

Comparing float pCO₂ profiles in the Southern Ocean to ship data reveals discrepancies

Summary: The authors compare the means of pCO₂ data measured by ships (1972-2021) with pCO₂ data derived from Argo floats (2014-2023) in the Southern Ocean. They find an increase in the float pCO₂ values compared to the ship pCO₂ and explain the mean difference by seasonality, trends in atmospheric CO₂, differences in sampling location, errors in TA, and the choice of carbonate system constants. Consequently, they adjust the pCO₂ values by removing the influence of these factors on the mean (e.g., normalizing the data to a reference year). They attribute the remaining difference in mean pCO₂ to quality issues. While I appreciate the concept of comparing float pCO₂ data with ship pCO₂ data, I have major issues with this study. My greatest concerns are as follows:

1. **Content:** The study falls short in making a meaningful contribution to the existing knowledge base. It lacks the presentation of any novel findings. (Higher pCO₂ observed in Argo float data than ship data → partially caused by seasonality, different sampling location etc). The conclusion of "bad data quality" appears inadequate given the methodology and is insufficiently discussed)
2. **Methodology:** I fail to understand the rationale behind comparing data from various time periods, seasons, and sampling locations in the first place, particularly when focusing solely on the mean values. In my opinion, this approach is simply not acceptable, as e.g. ocean biogeochemistry undergoes changes over time, leading to higher CO₂ levels in more recent float data. While the authors acknowledge this in their later analyses presented in the discussion section, the results in the results section of the study are therefore not comparable. Additionally, the study does not quantify the sources of uncertainty in float pCO₂ data, rendering the conclusions regarding data quality issues questionable.
3. **Structure:** Result sections 3.1, 3.2 and 3.3 should be merged as the subsections merely contain different plots. The discussion section comprises the presentation of additional analyses, thereby resembling more of a result section.
4. **Choice of visualization:** The content in Figure 2, 3, S1 as well as 4, 5, S2 could be merged (remove scatterplots, add error bars to line plots).
5. Authors doubt/question data quality without further arguments (l.200-206). After adjusting the means, they did not go into "float pH data quality issues". I would have appreciated a discussion on why the quality is perceived as poor and how it could be improved etc

Minor, detailed comments

l. 9: Please specify why/what role

l.12: It reads like "pCO₂ estimates obtained... from ship-collected data" which would imply that pCO₂ is also estimated from ship data (and not directly measured). Please rephrase it.

l. 14: Regarding the term "statistically significant differences," if not statistically significant, perhaps "substantial differences" might be more appropriate.

I.14: It might be beneficial to consider removing the phrase "A preliminary analysis, ignoring other factors," as it could prompt questions about the factors being ignored and why.

I.17 What "other factors"

Figure 1: Could you please specify the time period covered? Is it from 2014 to 2023?

I. 82: missing word ("is proposed"). Missing comma: "Here,"

I.87: The Method section appears incomplete and should be expanded to include the methods used to obtain the results presented in the discussion section.

I.101: I would suggest merging the sentence to "Only data where the parameters nitrate, oxygen, and pH had a quality control (QC) flag of "1" equivalent to good quality according to GDAC were used here (Schmechtig and Thierry, 2016)"

I.104: Were data from 1972-2021 used?

I. 115: missing blank space before "For the ship data"

I. 124: interesting word choice "recommended for researchers". Are there more data recommended for other groups?

I.124 What data center? Consider rephrasing or clarify whether it is necessary to include this detail. Maybe rephrase to "The float pH data are calibrated:" Or is the data center relevant?

I.128: Could you provide more context for choosing 1500m?

I. 130: Please remove blank space

I. 130-134: The sentence appears incomplete, and the bullet points merely reiterate the equation. It may be clearer to remove the bullet points and rephrase as follows: "We adjusted the in-situ pH measured by floats according to the method outlined by Williams et al., 2017, using the equation [equation]."

I.135: While "Interrogating" may be technically correct, it sounds very awkward/formal. You might want to opt for a more common verb like Instead investigating

I. 138 Perhaps consider using "visualizing" instead of "elucidating"

I. 143: Add a comma after "However",

I.144: missing word (is suggested)

I.188: Please consider removing the first sentence as it lacks meaningful information.

I.188: add blank space (Fig. 2)

Figure 2: The reader can't see the blue float data points underneath the red ship data (another plots with reversed order seems redundant).

One option I would suggest would be removing the black outline of the markers for better visibility and increasing the transparency. This would also make Supplement Figure S1 redundant. Since Figure 2 and 3 are closely related, I would suggest merging it. Add the lines from Figure 3 in Figure 2.

Option 2 (preferred option): plot mean/both lines (red-blue and black) in one plot and add error bars.

I.195: I would suggest merging the 3 sentences: We observe that mean pCO₂ values derived from floats exceed those than from ships with highest differences occurring at depths shallower than 1500 m.

I.201-202: If the two values are that important, I would suggest adding them to the plot. I don't think it's necessary to announce how important they are here without explaining further. I would suggest removing both sentences.

I.202: I would suggest removing "as is apparent in Fig. 2" and just add "(Fig. 2)" at the end.

I.204: Why should we assume the data is incorrect simply because it exhibits high values? Shouldn't the flag 1 remove any data considered "incorrect"? The authors should not automatically assume poor quality and remove data simply because the values are unusually high, especially if the quality check did not flag them.

I.204: I suggest rephrasing it: Some floats recorded unusually high pCO₂ values, as pCO₂ values in the Southern Ocean typically range from X to Y (ref). However, considering the "good" quality classification of these data by GDAC standards, and the minimal impact on pCO₂ averages (<0.1 uatm) upon their removal, we opted not to exclude them from our analyses." Did you investigate where the high pCO₂ data were measured? What regions? Do the authors discuss the increased values later in the manuscript?

L.208: For consistency, I recommend rephrasing it: Increases/decreases in O₂ are often accompanied by decreases/increases in pCO₂ due to photosynthesis/respirations

I-208-210: Keep your sentences simple. I recommend rephrasing it, e.g.: Although ship-based and float-based O₂ values are similar, pCO₂ values derived from floats are higher than those derived from ships.

Figure 4: Same comment as for Figure 2. Remove this figure and add error bars to figure 5.

I.215-217 Please remove these (redundant sentences), as the information they convey is already visible in the figure.

I. 218: If the difference is not statistically significant, it may be more appropriate to describe it as "substantial" rather than "significant."

I.223: Points mask axes labels and best-fit lines. Please just add error bars.

I.225: belongs to discussion

I. 232: from “the” atmospheric equilibrium

I.249: I suggest explaining in the method section why the Drake Passage was chosen and consider examining the distribution of the data rather than solely comparing averages.

I.270: “is” reduced

I.271: By removing “a seasonal bias” or “seasonal biases”

I.283: remove “some”

I.289: “substantial” and not “significant”

I.290: Could you clarify why Lueker et al.’s constants were used instead of others?

I. 298: missing blank space

I.318: if you are not accounting for a spatial bias, then you can’t attribute the difference to data quality issues

I.325: “2. Seasonal bias in ship-collected data” I would suggest using another word than “bias”

I. 335 & 346: I recommend using "e.g., seasonal bias" instead of "factors 2 & 5". It's only a few more words but significantly improves clarity.

I.352: incomplete sentence

I.354-361: Consider removing as it contains repeated information.

I.385: 4.5 Implications might be more suitable for Introduction and/or Discussion/Conclusions

I.388: Implication (1) does not directly stem from the presented results, so consider revising it accordingly.

I.399: The sentence is misleading. Bushinsky et al 2019 conducted 4 μatm offset experiments/tested whether introducing a 4- μatm offset to the float $p\text{CO}_2$ estimates improves it. However, they found “that the fit of the mapped products to the observations cannot confirm or disprove the existence of a 4 μatm bias in the SOCCOM observations.” While Bushinsky did artificially reduce the float $p\text{CO}_2$ by 4 μatm , I would add that they did this as a test of sensitivity to a possible bias in float $p\text{CO}_2$

I.411: Please add the time periods of the data

I.426. Please add that you are talking about means

I.419 As previously stated, I do not believe this conclusion is supported by the analysis provided. The analysis appears insufficient to justify such a claim.

I.421 You propose that the calculation scheme of pCO₂ from float pH requires further refinement and that all sources of uncertainty should be considered, yet you do not provide suggestions for a refinement.