Responses to comments on Manuscript Preprint egusphere-2023-1745 (Minor Revisions) by Clayer et al.

Title: Technical Note: Preventing CO2 overestimation from mercuric or copper (II) chloride preservation of dissolved greenhouse gases in freshwater samples

Referees' comments are underlined for clarity and our responses highlighted in gray. Line numbers refer to line in the revised (track-change) version of the manuscript, if not specified otherwise.

REFEREE REPORT(S):

Referee #1:

The figure legends are all difficult to understand at the moment. The captions are good. It would help if there were more words and less formulas/technical terms in the legends itself. They can then be explained further in the caption, if there is not enough space in the legend for a complete explanation of what each symbol means exactly.

Thank you for this suggestion. We updated the legend of Fig. 4 (now Fig. 5) which is now believed to be more explicit. Fig 1 doesn't have a legend. The legends in Fig. 2, 3 and 5 are concise and in-line with our terminology in the text. Note that small adjustments were made in Fig. 3 (now Fig. 4) legend changing "pH" for "*in situ* pH".

Referee #2:

Thanks to the authors for taking into account the suggested changes, which improved the manuscript and made the experimental setup clearer. I still have some comments with further suggestions. The methods for the two main experiments are described in more detail now, but I am still not understanding parts of them. After some more general concerns, I provide further comments below. The line numbers refer to the document with marked changes.

Thank you for another thorough assessment. Below we provide point by point responses to your suggestions. We have added an overview figure as suggested and have performed minor edits which are believed to significantly improve the clarity of the methods.

General comments:

I would like to reiterate the importance of the Methods section in a technical note. I suggest to add a table or graphic showing the two main experimental setups at a glance, together with the main parameters examined, the main instrumentation used, etc. I generally suggest to restructure the Methods section. For example, the GC was used for both experiments, right? However, the DIC analysis only applies to the second experiment if I'm not mistaken. Therefore, I suggest to simply get rid or rename some subsection headings, to clearly distinguish experimental setups. I leave this up to the authors, but currently I have difficulties fully understanding your Methods and experimental setups.

We agree that the Methods section is important. Thank you for a nice suggestion regarding the figure. We have added one (now new Fig. 1) which, we believe, makes the experimental procedures much clearer. The figure highlights where e.g., GC analyses were involved as well as calculations with

PHREEQC. We now start the methods section with one introductory sentence, see L. 116-118 and we also performed minor edits such as line 329 and 342 to improve clarity.

I'm still unsure about the usage of DIC analysis to estimate CO2 concentration. From what I understand this method introduces many uncertainties in estimating CO2 (e.g. Golub et al., 2017, https://doi.org/10.1002/2017JG003794), and I am not really convinced that it is a good way to compare with HgCl2 preserved samples for validation.

Among the various methods available, estimating pCO2 from DIC and pH is the least uncertain method according to Golub et al., 2017. They showed that the relative standard error of pCO2 determination based on DIC and pH was maximum 5.5%. We added this statement in the methods, see L. 210-211.

I also have difficulties with the section about the PHREEQC model and the resulting output. I still think it remains a bit vague what the rationale behind using this model in this study is, how it fits in with the remaining experimental design, and how it contributes to the findings of this study. Should the model outcome be considered essential for this study or is it rather an addition? I would favor moving (at least) the technical description of the model to the appendix and focus on the main components in the Methods section. Can you give references to studies where PHREEQC was used for similar purposes?

As stated in the beginning of the "*Chemical speciation (…)*" in the Methods, PHREEQC was used to estimate effects of preservative on pH, shift in carbonate equilibrium as well as carbonate precipitation (see L.321-323). To support this application, we now refer to Atekwana et al. (2016), Clayer et al. (2016) and Klaus et al. (2023) which have used PHREEQC for similar purposes.

Were results shown in section 3.2 obtained with the PHREEQC model? If so, that should be clarified. Is the pH data shown in Fig. 2 modeled and not directly measured? Are these findings reliable? Does a single comparison as in l. 508 suffice for validation?

These are observations which have been added upon request during revision. See L. 182-195. You are right, several comparison points need to be included, this is now included in section 3.3, see L. 472. Please also note the clarification in the caption of Fig. 3.

I think the additions of explanations in the Discussion helped clarify some aspects of this study and made it much rounder.

Thank you.

Specific comments:

l. 57 rewrite to something like: because it proved effective at very low concentrations…

Done, thank you.

l. 63-64 rewrite to: Previous studies showed…

Done.

l. 66 rewrite to: An alternative to using biocides is to collect in-situ water samples, extract the headspace in the field, and analyze the headspace in a laboratory…

Agreed, thank you.

l. 128 and 220: coordinates are missing the degree sign

These coordinates are not expressed as decimals, they do not need a sign since we added the “E” for East and “N” for North.

l. 155 and 205: replace “submitted”, e.g. by “lake water from the 25 L bulk sample was subjected to four treatments“

Agreed, thank you.

Tab 1: Could you add the uncertainty to the sample concentrations or can you not quantify it easily?

Yes, we considered the error of the scale (0.01 g) and the error of the flask (0.1 mL) when preparing the solutions.

l. 291: Was the system calibrated before analysis or several times?

The system is calibrated before each run and then standards are run every 5 or 6 samples. This is now added see L. 268.

l. 438: add: samples

Done, thank you.

l. 439: I think it is worth to be precise here. If you state t=24h, this refers to a 24 h incubation time, which was achieved after an equilibration period of the bulk water sample of 24 h, then distributing the bulk water into the bottle volume and leaving it incubate for 24 h, is that correct? Maybe you could briefly clarify that when you first mention this?

As stated L. 159, the filling of the 120mL bottles happened within 3h, the bulk water sample did not equilibrate for more than 1-2h since the lake water was 18.5C and room was at 21C. The total time between sampling and analysis was somewhere around 27 hours, 24h incubation plus 3h of temperature equilibration for the bulk water sample. So given the short equilibration time of the bulk sample, we don’t think it is necessary to repeat this information here.

l. 474 and 503: reword to: the whiskers display minimum and maximum

Done, thank you.

l. 498: replace “that” with “the”

Done, thank you.

Fig. 3 lower panel: could the differences we see for CO_2 be caused by differences in analysis techniques rather than fixation vs no-fixation? Or how can one be sure this was not the case?

Given that the relative error should be less than 6% (see l. 211), these differences spanning several 100’s or even 1000’s of µatm cannot be related to analytical biases.

l. 612-613: I do not understand what you mean by “required”

We rephrased this sentence to clarify see l. 532-533.

l. 620 replace “that” with “than”

Done, thank you.

l. 652: please reword to avoid repetition

We rephrased, thank you.

l. 678ff: I am unsure if this is the right location for this explanation or if it belongs in the Results section.

This paragraph fits better in the discussion given the high degree of data analysis and post-processing.

l. 735: what is natural water pH?

We rephrased for clarity, see l. 656-657.

l. 737ff: In terms of impact of this study, I think it would be interesting if you could give the reader an impression of how many studies there are that estimate CO2 concentration or fluxes using HgCl2 (or CuCl2) preservatives in the mentioned regions for freshwater samples? How large would be the overall error of current estimates?

Thank you for a nice suggestion. It is out of the scope of this study to quantify the overall error of current estimates, however, we now point towards a few studies from boreal lakes, but also from tropical aquatic environments where HgCl2 (studies using CuCl2 are rather rare) preservation might be a source of error to raise attention on this problem, see L. 664-669.

l. 767 greenhouse gas concentration

Added, thank you.

Comments applicable throughout paper:

Please use consistent names for the chemical addition: you use inhibitors, amendments, biocides, preservative, treatment, etc.

We removed the term “biocides” and only refer to “amendment” for the small volume of preservative solution added to improve clarity.

Replace Figure by Fig. and Table by Tab.

Done, thank you.

I think the overall paper could be shortened and be made more precise to further facilitate readability, but I leave this up to the authors.

We did not find any part that could be shortened and thus decided to leave the manuscript mostly as it is.