

Comment on **essd-2021-82**

Wouter Dorigo (Referee)

Referee comment on "ERA5-Land: a state-of-the-art global reanalysis dataset for land applications" by Joaquín Muñoz-Sabater et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-82-RC1>, 2021

This manuscript provides an excellent overview of the ERA5-Land dataset and its performance against several reference datasets. Due to its uniqueness and skill, I am convinced that this dataset will be heavily used and cited. Therefore, I recommend the manuscript publication after addressing a few issues.

- In line 6 of the abstract you mention that ERA5-Land enables the characterisation of trends. Yet, the manuscript does not include any analysis of trends. Why were trends not assessed, given that likely the product will be intensively used for trend analyses?
- The benefits of ERA5-Land wrt ERA5 are not ubiquitous. Clearly, the resolution is higher, but this doesn't always lead to better skill. To convince the user of using a ~10fold larger dataset and exponentially increasing subsequent computation times for further use, could you summarise at the end of the manuscript where, for which variables, and for what applications ERA5-Land shall be favoured over ERA5?
- I appreciate that the new product can be comprehensively assessed for all fields, but can you at least justify why these fields were evaluated and others not?
- Similarly, one needs to make choices regarding the reference data to be used. The choice seems slightly arbitrary and hence, the conclusions drawn can be inconsistent between fields. Especially, as ERA5-Land is embedded in the C3S Climate Data Store I would have expected that the assessment would draw on the various (satellite) products provided by this service. Vice versa, for LST you used a satellite product but you could also have used soil temperature data from the ISMN.
-

Minor issues:

- Line 227: please also cite the recently published paper on the ISMN: <https://doi.org/10.5194/hess-2021-2>
- Did you also apply the quality flags of the ISMN data?

<https://doi.org/10.2136/vzj2012.0097>

- Line 391: From the boxplots one cannot tell that ERA5-Land on average performs slightly better than ERA5, as the results for the different regions are based on a different number of stations. Does this statement still hold if you perform the analysis for all stations in these regions together?
- URL of theISMN: change into ismn.earth
- As this is a dataset paper, it would be useful to include a list and short description of the input and output fields
- For ET: briefly describe the evaluation metrics, even as the fields sshf, slhf (1492), and bowen ratio