RC2

Received and published: 24 March 2020

General comments:

This is a very interesting and useful study of the representation of the land surface energy budget in European global atmospheric reanalyses. A large number of diverse in situ observations are used to benchmark several simulations at a global scale. Overall, the paper is well written, apart from the mixing of results and discussion/interpretation. Quality of some Figures could be improved. Colour scales are sometimes confusing as "green" tends to look blue. Violin plots are useful but do not provide a point by point comparison. Could all the corresponding scatter plots be given in a Supplement? A discussion on the impact of land cover is lacking. Recommendation: minor revisions.

We would like to thank the referee for reviewing the paper and giving some interesting comments and feedback. Below, we give a point-to-point reply to the comments posted by the reviewer and list the changes that will be implemented in the manuscript.

Regarding the scatterplots, the authors believe that the violin plots together with the discussion of the results should give a sufficiently detailed understanding of the results and prefer not to include these addition figures.

Particular comments:

1. P. 1, Title: should be more specific. For example: "Evaluating the land surface energy partitioning in European global atmospheric reanalyses".

   We believe the reviewer means 'less specific'. The authors prefer to keep the title as is, as we want to emphasise that the focus of the paper is on the evaluation of the newest state-of-the art ERA5 reanalysis, rather than European reanalyses in general. Note that ERA-Interim only serves as a benchmark here to show the improvements. We believe mentioning the model/dataset in the title is in addition a requirement at GMD.

   <u>Changes in manuscript:</u> no changes

2. P. 3, L. 1-2 ("perform better than ERA5"): any reference on this?

   As this sentence builds upon the previous sentence, the reference supporting this statement is Urraca et al. (2018).

   <u>Changes in manuscript:</u> the reference will be cited again to make this clear.

3. P. 4, L.7: I would be more specific. For example: "the more evolved HTESSEL land surface model in ERA5".

   We agree with the referee that this sentence should be updated.

   <u>Changes in manuscript:</u> the sentence will be updated per suggestion of the reviewer.

4. P. 4, L. 19: could you explain how these anomalies are defined and calculated?

   The standardised anomalies are simply calculated by subtracting from the raw time series (1) the climatological expectation (i.e. the mean of the variable under consideration over at least 5 years for a certain time step) and (2) dividing by the standard deviation of that climatological expectation.

Changes in manuscript: a brief description of this procedure will be included in the revised version of the manuscript.

5. P. 5, L. 1-3: It seems that a key issue was not addressed. Land cover type in ERA5 may not correspond to the tower's one. E.g. a grassland Fluxnet site may be located in an ERA5 grid cell mainly covered by forests. How did you handle this?
   We agree with the referee that this is an issue in the in situ evaluation strategy, as it is always the case in comparisons of grid cell values to in situ data. The mismatch in spatial footprint between the in situ measurement and the grid cell of the models typically leads to an overestimation of the actual error, often referred to as the representativeness error. In this study, we did not apply any filtering to maximise the representativeness of the in situ measurement – and hence to minimise this representativeness error – for the grid cell of the model. However, we do agree that this issue should be explicitly mentioned in the discussion of the results.

   Changes in manuscript: this issue will be highlighted in the discussion of the results.

6. P. 6, L. 5: could you define "non-overlapping moving windows"?
   Non-overlapping (moving average) windows refer to the fact that the time windows used for calculating the averaged quantities do not intersect (e.g. Dehghani et al., 2019). Moving windows are commonly calculated for all data points of a time series, i.e. for a simple example with a window length of 5 months, the centered moving average (window) of March contains data from January, February, March, April and May, whereas the moving average centered on April is based on nearly the same data, except ranging from February to June (i.e., adjacent moving averages share some data, and are thus 'overlapping'). In this given example, the 'next' non-overlapping window would be centered on August (June–October), as the time window used for the centered average of March and this time window do not intersect.

   Changes in manuscript: this will be clarified in the text to make this more clear.

7. P. 6, L. 12 (G as a fixed fraction of Rn): Could this explain the poor scores obtained for sensible heat flux in Figure 8? The soil heat flux is related to soil properties and can be influenced by sensible heat exchange with rainwater (e.g. Zhang et al. https://doi.org/10.5194/acp-19-5005-2019).
   As described in the manuscript at page 12 lines 3–7, the results for the sensible heat flux should indeed be interpreted with care as they are (among others) affected by this specific assumption. However, the magnitude of the ground heat flux at daily scales is often substantially smaller than that of the other fluxes, so the authors expect only a minor impact of this assumption on the analyses.
   In addition, the approximation used in this study to calculate the ground heat flux is not that uncommon as the ground heat flux is typically strongly correlated to net radiation (see e.g. Kustas and Daughtry, 1990; Santanello and Friedl, 2003). Also note that, although we do not explicitly account for the effect of soil properties on the ground heat flux, we do account for the land cover type, as described in Miralles et al. (2011) and Martens et al. (2017) and the response to comment #8.

   Changes in manuscript: the calculation of the ground heat flux will be described in more detail in the revised version of the manuscript and references will be added.

8. P. 6, L. 12 ("land cover"): which land cover? Is it the land cover used in the model?
   The ground heat flux is calculated in this paper as described in Miralles et al. (2011) and Martens et al. (2017). In essence, the ground heat flux is calculated as a fixed fraction of net radiation, depending on the sub-pixel land cover heterogeneity. The latter is parameterised by the MOD44B Vegetation Continuous Fields product, describing each grid cell as a fraction of tall vegetation (e.g. forests), low vegetation (e.g. grasslands), and bare soil. For the fraction of tall vegetation, the ground heat flux is 10% of the net radiation, while for the fractions of low vegetation and bare soil the corresponding percentages are 20% and 35% (Miralles et al., 2011). In the end, the fraction of net radiation assumed to be converted into the ground heat flux is the weighted average of the former percentages with the fractional land covers.

   <u>Changes in manuscript:</u> the calculation of the ground heat flux will be described in more detail in the revised version of the manuscript and references will be added.

9. P. 8, L. 3: Is there an impact of the land cover type?
   The authors have tried to relate improvements/degradations from ERA-Interim to ERA5 to different ancillary data sets like land cover, elevation, and climate, but no conclusive results were obtained. Given the uncertainties in such an analysis, the authors have chosen not to further discuss these results.

   <u>Changes in manuscript:</u> no changes

10. P. 8, L. 17: Seasonality removal should be described is chapter 2.
    As replied to comment #4, the procedure will be described in more detail in the revised version of the paper.

    <u>Changes in manuscript:</u> a brief description of this procedure will be included in the revised version of the manuscript.

11. P. 8, L. 28: what about Fluxnet site distribution in terms of vegetation types?
    As replied to comment #8, the FLUXNET sites are indeed not well-distributed across vegetation types and climate.

    <u>Changes in manuscript:</u> this issue will be further emphasised in the revised version of the paper.

Editorial comments (Figures):

1. Figure 1: Sites cannot be easily spotted. Colors of dots and background should be changed. What about land cover types? Format of the subfigure on the right should be consistent with format of Figure 3.
   We agree with the reviewer that the details on the figure are hard to read.

   <u>Changes in manuscript:</u> the figure will be updated to improve the readability.

2. Figure 6: green or blue?
   We agree with the reviewer that the colours might be confusing for some readers, but we think this is comment is purely linguistic and that the figures are clear.

Changes in manuscript: no changes

3. Figure 10 (top subfigures): meaning of the red lines? These metrics are a bit obscure. Why not comparing scatterplots of ABL heights?
As described in the caption, the solid lines represent the median and inter-quartile range (green for ERA5 and red for ERA-Interim). We appreciate the suggestion. The reason why diurnal changes in temperature, humidity and ABL height are compared – as opposed to afternoon temperature, humidity and ABL height – is to reduce the influence of errors in the morning initial conditions. This in addition increases the comparability of the results reported by Wouters et al. (2019).

Changes in manuscript: a legend will be added to the figure.

4. Figure 11: Not readable. Difference figures should be expanded. Green or blue?
The authors agree that the figures were too small.

Changes in manuscript: the size of the figure will be increased so it covers the entire two columns of the manuscript.

References:

1. Dehghani, A. et al.: A Quantitative Comparison of Overlapping and Non-Overlapping Sliding Windows for Human Activity Recognition Using Inertial Sensors. Sensors, 19, 5026, 2019.
2. Kustas, W. P. and Daughtry, C. S. T.: Estimation of the soil heat flux/net radiation ratio from spectral data, Agric. For. Meteorol., 49, 205–233, 1990.
3. Martens, B. et al.: GLEAM v3: Satellite-based land evaporation and root-zone soil moisture, Geosc. Model Dev., 10, 1903–1925, 2017.
4. Miralles, D.G. et al.: Global land-surface evaporation estimated from satellite-based observations, Hydrol. Earth Sys. Sc., 15, 453–369, 2011.
5. Santanello, J. A. and Friedl, M. A.: Diurnal Covariation in Soil Heat Flux and Net Radiation, J. Appl. Meteor., 42, 851–862, 2003.
6. Wouters H., et al.: Atmospheric boundary layer dynamics from balloon soundings worldwide: CLASS4GL v1.0, Geosc. Model Dev., 12, 2139–2153, 2019.