2.

Reichle, R. H., et al. (2011), Assessment and enhancement of MERRA land surface hydrology estimates, *Journal of Climate*, 24, 6322-6338, doi:10.1175/JCLI-D-10-05033.1.

b) P2/L26-30: The text here is about very geographically limited results (southern Antarctic peninsula) or sea-ice, which is not the focus of the present paper. I suggest deleting this text or moving it further down. It confuses the reader by distracting from the focus of the paper on the global land surface turbulent fluxes.

c) P4/L19: On first reading, I completely missed the term "standardised" here and later got confused about the lack of dimensions/units in the graphics. In many papers, anomalies from the seasonal cycle are examined without standardisation, i.e., they are dimensional anomalies. There is nothing wrong per se with the standardised anomalies, but please make it clearer that you are focusing on dimensionless anomalies.

d) P10/L31: typo: "we should emphasis" → "we should emphasise"

e) P12/L16+L20: The numbers referenced here contradict the numbers in the graphic.

f) Caption Fig 2: Replace "For MD, the distribution of beta..." with "The distribution of the MD of beta..."???

g) Caption Fig 7: "...between the absolute bias in ERA5 and ERA-I;"??? Maybe I'm misunderstanding this, but I think the bottom panel shows $\text{abs}(\text{bias}(\text{ERA-I}))$ minus

C5

$\text{abs}(\text{bias}(\text{ERA5}))$, that is, the sign of the abs bias difference is different from what the caption suggests.

h) Caption Fig 10: "...versus ERA-I (squares)"??? Should this read "...versus ERA-I (triangles)"???

i) Figure 11 needs units on the colorbars. I also suggest making the graphic bigger so it can be read more clearly in a hard copy for the next round of reviews.

Interactive comment on *Geosci. Model Dev. Discuss.*, <https://doi.org/10.5194/gmd-2019-315>, 2020.

C6