Review for 'Atmospheric CO2 estimates for the Miocene to Pleistocene based on foraminiferal δ11B at Ocean Drilling Program Sites 806 and 807 in the Western Equatorial Pacific'

The authors have done a good job at dealing with the previous round of reviewers comments on this manuscript. The discussion and conclusions are more nuanced and more appropriate for the data as presented in this iteration.

I am pleased to see that most, if not all the suggestions from the reviewers have been taken on board as they were important for the accuracy and longevity of this paper. However, I have one new major concern with the manuscript that requires addressing prior to publication, alongside several minor points which are listed below.

My major concern is with the implications from the cross plot presented in figure 3 (in particular E), these were added as part of the revision which is great, but they are not represented well in the text and the data need a significant reassessment. The r^2 of the 'validation' dataset presented in 3E is 0.09 and there is a p-value of .3 which is nonsignificant and cannot be used to reject the null-hypothesis. I understand the frustration of such figures, but they should be presented in the main test where the plot is mentioned and their ramifications discussed. Numerics like this would typically present difficulty for the rest of the study, but in this case I think there are some alternate options as I do agree that the data are probably better than these performed statistics would imply.

The authors should investigate other methods for assessing their ice core overlapping data (e.g. root mean square error etc.), perhaps focussing on identifying and removing outliers and further exploring the knock-on effects on the rest of the record. I.e. given the prediction error envelope (rather than just a fit error or an analytical and calculation error), what effect does that have on the Pliocene and Miocene CO_2 reconstructions? Is there an impact from using the Mg/Ca corrected script of Gray and Evans or the calibrations used (e.g. Raitzsch 2018 vs. Henehan 2013)? This is really crucial for our community understanding and confidence in the proxy data.

I am confused as to the point of the other two cross plots in figure 3, they are all mentioned in the test together but with little explanation as to why or what they show. The layout of the axes is also confusing, why is the same parameter plotted on the x axis in 3C and the y axis in 3d?

In the text (line 275) it is mentioned that only 2 of the data points in the validation dataset (n=16) are outside the uncertainty, this does not appear to be the case with the whole record in figure 3B or in 3E.

Minor points:

Line 33-35: There is a lot of detail here regarding the MPT in the abstract which is still not for me the major finding or thrust of the paper, would suggest slimming this down.

Line 42: some formatting inconsistencies throughout with italics on pCO2.

Line 49: Would 'Neogene' be a better key word than 'Miocene'

Line 70: In reference list like this which are incomplete add 'e.g.' in front.

Line 90: The TE part of the acronym TE-NTIMS is not defined.

Line 95: or if the disequilibrium is known. Not many of the sites used so far are in perfect equilibrium, but the stability is important.

Line 121: Is this 17ppm 1SD or 2SD?

Line 138: insert here if you applied the pH dependency.

Line 144: Hopefully all the figures are of merit, suggest changing to 'principal figures'.

Lines 176-182: In the discussion of age models it is apparent that the two sites have very different qualities of age model, what are the implications of this for age uncertainty?

Line 226: Was the JCp analysed clean or uncleaned?

Line 248: is the capital needed in Alkalinity?

Line 257: I am not clear precisely what time-adjacent means here, please be specific.

Line 264-265: This is a very important point, why can they be lower and yet produce similar estimates of CO_2? What is inside the calculation that yields this result? (2nd carbonate parameter?)

Line 270: n=16 here.

Lines 281: missing space between reconstructions and in

Line 312: odd bold formatting.

Line 339: Specify the subplot of Fig 6.

Line 344: Please point to the plot here. Greenop et al. 2017 (e.g. plot 6E)

Line 376-379: What is the point made here, who are we supposed to believe? Please detail the differences. The difference in the calculation of two papers from the same year is huge here >400ppm, and of crucial importance for the viability of the proxy.

Line: 461: MPWP format

Line 468: i.e.., format

Line 488: clarify ice sheet stability over multiple obliquity cycles, or just large sheet generation. They are still not inherently stable.

Line 529: reprocessing data?

Line 542: specify the MCO record produced in this study. Also another line of reasoning is required to explain the difference seen between sites, that cannot be due to the basalts unless you are implying a local effect/ teleconnection

Line 553: typically ESS or ECS not earth system climate sensitivity.

Line 817: format

Figures and captions:

Figure 1: Contours would be good to add to this plot, the rainbow colour bar is not good for colour blind accessibility.

Figure 3: It is unclear to me what 3C and 3D are supposed to show and why they have been plotted. More explanation required as above. Also it would be easier to see and trends if the $\delta^{18}O$ ruber were plotted on the same axis in both figures.

Figure 4: What is orange?

Figure 6: The ice core reference is not correct.

Figure 7: The site 872 data is calculated by Rae et al. 2021 not from it.

Figure 9: Again ice core reference is not correct.

Figure 10: Rae's compilation is cited throughout these figures and while I appreciate it is difficult to cite all the original references on the longer timescale plots, I think it is unfair to not credit the original authors on any of these plots. None of the data is created by Rae et al. e.g. Figure 10 should credit Bereiter, Hoenisch, Chalk, Dyez, and Yan studies.