In the technical note “A view from space on global flux towers by MODIS and Landsat: The FluxnetEO dataset”, Walther et al. presents a standardized procedure to extract, gap-fill and quality control remote sensing observations around >300 flux sites. This contribution is critical to the reliable integration of remote sensing and eddy covariance measurements for understanding ecosystem functions and changes. I am in support of its publication, and my comments are meant to help improve the note and make it more clear to the audience.

> We appreciate your comments and your effort to help us improve clarity and thank the reviewer a lot! We answer to your questions point-by-point below in the indented blocks. Any sentences after re-phrasing we cite without line numbers as the final revised manuscript is not completed yet.

L34: As gap-fill is a key step in producing the dataset, perhaps it would be helpful to further clarify the general assumptions under these categories of methods. Other than the realistic considerations (i.e., generalizable, no need to use ancillary data) to do gap-fill only based on the remote sensing time series themselves, are there studies that suggest this method produce comparable results to complicated ones (i.e., the one that use ancillary meteo data).

> We understand that further benchmarking of FluxnetEO is absolutely necessary to understand the characteristics compared to other comparable data sets. In response to this and a similar question of reviewer 1, we will include in the revised manuscript:

- a paragraph in the introduction outlining key characteristics of FluxnetEO in comparison to other products
- a table in the SI
- a comparison of FluxnetEO Landsat with a gap-filled Landsat product from Moreno-Martinez et al. 2020
- experiments with an automated imputation by a missforest approach (Stekhoven et al. 2012) in the actual and also in artificially introduced gaps.

For the sake of space we do not include the same example plots here a second time, but would like to refer to our answer to reviewer 1 question 1.

L38: “contribution” means “study”?

> Yes, “contribution to the scientific knowledge and tool sets” was meant, but we changed it to “manuscript” to enhance clarity and fluency of the reading experience.

L51: reference for “view zenith angles”.

> We are afraid we do not understand what is meant here. In case a definition of a view zenith angle is wanted, it is the angle between the line of sight of a satellite instrument to the surface and the vertical line nadir above the observed point on the surface.

L136-137: I have some difficulties in understanding “The idea was to….instead of….”. I feel the authors are arguing that their method is appropriate for the study though I cannot understand the second part of the sentence. “Valid data” means ancillary data or just the good quality data of the time series.

> We split this sentence and this part now reads as: *“ A number of possible applications will require the analysis of actual observations, and consequently approaches that fit smooth functions to available good quality data (e.g. Jonsson and Eklundh 2002, Gonsamo et al. 2013) to represent a gap-free time series are not suitable. The idea therefore was to retain the good quality data and make as realistic estimates as possible for the gaps between them. “*

L154: it is not easy to understand the scaling method without carefully looking into some equations in ANN C. Perhaps it is helpful to insert some equations here, such as y = ax + b, where x means MSC while y is the non-gap filled time series. Then we can get a and b from the equation for each time window, and then apply a and b back to MSC for gap filled y.

> Thanks for your suggestion. We agree that equations strongly help to understand the processing in detail, and decided to add equations in appendix B. We also slightly rephrased point 4 of the conceptual description of the gap-filling procedure in the main manuscript for clarification. It now reads: *“Linearly regress the time series on its own median seasonal cycle. Compute a re-scaled median seasonal cycle with the obtained regression parameters and use it to fill longer gaps. Execute the regression and re-scaling in temporal moving windows as this guarantees more flexibility to correctly represent inter-annual variations in the time series and even partly accounts for changes in the shape of the seasonal cycle due to disturbances. It is, however, not suited to fill regularly recurring gaps at a certain time of the year, e.g. during rain seasons (Verger et al., 2013).”*

L175. Out of curiosity why do not use quality flag of MODIS here, any issue with the flag? By using statistical method only to remove the so-called outliers, are we risking removing some true extreme values?

> We had applied the MODIS quality flag (filtering for good quality in the mandatory flag, or other quality if the estimated LST error was smaller than 2K), and found that this filter barely removed data points. Restricting the second filter criterion to the LST error <1K systematically removed lower LST values within the variability range of LST, especially during summer and daytime, but obvious outliers were still kept. In order to keep data availability high and remove the obvious measurement or retrieval artifacts we opted for the outlier filter.

L210. See my comment above regarding the description of the scaling method.

> Thank you, also here we rephrased this sentence to: *“In temporal windows, find a linear scaling between one LST time series and its own MSC. Use the slope and intercept parameters to compute a re-scaled MSC which fills gaps in the time series for days of year when the MSC is valid. “*

L336. From Fig. 6a it is not accurate to say LST is consistently 30% higher, it is only the slope that is around 1.3.

> Correct, thanks a lot for this. We will also modify this analysis part in response to a suggestion from reviewer one and will correct this part accordingly in the revised manuscript.

L338. Do we really see the “slope decreases markedly for the highest temperature”? The figure only shows that slope increases a bit with the height.

> We did not intent to refer to the measurement height here with the word “high” but to the peak temperatures. We will rephrase (see also the last comment) and hope to clarify.

L389. For those sites with footprint less than 1km (which I think many sites are), how to define this aggregated snow flag. Are they either 0 or 1?

> The snow flag is 0 or 1 for each 500m subpixel in the cutout. Any aggregation across a selection of subpixels in the cutout (within 1km, across a flux footprint,etc.) will follow the procedure described by averaging the snow flag across the selected subpixels, resulting in a value between 0 and 1, or missing if less than 50% of the selected subpixels had valid information on the snow status.

L401 – 404. I am also wondering the rationale for choosing mean seasonal cycle and median seasonal cycle in different datasets. I also have to say in FLUXCOM mean seasonal cycle of remote sensing data was used but here the use of median seasonal cycle seems to be prevailing.

> This is an error and should read median seasonal cycle for Landsat as well. We figured that taking a median is advantageous over a mean to reduce the influence of undetected outliers, e.g. residual snow contamination in the reflectance-based processing.

L418. Valid snow cover < 60 days = snow does not occur at the site? I have a feeling the threshold is a bit large, e.g., a site with almost two months of valid snow cover might be considered to have no snow by this filter.

> This criterion is intended to identify sites with more or less regular snow cover to which the gap-filling step with a constant baseline value is applied. The criterion tests whether across the whole period that FluxnetEO currently covers (i.e. 21 years for MODIS and more than 30 years for Landsat), a certain number of snow days occurs. Admittedly, the thresholds are arbitrary, but based on investigation and testing. Sites with more or less regular snow cover cross this threshold easily. The MODIS snow flag occasionally (and supposedly wrongly) assigns snow for some days to weeks here and there to sites that do not typically experience snow precipitation. Most of those sites are identified with this filter and we can prevent to wrongly fill those gaps (often in the middle of the growing season!) with a constant baseline value. This benefit clearly outweighs the rare occasions that we might not fill actual snow gaps with a constant baseline value at sites that do not typically experience snow.

L450. To double check, do you mean for each time window we get a m and n?

> Yes, indeed.

L455. There is a redundant “[]” in the equation. Perhaps also would be helpful to explain the terms in the equations.

> The redundant brackets were removed and the description reformulated to make the meaning of the terms clear.

Álvaro Moreno-Martínez, et al., Multispectral high resolution sensor fusion for smoothing and gap-filling in the cloud, Remote Sensing of Environment, Volume 247, 2020, https://doi.org/10.1016/j.rse.2020.111901.

Daniel J. Stekhoven and Peter Bühlmann, MissForest—non-parametric missing value imputation for mixed-type data, Bioinformatics, Volume 28, Issue 1, 1 January 2012, Pages 112–118, https://doi.org/10.1093/bioinformatics/btr597