Rebuttal to the review by Anonymous Referee #1

We thank the reviewer for their comments on the manuscript and would hereby like to address the concerns they raised. Comments in italics, below our rebuttal. Page and line numbers refer to the revised manuscript.

> *The study by Berends et al. - if I understand it correctly - does not need either of the two to get the transitions and the change in cyclicity right. The entire record (including changes in amplitude and cyclicity) is entirely controlled by CO2. In particular, it has a constant relationship between d18O_benthic and CO2 over time and does not change the flow conditions at the bed. I highly recommend that the authors elaborate on this extensively in the Discussion and the Conclusions and discuss what this potentially implies for our understanding of the glaciation history.*

We use the "inverse modelling method" that does indeed use the benthic d18O record as forcing to calculate a CO2 history, but this is not based on a constant relationship. Eq. 1 relates the *rate of change of modelled CO2* to the *discrepancy between modelled and observed d18O*. The fact that it is *the rate of change of CO2* rather than the value itself mostly serves to suppress high-frequency oscillations in the benthic d18O record, resulting in a more "smooth" CO2 reconstruction. What's more important is the fact that this rate of change is determined by *the difference between modelled and observed benthic d18O*. Our modelled benthic d18O value is derived from both the deep-sea temperature (which is derived from the mean annual surface temperature, which depends on modelled CO2), and from modelled global ice volume. If, for example, the modelled benthic d18O value is too low (not "glacial" enough), then the inverse routine described by Eq. 1 will slowly decrease the modelled CO2. This will increase modelled d18O first by cooling the climate, lowering the deep-sea temperature (with a prescribed time-lag, using a moving average), and also due to the increase in global ice volume resulting from that cooling. When modelled benthic d18O is in line with the observed value, the downward trend in CO2 stops and the model stabilises, until the observed d18O value changes again.

We will extend and clarify the text describing the inverse routine, so that the reader will be able to comprehend the concepts behind it without having to read our earlier publications to which the manuscript referred.

P5 L10 – P7 L7: extended Sect. 2.1 (Methodology – Inverse modelling) so that it can be understood by readers unfamiliar with our previous work.

Regarding the implications of our CO2 reconstruction: the reviewer is correct in stating that our model does not account for possible changes in basal conditions during the MPT. This means that, in our view, our CO2 reconstruction represents how CO2 *should have evolved* over time in order the produce the observed d18O record, if no changes in basal conditions occurred. If we were to repeat our simulations and prescribe more basal sliding in pre-MPT Eurasia and North America, the reconstructed

CO2 would look different, likely both in terms of the glacial-interglacial amplitude, and the background trend. This is something we plan to investigate in future work, as it would provide useful context for interpreting the expected new ice-core record.

We agree that this is something that needs to be discussed in the manuscript. We will add a paragraph to the Discussion section.

P19 L4 – P19 L15: added a paragraph to the Discussion section.

> *Related to this, the authors set out in the (very nice!) introduction that these experiments will help us to understand and quantify Earth System Sensitivity. I regard it a missed chance that they do not pick up on this issue in the Conclusions as they seem to have all data at hand to contribute to this discussion. Given that the paper in its current form is relatively short, there is enough space to elaborate* on this.

The relation between atmospheric CO2, ice sheet geometry, and global climate is not explicitly included in our model. Rather, it is a result of the model physics of HadCM3, the GCM which was used to generate the "snapshots" included in our climate matrix. The Earth System Sensitivity that would be derived from our results would therefore just be that of HadCM3. Our model mainly provides insights into the long-term relation between CO2 and sea-level, which we believe we adequately discuss in the manuscript. The mention of earth system sensitivity in the abstract of our manuscript was inaccurate, we will change this.

P1 L9: changed phrasing in abstract to remove inaccurate mention of "Earth System Sensitivity".

> *I am puzzled by the way the deep ocean temperature, which influences the d18O_benthic signal, is calculated. In the manuscript and in Berends et al. (2019) the authors say that they used the global average of the surface temperature anomaly. Either there is some detail missing here (some scaling) or this seems to be at odds with the measured deep ocean temperature today. Here it is important to take into account that deep ocean temperatures have a strong bias towards the sea surface temperatures at deep water formation sites, which are located in high latitudes. There is also eddy diffusive transport of heat in lower latitudes, but the low deep water temperatures clearly point to a dominating role of deep water formation. In fact, the resulting deep ocean temperature, which is caused by the balance between advective transport of cold water from deep water formation and the diffusive entrainment of heat downwards, is also dependent on the strengths of the Atlantic Meridional Overturning Circulation (Galbraith et al., GRL 2016), which is also not included in the approach by Berends et al. The value of the deep ocean temperature used in the study by Berends is not mentioned in the paper. It is likely too warm, but this is also likely compensated by the CO2 sensitivity (120 ppm/permille) with which their approach is optimized. Even if there is some scaling involved that is not described*

in the manuscript, using the global average surface temperature appears to be an oversimplification. Accordingly, I think it is important to discuss this issue and use some alternative sensitivity runs, where the deep ocean temperature is parameterized by high latitude temperatures and the model CO2 sensitivity is recalibrated to show that the final result is not sensitive on the choice of the deep water temperature template.

Our modelled deep-sea temperature anomaly, which is used to calculate the modelled d18O, is derived from the northern high-latitude temperature anomaly (which is calculated as the mean temperature anomaly over the North America and Eurasia model grids), multiplied with a scaling factor (0.25 in our model) and smoothed over a 3,000 yr moving time window. This was not made clear in the manuscript; we will rectify this. The actual values this yields for deep-sea temperature changes are discussed in two of our earlier papers (Berends et al., 2018, 2019), and comparable with other studies showing a deep-sea cooling of 2 - 2.5 K during the last glacial maximum, which is agreement with proxy-based results (Shakun et al., 2015) We will add these numbers and references to the text.

P6 L19 – P6 L21: explained the relation between surface temperature, deep-sea temperature and d18O.

We also agree our method greatly simplifies the realistic relation between the global climate and benthic d18O. A more elaborate parameterisation based on ocean currents (which could be modelled ocean currents from the GCM snapshots, so that the relation can change over time using the same climate matrix) and global, spatially variable temperature anomalies, might be a significant improvement on this, without going so far as to run a fully isotope-enabled GCM. However, we believe that such work, while undoubtedly very interesting, lies beyond the scope of the current study. We will add a paragraph to the discussion section discussing this.

P18 L30 – P19 L2: added a paragraph to the Discussion section.

abstract line 2: "...over geological time scales... "

Added this.

abstract line 3: "...past CO2 concentrations, thus its radiative forcing, only... "

Added this.

abstract line 16-17: please do not use unexplained abbreviations such as KM5c and M2 in the abstract. Please also indicate such stages after you introduced them in the main text in the figures.

We have removed M2 and KM5c from the abstract. Since they were not mentioned anywhere else in the text, we did not add them to any figures.

page 2 line 4: "... atmosphere from the time..."

Changed this.

page 2 line 6 and throughout the manuscript: Here you cite Bereiter et al. (2015) for the CO2 record, but later you cite Lüthi et al., 2008. Note that in Bereiter et al., a correction of the EPICA Dome C CO2 values by Lüthi et al. was introduced for ice older than about 600 kyr. You should use the corrected Bereiter et al. data throughout the manuscript.

All figures, model-data comparisons, reported correlations, and citations have been updated to use the corrected Bereiter et al. data.

page 2 line 9: "... have measured d11B..."

Changed this.

page 4 line 14: "... can help interpreting..."

Changed this.

page 6 line 16: Here the d18O/CO2 scaling parameter is introduced. While it has been mathematically introduced in equation 1, it would be helpful to discuss the meaning of this parameter in more detail and also discuss what it implies if this parameter is assumed to be constant over time.

The meaning of this scaling parameter is explained in the new, extended section described the inverse modelling method.

page 8 line 14: it is not entirely clear to me what you mean by "combining the GCM snapshots according to position of the ice-sheet model in the climate matrix". I am sure there is an easier way to explain this.

Following a suggestion from reviewer #2, a short appendix has been added where the equations of the matrix method are presented and (briefly) explained.

page 10 line 6: "... which cover the last 500 kyr and 5.5 Myr, respectively."

Changed this.

page 11, line 2-4: here you say that the model study by Willeit does not include the Antarctic ice sheet, but the results look quite similar. What does this imply?

Reviewer #2 informed us that we missed a line in Willeit et al. (2019), stating that the Antarctic sea-level contribution is assumed to be 10 % of the of the northern hemisphere ice sheets. This is a reasonable assumption during the Pleistocene, but might produce larger errors during the warmer Pliocene. We will update the text accordingly. The reduced reliability of our own results for the late Pliocene (which we believe results in similar lower-than-expected sea levels during this period) is discussed in the newly added paragraph in the Discussion section.

page 11 line 10: "... close to those..."

Changed this.

page 12 line 9: "d11B"

Changed this.

page 12 caption Figure 6: The references are assigned wrongly in the caption. Hönisch et al. (2009) and Bartoli et al. (2015) use d11B, not alkenones. It is correctly referenced in the main text.

Changed this.

page 13 line 10: You write "which they prescribed". Who is "they" in this case, please provide the reference.

"They" are Willeit et al. Added this reference to the text.

page 13 line 19: "... different boron isotope based records is such..."

Changed this.

page 15 Figure 8: It would be helpful if the data younger than 800 kyr and those older than 800 kyr could be discerned in the model runs. Use different symbols or colors.

Changed this.

page 15 line 12: Here you say that the reconstructions by van de Wal and Willeit have a smaller spread. Do you mean in CO2 or sea level? In particular, for the reconstruction by Willeit I do not see a significantly smaller spread except that the sea level is capped.

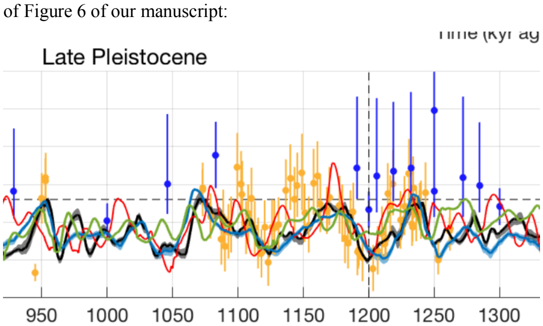
van de Wal and Willeit have less spread in sea level for each CO_2 value (or less spread in CO_2 for each sea level value), though this is indeed more obvious for van de Wal. We will clarify this in the text.

page 16 line 21: "d11B"

Changed this.

page 16 line 22-24: Here you say that the data/model comparison in terms of CO2 is not conclusive, but before you showed that the highest resolution d11B data by Chalk et al., show perfect agreement with your reconstructions. I think you undersell the d11B values. Clearly each individual d11B based CO2 value has an analytical uncertainty on the order of 20 ppm, but measured in high enough resolution/replication, these data are quite useful to validate your model results.

Chalk et al. (2017) published a collection of boron isotope proxy data around the MPT, shown in orange in the middle panel of Figure 6 of our manuscript:



While our results (and those of the other models) agree well with data points below 280 ppmv, there is a clear mismatch in both magnitude and timing of the high interglacial values around 1150 and 1100 kyr (which probably is related to our model's poor performance for warmer-than-present climates). We therefore do not believe that the agreement between these proxy data (yellow dots) and our model results (black line) can be called "perfect". Even if we assume that the autocorrelation of the individual proxy data points allows us to assume a lower measurement uncertainty, the difference between the median of the proxy data and the multi-model mean is much larger than the differences between the different models. This justifies our assertion that these proxy data cannot be used to choose one model reconstruction over the others.

We will clarify this line of reasoning in the manuscript.

P18 L15 – P18 L28: added a paragraph to the Discussion section.

page 16 line 32: "... for a colder-than-present..."

Changed this.

page 17 line 4: The last sentence is weak and does not give credit to the work performed in this study. The authors should elaborate much more on this, as outlined in my general comments above.

We have added two paragraphs to the Discussion section, discussing the various sources of uncertainty in our results, and how we believe our reconstruction should be interpreted.

Rebuttal to the review by Andrey Ganopolski

We thank the reviewer for their comments on the manuscript and would hereby like to address the concerns they raised. Comments in italics, below our rebuttal. Page and line numbers refer to the revised manuscript.

> *Method description. One of the problems for the readers of this manuscript is that the method used in this study has been developed over a long period time and its comprehensive description are scattered among a number of previous publications. Even although I was familiar with some of them, it took me a lot of time to get a more or less clear understanding of what authors are doing. Of course, one cannot expect such efforts from a typical reader. However, without a proper understanding of the method, the results presented in the manuscript are not very useful. This is why, I would suggest to make a more detailed description (including the key equations) in the appendix or supplementary information. In particular, I am curious how the effect of orbital forcing has been accounted for by the "matrix method".*

We agree that a thorough understanding of our methodology relied too much on concepts explained in earlier publications. We will revise and extend the sections describing both the inverse modelling routine and the matrix method.

Orbital forcing is included in our model in two separate ways. First, it is included as a term in the calculation of surface melt in the mass balance module (the insolation-temperature model). Second, it is included in the matrix method in the interpolation equation. The matrix method contains two separate interpolation equations that are used to combine the GCM snapshots; one for temperature, and one for precipitation. The one for precipitation is based on ice-sheet geometry, to account for orographic forcing of precipitation. The interpolation routine for temperature is based on "absorbed insolation": the product of insolation at the top of the atmosphere, and (1 – albedo). This accounts for both the changes in albedo resulting from ice-sheet advance and retreat, and the changes in insolation resulting from orbital forcing. While an extension of the climate matrix that includes GCM snapshots that have been calculated for different orbital configurations (so that the direct effect of insolation changes on surface temperature is also include) is planned, it is not included in the model version described here.

We will make sure that this is properly explained in the Methodology section.

P5 L10 – P6 L21: extended and clarified Sect. 2.1 (Methodology – Inverse modelling)
P19 L19 – P21 L8: added Appendix A, which presents and (briefly) explains the equations governing the matrix method

> *The model validation is based on the comparison of reconstructed CO2 over the past 800 kyr with the ice core data. The authors compare the results of their current study with several others and conclude that they are the best.*

However, it is obvious that comparison results of inverse modelling with forward modelling presented in Willeit et al (2019) is the same as comparison of apples with cucumbers. The inverse model is forced by benthic d18O which is already highly correlated with CO2 (correlation coefficient is 0.86). The authors should make this point very clear. The only surprising thing in this table is the extremely poor performance of Stap et al. (2017). Unfortunately, the authors themselves admit on page 15 that they cannot explain this fact. In fact, it is much more instructive to compare the result of a rather complex inverse modelling approach used by the authors to a simple linear regression CO2=175+50.2 (5.2-s), where "s" is 5000-years averaged d18O from LR04 stack. Surprisingly (or maybe not) this simple "model" outperforms Berends et al. Indeed, it has R2=0.71 (versus 0.68 in Berends et al.) and rms=13.8 ppm (vs. 15.3) for "simulated" CO2 concentration over the last 800 kyr. After such a comparison, the numbers in Table 1 do not look very impressive. For the rest of Quaternary, results of Berends et al. also do not differ much from this simple regression model. After all, it is rather expectable (and have been demonstrated by Willeit et al., 2019) that CO2 also followed ice volume variations during 41-kyr world but with a smaller amplitude.

We apologise to the reviewer if our manuscript seemed to suggest that our own method is "best"; we agree that there is no meaningful way to declare one modelling approach "better" than another. We also agree that the results of the simple linear regression proposed by the reviewer should be included in the comparison. Indeed, this helps to illustrate what we believe is the main conclusion from this comparison; more complex models, including more elaborate physics and describing more components of the Earth system, are useful for studying large-scale relations between these components, but are not necessarily better at resolving the evolution of a single component or parameter. We will clarify this in the text.

P13 L24 – P13 L31: added a simple linear regression to the statistical comparison

We will also follow a suggestion from Matteo Willeit (the main author of Willeit et al., 2019, who contacted us shortly after the discussion version of our manuscript was published, with a similar question about the comparison of correlations to the ice core record), to include an optimised time lag for each model CO2 reconstruction before calculating the correlations with the ice core record. This mainly affects the results of Stap et al. (2017) and Willeit et al. (2019), where the coefficients of determination R^2 for both studies increase from ~0.25 to ~0.45. While this extra step does not alter the conclusions we draw from this comparison, it does more properly give credit to the results of the different studies.

Following a comment by Anonymous Reviewer #1, we have also updated all figures and numbers to use the more recent ice-core CO2 record by Bereiter et al. (2015), rather than Lüthi et al. (2008).

The real question is what was CO2 concentration at the end of Pliocene. And here I see a real problem with the results presented in Berends et al. Indeed, if during the entire Pleistocene, CO2, ice volume and d18O variations *were essentially identical, during the late Pliocene CO2 get really wild. Figure 4 shows several CO2 oscillations with the amplitude above 100 ppm. Of course, this is not 200+ ppm as in Stap et al (2017) but still a lot. As the scientist who has been heavily involved in explaining glacial-interglacial CO2 variability, I must confess that it is extremely difficult to explain 80 ppm change in CO2 concentration even for the full glacial cycles of the late Quaternary. What could cause even larger Pliocene variations in CO2 without any obvious external forcing, the authors do not explain. This is why I strongly suspect that the reason for such weird behaviour of CO2 before Pliocene-Pleistocene transition is that the inverse modelling of CO2 concentration based on benthic d18O beyond Quaternary represents an ill-posed problem.*

We agree that our results for the late Pliocene are by no means the definitive answer to the question of how the world looked like in terms of CO2, global climate and ice sheet geometry. In our view, the main problem here is the relatively large changes in benthic d18O. Explaining these requires either very large changes in deep-sea temperature, moderately large changes in global ice volume, or (more likely) a mix of both. Our model results tend towards the "temperature" end of this spectrum, resulting in large changes in CO2: almost 100 ppmv difference between the coldest point of the Pliocene during M2, 3.3 Myr ago, and the warmest point during KM5c, 3.205 Myr ago (for our default simulation; in the low-CO2 end member, this difference reduces to about 85 ppmv).

We suspect that the relative sparsity of our climate matrix for warm worlds might result in a bias towards lager ice sheets in warm climates. This means that increasing modelled CO2 above present-day levels does not cause as much ice-sheet retreat as it maybe should, so that benthic d18O does not decrease so much. In order to reproduce the observed d18O record, the inverse routine will compensate by increasing CO2 until the resulting deep-sea temperature change is enough to produce the required change in d18O. This also explains the very large uncertainty range resulting from our sensitivity analysis; an additional change of 0.1 per mille in d18O, for constant ice volume, requires a very large change in deep-sea temperature, mean annual surface temperature, and CO2. Changing our climate matrix such that warm climates will lead to more ice-sheet retreat will essentially shift the blame for the high d18O variability from the temperature end of the spectrum towards the ice volume end; it will reduce the modelled CO2 variability during the late Pliocene, but it will also increase the sea-level high stands.

Lastly, looking at proxy-based reconstruction, the boron isotope data by Martínez-Botí et al. (2015), which has both the highest temporal resolution and longest temporal range of all available proxies, shows a variability during the late Pliocene of about 150 ppmv (albeit with an uncertainty of about 100 ppmv in either direction). While certainly not definitive, especially considering the large discrepancies between difference boron-based reconstructions (as also discussed in our manuscript), their data does seem to suggest strong CO2 variability in warmer-than-present worlds.

We will extend the Discussion section of the manuscript to reflect these thoughts.

P18 L15 – P18 L28: added a paragraph to the Discussion section, discussing the variability and uncertainty in our CO2 reconstruction during the late Pliocene.

> *The authors wrote on page 9 that "uncertainties are conservative in this study". What the authors mean under "conservative" is not clear to me. To me, the estimate of uncertainties in this study is overoptimistic at best. Even if the maximum error in benthic d18O is indeed only 0.1 promile, the methodology has a number of other uncertainties related both to forward model and to conversion between climate characteristics (ice volume, temperature) and d18O. For the large glacial cycles of Quaternary even a larger uncertainty still does not prevent a reasonable estimate of CO2 but the situation is very different prior to 2.7 Ma. Before Quaternary, the model "assumes" very little variability in global ice volume and thus most of d18O variability has to be attributed to CO2 change and this is precisely what the model does. However, in this case, even uncertainty of +-0.1 promile already constitutes a serious problem. Indeed, 0.2 promile correspond to about 1C change in the deep-water temperature which in turn corresponds to 1.5C in global air temperature. The later number corresponds to change of CO2 (assuming climate sensitivity =3C) from 280 to 400 ppm. Thus, even with a very optimistic estimate of the method uncertainty, for pre-Quaternary climates this method cannot distinguish between a possibility that CO2 was as low as the preindustrial one or that it was as high as the current one. Obviously, such "reconstruction" is not very helpful.*

The statement that "uncertainties are conservative" was intended to indicate that the uncertainties we report are only those that arise from the sensitivity analysis described in the manuscript. This is simply the sensitivity of the model to uncertainties in the d18O record, which we showed in our 2019 publication to be larger than the sensitivity to uncertainties in other model parameters. We agree with the reviewer that this is not at all the same as the real uncertainty in our results; there are many other factors introducing uncertainties that cannot be quantified through such sensitivity analysis, and these are likely to be larger still than the numbers we report.

We will extend the Discussion section of the manuscript to reflect this, especially in relation to the previous comment about the difficulty of interpreting the d18O record in the late Pliocene.

P18 L15 – P18 L28: added a paragraph to the Discussion section, discussing the variability and uncertainty in our CO2 reconstruction during the late Pliocene.

> *"80/120 kyr cycles". Although this is not very essential for the manuscript under consideration, but the authors used the expression "80/120 kyr cycles" (actually it should be 82/123) several times in this and previous papers which*

provokes me to make the following comment: The durability of "two or three obliquity cycles" myths is amazing since it is not supported by real data! Glacial cycles of the late Quaternary have average periodicity close to 100 kyr which explains strong 100 kyr peak in the frequency spectra of ice volume. It is true that the durations of individual glacial cycles deviate significantly from 100 kyr but they also do not cluster around 80 and 120 kyr (see for example Table 1 in Konijnendijk et al., 2015). In fact, durations of individual glacial cycles are relatively uniformly distributed between 80 and 120 kyr with half of the cycles been closer to 100 kyr than to 80 or 120 kyr.

While we believe that declaring the 80/120 kyr hypothesis to be a "myth" is overly dismissive of the studies supporting this hypothesis (especially when considering the difficulties in constructing insolation-independent age models, described by Huybers and Wunsch, 2004), we agree with the reviewer that it is important to mention the ongoing discussion about the nature of the late Pleistocene glacial cycles. We will clarify this in the manuscript.

P14 L11 – P14 L17: added a few lines about the 100 vs 80/120 kyr discussion.

P.3, L.8 "proxies for global mean temperature"? Greenland and Antarctic records present proxies only for local temperatures which differ significantly from the global one

Changed this.

P3., L.10. "In that case ocean water temperature can be resolved as closure term from the benthic signal" This is not clear

Changed this.

P. 4, L.9 The definition of "entire climate system (atmosphere, ocean, cryosphere, car- bon cycle, etc.)" is not consistent with contemporary terminology. Such system is named Earth system and Earth system models describe not only "physical processes" (L. 10).

Changed this.

p. 4, L. 21 "the known relations between atmospheric CO2, global temperature and climate, and ice-sheet evolution". Why authors think that these relations are "known". Even the relation between CO2 and global temperature is still not well-known.

We agree that our phrasing was unclear. We will clarify this in the manuscript.

p. 5. I am not sure I understand why the authors put "data" and "model" in quotes.

These words are put in quotes because they constitute rather informal, but also obvious, descriptions of the aim and general approach of the studies we describe. While it is not possible (or desirable!) to draw a clear line between data studies and model studies, many publications about paleoclimate either rely mostly on the presentation and interpretation of proxy data, or on the development and application of modelling methods. We feel that it is important to explain the distinction between them, and how our work relates to both.

P.7 L. 11 "The reconstruction by Laskar et al. (2004) is used to prescribe time- and latitude-dependent insolation". Insolation is not reconstructed by computed using physical laws. This is why orbital forcing can be calculated for the past and future with the same (very high) accuracy.

We will replace "reconstruction" by "solution", in line with the phrasing by the authors of Laskar et al. (2004).

P. 11, L. 2. "so any possible contribution from Antarctica to changes in sea-level ... is not accounted for in their reconstruction". This is an incorrect statement. It is written in Willeit at al. (page 6) "Sea level is computed from the volume of modeled NH ice sheets assuming an additional 10% contribution from Antarctica".

We will correct this in manuscript.

P. 15, L. 15. "... show a CO2 "threshold" for glaciation and sea-level drop around 250 ppmv". Our studies (e.g Ganopolski et al., 2016) do not support the existence of a single CO2 threshold for glaciations. To the contrary, we found that glacial inception is determined by a combination of insolation and logarithm of CO2 concentration.

We will correct this in manuscript.

Fig. 4. It is not explained what shading shows in this figure.

Shaded areas indicates the uncertainty in the LR04 benthic d18O stack, and the resulting uncertainty in the reconstructed CO2 and sea level. We will clarify this in the manuscript.

The reference Stap et al. (2017) is not in the reference list.

Added this reference.

Rebuttal to comments by the editor, mr. Brook

Dear mr. Brook,

Thank you for your helpful comments on our manuscript. We will here provide a point-by-point answer. Comments in italics, below our rebuttal. Page and line numbers refer to the revised manuscript.

1) *I'd like to emphasize the point from A. Ganapolski that the modeling strategy should be outlined more comprehensively. My own concern is to understand what happens at each time step in the model vs. what happens as time progresses.*

We agree with the editor, and with mr. Ganopolski, that the description of our modelling methodology in the previous version of the manuscript relied too much on references to previous work, making it difficult to follow for someone not as intimately familiar with those publications as we are. To remedy this, we have altered and extended the sections describing both the inverse forward modelling method, and the climate matrix method. The inverse modelling section should now provide the reader with a qualitative understanding of what exactly it is that this method does, i.e. solving the inverse problem of how atmospheric CO_2 should have evolved in the past in order to produce the observed benthic $d^{18}O$ record. Regarding the climate matrix method, we have added a short appendix to the manuscript where the key equations of this method are presented and explained. This should make it clear to the reader how the matrix method translates changes in modelled ice-sheet geometry to changes in applied climate forcing.

2) *The term "inverse-forward modeling" is also somewhat confusing to me. If this is established in previous work then it is fine.*

This term has indeed been consistently used in all of our previous publications using this methodology.

3) *Page 2, line 5. Technically the air is not in bubbles in the entire ice core, in deeper sections it is in hydrate form.*

Changed this line to read "…oldest ice core presently available, the EPICA Dome C core, contains ice, and air bubbles and hydrates, dating back 800 kyr…".

4) *Page 2, line 13. I believe it is the 13C/1C ratio in alkenones that is used to reconstruct CO2, not ratios of different alkenones (which is used for temperature). I Amy have misunderstood the intent of the sentence.*

The alkenone temperature proxy is indeed based on the $d^{13}C$ of the alkenones. We will correct this in the manuscript.

5) *Page 2, line 16-21. Could mention work of Franks et al. 2014, GRL on alternate ways to use leaf data to reconstruct CO2.*

This is indeed a valuable reference to cite. We will include it in the introduction section.

6) *Page 8, line 16. Something missing in this sentence.*

Corrected this.

1) *Page 9, line 3. Off?*

Corrected this.

2)
3) *Page 9, line 8-14. This sentence is difficult to follow.*

Corrected this.

4) *Page 19, line 9. Here and elsewhere I believe d12B should be d11B.*

Corrected this.

Reconstructing the Evolution of Ice Sheets, Sea Level and Atmospheric CO_2 During the Past 3.6 Million Years

Constantijn J. Berends[1], Bas de Boer[2], Roderik S. W. van de Wal[1,3]

[1] Institute for Marine and Atmospheric research Utrecht, Utrecht University, Utrecht, The Netherlands
[2] Earth and Climate Cluster, Faculty of Science, Vrije Universiteit Amsterdam, The Netherlands
[3] Faculty of Geosciences, Department of Physical Geography, Utrecht University, Utrecht, The Netherlands

Correspondence to: Constantijn J. Berends (c.j.berends@uu.nl)

Abstract. Understanding the evolution of, and the interactions between, ice sheets and the global climate over geological time scales is important for being able to project their future evolution. However, direct observational evidence of past CO_2 concentrations, and the implied radiative forcing, only exists for the past 800,000 years. Records of benthic $\delta^{18}O$ date back millions of years, but contain signals from both land ice volume and ocean temperature. In recent years, inverse forward modelling has been developed as a method to disentangle these two signals, resulting in mutually consistent reconstructions of ice volume, temperature and CO_2. We use this approach to force a hybrid ice-sheet – climate model with a benthic $\delta^{18}O$ stack, reconstructing the evolution of the ice sheets, global mean sea level and atmospheric CO_2 during the late Pliocene and the Pleistocene, from 3.6 million years (Myr) ago to the present day. During the warmer-than-present climates of the Late Pliocene, reconstructed CO_2 varies widely, from 320 – 440 ppmv for warm periods, to 235 – 250 ppmv for the early glacial excursion ~3.3 million years ago. Sea level is relatively stable during this period, with maxima of 6 – 14 m, and minima of 12 – 26 m during glacial episodes. Both CO_2 and sea level are within the wide ranges of values covered by available proxy data for this period. Our results for the Pleistocene agree well with the ice-core CO_2 record, as well as with different available sea-level proxy data. During the early Pleistocene, 2.6 – 1.2 Myr ago, we simulate 40 kyr glacial cycles, with interglacial CO_2 decreasing from 280 – 300 ppmv at the beginning of the Pleistocene, to 250 – 280 ppmv just before the Mid-Pleistocene Transition (MPT). Peak glacial CO_2 decreases from 220 – 250 ppmv to 205 – 225 ppmv during this period. After the MPT, when the glacial cycles change from 40 kyr to 80/120 kyr cyclicity, the glacial-interglacial contrast increases, with interglacial CO_2 varying between 250 – 320 ppmv, and peak glacial values decreasing to 170 – 210 ppmv.

Deleted: constrain earth system sensitivity

Deleted: such as Marine Isotope Stage (MIS) KM5c

Deleted: MIS M2

Deleted: a high stand

Deleted: a drop

Deleted:

Deleted: MIS M2

1 Introduction

Understanding the response of ice sheets and the global climate as a whole to changes in the concentrations of atmospheric CO_2, is important for understanding the future evolution of the climate system. Since large-scale changes in ice sheet geometry typically occur over thousands to tens of thousands of years, sources of information other than direct observational evidence are required. In order to gain more insight in the relation between these components of the Earth system, studying their evolution during the geological past is useful.

One particularly rich source of information is presented by ice cores. The chemical and isotopical content of the ice itself can provide valuable information on the state of the global climate, and in particular on the temperature, at the time the ice was formed (Dansgaard, 1964; Jouzel et al., 1997; Alley, 2000; Kindler et al., 2014). Air bubbles, trapped when the snow compresses first into firn and ultimately into ice, contain tiny samples of the atmosphere from the time the air got trapped. The oldest ice core presently available, the EPICA Dome C core, contains ice, and air bubbles and hydrates, dating back 800 kyr (Bereiter et al., 2015). The information obtained from these ice cores has greatly improved our understanding of the dynamics of the glacial cycles of the Late Pleistocene.

Different methods of relating chemical and isotopic properties of ocean sediments to the atmospheric CO_2 concentration have been used to create proxy data extending back further in time than the ice core record. Many studies have measured $\delta^{11}B$ of different fossil foraminifera, using the observed relation to seawater pH to calculate atmospheric CO_2 concentrations (Hönisch et al., 2009; Seki et al., 2010; Bartoli et al., 2011; Martínez-Botí et al., 2015; Foster and Rae, 2016; Stap et al., 2016; Chalk et al., 2017; Dyez et al., 2018; Sosdian et al., 2018). Another line of evidence has focused on the $\delta^{13}C$ of alkenones in fossil foraminifera (Seki et al., 2010; Badger et al., 2013; Zhang et al., 2013), although the reliability of this proxy for CO_2 concentrations lower than ~400 ppmv has recently been called into question (Badger et al., 2019). A different line of work relates the density of stomata on fossil plant leaves to atmospheric CO_2 concentration, providing data throughout the Cenozoic (Kürschner et al., 1996; Wagner et al., 2002; Finsinger and Wagner-Cremer, 2009; Beerling and Royer, 2011; Stults et al., 2011; Bai et al., 2015; Hu et al., 2015). This proxy has been the subject of discussion regarding its reliability (Indermühle et al., 1999; Jordan, 2011; Porter et al., 2019), based on its discrepancies with the ice-core record, and presently unresolved problems due to the influence of other effects such as evolution, extinction, changes in local environment, and immigration of species, although recent work has gone some way to resolving these issues (Franks et al., 2014).

Information about the global glacial state is furthermore obtained from the $\delta^{18}O$ of fossil benthic foraminifera, which is influenced by both total ice volume and deep-sea temperature. Ocean sediment cores containing foraminiferal shells have been used to create stacks of benthic $\delta^{18}O$ records dating back 65 Myr (Lisiecki and Raymo, 2005; Zachos et al., 2001, 2008). Fig. 1 shows the LR04 stack of benthic $\delta^{18}O$ records (Lisiecki and Raymo, 2005), together with the EPICA Dome C CO_2 record (Bereiter et al., 2015) and the Earth's orbital parameters and Northern hemisphere summer insolation (Laskar et al., 2004), during the last 800 kyr.

Deleted: world's

Deleted: as it was at

Deleted: the

Deleted: ratio of different types

Deleted: .

Deleted: Lüthi et al., 2008

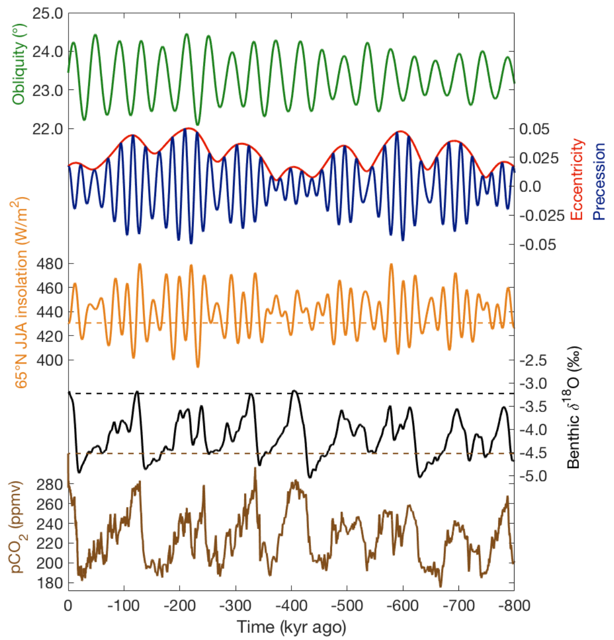


Figure 1: Earth's orbital parameters (obliquity, eccentricity and precession) and the resulting Northern summer insolation (Laskar et al., 2004), benthic δ^{18}O (Lisiecki and Raymo, 2005) and ice core CO_2 (Bereiter et al., 2015) during the past 800 kyr.

Since benthic δ^{18}O contains both an ice-based and a climate-based signal, using it to understand the evolution of either of these two Earth system components is only possible after disentangling the two contributions. Several studies have aimed at performing this separation by deriving either of the two signals from independent other proxies. One approach has been to derive deep-water temperatures from foraminiferal Mg/Ca ratios (Sosdian et al., 2009; Elderfield et al., 2012; Shakun et al., 2015). Global mean sea level is then solved as a closure term. Alternatively, temperature proxies have been obtained from records of oxygen and hydrogen isotope abundances from ice cores from Greenland (Alley, 2000) and Antarctica (Jouzel et al., 2007). Focusing on the signal from ice volume rather than temperature, Rohling et al. (2014) used planktic δ^{18}O in the Mediterranean Sea as a proxy for sea-level at the Strait of Gibraltar, translating the result into a global mean sea-level record spanning the last 5.5 Myr. Grant et al. (2014) applied the same method to the Red Sea, producing a record of sea-level at the Bab-el-Mandab Strait with a higher accuracy, but going back only 500 kyr.

Deleted: Lüthi et al., 2008

Deleted: for global mean temperature

Deleted: In that case ocean water temperature can be resolved as closure term from the benthic signal.

These studies are generally viewed as "data" studies; observed variables (chemical concentrations, isotope ratios, etc.) are used to derive climate parameters (ice volume, global mean temperature, pCO_2, etc.) using relatively simple, time-independent concepts. Discussions about the validity of the results generally focus less on the physical processes described by these simple models, and more on the properties of the data, such as sample contamination and diagenesis, statistical limitations and measurement uncertainty.

A different family are the "model" studies, where physics-based models are used to describe the evolution of (components of) the Earth system (atmosphere, ocean, cryosphere, carbon cycle, etc.) through time. Typically, the aim of a model study is to determine if our understanding of a physical process is good enough to explain the observations, or to use that understanding, to make predictions of that process in the future. Several recent studies have aimed to reproduce the evolution of global climate and the cryosphere throughout the Pleistocene, using primarily insolation as forcing (Brovkin et al., 2012; Willeit et al., 2015, 2019). By studying the differences between model results and observations from proxy data, such studies can help with interpreting these data, providing insight into the physical processes that govern the relation between observed and derived variables. In turn, the differences between data and model results can help to assess the importance of physical processes that are not included in the model.

Inverse modelling is a hybrid method, using an approach that combines elements from both these families of research. This method aims to derive the evolution of the entire global climate-cryosphere system through time, based on observations of benthic $\delta^{18}O$ (Bintanja et al., 2005; Bintanja and van de Wal., 2008; de Boer et al., 2010, 2012, 2013, 2014, 2017; van de Wal et al., 2011; Stap et al., 2017, 2017; Berends et al., 2018, 2019). This is done by using ice-sheet models, with climate forcing components varying in complexity from simple parametrisations to fully coupled GCMs, to reconstruct ice-sheet evolution. Such models can be used to calculate the contributions to the observed benthic $\delta^{18}O$ signal from both ice volume and deep-water temperature over time. By using a tool called an "inverse routine", the model can be forced to (almost) exactly reproduce the benthic $\delta^{18}O$ signal, providing a reconstruction of global climate *as it should have evolved* in order to explain the observations. The advantage of this approach over other proxy-based reconstructions is that the simulated changes in global climate and ice-sheet evolution are mutually consistent, building on the physical equations within the model framework. Earlier studies adopting this approach used simple, one-dimensional ice-sheet models to represent all ice on Earth or on a single hemisphere, and different parameterisations of the relation between ice volume and climate (e.g. de Boer et al., 2010; Stap et al., 2017). More recent studies have used more elaborate 3-D ice-sheet models covering different regions of the Earth, using more comprehensive mass balance parameterisations and representations of the global climate (e.g. Bintanja and van de Wal, 2008; de Boer et al., 2013, 2017; Berends et al., 2018, 2019). The most recent of these studies, by Berends et al. (2019), used a 3-D ice-sheet-shelf model, forced with output from several different GCM simulations by a so-called "matrix method" of model coupling, to simulate the last four glacial cycles. They showed that their results agreed with geomorphological and proxy-based evidence of ice-sheet volume and extent, benthic $\delta^{18}O$, deep-water temperature, ice-sheet temperature and atmospheric CO_2.

Deleted: al

Deleted: entire climate system

Deleted: nothing but reconstructions of

Deleted: of

Deleted: , to describe the known relations between atmospheric CO_2, global temperature and climate

Deleted: and

Deleted: 6

The work presented here builds on the work by Berends et al. (2019), extending their results to 3.6 Myr ago to produce a time-continuous, self-consistent reconstruction of atmospheric CO_2, global climate, ice-sheet geometry and global mean sea-level, over the period where ice core records of CO_2 are not available. The inverse modelling approach that was adopted to force this model with a benthic $\delta^{18}O$ record is described in Sect. 2.1. The ice-sheet and climate models are described in Sects. 2.2 and 2.3, respectively, while the matrix method used to couple these two models is described in Sect. 2.4. The resulting reconstructions of CO_2, ice volume, temperature and sea-level over the past 3.6 Myr are presented in Sect. 3 and discussed in Sect. 4.

2 Methodology

2.1 Inverse modelling

In order to disentangle the contributions to the benthic $\delta^{18}O$ signal from global land ice volume and deep-water temperature changes, we use the inverse modelling method proposed by Bintanja and van de Wal (2008) and refined by de Boer et al. (2013, 2014, 2017) and Berends et al. (2019). The two contributions to the benthic $\delta^{18}O$ signal are not independent; both result from changes in the Earth's climate, driven by changes in insolation and CO_2. A cooling of the global climate will, over time, affect the temperature in the deep ocean, especially when the cooling occurs at high latitudes, where deep water formation occurs. Simultaneously, such a global cooling leads to an increase in ice volume, which affects the $\delta^{18}O$ of the sea water itself. The inverse modelling method is a tool that determines how modelled CO_2 *should have evolved over time* in order to affect the global climate in such a way that the observed benthic $\delta^{18}O$ record is reproduced. The modelled value for CO_2 is used, together with the ice sheets simulated by the ice-sheet model, to determine the global climate by using a matrix method of model forcing. This method was presented by Berends et al. (2018) and refined by Berends et al. (2019), and is described in more detail in Sect. 2.4. The resulting climate is used to determine the surface mass balance over the ice sheets, which forces the ice-sheet model. The resulting global ice volume and deep-water temperature anomaly (which is derived from the high-latitude annual mean surface temperature anomaly, using a moving average of 3,000 yr) are used to calculate a modelled value of benthic $\delta^{18}O$. This approach is visualized conceptually in Fig. 2.

Formatted: Font: Italic

Deleted: (in terms of temperature and precipitation, which control the evolution of the ice sheets)

Deleted: changes in

Deleted: are

Deleted: In the inverse modelling method, t

Deleted: the

Deleted: global

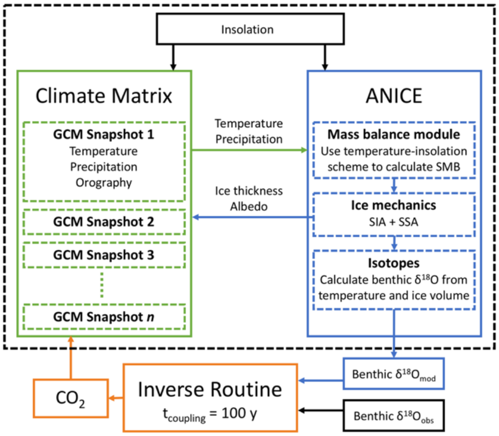


Figure 2: A conceptual visualization of the inverse forward modelling approach. The model is forced externally by an insolation reconstruction and a benthic δ^{18}O record (black boxes). pCO_2 is changed over time based on the difference between observed and modelled δ^{18}O. The climate matrix interpolates between the pre-calculated GCM snapshots, based on the prescribed pCO_2 value and the modelled state of the cryosphere (ice thickness and albedo), to determine the climate that is used to calculate the surface mass balance over the ice sheets. The modelled ice sheets and climate are then used to calculate a modelled benthic δ^{18}O value, which is used to update modelled CO_2 in the next time step. Figure adapted from Berends et al. (2019).

The modelled value $\delta^{18}O_{mod}$ is compared to the observed value $\delta^{18}O_{obs}$ at the model time *t*. A modelled value that is too low indicates that the modelled climate is too warm, or the ice sheets are too small. pCO_2 is then decreased in the next time-step (with an increment proportional to the δ^{18}O discrepancy). This relationship is quantified by the following equation:

Deleted: , leading to both a cooling and a subsequent increase in ice volume.

$$pCO_2 = \overline{pCO_2} + \alpha\left(\delta^{18}O_{mod} - \delta^{18}O_{obs}\right). \tag{1}$$

Here, $\overline{pCO_2}$ is the mean modelled pCO_2 over the preceding 8.5 kyr, and α is a scaling parameter which controls how fast modelled CO_2 is allowed to change (serving as a low-pass filter, suppressing overshoot that could result from large changes in the benthic δ^{18}O record). In order to calculate $\delta^{18}O_{mod}$, the contributions from modelled ice volume and deep-sea temperature are calculated separately. The spatially variable isotope content of the individual ice-sheets is tracked through time, with the surface isotope balance based on the observed present-day relation between precipitation rates and isotope content according to Zwally and Giovinetto (1997). Benthic δ^{18}O is assumed to be linearly dependent on the global mean deep-water temperature anomaly, which is calculated by temporally smoothing the high-latitude annual mean surface temperature anomaly over the preceding 3 kyr and multiplying with a scaling factor of 0.25. Berends et al. (2019) showed that this yields a deep-sea temperature anomaly of 2 – 2.5 K at LGM, in general agreement with proxy-based reconstructions (Shakun et al., 2015). The

Deleted: global

values of 8.5 kyr for the length of the CO_2 averaging window, 3 kyr for the deep-water temperature averaging window, and 120 *ppmv ‰⁻¹* for the $\delta^{18}O$-CO_2 scaling parameter, were determined experimentally to accurately reproduce the observed benthic $\delta^{18}O$ record.

The combination of this inverse modelling method to reconstruct pCO_2 and the matrix method to determine the global climate has been shown to accurately reproduce changes in benthic $\delta^{18}O$ and the individual contributions from both global ice volume and deep-water temperature, as well as ice-sheet volume and extent, ice-sheet surface temperature and pCO_2 during the last four glacial cycles (Berends et al., 2019).

2.2 Ice-sheet model

The evolution of the ice sheets is simulated using ANICE, a coupled 3-D ice-sheet-shelf model (de Boer et al., 2013; Berends et al., 2018). ANICE uses a combination of the shallow ice approximation (SIA; Morland and Johnson, 1980) for grounded ice and the shallow shelf approximation (SSA; Morland, 1987) for floating ice to solve the ice mechanical equations. Basal sliding is described by a Coulomb sliding law and solved using the SSA, using the hybrid approach by Bueler and Brown (2009), where basal friction is determined by bedrock elevation. Internal ice temperatures, used to calculate ice viscosity, are calculated using a coupled thermodynamical module. The surface mass balance is parameterized based on monthly mean surface temperature and precipitation, where ablation is calculated using the surface temperature-albedo-insolation parameterization, as explained in more detail by Berends et al. (2018). The solution by Laskar et al. (2004) is used to prescribe time- and latitude-dependent insolation at the top of the atmosphere. A combination of the temperature-based formulation by Martin et al. (2011) and the glacial-interglacial parameterization by Pollard & DeConto (2009), is used to calculate sub-shelf melt underneath the Antarctic ice shelves, calibrated by de Boer et al. (2013) to produce realistic present-day Antarctic shelves and grounding lines. A more detailed explanation is provided by de Boer et al. (2013) and references therein. A simple threshold thickness of 200 m is used to describe ice calving, where any shelf ice below this threshold thickness is removed. The model is run on four separate grids simultaneously, covering North America, Eurasia, Greenland and Antarctica, as shown in Fig. 3. The horizontal resolution is 40 km for Antarctica, North America and Eurasia, and 20 km for Greenland.

Deleted: The SIA and SSA ice velocities are averaged near the grounding line according to the approach by Winkelmann et al. (2011), as explained in detail by de Boer et al. (2013).

Deleted: reconstruction

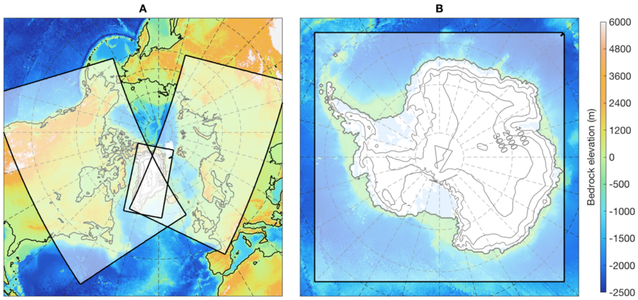


Figure 3: The areas of the world covered by the four model domains of ANICE. In the North America and Eurasia domains, Greenland is omitted. Figure adapted from Berends et al. (2018).

2.3 Climate model

HadCM3 is a coupled atmosphere-ocean GCM (Gordon et al., 2000; Valdes et al., 2017), which has been shown to accurately reproduce the present-day climate heat budget (Gordon et al., 2000). It has been used for future climate projections in the IPCC AR4 (Solomon et al., 2007), and paleoclimate reconstructions such as PlioMIP (Haywood and Valdes, 2003; Dolan et al., 2011, 2015; Haywood et al., 2013) and PMIP2 (Braconnot et al., 2007). Atmospheric circulation is calculated at a resolution of 2.5 ° latitude by 3.75 ° longitude. The ocean is modelled at a horizontal resolution of 1.25 ° by 1.25 °, with 20 vertical layers. We use results from several different steady-state time slice simulations with HadCM3 of different climate states to force our ice-sheet model, using the matrix method explained in Sect. 2.4.

2.4 Matrix method

According to Pollard (2010), a climate matrix is a collection of pre-calculated output data from several steady-state GCM simulations, called "snapshots". These snapshots differ from each other in one or more key parameters, such as orbital configuration, prescribed atmospheric pCO2, or ice-sheet configuration. Each of these constitutes a separate dimension of the matrix. When using an ice-sheet model to simulate the evolution of an ice sheet over time, the prescribed climate is determined in every model time-step by combining the GCM snapshots according to the position of the ice-sheet model state in the climate matrix, which is determined by the modelled values of the parameters describing the snapshots. This approach occupies the middle ground between offline forcing and fully coupled ice-sheet – climate models. The different GCM snapshots contain the key feedback effects of the altitude and albedo on the temperature. In addition, the matrix method captures the effect of ice sheet geometry on large-scale atmospheric circulation and precipitation. The matrix method creates a spatially variable linear interpolation of these snapshots providing a first order approach to the strength of the feedback and the effects of ice-sheet geometry on atmospheric circulation and precipitation.

Deleted: -temperature

Deleted: -

In this study, we use the matrix method developed by Berends et al. (2018), where temperature fields from the different climate states are combined based on a prescribed value for pCO_2 and on the internally modelled ice-sheets. The feedback of the ice sheets on the climate is calculated via the effect on absorbed insolation through changes in surface albedo. This interpolation is carried out separately for all four ice sheets. The altitude-temperature feedback is parameterized by a constant lapse-rate derived from the GCM snapshots. Precipitation fields are combined based on changes in surface elevation, reflecting the orographic forcing of precipitation and plateau desert effect caused by the presence of a large ice-sheet. The equations describing these calculations are presented in Appendix A, and explained in more detail by Berends et al. (2018), who demonstrated the viability of this method by simulating the evolution of the North American, Eurasian, Greenland and Antarctic ice-sheets throughout the entire last glacial cycle at the same time. They showed that model results agree well with available data in terms of ice-sheet extent, sea-level contribution, ice-sheet surface temperature and contribution to benthic $\delta^{18}O$. Here, we apply this matrix method to the climate matrix created by Berends et al. (2019), consisting of eleven pre-calculated GCM snapshots, created with HadCM3. Two of these, produced by Singarayer and Valdes (2010), respectively represent the pre-industrial period (PI) and the last glacial maximum (LGM). The other nine, produced by Dolan et al. (2015), represent the global climate during Marine Isotope Stage (MIS) M2 (3.3 My ago), for four different possible ice-sheet geometries and two different pCO_2 concentrations, plus one Pliocene control run. The total set of eleven snapshots allows the climate matrix to disentangle the effects on climate of changes in pCO_2 and changes in ice-sheet extent, and provides information on climates that are both colder and warmer than present-day.

Deleted: f

3 Experimental set-up and results

Here, we describe a set of simulations from 3.6 Myr ago to the present day. The choice for this starting point aims to include the end of the Late Pliocene and the inception of the Pleistocene glacial cycles. The model was initialized with the same PRISM3 ice sheets that were used as boundary conditions in several of the HadCM3 simulations by Dolan et al. (2015). The model was forced with the LR04 stack of benthic $\delta^{18}O$ records (Lisiecki and Raymo, 2005). We chose here to perform three simulations: one "default" run, one with the $\delta^{18}O$ forcing adjusted upwards by 0.1 ‰ and one with the forcing adjusting downwards by 0.1 ‰, as this uncertainty dominates the total uncertainty in the reconstructed CO_2 and global mean sea-level. Since simulating such long periods of time is very computationally demanding, we decided that the added value of including other simulations (as in the ensemble by Berends et al., 2019) did not outweigh the computational cost. As a result, uncertainties are conservative in this study, describing only the uncertainty arising from the uncertainty in the benthic $\delta^{18}O$ record.

Deleted: that were included in the climate matrix

The resulting reconstructions of CO_2 and global mean sea-level, together with the $\delta^{18}O$ forcing, are shown in Fig. 4. The 40 kyr glacial cycles of the Early Pleistocene, between 2.6 and 1.2 Myr ago, are clearly visible, changing to 80/120 kyr cycles after the MPT. The visibly larger uncertainty for higher CO_2 concentrations, occurring during the warm periods of the late Pliocene, is likely due to the relative sparsity of the climate matrix for warmer-than-present states (three snapshots only) compared to the colder-than-present part (seven snapshots), which might lead to a bias towards larger ice sheets in warm worlds.

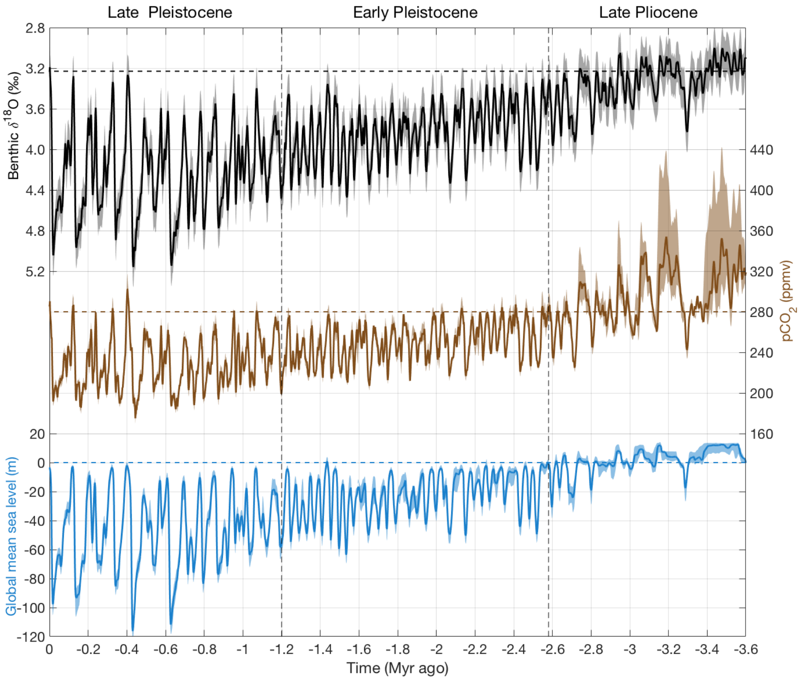


Figure 4: Observed benthic $\delta^{18}O$ (black; Lisiecki and Raymo, 2005) and reconstructed CO_2 (brown) and global mean sea level (blue) for the entire 3.6 Myr simulation period. The present-day values for the three variables (pre-industrial value of 280 ppmv for CO_2) are shown by horizontal dashed lines. Shaded areas indicates the uncertainty in the LR04 benthic $\delta^{18}O$ stack, and the resulting uncertainty in the reconstructed CO_2 and sea level.

The modelled sea-level is compared to several other reconstructions in Fig. 5. Shown are the reconstructions based on planktic $\delta^{18}O$ in the Red Sea (Grant et al., 2014) and in the Mediterranean $\delta^{18}O$ (Rohling et al., 2014), which cover the last 500 kyr and 5.5 Myr, respectively. The reconstruction by de Boer et al. (2014) is based on an ice-sheet model-based decomposition of the benthic $\delta^{18}O$ signal, very similar to the work presented here. The main difference is the climate forcing applied to the ice-sheet model, which by de Boer et al. (2014) was described by a simple globally uniform temperature offset, and which was replaced by the climate matrix approach plus inversely modelled CO_2 in our work. Despite this improvement in the climate forcing, which resulted in a simulated ice-sheet at LGM which agrees better with geomorphological evidence (Berends et al., 2018),

Deleted: respectively

the sea-level reconstructions are virtually indistinguishable. The reconstruction by Willeit et al. (2019) shown in Fig. 5 is based on a fully coupled ice-sheet – climate – carbon cycle model, forced only with insolation. However, their ice-sheet model only simulated the Northern Hemisphere; the contribution from Antarctica is assumed to be 10 % of that of the northern ice sheets. While this assumption seems to produce good results during the Pleistocene, its validity during the warmer-than-present late Pliocene is questionable. The reconstruction by Elderfield et al. (2012) was made using a Mg/Ca-based reconstruction of sea surface temperature, which was then used to disentangle the ice volume and deep-sea temperature signals in the benthic δ^{18}O record. Lastly, the reconstruction by Naish et al. (2009), adjusted by Miller et al. (2012) to match results from geological backstripping in New Zealand, suggests a relatively stable sea-level during the late Pliocene, 3.4 – 2.58 Myr ago, which agrees well with our results. Rohling et al. (2014), arguing that geological backstripping can provide information on relative changes in sea-level but not on absolute values, added a +20 m offset to the results by Miller et al. (2012), which made those results agree better with their own reconstruction from the Mediterranean Sea. However, our own results are reasonably close to those by Miller et al. (2011) over the period between 3.4 and 2.6 Myr ago without requiring such a correction. The strong variability in global mean sea-level visible in the results of Rohling et al. (2014) during this period, which suggests the repeated disappearance and reappearance of most of the East Antarctic ice sheet, is not visible in the data of Miller et al. (2012), nor in our own results.

Deleted: , so any possible contribution from Antarctica to changes in sea-level, especially during the warm Late Pliocene, is not accounted for in their reconstruction

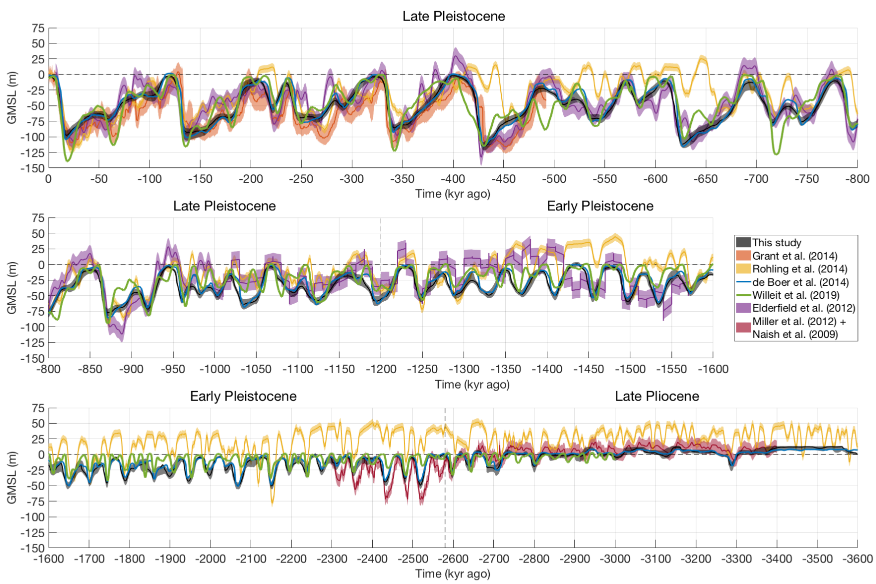


Figure 5: Reconstructed global mean sea-level for the entire simulation period (black), compared to reconstructions based on Red Sea $\delta^{18}O$ (Grant et al., 2014; red), Mediterranean $\delta^{18}O$ (Rohling et al., 2014; yellow), an ice-sheet model-based inversion of the global benthic $\delta^{18}O$ similar to this study (de Boer et al., 2014; blue), an insolation-forced, fully coupled ice-sheet – climate – carbon-cycle model (Willeit et al., 2019; green), a separation of the ice- and temperature-induced $\delta^{18}O$ signals based on Mg/Ca ratios (Elderfield et al., 2012; purple), and a direct scaling of benthic $\delta^{18}O$ scaled to match results from geological backstripping (Miller et al. 2012; Naish et al. 2009; dark red). The present-day value of zero is shown by a dashed line.

The CO_2 reconstruction for the past 800 kyr is compared to the EPICA Dome C ice core record (Bereiter et al., 2015) in Fig. 6, as well as to several different model reconstructions and proxy-based reconstructions. The three other model-based reconstructions shown were created by decoupling the benthic $\delta^{18}O$ signal using a 1-D ice-sheet model (van de Wal et al., 2011; Stap et al., 2017), or by using an insolation-forced, fully coupled ice-sheet – climate – carbon cycle model (Willeit et al., 2019). The geological proxies are based either on alkenones (Seki et al., 2010; Badger et al., 2013; Zhang et al., 2013) or $\delta^{11}B$ ratios (Hönisch et al, 2009; Seki et al., 2010; Bartoli et al., 2011; Martínez-Botí et al., 2015; Stap et al., 2016; Chalk et al., 2017; Dyez et al., 2018; Sosdian et al., 2018) derived from benthic foraminifera, or based on stomata (Kürschner et al., 1996; Stults et al., 2011; Bai et al., 2015; Hu et al., 2015).

Deleted: Lüthi et al., 2008

Deleted: 12

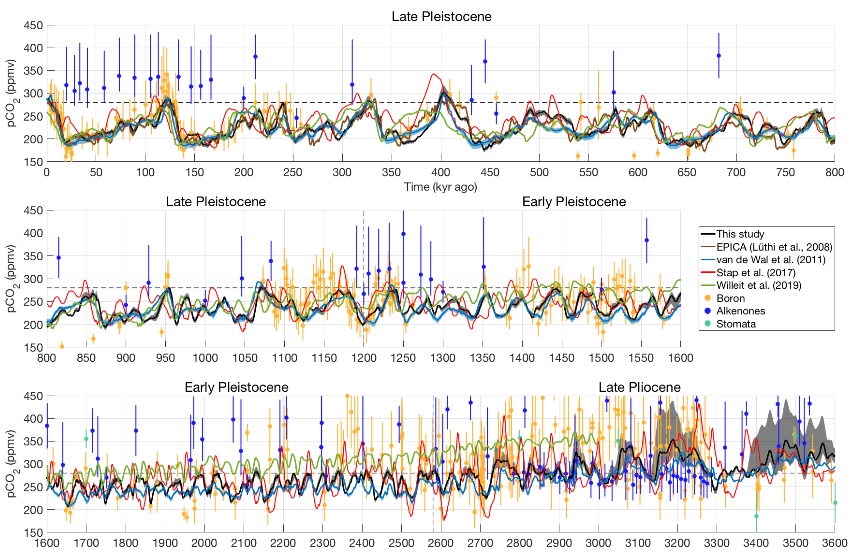


Figure 6: Reconstructed atmospheric CO_2 for the entire simulation period, compared to the EPICA Dome C ice core record (Bereiter et al., 2015), three different reconstructions based on ice-sheet-climate models (van de Wal et al., 2011; Stap et al., 2017; Willeit et al., 2019), as well as to chemical proxies based on ^{11}B ratios (Hönisch et al., 2009; Seki et al., 2010; Bartoli et al., 2011; Martínez-Botí et al., 2015; Stap et al., 2016; Chalk et al., 2017; Dyez et al., 2018; Sosdian et al., 2018), alkenones (Seki et al., 2010; Badger et al.,

Deleted: Lüthi et al., 2008

Deleted: Hönisch et al., 2009;

Deleted: Bartoli et al., 2011;

2013; Zhang et al., 2013) and stomata (Kürschner et al., 1996; Stults et al., 2011; Bai et al., 2015; Hu et al., 2015). The pre-industrial value of 280 ppmv is shown by a dashed line.

Our results broadly match those of van de Wal et al. (2011), who reconstructed CO_2 using 1-D ice models, forced with a simple spatially uniform temperature offset (assuming the strength of all feedbacks in the CO_2-temperature relation to be constant in time) calculated from benthic $\delta^{18}O$ using a similar inverse forward modelling approach. Both our results and those of van de Wal et al. (2011) show values during the late Pliocene that were higher than pre-industrial, but lower than present-day (412 ppmv at the date of writing), though the values by van de Wal et al. (2011) are at the low end of the uncertainty of our own results, as can be seen in Fig. 6. Stap et al. (2017) used a 1-D ice-sheet model set-up very similar to that by van de Wal et al. (2011), but used a more elaborate energy-balance model to represent the global climate. Their results differ markedly from ours, and those of van de Wal et al. (2011). The reconstruction by Willeit et al. (2019), who used a coupled ice-sheet – climate – carbon cycle model, agrees with ours and that by van de Wal et al. (2011) in terms of the glacial-interglacial amplitude both before and after the MPT. However, the downward linear trend in pCO_2, which Willeit et al. (2019) prescribed manually to induce the inception of northern hemisphere glaciation at 2.6 Myr ago, and the resulting high values during the Early Pleistocene, are not visible in the other reconstructions.

None of the model-based reconstructions agrees better than the others with the available proxy data. A recent study by Badger et al. (2019) showed that the alkenone proxy has only a very low sensitivity to atmospheric CO_2 at low to moderate CO_2 concentrations, as is clearly visible in the top panel of Fig. 6, where the alkenone proxy data show a nearly constant CO_2 concentration throughout the last glacial cycle. They state that the reliability of this proxy for values lower than ~350 ppmv (the entire Pleistocene and large parts of the Late Pliocene) is doubtful. In addition, several studies (Indermühle et al., 1999; Jordan, 2011; Porter et al., 2019) have questioned the reliability of the stomatal proxies. However, even when we disregard both of these proxies, the level of disagreement between the different boron isotope-based proxies is such that using them to choose one model-based reconstruction over another is not possible. The only available data record which is reliable enough to allow such a comparison is the ice-core record (Bereiter et al., 2015). In Table 1, we compare the performance of the different model-based studies at reproducing this CO2 record. All reconstructions have been prescribed an "optimised time lag" to find the highest possible correlation with the ice core record. Our results show the best agreement in terms of the coefficient of determination $R^2 = 0.74$, the root-mean-square-error RMSE = 13.5 ppmv (both taken over the entire 800 kyr ice-core period), and no time lag required to obtain these values. To put these numbers into perspective, the same numbers are listed for a simple least-squares linear fit between the LR04 benthic $\delta^{18}O$ stack and the EPICA Dome C record. This statistical "model" yields approximately the same results as our reconstruction in terms of correlation and RMSE, although the resulting post-MPT glacial-interglacial amplitude is about 23 ppmv too small, indicating that the increased glacial-interglacial variability is not properly captured.

Table 1: Statistical comparison of the different model-based CO_2 reconstructions (this study; van de Wal et al., 2011; Stap et al., 2017; Willeit et al., 2019) to the EPICA Dome C ice core record (Bereiter et al., 2015), as well as the glacial-interglacial amplitude in CO_2 both before and after the MPT for the different reconstructions.

Deleted: they

Deleted: linear correlation

Deleted: 68

Deleted: 0

Deleted: and of

Deleted: 15.3

Deleted: 14.7

Deleted:

Deleted: the

Deleted: of all four studies

Deleted: non-physical

Deleted: **Lüthi et al., 2008**

	EPICA	Linear regression	This study	Wal2011	Stap2017	Willeit2019
Optimised time lag (kyr)	-	-	0.0	2.0	10.0	9.0
Model-EPICA R^2	-	0.70	0.74	0.64	0.48	0.45
Model-EPICA RMSE (ppmv)	-	13.9	13.5	15.8	24.9	19.9
Pre-MPT amplitude (ppmv)	-	32	48	34	83	37
Post-MPT amplitude (ppmv)	91	68	87	79	104	72

Fig. 7 illustrates the difference between our modelled CO_2 values for the Early and Late Pleistocene. Interglacial CO_2 decreases from 280 – 300 ppmv at the beginning of the Early Pleistocene, to 250 – 280 ppmv just before the MPT. Glacial CO_2 decreases from 220 – 250 ppmv to 205 – 225 ppmv during this period, indicating a background trend of about -20 ppmv in 1.5 Myr. After the MPT, interglacial CO_2 varies between 250 – 320 ppmv, and glacial values decrease to 170 – 210 ppmv, with neither range showing a clear trend. The reduced glacial-interglacial CO_2 difference of 48 ppmv (averaged over the entire period) that we find in our results agrees with the value of 43 ppmv observed in the boron isotope data by Chalk et al. (2017). The values for all four model-based reconstructions, both before and after the MPT, are shown in Table 1. After the MPT, we find a glacial-interglacial difference of 87 ppmv, which agrees very well with the value of 91 ppmv over the past 800 kyr from the EPICA Dome C ice core record (Bereiter et al., 2015).

Deleted: Lüthi et al., 2008

The discussion about the nature of the Late Pleistocene glacial cycles is still ongoing. Different studies have proposed that they represent either a strongly amplified response to a weak 100 kyr forcing (Ganopolski and Calov, 2011), a response to the same 40 kyr forcing as in the Early Pleistocene, with an additional threshold mechanism that causes some insolation maxima to be "skipped", leading to 80/120 kyr cyclicity (Huybers, 2011; Tzedakis et al., 2017), or even as the result of some stochastic process with a collapse threshold, containing no true periodicity (Wunsch, 2003). Since the timing of glaciations and deglaciations in our model is forced by the LR04 benthic $\delta^{18}O$ stack, and thus its depth-age model, no meaningful conclusions regarding this discussion can be drawn from our results.

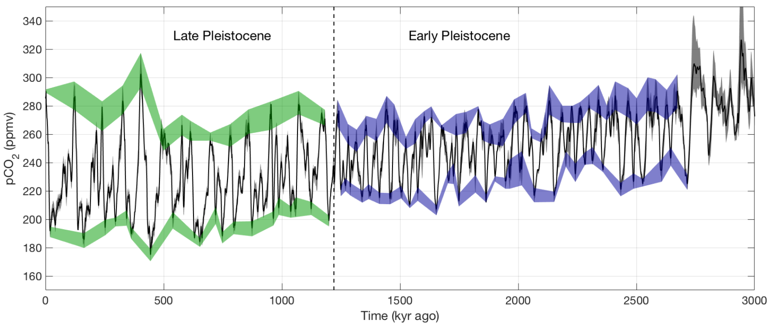


Figure 7: Reconstructed atmospheric CO_2 during the Pleistocene. The coloured regions indicate the glacial and interglacial ranges for the Early (2.6 – 1.2 Myr ago; blue) and Late (1.2 – 0.011 Myr ago; green) Pleistocene.

The non-linear, time-dependent relation between CO_2 and sea level is visualised in Fig. 8, both for the different proxy data (panel A) and the different model reconstructions (panels B-F). While the lag between changes in CO_2 and changes in global climate is not included in any of the model studies except Willeit et al. (2019), it is too short to be visible in the figure. Instead, the spread of the data points is due to both the slow response of the ice sheets to climate change, and the intrinsic hysteresis of ice sheet volume under a changing climate.

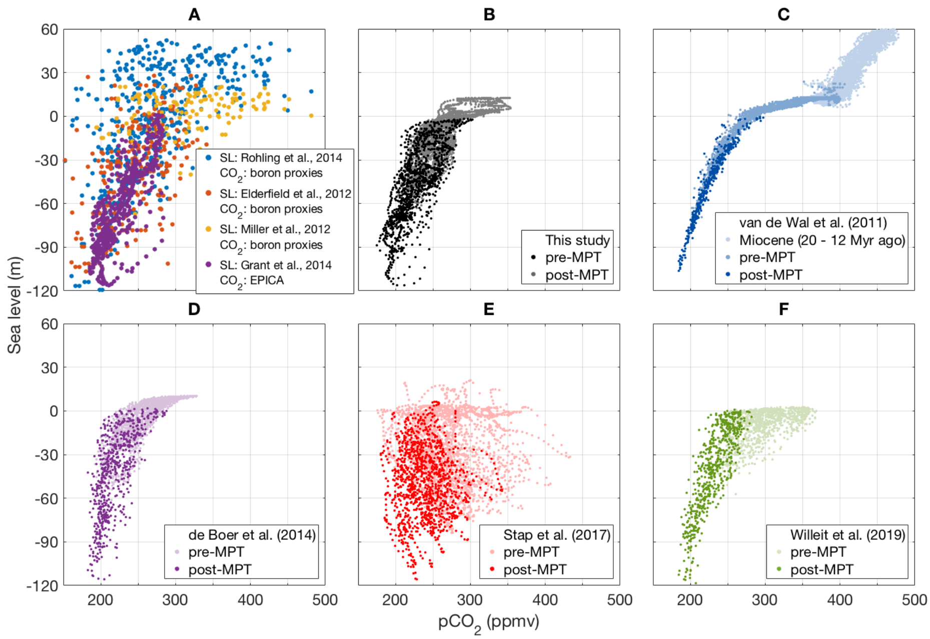


Figure 8: A: Sea level vs atmospheric CO_2 from proxy data. Blue: sea-level data from Rohling et al. (2014; last 5 Myr) vs boron isotope CO_2 data. Red: sea-level data from Elderfield et al. (2012; last 1.5 Myr) vs boron isotope CO_2 data. Gold: sea-level data from Miller et al. (2012) and Naish et al. (2009; 3.4 – 2.3 Myr ago) vs boron isotope CO_2 data. purple: sea-level data from Grant et al. (2014; last 500 kyr), vs the EPICA Dome C ice core record (Bereiter et al., 2015). B-F: reconstructed sea level and atmospheric CO_2 from different modelling studies (B: this study; C: van de Wal et al., 2011; D: de Boer et al., 2014; E: Stap et al., 2017; F: Willeit et al., 2019), separated into pre- and post-MPT values. de Boer et al. (2014; panel D) did not explicitly model CO_2, calculating only a global mean temperature change. We converted this to CO_2, using the logarithmic relation from van de Wal et al. (2011).

The combination of ice core CO_2 data (Bereiter et al., 2015) and Red Sea sea-level data (Grant et al., 2014), regarded as the two most reliable proxy reconstructions for the colder-than-present worlds of the Late Pleistocene and shown by the purple data points in Panel A, displays a pattern that is very similar to the model results by de Boer et al. (2014), and by this study. The reconstruction by van de Wal et al. (2011) shows a narrower range of sea levels for each CO_2 value, indicating that their ice sheets respond faster to changes in climate. The reconstruction by Willeit et al. (2019) also show a relatively narrow range of sea levels in the 180 – 250 ppmv CO_2 range, but a wider spread in the 250 – 350 ppmv range. The reconstruction by Stap et al. (2017) shows a vast spread, indicating that their ice sheets respond very slowly. The cause of this is not known (*personal communication with authors*). The results by van de Wal et al. (2011), de Boer et al. (2014), and this study all show a CO_2 "threshold" for glaciation and sea-level drop around 250 ppmv, slightly lower than what is indicated by the Grant/EPICA data.

Deleted: Lüthi et al., 2008

Deleted: 2

Deleted: Lüthi et al., 2008

Deleted: s

Deleted: and Willeit et al. (2019)

Deleted: smaller spread in data points

Deleted: Willeit et al. (2019)

Willeit et al. (2019) report a threshold for glaciation that depends on both insolation and CO_2. The models used by van de Wal et al. (2011), de Boer et al. (2014), and this study do not explicitly account for the effect of insolation on temperature, but only include an insolation term in the mass balance calculation (and, in this study, in the interpolation in the climate matrix), which might explain this difference.

The much weaker response of sea level to changes in CO_2 for concentrations higher than 280 ppmv, which is displayed by all models except Stap et al. (2017), seems to be confirmed by the sea-level data by Miller et al. (2012) and the boron isotope CO_2 data. The sea-level data by Rohling et al. (2014) show much more variation, especially during the early Pleistocene. The results by van de Wal et al. (2011) show a very clear threshold around 375 ppmv CO_2, beyond which sea level rises rapidly with increasing CO_2. As indicated by the colours, they only find these warm greenhouse worlds during the first half of the Miocene (20 – 12 Myr ago), which explains why none of the other models, nor the proxy data, display this behaviour. Although some of the boron isotope proxies suggest CO_2 concentrations higher than this 375 ppmv threshold value, the sea level data do not show the associated strong rise.

5 Conclusions and discussion

We have presented a new, time-continuous, self-consistent reconstruction of atmospheric CO_2, ice sheet evolution and global mean sea-level of the last 3.6 Myr, based on the benthic $\delta^{18}O$ glacial proxy. This reconstruction was created by using a hybrid ice-sheet – climate model (Berends et al., 2018) to decouple the contributions from ice volume and deep-water temperature to the benthic $\delta^{18}O$ record (Lisiecki and Raymo, 2005). Our sea-level reconstruction agrees well with similar model-based results (de Boer et al., 2014), the reconstruction based on Red Sea $\delta^{18}O$ by Grant et al. (2014) and the combined results of benthic $\delta^{18}O$ (Naish et al., 2009) and geological backstripping (Miller et al., 2012). Our results agree less well with the reconstruction based on Mediterranean Sea $\delta^{18}O$ by Rohling et al. (2014), which goes back further in time but has a lower signal-to-noise ratio, and with the reconstruction based on benthic $\delta^{18}O$ and Mg/Ca ratios by Elderfield et al. (2012). Although previous studies (van de Wal et al., 2011; de Boer et al., 2014; Stap et al., 2017) used more simple, parameterized ice-sheet and climate models, the resulting sea-level reconstructions are very similar.

Our CO_2 reconstruction agrees well with the EPICA Dome C ice core record (Bereiter et al., 2015), showing the strongest correlation of all four model-based reconstructions. Our results do not agree well with different chemical proxy-based reconstructions based on foraminiferal alkenones (Seki et al., 2010; Badger et al., 2013; Zhang et al., 2013), foraminiferal $\delta^{11}B$ ratios (Seki et al., 2010; Martínez-Botí et al., 2015; Stap et al., 2016; Chalk et al., 2017; Dyez et al., 2018; Sosdian et al., 2018) or fossil plant stomata (Kürschner et al., 1996; Stults et al., 2011; Bai et al., 201; Hu et al., 2015). However, the strong spread between those different proxies and the large uncertainty of each individual proxy record do not allow to conclude whether the difference between those proxies, and our model reconstructed values is significant.

Deleted: Lüthi et al., 2008

Deleted: 12

Between 2.8 and 1.2 Myr ago, during the early Pleistocene, we simulate 40 kyr glacial cycles with glacial-interglacial sea-level changes of 25 – 50 m. During this period, CO_2 varies between 270 – 280 ppmv during interglacials and 210 – 240 ppmv during glacial maxima, with both values decreasing by about 20 ppmv over the 1.4 Myr course of the Early Pleistocene. After the Mid-Pleistocene Transition (MPT), when the glacial cycles change from 40 kyr to 80/120 kyr cyclicity, their sea-level amplitude increases to 70 – 120 m, with CO_2 varying between 250 – 320 ppmv during interglacials and 170 – 210 ppmv during glacial maxima. This implies a glacial-interglacial contrast of about 45 ppmv pre MPT and 85 ppmv post MPT. The two most reliable proxy records, the sea-level reconstruction by Grant et al. (2014) and the EPICA Dome C ice core record (Bereiter et al., 2015), display a relation between sea-level and CO_2 for colder-than-present worlds that matches our model results. For warmer-than-present worlds, the large spread and uncertainty in available proxy data for both sea-level and CO_2, as well as the much larger uncertainty in our model results during these warm periods, prevent us from drawing conclusions on the reliability of our methodology in this regard. We believe that the inverse modelling method of using benthic $\delta^{18}O$, combined with coupled or hybrid ice-sheet – climate models, as a proxy for past CO_2 and sea-level, can add valuable insights into the evolution of the Earth system during these warm episodes.

Deleted: Lüthi et al., 2008

Deleted: model results

During the warm late Pliocene, we find a large variability in CO_2, with a difference of about 100 ppmv between the cold glacial excursion 3.3 Myr ago, and the warm peak around 3.2 Myr ago. The boron isotope-based data by Martínez-Botí et al. (2015), which has the highest temporal resolution and longest temporal range of all proxy records, shows a variability of about 150 ppmv during this period (albeit with an uncertainty of about 100 ppmv in either direction). While the large discrepancies between different proxy records are too large to draw any definitive conclusions, their findings do seem to support strong CO_2 variability during these warm periods. Our results for this period also show a much larger uncertainty range than during the Pleistocene. We suspect that the relative sparsity of our current matrix in this warm regime biases the model towards large ice sheets in warm worlds. This could explain the high variability in our reconstructed CO_2 during the late Pliocene; if little or no ice-sheet retreat occurs, then the observed variability in the $\delta^{18}O$ record must be attributed entirely to changes in deep-water temperature, which requires large changes in surface climate, and therefore in CO_2. By adding additional GCM snapshots for worlds with > 400 ppmv CO_2 and smaller-than-present ice sheets (particularly a reduced East Antarctic ice sheet), ice-sheet retreat in warm climates might be increased, reducing simulated CO_2 variability during these periods. While this means that the modelled sea-level rise during this period will be larger, our current results are at the low end of the uncertainties of other reconstructions, implying that a higher modelled value will not immediately lead to a mismatch.

The parameterised relation between the global climate, deep-sea temperature and the resulting contribution to the $\delta^{18}O$ signal could be improved upon. That contribution is currently derived from the modelled high-latitude annual mean surface temperature anomaly, using a simple 3,000 yr moving average and a constant scaling factor. This constitutes a strong simplification of the complex relation between atmospheric and oceanic temperatures. Based on output from available ocean-enabled GCM simulations of the LGM, a more elaborate parameterisation could be constructed, perhaps using surface temperatures over specific downwelling regions. Instead of using a global benthic $\delta^{18}O$ stack, it might even possible to separate the records from different ocean basins.

Our results should not be interpreted as a realistic reconstruction of what the world looked like in terms of global climate, ice sheet geometry, sea level and CO_2, during these periods of geological history. Rather, we believe they should be viewed as scenarios, which can help to interpret an expected new ice core record. For example, one hypothesised mechanism behind the MPT is regolith erosion, which led to a change in basal conditions from mostly sliding ice to mostly non-sliding ice over North America and Eurasia during the MPT (Clark and Pollard, 1998; Willeit et al., 2019. Basal conditions in our model are constant over time, so that the MPT is entirely attributed to changes in CO_2. A simulation identical to ours, but including a change in basal conditions, would result in a different pre-MPT CO_2 history. Comparing the new ice core record to these two different scenarios could help determine if regolith erosion was a crucial factor in the transition. Other proposed physical mechanisms, such as regolith erosion leading to a different relation between ice sheet geometry and glaciogenic dust, changes in ocean circulation leading to a different dynamic equilibrium between atmospheric and oceanic carbon, or a slow background decrease in CO_2 caused by weathering or reduced volcanism, could be explored in a similar manner. We believe our modelling approach is very suited for this kind of exploratory analysis.

Data availability. The reconstructed global mean sea-level and atmospheric CO_2 concentration over the past 3.6 Myr are available online at doi.org/10.5281/zenodo.3793592

Appendix A

Here, we briefly summarise the equations governing the matrix method. A more thorough explanation of this approach can be found in the original publication by Berends et al. (2018).

Each GCM snapshot contains data fields describing global monthly mean 2-m air temperature, global monthly mean precipitation, and surface elevation. The interpolation routine for the temperature field uses modelled CO_2 and modelled "absorbed insolation" as weighting parameters to interpolate between snapshots. This approach captures the two most important physical processes through which a change in ice-sheet geometry affects the local and global temperature; the ice-albedo feedback, and the altitude-temperature feedback. Since the absorbed insolation changes not only through changes in albedo, but also through changes in incoming insolation, the effects of orbital forcing are also (indirectly) accounted for. The weighting factor w_{CO2} is calculated as:

$$w_{CO2} = \frac{pCO_2 - pCO_{2,LGM}}{pCO_{2,PI} - pCO_{2,LGM}}, \qquad \text{(A1)}$$

with $pCO_{2,PI}$ = 280 ppmv and $pCO_{2,LGM}$ = 190 ppmv. Multiplying the modelled surface albedo α with the prescribed insolation at the top of the atmosphere Q_{TOA} yields the absorbed insolation I_{abs}:

$$I_{abs}(x,y) = \left(1-\alpha(x,y)\right)\cdot Q_{TOA}(x,y). \tag{A2}$$

This value is scaled between the PI and LGM reference fields to obtain the weighting parameter w_{ins}:

$$w_{ins}(x,y) = \frac{I_{abs,mod}(x,y) - I_{abs,LGM}(x,y)}{I_{abs,PI}(x,y) - I_{abs,LGM}(x,y)}. \tag{A3}$$

This weighting field is partly smoothed with a 200 km Gaussian filter F, to account for both local and regional effects:

$$w_{ice,T}(x,y) = \tfrac{1}{7}w_{ins}(x,y) + \tfrac{3}{7}F\left(w_{ins}(x,y)\right) + \tfrac{3}{7}\overline{w_{ins}}. \tag{A4}$$

This is combined with the scalar pCO_2 weight w_{CO2} to yield the final temperature weighting parameter w_T:

$$w_T(x,y) = \frac{w_{CO2} + w_{ice,T}(x,y)}{2}. \tag{A5}$$

This weighting field is used to interpolate both the temperature and orography fields of the GCM snapshots:

$$T_{ref}(x,y) = w_T \cdot T_{PI}(x,y) + (1-w_T)\cdot T_{LGM}(x,y), \tag{A6}$$
$$h_{ref}(x,y) = w_T \cdot h_{PI}(x,y) + (1-w_T)\cdot h_{LGM}(x,y). \tag{A7}$$

Orography is interpolated as well, so that it can be used to downscale the temperature field from the low-resolution GCM grid to the higher-resolution ice-model grid, using a simple lapse-rate approach:

$$T(x,y) = T_{ref}(x,y) + \lambda_{LGM}(x,y)\left(h(x,y) - h_{ref}(x,y)\right). \tag{A8}$$

The spatially variable lapse rate $\lambda_{LGM}(x,y)$ is derived from the temperature field of the GCM snapshot.

The interpolation for precipitation is based on ice thickness rather than absorbed albedo. This reflects the fact that precipitation over the ice sheet is strongly affected by altitude:

$$w_{ice,P}(x,y) = \frac{Hi_{mod}(x,y) - Hi_{PI}(x,y)}{Hi_{LGM}(x,y) - Hi_{PI}(x,y)} \cdot \frac{V_{mod} - V_{PI}}{V_{LGM} - V_{PI}}, \quad \text{(A9)}$$

$$w_P(x,y) = \frac{w_{CO2} + w_{ice,P}(x,y)}{2}. \quad \text{(A10)}$$

$$P_{ref}(x,y) = e^{\left(w_P \cdot log\left(P_{PI}(x,y)\right) + (1-w_P) \cdot log\left(P_{LGM}(x,y)\right)\right)}, \quad \text{(A11)}$$

Finally, the precipitation field is downscaled from the low-resolution GCM grid to the higher-resolution ice-model grid, using the precipitation model by Roe and Lindzen (2001) and Roe (2002), to correct for the steeper slopes in the ice model:

$$P(x,y) = P_{ref}(x,y) \frac{P_{Roe}(x,y)}{P_{Roe_{ref}}(x,y)}. \quad \text{(A12)}$$

Author contributions. CJB, BdB, and RSWvdW designed the study. CJB created the model set-up and carried out the simulations, with support from RSWvdW. CJB drafted the paper, and all authors contributed to the final version.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The Ministry of Education, Culture and Science (OCW), in the Netherlands, provided financial support for this study via the program of the Netherlands Earth System Science Centre (NESSC). B. de Boer was funded by NWO Earth and Life Sciences (ALW), project 863.15.019. This work was sponsored by NWO Exact and Natural Sciences for the use of supercomputer facilities. Model runs were performed on the LISA Computer Cluster, we would like to acknowledge SurfSARA Computing and Networking Services for their support. Special thanks go to B. Hönisch for advice on paleo-CO_2 data usage.

References

Alley, R. B.: The Younger Dryas cold interval as viewed from central Greenland, Quaternary Science Reviews 19, 213-226, 2000

Badger, M. P. S., Schmidt, D. N., Mackensen, A., and Pancost, R. D.: High-resolution alkenone palaeobarometry indicates relatively stable pCO2 during the Pliocene (3.3-2.8 Ma), Philosophical Transactions of the Royal Society A 371, 2013

Badger, M. P. S., Chalk, T. B., Foster, G. L., Brown, P. R., Gibbs, S. J., Sexton, P. F., Schmidt, D. N., Pälike, H., Mackensen, A., and Pancost, R. D.: Insensitivity of alkenone carbon isotopes to atmospheric CO2 at low to moderate CO2 levels, Climate of the Past 15, 539-554, 2019

Bai, Y.-J., Chen, L.-Q., Ranhotra, P. S., Wang, Q., Wang, Y.-F., and Li, C.-S.: Reconstructing atmospheric CO2 during the Plio-Pleistocene transition by fossil Typha, Global Change Biology 21, 874-881, 2015

Bartoli, G., Hönisch, B., and Zeebe, R. E.: Atmospheric CO2 decline during the Pliocene intensification of Northern Hemisphere glaciations, Paleoceanography 26, 2011

Beerling, D. J. and Royer, D. L.: Convergent Cenozoic CO2 history, Nature Geoscience 4, 418-420, 2011

Bereiter, B., Eggleston, S., Schmitt, J., Nehrbass-Ahles, C., Stocker, T. F., Fischer, H., Kipfstuhl, S., and Chappellaz, J.: Revision of the EPICA Dome C CO2 record from 800 to 600 kyr before present, Geophysical Research Letters 42, 542-549, 2015

Berends, C. J., de Boer, B., and van de Wal, R. S. W.: Application of HadCM3@Bristolv1.0 simulations of paleoclimate as forcing for an ice-sheet model, ANICE2.1: set-up and benchmark experiments, Geoscientific Model Development 11, 4657-4675, 2018

Berends, C. J., de Boer, B., Dolan, A. M., Hill, D. J., and van de Wal, R. S. W.: Modelling ice sheet evolution and atmospheric CO2 during the Late Pliocene, Climate of the Past 15, 1603-1619, 2019

Bintanja, R., van de Wal, R. S. W., and Oerlemans, J.: Global ice volume variations through the last glacial cycle simulated by a 3-D ice-dynamical model, Quaternary International 95-96, 11-23, 2002

Bintanja, R., van de Wal, R., and Oerlemans, J.: Modelled atmospheric temperatures and global sea levels over the past million years, Nature 437, 125-128, 2005

Bintanja, R. and van de Wal, R. S. W.: North American ice-sheet dynamics and the onset of 100,000-year glacial cycles, Nature 454, 869-872, 2008

Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J.-Y., Abe-Ouchi, A., Crucifix, M., Driesschaert, E., Fichefet, T., Hewitt, C. D., Kageyama, M., Kitoh, A., Laîné, A., Loutre, M.-F., Marti, O., Merkel, U., Ramstein, G., Valdes, P. J., Weber, S. L., Yu, Y., and Zhao, Y.: Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum - Part 1: experiments and large-scale features, Climate of the Past 3, 261-277, 2007

Dansgaard, W.: Stable isotopes in precipitation, Tellus 16, 436-468, 1964

de Boer, B., van de Wal, R. S. W., Bintanja, R., Lourens, L. J., and Tuenter, E.: Cenozoic global ice-volume and temperature simulations with 1-D ice-sheet models forced by benthic d18O records, Annals of Glaciology 51, 23-33, 2010

de Boer, B., van de Wal, R. S. W., Lourens, L. J., and Bintanja, R.: Transient nature of the Earth's climate and the implications for the interpretation of benthic d18O records, Palaeogeography, Palaeoclimatology, Palaeoecology 335-336, 4-11, 2012

de Boer, B., van de Wal, R., Lourens, L. J., Bintanja, R., and Reerink, T. J.: A continuous simulation of global ice volume over the past 1 million years with 3-D ice-sheet models, Climate Dynamics 41, 1365-1384, 2013 de Boer, B., Lourens, L. J., and van de Wal, R. S. W.: Persistent 400,000-year variability of Antarctic ice volume and the carbon cycle is revealed throughout the Plio-Pleistocene, Nature Communications 5, 2014

de Boer, B., Dolan, A. M., Bernales, J., Gasson, E., Goelzer, H., Golledge, N. R., Sutter, J., Huybrechts, P., Lohmann, G., Rogozhina, I., Abe-Ouchi, A., Saito, F., and van de Wal, R. S. W.: Simulating the Antarctic ice sheet in the Late-Pliocene warm period: PLISMIP-ANT, an ice-sheet model intercomparison project, The Cryosphere 9, 881-903, 2015

de Boer, B., Haywood, A. M., Dolan, A. M., Hunter, S. J., and Prescott, C. L.: The Transient Response of Ice Volume to Orbital Forcing During the Warm Late Pliocene, Geophysical Research Letters 44, 10,486-10,494, 2017

Brovkin, V., Ganopolski, A., Archer, D., and Munhoven, G.: Glacial CO2 cycle as a succession of key physical and biogeochemical processes, Climate of the Past 8, 251-264, 2012

Bueler, E. and Brown, J.: Shallow shelf approximation as a "sliding law" in a thermomechanically coupled ice sheet model, Journal of Geophysical Research 114, 2009

Chalk, T. B., Hain, M. P., Foster, G. L., Rohling, E. J., Sexton, P. F., Badger, M. P. S., Cherry, S. G., Hasenfratz, A. P., Haug, G. H., Jaccard, S. L., Martínez-García, A., Pälike, H., Pancost, R. D., and Wilson, P. A.: Causes of ice age intensification across the Mid-Pleistocene Transition, Proceedings of the National Academy of Sciences 114, 13114-13119, 2017

Clark, P. U. and Pollard, D.: Origin of the middle Pleistocene transition by ice sheet erosion of regolith, Paleoceanography 13, 1-9, 1998

Dolan, A. M., Haywood, A. M., Hill, D. J., Dowsett, H. J., Hunter, S. J., Lunt, D. J., and Pickering, S. J.: Sensitivity of Pliocene ice sheets to orbital forcing, Palaeogeography, Palaeoclimatology, Palaeoecology 309, 98-110, 2011

Dolan, A. M., Haywood, A. M., Hunter, S. J., Tindall, J. C., Dowsett, H. J., Hill, D. J., and Pickering, S. J.: Modelling the enigmatic Late Pliocene Glacial Event - Marine Isotope Stage M2, Global and Planetary Change 128, 47-60, 2015

Dyez, K. A., Hönisch, B., and Schmidt, G. A.: Early Pleistocene Obliquity-Scale pCO2 Variability at ~1.5 Milion Years Ago, Paleoceanography and Paleoclimatology 33, 1270-1291, 2018

Elderfield, H., Ferretti, P., Greaves, M., Crowhurst, S., McCave, I. N., Hodell, D. A., and Piotrowski, A. M.: Evolution of Ocean Temperature and Ice Volume Through the Mid-Pleistocene Climate Transition, Science 337, 704-711, 2012

Finsinger, W. and Wagner-Cremer, F.: Stomatal-based inference models for reconstruction of atmospheric CO2 concentration: a method assessment using a calibration and validation approach, The Holocene 18, 757-764, 2009

Foster, G. L. and Rae, J. W. B.: Reconstructing Ocean pH with Boron Isotopes in Foraminifera, Annu. Rev. Earth Planet. Sci. 44, 207-237, 2016

Franks, P. J., Royer, D. L., Beerling, D. J., Van de Water, P. K., Cantrill, D. J., Barbour, M. M., and Berry, J. A.: New constraints on atmospheric CO2 concentration for the Phanerozoic, Geophysical Research Letters 41, 4685-4694, 2014

Ganopolski, A. and Calov, R.: The role of orbital forcing, carbon dioxide and regolith in 100 kyr glacial cycles, Climate of the Past 7, 1415-1425, 2011

Deleted: de Boer, B., Stocchi, P., and van de Wal, R.: A fully coupled 3-D ice-sheet-sea-level model: algorithm and applications, Geoscientific Model Development 7, 2141-2156, 2014

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B., and Wood, R. A.: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, Climate Dynamics 16, 147-168, 2000

Grant, K. M., Rohling, E. J., Bronk Ramsey, C., Cheng, H., Edwards, R. L., Florindo, F., Heslop, D., Marra, F., Roberts, A. P., Tamisiea, M. E., and Williams, F.: Sea-level variability over five glacial cycles, Nature Communications 5, 2014

Haywood, A. M., Hill, D. J., Dolan, A. M., Otto-Bliesner, B., Bragg, F., Chan, W.-L., Chandler, M. A., Contoux, C., Dowsett, H. J., Jost, A., Kamae, Y., Lohmann, G., Lunt, D. J., Abe-Ouchi, A., Pickering, S. J., Ramstein, G., Rosenbloom, N. A., Salzmann, U., Sohl, L., Stepanek, C., Ueda, H., Yan, Q., and Zhang, Z.: Large-scale features of Pliocene climate: results from the Pliocene Model Intercomparison Project, Climate of the Past 9, 191-209, 2013

Haywood, A. M. and Valdes, P. J.: Modelling Pliocene warmth: contribution of atmosphere, oceans and cryosphere, Earth and Planetary Science Letters 218, 363-377, 2003

Hönisch, B., Hemming, N. G., Archer, D., Siddall, M., and McManus, J. F.: Atmospheric carbon dioxide concentration across the mid-Pleistocene transition, Science 324, , 2009

Hu, J.-J., Xing, Y.-W., Turkington, R., Jacques, F. M. B., Su, T., Huang, Y.-J., and Zhou, Z.-K.: A new positive relationship between pCO2 and stomatal frequency in Quercus guyavifolia (Fagaceae): a potential proxy for palaeo-CO2 levels, Annals of Botany 115, 777-788, 2015

Huybers, P.: Combined obliquity and precession pacing of late Pleistocene glaciations, Nature 480, 229-232, 2011

Indermühle, A., Stauffer, B., Stocker, T. F., Raynaud, D., and Barnola, J.-M.: Technical Comment: Early Holocene Atmospheric CO2 Concentrations, Science 286, 1999

Jordan, G. J.: A critical framework for the assessment of biological palaeoproxies: predicting past climate and levels of atmospheric CO2 from fossil leaves, New Phytologist 192, 29-44, 2011

Jouzel, J., Alley, R. B., Cuffey, K. M., Dansgaard, W., Grootes, P., Hoffmann, G., Johnsen, S. J., Koster, R. D., Peel, D., Shuman, C. A., Stievenard, M., Stuiver, M., and White, J.: Validity of the temperature reconstruction from water isotopes in ice cores, Journal of Geophysical Research 102, 26,471-26,487, 1997

Jouzel, J., Masson-Delmote, V., Cattani, O., Dreyfus, G., Falourd, S., Hoffmann, G., Minster, B., Nouet, J., Barnola, J. M., Chappellaz, J., Fischer, H., Gallet, J. C., Johnsen, S., Leuenberger, M., Loulergue, L., Luethi, D., Oerter, H., Parrenin, F., Raisbeck, G., Raynaud, D., Schilt, A., Schwander, J., Selmo, E., Souchez, R., Spahni, R., Stauffer, B., Steffensen, J. P., Stenni, B., Stocker, T. F., Tison, J. L., Werner, M., and Wolff, E. W.: Orbital and Millenial Antarctic Climate Variability over the Past 800,000 Years, Science 317, 793-797, 2007

Kürschner, W. M., van der Burgh, J., Visscher, H., and Dilcher, D. L.: Oak leaves as biosensors of late Neogene and early Pleistocene paleoatmospheric CO2 concentrations, Marine Micropaleontology 27, 299-312, 1996

Kindler, P., Guillevic, M., Baumgartner, M., Schwander, J., Landais, A., and Leuenberger, M.: Temperature reconstruction from 10 to 120 kyr b2k from the NGRIP ice core, Climate of the Past 10, 887-902, 2014

Laskar, J., Robutel, P., Gastineau, M., Correia, A. C. M., and Levrard, B.: A long-term numerical solution for the insolation quantities of the Earth, Astronomy & Astrophysics 428, 261-285, 2004

Lisiecki, L. E. and Raymo, M. E.: A Pliocene-Pleistocene stack of 57 globally distributed benthic delta-18-O records, Paleoceanography 20, 2005

Martin, M. A., Winkelmann, R., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C., and Levermann, A.: The Potsdam Parallel Ice Sheet Model (PISM-PIK) - Part 2: Dynamic equilibrium simulation of the Antarctic ice sheet, The Cryosphere 5, 727-740, 2011

Martínez-Botí, M. A., Foster, G. L., Chalk, T. B., Rohling, E. J., Sexton, P. F., Lunt, D. J., Pancost, R. D., Badger, M. P. S., and Schmidt, D. N.: Plio-Pleistocene climate sensitivity evaluated using high-resolution CO2 records, Nature 518, 49-54, 2015

Miller, K. G., Wright, J. D., Browning, J. V., Kulpecz, A., Kominz, M., Naish, T., Cramer, B. S., Rosenthal, Y., Peltier, W. R., and Sosdian, S.: High tide of the warm Pliocene: Implications of global sea level for Antarctic deglaciation, Geology 40, 407-410, 2012

Morland, L. W. and Johnson, I. R.: Steady motion of ice sheets, Journal of Glaciology 25, 229-246, 1980

Morland, L. W.: Unconfined ice-shelf flow, in Dynamics of the West Antarctic Ice Sheet, edited by C. J. van der Veen, and J. Oerlemans, pp. 99–116, Kluwer Acad., Dordrecht, Netherlands, 1987

Naish, T., Powell, R., Levy, R., Wilson, G., Scherer, R., Talarico, F., Krissek, L., Niessen, F., Pompilio, M., Wilson, T., Carter, L., DeConto, R. M., Huybers, P., McKay, R. M., Pollard, D., Ross, J., Winter, D., Barrett, P., Browne, G., Cody, R., Cowan, E., Crampton, J., Dunbar, G., Dunbar, N., Florindo, F., Gebhardt, C., Graham, I., Hannah, M., Hansaraj, D., Harwood, D., Helling, D., Henrys, S., Hinnov, L., Kuhn, G., Kyle, P., Laufer, A., Maffioli, P., Magens, D., Mandernack, K., McIntosh, W., Millan, C., Morin, R., Ohneiser, C., Paulsen, T., Persico, D., Raine, I., Reed, J., Riesselman, C. R., Sagnotti, L., Schmitt, D., Sjunneskog, C., Strong, P., Taviani, M., Vogel, S., Wilch, T., and Williams, T.: Obliquity-paced Pliocene West Antarctic ice sheet oscillations, Nature 458, 322-328, 2009

Pollard, D.: A retrospective look at coupled ice sheet-climate modeiing, Climatic Change 100, 173-194, 2010

Pollard, D. and DeConto, R. M.: Modelling West Antarctic ice sheet growth and collapse through the past five million years, Nature 458, 329-332, 2009

Porter, A. S., Evans-FitsGerald, C., Yiotis, C., Montañez, I. P., and McElwain, J. C.: Testing the accuracy of new paleoatmospheric CO2 proxies based on plant stable isotopic composition and stomatal traits in a range of simulated paleoatmospheric O2:CO2 ratios, Geochimica et Cosmochimica Acta 259, 69-90, 2019

Rohling, E. J., Foster, G. L., Marino, G., Roberts, A. P., Tamisiea, M. E., and Williams, F.: Sea-level and deep-sea-temperature variability over the past 5.3 million years, Nature 508, 477-492, 2014

Seki, O., Foster, G. L., Schmidt, D. N., Mackensen, A., Kawamure, K., and Pancost, R. D.: Alkenone and boron-based Pliocene pCO2 records, Earth and Planetary Science Letters 292, 201-211, 2010

Shakun, J. D., Lea, D. W., Lisiecki, L. E., and Raymo, M. E.: An 800-kyr record of global surface ocean d18O and implications for ice volume-temperature coupling, Earth and Planetary Science Letters 426, 58-68, 2015

Deleted: ¶
Lüthi, D., Le Floch, M., Bereiter, B., Blunier, T., Barnola, J.-M., Siegenthaler, U., Raynaud, D., Jouzel, J., Fisher, H., Kawamura, K., and Stocker, T. F.: High-resolution carbon dioxide concentration record 650,000-800,000 years before present, Nature Letters 452, 379-382, 2008

Singarayer, J. S. and Valdes, P. J.: High-latitude climate sensitivity to ice-sheet forcing over the last 120 kyr, Quaternary Science Reviews 29, 43-55, 2010

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., and Averyt, K. B.: Climate Change 2007: the Physical Science Basis, Contribution Of Working Group 1 To The Fourth Assessment Report Of The Intergovernmental Panel On Climate Change, 2007

Sosdian, S. and Rosenthal, Y.: Deep-sea temperature and ice volume cange across the Pliocene-Pleistocene climate transitions, Science 325, 306-310, 2009

Sosdian, S. M., Greenop, R., Hain, M. P., Foster, G. L., Pearson, P. N., and Lear, C. H.: Constraining the evolution of Neogene ocean carbonate chemistry using the boron isotope pH proxy, Earth and Planetary Science Letters 498, 362-376, 2018

Stap, L. B., de Boer, B., Ziegler, M., Bintanja, R., Lourens, L. J., and van de Wal, R.: CO2 over the past 5 million years: Continuous simulation and new d11B-based proxy data, Earth and Planetary Science Letters 439, 1-10, 2016

Stap, L. B., van de Wal, R. S. W., de Boer, B., Bintanja, R., and Lourens, L. J.: The influence of ice sheets on temperature during the past 38 million years inferred from a one-dimensional ice sheet-climate model, Climate of the Past 13, 1243-1257, 2017

Stults, D. Z., Wagner-Cremer, F., and Axsmith, B. J.: Atmospheric paleo-CO2 estimates based on Taxodium distichum (Cupressaceae) fossils from the Miocene and Pliocene of Eastern North America, Palaeogeography, Palaeoclimatology, Palaeoecology 309, 327-332, 2011

Tzedakis, P. C., Crucifix, M., Mitsui, T., and Wolff, E. W.: A simple rule to determine which insolation cycles lead to interglacials, Nature 542, 427-444, 2017

Valdes, P. J., Armstrong, E., Badger, M. P. S., Bradshaw, C. D., Bragg, F., Crucifix, M., Davies-Barnard, T., Day, J. J., Farnsworth, A., Gordon, C., Hopcroft, P. O., Kennedy, A. T., Lord, N. S., Lunt, D. J., Marzocchi, A., Parry, L. M., Pope, V., Roberts, W. H. G., Stone, E. J., Tourte, G. J. L., and Williams, J. H. T.: The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0, Geoscientific Model Development 10, 3715-3743, 2017

Wagner, F., Aaby, B., and Visscher, H.: Rapid atmospheric CO2 changes associated with the 8,200-years-B.P. cooling event, Proceedings of the National Academy of Sciences 99, 12011-12014, 2002

van de Wal, R. S. W., de Boer, B., Lourens, L. J., Köhler, P., and Bintanja, R.: Reconstruction of a continuous high-resolution CO2 record over the past 20 million years, Climate of the Past 7, 1459-1469, 2011

Willeit, M., Ganopolski, A., Calov, R., Robinson, A., and Maslin, M.: The role of CO2 decline for the onset of Northern Hemisphere glaciation, Quaternary Science Reviews 119, 22-34, 2015

Willeit, M., Ganopolski, A., Calov, R., and Brovkin, V.: Mid-Pleistocene transition in glacial cycles explained by declining CO and regolith removal, Science Advances 5, 2019

Winkelmann, R., Martin, M. A., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C., and Levermann, A.: The Potsdam Parallel Ice Sheet Model (PISM-PIK) - Part 1: Model description, The Cryosphere 5, 715-726, 2011

Wunsch, C.: The spectral description of climate change including the 100 ky energy, Climate Dynamics 20, 353-363, 2003

Zachos, J. C., Pagani, M., Sloan, L., Thomas, E., and Billups, K.: Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present, Science 292, 686-694, 2001

Zachos, J. C., Dickens, G. R., and Zeebe, R. E.: An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics, Nature 451, 279-283, 2008

Zhang, Y. G., Pagani, M., Liu, Z., Bohaty, S. M., and DeConto, R. M.: A 40-million year history of atmospheric CO2, Philosophical Transactions of the Royal Society A 371, 2013

Zwally, H. J. and Giovinetto, M. B.: Areal distribution of the oxygen-isotope ratio in Greenland, Annals of Glaciology 25, 208-213, 1997