Ignoring carbon emissions from thermokarst ponds results in overestimation of tundra carbon uptake

Lutz Beckebanze, Zoé Rehder, David Holl, Christian Wille, Charlotta Mirbach, and Lars Kutzbach

Dear Reviewer,

Thank you for your feedback. Please see our detailed responses below.

Thank you for the revised manuscript. Nice job done. I think, however, that the manuscript still needs work to clarify the presentation. I found the text incoherent, in many cases. I suggest that the authors will put some effort to review the consistency of the used terminology (for instance, ponds and models) and the logic and order how the different modeling and upscaling steps appear in the text. I feel that the problem is the precision of description.

These include also not so precise use of "CO2 flux" (shouldn't it be net ecosystem exchange of CO2, or a component flux, with direction relative to the atmosphere).

Thank you very much. In terms of consistency, we streamlined the use of the following words: ponds (and their different categories), models, NEE, CO_2 flux. Regarding the CO_2 flux, we also added a general sentence in the beginning of section 2.3 to define the direction of the fluxes:

We performed the raw data processing and computation of half-hourly fluxes for open-path and enclosed-path fluxes CO_2, CH_4 and H_2O) using EddyPro 7.0.6 *(Licor, 2019).* ***The convention of this software is that positive fluxes are fluxes from the surface to the atmosphere, while negative fluxes indicate a flux from the atmosphere downwards.***

Regarding the modeling and upscaling steps, please see our comment below.

And the handling of the CH4 data raises questions. One reason for my stumbling through the text may be in the end of the introduction: "Due to the tower's position, fluxes from the merged polygonal pond are the dominant source of the observed EC fluxes under easterly winds. The observed EC fluxes are dominated by semiterrestrial polygonal tundra with only a low influence from small thermokarst ponds from the other wind directions. We aim to deepen our understanding of carbon emissions from thermokarst ponds and constrain their impact on the landscape carbon balance. To this end, we (1) compare the water body and tundra fluxes focusing on temporal and spatial patterns, and we (2) investigate the influence of the merged polygonal pond on the landscape carbon balance."
That paragraph should precisely describe the objectives and questions. I may be wrong, but should you state here something like that your tower was positioned so that you can assess fluxes from the merged polygonal pond (large enough to get a pure 'thermokarst pond' signal) and polygon tundra composed of small ponds and terrestrial tundra (centers and rims?). Then, the objectives are, 1) use footprint model and 2) model net ecosystem CO2 exchange using the footprint weights of terrestrial tundra and thermokarst ponds (polygon ponds?), 3) examine temporal and spatial patterns of NEE & CH4 from terrestrial and aquatic tundra/land cover types, and 4) investigate the influence of the small thermokarst ponds (merged and polygon centers?) on the landscape NEE of CO2 (and CH4?) over the xxx period.

Thank you for this detailed suggestion. We agree with your suggestion and have adapted the introduction accordingly.

This paper aims to deepen the understanding of carbon emissions from thermokarst ponds and constrain their impact on the landscape carbon balance.

We (1) examine the temporal and spatial patterns of NEE and the spatial pattern of CH_4 flux from semi-terrestrial tundra and thermokarst ponds, and (2) investigate the influence of the thermokarst ponds on the landscape NEE of CO_2 during the months June to September 2019. To this end, we use a footprint model and model net ecosystem exchange (NEE) of CO_2 using the footprint weights of semi-terrestrial tundra and thermokarst ponds.

Additionally, we also changed the pond terminology in the introduction to clarify the differences between polygonal ponds, merged polygonal ponds and the more general term thermokarst ponds.

*These **thermokarst** ponds are often only as large as one polygon **(polygonal ponds)**. When several polygons are inundated, this can cause larger shallow **thermokarst** ponds to form, which we term merged polygonal ponds (Rehder et al. 2021).*

Line numbering would help pointing the comments, but here just some examples from the text Usually the mat&met and results are written in past tense. Check some published papers for a reference. A language edit maybe.

We apologize for forgetting to insert line numbers. We followed your suggestion and changed the mat&met section to the past tense.

p. 4. Fig. 1 add somewhere that there are larger thermokarst lakes in the map (not included in the study)

While some of the large water bodies visible on the orthophoto are indeed thermokarst lakes, it is not clear if all the lakes on Samoylov formed through thermokarst processes. However, we add a sentence pointing that we exclude the large lakes.

*The location of the study site in Russia is shown in (a) and the location of Samoylov Island within the Lena River Delta is shown in (b). Samoylov Island is shown in (c); the surrounding Lena River appears in light blue. The outline of the river-terrace land-cover classification (Sect. 2.4.1 is indicated by the blue line. **We focus on the polygonal tundra, however, large lakes are excluded (circled in yellow)**. In (d),...*

p. 5. Add what's the area of upscaling somewhere in the mat&met

We were unsure what's meant, we changed two things to be on the safe side:

1) We added the following part in the section of the land-cover classification:

(3.0 km^2, area within the blue line in Fig. 1 c)

and we also added a description in the section of the open water CO_2 flux:

To evaluate the impact of thermokarst ponds on landscape CO_2 flux, we estimated a polygonal tundra landscape-CO_2 flux from the late-Holocene river terrace of Samoylov Island...

2) We included the absolute area of the land-cover classification in the section of the land-cover classification (2.4.1) and added the relative contribution of each land-cover class in this section.

*The land-cover classification has a resolution of 0.17 mx 0.17 m. It is projected onto WGS 1984 UTM Zone 52N and the land-cover classes include open water **(15.7%)**, overgrown water **(7.0%),** dry tundra **(65.1%)**, and wet tundra **(12.1%),** as defined by Muster et al. (2012)*

p. 6. There are some illogic in describing the models and missing equation references Explanation of the different grey shades is good in the figure legend, but no need in the main text.

We agree and moved the explanation of the different gray shades to the figure caption.

P. 6. Number the models according to their appearance in the text. i.e. NEE (1), ….

Thank you for pointing this out. We re-structured the section and start with the description of the NEE followed by the description of the NEE components. Now the order of the equations fits the order in which we reference them. Please refer to section 2.4.3.

p. 7. in addition to the R2, RMSE and mean absolute error would be great, because the R2 is not informative how close the absolute values are.

Thank you for this suggestion. We now provide information about the overall model performance in the form of the RMSE for both applications of the bulk-NEE model:

The final RMSE between the model input and the gap-filled NEE had a value of 0.29 g $m^{-2}d^{-1}$.

And

This gap-filling modeling of CO_2-C flux had an RSME of 0.31 g $m^{-2}d^{-1}$.

p. 7. Would it make any more sense to explain the extraction of pond values before the gap-filling?

Thank you for this suggestion. While we would like to keep the general order, we moved the section about the detailed application to the first usage of the model to 2.4.4 to better separate the model description and the model application.

Now section 2.4.3. only explains the general model. In section 2.4.4. we apply this model twice. In this process, we also re-named the section 2.4.4 from "The open-water CO_2 flux" to "*Separating CO2 fluxes from tundra and open-water*".

And be precise if question is about the merged pond or polygon ponds or something else
Why it is worth mentioning that others compared between different vegetation types and here between tundra and water?

We clarify the aim of this step by adding the following sentence in section 2.4.4:

Since our aim was to assess the impact of thermokarst ponds (both polygonal ponds and merged polygonal ponds) on NEE, we needed to eliminate the influence of polygonal ponds from our NEE estimate.

p. 8. At hand? How the others gap-fil CH4 fluxes? Can't you use the data even without gap-filling if enough data? Give a rationale why these wind directions were chosen for the comparison, now pretty random in the text.

Thank you for pointing this out. We changed the beginning of the section to make clear why we can't use a gap-filling model and added information on why we chose the wind sectors:

The CH_4 emissions from the heterogeneous landscape around the tower are less spatially uniform than to the CO_2 emission. Therefore, we can not use a gap-filling model for the CH_4 comparable to the bulk model we used for CO_2. Thus, we investigate CH_4 emissions differently. Based on preliminary results from our analysis and the aerial image of the study site, we focus on four wind sectors instead of extracting the fluxes from the landcover types.

p. 9. When comparing, variation is better to be expressed as the ratio of the standard deviation to the mean. It is coefficient of variation.

Since we think there is meaning in the absolute values, we use instead the median and the percentiles, which are also used several times in the manuscript. Our point is that the fluxes from the tundra reach higher positive and negative numbers, indicating a stronger diurnal cycle, which can also be seen in Fig. 2. However, we clarify that we are only comparing absolute values.

The CO_2-C fluxes from this pond sector show a smaller absolute variability ($0.09^{0.38}_{-0.33}$ g $m^{-2}d^{-1}$, $Median^{95\%\ Percentile}_{5\%\ Percentile}$) than the fluxes from all other wind directions ($-0.08^{0.87}_{-1.56}$ g $m^{-2}d^{-1}$, $Median^{95\%\ Percentile}_{5\%\ Percentile}$).

P. 9. In the permutation test section: Should there be something else than "one of the two groups" in the next sentence? pairs? "In this test, we randomly assign each 30-min flux to one of two groups and calculate both groups' median and their differences."

Thank you for this suggestion. We added "all possible combination of pairs" to the following sentence to make the pair-structure of the permutation test clearer. The new sentence reads:

We conducted six tests in total, using all possible combinations of pairs with the four wind sectors.

p. 10. About the shore sections. "These peaks did not correlate with any of the four land-cover classes" is unclear

Thank you for this suggestion. We tried to clarify this sentence by changing it in the following way:

These peaks did not correlate with a specifically large contribution of one of the land-cover classes to the footprint.

p. 10. How about the high mean values in the sector 300-360? Don't neglect if 120 deg gets attention.

Thank you for this comment. We also had a look at the wind direction 290-330° and the peak at 180-190° in our analysis. However, the focus of this manuscript lies more on the water body fluxes than on spatial heterogeneity of tundra fluxes. We will ~~to~~ focus on the tundra fluxes more in an upcoming manuscript which uses more data. By averaging over the tundra wind sectors, we can compare an 'average tundra' to the water body fluxes. Nevertheless, we are happy to include more information on the peak at 180°-190° instead, which we think is a more pronounced 'peak' and thus more comparable to 120° (see text below).

In summary, neither high wind speed nor high temperatures act as a driver for the high CH_4 emission from shore$_{120°}$. In contrast, the peak at 180°-190° can be explained reasonably well using air temperature and friction velocity in a multiple linear regression (R^2 = 0.44). Using the same predictors results in an R^2 of 0.20 for the peak at shore$_{120°}$.

Fig. 6. is really nice

Thank you!

p.14. Comparison with the other pond studies in the area is a bit confusing. Methodological differences are possible but note also whether those were same or different ponds and how large the samples were.

Thank you for this comment. We added the number of samples (46) used by Abnizova et al. and also explained that the samples were taken on ponds not within the footprint of our study site (section 4.1).

p. 14. Maybe specify that with emissions you mean NEE of CO2 (direction to the atmosphere)

Thank you for this comment. As mentioned in a previous response, we streamlined our use of the terms NEE, CO_2 flux and emission.

Fig A3. Lake? Do you mean the merged pond?

Yes, we changed the word to "merged polygonal pond".