UNIVERSITY OF CALIFORNIA, LOS ANGELES **UCLA**

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

DEPARTMENT OF EARTH AND SPACE SCIENCES
3806 GEOLOGY BUILDING
BOX 951567
LOS ANGELES, CALIFORNIA 90095-1567

TEL: (310) 825-3880
FAX: (310) 825-2779

Dear Dr. Fisher,

Thank you for considering our manuscript for publication in Climate of the Past (manuscript # cp-2020-158). We believe we have addressed the points raised by the reviewers.

We explore multiple regression methods and statistics on the comparison between ice core pCO_2 and our $\delta^{11}B$-derived pCO_2. We revised this section of the manuscript accordingly. We provide more context to the Li isotope comparison, on the hypotheses for the Miocene climate transition and the discussion of the Columbia River basalts and the Miocene climate optimum.

We want to thank the reviewers for their constructive comments.

Sincerely,

Maxence Guillermic and Aradhna Tripati on behalf of the co-authors

Dear authors

the revised version of your manuscript has now been seen by the two referees that also assessed the initial version. Both referees acknowledge the significant changes that you did in the revision, which over most parts meet the points of criticism initially raised by both the reviewers. Reviewer #2 has added a few more constructive suggestions for improvement in the second review, in particular on the long-term volcanic and weathering carbon imbalance, which I would ask you to include in your re-revisions. Both referees also make a few comments on language/textual improvements etc. (see also my editorial comments below). I would regard these points raised by the referees as minor revisions as they imply mainly textual changes and can be implemented in a straightforward manner.

However, referee #1 also raises a point on the overall accuracy/validity of your pCO2 reconstruction based on the validation against the ice core data and I have to admit this gives me some headache as well. According to the figure caption of Figure 3E your r^2 is only 0.09, i.e. this would imply that only 10% of the variance in your d11B-derived pCO2 could be explained by atmospheric CO2 as contained in ice cores. Is this correct or is the r^2 of 0.09 only a typo? Looking at this figure, I visually can see a clear correlation although with a regression slope larger than 1.

In the main text you also refer to some p values, but these numbers do not agree with the ones in the caption and I am unable to attribute the numbers to the respective subplots. Please double check all the numbers and change the text to make things crystalclear. Moreover, it is unclear to me what the uncertainties in the x and y component of the scatter plot 3E refer to.

In fact, some of the ice core CO_2 data (x-axis) have a larger uncertainty than the d11B-derived data points (y axis) which appears virtually impossible as long as this does not include dating uncertainties.

Please also mention in the caption which evaluation/corrections scenario you used for the d11B-derived pCO2 in Figure 3E.

Response 1: Following yours and the reviewers suggestions, we:

- Rechecked, and confirmed the original linear regression performed on the data gave an R^2 of 0.09 and a p-value of 0.25.
- Performed both an ordinary least squares regression and a Deming regression followed by bootstrapping (n=1000) in order to take into account the uncertainties in both ice core CO_2 and δ^{11}B-derived CO_2. The slope and intercept are still within error of the 1:1 line, but the p-value is however still low and this is likely due to the limited resolution of the data. We have rewritten this section of the manuscript.
- We removed the p-values comparing the slope and intercept to the 1:1 line which may have induced some confusion.

We note the uncertainties in pCO_2 from ice core were estimated based on the 2 SD of ± 1 kyr of the age from the sediment sample. Uncertainty for ice core pCO_2 is only larger than derived from boron isotopes at 227 ka because +/- 1 kyr in the record is tracking a steep pCO_2 transition. The uncertainty for the boron based data are from propagated uncertainties. Data include dating uncertainties, 2 SD of +/- 1 kyr.

Finally, your linear regression is suggesting a slope smaller than 1 although the range in the d11B-derived pCO2 is larger than the range in ice core CO2!!!??? This is likely due to you using a one-sided regression, which in case of considerably uncertainties in x and y underestimates the slope. In fact, you should use a 2-sided regression (York et al., American Journal of Physics, 2004) for this scatter plot. It is also puzzling that 50% of the data points do not fall within the prediction envelope, what is the error this envelope refers to? In any case, for the validation period of the last 800 kyr your data seem to be clearly related to atm. CO2 but seem to overestimate somewhat the true atmospheric CO2 concentration range either due to methodological, correction/evaluation, diagenesis or CO2 surface water saturation issues. Although this systematic overestimation must not hold back in time, it is important to openly communicate that the absolute levels of your pCO2 reconstruction may be biased. However, the relative changes on time scales of 1 Myr, where alkalinity changes do not play such a great role may be more reliable. The issue of uncertain absolute CO2 levels in the past is also nicely illustrated in your evaluation scenarios, which show that the overall Miocene CO2 levels strongly depend on the assumptions made during the evaluation/correction of the data in particular on alkalinity. This is clearly an issue the d11B community will have to work on in the future.

Response 2: Thank you for this comment. In order to develop this section we now present in Fig. 3 a Deming regression with a bootstrapping method taking into account the uncertainties for both x and y axes. You can also find a table summarizing the outputs of the regressions (linear and Deming) in Table S6. The p-value is still not significant ($p=0.25$). However, the absolute values for reconstructed $\delta^{11}B$ overlap with the ice core record and the amplitude of glacial-interglacial variability is reproduced. Having a larger n with a higher resolution record would make this comparison with the ice core record more statistically robust, as would having more robust constraints on temperature.

In summary, I think the low correlation/weak validation of your data for the ice core period does not justify rejection of the manuscript as the paper, nevertheless, adds substantial information on pH and CO_2 changes over the last 20 Myr. However, I strongly feel that the discussion should be more open to the limitations of the reconstruction and the large uncertainties in the overall reconstructed CO2 levels. This is in my eyes overly optimistic in the current version. A more critical assessment of the overall levels especially in the conclusions and the abstract, would likely also satisfy the criticism of referee #1.

Response 3: We have revised this section as suggested, and the abstract and conclusions are also revised. The main section in the text now reads Line 261: "We sought to assess if there is evidence for air-sea equilibrium or disequilibrium in the WEP during the large amplitude late Pleistocene glacial/interglacial cycles, in order to validate our approach. We reconstructed pCO_2 for the last 800 kyr (n=16, Fig. 3). For the last 800 kyr, reconstructed pCO_2 values for Sites 806 and 807 are in the range from ice cores (Fig. 3, Petit et al., 1999, Siegenthaler et al., 2005, Lüthi et al., 2008; compilation from Bereiter et al., 2015). The two critical diagnostics we used for method validation are: 1) that the $\delta^{11}B$-based reconstruction of pCO_2 is consistent with ice core atmospheric CO_2 and 2) the boron-based reconstruction empirically reproduces interglacial-glacial amplitudes from ice cores. Fig. 3B shows that both of these criteria are met. We also created a crossplot comparing these two independent constraints on pCO_2 (Fig. 3C). Two regressions between ice core pCO_2 and boron-based pCO_2 are shown, with a simple linear regression (grey line) and a Deming regression that factors in error in variables (blue line), and bootstraps of outputs shown (n=1000, Figure 3C, Table S6). While slopes and intercepts are not statistically different from a 1:1 line, the regressions do not reach a high significance level ($p=0.25$); boosting the resolution of the record could help provide better constraints for this type of comparison. The age models for the site do not provide an explanation for this variability based on comparison of the benthic $\delta^{18}O$ records for both Sites 806 and 807 (Fig. 3A, Zhang et al., 2007; Lear et al., 2003; Lear et al., 2015) to the published isotopic stack (Lisiecki and Raymo, 2004). No significant difference in variability was observed at either site. We also note that reconstructed pCO_2 uncertainties (both accuracy and precision) could potentially arise from Mg/Ca-derived estimates of temperature; these uncertainties could be reduced using independent temperature proxies for the WEP such as clumped isotope thermometry (Tripati et al., 2010; 2014), a technique that is not sensitive to the same sources of error as Mg/Ca thermometry, and therefore is an area planned for future work. Other sources of uncertainty that have a larger effect on pCO_2 calculations are the weight/shell correction, while the TA and seawater boron isotope composition have a minor effect over this time interval."

I therefore would urge you to include such a more differentiated discussion in the re-revised version. Please add a point-to-point reply to the points raised by the referees and me in this second round of reviews and a track change version of any re-revised manuscript relative to the current revised version. Your replies will decide whether another independent review of the paper is needed. Please find below a few more technical editorial comments that I would ask you to include as well:

- at several instances in the paper you refer to the ice core data as "Vostok data". This is incorrect as only the data younger than 420 kyr are from Vostok, the older data is from the EPICA Dome C ice core. Please change this throughout the manuscript by just referring to "ice core data" and providing the relevant references (Petit et al., 1999, Lüthi et al., 2008, Bereiter et al., 2015). In the main text a reference such as "as compiled by Bereiter et al., 2015" may be sufficient in many cases but at least in the figure captions the original data references should be provided. A similar point was raised by referee #1 for the d11B compilation by Rae et al. Again, provide the full original references at least in the figure captions. Note also that in several figures you refer to the ice core data as Bereiter et al, but this data set only covers the last 800 kyr. Earlier ice core data comes from blue ice samples. Please correct accordingly and provide the correct references (see also comments on Figures below)

Response 4: We changed "Vostok" to "ice core" in the main text. We also changed in the Fig. 3 legend, and axis and acknowledge the original publications. We also added the references from Higgins et al. (2015) and Yan et al. (2019) when those data were presented.

- in line 118 you refer to an air sea flux of xx ppm, this is not a flux but an air sea difference

Response 5: We changed for air sea difference. Line 114: "annual air-sea CO_2 difference of +28 ppmv (Takahashi et al., 2014). The pre-industrial air-sea CO_2 difference is calculated to be +16 ppm, (GLODAP database corrected from anthropogenic inputs),"

- line 172: there is something wrong with this sentence

Response 6: We changed line 166: "Further, despite different sedimentation rates, our $\delta^{11}B$ and Mg/Ca results are consistent between Sites 806 and 807 , and with data from Site 872 (Sosdian et al., 2018), which implies that diagenesis is not a primary driver of the reconstructed trends. Comparison of raw data, and derived parameters, is shown in Figs. 2 and 7.

- throughout the manuscript add a space between numbers and units

Response 7: We changed this through the text.

- line 411. Here you refer to a pCO2 peak at 9 Myr in Sosdian. I couldn't find this point, maybe refer to the age more precisely?

Response 8: Sosdian et al. (2018) reported a peak at ~9 Ma, this peak is discrete and relatively small in comparison the MCO CO_2 concentration.

In section 3.4 from Sosdian et al. (2018): «Regardless of the chosen $\delta^{11}B_{sw}$ scenario, all our CO_2 reconstructions shows CO_2 peak at ∼9 Ma that is not seen in the climate records (Fig. 5). This peak in CO2 at ∼9 Ma is also seen in recent cell-sized corrected alkenones and pennate diatoms (diffusive and ACTI-CO models) CO_2 records (Bolton et al., 2016; Mejía et al., 2017; Supplementary Fig. S12). »

- line 517: there is something wrong with this sentence

Response 9: We changed from "Before 18.5 Ma, the pCO_2 is relatively stable, δ^7Li is increasing representative of a non-steady state weathering. From 18.6 to 16.7 the δ^7Li decrease of about 2 ‰, this decrease can inform on decreasing weathering rate and this decrease is associated with an increase in pCO_2." to line 540: "Before 18.5 Ma, the pCO_2 is relatively stable, δ^7Li is increasing, suggesting non-steady state / incongruent nature of continental chemical weathering. From 18.6 to 16.7 Ma, the δ^7Li record decreases by ~2 ‰, consistent with decreasing weathering rates and an associated increase in pCO_2."

Figure 2: There are brown triangles in figure 2B and a cross in figure 2C but not in the legend

Response 10: Brown triangles are data from *G. ruber*. It is in the figure caption from Fig. 2 "brown open symbols are for *G. ruber*."

Figure 3: see comments on Figure 3 above and in review #1. Please delete the subscript Rae et al in the legend and cite the data and compilation papers in the caption

Response 11: We are now calling in the legend the compilation from Sosdian et al. (2018) "compilation A" and the one from Rae et al. (2021) "Compilation B", references from those compilations are now added in the caption."

Figure 4: there are orange circles in this figure but not in the legend

Response 12: Caption of Figure 4: "Orange open circles are SST data calculated with our framework from the species *D. altispera* at ODP Site 806 (Sosdian et al., 2020) with an offset of +8°C."

Figure 5: Please delete the subscript to individual papers in the legend and cite the data and compilation papers in the caption

Figure 6: Please delete the subscript to individual papers in the legend and cite the data and compilation papers in the caption. There are bold and thin triangles/squares in the figure but not in the legend.

Responses 13 and 14: We added the original references in the figure legend for compilation A (Sosdian et al., 2018) and compilation B (Rae et al., 2021).

Figure 7: Please delete the subscript to individual papers in the legend and cite the data and compilation papers in the caption. There is no legend for the SST and d11B-pH subplots

Response 15: We are now calling in the legend the compilation from Sosdian et al. (2018) "compilation A" and the one from Rae et al. (2021) "compilation B" in the legend, references are described in the caption.

Figure 8: Please delete the subscript to individual papers in the legend and cite the data and compilation papers in the caption. There is no legend for the SST subplot

Figure 9: Please delete the subscript to individual papers in the legend and cite the data and compilation papers in the caption. There is no legend for the SST subplot

Figure 10: Please delete the subscript to individual papers in the legend and cite the data and compilation papers in the caption. There is no legend for the SST subplot

Figure 11: Please delete the subscript to individual papers in the legend and cite the data and compilation papers in the caption.s.

Response 16, 17, 18 and 19: We deleted the subscript from those figures and added legends for temperature.

Reviewer 1

Overall I would like to praise the authors for the careful revision of the figures which has really clarified the presentation of the new data and its comparison with previously published records. The text now reads more clearly in many sections as well. There are a few final issues I would propose the authors clarify, and some proofing issues that I identify.
The text at the end of methods (lines 242-243) summarizes the processing of Mg/Ca to SST, noting that it is corrected for influences of pH and salinity. However, additionally, the SST has been corrected for inferred changes in the Mg/Ca of seawater (detailed in the supplement). This point also needs to be mentioned in lines 242-244 of the main text, so the reader is aware of it, and is directed to the supplement for further details.

Response 20: We added line 243: "Detailed calculations can be found in the supplemental materials. Briefly, Mg/Ca was used to reconstruct sea surface temperature (SST) using the framework from Gray and Evans. (2019) correcting for influences of pH, salinity and secular variation in seawater Mg/Ca.".

The figures include detailed dashed lines to show correspondance to specific events such as MMCT defined from the benthic d18O stack of Zachos 2008 (and in younger part Lisieki-Raymo). I would suggest also including the low resolution benthic 18O from Lear et al 2015 from 806, to help document the comparison with these events within the same site, because the benthic 18O at 806 and the d11B-based CO2 and SST can therefore be compared within 806 without the complication of uncertainty in age models due to the normal limitations of precision in biostrat.

Response 21: We added the benthic $\delta^{18}O$ from Lear et al. (2003) and (2015), reprocessed in 2020. Even if lower in resolution the record is in line with the benthic compilation from Lisiecki and Raymo, (2005) and Zachos et al. (2008) on the different timescales studied here which supports our age model for 806 and our interpretations (Figure S3). The linear regressions for the cross plots Figs. 3C and D are however still non-significant.

line 406 makes a rather broad statement about the shift in $\delta^{13}C$ and $\delta^{18}O$ over the MMCT supporting drawdown of CO_2 by organic carbon burial, with no reference. As these benthic data were previously published and interpreted, and this coincidence and organic burial hypothesis advanced long ago, I would suggest referencing a number of papers which have produced the benthic $\delta^{13}C$ and $\delta^{18}O$ results and originally provided this interpretation, to support your link of organic carbon burial with CO_2 drawdown. The recent discussion paper in Climate of the Past by Raitzsch et al. provides detail over this interval and may be worth citing as well.

Response: 22: We are now referencing a couple of relevant papers associated with the middle Miocene climate transition. We also added the pCO_2 from Raitzsch et al. (2021) in Figure 8. Because our resolution is low we stress to overinterpret our data for this transition however we discuss finding from Raitzsch. (2021).

Line 403: "The synchronous shifts in the $\delta^{13}C$ and $\delta^{18}O$ of benthic foraminifera are consistent with increased carbon burial during colder periods, thus feeding back into decreasing atmospheric CO_2, and supporting the hypothesis that the drawdown of atmospheric CO_2 can in part, be explained by enhanced export of organic carbon (Flower and Kennett, 1993, 1996). However, given the limited sampling of this study, we are only able to resolve a pCO_2 decrease toward the end of the MMCT (~13.5 Ma). The higher resolution $\delta^{11}B$-pCO_2 from Site 1092 for the MMCT (Raitzsch et al. 2021) reports eccentricity-scale pCO_2 variability; the authors reported that low pCO_2 during eccentricity maxima was consistent with an increase in weathering due to strengthened monsoonal circulation, which would increase nutrient delivery and supporting higher productivity that in turn would impact carbon drawdown and burial, in line with modeling from Ma et al. (2011)."

The section 3.6 makes a few general statements that could be better balanced with context and caveats. The linking of the MCO pCO2 increase with the Columbia River is interesting but I think it is better balanced to comment that this is superimposed on temporal variation of the large volcanic CO2 flux from seafloor production, which has been estimated by recent tectonic models with increasing precision (eg Muller et al 2016). Likewise the estimated volume and CO2 flux from the Columbia River Basalts is very modest compared to other LIP in compilations (Courtillot et al 2003), suggesting that if it is contributing to high CO2 at the MCO, a better understanding of eruptive volume and volatile emissions is needed.

Likewise a bit of clarification about the d7Li is required - eg. note that the authors favored alkalinity model is Caves et al, but Caves is discussing specifically why d7Li is not necessarily a proxy for silicate weathering flux. I agree that it is interesting to note the coincidence of inflections in the d7Li curve and features of the pCO2 record, but I find the current lines 508-524 have gone too far in oversimplifying the d7Li and a few sentences to provide better context to the interpretation of d7Li and debates would make this final section much more useful.

Response 23 and 24: Thank you for those references and these comments. We incorporated your comments in section 3.6.

Line 498: "On million-year timescales, atmospheric CO_2 is controlled by its input through mantle degassing in the form of sub-aerial and sub-aqueous volcanic activity and its removal by chemical weathering of continental silicate rocks. Over the last 16 Myr, two relative maxima in atmospheric pCO_2 are observed in our record, one during the MCO (at 15.67 Ma) and a second around the late Miocene/early Pliocene (beginning at 4.7 and 4.5 Ma) (Fig. 11), though the timing for the latter is not precise. The strong pCO_2 increase from the early Miocene to MCO is timely with increasing volcanic activity (Foster al. 2012), associated with the eruption of the Columbia River Flood Basalts (Hooper et al., 2002; Kasbohm and Schoene, 2018), with recent geochronologic evidence published supporting higher eruption activity between 16.7 and 15.9 Ma (Kasbohm and Schoene, 2018) reinforcing the idea of an episodic pCO_2 increase during the MCO due to volcanic activity. Underestimation of net CO_2 outgassing from specific continental flood basalt eruption is possible, as both sub-aqueous and sub-aerial flood basalts, under right climatic conditions, are prone to enhanced chemical weathering. For example, the 4-5‰ drop in δ^7Li record at the K-Pg boundary (Misra and Froelich, 2012) is attributed to rapid quasi-congruent weathering of Deccan Traps (Rene et al. 2015) during their eruption. Courtillot and Rene (2003) estimates about 50% of emitted CO_2, roughly equivalent to the amount emitted by the eruption of a million cubic kilometers of Deccan Traps, is missing due to chemical and physical weathering. Additionally, the early Eocene (at ~50 Ma) 3-4‰ rise in seawater δ^7Li at a time where there is not significant uplift of the Himalayas (Misra & Froelich, 2012) is also attributed to incongruent weathering of previously erupted Deccan Trap basalts as the Indian subcontinent moved from arid mid-latitudes to the wet low latitudes (Kent and Muttoni, 2008, PNAS). Thus, a significant part of the outgassed CO_2 can be consumed by chemical weathering of freshly erupted hot basalts (Courtillot et al., 2003). However, the congruency of chemical weathering of basalts, depending on regional climatic conditions (warm-wet vs. cold-arid), will determine the nature of observable inflection in the seawater δ^7Li. The possible quantification of increased rates of silicate weathering inferred from δ^7Li (mentioned below) can be utilized to determine total eruptive volume (missing + existing) and volatile emissions from the Columbia River Flood Basalts. At the same time as continental flood basalt emissions, enhanced seafloor production could also be a second possible source of CO_2; however, we note there is evidence that the rate of seafloor production has remained virtually invariant over the last 60 million years (Rowley, 2002; Muller et al. 2016)."

Detail: Lear 2015 is missing in the reference list (it is cited in text line 179).

Response 25: We added the reference to the list.
Detail: Supplemental Information (there is no "s" needed at the end of "Information")

Response 26: We removed the "s".
Detail: the citation of Tanner et al 2020 in line 360 does not make sense as that sentence refers to MMCT and Tanner et al cover only the Late Miocene after 8 Ma. The citation in line 414 is appropriate although the proposed reason for differences is unclear attributing it to " the oceanographic setting of 1088" without indicating what process about the oceanographic setting is proposed to explain the difference.

Response 27: We removed the citation of Tanner from line 360.

Line 421: "pCO_2 differences between our reconstruction and that of Sosdian et al. (2018) and Raitzcsh et al. (2021) (Fig. 8) likely reflect assumptions made for calculations (of δ^{11}B, TA) and the specific mono-specific calibrations used for each study, as well as potential geographic differences in air-sea pCO_2. These differences do not invalidate the boron isotope proxy but illustrate the impact that specific seawater parameters and calibrations can have on reconstructed pCO_2 values."

Detail - line 436 either "during" or "from" - one of these words needs deleting

Response 28: We changed line 436 for: "We calculate high pCO_2 values of 419 ± 119 ppm (2 SD, n=3, Table 2) between 4.7 to 4.5 Ma during the Early Pliocene warm interval (Figure 9)."

Detail - line 483-484, use consistent formatting for the +/- values of CO_2.

Response 29: We removed the parentheses to be consistent with the rest of the manuscript.

Raitzsch, Markus, Jelle Bijma, Torsten Bickert, Michael Schulz, Ann Holbourn, and Michal Kučera. "Atmospheric carbon dioxide variations across the middle Miocene climate transition." Climate of the Past 17, no. 2 (2021): 703-719.
Müller, R. Dietmar, Maria Seton, Sabin Zahirovic, Simon E. Williams, Kara J. Matthews, Nicky M. Wright, Grace E. Shephard et al. "Ocean basin evolution and global-scale plate reorganization events since Pangea breakup." Annual Review of Earth and Planetary Sciences 44 (2016): 107-138.
Courtillot, Vincent E., and Paul R. Renne. "On the ages of flood basalt events." Comptes Rendus Geoscience 335, no. 1 (2003): 113-140.

Reviewer 2

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 nullhypothesis.

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 CO2 reconstructions?

Response 30: We replied in a previous comment, but we used also a Deming regression to take into account the both x and y uncertainties (Fig. 3C). The statistics do not improve (p=0.25). To make sure we didn't have a gap due to the age models we are now plotting in Fig. 3 the benthic record from site 807 (Zhang et al., 2007) and from 806 (Lear et al., 2003, 2015, 2020). Those records are consistent with each other which exclude an issue with the age models (Fig. S3). When plotted together, data from Chalk et al. (2017) and Hönisch et al. (2009), we can see in that our data present more variability (Fig. S4A), this variability doesn't seem characteristic of one site (Fig. S4A) but could be due to the reconstruction itself (weight/shell correction). We still are able to reproduce the interglacial/glacial variability which is encouraging. We do not think those results impact our paper which focus on a low resolution pCO_2 across the Neogene but a high resolution IG/G cycle would be needed to identify if the variability is truly associated with the reconstruction or if it representative of a localized CO_2 change and this is something we will probably explore in the future.

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.

Response 31: The sensitivity of $\delta^{11}B_{carbonate}$ to $\delta^{11}B_{borate}$ is really similar (0.82 ± 0.22) reprocessed by Henehan et al. (2016), (0.80 ± 0.17) for Raitzsch et al. (2018) and (0.83 ± 0.44) here from (Guillermic et al., 2020). The Mg/Ca from Gray and Evans does not play a significant role in our record because only *G. ruber* is impacted by pH not *T. sacculifer* (no evidence for now) from which most of the record is derived.

For the Miocene, alkalinity is actually of main importance but also $d^{11}B_{sw}$. Rae et al. (2021) used a constant alkalinity and an average of the 3 seawater $\delta^{11}B$ scenarios.

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?

Response 32: We removed those cross-plots as they were not adding information to the manuscript, instead we present other cross-plots in Figure S4 to add information on the differences between ice core and boron isotopes derived pCO_2.

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.

Response 33: We removed this sentence as the comparison with ice core data is better illustrated with the Figs. 3C and S4.

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.

Response 34: Line 32 "During the Mid-Pleistocene Transition there is a minimum in pCO_2 at MIS 30."

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

Response 35: We changed for "pCO_2" and it is now consistent through the text.

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

Response 36: We provide new data until the past 16 Myrs even if the Miocene is included in the Neogene, we don't want the reader to think we provide measurement for the early Miocene.

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

Response 37: We added a couple more recent references in order to complete the list, we still put the 'e.g.' in front of the references.

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

Response 38: Instead of TE-NTIMS (Ni et al., 2010) we changed for N-TIMS as the pCO_2 reconstructions are based on N-TIMS or MC-ICP-MS data. We changed line 74 for "N-TIMS" instead of "TIMS" and Line 85: "The marine CO2 proxy that appears to be subject to the fewest systematic uncertainties, based on our current understanding, is the boron isotopic composition ($\delta^{11}B$) of planktic foraminifera as measured using MC-ICP-MS and N-TIMS (Hain et al., 2018)."

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.

Response 39: We added line 91 "Atmospheric pCO_2 can then be constrained if the site being examined is in air-sea CO_2 equilibrium or if the disequilibrium is known and stable through time.".

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

Response 40: It is 2SD, line 117: "an average of ~17 ppm (2 SD) for the youngest samples".

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

Response 41: Line 132: "For temperature estimation, we utilize a multi-variable model for Mg/Ca correcting from salinity, pH and seawater Mg/Ca (Gray and Evans, 2019) […]".

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

Response 42: We changed for "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?

Response 43: To explore this, we now provide in Figure 3A the comparison between benthic $\delta^{18}O$ from site 806 and 807. Both age models are based on biostratigraphy but are in line with each other (only 0.55 Myr constrained) and the stack from Lisicki and Raymo, 2005.

Line 226: Was the JCp analysed clean or uncleaned?

Response 44: The JCp-1 was uncleaned.

Line 248: is the capital needed in Alkalinity?

Response 45: we removed the capital.

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

Response 46: We changed for line 250: "Although the record we generated does not overlap with Site 872, they are 1 myr apart (15.7 and 16.7 Ma); there is a good correspondence between our Mg/Ca data and the published Mg/Ca record from *T. trilobus* at Site 872 (Sosdian et al., 2018). Mg/Ca from a different species, *D. altispira* (Sosdian et al., 2020), is also plotted with an offset, for comparison."

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?)

Response 47: There is a strong size-dependence at site 806 for $\delta^{11}B$, reported in Hönisch and Hemming, 2004 and Ni et al., 2007, this is why we applied a size (weigh/shell)-$\delta^{11}B$ correction in order to adjust the intercept of the calibration for each individual and correct from this effect. This is a parameter that is not propagated in the error.

Line 270: n=16 here.

Response 48: We added line 263: "for the last 800 kyr (n=16, Fig. 3)."

Lines 281: missing space between reconstructions and in

Response 49: We added the space.

Line 312: odd bold formatting.

Response 50: We fixed this.

Line 339: Specify the subplot of Fig 6.

Response 51: We added the subplots Figs. 6B and 6E.

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

Response 52: We added the subplot 6E. Line 342: "by Greenop et al. (2017) (e.g. Fig. 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.

Response 53: Rae et al., 2021 used a constant alkalinity + an average of the three $\delta^{11}B_{sw.}$

This study used the alkalinity from Caves et al. (2016) + $\delta^{11}B_{sw}$ from Greenop et al. (2017).

Even if depending on the timescale studied the assumptions are reasonable between studies, difference in absolute values can be important.

Line 375: "because of the different assumptions used in their calculations. This difference is important because the assumptions from Rae et al. (2021) would imply a relatively high and stable pCO_2 from the early Miocene to MCO (Fig. S2), which would imply a decoupling between pCO_2 and temperature with no pCO_2 change during an interval of decreasing benthic $\delta^{18}O$. However, our reconstructed pCO_2 increase towards the MCO is in line with the observed benthic $\delta^{18}O$ decrease and $\delta^{13}C$ increase and suggest a coupling between temperature and pCO_2 over this period. This highlights the critical need for the use of a common set of assumptions for studies. Assumptions may vary between studies depending of the timescales studied, but a common framework is needed. In addition, further constraints on the second carbonate system parameter and on secular changes in seawater $\delta^{11}B$ will reduce uncertainties in reconstructed pCO_2, with improved precision."

Line: 461: MPWP format

Response 54: We changed for "mPWP".

Line 468: i.e.., format

Response 55: We changed for :" (i.e, ~100 kyr)".

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

Response 56: Line 494: "In our record for the last 16 Myr, the lowest pCO_2 is recorded at MIS 30 during the MPT, with values of 164 ($\pm^{44}_{35}$) ppm, which supports an atmospheric CO_2 threshold that leads to large sheet generation. During this transition, the pCO_2 threshold needed to build sufficiently large ice sheets that were able to survive the critical orbital phase of rising obliquity to ultimately switch to a 100 kyr world, was likely reached at MIS 30, but a higher pCO_2 resolution of the MPT is needed for confirmation."

Line 529: reprocessing data?

Response 57: We changed for "reprocessing".

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

Response 58: We developed this section following comment from reviewer 1. I guess our perspective on this is that variability between Sites are obscured by the assumptions of the reconstruction themselves. However, reprocessing the data with a common framework as expected from the boron workshop may highlight significant differences between Sites.

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

Response 59: We changed for "Earth climate sensitivity".

Line 817: format

Response 60: We fixed the format of the reference.

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.

Response 61: Figure 1 now have contours.

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 δ18O ruber were plotted on the same axis in both figures.

Response 62: Instead different of those figures we are now presenting other cross-plots to add information on the comparison between ice core and boron based pCO_2 (Figure S4).

Figure 4: What is orange?

Response 63: We added to the legend and details in the caption of Fig. 4, orange open circles are the reprocessed data from Mg/Ca of *D. altispera* from Sosdian et al. (2020) following our framework, those date are however plotted with an offset (8°C) calculated to match overlapping data from *T. sacculifer* Site 806 (this study).

Figure 6: The ice core reference is not correct.

Response 64: We changed Bereiter et al. 2015 for "Yan et al. 2019".

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

Response 65: We changed for: "asterix symbols are calculated pCO_2 at site 872 by Rae et al. (2021)."

Figure 9: Again ice core reference is not correct.

Response 66: We changed Bereiter et al. 2015 for "Yan et al. 2019".

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.

Response 67: We agree, we now present the original references in the different captions. "Data for compilation A are from: Hönisch and Hemming, 2009; Seki et al., 2010; Foster et al., 2012; Badger et al., 2013; Greenop et al., 2014; Martinez-Boti et al., 2015a; Chalk et al., 2017; Sosdian et al., 2018. Data for compilation B are from: Foster et al., 2008; Hönisch and Hemming, 2009; Seki et al., 2010; Foster et al., 2012; Badger et al., 2013; Greenop et al., 2014; Martinez-Boti et al., 2015a; Chalk et al., 2017; Dyez et al., 2018; Sosdian et al., 2018; Greenop et al., 2019; de la Vega et al., 2020."

Figure 10 should credit Bereiter, Hoenisch, Chalk, Dyez, and Yan studies.

Response 68: We now cite all these references.