Response to Reviewer 3's comment:

In this study, the authors report a multiproxy dataset analyzed on two sediment cores from the shelf slope and outer shelf of the Beaufort Sea. These records span the Holocene, and indicate that—similar to other proxy-based reconstructions in other sectors of the Arctic Ocean—the sea-ice extent in the Beaufort Sea has been increasing since the early Holocene.

The manuscript is well-written, and the dataset is a welcome addition to the literature. I applaud the authors for producing such a diverse dataset. The main conclusion about the Holocene trend in sea ice is supported by the IP25 and PIP25 data. However, I am not fully convinced yet by some of the discussion based on individual or few data points, and by some interpretations of the biomarker data. In addition, I feel that the organization of the figures and parts of the discussion can be further improved. I hope my suggestions and comments below will be helpful in revising the manuscript to improve its clarity and accessibility. Overall, I recommend **moderate to major revisions**.

We thank the reviewer for their time and constructive comments. We answer the comments and indicate the planned revisions below in blue.

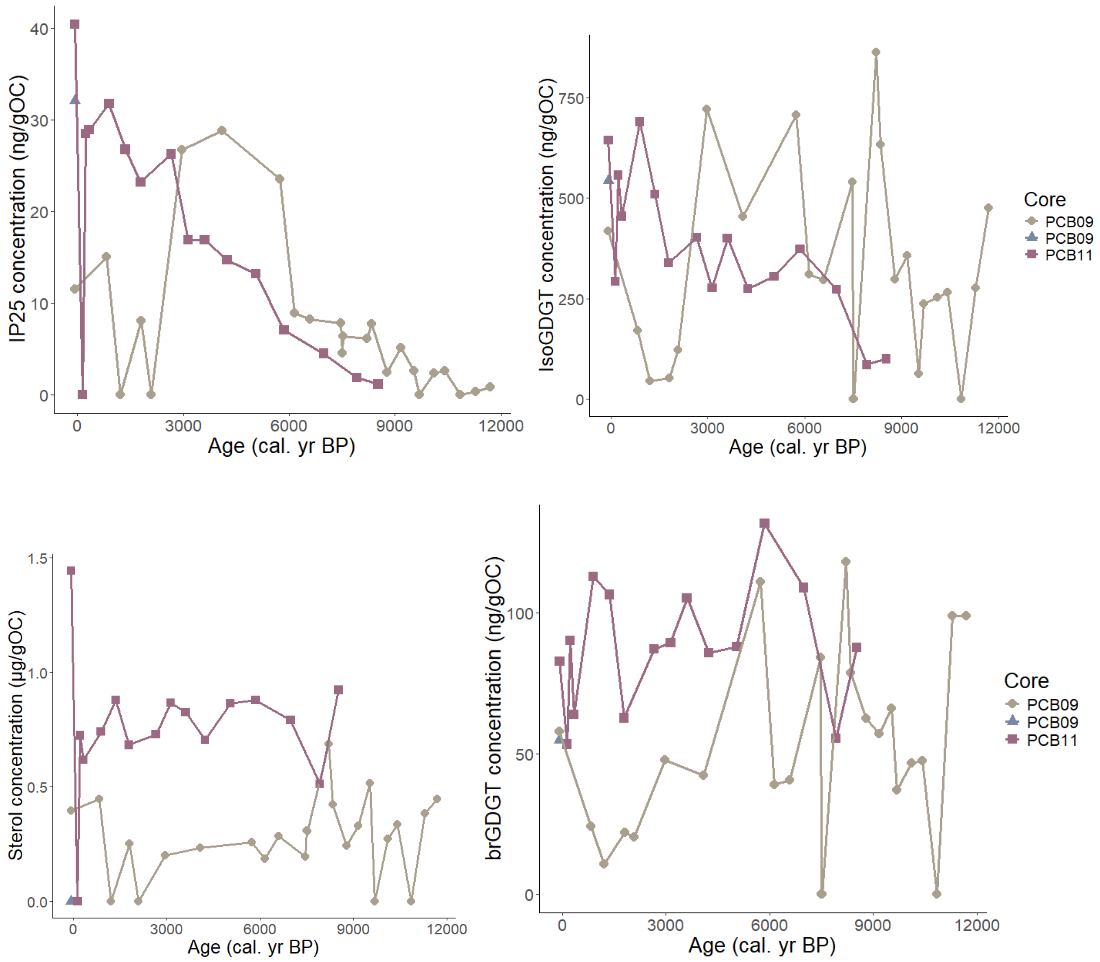
General comments

[1] Interpretation of biomarker abundance

The biomarker data are presented as concentration (ng/g sediment extracted). In a region where large shifts in sedimentation regime occur, the concentration may reflect not only biomarker production but also dilution by clastic material and oxic degradation. For example, is it possible that the downcore decrease in almost all biomarkers in Figure 3 reflects oxic degradation post burial? TOC also shows such a downcore decrease; normalizing the biomarker data using TOC, as some workers prefer, would probably eliminate the Holocene "increase" in other biomarker data (e.g., HBIs, isoGDGTs, and OH GDGTs). I understand that it is not always possible to fully tease apart these factors, but I encourage the authors to discuss this possibility and, if applicable, caveat their discussion to acknowledge the uncertainty of these data.

We appreciate the reviewer's comment regarding the potential influence of dilution by clastic input and post-depositional oxic degradation affecting biomarker concentrations when expressed per gram of sediment. We agree that such effects can, in some settings, obscure original variations in biomarker production, and we acknowledge that normalization to organic carbon (OC) content is one way to reduce the influence of variable mineral dilution. In our cores, however, TOC contents are relatively stable downcore (ranging only from 0.9 to 1.3% for PCB09 and from 1.1 to 1.3% for PCB11 after the initial increase; Figure 3a). Given this limited variation, normalization to TOC would have only a minor effect on the relative biomarker profiles (see figures below for selected biomarkers). Moreover, the different downcore behaviors observed among biomarker groups (Figure 3) suggest that their trends are not governed by a single control such as degradation or dilution, but rather reflect differences in source or preservation processes. We have added a statement in the revised results acknowledging that post-depositional degradation may contribute to some of the observed downcore decrease, but that the overall patterns are unlikely to be an artifact of dilution or TOC normalization. L313-317 "*Although biomarker concentrations expressed per gram of sediment may be influenced by mineral dilution or post-depositional degradation,* *TOC contents in our cores varied only slightly (0.9-1.3 % in PCB09 and 1.1-1.3 % in PCB11; Figure 3a). Such limited variability indicates that TOC normalization would not substantially alter the observed trends. The distinct downcore patterns among lipid biomarker classes therefore likely reflect differences in source input and preservation rather than a uniform effect of degradation or dilution.*

.”.



[2] BrGDGTs as terrestrial indicator

The authors interpret brGDGT concentration as terrestrial input. However, recent studies have proposed in situ production of brGDGTs (e.g., Xiao et al., 2016, Biogeosciences). In particular, based on the IIIa/IIa ratios, Singh et al. (2025, OG) suggested that brGDGTs in the central Arctic are of marine origin and transported from the shelf. Might this explain the differing temporal patterns in the concentrations of brGDGTs and sterols? I suggest that the authors first discuss the origin of brGDGTs in the Beaufort Sea before interpreting brGDGT concentration as terrestrial input.

We agree that brGDGT can be produced in situ in the marine environment as was described also by Peterse et al., 2009, in Svalbard or Sinninghe Damsté et al. (2016) in the Berau Delta (just to cite these two additional studies). As described in the latter, a good indicator for in situ production is the use of #ring$_{tetra}$ (rather than IIIa/IIa) in coastal sediments. In our records, #ring$_{tetra}$ is always below 0.7, which has been defined by Sinninghe Damsté et al. (2016) as the cutoff values for in situ production in most coastal shelves. We added this information L248-251 "*In addition, in situ marine production of brGDGT can occur in coastal sediments between 50 and 300 m water depth (Peterse et al., 2009; Sinninghe Damsté, 2016). To assess the potential for brGDGT to be in situ produced we calculated #ring$_{tetra}$ [Eq 9]*" in the methods and in the results L306-307 "*BIT values varied from 0.1 to 0.4 in PCB09 and from 0.1 and 0.5 in PCB11. #ring$_{tetra}$ values were always < 0.7 for both cores.*".

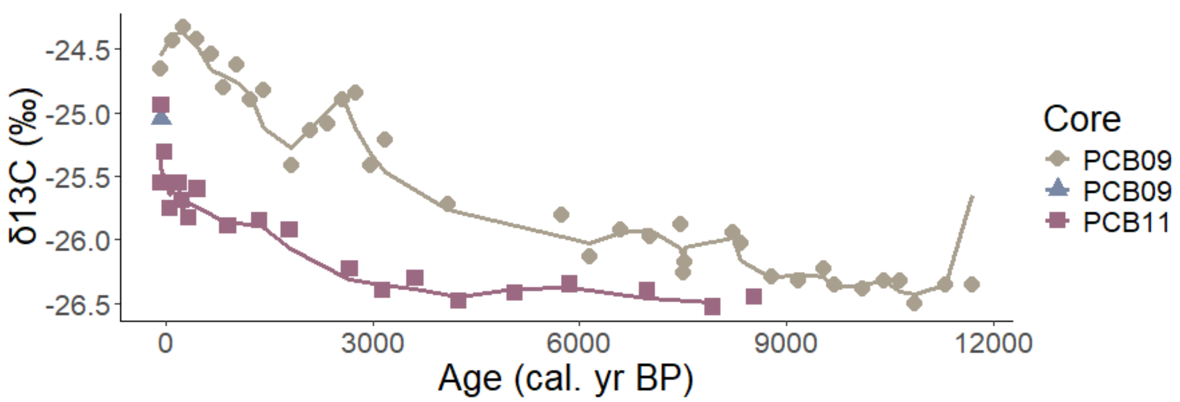
[3] Discussion not supported by the temporal resolution and precision of records

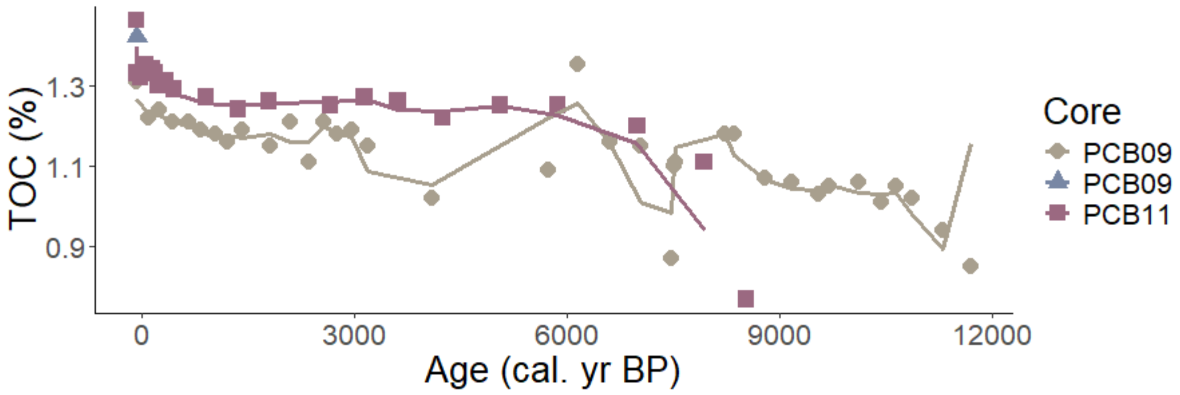
The temporal resolution of proxy records is relatively low, and there are large fluctuations in the data. In some parts of the manuscript, a single data point is interpreted as representing a climate event (e.g., "A sharp decrease in PIP25 at 1.5 ka…"). Given the temporal resolution of the proxy records (e.g., 15 HBI III data points over 9 ka, corresponding to ~600 years between data points) and the overall noise in the data, I am not convinced that the change between two data points (1.5 ka to 0.4 ka) can be robustly linked to the Little Ice Age—especially considering age-model uncertainty.

We rephrased this interpretation in the revised text, see for example L439-440 "*These changes are broadly consistent with the timing of the regional cooling associated with the Little Ice Age (Mann et al., 2009), though the resolution of the biomarker record does not allow precise attribution to centennial-scale events.*".

More generally, proxy records are subject to non-climatic noise from sedimentation and analytical error; unless demonstrated otherwise, interpreting absolute point-by-point values seems overconfident. It would be more robust to interpret overall trends or time-slice means instead of individual fluctuations. This can be achieved by applying a running average, smoothed series, or trend lines for defined intervals (e.g., Early, Middle, and Late Holocene).

We agree with the reviewer and we generally aim to not interpret single data points if they are not part of a general trend. However, when the changes between two data points are particularly large and reflected by several other proxies (and not readily explainable by changes in deposition – sedimentation rate, or degradation – or TOC) we at least discuss the transition without overstating their significance. In addition, we considered presenting running average for this dataset but we believe we have too low resolution for this as this study focused on presenting many different proxies rather than focusing on obtaining high resolution profile for 1 proxy (see example of running average for our dataset with higher resolution, TOC, d13C).

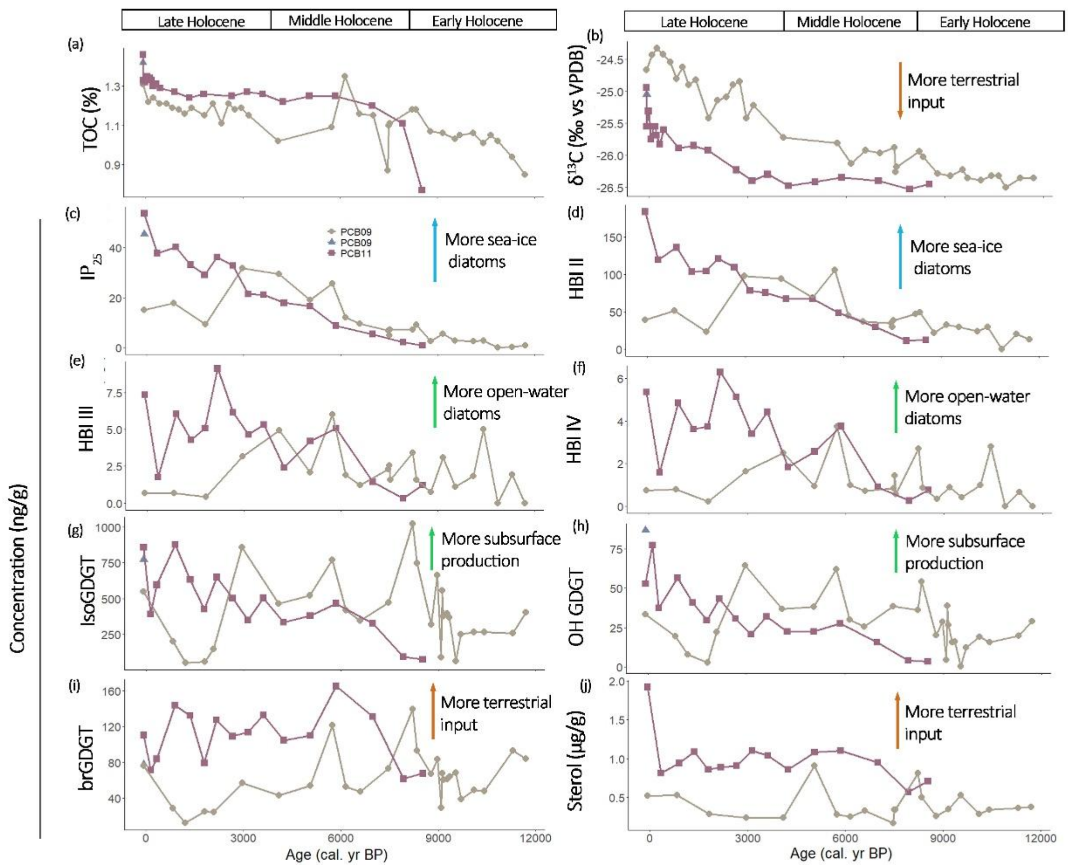




[4] Overall clarity and accessibility

At several points, it was challenging to follow the authors' reasoning, as it was unclear how some conclusions were reached. I've provided line-specific examples below, but the authors may want to check for this throughout. Readability could also improve if the authors indicate biomarker interpretations directly in Figures 3 and 4 (e.g., label "Terrestrial input" next to sterols).

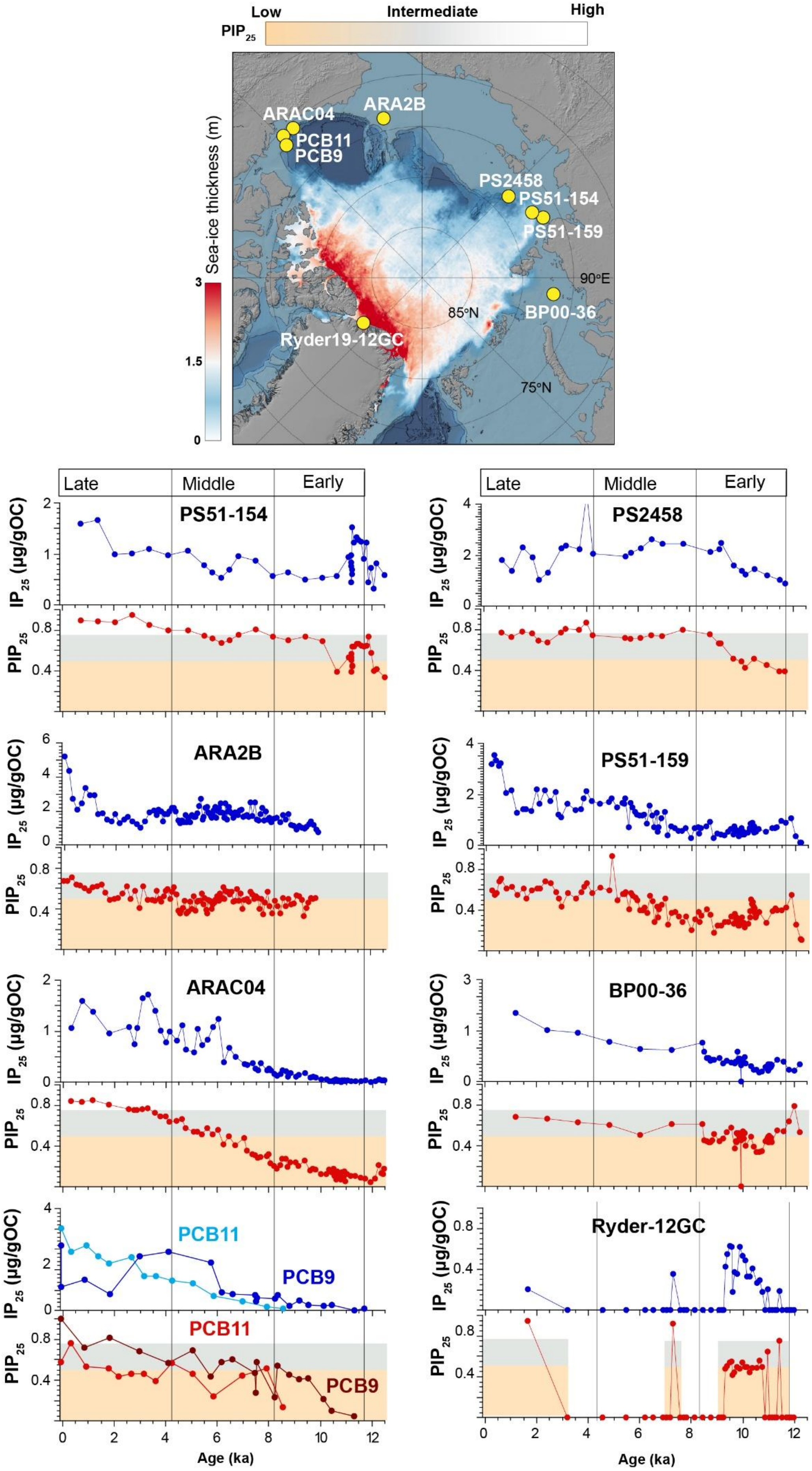
We appreciate the reviewer's suggestions for improving readability as detailed in our responses to the specific comments below. Following a suggestion by another reviewer, we also tried to guide the reader through our complex Figure 3 by adding arrows and interpretation on the figure (see revised version below).



In addition, the age unit is given in years in all figures but in ka in the text—please make this consistent. It would also aid comparison if Figure 4 panels were stacked vertically rather than side-by-side, and if time intervals were shaded in color to highlight different Holocene intervals.

We homogenized the ages to ka in the text and figures.

We partially adapted figure 4 to the reviewer: we aligned the records into 2 columns, and instead of shading different time periods, we added the time periods above the stacked figures (similar to Figure 3).



Low
Intermediate
High
PIP25
Sea-ice thickness (m)
3
1.5
0
ARAC04
ARA2B
PCB11
PCB9
PS2458
PS51-154
PS51-159
90°E
85°N
75°N
BP00-36
Ryder19-12GC
Late
Middle
Early
PS51-154
PS2458
ARA2B
PS51-159
ARAC04
BP00-36
PCB11
PCB9
Ryder-12GC
IP25 (µg/gOC)
PIP25
Age (ka)

Specific comments

L20: "…IP25, and other HBIs,…" → suggested phrasing "HBIs including IP25."'

We revised the text accordingly

L42: "…renewed…" needs context—did the interest lapse previously?

We think that the topic has attracted more attention in the last decade when the absence of sea ice in the Arctic was forecasted. However, we see the point of the reviewer and deleted that term.

L56: Add reference.

The references have been added L56-58 "*Numerous studies on Arctic sea ice variability have focused on offshore locations highlighting heterogeneity in sea-ice cover history and the importance of local currents (Belt et al., 2010; Detlef et al., 2023; Fahl & Stein, 2012; Hörner et al., 2016, 2018; Stein et al., 2017; Stein & Fahl, 2012; Vare et al., 2009; J. Wu et al., 2020).*"

L58: Use "Lipid biomarkers" for consistency with the title and throughout.

The text was changed L70 and throughout.

L62–63: "Several proxies for sea temperature exist using microfossils…" reads awkward; please rephrase.

The sentence was rephrased L74-77 "*In contrast to salinity, several established proxies exist for reconstructing sea temperature, including microfossil assemblages (e.g., dinocyst, Richerol et al., 2008), inorganic ratios (e.g., Mg/Ca of foraminifera, Barrientos et al., 2018; Kristjánsdóttir et al., 2007) and lipid biomarkers (Ruan et al., 2017; Varma et al., 2024).*"

L65–66: The phrase "usually include hydroxylated…" is too strong; OH-GDGT proxies are not yet widely used enough for "usually."

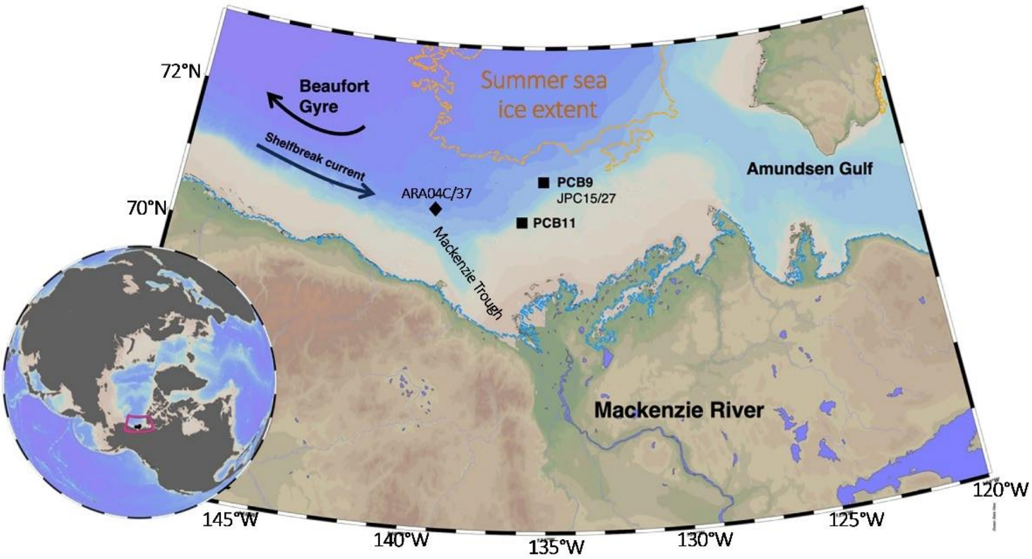
For the case of polar environments, they have been used for the last 10 years (since the discovery of Liu et al., 2012) which is, in our opinion, long enough to warrant the term "usually".

L80: "Material" → "Materials."

This has been changed in the revised text.

L83: Include and label Mackenzie Trough in the map for non-expert readers.

We labeled the Mackenzie Through in the revised Figure 1, see below.



L94: "…an open water flaw leads occur…" — if singular, use "occurs"; if plural, remove "an."

The "s" was moved to now read "an open water flaw lead occurs".

L99–102: Clarify whether multicores were used to generate proxy records. If not, remove.

The top slice of the multicore was used and this will be mentioned in the revised text (L117-119).

L103: "…as in (Matsuoka et al., 2012)." → "…as in Matsuoka et al. (2012)." Check and correct similar formatting errors throughout.

The formatting was checked and corrected throughout.

L110: "extend" → "extent."

This has been changed in the revised text.

L113: "…scanned shipboard on a Geotek…" → reword for clarity.

We disagree with the reviewer as this is common writing for this instrument.

L121–123: Add more explanation of the assumptions underlying these interpretations.

We added more details on the use of these ratios L138-139 "*Zr/Rb was used as a proxy for grain size variations as Zr content is elevated in coarse mineral, while Rb is associated with clay minerals (L. Wu et al., 2020) and Br/Cl as a proxy for marine organic matter as Br usually correlates with OC content (Wang et al., 2019).*"

L132–133: Define "well-preserved"; consider adding SEM images of foraminifera.

We added images taken on a binocular in Figure S6.

L156–157: Rephrase to clarify that 5 g refers to each depth/sample, not the whole study.

We added L174-175 "*For each sample, 5 g of homogenized freeze-dried sediment was extracted using an Energy Dispersive Guided Extraction (EDGE)*"

L165: C46 GTGT

We corrected to L184 "and C_{46} GDGT-like compound"

L170: "Concentration of IP25 were…" → "was."

This has been changed to "*Concentrations of IP_{25} were*" in L191.

L173: The PTFE filter should likely be 0.45 µm, not 45 µm. Please check.

It has been corrected in L194.

L175: "according to Hopmans et al. 2016 and following Lattaud et al. 2021" → clarify which method was actually used.

The Lattaud et al. 2021 method was used, which is based on Hopmans 2016. This was revised in the text L196-197.

L176: C46 GTGT (check notation).

L197-198 "*GDGTs were quantified using the C_{46} GTGT internal standard assuming the same response factor*."

L177: "sililated" → "silylated."

Corrected accordingly.

L178: "C22 5,16" → add "-diol."

Added accordingly.

L180: Define "IRMS" at first mention.

We added the definition of the acronym (L203).

L208–209: Use consistent terminology for "salinity" vs. "S."

We changed it in the equation accordingly.

L211: Define "SST" here, not later (L216).

We defined SST upon first use (now L234).

L218–219: "SST" is italicized in L218 but not L219—be consistent.

We removed the italic formatting throughout.

L227: Only N. pachyderma counts are presented, yet the methods mention all species. Please clarify.

Yes, only *N. pachyderma* is presented in the manuscript as this was the only species we found. However, we looked also for other species that could have been present, so the method description is open.

L241–243: Plot sedimentation rate in Figure 2 for clarity.

We do not think sedimentation rate are needed on top of the age model in Figure 2 and prefer to keep the Figure as is.

L248–254, 257–259: These sentences read more like discussion than results—consider moving or rephrasing.

We removed it here as this is already mentioned in the discussion.

L260–262: Add figure citations so readers can locate corresponding results.

We added the figure citation to this paragraph.

L273–274: See General Comment [3]. It's difficult to locate data at 1 ka; consider adding minor ticks and grid lines.

We added grid lines in Figure 4 and 5.

L285: Consider moving Fig. S3 to the main text (only 5 figures currently).

We think Figure S3 should remain in the supplement as it mainly deals with the age model and the geochronology of the study cores is not the main focus of the paper. We will rather consider Figure S4 which deals with the core-top calibration of two of the most novel proxy for salinity and temperature.

L301–308: Same as above. Alternatively, briefly state in Methods that RI-OH′ was not used because it yielded unrealistic SSTs.

We would like to discuss the data and state that the RI-OH′ yielded unrealistic SST, and therefore prefer to keep it in the results section.

L349–351: The results do not clearly show that SST follows summer insolation. To me, SST declines at the end of the Early Holocene while PIP25 remains low, suggesting weak coupling.

We only claim that SST were elevated at the early Holocene, not that it was tied to insolation during the whole Holocene L391-393: "*During 12 – 8.5 ka, SST are elevated in comparison with the rest of the Holocene (Fig. 4d) which coincided with peak 21 June insolation*"

L363–364: This interpretation assumes Ti is relatively constant—please justify or provide supporting evidence.

Ti counts are stable throughout the core, we added all XRF data into a supplementary Table S3.

L365: “slop” → “slope.”

Corrected accordingly.

L372: “have” → “has” (subject is singular).

Corrected accordingly.

L377–378: The conclusion “stable sea ice-edge or polynya conditions” is not well-supported—please elaborate.

At 7-6 ka PIP25 in PCB9 is above 0.5 indicating marginal sea ice zone conditions. We rephrased L419-422 “*At PCB09, SSTs cooled from ~6 °C to 3 °C (Fig. 4d), while steadily increasing sea-ice biomarker concentrations led to PIP_{25} > 0.5 by 7-6 ka, indicating the establishment of stable ice-edge or polynya conditions at the slope. The higher salinity (Fig. 4b) and greater distance from the coast at this site suggest enhanced influence of offshore Pacific-derived waters and reduced terrestrial input.*”

L386–400: See General Comment [3].

See our earlier response.

L413: “close by” → “close-by”

Corrected accordingly.

L417: “Norther” → “Northern”

Corrected accordingly.

L440: “Conclusion” → “Conclusions”

Corrected accordingly.

L451–453: This claim is not supported by the lack of correlation between the PIP25 and SST records.

We build our claim on the observation of the higher SST during the Holocene in link with the absence of sea-ice cover in all marginal seas.

Overall assessment:

This is a valuable contribution to Arctic paleoceanography, and I commend the authors for assembling such a comprehensive dataset and transparency reporting reporting all results, including proxies that did not yield clear signals. Addressing the points above—particularly regarding biomarker interpretation, temporal resolution, and figure clarity—will greatly improve the robustness and accessibility of the paper.

We appreciate this positive assessment of our study and the constructive criticism. With this revised version, we hope to have satisfactorily addressed the reviewer's concerns.