Reviewer 1:

"The Marine Isotope Stage 7: A relic of the 41-ka world?" by Legrain et al provides a valuable compilation and summary of sea surface temperature (SST) spatio-temporal variability during MIS 7, which after revision would make an excellent contribution to Climate of the Past. The comparison with the Clark et al (2024) SST compilation is particularly interesting. Below are some details I would like to see clarified in the manuscript before publication.

We thank Reviewer 1 for their positive assessment of the study.

Major comments:

Methods:

Line 151: The LR04 benthic d18O curve is affected both by ice volume and deep water temperature. Due to a lag between temperature and ice volume, benthic d18O is not synchronous with ice volume change. The Spratt & Lisiecki (2016) sea level stack or the new Clark et al sea level reconstruction would be more suitable to use for ice-volume correction.

We thank Reviewer 1 for the careful attention on SST calibration and agree with the idea of using more recent information on ice volume. However, the ice volume correction included in the BAYFOX calibration requires a $\delta^{18}O_{ice}$ value (estimate of global ice volume change in $\delta^{18}O$ for seawater) rather than a sea-level value (see Appendix B in Malevich et al., 2019). We therefore decided to retain the LR04 value as the reference value for correcting the ice volume for $\delta^{18}O_p$ calibration.

Lines 241-245: Similarity between the results of "selected" and "all" records is not a particularly good argument in favor of using the selected records compilations because the set of all records is acknowledged to be impacted by bias, particularly from the fact that many planktonic d18O records may be impacted by salinity changes. The compilation that excludes all planktonic d18O records produces a noticeably sharper and warmer peak for MIS 7e, especially in the SH stack. Is there anything particular about the spatial distribution of records that have been removed that makes you consider the d18O_p-excluded compilation inaccurate? Would using the smaller compilation alter any of your main conclusions?

We thank Reviewer 1 for their comment. We agree that the similarity between the "all" and "selected" stacks is not a sufficient argument to justify the use of the selected compilation. The records excluded from the "selected" compilation were removed because they exhibit trends that are markedly inconsistent with regional patterns. From this perspective, we consider their exclusion justified. The similarity between the two approaches is noted to underscore the minimal influence of selecting one stack rather than the other. Regarding the compilation excluding all $d^{18}O_p$-derived SST records, we acknowledge that it produces a sharper and warmer peak for MIS 7e in the Southern Hemisphere stack. This difference is primarily related to the strong reduction in data coverage in the South Atlantic. When $\delta^{18}O_p$-derived SST records are excluded, the number of annual SST records in this region decreases from 15 to only 5. The resulting stack is therefore based on a much more limited spatial sampling, which amplifies the influence of a few individual records and likely explains the sharper and warmer peak observed in the Southern Hemisphere. Using the smaller compilation excluding $\delta^{18}O_p$ records would not alter our main conclusions regarding the relative structure, timing, and large-scale features of MIS 7e. The differences mainly affect the amplitude and sharpness of the MIS 7e peak in specific regions, rather than the overall MIS 7 evolution or inter-basin phasing. To ensure full transparency and facilitate broader use of the dataset, we will make all three stacking approaches ("all", "selected", and $\delta^{18}O_p$-excluded) available in the final publication and associated database. This approach will allow readers to explicitly assess the sensitivity of the results to the inclusion or exclusion of specific proxy types and to select the compilation most appropriate for their intended application. A similar strategy was adopted in our recent MIS 9 synthesis (Stevenard et al., 2025), where the three different stacking approaches were also provided.

We have revised the manuscript to clarify these points and to better justify the rationale for retaining the "selected" compilation as the primary reference stack. Now read in the revised manuscript, lines 268-275:

(…) Excluding all $\delta^{18}O_p$-based SST records (2) nearly halves the number of available records and reduces the spatial coverage of the synthesis prior to stacking, particularly in the South Atlantic sector of the Southern Hemisphere. The "selected" stack (3) represents a compromise approach. Records were excluded only when they displayed trends markedly inconsistent with regional patterns, while most of the available data were retained. The similarity between the "all records" (1) and "selected" stacks indicates that the main large-scale features, structure, and timing of MIS 7 are robust to the choice of compilation. We therefore use the "selected" stack as our primary reference for description and interpretation, while making all three stacking approaches ("all", "selected", and $\delta^{18}O_p$-excluded) available to allow explicit assessment of the sensitivity of the results to proxy selection. (…)

Chronologies:

To what extent is the relative timing of NH SST change and SH SST change set directly by the Greenland and Antarctic alignment targets? To help readers evaluate this, please add the GLT_syn time series to right panel of Fig. 1 and summarize the assumptions inherent in its construction. Are similar regional shifts in SST timing observed when SST and benthic d18O are compared in individual cores? (The timing of benthic d18O changes can vary by 2-4 kyr during terminations also, which would be useful to also mention.)

We acknowledge that the chronological framework is anchored to Greenland and Antarctic ice core records, which necessarily constrains the timing of major climatic transitions between hemispheres. However, the alignment tie-points are defined at large-scale transitions and do not prescribe the internal evolution of SST within the interglacial interval. The timing of peak SST conditions in each hemisphere therefore represents the marine records themselves rather than being directly imposed by the targets. Importantly, age-depth uncertainties are propagated using a Bayesian model. The inferred NH–SH phasing is thus evaluated across a distribution of plausible chronologies, ensuring that the interhemispheric offset is not an artefact of a single deterministic alignment. This alignment strategy is based on the working assumption of near-synchronous high-latitude air–sea temperature changes, consistent with previous interglacial syntheses (e.g. Capron et al., 2014, Hoffman et al., 2027; Stevenard et al., 2025).

Are cores from the North Pacific aligned to the synthetic Greenland record? If so, the authors should discuss the degree to which North Pacific SST is expected to be synchronous with Greenland air temperature. For example, do they vary synchronously with one another during T1 and the Holocene?

We thank Reviewer 1 for highlighting this important point regarding the alignment strategy, which was insufficiently documented in the previous version of the manuscript. A dedicated paragraph has now been added to the Supplementary Materials to clearly describe the alignment targets for each oceanic basin. Now read in the revised Supplementary Materials, lines 133-143:

(…) Chronologies were established by aligning oceanic basins onto ice core and speleothem archives, that differs among the considered basins. In the North Atlantic and Mediterranean basin, the records are aligned to the Greenland Synthetic Curve (Barker et al., 2011). In the North and Equatorial Pacific basin, the alignment procedure differs slightly between sub-basins. The North- and Equatorial-East Pacific records are aligned to the synthetic Greenland temperature record (GLT_syn). In contrast, the North- and Equatorial-West Pacific records are aligned to a reference record, itself aligned to the Sanbao speleothem record (Cheng et al., 2016). In the South Atlantic, South Pacific and Indian Ocean basins, the records are aligned onto reference record, themself aligned to the EPICA Dome C ice isotopes records (Jouzel et al., 2007). Finally, the Caribbean records are aligned through a stepwise procedure. Core MD02-2575 was first aligned to U1385 using benthic $\delta^{18}O$, as direct alignment of SST signals with North Atlantic SST records was not feasible due to large discrepancies between the records. Subsequently, ODP-999 and ODP-1002 were aligned to MD02-2575 using both Mg/Ca and $\delta^{18}O$ SST records. (…)

For Termination I and the Holocene, some studies suggest broadly coherent millennial-scale variability between the North Pacific and Greenland (Praetorius et al., 2014). The relationship at sub-millennial scale remains more elusive. However, our strategy of alignment do not assume perfect synchronicity and rather relies on the reproducibility of major transitions and large-scale variability patterns. To our knowledge, there are no dedicated studies specifically assessing the phase relationship between North Pacific variability and Greenland air temperature over MIS 7. However, based on our observations for both stages (MIS 7 and 9), the North-East Pacific SST records display variability that broadly mirrors the large-scale structure seen in GLT_syn, within the limits imposed by marine sediment core resolution and chronological uncertainty.

Fig. S4 – Alignments to core U1429: I have concerns that U1429 is used as an alignment target for other cores given that it has a large gap immediately before (during?) TIII. Could the authors align those cores to a different core? Also, Fig S4 suggests that the age model for U1429 is based on alignment to Sanbao, which isn't mentioned as an age model target in the main text.

We thank Reviewer 1 for pointing out this omission. We have now mentioned the Sanbao $\delta^{18}O$ stack as an age model target in the main text. Now read in the revised manuscript, lines 183-188:

(…) At Site U1429 (the reference for Northwest Pacific), we synchronized the planktic $\delta^{18}O$_notched record of G. ruber, which is the $\delta^{18}O$ signal with eccentricity- and obliquity-band components removed (Clemens et al., 2018) interpreted as a proxy for East Asian monsoon variability, with the $\delta^{18}O_{calcite}$ stack from the Sanbao Cave (Cheng et al., 2016). The Sanbao $\delta^{18}O_{calcite}$ series was already been placed on the AICC2023 timescale by linear interpolation between the tie points defined by Bouchet et al. (2023). (…)

Regarding the data gap around ~245–248 ka, no tie points are defined within this interval. Thus, this gap does not affect the alignment strategy, as tie points are placed outside this interval and the alignment is constrained by surrounding sections.

Interpretation:

Line 446: Is the time resolution of the data sufficient to evaluate the SST response to a 2-kyr CO2 overshoot? Low temporal resolution in many records and the smoothing effect of bioturbation may prevent you from observing a short-lived warming response to the CO2 overshoot.

We agree that limited temporal resolution and bioturbation may attenuate short-lived (~2-4 kyr) signals in some marine records. However, in the highest-resolution SST records (<1 kyr resolution), we do not observe a distinct temperature overshoot synchronous with the CO_2 peak (Figs. S1, S2 and S3). While some smoothing cannot be excluded, the absence of a clear warming excursion in the best-resolved records suggests that a pronounced and widespread SST overshoot is unlikely.

Lines 580-585: The manuscript discusses whether the mode of glacial climate dynamics shifted during MIS 7 due to the particularly strong obliquity forcing at the time. One way to evaluate whether the double peak in MIS 7 actually implies a change in dynamics is to analyze the outputs of simple models that fit the pattern of glacial cycles using a single set of equations and parameter values for the entire late Pleistocene. For example, is a double peak for MIS 7 but not other glacial cycles predicted by the Parrenin & Paillard (2012) model or others? It might also be useful to compare the performance of models that are forced by 65N summer insolation with those that separately optimize sensitivity to precession and obliquity orbital parameters.

We thank Reviewer 1 for this suggestion. Literature provides mixed results on this question. Some conceptual models using a single set of equations and parameters for the late Pleistocene (e.g., Leloup et al., 2021; Imbrie et al., 2011) are able to generate complex or multi-peaked interglacials under particular orbital configurations, suggesting that MIS 7-like double peaks can arise without a fundamental shift in glacial dynamics. However, other studies (e.g. Parrenin and Paillard, 2012, Legrain et al., 2023) find MIS 7 more difficult to reproduce and emphasize its strong obliquity forcing as potentially requiring enhanced sensitivity to obliquity or modified feedback strengths. We understand that the comment points toward a model–data comparison. However, such an analysis would require a dedicated framework and a more detailed treatment of the model outputs, which is beyond the scope of the present study.

Minor comments:

Line 206: Rather than "inverse" (which often means 1/x), it would be better to say "multiplied by -1"

We applied the suggestion of Reviewer 1.

Line 256: awkward phrase – “display a very pronounced regional variability of the MIS 7”

Thanks for pointing this out. Now read in the revised manuscript, lines 286-289:

(…) *The regional ΔSST stacks (Fig. 4e) of the different oceanic basins ~~display a very pronounced regional variability of the MIS 7~~ show marked differences in the timing and magnitude of MIS 7 at regional scale: their structure ~~it~~ can vary from a structure associated to three temperature peaks clearly identified as the warm substages MIS 7e, 7c and 7a (e.g. North Atlantic) to a weakly amplified signal (e.g. Equatorial Pacific).* (…)

Line 258: I think “low amplitude” would be clearer than “weakly amplified”

We applied the suggestion of Reviewer 1.

Line 307-308: “North of 23N” or “23-90N” would be clearer than “North to 23N” (same for south)

We applied the suggestion of Reviewer 1.

Line 309: This should be “Southern Hemisphere stack”

Thanks for pointing this out, we modified accordingly the manuscript.

Line 564-565: Missing words? “a global synthesis evidence warmer deep ocean temperature during MIS 7C, similar to what observed during Holocene”

Thanks for pointing this out. We modified the sentence accordingly. Now read in the revised manuscript, lines 617-619:

(…) *However, a global synthesis indicates that deep-ocean temperatures were warmer during MIS 7C, comparable to those observed during the Holocene.* (…)

Line 565: Insert a paragraph break before “Regarding termination mechanisms…”

We applied the suggestion of Reviewer 1.

Line 614-624: English usage needs revision

We rephrased the paragraph that now reads (lines 670-679):

(…) *In the Clark et al. (2024) synthesis, all terminations of the past 500 ka registered an early warming in the Southern Hemisphere compared to the Northern Hemisphere (Fig. 9b). In contrast, our results for TIIIa, which leads into MIS 7c, are marked by a near-synchronous SST rise across all latitudes, with no detectable interhemispheric phasing at the onset of warming (Fig. 9b). This globally uniform warming signal differs from the youngest terminations and suggests the involvement of other underlying dynamics, such as the role of the Agulhas Leakage (Stevenard et al., 2025; Fig. 9b). However, the analysis of recently published hemispheric SST stacks spanning the entire Pleistocene (Clark et al., 2024) reveals that a warming sequence characterized by an early Southern Hemisphere warming also appears to be a dominant feature of the 41-ka world. Some exceptions may occur during MIS 53 and MIS 41 (Fig. 9a), but chronological uncertainties for this interval remain large, and the analyses rely on only a small number of records (Clark et al., 2024).* (…)