

Earth Syst. Sci. Data Discuss., author comment AC2
<https://doi.org/10.5194/essd-2021-181-AC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC2

Anna M. Ukkola et al.

Author 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-AC2>, 2021

RC2: This manuscript synthesizes a global network of site level meteorological and flux variables that draws from the existing flux tower networks of FLUXNET2015, OzFlux and LaThuile. For meteorological data the FLUXNET2015 data product is gap-filled using ERAinterim downscaled estimates, whereas OzFlux and LaThuile provided statistical gap filling techniques. For flux variables all data sets used statistical gap-filling approaches. If gap filling exceeds 10% of the data set, then the entire year is thrown out. A single format gap-filled data site level data product is helpful for a regional or global network of land model simulations, as well as for model-data fusion (data assimilation) purposes where a network of data could be ingested into a model to provide a regional analysis. Synthesized data sets like these are certainly welcome to the modeling and data assimilation community. This reviewer would like to encourage the authors to think about more quantitative inclusion of uncertainty with the gap-filling routines. Robust uncertainty estimates are an often overlooked but important piece included within data products both for model validation and model-data fusion exercises. See next section for more detailed feedback.

We are obviously pleased that the reviewer thinks this work will be useful for the community.

RC2: Sections 2.2.1: Would prefer a little more detail on the flux variable gap filling approach, where statistics on the skill of the filling methodology (linear regression against shortwave, temp and humidity) could be included from Ukkola et al., 2017?

A detailed analysis of the performance of this particular approach is a little out of scope for this paper, given this approach did not play a major role here. Nevertheless we appreciate the need to allow readers to understand exactly what has been done. There are several existing papers that do explicitly focus on the skill of this approach, so we have rewritten this section to briefly summarise their results and include references to these studies:

"Flux variables were gap-filled using statistical methods for all datasets. As per meteorological variables, short gaps of up to 4 hours were gap-filled using linear interpolation. Longer gaps (up to 30 days for OzFlux and FLUXNET2015, and 365 days for La Thuile) were gap-filled using a linear regression of each flux variable against incoming shortwave radiation, air temperature and humidity (relative humidity or vapour pressure deficit). This approach was demonstrated to outperform a range of LSMs in a broad range of metrics in out of sample tests (see Abramowitz, 2012; Best et al, 2015). In the absence

of air temperature or humidity data, the linear regression was constructed against shortwave radiation only. A separate linear model was created for day- and night-time data. Further details of all gap-filling methods can be found in (Ukkola et al., 2017)."

RC2: Table 1: Isn't NEE uncertainty reported in FLUXNET2015, in addition to GPP and ER partitioning uncertainty? Why not provide here?

Yes indeed it is, and in fact it was provided here, but omitted from the variable table in the manuscript. This has now been rectified, in addition to providing the original Fluxnet names for each of these variables and a reference to Fluxnet documentation, so that readers can find their definition and derivation in the Fluxnet database.

RC2: Figure 2: It appears for the site BE-Bra that SW_down and T_air are labeled as gap-filled (red-line), because in the text it is mentioned that year 2003 is removed for that site. This should also be stated in the figure caption and this time period should not be listed as gap-filled (red line). US-Tw2 is same thing, state in caption this was discarded.

Yes, this could be confusing. We have amended the caption to clearly state that these plots are pre-screening, and that some section were removed:

"Figure 2: Examples of meteorological data pre-screening plots for three sites (AU-Lit, BE-Bra and US-Tw2). For each site different processing approaches were used and sections of this data discarded."

RC2: "The majority of sites are located in grassland (40), forested (89) and cropland (17) ecosystems. 22 sites are located in savanna and shrubland ecosystems and 10 sites in wetlands." It would be preferable to see a more specific breakdown of these sites into plant species, or perhaps into more specific classifications (e.g C3 or C4 grasses, boreal/temperate evergreen, deciduous, crop types etc). Figure 3c has some of this information, but it would be helpful to spell it out here a bit more in the text, or at least refer to Figure 3c at this point.

From an ecological study perspective this is of course true, but keep in mind that this data set is being produced primarily for a study evaluating land models that are run inside climate models, that typically select vegetation characteristics based on a global grid of vegetation types. Many of these models would be unable to utilise this information, in particular the DGVMs. Nevertheless, where this information was easily accessible we already did include it, and indeed some soil and disturbance information, using the "vegetation_description", "soil_type" and "site_description" global attributes in the final netcdf files. Rest assured if this information were universally available for all of these sites we would have included it. Information on the C4 fraction at fluxnet sites is a long standing knowledge gap, with previous multi-site syntheses replying on satellite derived products to attempt to bridge this uncertainty (e.g. Medlyn et al. 2017 New Phyt.; De Kauwe et al. 2017 Biogesci.).

RC2: Figure 3: It is helpful to show the coverage of the sites in terms of MAT and MAP, and spatially. Might it also be useful to present site level location in terms of regions with the most productivity/biomass? Presenting a climate envelope is not as necessarily important as perhaps locating regions which have the most influence upon carbon cycle, for example. T.

That really depends on the purpose of a particular study. As noted above, this data set is being produced primarily for a study evaluating land models that are run inside climate models that are primarily focused on climate biophysics. Consequently, Figure 3 simply represents an overview of where sites are drawn from and as there is no attempt to carry out detailed analysis here, information about maximum productivity (etc) would be

beyond the scope or aims of this paper.

RC2: "Model evaluation, particularly at shorter time scales, should thus be avoided against long periods of gap-filled data." This statement could be made more quantitative by providing uncertainty bounds with the linear regression model so that the user, can make a quantitative evaluation of the model run. Simply, providing a recommendation that long gap-filled time periods should be avoided seems a bit qualitative and not as helpful as providing an uncertainty estimate.

It's a little unclear how one might make sure of such a quantification, except perhaps in a data assimilation exercise (which is not the purpose of the study for which this data set was derived). Either way, given that only a fraction of the gap-filling in this data set was performed by us, we are unable to derive this. While the statement could potentially be made more quantitative, it still serves an important qualitative role. We have opted not to amend the text.

RC2: " it should be noted that the remotely sensed LAI estimates are highly uncertain at the site scales, with large differences between Copernicus and MODIS LAI at many sites." Could you discuss/quantify what is meant as 'highly uncertain' at site scales, and where does this uncertainty derive from? Is it a case of algorithm uncertainty directly from the MODIS or Copernicus raw data – or is it a case of representation mismatch between the coarse product spatial resolution and the footprint of the site (1 km²)?

It is indeed both of these. We've added some additional text to make this clearer:

"However, it should be noted that the remotely sensed LAI estimates are uncertain at site scales, with large differences between Copernicus and MODIS LAI at many sites. This is both because of the difficulties inherent in estimating LAI from satellites (methodological) and the fact the satellite data may be drawn from a different footprint from the one that influences the site scale measured fluxes (De Kauwe 2011 RSE)."

RC2: Table S1: The manuscript states the flux data covers the period 1992-2018, but the vast majority of sites are (FLUXNET2015) and thus only available through 2014. I guess that is to be expected given you are drawing from FLUXNET2015, but a bit misleading.

We have made this clearer where stated in the introduction:

"The dataset covers the period 1992-2018 (although the majority of site records end in 2014) with individual sites spanning from 1 to 21 years, with a total of 1040 site years."