

Earth Syst. Sci. Data Discuss., referee comment RC2  
<https://doi.org/10.5194/essd-2021-233-RC2>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on **essd-2021-233**

Anonymous Referee #2

---

Referee comment on "The ABCflux database: Arctic–boreal CO<sub>2</sub> flux observations and ancillary information aggregated to monthly time steps across terrestrial ecosystems" by Anna-Maria Virkkala et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-233-RC2>, 2021

---

### General Comments:

This manuscript describes a new database (ABC Fluxes) of CO<sub>2</sub> flux measurements in arctic and boreal ecosystems. Overall the manuscript is well written and clear. However, the process of downloading data needs to be clarified – as I explain below, the big green "Download Data" button on the ORNL DAAC website did not give me a complete data file. One apparently needs to scroll down to the bottom to request the entire dataset, but this is not at all apparent at first pass and could lead to users missing data.

I also question the decision to exclude studies with limited measurements during the summer (while including limited measurements in off-seasons due to data scarcity during those time periods). While I agree that data sets with limited repetition are more uncertain than data with many repetitions, ideally ABC Fluxes users could make their own judgements about whether to include this data in their work. I recognize that going back to include this data is likely unfeasible at this point, so please instead provide an idea of how many studies were excluded. Then, perhaps future versions of ABC Fluxes could include these additional studies.

### Specific Comments:

Line 208: "fluxes from plants and soils to the atmosphere."

Line 216: Since you have included the measurement scale for eddy covariance and chamber measurements, please comment on the measurement scale for snow diffusion.

Line 250: At some point in paper, please quantify (at least approximately) how much data you lose by using monthly values and excluding papers that do not report data on a monthly basis.

Lines 304 – 309: the sub-plot labels b, c, d, and e are mixed up

Line 313-314: Here you only cite 5 prior synthesis efforts, but your Table has 7 other efforts. Is this an oversight, or is there a reason you did not look at two of the synthesis papers to identify potential data?

Line 323-328: I am curious to hear more about why you decided to exclude summer measurements if there weren't many replicates. While fewer replicates will make the uncertainty higher, eliminating these data sets entirely could throw away potentially valuable information that ABCfluxes users may want. It seems to me that, in an ideal world, ALL data of sufficient quality (if not quantity) would be included, and then database users can decide whether or not to include these small studies in their results. I am not suggesting you re-do the database now to include all these disregarded summer points, I recognize that would likely be a huge time investment, but I am curious to know approximately how many datasets were discarded.

Line 364-368: Please clarify here that in FLUXNET2015, "\_QC = 0" means measured values and "\_QC = 1" means good quality gap-filled values. Otherwise, your phrase "0 = extensive gap-filling, 1=low gap-filling" could be interpreted to conflict with the

FLUXNET2015 QC designations and may confuse people. You could write something like "indicating percentage of 366 measured (quality flag QC = 0 in FLUXNET2015) and good-quality gap-filled data (quality flag QC = 1 in FLUXNET2015); average from daily data; 0=extensive gap-filling, 1=low gap-filling).

Line 397: Ah! I now see that the dataset is supposed to have 6309 rows. The first time I downloaded this dataset I went to <https://doi.org/10.3334/ORNLDAAC/1934>, logged in, and clicked on the big green "Download data" button near the top of the page. However, the file I get from this only has 1408 rows. I now see that if I scroll down the webpage I can request a much larger file. Why does the "download data" button not provide the full dataset? Can this be changed? If not, you may want to warn the user about this in your text.

Table 3: please clarify that by "Number of observations" you mean # of months of data.

Line 515: Please clarify that the likely reason you have less data from 2015-present is because of a reporting lag, not because eddy covariance towers are measuring less data now.