

Review of *Stratification-induced hypoxia shapes greenhouse gas dynamics in ponds* by Esposito et al.

General comments

This study presents a valuable dataset characterizing the development and frequency of thermal stratification and associated hypoxia, as well as their effects on CH₄ concentrations and fluxes over an entire growing season in six small, shallow ponds. Although the main findings generally support previous observations from lakes and ponds, the study provides a novel contribution through its detailed, depth-resolved measurements of water temperature and dissolved oxygen, combined with measurements of CH₄ and CO₂ concentrations in both surface and bottom waters. This approach allows the authors to assess how short-term transitions between stratification and mixing, and the resulting alternation between oxic and hypoxic conditions, affect methane concentrations and fluxes. The study therefore offers valuable insight into greenhouse gas dynamics in small ponds.

Overall, I found the manuscript valuable and relevant. However, several aspects would benefit from further clarification or more detailed consideration. In particular, I consider it important to address the uncertainty associated with the CH₄ ebullition estimates and its potential implications for the reported total CH₄ fluxes and the relative contribution of ebullition. In addition, the description of the statistical analyses should be expanded to provide sufficient information about the model structures and how repeated measurements and temporal autocorrelation were addressed. Finally, reporting the actual diffusive CH₄ fluxes would provide valuable context and facilitate comparisons with other studies (see the specific comments below for further details).

Specific comments

Lines 62–63: Could the authors briefly explain the mechanism through which shifts in primary producers may enhance greenhouse gas emissions?

Lines 123–125: Did this calculation account for the background concentrations of CH₄ and CO₂ in the air used for the headspace method described in lines 118–120?

Lines 126–137: Except for the relative contributions of ebullition and diffusion, I could not find any results on diffusive fluxes in the Results or Discussion sections. I understand that the diffusive fluxes were calculated from dissolved gas concentrations, whose temporal patterns are described in detail in the Results. Nevertheless, reporting the actual diffusive flux values would provide useful additional context and facilitate comparisons with other studies. I therefore suggest including, at a minimum, the mean and range of the diffusive fluxes.

Lines 151–153: The sampling design likely leads to some underestimation of CH₄ ebullition because of CH₄ exchange between the chamber headspace and the pond water. Given that the magnitude of this uncertainty was quantified in the cited study, could the authors clarify why it was not taken into account when reporting the results? Is the uncertainty range reported in that study also applicable to the ponds investigated here? If so, I suggest presenting values corrected using this uncertainty range alongside the original, uncorrected estimates. Such a

correction could affect both the estimated total CH₄ fluxes and the relative contribution of ebullition to the total flux, which appears comparatively low relative to values reported for other ponds when based on the current conservative estimates. The potential effects of the selected methodology on the results, including a comparison of corrected and uncorrected CH₄ ebullition estimates, should also be considered in the relevant part of the Discussion.

Lines 163–173: The description of most of the statistical analyses is rather general and would benefit from additional detail. More specifically, for the analysis described in lines 168–172, which aspect of hypoxia was tested: its occurrence, its duration, or the dissolved oxygen concentration? Please also specify which other variables were included as fixed effects and which, if any, were included as random effects. Was the final model assessed for temporal autocorrelation? This could influence the results because repeated measurements were conducted in the same ponds.

Similarly, for the analysis described in lines 172–173, please specify the type and structure of the linear model used. Did the model account for repeated measurements within ponds and potential temporal autocorrelation? These observations cannot be considered fully independent.

Lines 279–281: Three different terms appear to be used for CH₄ ebullition within these three lines, including bubble flux and CH₄ bubble emissions. This terminology may be confusing because, in some studies, bubble flux refers only to the volume of released gas, irrespective of its CH₄ content. To avoid ambiguity and ensure consistency, I recommend using the term CH₄ ebullition throughout the manuscript.

Lines 293–295: Please also report the ranges (minimum–maximum) for the relative contributions of diffusive fluxes and ebullition. Could the comparatively low contribution of ebullition observed in these ponds be, at least partly, an artefact of the selected methodology, which may substantially underestimate CH₄ ebullition?

Lines 291–300: Given that ebullition was generally higher in ponds experiencing more frequent and prolonged hypoxic events, was this pattern also reflected in a greater relative contribution of ebullition to the total CH₄ flux? Alternatively, were diffusive fluxes also higher in these ponds, resulting in a similar relative contribution of ebullition across ponds?

Lines 326–328 and 334–335: Can the available data clearly distinguish whether CH₄ accumulation in the bottom waters of these ponds was driven primarily by enhanced CH₄ production rather than by reduced CH₄ oxidation? One could argue that the sediment layers where methanogenesis occurs remain anoxic immediately below the sediment–water interface, irrespective of whether hypoxia develops in the overlying bottom water. This interpretation may also be supported by the observations from pond RH3, where ebullition increased before hypoxic conditions developed in the bottom water (lines 276–277). What evidence indicates that the observed CH₄ accumulation resulted mainly from enhanced methanogenesis rather than reduced methanotrophic activity? This distinction should be clarified or the interpretation presented more cautiously.

Line 348: Please specify the duration of the hypoxic events described here as “short.”

Technical corrections

Lines 80–86: Table 1 does not include the TN ratio, although this variable is mentioned in both the main text and the table caption.

Line 178: Figure S3 does not appear to present the manually measured temperature profiles. Please check and correct the figure reference

Figures 3 and 4: These figures are not boxplots, as stated in their captions. Please revise the figure descriptions accordingly.