Predicting trend in atmospheric carbon dioxide using existing climate archives.
POST REVISION REVIEW RESPONSES
JRW. Martin, J. Pedro, T. Vance

We thank the reviewers for their valuable suggestions on our paper. In response, we have significantly revised this manuscript. As requested, we have greatly increased the number of external data sources used to compare to our simple GLS-based carbon dioxide reconstruction over the mid Pleistocene transition and addressed all other reviewer suggestions. Following these revisions, we find no clear evidence to reject our predicted carbon dioxide record and inherent null hypothesis. This combined review response should be read together with our earlier responses to the individual reviews, as requested by the Copernicus editorial team. The revised manuscript was prepared after completing the earlier individual review responses. This combined review response contains the specific changes to the manuscript in response to the reviews, whilst the earlier review responses were more general.

RC1:

Comment: *"The analysis appears to be performed well, and I have only a few concerns about the interpretation of the results. However, my main concern is that the work is too simple. I would encourage the authors to add more intellectual substance to the paper by exploring perhaps nonlinear regression between benthic $\delta^{18}O$ and CO_2".*

Our objective with this manuscript was to generate the simplest reasonable model to predict CO_2 from the δ^{18}O benthic stack and to test the predictions against available observations. In response to this review and comments from other reviews we have substantially expanded the work in three main areas:

1. We bring in additional data to evaluate our predicted record CO_2, specifically: Allan Hills Blue Ice CO_2 data from Yan et al., (2019) at 1.5 Mya; CO_2 proxy reconstructions from leaf wax from Yamamoto et al., (2022); δ^{11}B reconstructions by Dyez et al., (2018), and Guillermic et al., (2022) and a high-resolution CO_2 reconstruction by van de Wal et al. (2011).

   See revised fig. 03 (Line 231), and revised fig. 04 (Line 337), which show these records. Further, additional text has been added:

   - The Results section (line 193-262) now primarily focuses on the two blue ice data sets; specifically, their comparison to each other, the current continuous CO_2 records, and to our modelled predictions.
   - Additional text on other proxy-based, and model-based estimates of CO_2 across the MPT is now provided in the Discussion (see manuscript lines 299-335).

2. In light of the additional comparisons with blue and proxy CO_2 records from (1.) we find that our regression-based prediction of pre-MPT CO_2 is not rejected by the existing data and we discuss the implications (see Discussion lines 322 to 389). Further, see line 266-285 for additional justification and discussion of our use of the GLS regression model over more complex, non-linear alternatives:

> "It is possible that the fit between observed and our predicted CO_2 data could be further improved using a non-linear approach. However, we refrain from a non-linear approach for several key reasons. First, a scatter plot of the LR04 $\delta^{18}O$ benthic stack versus observed ice core CO_2 over the past 800 kyr yields a Pearson's correlation coefficient (R) of -0.82 (Fig. 2), indicating that ~68% of the variance in observed CO_2 is shared with the benthic stack. Importantly, there is no evidence in this scatter plot for departure from the linear relationship at high or low CO_2 or benthic $\delta^{18}O$ levels. Second, following the approach of Chalk *et al.*, 2017 and interpreting the upper 25th percentile of CO_2 data as representing mean interglacial stage CO_2 and the lower 25th percentile of CO_2 data as representing mean glacial stages CO_2 levels, we see that our predicted interglacial mean value for the past 800 kyr (253.1 ± 2.3 ppm) closely overlaps with the observed interglacial mean value (253.9 ± 4.1 ppm) and similarly, the predicted glacial stage mean (199.7 ± 1.7 ppm) closely overlaps with the observed glacial stage mean (202.0 ± 3.2 ppm). Third, the predictions are remarkably insensitive to bootstrap analysis in which 50 % of that data are omitted with each iteration of the GLS model (Fig 1). Such insensitivity to the bootstrap analysis and accurate prediction of glacial *and* interglacial state CO_2 values would be unlikely in the case of major non-linear dependencies between the LR04 predictor and CO_2 response variables. Fourth, non-linear approaches would risk generating an improved fit due to statistical artefacts that do not meaningfully relate to any dependence between benthic $\delta^{18}O$ and CO_2. Finally, the specific causes and sources and sinks involved in glacial to interglacial and millennial-scale CO_2 variations still remain poorly constrained (e.g. Archer *et al.*, 2000; Sigman *et al.*, 2010; Gottschalk *et al.*, 2019). Given this process-uncertainty, the GLS model fits our criteria of the simplest reasonable model. Further, the use of benthic $\delta^{18}O$ to predict atmospheric CO_2 has precedence; in response to the EPICA challenge (Wolff et al., 2004), N. Shackleton used this method to predict atmospheric CO_2 out to 800 kyr (Wolff, 2005). Furthermore, inverse modelling of CO_2 forced by the LR04 benthic stack has been undertaken by Berends et al. (2021a) and van de Wal et al. (2011)"

Comment: *"…discussing in more depth the underlying mechanisms relating benthic $\delta^{18}O$ and CO_2 to say more about the implications of potential misfit between CO_2 and the regression-based estimate."*

See lines 121-132 for discussion of mechanisms relating benthic $\delta^{18}O$ and CO_2:

> "… Mechanistically, multiple processes are expected to contribute to the shared variance. A first order factor is the dependency of CO_2 solubility on ocean temperature (e.g. Millero, 1995). From the simple solubility perspective, colder climate states with increased ice volume and colder ocean temperatures will drive increased ocean uptake of CO_2 (Berends *et al.,* 2021). However, the solubility effect only accounts for a portion of observed glacial CO_2 drawdown (Archer *et al.*, 2000). Multiple additional contributors to the shared variance are proposed in the literature. These include (not exhaustively), direct radiative forcing of ice volume changes by CO_2 (e.g. Shackleton *et al.*, 1985); the impact of ice volume/sea level changes on

> atmospheric CO_2 via ocean productivity and carbonate chemistry changes (e.g. Broecker, 1982; Archer *et al*., 2000; Ushie and Matsumoto, 2012); CO_2 drawdown during periods of high ice volume by increased iron fertilization (e.g. Röthlisberger *et al*., 2004; Martinez-Garcia *et al*., 2014) and enhanced sea ice extent during periods of high ice volume capping the ventilation of CO_2 from the ocean interior at high latitudes (Stephens and Keeling, 2000)."

Comment: *"Abstract, line 18: I think the authors meant benthic foraminiferal stable isotope (δ18O). The $\delta^{18}O$ data used is from foraminiferal calcite, not "water.""*

Revised to "LR04 benthic $\delta^{18}O_{calcite}$ stack (Lisiecki & Raymo, 2005) from marine sediment cores" at line 22 (abstract)

Comment: *"Line 118: It is not clear what the authors mean by "This trend is seen in our predicted record, and in the filtered BI-CO_2 and BOR-CO_2 data (Fig. 1C)." The previous sentence describes glacial stage CO_2 draw-down and the absence of an interglacial draw-down. In Fig. 1C, it appears that this description holds for the predicted CO_2 record (i.e., glacial draw-down but steady interglacial values). However, the BI-CO_2 and BOR-CO_2 data show a change in BOTH glacial and interglacial CO_2 compared to the post-MPT average. The text should be revised to make clear which trends are similar between the predictions and observations and which are different."*

In our revised manuscript we consider blue ice as the most reliable (currently available) measure of CO_2 concentration across the MPT for reasons noted in line 299-320 (see below). We quantitatively compare glacial and interglacial thresholds for blue ice data with our predictions and we no longer use BOR-CO_2 for these comparisons (see figure 3 – line 231). Our results section (lines 192-262) discusses glacial and interglacial trends exhibited by the two blue ice data sets (Yan *et al*, Higgins *et al*) through comparison to the current continuous records and we have worked to clarify the text about which trends are similar and which different between trends and observations:

> "We now consider long-term trends in interglacial and (separately) glacial CO_2 levels across the past 1.5 Myr in PRED-CO_2 and in the existing ice core CO_2 data. For PRED-CO_2 there is no significant difference between CO_2 concentrations in the interglacial stages of the 1.5 Mya ± 213 kya, 1000 ± 89 kya and 0–800 kya windows (Fig 4 D, blue bars). In the ice core observations, interglacial levels at 1.5 Mya in BI-CO_2 are also within the uncertainties of those in the 0–800 kya interval. Notably, the BI-CO_2 concentrations in the 1000 ± 89 kya interval appear elevated with respect to the 0–800 kyr and 1.5 Mya ± 213 kya intervals, however this elevated (ca. 271 ppm) level is consistent with the observed interglacial CO_2 concentration during interglacials 5, 9 and 11 (Fig 3B). Overall, there is no indication in the observed ice core CO_2 data or in PRED-CO_2 for a long-term trend in *interglacial* CO_2 levels across the past 1.5 Myr.
>
> In comparison, there are significant declines in glacial CO_2 levels across the MPT in PRED-CO_2 and the observed ice core data. For PRED-CO_2, glacial CO_2

> concentrations are not significantly different during the 1.5 Mya ± 213 kya and 1000 ± 89 kya windows. However, across the MPT, PRED-CO_2 glacial concentrations drop by ~18 ppm. This pattern is consistent with the observed data, where glacial CO_2 levels are also not significantly different between the 1.5 Mya ± 213 kya and 1000 ± 89 kya windows (217.6 ± 2.3 and 226.2 ± 4.0 ppm, respectively) and then fall by 24 ppm to the 0–800 kyr observed glacial mean of 202.0 ± 3.2 ppm. Glacial-stage draw-down of CO_2 across the MPT in the absence of interglacial draw-down is consistent with previous observations based on the boron-isotope-based CO_2 reconstructions (e.g., Chalk *et al*., 2017; Hönisch *et al*., 2009 and see Discussion). In the following section we also compare PRED-CO_2 data to boron-isotope-based and other CO_2 proxy records covering the 0 to 1.5 Myr interval."

Lines 299-320 (Discussion) and Fig. 04 (line 337) quantitatively discuss boron-based CO_2 estimates in comparison to our predicted CO_2 record and their limitations:

> "We consider the BI-CO_2 date to provide the most reliable measurements of CO_2 concentration, in the absence of a continuous ice core record across the MPT. However, further comparison of our CO_2 predictions can also be made against CO_2 proxy data from non-ice core archives (Fig 4A). We consider here δ^{11}B-based atmospheric CO_2 reconstructions (Chalk *et al.,* 2017, Dyez *et al.* 2018 and Guillermic *et al.* 2022) and a recent atmospheric CO_2 reconstruction from δ^{13}C of leaf wax (Yamamoto *et al.,* 2022). The continuous δ^{11}B-based reconstructions of Dyez *et al.*, (2018) overlap PRED-CO_2 from ~1.38 – 1.5 Mya while the Chalk *et al*., (2017) reconstruction overlaps PRED-CO_2 from 1.09 – 1.43 Mya. Discrete reconstructions from Guillermic *et al.* (2022) are distributed non-uniformly across the 800 to 1.5 Mya interval. For the two continuous δ^{11}B-based reconstructions (Chalk *et al.,* (2017) and Dyez *et al.,* (2018)) the glacial CO_2 levels appear consistent with the PRED-CO_2 record, within their reported 30 – 60 ppm uncertainties. However, δ^{11}B-based interglacial stages in these reconstructions exceed those of the PRED-CO_2 record (Fig. 4A). The Guillermic *et al.* (2022) reconstructions suggest a larger range of CO_2 concentrations than the overlapping intervals of PRED-CO_2 and of the two continuous δ^{11}B-based reconstructions (Fig. 4A). The large range of the Guillermic *et al.* (2022) data and the high interglacial maxima in the Chalk *et al* (2017) and Dyez *et al*., (2018) data, all significantly exceed the range and interglacial maxima from the BI-CO_2 estimates. These discrepancies internally between different δ^{11}B-based CO_2 reconstructions and between the δ^{11}B-based reconstructions and the BI-CO_2 data, may be due to uncertainties associated with the δ^{11}B proxy transfer function. The δ^{11}B-based CO_2 reconstructions are dependent on assumptions about multiple components of the carbonate system, including local marine carbon chemistry and the CO_2 saturation state in the past and (Hönisch *et al*., 2009). Evidence that δ^{11}B-based reconstructions may overestimate interglacial stage CO_2 is also seen in data from Chalk *et al*., (2017) spanning ca. 0–250 kya, where the δ^{11}B-based interglacial CO_2 levels exceed the continuous ice core CO_2 record by ca. 30 ppm (not shown)."

Comment: *"Line 185-186: The authors need to explain why out-of-phase responses in northern and southern ice before the MPT (as proposed by Raymo et al., 2006) would lead them to expect "large discrepancies" between their regression-based CO_2 prediction and the* *realized data. This inference seems to rely on the assumption of a certain relationship between CO_2 and northern or southern ice sheets, but I'm not sure what relationship the authors are assuming. Section 4.2 overall is quite short and would benefit from a more in-depth, process-based discussion of implications of the anti-phased hemisphere hypothesis for pre-MPT CO_2 variability"*

Lines 364-375 now offer a brief, but more in depth summary of the phase locking hypotheses:

> "The phase locking hypothesis is proposed to explain the absence of precession-related (23 kyr) periods in the LR04 benthic stack prior to the MPT (Fig 1), despite the strong precession cycle in insolation (Raymo *et al*., 2006, Morée *et al*., 2021). The key concept is that prior to the MPT the Northern Hemisphere and Antarctic ice sheets were responsive (in ice volume) to insolation changes in the precession band, but because precession forcing is out of phase between the hemispheres, the ice volume changes were opposing between the hemispheres and therefore cancelled in the benthic stack. This cancellation of the precession signal left insolation forcing in the 41 kyr obliquity band to dominate globally integrated ice volume changes expressed in the benthic stack. A transition from a smaller and more dynamic terrestrial-terminating Antarctic ice sheet to a larger and more stable marine-terminating ice sheet with cooling climate across the MPT (e.g. Elderfield *et al*., 2012) is then proposed to remove sensitivity of Antarctic ice volume to precession forcing and to suppress ice sheet sensitivity to the obliquity band in favour of quasi-100kyr ice volume changes that are in phase between the hemispheres (Raymo *et al*., 2006)."

We address/expand why out of phase responses in northern and southern hemisphere ice sheets would lead to large discrepancies between our model and a future realised CO_2 records from line 377-389:

> "Recently presented data from Yan *et al*. (2022), lend some support to the phase locking hypothesis, specifically with evidence that pre-MPT Antarctic temperature (and by extension ice volume) is positively correlated with a local precession-band insolation proxy based on the oxygen to nitrogen ratio of trapped air (Yan *et al*., 2022). Whereas the correlation becomes negative in the blue ice and continuous ice core data in the post-MPT record. If Yan *et al*., (2022) is correct and the phase locking hypothesis holds, then an implication is that prior to the MPT, Antarctic climate, Antarctic ice volume and by extension Southern Ocean climate conditions, would fall out of phase with the LR04 benthic stack. To now extend the argument to potential impacts on CO_2 exchange, if the phase locking hypothesis holds, then prior to the MPT the Antarctic and Southern Ocean climate conditions and by extension the Southern Ocean mechanisms of CO_2 exchange described earlier, would also be expected to fall out of phase with the benthic stack. Since our regression model assumes continuation of the in-phase relationship between the benthic stack and Antarctic and Southern Ocean climate conditions (as inherited from the post-MPT training data) we would expect to see major disagreement between our pre-MPT CO_2 predictions and a realised oldest ice continuous ice core CO_2 record."

RC2:

Comment: *"line 16: "is to make""*

Accepted and revised

Comment: *"line 17 and throughout the manuscript: Myr instead of myr"*

Accepted and revised

Comment: *"line 25: the authors state that the null hypothesis should be rejected, however, without laying out the causal relationship between the regression parameters and potential reasons why the regression may not hold back in time, this statement is not entirely satisfying."*

We have undertaken further research and work for this significantly revised manuscript that has resulted in changes to our conclusions:

> 1: We have examined and include the more recently published blue ice core data at 1.5 Mya (Yan *et al*. (2022)) along with the blue ice data from 1 Mya (Higgins *et al.,* 2015).
>
> 2: To infer long term glacial and interglacial trends in observations compared to our predictions, we primarily focus on these two blue ice data sets.
>
> 3: We now suggest there is no clear evidence to reject our null-hypothesis.(line 399-406:
>
>> "We made initial tests of the null hypothesis by comparing our predicted CO_2 record to existing discrete blue ice CO_2 records and other non-ice-core proxy-CO_2 records from the 800–1500 kyr interval. Our predicted CO_2 concentrations do not show any systematic departure from observed blue ice CO_2 concentrations. The predictions are marginally lower (during glacial *and* interglacial stages) than those observed in blue ice from 1000 ± 89 kya and marginally higher than observed in blue ice data from 1.5 Mya ± 213 kyr. Our predictions were generally lower than interglacial δ^{11}B-based-CO_2 reconstructions, but higher than recent δ^{13}C of leaf-wax based CO_2 reconstructions. Overall, we do not find clear evidence from the existing blue ice or proxy CO_2 data to reject our predictions nor our associated null-hypothesis."
>
> 4: We have included more information on the relationship between the model parameters (See lines 117-132):
>
>> "Fig. 2 shows a scatter-plot of the LR04 δ^{18}O benthic stack versus observed ice core CO_2 over the past 800 kyr. Both data sets are binned to equivalent 3-

> kyr time steps (Methods). The Pearson's correlation coefficient (r) between the data sets is -0.82 ($p < 0.05$) indicating that ~68% of the variance in observed CO_2 is shared with the LR04 $\delta^{18}O$ benthic stack. This strong relationship provides an initial rationale for using the LR04 $\delta^{18}O$ benthic stack as an input parameter to predict CO_2 beyond 800 kyr. Mechanistically, multiple processes are expected to contribute to the shared variance. A first order factor is the dependency of CO_2 solubility on ocean temperature (e.g. Millero, 1995). From the simple solubility perspective, colder climate states with increased ice volume and colder ocean temperatures will drive increased ocean uptake of CO_2 (Berends *et al.,* 2021). However, the solubility effect only accounts for a portion of observed glacial CO_2 drawdown (Archer *et al*., 2000). Multiple additional contributors to the shared variance are proposed in the literature. These include (not exhaustively), direct radiative forcing of ice volume changes by CO_2 (e.g. Shackleton *et al*., 1985); the impact of ice volume/sea level changes on atmospheric CO_2 via ocean productivity and carbonate chemistry changes (e.g. Broecker, 1982; Archer *et al*., 2000; Ushie and Matsumoto, 2012); CO_2 drawdown during periods of high ice volume by increased iron fertilization (e.g. Röthlisberger *et al*., 2004; Martinez-Garcia *et al*., 2014) and enhanced sea ice extent during periods of high ice volume capping the ventilation of CO_2 from the ocean interior at high latitudes (Stephens and Keeling, 2000)."

Comment: *"line 58-59: $\delta^{18}O$ is not just a sea level proxy but also influenced by deep ocean temperature. A process-based discussion of why LR04 is a viable input parameter to predict CO_2 is required."*

See lines 117-132 of the revised text (Introduction/above) for potential mechanistic relationships between CO_2 and $\delta^{18}O$, and lines 264-288 for a process based discussion of why LR04 is a viable input parameter to predict CO_2:

> "Our objective with this manuscript was to generate the simplest reasonable model to predict CO_2 from the LR04 $\delta^{18}O$ benthic stack and to test the predictions against available observations. It is possible that the fit between observed and our predicted CO_2 data could be further improved using a non-linear approach. However, we refrain from a non-linear approach for several key reasons. First, a scatter plot of the LR04 $\delta^{18}O$ benthic stack versus observed ice core CO_2 over the past 800 kyr yields a Pearson's correlation coefficient (R) of -0.82 (Fig. 2), indicating that ~68% of the variance in observed CO_2 is shared with the benthic stack. Importantly, there is no evidence in this scatter plot for departure from the linear relationship at high or low CO_2 or benthic $\delta^{18}O$ levels. Second, following the approach of Chalk *et al*., 2017 and interpreting the upper 25th percentile of CO_2 data as representing mean interglacial stage CO_2 and the lower 25th percentile of CO_2 data as representing mean glacial stages CO_2 levels, we see that our predicted interglacial mean value for the past 800 kyr (253.1 ± 2.3 ppm) closely overlaps with the observed interglacial mean value (253.9 ± 4.1 ppm) and similarly, the predicted glacial stage mean (199.7 ± 1.7 ppm) closely overlaps with the observed glacial stage mean (202.0 ± 3.2 ppm). Third, the predictions are remarkably insensitive to bootstrap analysis in which 50 % of that data are omitted with each iteration of the GLS model (Fig 1). Such insensitivity to the bootstrap analysis and accurate prediction of glacial *and* interglacial state CO_2 values

> would be unlikely in the case of major non-linear dependencies between the LR04 predictor and CO_2 response variables. Fourth, non-linear approaches would risk generating an improved fit due to statistical artefacts that do not meaningfully relate to any dependence between benthic $\delta^{18}O$ and CO_2. Finally, the specific causes and sources and sinks involved in glacial to interglacial and millennial-scale CO_2 variations still remain poorly constrained (e.g. Archer *et al*., 2000; Sigman *et al*., 2010; Gottschalk *et al*., 2019). Given this process-uncertainty, the GLS model fits our criteria of the simplest reasonable model. Further, the use of benthic $\delta^{18}O$ to predict atmospheric CO_2 has precedence; in response to the EPICA challenge (Wolff et al., 2004), N. Shackleton used this method to predict atmospheric CO_2 out to 800 kyr (Wolff, 2005). Furthermore, inverse modelling of CO_2 forced by the LR04 benthic stack has been undertaken by Berends et al. (2021a) and van de Wal et al. (2011)."

Comment: *"Line 66: please include also the record by Dyez et al., Paleoceanography 2018"*

We have included the record presented by Dyez et al. in Fig. 04 (line 337) as a further boron-based CO_2 estimates. We discuss it in relation to our prediction about line 301-311:

> "We consider here $\delta^{11}B$-based atmospheric CO_2 reconstructions (Chalk *et al.,* 2017, Dyez *et al.* 2018 and Guillermic *et al.* 2022) and a recent atmospheric CO_2 reconstruction from $\delta^{13}C$ of leaf wax (Yamamoto *et al.,* 2022). The continuous $\delta^{11}B$-based reconstructions of Dyez *et al*., (2018) overlap PRED-CO_2 from ~1.38 – 1.5 Mya while the Chalk *et al*., (2017) reconstruction overlaps PRED-CO_2 from 1.09 – 1.43 Mya. Discrete reconstructions from Guillermic *et al.* (2022) are distributed non-uniformly across the 800 to 1.5 Mya interval. For the two continuous $\delta^{11}B$-based reconstructions (Chalk *et al.,* (2017) and Dyez *et al.,* (2018)) the glacial CO_2 levels appear consistent with the PRED-CO_2 record, within their reported 30 – 60 ppm uncertainties. However, $\delta^{11}B$-based interglacial stages in these reconstructions exceed those of the PRED-CO_2 record (Fig. 4A). The Guillermic *et al.* (2022) reconstructions suggest a larger range of CO_2 concentrations than the overlapping intervals of PRED-CO_2 and of the two continuous $\delta^{11}B$-based reconstructions (Fig. 4A)."

and in relation to blue ice about line 311:

> "The large range of the Guillermic *et al.* (2022) data and the high interglacial maxima in the Chalk *et al* (2017) and Dyez *et al*., (2018) data, all significantly exceed the range and interglacial maxima from the BI-CO_2 estimates"

Comment: *"line 68: The very old ice at Allan Hills is not really from the surface but from a shallow ice drilling of more than 100 m depth."*

In this revised manuscript we review our initial description of blue ice about line 150-153:

> "We use the term blue ice to describe deep, ancient glacial ice that has been brought nearer to the surface of an ice sheet by ice flow. Blue ice is sampled by cutting

> trenches or shallow drilling of up to several hundred meters (e.g. Higgins *et al.,* 2015)"

Comment: *"Methods: the uncertainty in the regression connected to the independent age scales should be discussed"*

Within our methods section we discuss the age uncertainties associated with the LR04 benthic stack and the composite CO_2 record from line 178-190:

> "Uncertainties in the independent age scales of both the LR04 stack and the compiled CO_2 record are inherited by our GLS model and its predictions. The LR04 stack includes 57 globally-distributed benthic $\delta^{18}O$ sediment core records. The age models for these cores are independently constructed from the average sedimentation rates of each core, assuming global sedimentation rates have remained relatively stable, and with tuning to a simple ice model based on 21 June insolation at 65°N (Lisiecki & Raymo, 2005). The authors estimate uncertainty of 6 kyr from 1.5 – 1.0 Mya and 4 kyr from 1 – 0 Mya (Lisiecki & Raymo, 2005). The observed CO_2 composite ice core record for the past 800 kya (Bereiter at al., 2015) uses six independent dating methods for various core locations both spatially across Antarctica, and stratigraphically for different sections of the same core. The age uncertainty in the gas timescale has a median over the 0 – 800 kya interval of 2 kyr, but individual uncertainties can reach up to 5 kyr (Veres *et al* 2013; Bazin *et al*., 2013). The relative age uncertainties between these input variables may diminish the regression or in some instances lead to spurious correlation. However, we expect any such effects are minor on the basis that our predictions show little sensitivity to the bootstrap analysis with 1000 iterations of re-computing the regression after removing 50% of data (see Fig. 3B, C and Discussion)."

Comment: *"line 85: not clear what r(226) means, please explain. Did you allow for lag correlation? (see also comment on age scales above)"*

R(226) represented the degrees of freedom of the test to determine the correlation coefficient. For simplicity we have removed it. Further, Our generalised least square model accounted for autocorrelation/lag between the predictor ($\delta^{18}O$) and CO_2 using an AR(1) correlation factor.

Comment: *"line 89: the limitations of blue ice CO_2 reconstructions and $\delta^{11}B$ reconstructions of CO_2 should be discussed as well."*

Caveats associated with blue ice CO_2 measurements is now discussed at line 153-155 (Introduction):

> "The vertical migration of blue ice is associated with high deformation making the ice samples stratigraphically complex and hard to date (Higgins *et al.,* 2015). As a result, blue ice records alone do not provide a continuous CO_2 record across the MPT."

and further at line 290-296 (Discussion):

> "There are several caveats with blue ice data that may affect its use to evaluate our GLS model predictions. The blue ice data may have been subject to diffusional smoothing of CO_2 (e.g. Yan *et al.,* 2019), which would act in the direction of elevating the (lower 25th percentile) assumed glacial concentrations above the glacial atmospheric values and reducing the (upper 25th percentile) assumed interglacial concentrations. There is also the potential for artificially elevated CO_2 concentrations in blue ice due in-situ respiration of CO_2 due to microbial activity in detrital matter. Respiration effects are screened for by measurements of $\delta^{13}C$ of CO_2, however it is difficult to demonstrate that all samples are unaffected (Yan *et al*., 2019)."

Line 313-320 (Discussion) also covers limitations of $\delta^{11}B$ CO_2 reconstructions:

> "These discrepancies internally between different $\delta^{11}B$-based CO_2 reconstructions and between the $\delta^{11}B$-based reconstructions and the BI-CO_2 data, may be due to uncertainties associated with the $\delta^{11}B$ proxy transfer function. The $\delta^{11}B$-based CO_2 reconstructions are dependent on assumptions about multiple components of the carbonate system, including local marine carbon chemistry and the CO_2 saturation state in the past and (Hönisch *et al*., 2009). Evidence that $\delta^{11}B$-based reconstructions may overestimate interglacial stage CO_2 is also seen in data from Chalk *et al*., (2017) spanning ca. 0–250 kya, where the $\delta^{11}B$-based interglacial CO_2 levels exceed the continuous ice core CO_2 record by ca. 30 ppm (not shown)."

Peter Kohler:

Comment: *"To be transparent in what has been done, the equation which calculates CO_2 out of the LR04 benthic $\delta^{18}O$ stack is missing. Plotting of the LR04 benthic $\delta^{18}O$, which is at the core of the approach is also missing."*

Line 171 now shows the equation used to calculate CO_2 from $\delta^{18}O$, and the LR04 benthic plot has been added to figure 3a (line 231).

Comment: *"Blue ice CO_2 data from Allan Hills have been extended in Yan et al (2019), now also containing snapshots of CO_2 at 1.5 and 2.0 Ma"*

Thank you. We have included the blue ice data at 1.5 Mya by Yan *et al.* (see figure 3 – line 231). In the manuscript lines 206-262 (Results) we compare both the Higgins *et al.,* and Yan *et al.,* blue ice data to each other, to the current 800 kya record and to our predictions: See manuscript lines 192 to 262:

> "We now consider long-term trends in interglacial and (separately) glacial CO_2 levels across the past 1.5 Myr in PRED-CO_2 and in the existing ice core CO_2 data. For PRED-CO_2 there is no significant difference between CO_2 concentrations in the interglacial stages of the 1.5 Mya ± 213 kya, 1000 ± 89 kya and 0–800 kya windows (Fig 4 D, blue bars). In the ice core observations, interglacial levels at 1.5 Mya in BI-CO_2 are also within the uncertainties of those in the 0–800 kya interval. Notably, the BI-CO_2 concentrations in the 1000 ± 89 kya interval appear elevated with respect to the 0–800 kyr and 1.5 Mya ± 213 kya intervals, however this elevated (ca. 271 ppm) level is consistent with the observed interglacial CO_2 concentration during interglacials 5, 9 and 11 (Fig 3B). Overall, there is no indication in the observed ice core CO_2 data or in PRED-CO_2 for a long-term trend in *interglacial* CO_2 levels across the past 1.5 Myr.
>
> In comparison, there are significant declines in glacial CO_2 levels across the MPT in PRED-CO_2 and the observed ice core data. For PRED-CO_2, glacial CO_2 concentrations are not significantly different during the 1.5 Mya ± 213 kya and 1000 ± 89 kya windows. However, across the MPT, PRED-CO_2 glacial concentrations drop by ~18 ppm. This pattern is consistent with the observed data, where glacial CO_2 levels are also not significantly different between the 1.5 Mya ± 213 kya and 1000 ± 89 kya windows (217.6 ± 2.3 and 226.2 ± 4.0 ppm, respectively) and then fall by 24 ppm to the 0–800 kyr observed glacial mean of 202.0 ± 3.2 ppm. Glacial-stage draw-down of CO_2 across the MPT in the absence of interglacial draw-down is consistent with previous observations based on the boron-isotope-based CO_2 reconstructions (e.g., Chalk *et al*., 2017; Hönisch *et al*., 2009 and see Discussion). In the following section we also compare PRED-CO_2 data to boron-isotope-based and other CO_2 proxy records covering the 0 to 1.5 Myr interval."

Comment: *"A recent paper by Yamamoto et al (2022) calculates CO_2 over the MPT from leaf wax $\delta^{13}C$ and finds that smaller glacial/interglacial amplitudes in CO_2 before the MPT are based on stable glacial CO_2, but smaller interglacial CO_2 before the MPT. This differs to the $\delta^{11}B$-based CO_2, and if I got it right might support the here defined Null Hypothesis, which then cannot easily be dismissed."*

Another great recommendation for data to examine. This has been included in our discussion sections in figure 4A (line 337). We discuss it specifically at line 322-325:

> "By comparison, the $\delta^{13}C$ of leaf wax data (Yamamoto *et al.,* 2022) has a similar glacial to interglacial range as PRED-CO_2, but a ca. 20ppm lower mean concentration than our predictions (Fig 4A). Hence, our PRED-CO_2 data fall lower than interglacial $\delta^{11}B$-based interglacial levels but are higher than the $\delta^{13}C$ of leaf-wax based estimate. Given the evidence that $\delta^{11}B$-based reconstructions are known to overestimate atmospheric CO_2 concentration in the continuous ice core record, we do not find cause from the existing CO_2 proxy data to reject our predictions nor our associated null-hypothesis."

Comment: *"New CO_2 data based on $\delta^{11}B$ from Pacific cores have recently been published (Guillermic et al., 2022). Ok, data coverage across the last 1.5Ma might be weak, but worth discussing it."*

Boron based CO_2 estimates by Guillermic *et al.,* have been added to figure 4A (line 337) and discussed with other boron based estimated at line 301-320:

> "We consider here δ^{11}B-based atmospheric CO_2 reconstructions (Chalk *et al.,* 2017, Dyez *et al.* 2018 and Guillermic *et al.* 2022) and a recent atmospheric CO_2 reconstruction from δ^{13}C of leaf wax (Yamamoto *et al.,* 2022). The continuous δ^{11}B-based reconstructions of Dyez *et al*., (2018) overlap PRED-CO_2 from ~1.38 – 1.5 Mya while the Chalk *et al*., (2017) reconstruction overlaps PRED-CO_2 from 1.09 – 1.43 Mya. Discrete reconstructions from Guillermic *et al.* (2022) are distributed non-uniformly across the 800 to 1.5 Mya interval. For the two continuous δ^{11}B-based reconstructions (Chalk *et al.,* (2017) and Dyez *et al.,* (2018)) the glacial CO_2 levels appear consistent with the PRED-CO_2 record, within their reported 30 – 60 ppm uncertainties. However, δ^{11}B-based interglacial stages in these reconstructions exceed those of the PRED-CO_2 record (Fig. 4A). The Guillermic *et al.* (2022) reconstructions suggest a larger range of CO_2 concentrations than the overlapping intervals of PRED-CO_2 and of the two continuous δ^{11}B-based reconstructions (Fig. 4A). The large range of the Guillermic *et al.* (2022) data and the high interglacial maxima in the Chalk *et al* (2017) and Dyez *et al*., (2018) data, all significantly exceed the range and interglacial maxima from the BI-CO_2 estimates. These discrepancies internally between different δ^{11}B-based CO_2 reconstructions and between the δ^{11}B-based reconstructions and the BI-CO_2 data, may be due to uncertainties associated with the δ^{11}B proxy transfer function. The δ^{11}B-based CO_2 reconstructions are dependent on assumptions about multiple components of the carbonate system, including local marine carbon chemistry and the CO_2 saturation state in the past and (Hönisch *et al*., 2009). Evidence that δ^{11}B-based reconstructions may overestimate interglacial stage CO_2 is also seen in data from Chalk *et al*., (2017) spanning ca. 0–250 kya, where the δ^{11}B-based interglacial CO_2 levels exceed the continuous ice core CO_2 record by ca. 30 ppm (not shown)."

Comment: *"CO_2 as function of benthic $\delta^{18}O$ has in an inverse modelling approach already been calculated by Stap et al (2016). This approach has been updated by Berends et al. (2021a). So comparison to their results might tell, how (if at all) this study shows something new."*

Our CO_2 prediction, like Berends et., 2021a, was trained on data from the recent 800 kyr and motivated by comparison to the upcoming oldest ice core records. Our simple model yields a high correlation to the observed 800 kyr Bereiter et al., CO_2 record ($r^2 = 0.68$) and our CO_2 predictions out to 1.5 Myr can not be confidently excluded by the available blue ice and CO_2 proxy reconstructions. From what we understand Berends et., 2021a does not make any evaluation or comparison to the discrete δ^{11}B and blue ice data over the MPT.

Comment: *"Maybe also discuss other approaches of CO_2 across the MPT, eg C cycle simulation results (apart from those in Willeit et al, 2020, which are cited) of Köhler & Bintanja (2006), or the compilation of at that time available CO_2 data and the calculation of a continous high-resolution CO_2 record in van de Wal et al. (2011), updated in Stap et al. (2018)."*

Figure 4B (line 337) now includes the modelled record by van de Wal (*et al*) and Willeit (*et al*) and these are discussed, relative to our predictions at line 329-335:

> "We also compare our predictions to existing more complex model simulations (Fig 4B.). First, against a transient simulation using an intermediate-complexity earth system model (CLIMBER-2) by Willeit *et al.* (2019). This study suggests a combination of gradual regolith removal and atmospheric CO_2 decline can explain the long-term climate variability over the past 3Myr. Second, against a longer-term reconstruction by van de Wal *et al.* (2011) that utilises deep-sea benthic isotope records to reconstruct a continuous CO_2 record over the past 20 Myr. Our simple GLS model demonstrates a similar long-term trend and timing of glacial-interglacial signals and an atmospheric CO_2 level that sits approximately mid-way between the two more complex models."

Comment: *"The recent review on the MPT (Berends et al., 2021b) gives also an idea about processes including a collection of CO_2 data and discusses a potential influence of the carbon cycle on the climate transition."*

We have included a reference to this excellent review (line 82) and have also referenced it throughout the paper.

Comment: *"While mentioning the call for the EPICA challenge, maybe also cite / discuss its results (Wolff et al., 2005). They have been shown on 2 posters at AGU fall meeting in 2004 (PDFs for download at: https://epic.awi.de/id/eprint/11721/, https://epic.awi.de/id/eprint/11722/), on which you see, that one of the participants to the challenge (N Shackleton) also used $\delta^{18}O$ to predict CO_2 for the 400-800 ky time window."*

We refer to the results of the EPICA challenge (particularly the precedence set by N Shackleton) at line 285 of the manuscript:

> "… the use of benthic $\delta^{18}O$ to predict atmospheric CO_2 has precedence; in response to the EPICA challenge (Wolff et al., 2004), N. Shackleton used this method to predict atmospheric CO_2 out to 800 kyr (Wolff, 2005). Furthermore, inverse modelling of CO_2 forced by the LR04 benthic stack has been undertaken by Berends et al. (2021a) and van de Wal et al. (2011)."