Answer to Reviewer 2's comment:

Santos et al. present two new multi-proxy paleoceanographic records from the eastern Beaufort Sea. The unique sampling sites and wide range of proxies analyzed have the potential to enhance our understanding of Holocene sea-ice regime in this underexplored region. However, the current version suffers from methodological inconsistencies and a lack of focus in presentation (e.g., trying to do both analysis of site-specific and pan-Arctic trends), which limit its suitability for publication at this stage. While many proxies are analyzed, few are discussed in sufficient depth. The authors are encouraged to narrow their focus to the regional paleoceanography of the Beaufort Sea and to compare their results more thoroughly with nearby cores - especially Wu et al., 2020, which employed a very similar set of proxies. It is also important to critically evaluate the suitability of each proxy for this region. Indices such as RI-OH′ are still relatively new, and their environmental interpretations should be treated with appropriate caution until further validation is available. Below are more detailed suggestions to help strengthen the manuscript.

We thank the reviewer for their thoughtful and constructive comments. We will respond point-by-point below (in blue).

We acknowledge that some methodological details and proxy comparisons require clarification and we will revise the methods section to ensure consistency and transparency.

Concerning the use of recently developed proxies (e.g., RI-OH′, TEX-OH, δ^2H of C_{16}), we agree that their interpretation should be approached with caution. However, given the scarcity of validated Arctic-specific proxies, testing and cross-evaluating these indices remains valuable. We will clarify their potential and limitations more explicitly in the revised discussion.

We appreciate the reviewer's suggestion to narrow the focus of the study. We agree that the regional perspective is central, and we will ensure that the discussion clearly emphasizes the Beaufort Sea record. At the same time, we believe that retaining a concise Arctic-scale overview adds essential context. Regionally, our coastal core provides new insight into the Holocene sea-ice history of the Beaufort Sea, complementing the offshore record of Wu et al. (2020). To our knowledge, this is the first coastal record available for this region. Placing these results within a broader Arctic framework allows us to illustrate how local conditions relate to large-scale climate forcing, such as variations in insolation, ocean circulation, and bathymetry, which together shaped spatial differences in sea-ice cover during the Holocene.

INTRODUCTION

Overall, this introduction is brief, glances through many proxies without properly explaining their mechanism, limitations, and justifying their applicability to your specific study area. It also doesn't establish well the theoretical link between numerous oceanographic conditions and sea ice. The authors fail to identify knowledge gaps in the existing Holocene sea ice literature. If Holocene sea ice records have no regional heterogeneity among them, there is then no need for another paleo sea ice reconstruction. I understand the purpose of the statement in lines 55-56, but this present study does not offer more cores or a more coastal location than other average paleo sea ice studies. Regarding line 68-69 identified a valid gap in regional calibration for GDGT proxies, and this led naturally to the need for calibration from surface sediment samples, which was briefly mentioned in line 74-75, but this

connection is textually hard to see. The authors should put more effort into highlighting this connection and use citations to support their claim.

We amended the introduction L55-58 to highlight the difference in sea-ice history reconstruction in the Arctic "*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; Wu et al., 2020).*".

We disagree with the statement that our study does not offer more core/more coastal location. The strength of this study is to analyse both cores (on the slope and on the shelf) which has only been done in one other location in the Laptev Sea. To add, most of these aforementioned Holocene sea-ice reconstructions from the Arctic have been derived from offshore settings, where records predominantly reflect pack-ice dynamics. However, comparatively few records exist from coastal or inner-shelf environments, where landfast ice and seasonal polynya activity exert dominant controls on sea-ice conditions.

Regarding the proxies, we added L81-83 "*These limitations highlight the need to further develop and test Arctic-specific proxies for both salinity and sea temperature."*

line 47-51: As one of your cores covers the end of the last deglacial, discussing some pre-/early Holocene warming events that you expect to see in your record in chronological order might help streamline your narrative.

We currently describe cooling events in chronological order but not the warm events pre-Holocene, so we amended the text L49-52: "*Throughout the Holocene, Arctic sea ice has responded to changes in orbital forcing, ocean circulation, and ice sheet dynamics (Park et al., 2018; Stein et al., 2017). During the last deglacial, abrupt climatic events such as Bølling-Allerød (~14 – 12.8 ka) and Younger Dryas (~12.8–11.7 ka), contributed to the instability of the Arctic cryosphere. In the Canadian Arctic, the enhanced meltwater discharge and re-routing following the retreat of the Laurentide Ice Sheet (LIS) contributed to oceanographic shifts and transient cooling events (Broecker et al., 1989)."*

line 52-54: The summary of Holocene sea ice trend in line 52-54 might be too simplistic and overlook the regional disparity among sea ice records. The narrow age constraint on the Holocene thermal maxima is questionable. If your claim only describes the Beaufort Sea, please specify that, as it has not been made clear.

We go into more details about spatial heterogeneity of Arctic sea-ice records in the last paragraph of the discussion.

Thanks for pointing out that the timing of the HTM is too narrow, we updated the time period for the western Beaufort Sea. We now followed Kaufman et al. 2004 timing and revised for 11 to 6 ka.

line 55-56: Which studies? Need citation or need to reformulate the sentence.

We added these references L57-58.

line 57-69: The authors presented the current arctic paleoceanography proxy toolbox in a confusing order, while leaving out the key HBI-based proxies and index (IP25, PIP25), which definitely deserve some discussion. There are numerous GDGT-based paleothermometers and indices; the authors should name them directly in the paragraph to avoid confusion. Please consider adjusting the use of "e.g." from line 60 onward, especially for line 63. Regarding line 69 needs citations to support the claim.

We tried to limit the use of "e.g." when citing references when possible. We added one paragraph on sea-ice proxies including IP25 and PIP25 L60-68 *"Sea-ice cover can be reconstructed from microfossil and lipid biomarker evidence preserved in marine sediments. Remains of sea-ice organisms such as dinocysts (de Vernal et al., 2013) and diatoms, the latter producing a specific biomarker known as IP_{25} (Belt et al., 2007), provide valuable records of past sea-ice conditions. This highly branched isoprenoid (HBI) and its isomer HBI diene (HBI-II) are used to trace the presence of spring sea-ice in modern and geological settings. However, because the absence of these two HBIs may reflect either a permanent sea-ice condition (due to the absence of light) or completely sea-ice free waters, the PIP_{25} ratio was developed (Müller et al., 2011). This ratio includes a phytoplankton biomarker (typically dinosterol, brassicasterol or HBI-III) that represents open-water productivity. PIP_{25} values have been used to distinguish between seasonal sea-ice (>0.5) and permanent sea-ice cover (>0.75).*".

For the other proxies. we follow the order stated in L60 "*salinity, sea temperature and freshwater influence*". Specifically, for sea temperature proxies the order is historical with first, the description of microfossil use (here we use e.g. as this is not the focus of the study), the inorganic ratio (again e.g. as this is not the focus), ending the list with biomarkers. We spelled out the names of the GDGT ratios used in SST reconstruction L77-79 "*Among biomarker proxies for cold water (< 15°C) environments, hydroxylated glycerol dialkyl glycerol tetraether (OH-GDGT) are particularly useful, with RI-OH' and TEX-OH identified as promising temperature indices (Lü et al., 2015; Varma et al., 2024)*".

line 70-79: The objectives 2 and 3 are bold statements; some reformulation while keeping the limitation of your proxies in mind is needed.

We rephrased the objectives L90-93 "*The primary objectives are to (1) reconstruct past variations in sea-ice cover on the Beaufort Shelf throughout the Holocene, (2) explore the potential roles of insolation changes, meltwater input, and oceanic conditions in shaping regional sea-ice variability, and (3) place the Beaufort Sea record within a broader Arctic context to provide insights into past and present climate variability.*"

MATERIALS AND METHODS

Overall, the biomarker workflow is unconventional. The authors attempt to analyze too many biomarker classes from a single extract, which may compromise the analytical quality of each individual proxy. The solvent system, choice of standards, and selected m/z values raise concerns about data robustness. Methodologically, the section reads as a workflow appears fragmented and could benefit from clearer structure with uneven detail across proxies. While I appreciate the challenge of condensing complex workflows into a limited space, a clearer structure, such as a summarized workflow diagram or table, would greatly help readers follow the analytical sequence and evaluate reproducibility.

We disagree with the reviewer on the unconventional workflow, these protocols have been used in many other studies (see the interlaboratory comparisons of Belt et al., 2014; Bijl et al., 2025; De Jonge et al., 2024). The solvent system for solid-phase chromatography separation is ideal to separate alkenones from the rest of the lipids (in the DCM fraction). We agree that the choice of standard for the sterol is unusual and will not give a perfect quantification due to their different structure, this is also why we do not focus on absolute concentrations but rather the relative difference within the sterols. The sterols are not used to calculate the PIP25 or other ratios where this could be an issue.

It is unusual to have a table of workflow in such manuscripts, especially for established methods following interlaboratory recommendations (Belt et al., 2012, 2014), (Bijl et al., 2025; De Jonge et al., 2024).

Figure 1: Where are the surface sediment samples? Please consider using a legend for the labeling of sea ice extent. Country names are not necessary; authors should use a brighter color to indicate the study area in the global map, or else the small map will not be very useful.

The surface sediment samples are presented in Figure S1 as they are not part of the main discussion and serve as support for some of the newer proxies used in this study. This is mentioned L117-119 "*The core tops (0-1cm) from 22 multicores collected during PeCaBeau, were used to ground truth the hydrogen isotope ratio of $C_{16:0}$ fatty acid proxy for reconstructing salinity and test the applicability of the temperature reconstructions (Fig. S1).*"

We will remove the country names and change the color for the study area in the revised figure 1 (see revised figure below).

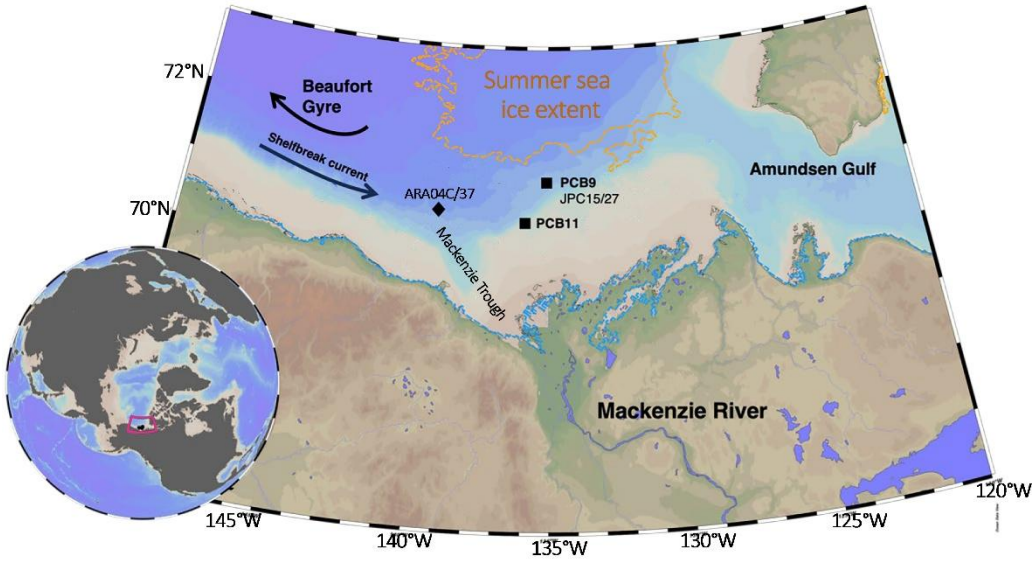

Figure 1

line 156: Could you clarify what is the resolution of your biomarker analysis?

We added one sentence to the revised text to define the number of samples and resolution in the core L173-174: "*Lipid biomarkers were analysed from 42 samples (every 10 cm for the*

first 143 cm, then every 20 cm) for PCB09 and 21 for PCB11 (every 20 cm). Core top samples from the MC's were also analysed for lipid biomarkers.".

line 158: Could you clarify what is the concentration and the composition of the alkaline solution used for saponification?

We will add this detail in the revised text L178 (stated in the reference study Lattaud et al., 2021), we used a KOH in methanol solution at 0.5M.

line 163: Could you explain why internal standards added post-extraction?

We acknowledge that it would have been better to add the internal standard before extraction to assess loss during laboratory work and to assess extraction efficiency. However, most losses occur during workup after extraction so the effect of adding the internal standards after extraction only add a small uncertainty.

line 164: The choice of C22 5,16-diol as an internal standard for sterols is unconventional. As a long-chain diol, its polarity, structure, and chromatographic behavior may differ substantially from sterols, and it is unlikely to mimic sterol recovery or derivatization efficiency.

We agree that is it unconventional for this standard to be used for sterol quantification, however it is still a complex alcohol molecule. Here we do not stress the absolute concentration of the sterols in our records but rather the relative changes.

line 170: HBI IV is an isomer of HBI III with the same degree of unsaturation. Why is it monitored at 348? The authors should include ion monitoring values for the standards as well, since previously there have been a few different fragments monitored for the same standard (For example, 7-HND can be monitored with m/z 99 and 266).

We added the ion monitoring values for the standards to the revised methods, here 266 for 7-HND and 350 for 9-OHD. We agree that HBI IV and HBI III are monitored at the same m/z, this was a writing mistake, which is now corrected in the revised text L192.

line 172: Please provide comparison data in the supplementary material.

We will provide the concentration of the reference material as supplement (Table S2).

line 173: Saponification is not a viable strategy for GDGTs work. Assuming your post-saponification liquid-liquid extraction is between a methanolic KOH and hexane:DCM mix, there is a possibility that quite some GDGTs will be lost in the aqueous phase.

We disagree with the reviewer here. Saponification is a viable strategy for GDGT analysis as was shown in the latest recommendation for handling GDGT (Bijl et al., 2025). It has also been shown to be the ideal method to quantify sterols (Fu et al., 2025).

line 177: This raises concerns and warrants clarification if the sterols would end up in F3, as their polarity is different than GDGT and according to this protocol, would probably mostly elute in F2 (DCM) and potentially partially elute in F1. Also, how is the F3 used both for GDGTs and phytosterols? Is it a split of the fraction? This point please clarify.

Previous protocols for GDGT analysis (Bijl et al., 2025, Lattaud et al., 2021) show that GDGTs end up in fraction 3. No sterol eluted in F1 or F2 which were monitored for other compounds (for another project these fractions were screened for alkanes and alkenones). F3 was split in 2 with one fraction analysed for GDGTs and one fraction ran for sterol analysis. This is added in the revised methods L194.

line 178: What are the m/z ratios, and which sterols are you quantifying?

We used the total ion current for quantification, m/z = 129 and specific ions were used for identification of brassicasterol, stigmasterol, β-sitosterol, campesterol, see revised L202.

line 203: c factor should be reported here.

Initially the c factor was only in the supplementary, we added it to the revised method section L227.

line 202: The statement of not being able to detect dinosterol in the samples is concerning, since dinosterol should be a regionally abundant biomarker. In Wu et al. 2020, the core, which is nearby, the PIP25 index was calculated solely based on dinosterol. The complete absence of dinosterol raises questions about the sterol recovery in this study under review.

Dinosterol has been absent from other Beaufort Sea studies (Belt et al., 2013; Fu et al., 2025) and its absence in our study sites is expected. Instead, its presence in Wu et al. site is surprising and could be due to local currents bringing different amount of nutrients. We now add this information L225"*Dinosterol was not detected in the samples which is common in the Beaufort Sea (Fu et al., 2025),"*

RESULTS

Figure 3: The practice of including brassicasterol as a terrestrial sterol is potentially problematic and warrants reconsideration, even regionally, brassicasterol is primarily of terrestrial origin, but it's still a sterol that has a mixed source.

We agree with the reviewer that brassicasterol needs to be interpreted with caution and can for our study region be used as a primarily terrestrial biomarker. This is indicated L224-225: "*brassicasterol has been shown to derive mainly from terrestrial input in the region* (Wu et al., 2020)."

Figure 3&4: The authors should consider grouping proxies that are reconstructing the same information in the same figure. (such as PIP25 with the HBIs, BIT with terrestrial sterols)

For figures 3 and 4 we prefer to present all biomarker concentrations in figure 3, and the biomarker ratios in figure 4.

DISCUSSION

line 335: The authors claim that the comparison between the slope core and outer shelf core is the focus of this study, but the comparison is weak throughout the discussion. For most of the discussion, the authors seem to treat both cores as one single record without emphasizing the difference in their depositional environments. In the introduction (line 74) and results (line

289), the authors claim that they conducted further calibration work with surface sediment, but these samples weren't included in the map, nor do we see the data presented or discussed anywhere else.

We appreciate the reviewer's suggestion to clarify the comparative aspect between the slope (PCB09) and outer shelf (PCB11) cores. We would like to note that our comparison is constrained by the temporal coverage of the records: the Deglacial–Early Holocene interval is only preserved at PCB09, preventing a direct comparison for this period. However, for the overlapping Middle–Late Holocene interval, we explicitly highlight the spatial contrasts between the two sites, for example: differences in the timing of sea-ice stabilization (PCB09 at 7–6 ka versus PCB11 at 4 ka), the distinct depositional settings (flaw-lead versus stable ice-edge), and associated productivity trends. We rewrote part of the discussion to emphasize this part of the study (see paragraph 4.1. and 4.2.).

The core-top sediments are presented in Figure S1 and the results for the calibration of the salinity proxy are described L319-324 and reported in figure S4. We added some details on the SST reconstruction in the surface sediment in the revised text L338-340 "*RI-OH' in the surface sediments varies from 0.05 to 0.17 while TEX-OH varies from 0.08 to 0.32. Both indexes plot in the global calibration curves from (Varma et al., 2024) and the reconstructed SST varies from 0.9 to 4.0 °C and -0.1 to 11.6 °C, respectively*.", we also added a figure for the calibration of the SST indexes (see below).

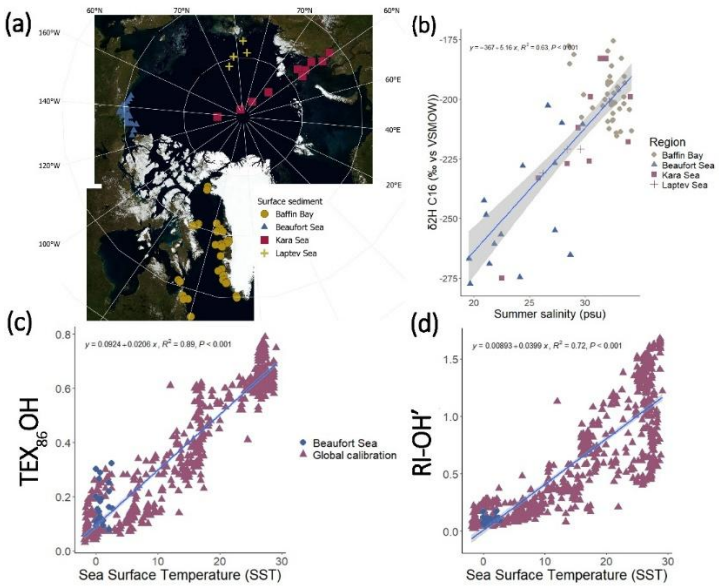

Figure S4

line 337: The use of HBI II as an arctic paleo sea ice proxy is far less common than HBI I; it does not add anything to the author's narrative. It is redundant and adds to the confusion. The same goes for HBI IV, since this paper isn't exploring the HBI TR25 index, there is no real reason to present the HBI IV as a separate record.

We do not agree with the reviewer, although new, the use of HBI-II can bring additional information for future studies and as such is included in our results. Its behaviour in both cores reflects clearly IP25 (Fig. 4) hence we do not find it confusing but rather a valuable

information. Moreover, previous studies have identified HBI-II to be at least a useful alternative when IP25 is absent or falls below detection limits (Andrews et al., 2018).

line 345: Please consider reformulating “some sea ice coverage”.

We changed it for “intermittent” L387.

line 348: Could you clarify what is the reasoning behind this heterotrophic production claim? Citations are needed.

Here we use isoGDGT as a tracer for ammonia -oxidizers as mentioned L389-390 “*Heterotrophic production in the shelf slope region during this period is relatively low (as suggested by the presence of ammonium oxidizer Thaumarchaea-derived isoGDGTs*”. We added a reference for the production of GDGT by Thaumarchaeota.

line 399-400: This claim needs some further support.

We reformulated for clarity L441-444 “*In contrast, on the outer shelf, seasonal sea-ice conditions persisted longer and sea ice cover expanded gradually and became well established after about 2 ± 0.6 ka. Even as sea-ice biomarkers increased, open-water diatom markers remained relatively abundant, implying continued flaw-lead or marginal-ice-zone productivity sustained by intermittent open-water formation and coastal influence.”.*

Figure 5: All records should have a shared time axis. If authors insist on presenting both IP25 and PIP25, they should separate them into two clear columns.

We adapted the time axis to make sure aligned records share the same axis. We do not think separating IP25 and PIP25 more than it is now done (see revised figure below) is needed.

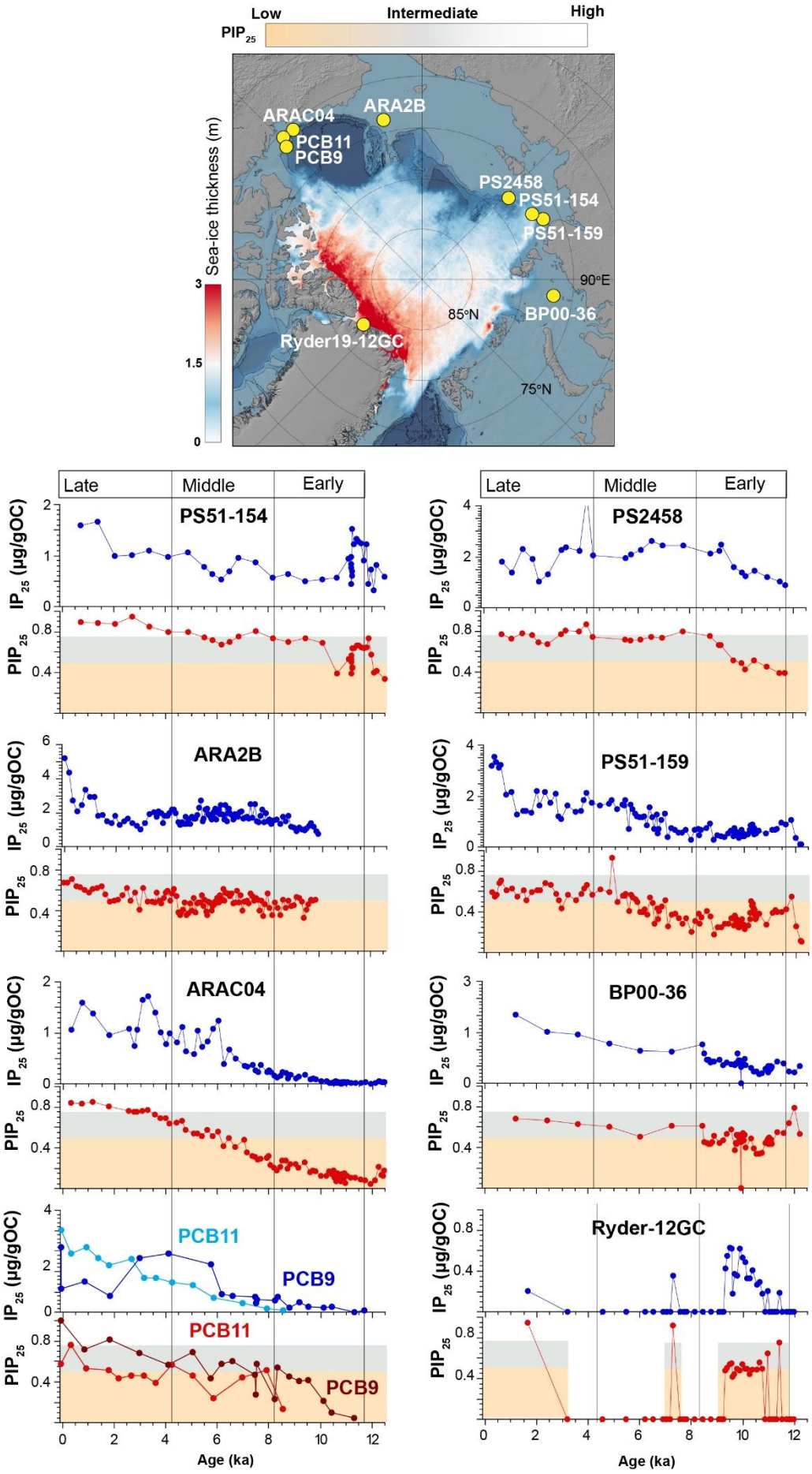
Low
Intermediate
High
PIP25
Sea-ice thickness (m)
3
1.5
0
ARAC04
ARA2B
PCB11
PCB9
PS2458
PS51-154
PS51-159
90°E
85°N
BP00-36
Ryder19-12GC
75°N
Late
Middle
Early
PS51-154
PS2458
ARA2B
PS51-159
ARAC04
BP00-36
PCB11
PCB9
Ryder-12GC
IP25 (µg/gOC)
PIP25
Age (ka)

Figure 5

line 450: The resolution of biomarker analysis in this study doesn't allow the authors to make claims about centennial events like the Little Ice Age.

We rephrased this part in the revised text to reflect the reviewer's comment 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.*".