Increased abyssal ocean density stratification across the Middle Pleistocene Transition

Nicola C. Thomas[1], Heather L. Ford[2], Mervyn Greaves[1], and David A. Hodell[1]

In the following, Reviewer's comments are written in black text; our previously submitted responses are given in blue; and the actual changes made to the manuscript are provided in red text. We provide the line, section or figure numbers where those changes have been instituted in the revised manuscript.

Point-by-point reply to the Editor's comments:

In addition to implementing the changes described in your responses, please also clarify how weights for the weighted stack were applied where data were not available for all three regions. The current explanation (lines 250-252) is difficult to understand. What are the impacts of large data gaps in the South Atlantic?

For weighted stacks, basin weights were defined using fixed deep ocean volumes following the approach of Lisiecki and Stern (2016) (see their Table S2) and renormalised to unity at each time step to reflect the ocean volume represented by the available records. This ensures that stacked estimates reflect the relative deep ocean volume of the contributing basins without introducing bias when one or more basins are absent. Ocean basin boundaries and volumes are presented in Table S1.
Lines 312–315 and Equations 5–7 on Lines 322, 327 and 335.

We have also added a new supplementary figure (Fig. S10) showing the impact of data gaps on the global stack when using the Site 1094 record, which has the greatest number of gaps. The results suggest the influence on both weighted MDOT and $\delta^{18}O_{seawater}$ reconstructions is minimal.
Stated in Line 317.

Also, to improve figure clarity: Please adjust the colors in panel b of Figure 10; the two shades of blue are too difficult to distinguish.

The colours in panel b have been revised to improve the distinction between the two parameters, which is now presented in Fig. 9 in the revised manuscript. Colours used in the presentation of figures have been checked for colour blindness.

In Figure S10, please create separate panels for plotting the MDOT and d18O_sw stacks.

Former Fig. S10 has been updated to present MDOT and $\delta^{18}O_{seawater}$ in separate panels for improved clarity. This figure is now Fig. S12.

Point-by-point reply to Reviewer comments

AC1 'Review', Anonymous Referee #1, 27 Oct 2025

This paper measures and analyses Mg/Ca-based deep ocean temperature (DOT) and d18O_benthic and the previously obtained DOT to determine d18O_seawater of the Iberian Margin site U1385 across the last 1.5 Ma covering the Mid-Pleistocene Transition (MPT) and sets results into context of others by (a) making basin- and global-wide averages using 4 other sites, and (b) discussing especially difference to the recent finding of Clark et al. (2025) on mean ocean temperature (MOT) and the deconvolution of d18O_benthic into temperature and seawater components. It concludes that while there is hardly any change in DOT across the MPT there is a drop in d18O_seawater around 900 ka, rather the opposite of what is proposed in Clark et al. (2025).

I cannot judge the methods, so my comments will solely be on the interpretation.

I think this is a solid piece of work on a new data set across the MPT. The discrepancy with the Clark et al interpretation somehow illustrates the state of knowledge that we are facing right now. So, more details here on what is compared and why things differ would help the reader to set things better into context. See my comments below for improvements:

We agree and will provide more context to highlight the different interpretations of deep ocean temperature and $\delta^{18}O_{seawater}$ across the MPT. This paper cannot resolve the issue but we note noble-gas proxies of mean ocean temperature (MOT) across the MPT in the Beyond EPICA-Oldest Ice core will soon provide much needed evidence to test the different interpretations.

We now include two additional Sections (4.3 and 4.4) and two improved figures (Fig. 10 and 11) in the revised manuscript to provide more context to the different interpretations of deep ocean temperature and $\delta^{18}O_{seawater}$ across the MPT. We now compare our MDOT with the recently published noble-gas-derived mean ocean temperature (MOT) record of Shackleton et al. (2026). The two agree well and show no major change in MOT or MDOT across the MPT, contrary to the reconstruction of Clark et al. This was highlighted in a News and Views article in Nature by Eric Wolff (*doi: https://doi.org/10.1038/d41586-026-00636-3)*

Major issues:

1. - The used equations for calibration (Eq. 2-4) come with uncertainties, which are completely ignored here. For example, in Weldeab, Arce and Kasten (2016) the full equation is given as Mg/Ca [mmol/mol] = (0.36 ± 0.02) ∗ BWT [°C] + (2.22 ± 0.19). Furthermore, both the Mg/Ca-T calibrations are rather scatted in the temperature range used here (-1 to +6°C). Therefore I suggest to add discussions how these ignored uncertainties impact the results, at best recalculate propagated errors using uncertainties on all the parameters that go into the calibration curves. The 3 equations 2-4 should also be reworked to properly state units (so far it is only mentioned, that T comes as °C).

We thank the reviewer for highlighting the importance of accounting for uncertainty in Mg/Ca–temperature calibrations. To evaluate the impact of calibration uncertainty on our reconstructions, we used an updated version of PSU Solver that now includes calibration error (thank you to K. Thirumalai, who kindly provided it), and explicitly propagates uncertainties in the slope and intercept of published Mg/Ca–temperature equations.

We now provide a full list of the species-specific Mg/Ca–temperature calibrations and error used in this study in Table 2 of the revised manuscript. The calibration error has been considered using an updated version of PSU Solver.

Calibration uncertainty was incorporated by Monte Carlo sampling of the reported regression coefficients (e.g., Mg/Ca = (0.36 ± 0.02)* BWT + (2.22 ± 0.19); Weldeab et al., 2016) and for laboratory uncertainty we used a standard uncertainty of 0.08 ‰ for $\delta^{18}O$ and 0.05 mmol mol^{-1} for Mg/Ca. Comparison of reconstructions with and without calibration error demonstrates that inclusion of calibration uncertainty produces only modest changes in temperature uncertainty envelopes and has a negligible effect on $\delta^{18}O_{seawater}$ estimates.

These results indicate that while calibration uncertainty does not materially affect the conclusions of this study. We are happy to include calibration error, where available, for completeness in the revisions.

Full details of our revised approach are provided in Section 2.5 including:

Revised Methods sentence:
Uncertainties associated with Mg/Ca–temperature calibration slopes and intercepts were propagated explicitly within a Monte Carlo framework by sampling calibration parameters within their reported uncertainties during each iteration (n = 1000).
Lines 268–270

Equations 2-4 have now been corrected to properly state units.
Between lines 275–286

2. - Several times Raymo et al. (2006) is discussed as an probably cause for the deduced salinity changes. However, the global SST compilation in Clark et al (2024) had no power in precession which is the dominant frequency in the Raymo et al (2006) hypothesis. This is discussed in part 3 of the Clark et al papers, just published in Science (doi: 10.1126/science.adv8389). I therefore think this hypothesis can be drop, or at least add this discussion point made in Clark et al to this idea.
3.

The global SST compilation of Clark et al. (2024) shows little to no power in the precession band, because the stacking, low-resolution records, and age model uncertainties have likely suppressed the power of precession in the Clark reconstructions (Liautaud et al., 2020; Liautaud, 2021). The absence of detection of significant precession in global mean SST does not preclude precession-paced influences on regional surface salinity or ice–ocean interactions.
We therefore retain Raymo et al. (2006) as one of several possible mechanisms as partially supported by McKay et al. (2012), alongside alternatives such as Adkins (2013), which are now more fully discussed.
Lines 490-498

4. - In the comparison to Clark et al (Figure 12) it is important to plot both, the own and the Clark data with uncertainties. Only then can we see if there is overlap or complete disagreement in certain periods. I actually see at least two main differences in T or both reconstructions. First, the drop in T across the MPT in Clark, missing here (discussed by the authors). Second, the much larger G/IG amplitudes in T after the MPT in the Clak data (not discussed). Thus, both T data sets never show similar things. For the part after the MPT Clark et al. used a ratio of GMSST:MOT=1.1, thus no information from deep ocean T are used which could explain the offset between both T. Furthermore, the d18O_sw from Clark contain the removal of a long-term increase of 0.083 permil/My in d18O_benthic, which is not included in d18O_sw of this study. Although this effect is small (1.5 Myr + 0.083 = 0.12 permil) it needs to be mentioned to understand the comparison.

We agree that additional contextual detail and the explicit inclusion of uncertainties in Figure 12 are essential for a meaningful comparison between our reconstructions and those of Clark et al. (2025). In the revised manuscript, we will include both 1σ and 2σ uncertainty envelopes in Figure 12 and expand the accompanying discussion to clarify methodological differences between the two approaches.
Previously Fig. 12 included comparisons of our reconstructed deep ocean temperature and $\delta^{18}O_{seawater}$ with those of Clark et al., 2025. We have added the 1σ and 2σ uncertainty envelopes of our data and that of Clark et al. to Figs. 10 and 11.

As noted by the comment by Peter Clark, uncertainties should be shown when comparing the temperature reconstructions. In the revised figure, we will therefore display the full propagated uncertainties for both records. We also identified a small but relevant difference in the reference used for the Early Holocene temperature adjustment. In our original submission, temperatures were referenced to 0 °C at 10 ka using a −1.5 °C offset, whereas Clark et al. (2025) applied a −1.73 °C offset. To ensure consistency, our temperature record will be adjusted accordingly in the revision. In addition, following the Reviewer's suggestion and subsequent discussion with Peter Clark, we will compare our Mg/Ca-based deep-ocean temperature estimates with the ΔT > 2000 m reconstruction of Clark et al. (2025), rather than with MOT, as this provides a more realistic comparison.
Fig. 10 now compares our MDOT record with the ΔT > 2000 m reconstructions of Clark et al. (2025) instead of MOT. The two deep-water temperature records continue to show substantial disagreement prior to ~900 ka as well as during some glacial stages thereafter (MIS 18, 12 and 2). We suggest this discrepancy is real and further elaborate on this in response to point 5 below.

In our study, uncertainties are fully propagated for each individual Mg/Ca-based temperature and $\delta^{18}O_{seawater}$ estimate using the bootstrap resampling approach implemented in PSU Solver (Thirumalai et al., 2016), based on the input data listed in Table 2. The uncertainties associated with ΔMOT and $\delta^{18}O_{seawater}$ are taken directly from the published data tables of Clark et al. (2025); we refer the reader to that study for details of the underlying assumptions and error analysis.
We also acknowledge the Reviewer's point that the $\delta^{18}O_{seawater}$ reconstruction of Clark et al. (2025) incorporates the removal of a long-term increase of 0.083 ‰ Myr^{-1} in benthic $\delta^{18}O$, which is not applied in our reconstruction. We think such an arbitrary adjustment for diagenesis is unproven but even if it were applied to our data, the maximum adjustment to benthic $\delta^{18}O$ would be 0.12 per mil at 1.5 Ma and would decrease with

time. This adjustment is small considering the error of the $\delta^{18}O_{seawater}$ compilations and, if applied, would result in an increase in the divergence between the two mean $\delta^{18}O_{seawater}$ curves shown in Figure 12.
We have added this point to the caption of Fig. 11.

We recognise that the construction of global or near-global stacks involves substantial uncertainty, and that methodological choices—such as reference temperature, treatment of long-term trends in benthic $\delta^{18}O$, and assumptions about deep-ocean heat storage—exert first-order control on the resulting reconstructions. Nonetheless, a key distinction remains: our Mg/Ca-based estimates suggest relatively modest changes in mean deep-ocean temperature across the MPT (with possible regional variability), whereas the reconstruction of Clark et al. (2025) indicates a pronounced cooling across this transition. Differences in inferred temperature histories necessarily propagate into differences in reconstructed $\delta^{18}O_{seawater}$, given the well-established coupling between temperature and benthic $\delta^{18}O$.
Lines 572- 585
Discussed further in point 5 below.

These contrasting interpretations underscore the importance of independent constraints on deep-ocean temperature, such as those anticipated from noble-gas measurements in blue ice (COLDES Project) and the Beyond EPICA–Oldest Ice core.

5. - How can we learn even more from the comparison to the Clark et al. paper? Your stacked d18O_benthic compares very good to the Prob-stack d18O_benthic for the last 400 kyr, but differs somehow before (Fig. 9a). Maybe this illustrates, that for the last 400 kyr your stack is representative of the global signal, but not so much before. Maybe calculate the difference between both d18O_benthic time series to use as a proxy of global representativeness. If you plot this difference then together with the T and d18O_sw comparison to Clark (Figure 12), the reader would see, how much your stack are (or are not) global signals. Maybe this difference can also be exploited as a correction for d18O_sw and/or T, but I am not sure it will work or if it is a good idea. Furthermore, Fig 12 needs different y-axis labels. Your stack of d18O_sw is NOT a global stack (while d18O_sw of Clark is), it is a stack of globally distributed cores from waters below 2km. Actually, maybe at best you should not compare to the MOT from Clark but to the time series of T changes below 2km (Fig 16d in Clark et al., 2025, data should be available from the authors), which uses GMSST and a change in heat storage efficiency (=Delta MOT / Delta GMSST). This has large uncertainties 1.5-0.9 Myr, but basically allows (more or less) a long-term flat line across the MPT, similarly as here (no trend in temperatures).

We have replaced the term 'global' with 'globally distributed' records when referring to our MDOT and mean $\delta^{18}O_{seawater}$ stacks.

We appreciate the Reviewer's constructive suggestion regarding how further insight might be gained from comparison with the Clark et al. (2025) reconstructions.

As suggested, we have obtained the deep-ocean temperature (DOT) reconstruction (T > 2000 m) from Peter Clark and will replace the previous comparison to MOT with the DOT in a revised version of Figure 12. This comparison provides a more appropriate basis for evaluating deep-ocean temperature evolution across the MPT. We have also separated the y-axes and revised labels to clearly distinguish between the globally averaged $\delta^{18}O_{seawater}$ reconstruction of Clark et al. (2025) and our volume-weighted stack derived from globally distributed sites below 2 km depth. A preliminary comparison Figure 12 is shown below which will be updated after our revised analyses. While the deep ocean temperature records show better correspondence after the MPT, compared to the MDOT stack against MOT, there remains a notable difference in temperature during MIS 22.
As suggested by the reviewer, we now compare our Mg/Ca-derived deep-ocean temperature estimates with the $\Delta T > 2000$ m reconstruction of Clark et al. (2025) in Fig. 10d. The discrepancy in temperature between our MDOT stack and Clark's DOT ($\Delta T > 2000$ m) is evident between ~900 ka (~MIS 22) and the base of the records at 1500 ka. Differences also exist in glacial MIS 18, 12 and 2 when DOT ($\Delta T > 2000$ m) shows greater glacial-interglacial amplitudes than MDOT.

As observed by Reviewer 1, there is a small difference between our benthic $\delta^{18}O$ stack based on 5 records and that of Ahn et al. (2017) based on over 100 records, raising the question how representative our volume weighted benthic $\delta^{18}O$ stack is of the global signal. Our benthic $\delta^{18}O$ stack is within error of Prob-stack for the last 800 ka and displays some kyr differences in the older record (see Fig. Xa below). We attribute this to several factors. Our stack is weighted by ocean volume whereas Prob-stack is not. Because the deep Pacific has the greatest volume our stack is weighted towards the Pacific. In contrast, Prob-stack is more heavily influenced by Atlantic records because there are more of them and all ocean basins are weighted equally. Because the amplitude of benthic $\delta^{18}O$ variation is greater in the Atlantic than in the Pacific, we would expect Prob-stack to show a greater amplitude than a Pacific-weighted stack (Fig. Xb). We have not attempted to correct for any differences between our volume-weighted stack and the equally-weighted Prob-stack because the differences are relatively small considering error and these differences cannot explain the rather large differences in temperature (independently determined by Mg/Ca) and calculated $\delta^{18}O_{seawater}$ (affected by benthic $\delta^{18}O$) with Clark et al. (2025).

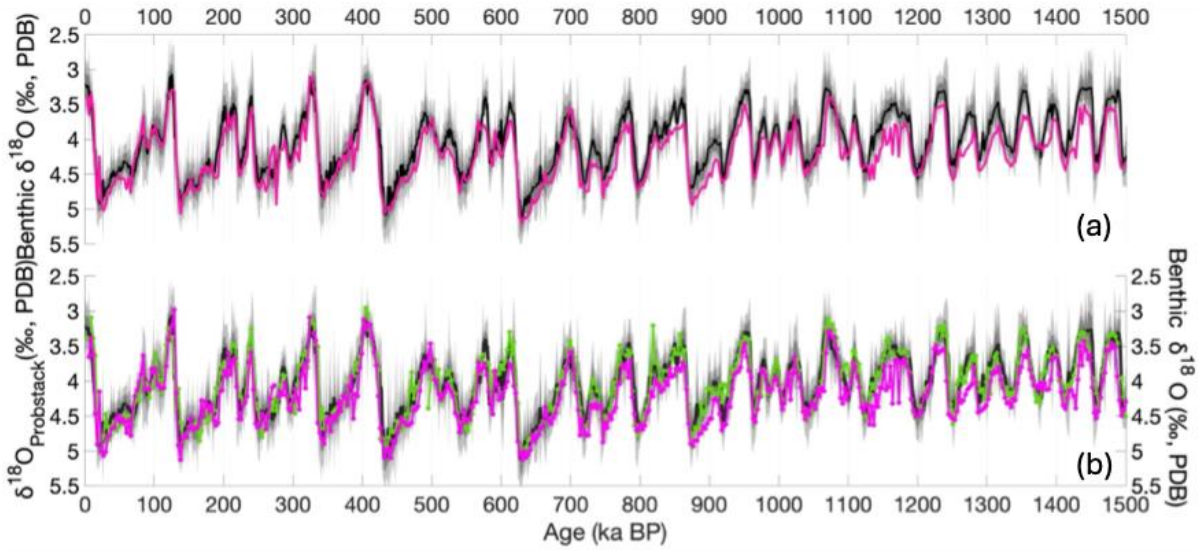

Fig. X (a) Comparison between our basin-volume-weighted benthic $\delta^{18}O$ stack (pink) and the Prob-stack (black) with 1 and 2 σ uncertainty given by (Ahn et al., 2017). (b) Atlantic (green) and Pacific (pink) benthic $\delta^{18}O$ stacks against the Prob-stack (black) (Ahn et al., 2017). Note the correspondence between our Atlantic stack and Prob-stack and the offset with our Pacific stack.
Panels in these figures are presented in Figs. 7 and 8 of the manuscript.

6. - So all together, (a) considering if a long-term trend in d18O_benthic is included in plotted data or not, (b) which kind a T curves are compared, (c) highlighting that your approach is a least partially not representing global signals, (d) having some more emphasis on (missed) uncertainties, both your data and the Clark data seem to be not so much in disagreement as so far discussed.

Whilst the uncertainties associated with both the Clark et al. (2025) compilation and our reconstructions are considerable, the differences are difficult to reconcile based on methodological grounds. We note that the recent noble-gas-derived mean ocean temperature (MOT) record from Shackleton et al. (2026) is consistent with relatively stable temperatures over the MPT and an increase in ice volume (Wolff, 2026), in contrast to the reconstruction of Clark et al. (2025). We have added the ice core MOT to the manuscript but emphasized the critical importance of obtaining an MOT record from BE-OI across the MPT.

Minor Details:

- Not all SI figures are cited in the main text. Either cite (and order the SI figures according to citation) them or delete them.

We will ensure all supplementary figures are cited in the main text and ordered accordingly.
Done

- Maybe I missed it, but why does Fig. 2 only show the time 700-1500 ka. Is this the only new part and 0-700 ka was already published elsewhere? Details can probably be extend in section 2.

As only four samples fall within the 0–700 ka interval, these data were initially excluded from the main plot to improve figure clarity. They will now be presented for completeness.

Fig. 2 in the original manuscript has now been revised and moved to the Supplement as Fig. S3. It now includes the full comparison of deep-water temperature and $\delta^{18}O_{seawater}$ estimates derived from $\delta^{18}O$ and Mg/Ca measurements of *Uvigerina peregrina* and *Globobulimina affinis* at Site U1385. The records only cover 400–1450 ka because there is no comparative data for *G. affinis* for the interval 0–400 ka.

- Line 435 or so: I believe Pacific bottom water temperature being near-freezing throughout the MPT and thereafter has first been shown in Siddall et al. (2010: doi:10.1016/j.quascirev.2009.05.011).
We revised our discussion to explicitly acknowledge that near-freezing Pacific bottom water temperatures during and after the MPT were first demonstrated by Siddall et al. (2010).

In the revised manuscript we now cite Siddall et al. (2010).
Line 582

- Line 668: Raymo et al (2006) was published in Science, not Nature.
This journal reference has now been corrected.
Line 952

AC2 'Review', Anonymous Referee #1

In a paper published yesterday in this journal

Larsson, V. and Jung, S.: Persistent contamination in benthic-foraminifera-based Mg/Ca thermometry using standard cleaning methods, Clim. Past, 21, 1871–1894, https://doi.org/10.5194/cp-21-1871-2025, 2025.

Present a different calibration for Uvigerina peregrina than the one used here (their Fig. 12) and some contamination issues are discussed. Please add some thoughts, if and how their findings are relevant for your study here.

Recent work has highlighted the sensitivity of *U. peregrina* Mg/Ca to residual contamination under certain cleaning protocols (Larsson and Jung, 2025); however, the stringent data quality evaluation employed in this study (Section 2.4) and the absence of covariance between Mg/Ca and Fe/Ca or Mn/Ca in our dataset indicates that contamination does not significantly influence the Mg/Ca-derived temperatures presented here.

The alternative calibration for *U. peregrina* shown in Fig. 12 of Larsson and Jung (2025) is based on earlier work by Elderfield et al. (2006) which relied on a relatively small number of core-top samples spanning ~1.5–12 °C. In contrast, the calibration of Elderfield et al. (2010, 2012) used in this study incorporates independent constraints on glacial-interglacial temperature change across the LGM-Holocene, yielding a Mg/Ca–temperature sensitivity of ~ 0.1 ± 0.013 mmol mol^{-1} °C.
We now cite both Elderfield et al., 2006 and Larsson and Jung, 2025. Lines 205 and 212

Related discussion of the calibration choice and its applicability at low temperatures is provided in our response to Reviewer 2 (Equation 2).

AC3 'Review', Anonymous Referee #2

Thomas et al. present combined benthic foraminiferal Mg/Ca and oxygen isotope data from IODP Site U1385 at the Iberian Margin to reconstruct $\delta^{18}O_{sw}$ (seawater oxygen isotopes) covering the past 1.5 Myrs. The record is compiled with similar datasets from the Atlantic and Pacific sectors of the Southern Ocean as well as from the North Pacific, allowing the authors to produce a deep ocean $\delta^{18}O_{sw}$ and temperature stack that reflects global changes. The study uses a (to my knowledge) novel approach by applying a volume-weighted averaging scheme to calculate the global $\delta^{18}O_{sw}$ stack, that accounts for the relative sizes of the major ocean basins. This method is contrasted with previous approaches that did not include ocean volume-based corrections, providing new insights into the interpretation of global signals. The main findings of the study indicate increased salinity in the deep Pacific and reduced deep-water temperatures in the Atlantic during glacial periods after 900 ka, compared to the interval between 1.5 and 0.9 Ma. The authors interpret these results as evidence for enhanced glacial water mass stratification in the Pacific after 0.9 Ma, leading to a greater sequestration of atmospheric carbon dioxide in the deep ocean, as well as increased salinity in surface waters within regions of southern component water formation around the Antarctic margins.

The study draws on an extensive dataset, and the methodology appears sound and appropriate to the study's objectives. The resulting dataset has the potential to make a valuable contribution to our understanding of global ocean temperature, salinity and ice volume changes across the Mid-Pleistocene Transition (MPT). However, the strength of the methodological and analytical work is not matched by the clarity and depth of the data presentation and interpretation. In its current form, the manuscript does not fully exploit the scientific potential of the dataset, and several key aspects of the analysis and discussion require revision and elaboration. Below I outline my major and minor points of criticism in the hope that the authors find them useful in improving and revising their manuscript.

We thank the Reviewer for their detailed review of our paper and appreciate their helpful comments.

We have modified the introduction and discussion to emphasise the implications of our results to understanding changes in global ocean temperature, salinity and ice volume across the Mid-Pleistocene Transition (MPT).
Lines 24–60; Sect. 4.1 464–498; Sect. 4.2 Lines 523-547

In the following, I will first outline several broad, major concerns that affect the overall interpretation and structure of the manuscript. I will then provide detailed, section-by-section comments and suggestions aimed at improving the scientific robustness and overall comprehensiveness of the manuscript.

1. The manuscript warrants more transparency in the many nuances of the interpretation and should clearly outline the limitation of the data interpretation. For instance, the authors base their conclusion of an enhanced density stratification throughout the entire deep Pacific after 900 ka (compared to 1.5–0.9 Ma) primarily on data from only one Site in the South Pacific and one in the North Pacific. This interpretation is presented as a definitive conclusion, rather than as a working hypothesis supported by limited regional evidence. Potential variations in salinity throughout the entire water column—and their corresponding effects on vertical stratification—are not discussed, and possible regional hydrographic differences are insufficiently addressed in my opinion.

   We acknowledge the spatial coverage in the Pacific is limited, and that our interpretation of enhanced deep Pacific density stratification after ~900 ka should be framed more cautiously. In the revised manuscript, we will explicitly rephrase this interpretation as a working hypothesis supported by available North and South Pacific sites, rather than as a definitive basin-wide conclusion.
   Lines 13–22; 480–520; 658–669

   We will also expand the discussion to acknowledge potential vertical and regional variability within the Pacific water column and clarify that our conclusions pertain to the sampled deep-water masses rather than the entire Pacific basin. Additional text will be added to discuss how salinity changes at different depths, as well as unresolved regional hydrographic differences, could influence the reconstructed stratification signal.
   For reference to Pacific deep-water see Lines 121–133

And for associated changes in salinity Lines 479–517

These revisions will improve transparency regarding the limitations of the dataset and the scope of the conclusions that can be drawn.
In the revised manuscript we provide a new section – Sect. 4.5 Limitations and Future Opportunities – outlining the main limitations of our study. Lines 633–651

Elderfield et al. (2012) argued that more than half of the flux of cold bottom water entering the major ocean basins transits the Southwest Pacific as part of the Pacific Deep Western Boundary Current (DWBC) (Elderfield et al., 2012), consistent with earlier work by McCave et al. (2008) and Whitworth et al. (1999). This view is supported by evidence from the Last Glacial Maximum (Insua et al., 2014) and, more recently, across the MPT (Li et al., 2026). Although limited to two sites, it is reasonable to assume that Pacific Sites 1123 and 1208 may more broadly reflect conditions in the deep Pacific. Similar records from other sites are needed to test if this assumption is correct.
The revised text now provides a clearer explanation of why the two Pacific sites are considered to be broadly representative of changes in Pacific deep water (Chandler et al., 2024 and references therein).
Lines 121–131

2. Throughout the manuscript, several sections lack a clear explanation, and important arguments are presented in a way that they appear somewhat ambiguous. This occurs in multiple places and affects the overall clarity of the manuscript (e.g. comments on lines 42, 127, 172-177, 191, section 4.1). I recommend using more precise wording in many sections to avoid confusion.

   As suggested by the reviewer we will revise the highlighted passages (e.g. lines 42, 127, 172–177, 191, and Section 4.1) to improve clarity, reduce ambiguity, and ensure that key arguments are stated explicitly and consistently.
   Addressed under Detailed Comments below

3. In the abstract and discussion, the authors attribute the observed increase in deep Pacific salinity after 900 ka to changes in the source regions of Southern Component Waters (SCW), i.e., Antarctic margin regions (e.g., the Weddell and Ross Seas). However, this interpretation (i.e. a key finding in the abstract) is not substantiated by any evidence presented in the study. In my view, the manuscript provides insufficient discussion of existing literature and data that would support such large-scale changes in SCW source regions during this interval. Revision is needed here. Moreover, as PDW is a mixture of several water masses, the possibility of temporal changes in PDW end-member composition—and their influence on reconstructed salinity—should be addressed, to support the suggested changes in SCW across the MPT. The authors could consider potential variations in inter-basin connectivity within the Southern Ocean, e.g. using the Antarctic Circumpolar Current strength reconstruction of Lamy et al. (2024), which could help assess whether large-scale circulation changes—rather than solely local shifts in AABW formation—may contribute to the observed Atlantic–Pacific differences across the MPT.

   We agree with the Reviewer that the attribution of increased deep Pacific salinity after ~900 ka requires more careful framing and broader contextualisation. In the revised manuscript, we present changes in Southern Component Water (SCW) source regions as a working hypothesis rather than a definitive conclusion.
   The Abstract has been revised accordingly, and we now present changes in freshwater fluxes to the source regions of Southern Component Water (SCW) as a working hypothesis that requires further testing rather than a definitive conclusion of this study. Lines 13–20; Sect.4.1, 4.2, 4.5 and 5.

   We will substantially expand the discussion to integrate recent and relevant literature that bears on Southern Ocean processes across the MPT, including evidence for changes in Antarctic ice-sheet configuration, abyssal stratification, and deep-ocean circulation (McKay et al., 2012; An et al., 2024; Wirths et al., 2025; Li et al., 2026; Scherrenberg et al., 2025). In addition, we now explicitly consider the multi-sourced nature of Pacific Deep Water and

discuss how temporal changes in PDW end-member composition could influence reconstructed deep Pacific salinity.
We now draw on a wide range of recent and relevant studies that support a role for the Antarctic Ice Sheet and Southern Ocean in affecting abyssal stratification. We cite modelling results and modern data supporting freshwater fluxes as a dominant control on deep-ocean stratification (Sun et al, 2016 and references therein; Tan and Thurnherr, 2023, 10.1029/2022GL102422).
Sections 4.1 and 4.2.
We infer temporal changes in end-member composition in Lines 535-537 and through notable observations in the temperature and $\delta^{18}O_{seawater}$ of Site 1094 in Lines 419–423.

To address the possibility that large-scale circulation changes, rather than only local changes in AABW source areas, may also have contributed to the observed Atlantic–Pacific contrasts, we have incorporated discussion of inter-basin connectivity within the Southern Ocean. In particular, we draw on recent reconstructions of Antarctic Circumpolar Current (ACC) strength (Lamy et al., 2024) to assess whether changes in ACC strength may have modulated deep-ocean exchange and contributed to the large-scale patterns observed across the MPT including in the Southern Ocean (Williams et al., 2024).
We examined the correspondence between our records and ACC reconstructions in detail. Some intervals (e.g., MIS 31 and 24; ~1050-930 ka) show interesting relationships with ACC strength in the Pacific sector, we note that different sectors of the Southern Ocean seem to display asynchronous changes in the ACC (Wu et al., 2025) making it difficult to use one record as representative of the ACC as a whole.

Throughout the manuscript, statistical analyses are insufficiently applied when comparing different datasets and time intervals. Statements referring to “significant” differences are frequently made without presenting the results of any formal statistical tests. To substantiate these claims, appropriate statistical analyses (e.g., t-tests, ANOVA) should be conducted and reported. Including quantitative measures of variability and statistical significance would greatly strengthen the robustness of the interpretations and allow readers to assess whether observed differences between regions or time periods are statistically significant.

We agree with the Reviewer that statements referring to “significant” differences should be backed by formal statistical testing.

Where necessary, in the revised manuscript, we will apply appropriate statistical tests (e.g. t-tests or non-parametric equivalents, where applicable) when comparing different datasets, and we will report the results explicitly including quantitative measures of variability and uncertainty to allow readers to assess the robustness of observed differences. Where statistical significance cannot be demonstrated, we will revise the text to avoid overinterpretation.
We now use the word “significant” only where statistical testing has been applied. Statistical testing has been applied to changes in mean glacial $\delta^{18}O_{seawater}$ values in the ocean volume-weighted stack and is shown to be significant Lines 429-430

4. In my opinion, the introduction lacks a clear structure and does not fully establish the central research question. Integrating relevant key studies (e.g. Quin et al., 2022)—many of which are discussed in the discussion (e.g. Hines et al., 2024; Pena and Goldstein 2014, Farmer et al., 2019; Lear et al., 2016) —more prominently at the start would help frame the topic, clarify existing knowledge gaps, and better motivate the study within its scientific context. Moreover, a concise overview of the main hypotheses proposed to explain the MPT—including e.g. the regolith hypothesis (e.g. Clark and Pollard, 1998, Clark et al., 2006), the expansion of North American and/or East Antarctic ice sheets (e.g. An et al., 2024; Bintanja and van de Wal, 2008; Raymo et al., 2006), or feedback mechanisms involving changes of the marine carbon cycle (e.g. Chalk et al., 2017; Willeit et al., 2019)—would provide important context and more clearly situate the study within the broader landscape of MPT research.

In the revised manuscript, we have restructured the Introduction to provide a clearer conceptual framework and to better motivate the study.
Specifically, the Introduction now opens with a concise overview of the MPT and summarises the principal hypotheses proposed to explain it.

Key studies previously introduced later in the Discussion (e.g. Lear et al., 2016; Farmer et al., 2019; Qin et al., 2022; Peña and Goldstein, 2014; Hines et al., 2024) are now integrated into the Introduction to clarify existing knowledge gaps and to position our study more clearly within the current literature.

We then explicitly articulate our central research question: *Was there a shift in abyssal ocean density structure across the MPT?* This leads directly into our use of benthic $\delta^{18}O$ and Mg/Ca records to disentangle the temperature and $\delta^{18}O_{seawater}$ components of density stratification and to evaluate their implications for deep-ocean circulation across the MPT.
See the revised introductory paragraphs on Lines 24–61

5. Overall, the figures would benefit from improved clarity and more explicit information to help the reader follow the comparisons and interpretations of the many timeseries presented. It would be helpful to include the location of each record directly in the figure, for instance by labelling the regions (e.g., Pacific, Southern Ocean, Atlantic…) in the figures. For the Iberian Margin record, it should be clearly indicated which data were generated in this study and which were compiled from previous publications (different symbols or colours). Finally, it would be beneficial to show key time intervals discussed in the manuscript at higher temporal resolution, rather than referring exclusively to the long composite records. This would allow for a clearer evaluation of the main features and transitions emphasised in the text. If feasible, reducing the overall number of figures may improve the readability of the manuscript, as the large number of time series and figures can be somewhat overwhelming for the reader.

   We appreciate the Reviewer's observations and will revise the figures to improve clarity and readability.

   We will clearly indicate which Iberian Margin data were generated in this study and which were compiled from previously published records, using distinct symbols and/or colours. The regional origin of each record (e.g., Atlantic, Southern Ocean, Pacific) will be explicitly labelled within the figures.
   Fig. 2 showing the Iberian Margin data now clearly indicates the compilation records
   Figs. 3, 4, S7 and S8 have been annotated accordingly

   We will also reduce the number of figures where possible by combining or eliminating redundant panels. In addition, we will include more detailed plots over key intervals discussed in the text (e.g., pre-1.2 Ma, 1.2–0.65 Ma, and 0.65–0 Ma for the ocean stacks) to better highlight the main features and transitions contained in the records.
   We have reduced by the number of figures by one in the main manuscript but are reluctant to make further reductions. We produced some detailed plots focussing on specific time intervals but, upon reflection, decided not to include them because they added to the total number of figures yet did not add much to the discussion, which is centred around the full records. The full data set is provided should readers wish to examine specific intervals more closely.

6. Salinity on the Practical Salinity Scale (PSS-78) is a dimensionless quantity, and the use of the unit "PSU" (Practical Salinity Unit) is therefore incorrect. The authors should avoid using "PSU" throughout the manuscript. Instead, it should be stated upon first mention that all salinity values are reported on the Practical Salinity Scale (PSS-78) and are thus unitless. It would be preferable to adopt the TEOS-10 (Thermodynamic Equation of Seawater, since 2009) scale and thus report Absolute Salinity in g/kg. However, given that global datasets such as GLODAP and WOA report salinity on the PSS-78 scale, calculated from conductivity, temperature and pressure, clarification in the text that values are dimensionless and follow PSS-78 are sufficient.

We have corrected the terminology to avoid using "PSU" when referring to salinity values. A note has been added beneath Table 1 clarifying that all salinity values follow the Practical Salinity Scale (PSS-78) and are unitless.
The ODV Fig. S2 has also been corrected

Detailed comments:

Line 13, 14: The study would benefit from more cautious wording, as it presents no direct evidence for changes in the carbon storage capacity.

We thank the reviewer for this suggestion. We have revised the abstract to adopt more cautious wording and to avoid implying direct evidence for changes in deep-ocean carbon storage. Specifically, the discussion of carbon storage has been reframed as being consistent with existing hypotheses and prior studies, rather than demonstrated directly by our data. The abstract now emphasises that our results support a physical mechanism that may contribute to the MPT, while acknowledging that confirmation awaits independent evidence from ice cores (e.g., Beyond EPICA–Oldest Ice).

The abstract now states that enhanced abyssal stratification is "consistent with previous hypotheses and data supporting increased glacial carbon storage," rather than asserting a direct causal relationship.
In the revised manuscript we adopt more cautious approach and present a potential sequence of cause/effect events as a working hypothesis, rather than a definitive interpretation. We outline mechanisms from the literature that may have contributed to deep ocean carbon storage.
Lines 13–20
Section 4.2, Lines 522–547.

Line 23: One of the main conclusions of the study concerns a salinity change across the entire deep Pacific after 900 ka based on two Sites. However, the text states that $\delta^{18}O_{sw}$ only reflects variations in global ice volume and **local** hydrographic effects. This phrasing is somewhat ambiguous and should be clarified.

We have revised the text to more explicitly distinguish between the global and local components of $\delta^{18}O_{seawater}$.

The sentence now reads:
The interpretation of benthic $^{\delta18}O$ records is complicated, however, because the signal is dependent on both temperature and the oxygen isotopic composition of seawater ($\delta^{18}O_{seawater}$) which varies with global ice volume as well as regional hydrographic effects (Waelbroeck et al., 2002), such as salinity changes at sites of deep-water formation associated with changes in ocean circulation.
Revised to "However, interpretation of benthic $\delta^{18}O$ records is complicated because the signal is dependent on both temperature and the $\delta^{18}O_{seawater}$, the latter varying with global ice volume and regional hydrography including salinity changes at sites of deep-water formation (Waelbroeck et al., 2002)."
Lines 63–65

Line 32: This statement is too vague. Since this appears to be one of the main motivations of the present study, the divergent interpretations should be clearly introduced in the Introduction.

We have revised the Introduction to explicitly outline the divergent interpretations of Quaternary temperature and ice-volume evolution that motivate this study.
Lines 74–78

Line 40-43: Please specify what specific concerns are being referred to here and clarify whether "previous findings" refers to Yu & Broecker (2010) or Sosdian & Rosenthal (2009).

We have revised the text to identify the concerns raised by Yu and Broecker (2010) regarding the Mg/Ca-derived temperature record at Site 607 (Sosdian & Rosenthal, 2009), thereby clarifying the

reference to “previous findings.” Additional discussion of the carbonate ion effect remains in the Supplementary Information.
Lines 83–88

Table 1: Please use a consistent coordinate format. The authors should also cite the sources or datasets from which present-day temperature and salinity data were obtained.

All site coordinates have now been converted to decimal degrees, following the standard convention used in oceanographic databases.
We have also cited the sources and datasets from which present-day temperature and salinity values were obtained in Table 1.
Table 1 completed

Fig. 1: Please indicate in panel (a) where the transects shown in panels (b) and (c) were taken. Given potential differences between the eastern and western Atlantic, this spatial context is important. It would also be useful to include a single transect (covering the Pacific, Southern Ocean, and Atlantic) directly in the main manuscript rather than in the Supplement, as temperature gradients between these basins are a key aspect of the study.

We will revise Figure 1(a) to explicitly indicate the transect route shown in panel (b). The meridional profile shown in panel (c) was constructed as an idealised Pacific representation based on a very limited number of available deep ocean $\delta^{18}O_{seawater}$ estimates and does not correspond to a uniquely defined transect. We will clarify this explicitly in the figure caption and text, or to avoid potential confusion, we will either (i) revise panel (c) or (ii) remove this panel from the main figure and provide an indication of the transect route for Pacific temperature and salinity instead (i.e., panel (b) in Figs. S1 and S2).
Figure 1 has been revised to indicate in panel (a) the locations of the transects shown in panels (b) and (c).
Panel (c) has been revised using oxygen isotope data from LeGrande and Schmidt (2006) extracted along the line of longitude of Site 1123 (-171.50°).
Furthermore, transect locations for Atlantic and Pacific hydrographic profiles of temperature and salinity (Figs. S1 and S2) using GLODAPv2.2021 and GLODAPv2.2023 and GEOSECS in ODV are presented in Fig. S1.

In response to the Reviewer’s second point, we will move our meridional temperature profiles into the main manuscript to better illustrate large-scale temperature gradients that are central to our interpretation.
We explored this approach but retained the separate basin-scale hydrographic profiles (temperature and salinity) in the Supplement (Figs. S1–S2). Presenting the basins individually provides a clearer picture of spatial gradients and avoids overcomplicating the figures.

Fig. 3: Please use a consistent capitalisation for “Marine Isotope Stage” throughout the manuscript (either always “Marine isotope stage” as in line 81 or “Marine Isotope Stage,” as in line 103). In addition, clarify the basis of the 2σ error bars mentioned in lines 82-83; are these derived from replicate measurements or represent analytical uncertainty or both?

We have ensured consistent capitalisation of “Marine Isotope Stage” throughout the manuscript. In addition, the figures have been revised to incorporate calibration uncertainties; accordingly, the 1σ and 2σ error bars now represent the combined calibration and analytical uncertainties and are not derived from replicate measurements.
Done

Line 89 and 92: The phrase “and have similar seawater physical properties” is somewhat vague. Please specify which properties are referred to (e.g., temperature, salinity, density) or include a reference to Table 1 for clarification.

For clarification, we have included a reference to Table 1.
Line 121

Line 107-109: Do all studies use the same approach for their age model?

The records compiled in this study do not all use identical age-models; however, they follow consistent and well-established approaches. All records, with the exception of MD99-2334 and MD01-2444 for the last 50 kyrs, are based on alignment of benthic foraminiferal δ¹⁸O to the LR04 reference stack or closely related reference stacks. MD99-2334 and MD01-2444 were independently tuned to ice-core chronologies. We now clarify the age-model approaches used for each record in the Methods section and note that, for the interval considered here, these chronologies are mutually consistent within age uncertainties. Accordingly, no additional retuning was applied.
Lines 153–159

Suggested paragraphs:
2.1 Site details and chronology
All previously published records used to compile the ocean stacks are presented on their original published age models. Most are based on alignment of benthic foraminiferal δ¹⁸O to LR04, including Deep Sea Drilling Project (DSDP) Site 607 combined with Chain 82-24-23PC 23PC (Ford et al., 2016; Sosdian and Rosenthal, 2009), and Ocean Drilling Program (ODP) Site 1123 (Elderfield et al., 2012). ODP Site 1208 (Ford and Raymo, 2020) was aligned to the Prob-stack using the HMM-Stack MATLAB code (Lin et al., 2014; Butcher et al., 2017; Ahn et al., 2017) which is effectively identical in structure to LR04 over the interval considered here. The ODP Site 1094 age model (Hasenfratz et al., 2019) was derived by graphical alignment of benthic δ¹⁸O to LR04, with the original chronology of Jaccard et al. (2013) retained where benthic foraminifera were sparse (e.g. MIS 4).
Lines 153–159

The Iberian Margin piston cores MD99-2334 and MD01-2444 were independently aligned to Greenland ice-core records through synchronisation of Dansgaard–Oeschger events recorded in planktonic δ¹⁸O (Shackleton et al., 2000; Skinner et al., 2003; Skinner and Elderfield, 2007). Due to the close agreement in both structure and absolute benthic δ¹⁸O values with Site U1385 over the interval of overlap, no additional adjustments to these published age models were applied.

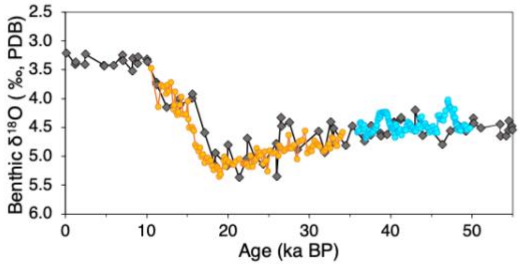

Figure SXX: Benthic foraminiferal δ¹⁸O records from Iberian Margin piston cores MD99-2334 (orange line and circles)) and MD01-2444 (blue line and circles) shown on their original published age models, together with the benthic δ¹⁸O record from Site U1385 (black line with dark grey diamonds) for comparison. The records show close agreement in both structure and absolute values over the interval of overlap, including the timing of major glacial–interglacial and millennial-scale features. This close correspondence indicates that the published age models for the Iberian Margin piston cores are consistent with the Site U1385 chronology and do not require further adjustment.

Line 122-123: To my knowledge the MD01-2444 record spans ~50–35 ka, which does not include the LGM.

We have revised the text to clearly differentiate between the early (50–35 ka; MD01-2444) and late (35–10 ka; MD99-2334, including the LGM) parts of the last glacial period.
Lines 184–187

Line 123-124: Please Specify whether this offset is based on overlapping intervals between the cores or on modern temperature differences (which are not listed in Table 1).

Following Skinner et al. (2007), we applied a constant temperature offset of +0.6 °C to Mg/Ca-derived temperatures from the deeper Site MD99-2334 to account for depth-related cooling and to render the record comparable to Site U1385, and consistent with the nearby piston core MD01–2444. This offset is based on modern deep-water temperature gradients along the Iberian Margin as confirmed in depth-transect observations, ((Hodell, 2014; Skinner et al., 2003, 2007; Skinner and Elderfield, 2007); Table 1) rather than from overlapping downcore intervals. Modern temperature values supporting this adjustment have now been added to Table 1, and the methodological basis for the offset has been clarified in the text. The offset was applied within PSU Solver after temperature estimation and prior to the reconstruction of $\delta^{18}O_{seawater}$.
Lines 187–189

Line 127: I cannot fully understand this sentence. If the text refers to a specific taxonomic level, please use the term "species" to avoid confusion.

The sentence now reads:

"Measurements of benthic foraminiferal $\delta^{18}O$ at Site U1385, paired (same sample) with *Uvigerina peregrina* Mg/Ca analyses to reconstruct $\delta^{18}O_{seawater}$, were predominantly obtained from the same specimens of *Uvigerina peregrina* where sufficient calcite was available prior to cleaning for Mg/Ca analysis (Barker et al., 2003) (n = 560 representing 64 % of Mg/Ca samples)."
The sentence has been removed.

Line 128: Please clarify how comparable the results are between the two genera when both were measured.

We now include the following paragraph:
"Comparison of Mg/Ca–derived deep-water temperatures from *Uvigerina peregrina* and *Globobulimina affinis* at Site U1385 generally agree in both mean values and variability (Fig. 2a), with corresponding $\delta^{18}O_{seawater}$ estimates exhibiting consistent structure (Fig. 2b). This supports the species-specific Mg/Ca–temperature calibrations applied in this study (Elderfield et al., 2006, 2010, 2012; Weldeab et al., 2016) and justifies combining the two datasets into composite records of deep-water temperature and $\delta^{18}O_{seawater}$. The individual *Uvigerina peregrina* and *Globobulimina affinis* time series were used as inputs for the PSU Solver calibration procedures described in Section 2.4 and the combined records form the basis of the deep-water temperature and $\delta^{18}O_{seawater}$ reconstructions (Fig. 3) analysed in Section 3."
A revised version of this paragraph starts from Line 202 to 208

Line 143-144: Please specify which error the authors are referring to. No corresponding uncertainty is shown in Fig. 2a, so it is unclear what uncertainty estimate is referred to or how it was derived.

In the submitted version, the uncertainty referred to in lines 143–144 corresponds to analytical and replicate variability in the Mg/Ca-derived temperatures; however, this uncertainty was not shown in Fig. 2a, which we acknowledge was unclear. In the revised manuscript, all records will be reprocessed consistently through PSU Solver and the stacking procedure, and uncertainties will instead represent the combined analytical and calibration uncertainty. These uncertainties will be explicitly shown in Fig. 2a and described in the Methods and figure caption.
This original Fig. 2a has been moved to the Supplement and renumbered as Fig. S3. Uncertainty envelopes are now included in the plot and described in the caption.

Line 146: It seems likely that the intended method was inductively coupled plasma–optical emission spectrometry (ICP OES), rather than "spectrophotometry".

Changed to spectroscopy. Line 194

Line 166: I think the intention was to refer to the excluded samples when using the term "the remaining samples." However, lines 166–169 read as if this refers to the samples actually used in the dataset. Please replace "remaining" with "excluded" to avoid confusion.

We have revised the text on Line 168 to clarify that the samples referred to are those retained in the final Mg/Ca dataset, and that sample exclusion was based on the identification of anomalously high co-occurring Fe/Ca–Mg/Ca values rather than the application of a fixed Fe/Mg threshold. We also clarify that the Fe/Mg values cited (<0.03 mol mol^{-1}) represent typical values reported in previous studies rather than exclusion criteria. The revised text now explicitly distinguishes between excluded and retained samples, resolving the ambiguity noted by the Reviewer.
Lines 219–225

Line 168: Please remove the "<" before 0.03 mol/mol^{-1} or instead use "< 0.1 mol/mol^{-1}," as Barker et al. (2003) only rejected samples with Fe/Mg < 0.1 mol/mol^{-1}, whereas the 0.03 mol/mol^{-1} value cited here represents the typical range of their samples rather than an exclusion criterion.

See response for Line 166 above.
As above

Line 171: units are missing.

Units have now been included.
Lines 227–228

Line 172-177: The reasoning in this paragraph is difficult to follow. The use of "rather" and "more likely" is confusing. "Ferromanganese oxide overgrowths" and "authigenic Mn–Fe-oxide coatings" refer to the same type of contamination.

The paragraph has been revised for clarity, consistency, and quantification. We now use a single term, "authigenic Mn–Fe-oxide coatings," to describe the observed Mn–Fe overgrowths, avoiding confusion caused by multiple phrases. The revised text also clarifies the interpretation of elevated Mn/Ca values and their implications for Mg/Ca.
Lines 228-237

Line 174: What is meant by elevated? Please quantify.

The term "elevated" is now quantified by explicitly providing Mn/Ca ranges for both *U. peregrina* (0.01–0.28 mmol mol^{-1}) and *G. affinis* (0.04–1.88 mmol mol^{-1}), with anomalous specimens noted. Mn/Ca values exceeding ~0.1 mmol mol^{-1} are discussed in the context of potential authigenic Mn–Fe–oxide coatings.
Line 231–233

Line 175: Please rephrase "this contamination" as not all Mn in foraminiferal calcite necessarily indicate contamination. It could also reflect (partly) natural incorporation of Mn during biomineralisation.

We clarify that higher Mn/Ca values do not necessarily reflect contamination alone; species-specific Mn incorporation during biomineralisation may also contribute (van Dijk et al., 2025). The absence of a positive correlation between Mn/Ca and Mg/Ca further supports that Mg/Ca variability is not significantly influenced by Mn–Fe oxides.
Lines 234–237

Line 191: This phrasing is ambiguous. Do the authors mean a 0.32 mmol/mol correction for $Mg/Mn_{coating}$ within formula 1, or a constant correction for all foraminiferal Mg/Ca data?

We have revised the text to explicitly state that 0.32 mmol mol^{-1} refers to the $Mg/Mn_{coating}$ term used in Equation 1, rather than a constant correction applied directly to Mg/Ca values. This clarification has been added to the manuscript.

Text now reads:
To maintain consistency, we applied an $Mg/Mn_{coating}$ correction following Equation 1, with the $Mg/Mn_{coating}$ term fixed at 0.32 mmol mol^{-1}, and applied uniformly to all Mg/Ca data (Hasenfratz et al., 2017; de Lange et al., 1992) (Fig. S5).
Line 251-252; Fig. S5 has been renumbered Fig. S6

Equation 2: How do the core-top Mg/Ca-derived temperatures compare with modern and/or pre-industrial deep-water temperatures at the study site? Please also justify the use of the Elderfield et al. (2012) calibration (originally introduced in Elderfield et al., 2010). The Elderfield et al., 2010 calibrations lower end is ~1°C. Roberts et al. (2016) specifically adjusted the Elderfield calibration for temperatures below 0 °C.

We apply the linear Mg/Ca–temperature calibration of Elderfield et al. (2012) for *Uvigerina peregrina* because it is species-specific, widely used in deep-sea Mg/Ca studies, and internally consistent with the calibration applied to the majority of records included in our regional stacks. Modern hydrographic data indicate that deep-water temperatures at all Iberian Margin sites examined here are ~2–3 °C (Hodell et al., 2014; Skinner et al., 2003). Consistent with these measured values, the most recent (core-top) portion of the Mg/Ca-derived temperature records shown in Figs. 2 and 3 yields temperatures of ~2.5 °C. As these values lie well within the calibrated range of the Elderfield et al. (2010, 2012) relationship, its application does not require extrapolation to sub-zero temperatures (Roberts et al., 2016). We therefore retain the Elderfield et al. (2012) calibration and clarify this in the Methods section.
Lines 277–279

Line 225: Is this uncertainty based on replicate measurements of the same standard material, or of multiple aliquots of the same sample?

The reported uncertainty is based on a pooled replicate error calculated from replicate analyses across multiple samples, rather than from repeat measurements of a single standard or multiple aliquots of the same sample. In this study, the pooled uncertainty is derived from 47 sets of replicate measurements and therefore reflects analytical reproducibility across the dataset as a whole.
Lines 197–200

Figure 4: Is there a reason why some interglacials and glacials are missing their labels in panel (a)? For example, MIS 13, 14, 17, and 18 are not indicated. Please clarify or add the missing MIS numbers for consistency.

We will ensure all MIS labels are added to this figure in the revised manuscript.
Done

Line 283: Please ensure consistent use of time units throughout the manuscript. Earlier, “Ma/ka” was used for specific ages or time periods in the past, whereas “Myr/kyr” was used for durations. Here, “1.5 Myrs” is inconsistent with that convention. The same issue appears in line 285. Please standardise the terminology.

In both cases, we have revised the terminology to ‘over the past 1.5 Myr,’ replacing the previous usage of ‘Ma’.
Corrected

Line 284 - 285: This statement is difficult to follow. Is the observed cooling occurring during interglacials, glacials, or consistently across both climate states? Are uncertainties associated with this trend quantified? How are these numbers calculated?

In the revised manuscript, we will explicitly state whether the inferred cooling occurs during glacial, interglacial, or both climate states, and we will quantify the associated uncertainties. Cooling magnitudes at Sites 607 and U1385 will be recalculated using the revised datasets and described explicitly in the text, including the method used to derive these values.
The text has been revised to clarify the nature of the observed cooling and to improve readability. We note that these values represent general trends inferred from the timeseries rather than formally separated glacial and interglacial estimates and are intended to provide a first-order comparison between sites. We refer the reader to Ford et al. (2016) for a more quantitative treatment of long-term cooling trends for Site 607.
Lines 377–378

Line 313-316: The comparison is difficult to follow, as it mixes different quantities: interbasin glacial–interglacial temperature differences for 1.5–0.9 Ma with glacial-only temperatures before and after 0.9 Ma. This makes the intended conclusion unclear. Please clarify the rationale for comparing these metrics and include uncertainties.

In the revised manuscript, we will ensure that like quantities are compared consistently, so that the intended conclusion is explicit. Uncertainties associated with these comparisons will be quantified and reported based on the revised data analyses.
The revised text now explicitly compares the same quantities when describing basinal temperature changes before and after 930 ka and provides values with their uncertainties.
Lines 411–413

Line 319-321: If the authors already write "We interpret the greater increase …," this should strictly not appear in the Results section.

We will move this interpretative statement from the Results to the Discussion in the revised manuscript.
Done

Figure 9: The purpose of the dashed lines in Figure 9 is unclear. In panel 9b, it appears that the upper line might represent the mean interglacial $\delta^{18}O_{sw}$ across all interglacials and the lower line the maximum values—but this is not explained. Were these lines calculated from the data, or were they arbitrarily chosen? Similarly, in panel 9c, it is not evident how the dashed line is intended to represent "constant deep-ocean temperatures (line 343)." It appears to be placed at ~0 °C rather than reflecting a calculated mean. If this line is meant to represent an average, please show the associated variability (e.g., 2 SD).

The dashed lines in Figure 9 will be clearly defined on the revised Figure 9.
In the revised manuscript the former Fig. 9 has been renumbered Fig. 8, in the caption to this figure dashed lines are now clearly defined

Line 351-353: Please add numbers for the Temperature decrease

We have added values for the decreasing temperatures and added more detail to the text.
Line 468

Line 357-360: There may also be an important interbasin difference between the western and eastern North Atlantic, which challenges the assumption that Iberian Margin records can be treated as representative of the entire North Atlantic basin (e.g., Chalk et al., 2019). This aspect should be discussed, especially in relation to the statement in lines **287–290** that the carbonate ion effect cannot fully explain the temperature differences between Sites 607 and U1385. If the carbonate ion effect is not the sole driver of interbasin- and species Mg/Ca temperature offsets, the authors should explicitly consider and discuss alternative explanations, including east–west basin-scale hydrographic differences, which are currently not addressed anywhere in the manuscript.

We agree that potential east–west differences within the North Atlantic warrant more explicit discussion and represent an important limitation of treating the Iberian Margin as representative of the entire basin. In the revised manuscript, we will expand the discussion to address evidence for hydrographic differences between the western and eastern North Atlantic (e.g. Chalk et al., 2019, Hines 2024) and their implications for interpreting Mg/Ca-derived temperatures.

Furthermore, we will clarify that while the carbonate ion effect may contribute to the temperature offset between Sites U1385 and 607, it is unlikely to be the sole driver. We will therefore explicitly consider alternative explanations, including basin-scale hydrographic structure and differences in water-mass provenance. This will allow us to more clearly articulate both the assumptions and limitations associated with extrapolating Iberian Margin records to the wider North Atlantic.
We have pointed out that east–west differences within the North Atlantic may contribute to the observed offsets between Sites U1385 (east) and 607 (west). Combining these two records into a North Atlantic stack provide an average of hydrographic conditions in the wider North Atlantic to compare with the Pacific stack.
Lines 375–386

Section 4.1: I would appreciate it if this entire section were given more structure. It is difficult to follow the connections between the different hypotheses and literature references and to identify the main message of the section. Throughout the paragraph, several mechanisms are mentioned (e.g., standing volume effect, NADW cooling, SCW salinity increase, ice sheet/shelf melting, and a stratification inversion in the North Atlantic), but they are not sufficiently integrated into a coherent narrative or overarching context in my opinion. For example, it is not clear to me whether the authors are arguing that the North Atlantic was bathed in NCW, in SCW, or favouring some other scenario.

Section 4.1 will be reorganised to explicitly separate competing circulation scenarios, outline their expected signatures, and state which interpretation is best supported by our data. In the revised section, we will more explicitly integrate the proposed mechanisms (e.g., standing volume effect, NADW cooling, SCW salinity changes, ice sheet and ice-shelf meltwater fluxes, and North Atlantic stratification changes) within a coherent interpretative framework. We will also make explicit which circulation scenario—whether a greater influence of NCW or SCW—we consider most consistent with our results, and clearly state the assumptions and limitations associated with this interpretation.
Section 4.1 has been substantially revised providing greater structure in the context of the discussion on the potential mechanisms contributing to the observed changes in deep ocean temperature and salinity. We now provide more weight to the argument favouring the scenario of a greater SCW influence on the deep North Atlantic during glacial periods after the MPT but acknowledge alternative interpretations of Nd and carbon isotopes.
Lines 470 -490

Line 400: Please remove strong, as it is relative.

Removed

Figure 11c: Please ensure that your own data points are clearly visible. They are currently obscured by the Grimmer et al. (2025) dataset, which prevents a proper comparison between the two records.

We will ensure our data points are visible in the revised figure.
Done

Line 425: The period mentioned (2.75–2.15 Ma, according to the time frame defined in line 38) is not covered by the dataset. Please rephrase.

We recognise that the phrasing in Line 425 may have caused confusion regarding the temporal coverage of the dataset. To avoid ambiguity, we will rephrase this sentence to clarify that the interval discussed spans the onset of the MPT, rather than a period preceding it.
We have rephrased the sentence to remove the ambiguity
Line 574-576

Line 427-429: Please clarify which MDOT dataset you are referring to and indicate the corresponding figure.

Done
We refer to Fig. 10d showing our volume weighted MDOT record.
Line 576

Line 431: See comment on line 400

Done

Line 431-433: I would strongly recommend being more cautious with the argument that the records from Sites 607 and U1313 are complicated by multi-species measurements, since the record presented in this study (U1385) is also a derived from three different species.

We have revised the text to avoid implying that multi-species measurements uniquely complicate the records from Sites 607 and U1313. While the Site U1385 record is derived from two benthic species rather than three, we have removed the phrase "…complicated by multi-species measurements…" to ensure a more balanced and consistent treatment of multi-species records across sites.
Removed

Figure 12: Please indicate uncertainties

Figure 12 will be revised following completion of the new calibration analyses and will show uncertainties in the final revised version. Please see the response to anonymous Reviewer 1 point 3 in which a provisional Figure 12 is provided.
The former Fig. 12 has now been separated in the revised manuscript into Figs. 10 and 11 to facilitate the discussion and enable separate comparisons with MDOT and mean $\delta^{18}O_{seewater}$.

Line 453-456: Are there any similar approaches applied to more recent, data-rich intervals? A comparison would help assess uncertainties.

Similar tandem Mg/Ca–benthic $\delta^{18}O$ approaches have indeed been applied to a more recent, data-rich interval over the late Pleistocene and Last Glacial Maximum (Martin et al., 2002). Comparison with this study may demonstrate that increased site density improves constraints on both temperature and $\delta^{18}O_{seawater}$ and facilitate a more robust assessment of regional versus global signals. We will now explicitly reference this example–and any others obtained through a deeper literature search–in the discussion to place our results in the context of better-constrained intervals and to clarify that the primary limitation of our study arises from sparse spatial coverage rather than from the tandem proxy approach itself.
We agree that comparisons with more data-rich intervals are valuable for assessing uncertainties. While a full quantitative comparison is beyond the scope of this revision, we clarify that the primary limitation of our reconstruction arises from sparse spatial coverage and acknowledge that better-sampled intervals with higher site density would improve constraints.
See 4.5 Limitations and Future Opportunities
Lines 634–651

Conclusion: The conclusion includes a substantial amount of important and relevant literature. I would appreciate it if these key studies were introduced in more detail in the Introduction, allowing the conclusion to focus more clearly on the study's own findings.

As addressed in our revision of the Introduction (see Major Comment 4), key studies will now be introduced in more detail earlier in the manuscript. This allows the Conclusion to focus more clearly on the primary results derived from our stacked temperature and $\delta^{18}O_{seawater}$ reconstructions.
As suggested, we have substantially revised the introduction to include key studies (see Major Comment 4) that were formerly in the conclusion.
The Conclusion is now focussed more clearly on the primary results derived from our stacked temperature and $\delta^{18}O_{seawater}$ reconstructions.

Line 465-467: Most of the studies cited in this context examine glacial–interglacial changes—often limited to the last glacial cycle—and do **not** address variability across the MPT. It is therefore unclear how these references support the conclusions drawn here or how they relate to the long-term patterns discussed in the study.

In the revised manuscript, we will prioritise references that explicitly address long-term variability across the MPT when drawing conclusions. Studies focused on glacial–interglacial variability over shorter timescales will be more clearly contextualised or removed where they do not directly support the long-term patterns discussed.
Some recently published studies across the MPT are highly relevant to this paper and are now cited.