



Comment on **essd-2021-181**

David Lawrence (Referee)

Referee comment on "A flux tower dataset tailored for land model evaluation" by Anna M. Ukkola et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-181-RC1>, 2021

This manuscript describes a set of 170 Tower Sites for use in land modeling. The set of sites have been quality-controlled and gap-filled, providing a consistent set of data that will be of wide utility to land modelers and other researchers. I commend the authors for their valuable work.

The paper is well-written, clear, and concise and as such I do not have much by way of critical comments. I find the manuscript to be acceptable for publication after minor revisions. This is/will be a valuable resource for the community.

Comments:

- Are there any plans to update the datasets going forward with new data as it is collected or with additional sites? The authors mention that other sites could potentially be included from the greater global Tower Site network, which partially answers this question, but it doesn't resolve the question about whether / how these sites will be updated as more data is collected.
- On the Flux datasets, there are variables with a `_uc` designation. The long name says that this is "Qh_cor joint uncertainty". I couldn't find anywhere in the text that explains what this variable is. A brief explanation in the text about what these variables are and how they were calculated would be useful.
- Also on the Flux datasets, a few variables have `_se` or `_DT` suffixes. The `se` refers to standard error, but I'm not sure exactly what error is being referred to .. measurement error or error induced by gap-filling or ??? Several of these variables also have long names that say something like "GPP of CO2 standard error (variable `ustar`, nighttime partitioning)". I'm not sure what is meant by variable `ustar` in this context. I can guess at nighttime partitioning, but it would be a guess. As above, some additional text in the main document explaining these variables would be helpful.
- Regarding the corrections for energy balance closure, I would like to direct your attention to this paper.

Swenson, S.C., D.M. Lawrence, and S. Burns, 2019: The Impact of Biomass Heat Storage on the Canopy Energy Balance and Atmospheric Stability in the Community Land Model. *JAMES*, doi.org/10.1029/2018MS001476.

In that paper, we found that when we included biomass heat storage in CLM, the need to correct the observed latent and sensible heat fluxes (at least during middle of day) seemed to go away (i.e., the model produced an LH+SH/Rnet ratio that was consistent with the uncorrected LH and SH), especially for forested sites. See in particular Figures 10 and 11 which show good agreement of model and obs with biomass heat storage against the uncorrected LH and SH.

Note that we weren't seeing such a big discrepancy between Rnet and LH+SH, at least compared to the numbers you state in your paper (average of 0.55 for LH+SH/Rnet+G). We were seeing an average of more like 0.75 (without the G term, which I think is generally pretty small). Anyway, you do mention unaccounted energy storage as one source of the lower LH+SH vs Rnet, but then it seems like you ignore the possibility that energy storage might be playing a role with the correction that you apply. The results from our paper suggest that biomass heat storage potentially shouldn't be ignored when making the correction, especially for forest sites. Anyway, this is a rambling comment and this really isn't my area of expertise, but I wanted to raise your attention to it. What to do is unclear to me, but maybe at least some additional discussion is warranted about the assumptions and implications and from the corrections. If the results in our paper are correct, and are being interpreted correctly, then including the correction could actually degrade the accuracy of the observed fluxes rather than improve them, leading modelers to target the wrong flux values.

- Line 30, page 4: "We used expert judgement to visually manually screen sites instead of an"; visually or manually or both?
- Do ANY of the sites report LAI (I thought some did, but not sure about that) and if they do, did you check them against the LAI from MODIS?