

Interactive comment on “Local tower-based merging of two land evaporation products” by Carlos Jiménez et al.

Anonymous Referee #3

Received and published: 28 December 2017

This manuscript describes work to combine two ET products, PT-JPL and GLEAM, using a weighted average, with weight determined by fit to tower observations. The resulting product is limited to the locations of the towers, and no attempt is made at extrapolation to other sites. While the manuscript is well written, and the analysis sound, the work itself is not well motivated and, as currently presented, does not represent a significant contribution.

The merged product presented in this manuscript does not add any value to the ET products that are already available. The motivation seems to be to merge the two ET products (PT-JPL and GLEAM) to produce a new product that is as close to the tower ET time series as possible. How is this new merged product then any more useful than the original tower ET time series? The merged ET product has not been independently

C1

evaluated. It is shown to be closer to the tower observations, but this is by design. Given that the tower observations also have errors (and given how closely the merged product has been fit to the tower obs), it does not follow that the merged product is necessarily more accurate. I am concerned that the merged product is over-fitting to the tower obs (weights calculated independently at each location, using a moving temporal window)

The work is not very well motivated. Why merge just these two products? Why not merge as many as are available, or as many as meet some pre-defined standard? The selection of just these two products is particularly awkward given that they are not independent.

To be publishable this work must i) provide a product that adds value in some way to the original products., and ii) the resulting data set must also be independently verified. The most obvious way to achieve this would be to spatially extrapolate the weightings. This could potentially provide a new product with (near-) global coverage that is more accurate than either of the original gridded ET data sets, and would also allow independent verification against withheld tower observations.

If this is not possible, I suggest that the manuscript be re-submitted and re-written (with additional discussion and conclusions) to focus on evaluating the GLEAM and PT-JPL products against tower obs.

MINOR COMMENTS:

Section 2: There is not enough information here for the reader to understand how the two products are calculated and what their main differences are. Please provide full details of the methodology of each product, rather than relying on previous work.

P5, L24: give the specific resolutions.

P8, L5: mention that the station coverage is not globally uniform, with nearly all stations in Europe and the US.

C2

P8, L20: ‘corrected fluxes are preferred’. Provide citation. Also, for the results provided in this paragraph for the corrected fluxes, how were they corrected?

Equation 1: add a sentence to describe what this metric is measuring (something like “the first term is the mismatch between the land cover at the tower and at the grid cell level, and the remaining terms are the net mismatch in land cover types across the two resolutions”).

P14, paragraph from line 10: the text here implies that the motivation is to match the tower ET as closely as possible, but the tower ET will also include errors. This paragraph should be re-written to acknowledge that the tower ET will also include errors (and the methodology perhaps adjusted to not over-fit to the ET data)

P15, L10. The use of the full seasonal cycle concerns me. In general, different ET products agree reasonably well in terms of the seasonal cycle (Jimenez et al. 2011; Mueller et al. 2011; Miralles et al. 2011). It is the anomalies that have more disagreement, and should then be the focus of efforts to improve / combine ET products. Also, using anomalies would be consistent with the assumption in the methodology that there are no biases. The reason given for not using anomalies is that there is insufficient tower data - if there really is insufficient data, this implies that ET cannot be trained on tower obs.

Jimenez, C., and Coauthors, 2011: Global intercomparison of 12 land surface heat flux estimates. *Journal of Geophysical Research: Atmospheres*, 116, D02 102, doi:10.1029/2010JD014545. Miralles, D., T. Holmes, R. de Jeu, J. Gash, A. Meesters, and A. Dolman, 2011: Global land- surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, 15, 453–469, doi:10.5194/hess-15-453-2011. Mueller, B., and Coauthors, 2011: Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations. *Geophysical Research Letters*, 38, L06 402, doi:10.1029/ 2010GL046230.

P18, L18: EC is known to under-estimate the fluxes. Using the sum of LH and SH as

C3

the incoming energy will almost certainly give an underestimate.

Figure 5: what is causing the sudden changes in the time series? The 91 day windows used shouldn’t suddenly change like this.

Figure 7: This sudden increase in the tower ET in the upper panels look incorrect (and seem to occur at the same time each year - unless these are preceded by significant rain events, this don’t look right). This time series needs to be checked, carefully QC-ed, and unusual features like this should be explained in the text.

The work would benefit from being placed within the context of other efforts to estimate ET with tower data / statistical methods. In particular the MTE product should be mentioned somewhere, as an example of using tower EC obs to estimate global ET.

Jung, M., M. Reichstein, and A. Bondeau, 2009: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, 6, 2001–2013, doi:10.5194/bg-6-2001-2009.

Interactive comment on *Hydrol. Earth Syst. Sci. Discuss.*, <https://doi.org/10.5194/hess-2017-573>, 2017.

C4