Reviewer 1

The addition of the statement on eddy covariance as a means of assessing high temporal resolution GHG fluxes from lakes is jarring in its current position in the manuscript (lines 380-382). Suggest moving to the end of the following paragraph which highlights the need to measure across discrete, hard to predict events such as turnover.

Moved to the final paragraph of the discussion

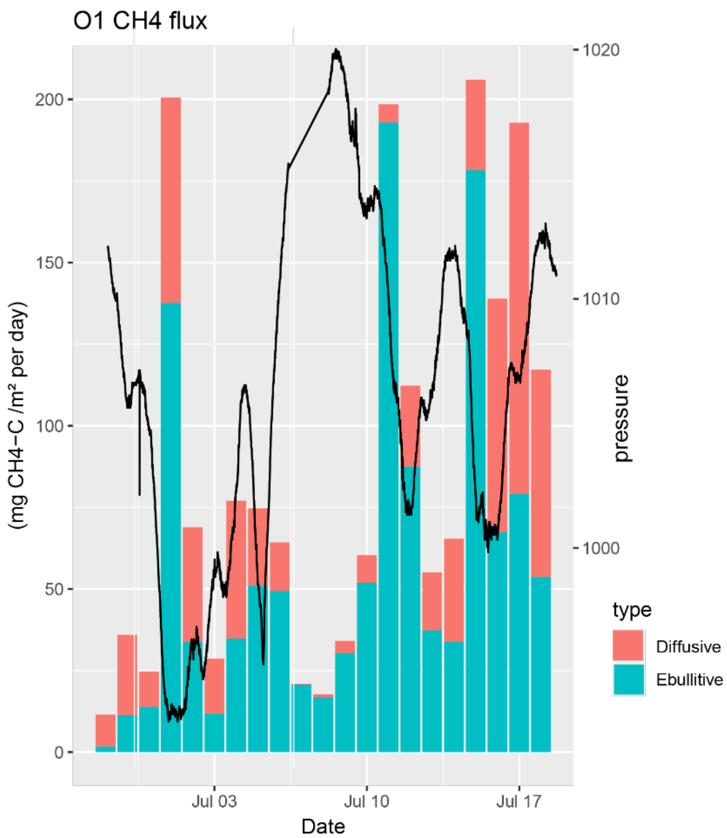
Reviewer 2

I appreciate the edits the authors made on this manuscript, and overall find it improved. I still believe this paper will be of broad interest to the scientific community. That said, I still have two concerns with the paper:

Ebullition estimates: I am still concerned about the ebullitive flux estimates, and do not think these fluxes should be included in the paper. My concern for the ebullition flux estimates is as follows. The authors use floating chambers sampled every two weeks to estimate ebullition. I am not aware of any published studies that use this method as opposed to bubble traps or estimating bubbles from portable gas analyzers (the latter is tricky as the authors point out because bubbles are so spatially and temporally variable), and the authors provide no citations for this method. I don't think the floating chamber method is appropriate and I worry that publishing this method will encourage its use among researchers. The problems I see are: (1) the large surface area of a bucket (compared to bubble traps with syringes or graduated cylinders) increases the likelihood of diffusive fluxes and then reciprocally, back-diffusion, (2) the 2-week sampling period will also increase back-diffusion. In contrast, bubble traps measure volume of bubbles directly and their smaller surface area reduces diffusion in and out, and sampling often occurs more frequently. Bubble traps are also easy to use and can be deployed over long time periods (weeks to months)—this is not that logistically or financially challenging (e.g., Burke et al. 2019 JGR, Ray et al. 2023 GRL, DelSontro et al. 2016 L&O, Baron et al. 2022). The authors have made sincere attempts to justify and estimate error with their estimates, including the new supplement comparing ebullition to automated flushing chambers. However, there are still some large assumptions: (1) they assume 0 ebullition if chamber concentration is similar to water concentration (I assume there would have been some bubbling over two weeks), and (2) assume high chamber concentrations account for most of the bubbles, acknowledging back-diffusion. The flushing chamber comparison shows huge differences on the daily scale (e.g., Supplemental Table 2), and we don't know the "true" ebullitive flux without bubble traps integrating multiple days.

#We appreciate the reviewers concerns and we also have concerns about the methods and would not recommend their use if other methods were easily available. However, bubble traps suffer from a number of limitations too, as they are harder to maintain in a long-term deployment of many months. In eutrophic systems they suffer from huge biofouling and can become blocked by algal growth and filamentous algae. The reviewer is correct that very low ebullitive emission might be missed by the approach here, but the data suggest this is not a big problem as a very small proportion of the samples (28 of 401 records) were recorded as having zero ebullition. Also, we show that short-term, a few days, or a week every month of a 'better' method is a poorer estimation of ebullition in terms of % error in the estimate (median 50% error) than the continuous deployment of the static chamber (15-30% error) (supplementary materials). We cautiously revised the text to emphasize the caveats involved in their use, highlighting the underestimation of ebullition but we also show there is value in the data, through the comparison with the 'true' data from the flushing chamber.

the flushing chamber does show huge daily variation with the peak in ebullition correlating with drops in atmospheric in pressure (see fig 1). These chambers do in fact measure the true ebullition as we separate the diffusive from ebullitive emission, by repeat resampling 3 minute periods and discarding periods where bubbles arrive (if R2 value is low, indicating a non-linear increase) and selecting the median beta of the regression equations that pass the test of linearity we can reliably determine diffusive flux. Then we can calculate the concentration in the chamber expected from diffusion after 4 hours and by taking this away from the final measured concentration we can separate ebullition from diffusion.



.

Fig 1. Ebullitive and diffusive flux over 3 weeks period in June/July 2023, showing that peak emissions are driven by drops in atmospheric pressure (black line).

In short, these flux numbers are unreliable, and I don't think they are useful if we can't trust them (even as underestimates). As stated above, it makes me nervous to publish this method as other researchers may then use it as opposed to using bubble traps. If the authors removed the ebullition estimates, I think there is still a compelling story that periodic mixing events result in high GHG fluxes from shallow lakes.

We altered the text in the section discussing the use of chambers to the following:
"Thus, the continuous monitoring of ebullition using chambers with known biases was deemed the least

worst method available, but we acknowledge the caveat that ebullitive emissions are underestimated, (see supplementary material). We further acknowledge that this approach of static chambers should, where possible, be replaced by other methods to estimate ebullition, such as automatic flushing chambers.

Convention on assessing stratification and depth of hypolimnion: On a less fundamental, but still important, note, I still believe the authors should follow convention with defining stratification by density, which is also what should be used to identify the area of the hypolimnion (e.g., see Gray et al. 2020). The authors currently use temperature for stratification thresholds, but the physical mechanism is density, which has a non-linear relationship with temperature. Given that there is a strong temperature gradient in this study (April – October), these relationships may become important (though will not likely dramatically alter results). Further, there are no citations for using oxygen over density (or temperature) for estimating anoxia in the hypolimnion. If the authors prefer to focus on oxygen, consider using a more common metric of "anoxic fraction," or the fraction of the sediment exposed to anoxic waters (see Rabaey et al. 2023, Frontiers in Env. Science from Nurnberg et al. 1995).

To address the concerns of the reviewer concerns we reanalyzed the data to define the thermocline depth and the area of the hypolimnion using density with the 'rlakemonitor' and 'lakeanalyser' R packages. The results are the same as our previous estimation of hypolimnetic area are not different than those we already used. We changed the methods section to include the use of these packages and methods of assessment, but the results remain unchanged.

Introduction

• Lines 43-44: I think Rosentreter says all aquatic systems contribute to half of global emissions (not lakes and ponds specifically)?

altered

• Lines 57-61: As macrophytes are not important to the current study, I suggest removing.

we are discussing controls of GHG in general in the intro so I think that it is relevant to mention macrophytes

• Line 83: Citation was removed between drafts—add back in?

#thanks - restored

• References to the Sondergaard paper (same lake and time frame have been removed)—it seems important to highlight this work in the Introduction, and explain how this work expands it.

Done in intro

• Lines 81-83: citation needed. Is there consensus here?

This section in well references

Methods

• Line 139: from 26 May 2020 through when in the fall?

#added

• Line 187: was the oxidation study done in this lake? Please provide more details.

Yes – now clarified

• Lines 216-217: It is impossible to know which is “better” as we don’t have true fluxes in this comparison—e.g., not compared to bubble traps.

See general comments the flushing chambers provide ‘true’ flux estimates as good or better than bubble traps

Discussion

• It would be helpful for the first paragraph of the Discussion to be an overarching statement of major findings and roadmap for the Discussion, rather than immediately comparing to three other studies.

We added a sentence clarifying the main aims and findings at the beginning of the discussion.

• I still find that the Discussion would be easier to follow with subheadings.

We think the discussion is clear and logically laid out, but have added subheading and would suggest the editor advises on whether they are needed.

• Lines 354-371: Other studies have also found higher ebullition linked with deeper locations in lakes, including Sø et al. 2023, Science of the Tot. Env.)

• Lines 365 – 371: See Ray et al. 2023, Sø et al. 2023—other studies examine the complexities beyond temperature

• Lines 373 – 381: Bubble traps are not that much more logistically or financially challenging – some bubble traps are deployed for long periods of time (e.g., Burke et al. 2019 JGR, Ray et al. 2023 GRL, DelSontro et al. 2016 L&O, Baron et al. 2022)—these papers would be useful to cite not just for methods, but also for interpreting ebullitive fluxes, should they continue to be included.

• Lines 450-453: for temporal sampling, the authors may want to reference Ray et al. 2023 (L&O), Natchimuthu et al. 2017 (JGR), and Wik et al. 2016 (GRL). These papers all examine temporal variability and emphasize the importance of more samples over an open-water season.

to address the above 4 bullets we added more extensive reference including Sø 2023, and Burke 2019 and Ray 2023, DelSontro et al. 2016 to the discussion of ebullition.

Supplement

• Table 2: Add units for flux, and note the date is DD-MM-YYYY. Caption should briefly explain timescale – what is differences from the mean? Mean floating chamber estimate of bubbles? This is also not clear in the text, and supplemental materials 2 and 3 are cited together. Please explain each individually, along with the take-home messages.

explanations added

added

Minor comments:

• I agree with Reviewer 1 of using “freshwaters” rather than “fresh waters”

Fresh waters are two words as a noun and one as an adjective, e.g. Freshwater ecology is the study of the ecology of fresh waters, which is how they are used in the paper. As this is correct English I prefer to keep it that way, but English evolves though use and it is increasingly common to see freshwaters as a noun, so it is not a hill I am prepared to die on if you insist it is changed.