Small waterbodies reduce the carbon sink of a polygonal tundra landscape

Beckebanze et al.

This manuscript describes a measurement campaign conducted to measure CO2 and CH4 fluxes in a polygonal tundra in northern Siberia, with a specific interest in fluxes of a merged polygonal pond. The 2-month data has been used to show that if emissions from ponds are neglected and those from the tundra surface only are used, the upscaling to landscape fluxes underestimates the landscape CO2 uptake rate. There is not much data existing from polygonal tundra or ponds in permafrost region, and therefore I think the manuscript adds an important piece in the understanding of GHG fluxes in tundra landscape. In general, I find the paper and the data important and I suggest acceptance of the paper after minor revisions. There are certain aspects in the calculations, explanations and interpretations which are not clear to me and I think the paper could be improved by clarifying these. In addition, there are inaccuracies in the text which need to be revised before acceptance of the paper.

I suggest using overall an alternative expression for ”decrease the landscape carbon sink”. In my mind the sink does not decrease. I think a more descriptive expression could be e.g. “not accounting for the pond fluxes results in an overestimation of the tundra surface CO2 uptake” or similar. I understand however that in the previous literature this might have been a typical way to express the phenomenon, but it is not too late to change this.

More specific comments

- open path ch4 analyzer has been used in the study. It is well known that relatively large “Webb-corrections” are typical for that type of analyzers. I think it would be necessary to shortly discuss the implications of these corrections and perhaps mention how large they are.
- The unit of flux rate used throughout the paper is g m-2 d-1. That’s OK even though not the most typically used; however it should be mentioned in the figures what do the points in the figures represent – 30 min flux, daily flux, or something else. ALso, it is crucial to tell in each figure, if just pure measurement data has been used, or if gap-filled fluxes are included also.
- I think it is always useful to see a time series of the original (screened) fluxes. Or if not, it is a good habit to tell how much data was available and if there were long gaps.

line 22: “…reduce the C sink…” consider revising (see my comment above)

line 23: impact on what?

lines 129-131: This means that there is a probability of 10% that fluxes observed at the EC tower originate from areas outside of the light gray area. Medium gray represents 50-70%, medium-dark gray 30-50%, and dark gray indicates that there is a probability of less than 30% that the observed flux originates from within the marked area.

Shouldn’t it be: 10% of the flux signal originates outside the target area/fetch? Now it seems to me that you are saying that in 1/10 of the cases the whole flux data signal originates outside the area. That cannot be the case.

chapter 2.4.3: you say you use the model to partition and gap fill the NEE data. However, I do not see a mentioning of GPP or Rtot anywhere after this chapter. Also, it is not clear for me, where you have used the gapfilled data, and where you have just averaged the accepted observations. For example in Fig. 2, is this gap filled or measured data?

Line 149: “We split the datasets into a training (70%) and a validation (30%) data set to test model performance” Where do you show or discuss these test results?

chapter 2.4.4: modeled CO2 flux represents the vegetated tundra. So, do you use purely modeled data here, or gap-filled? If modeled, why not gap-filled? How many gaps there are in the data?

line 160: I’m not fully convinced why do you need the fluxes from the mixed surface to conclude something about the CO2 fluxes from the merged pond? Don’t you have enough observations from that directions? Seeing the number of accepted observations and their distribution in time would help in understanding that.

line 161: should be >30 & < 150, right?

line 165: “Thus, we can calculate the observed CO2flux…” this formulation sounds weird (why do you need to calculate the observed flux?), please consider revising, perhaps replace with “express” or something

line 170: “To improve data quality, we exclude 30-min flux intervals of Fpond when apond<50%.” Now there seems to be a contradiction: in lines 162-163 you state that Fmodeled, mix includes only data with <30% of weighted footprint fraction of open water (apond?). But now you say that you exclude all Fpond values with apond < 50%. Perhaps this needs clarification.

lines 222-232: I have difficulties to follow the logic in this text. The chapter starts by stating that “To evaluate whether the differences in medians between the four wind sectors are significant, we apply a permutation test”. Then fluxes are randomly assigned to one of two groups (why two? Ok, this comes evident when one looks at the appendix figures. But not from the text). What is unclear to me is that how can you conclude from the test explained here and illustrated in Appendix figures that “no meteorological parameter acted as a driver for the high CH4 emission”?

Then, the CH4/CO2 ratios explained on lines 233-242: what is the conclusion from that analysis? I do not find any discussion about that.

Lines 271-272: “Our approach of combining a footprint model with a land cover classification to extract fluxes from different land cover classes allows us to determine the pond CO2flux.” This sentence is in the core of all my difficulties in understanding what has actually been done. Didn’t you use the direct pond fluxes (from sector 60-120) to infer the pond CO2 flux? At least this is what you mention on line 207, and in the table 1. And (in my understanding so far) you used the footprint model approach to estimate the flux from “tundra” (or semi-terrestrial tundra, vegetated tundra; are these same? If yes, please ease the reader’s pain and use uniform expressions here. If not, please explain more clearly what’s the difference.

Chapter 4.2: The observation of the CH4 spike in the shore120 is interesting, and the fact that it remains unexplained, is pity but not unexceptional in flux studies! It is also somewhat convincing how much effort you have had to explore the reasons for the higher emission

line 340: a somewhat similar approach has been used also earlier, see e.g. https://bg.copernicus.org/articles/16/255/2019/

Figures 2 and 5 (which are nice and indicative figures overall!): please indicate if the fluxes consist of purely measured values, purely gap-filled, or both. If just measured values are shown, how are the mean values (in red) used in the study? If there are missing values during the day, the mean does not represent the true daily NEE. Is the mean (red value) a mean of all fluxes from that direction during the 2-month period?

Figure 3: is each dot a 30-min flux? Please explain it. “Flux intervals at night time”? Why interval, aren’t these just fluxes?

Fig 4: please explain how the violin plot should be interpreted

Fig A2: what is the red line?

Figs A2 & A3 and line 310: there are no b’s or c’s in Fig. 6