Review No 1 (anonymous)

Raitzsch et al present a new planktic boron isotope record across the mid Miocene Climate Transition, and interpret this in terms of changing CO2 and carbon cycle changes. The record enables deeper interrogation of potential climate and carbon cycle feedbacks during this important interval of global cooling. This interval also spans the end of the Monterey Excursion, which contains the well-known "carbon maxima" events. Having more CO2 data across this interval is very exciting, and thus the submission represents a substantial contribution worthy of publication in CP.

However, the narrative of the manuscript could be much improved – e.g. the conclusions are that a particular carbon cycle model is supported by the new records, but this model (or the mechanisms within it) are not referred to in the introduction. Much of the discussion hinges on assumptions of processes that have not been observed but are presented here as fact (e.g., that middle Miocene CO2 variations are caused by changes in shallow water carbonate production). The manuscript also fails to cite some key recent studies. However, I believe that if the narrative of the paper is improved and a few specific scientific questions are addressed then the robust interpretations of this fantastic dataset will be clearer and this will become a really exciting and useful contribution. Below I list the specific scientific questions I would like to see addressed, followed by more minor comments.

AC: We thank the referee for his constructive and thorough review of our manuscript. It was really helpful to look at our data from different perspectives discussed in more detail below.

1. Is high eccentricity definitely associated with decreasing CO2? Would it be more robust to focus on the clear relationship between the d13C and CO2?

The key finding presented in the abstract of the paper is that "long-term pCO2 variations between ~14.3 and 13.2 Ma were paced by 400 k.y. eccentricity cycles, with decreasing pCO2 at high eccentricity and vice versa." I struggle to see this relationship in Figure 4 or Figure 5 – i.e. an inverse correlation between the pCO2 record and the thin grey line. I'm not saying the pCO2 record doesn't contain a long eccentricity signal, but the temporal relationship with the eccentricity curve could perhaps be demonstrated more robustly. While I struggle to discern a negative correlation between the authors' pCO2 record and the long eccentricity cycle, the relationship that does seem convincing is the positive relationship between d13C and CO2 across CM6 (Figure 4 and described in section 3.1). Perhaps therefore a clearer and simpler approach would be to set out different mechanisms in the introduction with their respective d13C-CO2 relationships. E.g., Holbourn et al 2007 ascribes the eccentricity signal in the longer d13C record to EITHER increased productivity and burial of Corg (the classic Monterey hypothesis) which would result in a negative correlation between d13C and CO2, OR a monsoon-driven increase in shallow water carbonate deposition, removing alkalinity from the oceans and releasing CO2, producing a positive correlation between d13C and CO2, as observed here. This explanation would be particularly useful because the increase in monsoon intensity variability has different impacts on the carbon cycle, but these are not discussed in the introduction. For example, an increase in nutrient delivery to the oceans and increasing Corg burial tends to increase d13C and decrease CO2. But in the Ma et al 2011 model this is outweighed by the concomitant increase in carbonate burial and associated CO2 release. It would be useful if this current manuscript could comment on the robustness of the relative importance of these processes in the model. However, see also my comment below about the uniqueness of CM6 - what might be expected to happen if for example the climate transition led to a general increase in marine productivity superimposed on these orbital variations? Should we expect the d13C-CO2 relationship for CM6 to hold true for all CM events? This complication should be clearly addressed in the introduction also.

AC: We still think that both d13C and pCO2 are linked to eccentricity variations, but there is a phase lag of d13C and pCO2 to long eccentricity in the order of 50 k.y., which is possibly related to the slow response of weathering feedbacks to orbital forcing. This is now clearer from the text (l. 15-16; 222-225; 329-331; 403-404). Also we further emphasize the close relationship between d13C and pCO2 (l. 13-15; 221-222; 256-257; 326-327). However, we also slightly changed the title to not declare eccentricity as the only possible mechanism influencing our record. The introduction has been restructured to better explain the different hypotheses and the different d13C-pCO2 relationships associated with them (l. 35-51), also including the new study by Sosdian et al. (2020). In addition, we introduce the reader to the uniqueness of CM6 and the effect of Antarctic glaciation on the carbon cycle (l. 52-64), which also includes the recent study by Leutert et al. (2020) and Ma et al. (2018). Moreover, as the model of Ma et al. (2011) we refer to was designed for an ice-free world, we discuss the potential effect of Antarctic ice-sheet expansion on the carbon cycle in more detail (l. 358-372).

2. What is the evidence for the increased shallow water carbonate deposition?
As far as I know there is no direct evidence for eccentricity paced changes in shallower water carbonate burial. There are dissolution cycles in deep sea carbonates, but I am not aware of any study that rules out the influence of bottom water mass ventilation on these? I believe Holbourn et al 2007 favoured the monsoon hypothesis due to the 50 kyr lag between eccentricity and d13C. But this lag is not discussed in this manuscript, and no supporting evidence is given for the statements that shallow water carbonate burial changed at this time.

AC: We provide a number of references that deal with regional studies of increased carbonate deposition, but also mention that it is difficult to assess global trends from those (l. 342-348).

3. Does the Site 1092 planktic d11B record global pCO2?
This relationship between d13C and CO2 is key, and is independent of age model issues because the d13C has been recorded at each site and is a global signal. So it is interesting therefore that the Malta record shows a decrease in d11B-derived pCO2 associated with the onset of CM6 (increasing d13C) whereas the 1092 d11B-derived pCO2 shown here shows an increase. This raises the important question – is Site 1092 recording global atmospheric pCO2, or a more localised signal? A more local signal showing C storage in the high latitudes of the south Atlantic over the MMCT would also be very interesting of course. The change in the frontal positions is acknowledged in the text (Kuhnert et al 2009) but the change in stratification at this time is not mentioned. Paulsen (2005) (Bremen thesis) shows an increase in stratification at Site 1092 starting at 13.85 Ma (using the divergence of surface:deep planktonic foraminiferal d18O). Could this stratification have been associated with an increase in surface water [CO2]? This possibility should be addressed in the manuscript.

AC: We have added an entire new section ("4.1 Origin of Site 1092 *p*CO_2 signal") to tackle this question, discussing whether Site 1092 could have recorded a regional signal or not, with the conclusion that we cannot completely rule out local processes (l. 255-282).

4. Is it appropriate to compare the records across CM6 with those of a model that does not consider the impacts of the MMCT itself?
CM6 immediately follows the ice growth of the MMCT and is the largest by far of the CM events, making its interpretation more complex than the other CM events. It follows a major sea level fall, which would have affected the shelf:basin burial of carbonate, d13C and CO2 (e.g., Mckay et al 2016, Ma et al 2018). Further, the ice advance was likely associated with changes in ocean circulation and ventilation of bottom water masses (with implications for both d13C and CO2). Is it therefore

appropriate to make a straightforward comparison of the d13C and CO2 records to the Ma et al 2011 model without any consideration of these other processes? The Ma et al 2011 model was constructed to explain the long eccentricity signal in the d13C record throughout the Miocene Climatic Optimum, rather than examine CM6 specifically. It is solely forced by ETP, and does not include processes triggered by cryospheric thresholds in the climate system and resulting impacts. On the other hand, the Ma et al (2018) model suggests that a significant cause of the d13C increase at CM6 was the increased weathering resulting from the sea level fall. By saying the results here support the Ma et al 2011 model for CM6, where does that leave our understanding of the importance of shelf-basin carbonate deposition on global d13C signals?

AC: The reviewer correctly criticized that we refer to the model of Ma et al. (2011), which was designed for an ice-free world. Hence, we discuss the potential effect of Antarctic ice-sheet expansion superimposed on weathering on the carbon cycle in more detail (l. 358-372).

5. How does the pCO2 record compare with the high-resolution B/Ca record of Sosdian et al 2020?

It is odd that the manuscript does not compare the d11B-CO2 record with the high resolution B/Ca records published by Sosdian et al 2020. Those authors suggest that the increasing d13C of the CM events is associated with decreasing surface water DIC at Site 761. This is consistent with an increase in Corgburial:Carbonate burial which would predict a negative relationship between d13C and global CO2 across CM events, rather than the positive relationship observed across CM6 at Site 1092 here. Although, before a direct comparison can be made the origin of the d11B signal at Site 1092 needs to be thoroughly addressed.

AC: We have added a discussion of the new study by Sosdian et al. (2020) (l. 43-48; 397-313).

6. What is the significance of the change in the Malta age model for the GSSP?

The GSSP for the Langhian-Serravallian boundary is placed in the Malta section. There should be some comment about this in the age model section as the authors have changed the Malta age model.

AC: We have added a statement to this section, stating that that the revised age model is valid, although we did not explicitly place the boundary in there (l. 100-103).

More minor comments:

Line 45: "*However, most proxy records for the history of pCO2 across the MMCT are incomplete or at low resolution, thus prohibiting resolution of the CM events (Pagani et al., 1999; Kürschner et al., 2008; Foster et al., 2012; Ji et al., 2018; Sosdian et al., 2018; Super et al., 2018) and making it difficult to identify the mechanism responsible for the major step into the "icehouse" world."*
This statement is a bit disingenuous, these published records clearly show a significant pCO2 decrease just prior to the MMCT (see also Sosdian et al 2020). What is interesting of course is the higher resolution CO2 changes, and questions such as: what caused the CM events? how do the CM events depend on background climate state? Why is CM6 (immediately following the major ice growth) the largest of the CM events?

AC: We do not contradict existing studies showing a pCO2 that is lower after EAIS expansion than before. Actually we see something similar, but our record shows more details across CM events 5b and 6 (Fig. 6). We thoroughly discuss what could have caused the observed pCO2 and d13C evolution (l. 284-398) and summarize our conclusions in section 5 (l. 400-416).

Line 35: *"In a recent study, it was proposed that a more sluggish meridional Pacific Ocean overturning circulation, due to reduced deep-water formation in the Southern Ocean, enhanced the weathering of 13C-enriched shelf carbonates".*
This is incorrect, the enhanced weathering in this model was ascribed to the sea level fall. This model (Ma et al 2018) is very different to the Ma et al 2011 model. The Ma et al 2018 model treats the MMCT as an "event" whereas the Ma et al 2011 uses continually evolving orbital parameters as forcing. This important point is not explained in the current manuscript.

AC: The summary of the Ma et al. (2018) study has been corrected (l. 60-64, 361-363).

I find it odd that pH and pCO2 are plotted as the same curve with the same uncertainty envelope in Figure 4. pCO2 clearly has additional uncertainties (e.g. estimate of TA, d11Bsw), and I think these should be better represented in Figures 4 and 5.

AC: This was also an issue raised by reviewer 2. We have re-calculated all pCO2 values, temperatures and associated uncertainties using a Monte Carlo approach, which propagates all specific uncertainties in B isotope and Mg/Ca measurements, calibration regressions, TA, Sal and d11B of seawater. Further, we account for the effect of Miocene [Mg,Ca] of seawater on the equilibrium constants and on Mg/Ca temperatures, as well as the effect of Miocene d11Bsw on the intercept of the applied B isotope calibration. The method part on this all has been edited, extended and restructured (l. 123-207). In addition, we created a new figure (S1) to show the effect of d11Bsw extremes on calculated pCO2.

Line 31: Need to also cite Foster et al 2012 as demonstrating climate-CO2 relationship across MMCT.

AC: Added (l. 33).

Careful with all references – e.g. Fig 3 caption "Badger et al 2015" should be "Badger et al 2013".

AC: Corrected (l. 716).

Line 168 – Fig 5 should be Fig 4.

AC: Corrected.

Review No 2 (anonymous)

GENERAL COMMENTS Obtaining more B isotope-pH and CO2 estimates for the middle Miocene climate transition is a long overdue goal, making this study very timely and of great importance. The authors provide a comprehensive view of CO2 evolution for this period and potential mechanisms overlying potential eccentricity driven variations. The paper would benefit from some re-organization and focus on clarity, incorporation of recent studies in the discussion and the data, and a more comprehensive propagation of uncertainties, beyond the sensitivity analyses performed for alkalinity and salinity.

AC: We thank the referee for his constructive and thorough review of our manuscript, particularly concerning the analytical and data handling part. We have addressed all questions and comments, which are listed in more detail below.

SPECIFIC COMMENTS Uncertainties: a) The analytical uncertainty reported is quite small compared to the uncertainty of replicate analyses. There may be differences between the use of IC vs. Faraday detectors in different studies and here. Nevertheless, the authors should provide more details here as well on how they calculate analytical uncertainty. For example, the consistency standard used to calculate long term precision should have been run at similar concentrations to those of samples, and the uncertainty of this should be larger at low B-levels. Additionally, the authors should provide more details on the B blank contribution (if any).

AC: Only looking at the control standards measured at 2-3 ppb such as in this study, the long-term reproducibility is not different from the of 0.3 ‰ (2*SD) we have reported in the previous version. What was not clear from the previous manuscript is that the control standard is normally run ~6 times or more within a session, also in order to condition and stabilize the system before measuring the samples. At the end of a session, the average of the measured control standards is calculated. The 2*SD of the averages between all sessions is ~0.3 ‰ (which we take as the minimum uncertainty), while it is ~0.6 ‰ when all single analyses are considered. We therefore added the information that the long-term reproducibility is calculated from “per-session” averages of a control standard (l. 120). We also added some brief information on the procedural blanks (l. 110-111).

b) Why is d11Bsw error systematic? If weathering is extremely pronounced, couldn’t this cause variations in d11Bsw across this time window, even if the average residence time for B may be longer? Even if so, because of the non-linearity of the d11B-pH proxy, at different d11Bsw the dpH and thus dCO2 could differ. Some could, thus, argue the uncertainty in d11Bsw encompasses both uncertainty in absolute value across the MMCT but also potential variations across the window. The authors should provide at minimum two scenarios based on minimum and maximum d11Bsw estimates for this period.

AC: The uncertainty in d11Bsw is supposed to be systematic, even if weathering varied a lot, due to the large reservoir of the ocean vs rivers, and the B concentrations of those (ca. 4500 vs 16 ppb, respectively). However, as we re-calculated all pH, pCO2 and temperature values plus their associated uncertainties, which are propagated from specific uncertainties in input parameters, we also applied an uncertainty of 0.2‰ for d11B of seawater to account for potential variations in weathering. In addition, as the reviewer suggested, we added a new figure (S1) to demonstrate the effect of extreme d11Bsw values on calculated pCO2 values.

c) The level of details in Fig. 7 with all sensitivity analyses is very much appreciated. However, could the authors provide more explanation on how they estimate the Alk and Temp uncertainties? If they compare to literature or proxy estimates, shouldn't they use the maximum uncertainty reported (i.e. ± 2C, and ± 130 umol/kg Alk)? Comparing to other studies: The authors should discuss their results in light of two recent publications for the middle Miocene, Leutert et al. 2020 (Nat. Geo) for both their SST and dpH estimates, and Sosdian et al. 2020 (Nat. Comm.) for C cycle in relationship to climate.

AC: We provide more details on estimates of uncertainties and error propagation in the completely restructured method section 1.4 (l. 123-207). We also included discussions of the new studies by Sosdian et al. (2020) and Leutert et al. (2020) (l. 43-48; 58-60; 276-279; 289-292; 297-307).

Comparing CO2 records: e.g. Fig. 5: what drives the differences between different d11B records for the target age-window? Section 4.1 needs some more discussion, with focus on how this new record could differ from previous d11B records. Could there be an upwelling signal at the study site driving those high CO2 estimates when they deviate from the other records? It could also be differences in the calibration used for d11B, or the assumptions for calculating carbonate system parameters. It may be wise to process the d11B records in the same manner, exclude the potential of any regional and variable CO2 disequilibrium, and then merge reliable d11B records into a single record with full propagation of uncertainties. If uncertainties are not propagated, and instead sensitivity runs are provided (i.e. d11Bsw), then better to display relative changes in pH/CO2 instead of absolute values. (Here it may be wise to remove the alkenone CO2 as they are not discussed enough beyond what is already available in the literature and thus do not contribute to the story).

AC: We have added an entire new section ("4.1 Origin of Site 1092 pCO_2 signal") to tackle this question, discussing whether Site 1092 could have recorded a regional signal or not, with the conclusion that we cannot completely rule out local processes (l. 255-282). Moreover, we have improved our approach to better compare or record with other boron-based pCO2 reconstructions by applying exactly the same calculation procedure and using the same boundary conditions. However, we did not merge the records into a single one because of the few published data in rather low resolution. Further, there seem to be in part differences between the records, which would have skewed the fitted running average. We still show the alkenone records for comparison, but added a brief discussion why those are possibly not reliable at low pCO2 levels (l. 292-296).

Focusing on the d11B records available and this new one, it would be also beneficial to display not only CO2 but also pH evolution across the MMCT, and how different records compare.

AC: As we re-calculated the boron-based pCO2 records, using identical boundary conditions (TA and d11BSw), we do not show the pH curves in extra plots as they just mimic the pCO2 curves.

Site setting: It is argued that the site is not affected by upwelling being north of the frontal systems in the Southern Ocean. However, can this be said with certainty for the middle Miocene? Is there any evidence for that?

AC: As mentioned earlier, we have added an entire new section ("4.1 Origin of Site 1092 pCO_2 signal") to tackle this question, discussing whether Site 1092 could have recorded a regional signal or not, with the conclusion that we cannot completely rule out local processes (l. 255-282).

Carbonate system calculations: The authors should consider the effect of Mg and Ca concentrations in seawater on carbonate system calculations (i.e. K1, K2, Ksp), such as in Hain et al. 2015; 2018 or Zeebe and Tyrrell 2019.

AC: Thanks for this important hint. We have performed all calculations using the effect of [Mg,Ca] on the equilibrium constants. The procedure for this is explained in the new subsection "1.4.2 Equilibrium constants" (l. 142-147).

Benthic-planktic pH records: Although the uncertainties are very large to make discernible conclusions about pH gradient values during the middle Miocene, it is interesting to further explore the dpH evolution and the surface-to-deep gradient evolution during the Miocene, and what drives this. If the benthic foraminiferal pH record is included, it should be discussed further.

AC: The resolution of the planktic and benthic pH records in Figure S4 is too low and uncertainties too large to draw substantiated conclusions concerning the temporal evolution of the surface-to-deep gradient. This is now better explained (l. 216-220), but we could use the data to possibly exclude decreased vertical mixing at this site after glaciation (l. 274-276).

Discussion on role of eccentricity and deep water ventilation: Here the section leaves us wanting more! It could benefit from some reorganization for clarity and flow, including recent studies such as those mentioned above.

AC: The discussion has been much extended and restructured to enhance clarity and flow (l. 284-398).