# Reply to reviewer 1

*"Leger et al. present simulations of the Greenland Ice Sheet's history spanning the last 24,000 years. The simulations rely on a newer generation ice sheet model run at 5 km resolution and forced by a climatology from a recent transient simulation of the last deglaciation (ICESM: 21-11ka), and a climate index scheme using snapshot periods to extend this climate history back to 24ka and from the early Holocene to the year 1850 CE. An ensemble of simulations is presented, with variations in key model parameters that influence ice flow, surface mass balance, ocean-ice interaction, and climate. The results of these simulations were compared against a recent gridded reconstruction constraining the timing of GrIS retreat across the last 24kyr, which is constrained by spatiotemporally varying available geologic proxies (PaleoGris).Generally, I found the manuscript to be well written, and in most places, easy to follow. I do consider the model-data comparison framework presented here a hallmark of this work. Prior ice sheet modelling studies across similar intervals have either compared results to point measurements or regional reconstructions past ice sheet retreat (margin reconstruction), relative sea-level records, or no comparison at all. This work constitutes to my knowledge the first attempt to compare simulated ice sheet modelling results to a complete reconstruction of deglacial ice margin change (of course acknowledging the uncertainties in PaleGrIS as well). By doing so the authors constrain the state of the GrIS at a few key intervals adding knowledge as to the possible GrIS areal extent and volume during the LGM (or local LGM), drivers of deglacial retreat when many portions of the GrIS resided on or near the continental shelf, and GrIS Holocene evolution up to 1850CE. The authors do a good job acknowledging shortcomings in their modelling approach, highlighting that no unique simulation exists that satisfies goodness of fit with the geologic reconstructions during all time intervals. Nevertheless, the results and model outputs are an important contribution and should have wide appeal. While I am supportive of publication, I do have some points below that should be addressed before publication."*

## General comments:

### General comment 1:

*"From my understanding of the text, it seems the simulations were found to be most sensitive to climate (temperature and precipitation and their offsets) and the ice flow enhancement factor. Other parameters seemed to have a reduced influence on the simulated ice state and retreat. As a reader, I took this to mean that climate biases and simulated ice rheology may produce large misfits in the simulated ice history. While these uncertainties are generally expressed well in the text, few things stick out to me that could be betterer acknowledged in the text (*Note, please also see line by line comments for specific places where the text could be more clear)."*

**Author response:**

We thank the reviewer for this comment, which links to several more specific comments from both Reviewers 1 and 2. Please refer to our replies to Reviewer 1 specific comments 24, 25, 28, 31, 34 (for discussion on potential climate biases), 26 (for the enhancement factor), and also to Reviewer 2 specific comment 20, for more information on the changes made according to these comments.

### General comment 2:

*"From my knowledge working with collaborators that use ICESM and older versions like Trace21ka, there are some potential warm biases in ICESM temperature during many intervals such as the LGM and Younger Dryas (https://ui.adsabs.harvard.edu/abs/2023AGUFMPP14B..01T/abstract). This bias was shown in Badgeley et al., (2020; https://doi.org/10.5194/cp-16-1325-2020). I point the authors to Figure 13 for an example of Trace21ka temperature vs. reanalysis."*

**Author response:**

We thank the reviewer for this valuable insight. We here copy our reply to Reviewer 1 specific comment 24:

There are several reviewer comments from both reviewers 1 and 2 suggesting to make comparisons of our climate forcing with the ice-core-derived and TrACE-21-derived temperature and precipitation reconstructions of Badgeley et al. (2020). Please refer to our reply to reviewer 2 specific comment 20: where we detail how we have now added a new analysis of our climate forcing, a new figure (now Figure 25) comparing directly to Badgeley et al. 2020, and some new text in two paragraph of the discussion which refer to this comparison and provide new insight.

It is worth adding here that the newest version of iCESM used in our study (e.g. with iTRACE) is a highly different and more modern GCM from the one used in TrACE-21ka, in many ways, and the two produce very different transient climates during the deglaciation as a result. For instance, the entire atmosphere model in the CESM is different between the two. When looking at our temperature forcing data in comparison with Badgeley et al. (2020) but also TrACE-21ka and Buizert et al. (2018), the differences are spatially heterogeneous. We do not find that our iCESM-derived mean annual temperature forcing has a warm bias during the LGM, Younger Dryas, or during the early Holocene throughout the GrIS. This differs in each region. On the central western GrIS margin, for instance, our forcing presents colder temperatures than most other datasets (see new figure 25 a). Towards GrIS summit, it presents temperatures towards the upper range of data from Badgeley et al. 2020 during the Younger Dryas, and during the Holocene, but towards the lower range during HS1 (see new figure 25 c). Regardless, temperature biases would likely not influence our model simulations prior to 12 ka since our modelling produces no melt until 12 ka and retreat is instead primarily caused by ocean warming. However, during the Holocene, the main misfit we see is not enough GrIS retreat, whilst a strong warm bias would tend to generate the opposite: too much retreat.

**General comment 3:**

*In paleo ice flow modelling, the climate forcing is one of the most uncertain pieces necessary to simulate ice history over paleoclimate timescales. I would recommend the authors plot their derived temperature forcing (such as that shown in Figure 5) against the Badgeley et al. (2020) product which is available here: https://arcticdata.io/catalog/view/doi%3A10.18739%2FA2599Z26M. This may provide additional information as to if the +/- 3.5 degrees Celsius magnitude temperature offset used in the ensemble was enough to capture climate uncertainty.*

**Author response:**

We thank the reviewer for raising these points and suggesting this comparison. Please refer to our replies to Reviewer 1 specific comments 24, 25, 28, 31, 34, and also to Reviewer 2 specific comment 20, for more information on the changes made according to these comments.

**General comment 4:**

*1850 CE ice extent: I was a bit surprised to see the 1850 CE ice extent in Northern Greenland be so extensive, although given the difficulties in simulating paleo ice history this is not a major criticism. Instead, I am curious if the simulated 1850 CE state being too extensive outside of present-day ice margin is a consequence of the large magnitude cooling following peak temperatures in the climate forcing (Figure 5). Ice core proxies and reanalysis suggest a lower magnitude of cooling following the HTM, which may limit regrowth from Holocene minimum.*"

**Author response:**

We thank the reviewer for this comment and for these thoughts. Yes perhaps the magnitude of cooling in the late Holocene is responsible for this, but note that this signal is not present throughout the entire GrIS in our forcing: as shown by new Figure 25 making the comparison with Badgeley et al. (2020)'s figures 8 and 13, which we produced in reply to Reviewer 1 specific comments 24, 25, 28, 31, 34, and also to Reviewer 2 specific comment 20. In this analysis, we find that in some regions, the magnitude of mean annual cooling in the late Holocene is instead low compared to Badgeley et al. (2020)'s reconstruction or Buizert et al. (2018)'s: e.g. the CW region : see new figure 25 panel a.

Our simulated ice extent in northern Greenland remains as extensive as during the PI even during the mid-Holocene, and does not change much during this late-Holocene period despite cooling (this is shown in new Figure 17 panels a and d), so a late Holocene cooling does not seem to impact the northern sector of the GrIS in terms of GrIS margins readvancing. However we fully agree with the reviewer that our iCESM-derived climate forcing must present biases, for the 1850 CE state being this extensive in the north. Perhaps more impactful are too high precipitations during the Holocene in iCESM simulations (see new Figure 25 panel d). On this point, we have added new text to the discussion sections on climate biases specific to too high Holocene precipitations: please refer to our reply to Reviewer 1 specific comment 29 and 31 for more detail.

**General comment 5:**

"*Simulated basal ice temperature: I did find it odd that the ice sheet remains temperate, for the exception of the margins throughout the entirety of the transient simulation. While PISM may have a warm bias from looking at MacGregor et al. (2016;2022), I wonder if biases in the simulated ice sheet temperature are contributing to the ensemble derived high sensitivity to ice flow enhancement. If the simulated ice sheet is too warm, that may require the low enhancement values (to make ice softer), which the authors showed were necessary to have a better match to PaleoGris during certain time intervals. If this is a possibility, it would be nice to see this better acknowledged in the text. See below in the line-by-line comments for specific location to clarify this if my assessment is logical.*"

**Author response:**

We thank the reviewer for this great insight. Please refer to our reply to Reviewer 1 specific comment 26 which addresses the same point and shows how we modified the text accordingly. Please also refer to our reply to Reviewer 1 specific comment 23, where we discuss the possible reasons why we think our model may be producing a warm ice temperature bias.

## Specific comments:

***1) Reviewer:***

*"Line 71: Why not cite Lecavalier et al. (2014) here? I believe this reconstruction is used for the GIA reconstruction in IMBIE assessments currently (HUY3 model) and is to some degree data constrained (RSL records). I believe these other simulations listed do not do any appreciable model-data comparison."*

**Author response:**

We thank the reviewer for this comment, and have now added Lecavalier et al. (2014) in the bracket of listed references. This list of references was merely a smaller example subset of the our literature review conducted later on and which lists papers providing estimates of modelled LGM GrIS volume in SLE, and from which we particularly focus on the work of Lecavalier et al., (2014) as well.

***2) Reviewer:***

Line 88: change "…with a paleoclimate, and ii)…" to "…with a paleoclimate reconstruction, and ii)…"

**Author response:**

This change was now made accordingly, thank you for spotting this.

***3) Reviewer:***

*"Line 215 (Figure 2): In the "Red Text" line near the bottom of figure, "(results/discussion)" is highlighted. There does not seem to be a defined results or discussion section in this paper. Maybe instead, highlight the specific sections here (e.g. 3.1-4.1)."*

**Author response:**

Thank you for spotting this. There initially was a results/discussion section but the structure was changed. This figure had not been updated accordingly, by mistake. This change was now made to Figure 2 by referring to sections 3-5.

***4) Reviewer:***

*"Figure 4: I leave it up to the authors here, but I would find temperature anomalies from a reference period (e.g. 1850 CE) to be more informative about the degree of spatial variability."*

**Author response:**

We thank the reviewer for this insightful comment. As we now have added an extra figure for temperature and precipitation anomalies comparisons with Badgeley et al., (2020) and other reconstructions, following comments from both reviewers, our data is now presented in both absolute model values : i.e. with Figure 4, and in anomalies with respect to 1850 AD (i.e. with the new figure but also as 2D fields of temp and precip differences with reference to the PI in Figure 7). We think leaving both is valuable as it may prove useful to scientists also interested in visualizing precipitation and temperature magnitudes over Greenland as modelled by iCESM.

***5) Reviewer:***

*"Line 483: What is reference height temperature? Is that the 2-meter temperature at the height of the ice sheet in the climate model output?"*

**Author response:**

Yes indeed, that is correct. Although we agree this is quite specific and not clear enough, and have now simplified all text and figures to only refer to "surface air temperature" instead, and removed the term "Reference height" altogether, also in response to a similar comment from reviewer 2.

***6) Reviewer:***

*"Line 488: "...we use data...". Replace 'data' to 'output'"*

**Author response:**

We thank the reviewer for spotting this, we agree 'output' is more appropriate and have replaced the term accordingly.

***7) Reviewer:***

*Line 492-501: I am confused with the climate forcing setup when looking at the text and Figure 5. If possible, perhaps this paragraph could be rewritten so that the climate forcing timeseries is written sequentially from older to younger time periods? For example, "....between 24ka-21ka we use..., between 21ka-11ka we use transient output from ICESM, between 11-9ka we use...., and 9ka to 1850CE we use."*

**Author response:**

We thank the reviewer for this comment. We agree this paragraph is not very clear in its current form and can be improved. However a purely chronological order can also turn into being confusing due to the complexity of our data multiformat. The key is that the reader understands there are a few snipets of time within our simulations' timeframe with actual climate model data : some are time slices (21, 9, 6, 3 and PI) and one transient data-frame (iTRACE), and other snippets of time for which we use a glacial index scheme to bridge the data gaps in time. We have now modified the paragraph to separate these two distinct sources in two distinct paragraphs, and within these then respect more a chronological order. We believe this new structure is clearer and easier to read:

"*Air temperature and precipitation*

SMB is forced with two-dimensional and time-dependent fields of surface air temperature and total precipitation (Figs. 4-7). We use pre-existing simulations from iCESM (Brady et al., 2019) versions 1.2 and 1.3, run globally at a horizontal resolution of 1.9° in latitude and 2.5° in longitude for the atmosphere and a nominal 1° for the oceans. We use simulations ran with full forcing, *i.e.* including ice sheet (from ICE-6G: Peltier et al., 2015), orbital (Berger, 1978), greenhouse gases (Lüthi et al., 2008) and meltwater forcings. Between 20 and 11 kyr BP, we use monthly-resolution output from the iTRACE experiment, ran with iCESM 1.3 (He et al., 2021a, b). Thanks to an improved climate model, higher resolution, and the addition of water isotopes, iTRACE simulates a climate over Greenland that is more data-consistent (He et al., 2021a) than the former CESM simulation of the last deglaciation TRACE-21 (Liu et al., 2009). Additionally, we use output from five equilibrium time-slice simulations ran at 21 kyr BP (iCESM 1.3), at 9, 6, and 3 kyr BP (iCESM 1.2), and at the PI (1850 AD, iCESM 1.3) (Fig. 4).

To create continuous forcing over remaining data gaps in time, we use a glacial index approach (Niu et al., 2019; Clark et al., 2022) and linearly scale our climate fields proportionally to variations in independent climate reconstructions in a space-dependent manner i.e. building a glacial index for each individual grid cell (Fig. 5). Between 24 and 21 kyr BP, we use surface air temperature and $\delta^{18}O$ reconstructions of Osman et al. (2021) to scale variations in temperature and precipitation fields, respectively. For data gaps between 21 kyr BP and the PI (*e.g.* 11 - 9 kyr BP), we use the seasonally-resolved Greenland-wide temperature and precipitation reconstruction of Buizert et al. (2018) as glacial index.

As a result, we produce time-dependent, two-dimensional fields of mean annual and mean summer (JJA) surface air temperature and mean precipitation rate, continuous between 24 kyr BP and PI (Fig. 4-7). From mean annual and mean summer temperatures, our SMB scheme reads a cosine yearly cycle generating an idealised seasonality signal."

***8) Reviewer:***

*"Line 505: The authors use sea-surface temperature as the oceanic forcing. This is not ideal as research supports that shelf depth temperature is more influential for ice-ocean interaction in Greenland. Likewise, the authors could have used simulated shelf depth temperature from ICESM, similar to Tabone et al., 2024 (https://www.nature.com/articles/s41467-024-50772-5), since they used modelled outputs of salinity bridged by linear interpolation in this modelling. While I recognize that there is a lack of palaeoceanographic reconstructions at depth, especially over such timeframe, this shortcoming should be at least acknowledged here in the text and discussion (which I will highlight further down)."*

**Author response:**

This valuable point was also raised by Reviewer 2, and lead us to further justify our choice in this paragraph with extra sentences. The main reason is that we considered the data-assimilation, proxy-based, and transient nature of the reconstruction by

Osman et al. (2021) from 24 kyr BP to present to be powerful, and to be a safer and a less biased forcing product than raw iCESM outputs which are not proxy-data corrected, even if they are shelf-depth. However we agree this may prove to be a shortcoming of our modelling design and discuss it more in section 5 when analysing potential reasons for model-data misfits.

"For the ocean temperature, we use the LGM-to-present ensemble-mean sea surface temperature (SST) reconstruction of Osman et al. (2021), yielding a 200-year temporal resolution and nominal 1° spatial resolution (Fig. 6). This re-analysis uses Bayesian proxy forward models to perform an offline data assimilation (using 573 globally-distributed SST records) on climate model priors; *i.e.* a set of iCESM 1.2 and 1.3 simulations (Zhu et al., 2017; Tierney et al., 2020). Whilst we acknowledge sub-shelf ocean temperature would be a more appropriate forcing than SST, their does not yet exist a Greenland-wide time- and space-dependent sub-shelf ocean temperature reconstruction which assimilates proxy data between 24 kyr BP and the PI. The transient and data-assimilated nature of the SST reconstruction by Osman et al. (2021) was thus preferred to iCESM outputs of shelf-depth ocean temperature (e.g. Tabone et al., 2024)."

***9) Reviewer:***

*"Line 505: What is the reason for starting from 24 kyr BP? Is it possible that staring at 21 kyr BP, when transient output starts for ICESM would have yielded similar results for the simulations?"*

**Author response:**

We thank the reviewer for this interesting question. The main reason for this choice is that the transient SST and Temperature reconstruction of Osman et al. (2021) starts at 24 kyr BP. This also has the added advantage to be a few more thousand years before the local LGM (17-15 kyr BP), than 21 kyr BP. This extra buffer time can let the model adjust to new ensemble-varying parameters and new boundary condition in each ensemble simulations prior to HS1, as all ensemble simulations initially start from the same spinup output at 24 kyr BP. As our input climate at 24 kyr BP is notably warmer than our iCESM-derived 21 kyr BP equilibrium climate simulation, using 24 kyr BP as a time start makes sure that we obtain the correct local LGM GrIS extent for the right reasons: i.e. we first build the LGM by modelling GrIS advance from a more deglaciated extent during ensemble simulations (unlike Tabone et al., 2024), as opposed to producing a spin up directly at the LGM extent, which is as we know from experience easier to achieve as one may let the model build for tens of thousands of hypothetical years to get to its LGM positions by tweaking parameters manually. We acknowledge it is however possible that results would have been similar by starting at 21 kyr BP, although this is hard to tell at this stage.

***10) Reviewer:***

*"Line 668: Make sure citations in chronological order (O Cofaigh)."*

**Author response:**

We thank the reviewer for spotting this mistake. This has now been corrected.

***11) Reviewer:***

"*Line 692: Doesn't Figure 9 show the periphery glaciers (Test 3)? The text states periphery glaciers are removed.*"

**Author response:**

We thank the reviewer for spotting this error in the text. Old Figure 9 (now figure 8) is accurate and periphery glaciers were only removed for the deglaciation extent test, as we then compare against PaleoGrIS which only reconstructs the GrIS. Periphery glaciers were indeed included in the Pre-industrial extent test. We have now updated the text to only mention the removal of periphery glaciers in the deglaciation extent test paragraph.

***12) Reviewer:***

"*Line 715: Can Table 1 be moved up closer to the text here?*"

**Author response:**

We agree it would be more appropriate to move Table 1 to just after the relevant section 2.3. and have now done this change accordingly.

***13) Reviewer:***

"*Line 846: I do wonder to what extent the poor fit in NE Greenland is due to the use of SST's from Osman et al.? Tabone et al. (2024; nature.com/articles/s41467-024-50772-5) simulate the ice margin in this region using a different ice sheet model, forced by Trace-21ka shelf depth temperature. Would it be possible to compare the ocean temperatures used in this study across this region to shelf depth temperatures from ICESM? Perhaps anomalies would be best to compare against.*"

**Author response:**

We thank the reviewer for this comment. As detailed in section 5, we believe there are multiple possible, and perhaps interacting reasons for these model-data misfits, possible biases in the use of SST from Osman et al (2020) being one of them, along with underestimation in precipitation forcing, inappropriate SMB or basal drag parameterisation etc. Proper attribution of these potential drivers of misfit would require a full study in itself, we feel, which would actually make a very interesting follow-on study. At this stage, it remains very hard to know for sure, however. Future simulations and studies that I plan to make to improve this initial modelling work will also help attribute these different causes.

Please find below here a supplementary figure from Tabone et al., (2024) indicating their shelf-depth ocean temperature forcing towards the outer shelf of NE Greenland. The linear interpolation seems very sudden and crude in the data, which is likely related to the poor global resolution of CESMs such as *Trace- 21ka or iCESM* which cannot properly resolve fine details of the bathymetry and thus do not actually resolve a proper continental shelf in Greenland. Regardless, their resulting ocean temperatures are both notably colder during full glacial and warmer during the Holocene than ours, offshore NE Greenland. We agree that this is worth noting, and have now added an extra sentence referencing specifically to Tabone et al., (2024) and their supplementary figure to the relevant discussion section where we mention, in detail, our SST forcing as a potential source of misfit during deglaciation in the NE. Note also that extra text was added to the paper relative to ocean forcing also to answer similar comments by reviewer 2 (see reply to reviewer 2 specific comment 18 for instance): and to also acknowledge that our experiment could have perhaps been improved by incorporating ocean temperature forcing offsets within our ensemble.

"Another potential source of model-data misfit could be biases in our input climate forcing causing either too low precipitation rates, or too high sea-surface temperatures (SST) across NO and NE Greenland. We do not expect biases in our air temperature forcing to have a meaningful impact at this stage, as despite our conservative ensemble parameter perturbations, we find no PDD-derived surface melt is produced until 12 kyr BP, several millennia after the local LGM and initial deglaciation, due to mean annual and summer temperatures remaining below 0°C (Figs. 4, 5). We note that during HS1 cooling, input mean-annual SST drops to lower minimum values (-2 to -3 °C) offshore SE and SW Greenland than offshore NE Greenland (-1.5 to -2 °C) (Fig. 6), which may highlight a possible overestimation of our sea-surface temperature forcing (from Osman et al., 2021) in NE Greenland during the local LGM. This 0.5-2°C drop in SST at around 18-17 kyr BP, which occurs in response to HS1, is a key driver of modelled GrIS expansion during the local LGM, as it is associated with sharp reductions in GrIS-wide sub-shelf melt rates and thus basal mass loss rates (Fig. 11). A small underestimation in HS1 sea-surface cooling offshore NE Greenland, in the order of 1-2°C for instance, may be enough to deter the modelled GrIS margins from advancing extensively. This hypothesis may also be reinforced by the general lack of spatial coverage of SST proxy records used in the data-assimilation scheme of Osman et al. (2021) north of 65°N, offshore Greenland coasts. Biases may also be introduced by our interpolation scheme used for resampling from the nominal 1° horizontal resolution of the original data (Osman et al., 2021), equivalent to a ~20 x 27 km grid offshore NE Greenland, to our 5 x 5 km model grid. This highlights that our experiment may be limited by a lack of variation in SST input fields between ensemble simulations. A future experiment using an ensemble-varying parameter introducing spatial and temporal perturbations to the input ocean forcing may help test this hypothesis and possibly increase model-data fit."

"Alternatively, the underestimated modelled GrIS retreat in NO, NE, and CE Greenland could be associated with a lower-than-needed ocean temperature increase during the last deglaciation (Osman et al., 2021; Figs. 6-7) offshore the present-day GrIS. It is important to note that our ice-ocean interaction model does not consider multiple ocean layers, which are important when poorly mixed sub-surface layers of higher temperatures increase sub-shelf melt at depth and towards the grounding line (Lloyd et al., 2023). It also does not consider a seasonal cycle of ocean water temperature change as forcing, which may be important to model the necessary magnitude of deglacial sub-shelf melt in these regions. We note that, for instance, TrACE-21ka-derived shelf-depth ocean forcing used in Tabone et al. (2024; Fig. S3 therein) reaches above 0°C (up to 2°C) towards the NE Greenland outer shelf, between 13 and 8 kyr BP, whilst our SST forcing does not produce values above -1°C in that region and timeframe.”

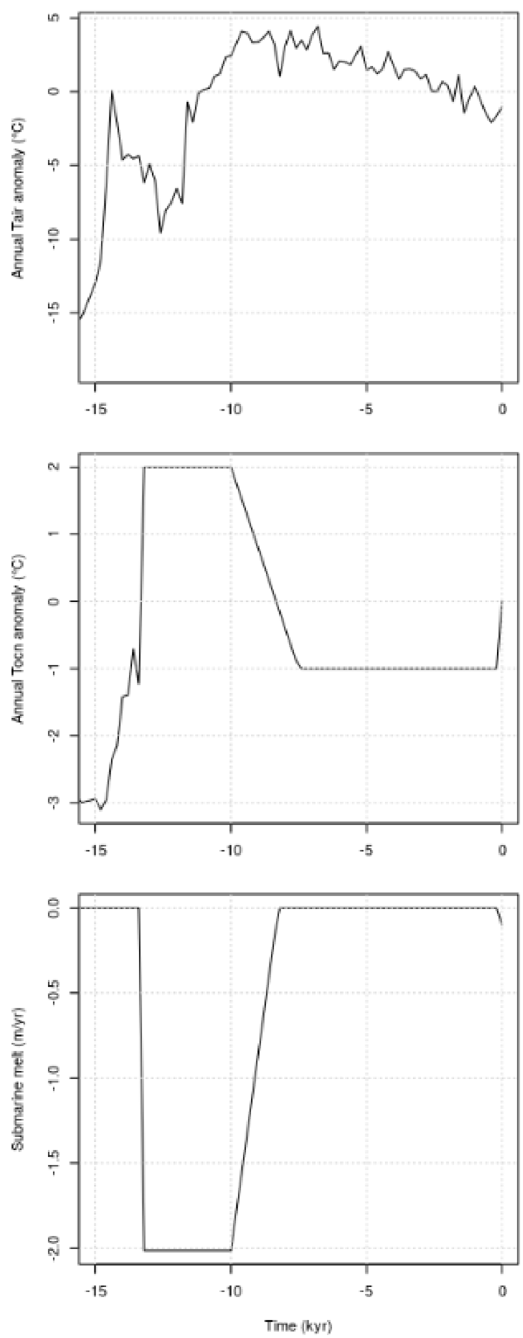



**Figure S3**. Atmospheric temperature anomaly (upper panel), ocean temperature anomaly (middle panel) and submarine melt (lower panel) used to force the ice-sheet model in the northeast outer shelf.

Tabone et al. (2024) supplementary figure 3.

*14) Reviewer:*

*"Page 24: There are references to a number of geographic locations. If possible, it would be nice if some of these locations could be placed on one of your Greenland maps (abbreviated), for example on Figure 14."*

**Author response:**

We thank the reviewer for spotting this, and we have now added labels for the locations of the few places / ice streams mentioned in these paragraphs, to old Figure 14 and 17, now figures 13 and 16.

*15) Reviewer:*

*"Line 940-950: Very interesting to see the relationship to model resolution. I would also acknowledge that each model used different climate forcings, ice flow approximations. Also, do these models use different model extents for the LGM mask (max. ice extent)?"*

**Author response:**

We thank the reviewer for this comment. Indeed, as empirical reconstructions have evolved, some of these studies which nudge their model to a specific extent mask (which is not the case in our study) will have used different LGM masks. We have added this extra information to the paragraph:

"All previous studies producing an ensemble of GrIS LGM-to-present model simulations with model-data comparison (Lecavalier et al., 2014; Simpson et al., 2009) used substantially coarser grid resolutions (15-20 km) than this study (5 km). Of these modelling studies, moreover, few include floating ice shelves in their models, which are known to often provide a buttressing effect leading to ice-flux lowering and thus increases in grounded ice-sheet thickness (Pritchard et al., 2012). Each of these studies also use different climate/ocean forcings and ice flow approximations, and those nudging the model to a specific ice extent may use different data-informed lLGM masks. Together, these differences may help explain the higher volumes obtained in our results."

*16) Reviewer:*

*"Line 1194 and 1270: For consistency, would it be better to change title in line 1194 to "….during the last deglaciation" instead of late glacial to better match the subtitle on line 1270?"*

**Author response:**

We agree that this would be more appropriate and now made this change to the title accordingly.

***17) Reviewer:***

*"Line 1205-1210: This result is similar to Tabone et al. (2018; https://cp.copernicus.org/articles/14/455/2018/), so would be useful to cite."*

**Author response:**

We thank the reviewer for this valuable insight, and have now added a reference to Tabone et al. (2018) towards the end of this paragraph:

"Similarly to results of Tabone et al. (2018), our ensemble thus suggests that during the late HS1 and B-A warming, between 16 and 14 kyr BP, ocean forcing likely caused the GrIS to retreat rapidly and lose most of its glaciated continental-shelf areas, despite air temperatures remaining too cold to produce any surface melt (Fig. 11)."

***18) Reviewer:***

*"Figure 16: Should the subtitle read "5 best lLGM simulations" instead of "5 best LGM simulations"?"*

**Author response:**

Yes indeed that would be more consistent. The change to old Figure 16 (now figure 15) was made accordingly. We thank the reviewer for spotting this.

***19) Reviewer:***

*"Line 1276: I cannot find the '13-12 kyr BP' 0me slice in supplemental figure 2."*

**Author response:**

Indeed, to be more consistent with the supplementary figure 2, this should read '14-12 kyr BP', and not '13-12 kyr BP'. The change was made to the text accordingly, thank you for spotting this.

***20) Reviewer:***

*"Line 1373: I think Briner et al., 2014 is the wrong citation here as that was more for west central Greenland. Larsen et al. (2015; https://doi.org/10.1130/G36476.1) and Larsen et al. (2011; https://doi.org/10.1016/j.quascirev.2011.07.022) would be more appropriate citations here."*

**Author response:**

We thank the reviewer for this comment. Indeed, we agree that Larsen et al. 2011; 2015 are more appropriate references for this sentence given the locations of where we model such retreat, and have now replaced the original reference by those.

***21) Reviewer:***

*"Line 1377-1382: It would be good to know what parameter combinations were responsible for the ~100 km retreat (high PDD factors?)."*

**Author response:**

As is shown in Figure 24 (previously 26), it can be difficult to isolate the role of a single parameter due to their complex interactions and the fact that we do not explore the parameter space enough, with 100 simulations, to find global minima in parameter value solutions maximizing model-data fit. However, when analysing parameter values for the 5 best-scoring simulations at the PI extent test (green dots on Figure 24), there are no clusters or trends in PDD melt factors, meaning that this 100 km retreat does occur despite a wide range of PDD melt values for both snow and ice (high or low): e.g. in simulations 78 and 31 (Supplementary Fig. 4). Following this reviewer comment, we have however added a sentence referring to the existing clusters of high temperature offset values and low precipitation offset values for these best-scoring PI simulations, as a correlation and likely causality:

"Although this result may well be an overestimation and should be interpreted with caution, our modelling suggests such a magnitude of retreat behind present-day margins (~100 km) in response to the HTM cannot be fully ruled out, in certain regions. This is correlated to, and likely caused by, PI best-fit simulations all presenting both positive (>+1.5°C) and negative (<40% of original) temperature and precipitation offset ensemble-varying parameter values, respectively (Fig. 24)."

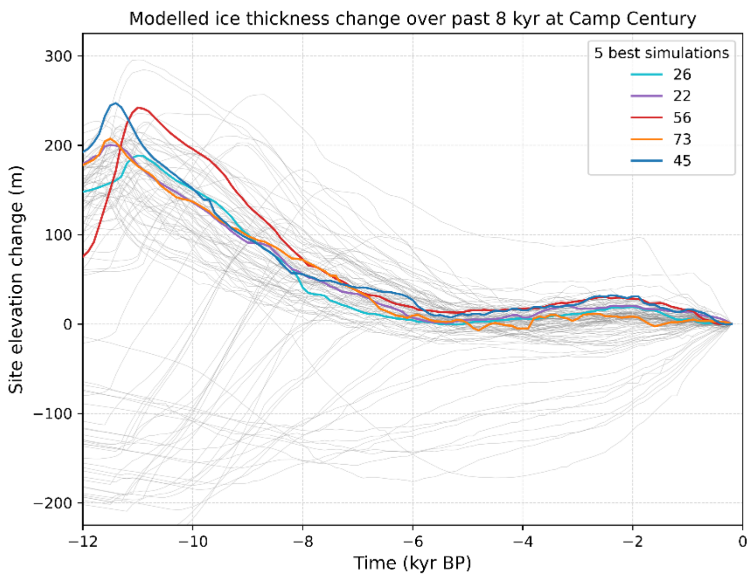
***22) Reviewer:***

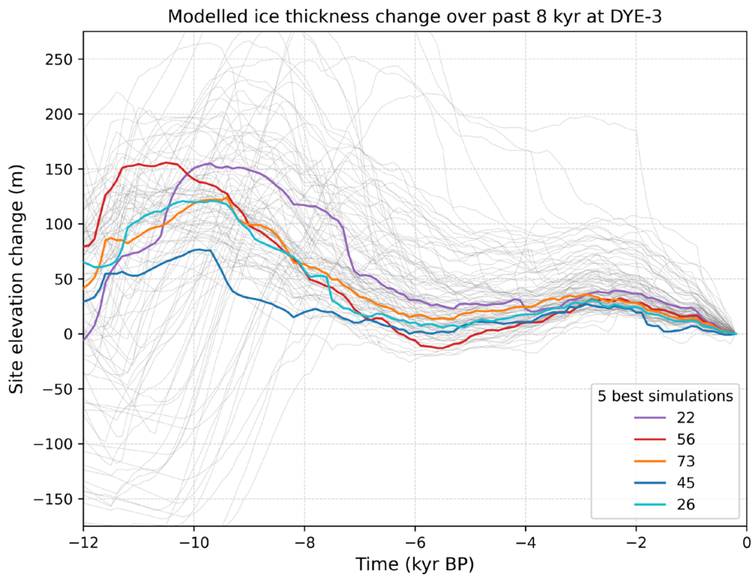
*"Lines 1521-1544: The thinning curves go back to ~11.5 ka. Any reason why the analysis was only for the 8kyr to 1850CE period and not back further since this section does include discussion about the early Holocene? It would have been nice to see if any models simulated the early Holocene thickening."*

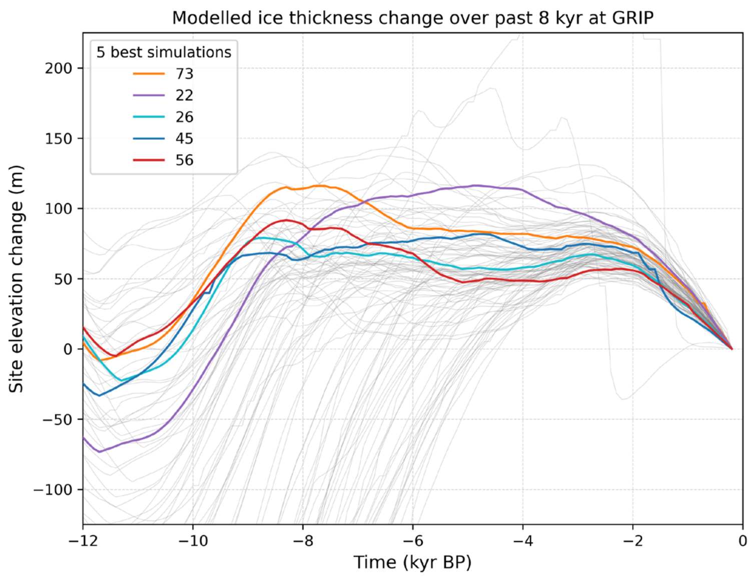
**Author response:**

Here we followed Lecavalier et al. (2013; 2014) who exclusively focus on the 8-0 kyr BP interval as thinning curves prior to 8 kyr BP are notably more uncertain, subject to more assumptions, and need more calibration (which involves glacier modelling and thus some circularity as far as we can tell: e.g. Lecavalier et al. 2017) : due to the possible influence of the Innuitian ice sheet demise across the Canadian Arctic on the altitude correction required to infer temperature from Agassiz ice during the early Holocene, which is then used to analyse the delta18O residuals from other ice cores used to estimate altitude changes of the ice surface through time.

Our simulations do indeed present some thickening during the early Holocene, however, but not a lot: in the order of ~150-50 m for DYE3, ~100 m for GRIP, ~50 m for NGRIP, and ~200 m for Camp century. Please see the below plots for the results:


Modelled ice thickness change over past 8 kyr at Camp Century
5 best simulations
26
22
56
73
45
Site elevation change (m)
Time (kyr BP)


Modelled ice thickness change over past 8 kyr at DYE-3
5 best simulations
22
56
73
45
26
Site elevation change (m)
Time (kyr BP)


Modelled ice thickness change over past 8 kyr at GRIP
5 best simulations
73
22
26
45
56
Site elevation change (m)
Time (kyr BP)

***23) Reviewer:***

*“Lines 1546-1555: It is interesting that the LGM and 1850 CE simulated basal regime has the majority of the ice sheet at pressure melting. I am assuming the margins are colder based because the ice is thinner there, but I would have expected to see some modelled transient behavior of the ice sheet through time. Looking at MacGregor et al., 2016;2022, it does seem compared to other ice sheet models, PISM may have more of a warm bias. However, at the ice divides where vertical advection dominates, it’s surprising not to see any cold based ice. Is there any explanation for this?”*

**Author response:**

We thank the reviewer for raising this interesting comment, which was also raised by the Editor. To be 100% honest, we are also a little puzzled by this result. And here are our current thinking on it after investigating.

-We are using the default enthalpy module and basal boundary condition parameterization of PISM, which was used by numerous studies before in both Greenland and Antarctica ice sheet wide simulations. So maybe there are differences related the enthalpy formulation of ice-sheet thermodynamics specific to PISM (Aschwanden et al. 2012). However we do not apply any changes from the default parameterization, so there isn’t a change specific to this experiment that can necessarily explain this possible warm bias.

- Like was pointed by the reviewer, we see a large diversity of basal thermal regimes even modelled just for the present day : in “GBaTSv2: revised Greenland basal thermal state” by Mcgregor et al (2022, see figure 2). In that figure, we notice PISM simulations tend to produce more temperate conditions towards the ice sheet interior than most other models, so there seems to be a trend.

- There are very large differences between different geothermal heat flux input datasets, The one we use (Martos et al. 2018), also used by Tabone et al. (2024), was chosen because the authors from that study argue their heat flux reconstruction better captures the former passage of the Iceland Hotspot through Greenland, and is thus more realistic. However the result in this geothermal heat flux produces substantially higher values than what other modellers may have used under the GrIS: with values of up to 65-70 mW $m^{-2}$ under the GrIS summit (see Fig. 1c), whilst the heat flux forcing used by SeaRISE and used in many other studies, for instance (Shapiro, N.M. and Ritzwoller, M.H., 2004) and which is inferred from a global seismic model, don’t produce values above 45-50 mW $m^{-2}$ under the GrIS (thus up to 1.5 times lower heat flux below the GrIS summit). This discrepancy may just be enough to influence basal conditions to the point that warm-based temperatures prevail.

-Another likely driver is our model initialization, i.e. it is likely our climate used for the constant climate spinup at 24 kyr BP is slightly too warm and, due to the length of that spinup (30 kyr), ends up being too warm for too long, thus introducing a bias in modelled ice temperature which has time to spread through the ice column, and which is then inherited by the transient ensemble simulations: i.e. a biased thermal memory related to our model initialization. As part of another study, we once conducted an equilibrium spinup at the PI using constant climate but parameters from the best scoring simulation from this ensemble (see EGU presentation of this here: https://doi.org/10.13140/RG.2.2.12786.11208). The resulting basal ice temperatures (pressure-adjusted) after reaching steady-state under our constant PI climate show a smaller region of warm-based ice under the GrIS, than our PI state at the end of the transient 24 ka-to-present simulation conducted using the exact same parameters. We take this as evidence that our initialization climate at 24 kyr BP is likely introducing a bias enhancing warm-based conditions underneath our modelled GrIS.

-From the literature, we have struggled to find paleo and LGM-to-present simulations with thermomechanical ice sheet models where the pressure-adjusted basal temperature outputs are provided for visualization/download, so it remains difficult to get a point of comparison from other simulations/models/parameterizations with our results on the modelled basal temperatures over the LGM-to-present timeframe at least. That being said, the best-scoring simulation from Tabone et al., (2024) is available for download (https://zenodo.org/records/12667358). We notice that in their model, vast regions of the ice sheet's interior are also at pressure melting point throughout the entire LGM-to-present timeframe, including underneath the GrIS summit, NEGIS region, Humboldt glacier region, and entire Southwest region. However much wider regions are cold-based towards the southeast and Northwest GrIS in their model, compared with ours.

- Finally, we do not believe this potential warm bias to represent a major issue, as otherwise, if our basal thermal state was so off that it would cause too much basal mass loss, too low basal yield stress, and too much sliding and ice discharge as a result, our model would perhaps not succeed to build the lLGM extent within our ensemble transient simulations, which it does in agreement with empirical data in almost all regions, and also retreat during deglaciation, which it does successfully in certain regions (not all).

- It must also be noted that basal thermal regimes are likely one of the most uncertain parts of the modelling we do, especially in the paleo, and that changes from cold-based to warm-based occur over a small threshold that is sensitive to subtle changes in conditions, whilst basal temperatures are often coupled with basal yield stress and sliding, thus introducing complex feedbacks.

In order to acknowledge that our model likely presents this warm ice temperature bias, however, we have added some text to section 4.3, which is also in line with reviewer 1 comment 26:

"This may imply that better model-data fit during maximum expansion requires to model a GrIS with harder, less deformable, and more viscous ice (or with lower impurity contents), than is modelled with default flow law constants (*E*=1, *n*=3). However, this may also represent a compensating adjustment from our modelled ice temperatures, which are warmer (thus possibly resulting in too soft ice) and produce more widespread warm-based conditions over greater proportions of the GrIS than most other GrIS models (e.g. Tabone et al., 2024; MacGregor et al., 2022) and this across all best-fit simulations (e.g. Supplementary Figs. 7, 8)."

***24) Reviewer:***

*"Lines 1610-1620: I really do wonder if the climate forcing used is really responsible for some of this mismatch. See some of my other comments, but ICESM has a warm bias compared to ice core proxies with it being warmer during the Younger Dryas than proxies show. Additionally, the trend in Holocene temperature is not that similar to proxies (or Data assimilated products like Badgeley et al., 2020; Buizert et al., 2018), which reconstruct the HTM earlier in the Holocene, followed by a slight cooling towards stable temperature, whereas the reconstructed climate used here has peak warming at ~ 6 ka followed by a large magnitude of cooling to 1850CE."*

**Author response:**

We thank the reviewer for this comment. There are several reviewer comments from both reviewers 1 and 2 suggesting to make comparisons of our climate forcing with the ice-core-derived and TrACE-21-derived temperature and precipitation reconstructions of Badgeley et al. (2020). Please refer to our reply to reviewer 2 specific comment 20: where we detail how we have now added a new analysis of our climate forcing, a new figure comparing directly to Badgeley et al. 2020, and some new text in two paragraph of the discussion which refer to this comparison and provide new insight.

It is worth adding here that the newest version of iCESM used in our study (e.g. with iTRACE) is a highly different and more modern GCM from the one used in TrACE-21ka, in many ways, and the two produce very different transient climates during the deglaciation as a result. For instance, the entire atmosphere model in the CESM is different between the two. When looking at our temperature forcing data in comparison with Badegeley et al. (2020) but also TrACE-21ka and Buizert et al. (2018), the differences are spatially heterogeneous. We do not find that our iCESM-derived mean annual temperature forcing has a warm bias during the LGM, Younger Dryas, nor during the early Holocene, throughout the GrIS. This differs in each region. On the central western GrIS margin, for instance, our forcing presents colder temperatures than most other datasets (see new figure 25 a). Towards GrIS summit, it presents temperatures towards the upper range of data from Badgeley et al. 2020 during the Younger Dryas, and during the Holocene, but towards the lower range during HS1 (see new figure 25 c). Regardless, potential warm temperature biases would not influence our simulations at all prior to 12 ka since our modelling produces no melt until 12 ka and retreat is instead primarily caused by ocean warming. However, during the Holocene, the main misfit we see is not enough GrIS retreat, whilst a strong warm bias would generate the opposite: too much retreat.

**25) Reviewer:**

*"Section 4.3: - The temperature and precipitation offset were shown to have a large influence. Doesn't this mean that the climate forcing uncertainty plays a major role? It would be interesting to see where the reconstruction for temperature used in this work compares against data assimilated products which are available (Badgeley et al., 2020)."*

**Author response:**

Please refer to our reply to specific comment 24 above, and to our reply to Reviewer 2 specific comment 20, where we present our new analysis and comparison with the climate of Badgeley et al. (2020). We fully agree with the reviewer that climate forcing but also ocean forcing uncertainties likely play a major role here regarding our misfits: despite us modifying the climate within our ensemble by using ensemble-varying temperature and precipitation offsets, and also varying PDD melt factors: thus resulting in quite different forcings between each of our 100 simulations. We believe that, along with the previously written discussion plus the addition of the new comparison with Badgeley et al., 2020, this point is now discussed extensively in sections 5.1 and 5.2: where a total of 1032 words are now written just on the potential impact of climate forcing biases on our misfits.

**26) Reviewer:**

*"Additionally, the flow enhancement was shown to be important, requiring 'more rigid' ice to expand during the LGM (line 1635). Could this be a consequence of the simulated ice temperature (at least we see the basal temps in supplemental figures) being too warm, and therefore needing adjustment in ice flow enhancement?"*

**Author response:**

We thank the reviewer for this valuable insight. We agree that this is a possibility and have now added a sentence to paragraph 4.3 in line with the observation of rather warm basal temperatures in our model, and referring to supplementary figures 7 and 8 for the reader to see this potential bias:

"This may imply that better model-data fit during maximum expansion requires to model a GrIS with harder, less deformable, and more viscous ice (or with lower impurity contents), than is modelled with default flow law constants (*E*=1, *n*=3). However, this may also represent a compensating adjustment from our modelled ice temperatures, which are warmer (thus possibly resulting in too soft ice) and produce more widespread warm-based conditions over greater proportions of the GrIS than most other GrIS models (e.g. Tabone et al., 2024; MacGregor et al., 2022) and this across all best-fit simulations (e.g. Supplementary Figs. 7, 8)."

**27) Reviewer:**

*"Line 1652: I agree an ensemble approach is beneficial, but future work could also make use of more climate reconstructions. Here I think the climate uncertainty is under sampled compared to the parameter uncertainty (personal opinion though)."*

**Author response:**

We thank the reviewer for this comment. We fully agree with this opinion, and we also believe that future experiment would benefit from producing several waves of ensembles, with perhaps a first ensemble primarily focused on using different climate and ocean reconstructions as forcings (i.e. external forcings) and exploring this more widely than we did in this study, prior to a second ensemble focusing on internal dynamics and parameterisations of mass exchange processes. However, this would come with a high computational cost. Our 100 member ensemble took just above 1 year of computation on HPCs to complete. We have now added an extra sentence to a relevant paragraph discussing climate exploration and ensembles (section 5.1 paragraph 5) to support this point further for future work:

"Alternatively, our ensemble may be too small to fully explore the full impacts of our climate correction parameters on grounded GrIS extent evolution. As a test, we conducted an additional simulation using default (mid-range) values for all ensemble-varying parameters excluding the precipitation scalar offset (Table 1), here set to 2.0 (+200% precipitation rate). This test simulation successfully produces an extensive HS1 advance of the grounded GrIS margin offshore NE Greenland, reaching a mid-shelf position. This modelled local LGM advance is more extensive than any of our ensemble simulations, and suggests a 100 simulation ensemble is too small to explore the parameter-space region that models this preferable GrIS behaviour. Therefore, although computationally unfeasible here, running a larger ensemble while keeping perturbed parameter ranges identical to our setup may likely produce simulations yielding a better model-data fits in ice extent, during the local LGM. Alternatively, future experiments running several ensemble waves (e.g. Lecavalier and Tarasov, 2025), with a first ensemble exclusively focused on more widely exploring different climate and ocean forcings with different perturbations schemes, may achieve more data-consistent GrIS LGM-to-present simulations."

**28) Reviewer:**

*"Line 1732: Again, I think ICESM has an LGM warm bias (or at least should be looked into). See Badgeley et al. (2020; Figure 13). At least Trace21ka had a warm bias during the LGM compared to data assimilated products and ice core proxies."*

**Author response:**

We thank the reviewer for this comment. Please refer to our replies to Reviewer 1 comments 24, 25, and to Reviewer 2 comment 20: where we present our new analysis and comparison with the climate of Badgeley et al. (2020) and other datasets presented therein. The output from iCESM's newest versions, which we use in our study, do not seem to produce a warm bias during the LGM compared to data assimilated products (at least in most regions): see new figure 25 a,c, unlike TrACE-21ka which did indeed.

**29) Reviewer:**

*"Lines 1785-1790: Downs et al. (2020; https://tc.copernicus.org/articles/14/1121/2020/) used data assimilation techniques, sampling climate and model parameter uncertainty,*

*to determine that higher precipitation than modern was likely needed to simulate Holocene margin migration in SW Greenland. Might be a good citation here to acknowledge that other modelling studies have come to a similar conclusion indicating a need for increased precipitation."*

**Author response:**

We thank the reviewer for this comment and for suggesting adding this highly relevant reference. Since this work finds a need for increased precipitation increases during the Holocene, we have now added this reference to our paragraph speaking about potential precipitation biases from our climate forcing during the Holocene, which we actually think may present a too-high precipitation bias impeding retreat (at least in some regions i.e. GrIS summit) during the mid-Holocene, visible when compared with Badgeley et al. (2020) and other datasets (see new figure 25 d).

"On the other hand, this comparison reveals that our iTRACE and iCESM - derived climate forcing results in significantly higher (up to ~100%) precipitation rates during the entire Holocene towards the GrIS summit and its vicinity, than is obtained in ice-core-data-informed reconstructions from Badgeley et al. (2020) and Buizert et al. (2018) (Fig. 25d). Although the HTM has been shown to likely be associated with higher-than-present precipitations (e.g. Downs et al., 2020), and although our experiment features an ensemble-varying precipitation offset scheme with possible reductions down to 20% input precipitations, this potential positive bias may be responsible for too high Holocene precipitation in many of our ensemble simulations, thus impeding GrIS retreat in certain regions and causing ice extent overestimations during the modelled deglaciation but also during the PI (Fig. 25d)."

**30) Reviewer:**

*"Lines 1795-1805: I am totally fine with the use of the Osman SST records as forcing for the ocean model. But it would be good to acknowledge here that it is not ideal, and shelf depth temperature would be more accurate (even though that does not exist)."*

**Author response:**

We agree this should be acknowledged. Please refer to Reviewer 1specific comment 8 and 13, where we explain the changes made accordingly.

**31) Reviewer:**

"Lines 1808-1815: Is there any information on the bias of ICESM precipitation for present day?"

**Author response:**

We thank the reviewer for this valuable comment. Yes indeed, for example Lenaerts et al. (2020) looked into this and show (e.g. Figure 5 therein) a positive bias in present-day modelled precipitations (mostly in snowfall) whereby CESM overestimates annual precipitation frequency relative to CloudSat observations (although not by more than 20%). We have now added a reference to this to the relevant paragraph:

"[...] this potential positive bias may be responsible for too high Holocene precipitation in many of our ensemble simulations, thus impeding GrIS retreat in certain regions and causing ice extent overestimations during the modelled deglaciation but also during the PI (Fig. 25d). Moreover, it is worth noting that CESM is known to also overestimate (by <20%) present-day snowfall precipitations over the GrIS relative to observations which my also explain our overestimations in ice extent during the PI (e.g. Lenaerts et al., 2020; Fig. 5 therein)."

**32) Reviewer:**

*"Lines 1820-1825: See comment above. Downs et al. (2020) may be a good citation supporting enhancement of precipitation anomalies needed to match geologic reconstructions."*

**Author response:**

Please refer to our reply to Reviewer 1 specific comment 29 where we address this and make changes accordingly.

**33) Reviewer:**

*"Section 5.2: This paper is long (and for good reason with all of the analysis), so maybe this would not make a big impact. However, since you did not simulate the GrIS with multiple mesh resolutions, perhaps this portion of the text could go in the supplemental, and you can instead shorten this by citing your analysis of bedrock topography at different resolutions (if you move the majority of this text and figure to the supplement)? Also, Cuzzone et al. (2019; https://tc.copernicus.org/articles/13/879/2019/) highlighted the influence of mesh resolution on paleo ice retreat (albeit in 1 region of Greenland) and ice mass change. I would think 5km should be good enough, but in areas of very complex bed topo, or where gradient in the surface mass balance is important to simulate, higher resolution would have a larger impact on mass flux and comparison to ice margin migration."*

**Author response:**

We thank the reviewer for pointing this out. We agree that valuable space could be gained by making this section significantly shorter. We have thus removed ~60% of the text of that section (250 words instead of 590 originally). We have now also moved the relevant figure (ex Figure 27) to the supplementary materials, and copied the original full text of the discussion on this topic to the supplementary materials.

We however feel like this point is important to mention as it is often forgotten about in paleo ice-sheet modelling. Bed resolution matters in some places, and is likely to introduce biases. We are likely to realize this more and more in the coming years as modelling at higher resolutions becomes more and more computationally feasible. Our work on the LGM European Alpine Ice Field modelling with IGM (Leger et al., 2025) and an upcoming paper currently in review in The Cryosphere from Helen Werner et al. ,also with IGM, is providing some interesting new insights on this question, I believe. The CE and NE regions of Greenland as well as its southernmost tip present much rougher topographies and topographically-constrained iceflow conditions, at present, than other flatter regions such as the CW, NW, and NO regions, where the impact of bed resolution is thus expected to be lower.

**34) Reviewer:**

*"Line 2019: If you plot your reconstruction against data assimilated products (Badgeley et al., 2020; Buizert et al., 2018), you might find that the +/- 3.5 degrees Celsius offset is not enough to account for uncertainty in temperature."*

**Author response:**

Please also refer to our replies to Reviewer 1 comments 24, 25, and to Reviewer 2 comment 20: where we present our new analysis and comparison with the climate of Badgeley et al. (2020) and other datasets presented therein. We have now also added a reference to the new figure 25 panels a and c, which may supports the next sentence that goes in line with what is suggested by the reviewer here:

"Our experiment features an ensemble-varying temperature offset parameter (Table 1) with maximum space-independent warming of up to +3.5 °C, along with ensemble-varying snow and ice PDD melt factors that can reach 5 and 12 mm *w.e.* $d^{-1}$ $°C^{-1}$, respectively. However, if significant input climate biases exist in the regions of concern, these perturbations may still underestimate the resulting surface melt during deglaciation (see Fig. 25a,c)."

#########################################################################

# Reply to reviewer 2

*"This study explores how the Greenland Ice Sheet has changed over the past 24 kyr by running 100 high-resolution simulations and comparing them with real-world data on past ice margins. The results reveal that a major retreat happened between 16 kyr and 14 kyr years ago, mainly driven by ocean warming, even though the air was still too cold to cause much surface melting. No simulation from the ensemble is able to well represent both the Last Glacial Maximum (LGM) extent, deglaciation history and pre industrial extent of the ice sheet. Still, by identifying regional mismatches in model-data agreement, especially in northern and eastern Greenland, and by investigating the causes behind this mismatch, the study helps to improve how we simulate past ice sheet changes.*

*This is a well-written and well-structured study that clearly reflects the extensive work behind it. As a modeler I know how hard might be this kind of exercises and I congratulate the authors for their work. To my understanding, this is the first comprehensive paleo-modelling effort that simulates the full deglaciation of the Greenland Ice Sheet since the LGM. Importantly, it compares these simulations against a spatially distributed set of paleo extent records, which adds significant value and originality to the work. Undoubtedly, this work stands out as a valuable contribution to the field.*

*Still, I recommend some modification before being considered for publication in TC."*

## General comments:

### General comment 1:

"*The use of the terms "ice-sheet extent" or "margin" could benefit from clarification: do you refer to the grounding line or the ice shelf front? Where and when are ice shelves present in your simulations? This distinction is important, particularly given that sub-shelf melting appears to be a key driver of mass loss in the early Holocene. While most of the discussion and figures clearly focus on the grounded ice sheet, knowing the extent and presence of ice shelves is essential. For example, plotting the ice shelf extent in Figure 17 would be helpful. The presence of ice shelves could delay the onset of deglaciation due to their buttressing effect, and this deserves more attention*."

**Author response:**

We thank the reviewer for this valuable comment. Indeed, this could have been made clearer. Our modelling does indeed include floating ice shelves and calving, as detailed in section 2.1.2. of the sub-paragraph entitled "*The ice-atmosphere interface*". It is true that throughout the paper, which focuses on model-data comparison with geological evidence providing constraints on grounded ice extent only for the GrIS (e.g. PaleoGrIS), our post-processing analyses thus focuses on the modelled grounded ice extent, and floating ice is removed from model outputs during post-processing (for the purpose of model-data comparison only).
The extent of modelled ice shelves is however shown by Figures 21 and 22: in dark Grey colour. These figures display our model outputs for one of the overall best-scoring simulation between the local LGM and the PI. Our model outputs from best-scoring simulations, shared within our Zenodo repository, also feature the raw modelled ice thickness fields including ice shelves, and so do the videos we shared within the same Zenodo.
In response to this comment, we decided to add text to make this clearer at the very start of section 2.4 :"*Model-data comparison scheme*":

"Isolating best-fit ensemble simulations requires a quantitative assessment of model-data agreement on past GrIS behaviour. Here, each simulation is scored using three chronologically distinct tests, described below. Before testing, we remove the IIS and ice thinner than 10 m from modelled thickness fields. Because former GrIS ice-shelf extent is poorly constrained, and empirical datasets used here only constrain grounded GrIS extent, we also exclude floating ice (post-simulation) and restrict all ice-extent analyses to grounded ice for the remainder of the study. Modelled ice shelf extent at selected time periods is nonetheless shown in Figures 21 and 22."

In our best-fit simulations, modelled ice shelves are highly extensive during the local LGM, at around 16.5 kyr BP. However they mostly disintegrate by 15 ka (quite early), and remain rather small after that, and thus do not seem to delay the onset of deglaciation. As presented in new Figure 10, and in section 4.1 "*Model agreement with empirical data*" paragraph 1; "In all simulations, the onset of modelled GrIS retreat also occurs earlier than is suggested by PaleoGrIS 1.0, with an offset of nearly 2 kyr (Fig. 10).”. Thus, we do not believe modelled ice shelves yield a buttressing impact that delay the onset of deglaciation in our best-fit simulations, as the issue (and this is the case across the whole ensemble) seem to be the opposite: i.e. the onset of deglaciation occurs too early, and is almost entirely forced by oceanic warming during the late HS1 and early BA warming.

**General comment 2:**

“*The manuscript is quite long, and the detailed descriptions of individual simulations (e.g., the best LGM and best deglaciation simulations) could be streamlined. You might consider focusing on the top five simulations after the full selection process and moving some other figures and discussions to the Supplementary Material. This would help improve the readability and flow for the reader.*”

**Author response:**

We thank the reviewer for this comment. Indeed, we agree that some more effort could be made to reduce the length of the paper and improve readability. To address this, we have:

1-Almost entirely moved the discussion section on the potential biases from coarse model resolution to the supplements (following also a specific comment from reviewer 1) as well as its associated figure (old figure 26).

2-Accordingly to specific comment 7 from reviewer 2, we have reworked the climatic figures and merged old Figures 6 and 7 to a single figure on ocean temperature forcing. Old figure 11 was also removed and replaced (see our reply to reviewer 2 specific comment 12 below). This has thus reduced the number of figures (although more comments on adding the comparison with Badgeley et al. (2020) has also contributed to adding a new figure) such that the total number of figures with respect to original is now 25 instead of 27.

3-We spent a good amount of time going through the entire paper to make the writing more concise without loosing any of the important information, details, references, number etc: but instead just by rephrasing and removing redundancies in language. This was done after all other reviewer comments were addressed, and resulted in making the paper 1472 words shorter.

We did not find that we could move entire paragraphs regarding ensemble-wide results to the supplements (as is suggested to consider in the above comment) without loosing what we consider to be some of the most insightful results from our experiment and without breaking the paper’s structure. However, we believe changes presented directly above have now notably improved readability and flow.

**General comment 3:**

"*I'm surprised that sub-shelf melting is not explored within the ensemble, despite the large uncertainties in SST forcing. Unlike for surface air temperature, no parameters in the sub-shelf melt parameterization are varied, nor is any uncertainty in SST explicitly addressed. This is particularly striking given that the retreat at the onset of the last deglaciation is attributed primarily to oceanic forcing. Statements such as: "Between 16 and 14 thousand years ago, the ice sheet lost most of its ice grounded on the continental shelf. This marine-sector demise, associated with up to seven times greater mass loss rates than observed today, was predominantly caused by ocean warming while air temperatures possibly remained too cold to generate surface melt" seem premature in the absence of a more thorough investigation of the sub-shelf melt scheme and its sensitivity to oceanic conditions. For example, how robust is the timing of SST increases around 16 ka? This uncertainty deserves more discussion. See also my specific comments.*"

**Author response:**

We thank the reviewer for raising this point. In line with this comment which aligns with specific comments made by both reviewers, we have now added more text in several locations of the manuscript to acknowledge the potential biases of Osman et al's SST reconstruction, and also to acknowledge more clearly the fact that using SST instead of sub-shelf temperatures as forcing may not be ideal, but that there are clear justifications for why we chose this strategy, and to make more comparison with sub-shelf temperatures used as forcing by Tabone et al. (2024), as suggested by Reviewer 1.

Please refer to our replies to Reviewer 2 specific comments 4 and 18, and Reviewer 1 specific comments 8 and 13 for more details. We also guide the reviewer towards this paragraph in our discussion (section 5.1 paragraph 2) where we indeed agree with the argument raised here that our ensemble may have benefited from an extra ensemble-varying parameter introducing perturbations to the ocean forcing:

"A small underestimation in HS1 sea-surface cooling offshore NE Greenland, in the order of 1-2°C for instance, may be enough to deter the modelled GrIS margins from advancing extensively. This hypothesis may also be reinforced by the general lack of spatial coverage of SST proxy records used in the data-assimilation scheme of Osman et al. (2021) north of 65°N, offshore Greenland coasts. Biases may also be introduced by our interpolation scheme used for resampling from the nominal 1° horizontal resolution of the original data (Osman et al., 2021), equivalent to a ~20 x 27 km grid offshore NE Greenland, to our 5 x 5 km model grid. This highlights that our modelling experiment may be limited by a lack of variation in SST input fields between ensemble simulations. A future experiment using an ensemble-varying parameter introducing spatial and temporal perturbations to the input ocean forcing may help test this hypothesis and possibly increase model-data fit."

Regarding this comment : "*no parameters in the sub-shelf melt parameterization are varied*".

One of our ensemble-varying parameter does vary the impact of calving and thus the extent of ice shelves, between simulations. However we fully agree that in future ensemble experiments of this nature, bringing perturbations to the ocean forcing and thus more directly sub-shelf melting as an ensemble-varying parameter, would be very important. This is for us a lesson of our study, not something we knew in advance. We admit that its impact was underestimated during our ensemble experiment design phase, prior to launching the runs, at the stage when we had to be restricted to 10 ensemble-varying parameters for computational cost reasons (our 100 member ensemble took 1.5 years to compute on HPC…). Such an experiment cannot, unfortunately, be re-run easily.

## General comment 4:

*"Many of the model-data misfits likely stem from uncertainties in the climatic forcing. I won't elaborate further on this point, as it has already been discussed in detail by another reviewer. However, I do agree that it would be helpful to compare your climate forcing, particularly air temperature and precipitation, with that used in Badgeley et al. (2020). This could provide useful context and help assess the robustness of your forcing choices (see also specific comments)."*

**Author response:**

We thank the reviewer and agree that adding this comparison represents a valuable addition to our discussions. Please see our responses to the Reviewer 2 specific comment number 20, in which we describe the additional analysis and figure, and text added to the paper and which provides a direct comparison with the climate reconstruction of Badgeley et al. (2020). Please also refer to our replies to Reviewer 1 specific comments 24, 25, 28, 31, 34 (for discussion on potential climate biases).

## General comment 5:

"*Your simulations do not show an expansion of the GrIS in the northeastern sector during the LGM, despite several studies suggesting that the margin likely reached the continental shelf break. It's true that many models struggle to reproduce this feature, and you do acknowledge this limitation in the text. However, I believe the discussion could be strengthened by incorporating some of the additional points I've outlined in the specific comments."*

**Author response:**

We thank the reviewer for this insightful comment. Please refer to our replies to Reviewer 2 specific comments 13, 14, 15, and 17, where we present our newly added text to strengthen the discussion on this point.

## Specific comments:

***1) Reviewer:***

*"-Line 233: please, write the equation with \cdot."*

**Author response:**

Thank you for spotting this mistake. This change to equation 1 was made accordingly.

***2) Reviewer:***

*"-Line 255: Could you describe the simple hydrologic model from Tulaczyk et al., 2000 with more detail? how is the N_till calculated?"*

**Author response:**

We thank the reviewer for this comment, following which we added more details to the description of the subglacial hydrology model including an extra equation:

"We account for space- and time-dependent basal yield stress, $\tau_c$, controlled by, firstly, a simple hydrology model (Tulaczyk et al., 2000) which determines the effective pressure, $N_{till}$, from the till-pore water content obtained by storing basal melt locally up to a threshold (here set to 2 m). With this simplified parameterisation, water is not conserved as water reaching above the threshold is lost permanently. The basal water thickness in the till layer, $W_{till}$, is computed from the basal melt rate, $m_b$, obtained from the enthalpy, as follows:

$$\frac{\partial W_{till}}{\partial t} = \frac{m_b}{\rho_w} - C_{dr}, \qquad (3)$$

where $C_{dr}$ is a simple decay rate parameter and $\rho_w$ is the density of fresh water."

***3) Reviewer:***

*"-Line 375: Please add in this paragraph that temperature and precipitation used to force the SMB model will be described in section 2.1.3."*

**Author response:**

Thank you for this valuable suggestion. We have now added this information within the relevant paragraph:

"To compute Surface Mass Balance (SMB) from two-dimensional fields of time-dependent reference height temperature and precipitation (see section 2.1.3), we use PISM's default Positive-Degree-Day (PDD) model (Calov and Greve, 2005; Ritz, 1997). Precipitation when temperature is above 2 °C and under 0 °C is interpreted as rain and snow, respectively, with a linear transition between. Temperature and precipitation fields used to force the SMB are further described in section 2.1.3. The fraction of surface melt that refreezes is set to 60% (EISMINT-

Greenland value; Ritz, 1997). Spatio-temporal variations in the standard deviation, $\sigma$, of daily temperature variability influences SMB (Arnold and MacKay, 1964)."

***4) Reviewer:***

*"-Line 391: in section 2.1.3 you describe the SST and salinity signal used to force the sub-shelf method, which is only briefly described here. Please, add a section describing the sub-shelf melting scheme in more detail (the basics of the three equation formulation of the ocean fluxes, if I understand correctly) and how are the SST and salinity taken into account in this formulation. I guess that what's especially interesting here is the fact that this method takes the SST as a forcing to compute temperature and melting at the base of the ice shelf. "*

**Author response:**

According to this comment, we added more details to the section describing the sub-shelf melt model and its three equations. We however assume that writing out all equations of the many sub-models used in our PISM setup for mass exchange processes would make the methods section far too long, and thus refer to the original papers for more details:

"For floating sectors of the modelled GrIS, sub-shelf melt is obtained by computing basal melt rate and temperature from thermodynamics in a boundary layer at the ice shelf base (Hellmer et al., 1998; Holland and Jenkins, 1999). This model, which does not consider sub-shelf circulation, uses three equations describing: 1) the energy flux balance, 2) the salt flux balance, and 3) the pressure- and salinity-dependent freezing point in the boundary layer. This sub-shelf melt parameterisation thus requires time-dependent two-dimensional fields of potential temperature and practical salinity (see section 2.1.3.). More details can be found in Hellmer et al. (1998) and Holland and Jenkins (1999)."

Please note that this sub-shelf melt model can take any 2D input ocean temperature forcing field: which could thus be sub-shelf temperatures as opposed to SST. However, we here chose to use SST as we do not have time- and space-dependent sub-shelf ocean temperature reconstructions which assimilated proxy data and that is transient between 24 ka and the PI: which we have for SST at 200 yr temporal resolution thanks to Osman et al. (2021)'s study. However we agree this is not 100% ideal as SST and sub-shelf temperatures can differ. We agree with the reviewer that this was not made clear enough in the text. In line with this, we added more text to justify this decision to section 2.1.3.:

"To compute sub-shelf melt, the chosen parameterisation (Holland and Jenkins, 1999) requires time-varying two-dimensional fields of potential ocean temperature and salinity (see section 2.1.2). For the ocean temperature, we use the LGM-to-present ensemble-mean sea surface temperature (SST) reconstruction of Osman et al. (2021), yielding a 200-year temporal resolution and nominal 1° spatial resolution (Fig. 6). This re-analysis uses Bayesian proxy forward models to perform an offline data assimilation (using 573 globally-distributed SST records) on climate model priors; *i.e.* a set of iCESM 1.2 and 1.3 simulations (Zhu et al., 2017; Tierney et al., 2020). Whilst we acknowledge sub-shelf ocean temperature would be a more appropriate forcing than SST, their does not yet exist a Greenland-wide time- and space-dependent sub-shelf ocean temperature reconstruction which assimilates proxy data between

24 kyr BP and the PI. The transient and data-assimilated nature of the SST reconstruction by Osman et al. (2021) was thus preferred to iCESM outputs of shelf-depth ocean temperature (e.g. Tabone et al., 2024)."

***5) Reviewer:***

*"-Line 423: Figures are generally far away in the text from where they are cited. Please, try to put them closer. Also, how do you "combine" the two GIA (local + non-local) signals? Simply by adding them up?"*

**Author response:**

Noted, we will try to re-arrange the figures' locations to minimize this issue, in particular during the eventual proof editing stage. Regarding how the two GIA signals are combined: we added more text to the end of the relevant paragraph clarifying this. The key is that the non-local GIA signal obtained by running an offline model prior to our PISM simulations is used to produce an independent input field of sea level offsets from the present-day sea level (at 500 yr temporal resolution). These offsets are then read by PISM which also incorporates the local bed deformation from its own GIA model updated with the evolving ice sheet; to obtain the final relative sea level used to calculate the floatation criterion (where is ice grounded vs floating):

"For the local GrIS signal, we use PISM's Lingle-Clark-type viscoelastic deformation model (Lingle and Clark, 1985; Bueler et al., 2007). […] To calculate the non-local sea level change across the region of interest, we run an offline GIA model. […] . This offline model is used to produce input two-dimensional sea level offsets from the present-day sea level between 24 kyr BP and the PI, at 500 yr temporal resolution. PISM uses these offsets to compute the final relative sea level after computing the local GIA deformation."

***6) Reviewer:***

*"-Line 483: what do you mean by "reference height air temperature and precipitation"? Do you mean you apply a lapse rate to correct the surface air temperature with respect to the topography? How is this lapse rate applied to precipitation if there is one?"*

**Author response:**

Thank you for spotting this mistake. The term "reference height" was initially used as it is the default name of the iCESM output variable used for the surface temperature forcing: 'TREFHT' (temperature at reference height): which refers to the temperature at a given height above the surface: usually set to 2 m above surface in CESM). Although I had initially described the temperature data using this specific term, after comments from co-authors, I had removed this term which was a bit overkill from the manuscript and figures altogether. However I must have forgotten to remove it from a few locations. I have now removed it from all the locations across the manuscript I had missed to simply write "surface air temperature" instead. In our PISM model setup we do not use lapse rate corrections as our 2D input fields of Temperature and precipitation are from CESM runs which includes an evolving surface topography, although there are potential biases from this surface being off, and which we mention in our discussion.

***7) Reviewer:***

*"-Figures 4-8: I would suggest to restructure the figures of climatic forcing so that they are separated from those from sea surface temperature. In that case you'd have figure 6 with precipitation only or you could add a third panel to figure 4 with precipitation fluxes and fig. 6 would only have SST."*

**Author response:**

We thank the reviewer for this valuable suggestion. Accordingly, we have now changed the figures such that figure 4 also features the precipitation fields panels and figure 6 rather only features the SST data but with also the time series from old Figure 7: resulting in the merging of old figures 6 and 7, and thus reducing the total number of figures and length of the paper.

***8) Reviewer:***

*"-Line 488: which is the temporal resolution of the transient iTRACE experiment?"*

**Author response:**

The iCESM data made available by NCAR from the iTRACE experiment has a monthly temporal resolution, which allows us to produce mean annual and mean summer products (JJA-mean). We added this information to the relevant line using the term 'monthly-resolution':

"Between 20 and 11 kyr BP, we use monthly-resolution data from the iTRACE experiment, ran with iCESM 1.3 (He et al., 2021a, b)."

***9) Reviewer:***

*"-Line 495: Buizert et al., 2018 provides transient spatially variable fields for temperature since 21 kyr BP. How do you build the 1D glacial index from these 2D fields? Do you compute a different climatic index for every 5x5 km cell across the transient 2D fields, otherwise I don't get how would you end up with a different signal for each location (Figure 5). Please specify. Also, could you show the precipitation time series for different ice core locations using the same index approach? How does this compare with Badgeley et al., 2020?"*

**Author response:**

We thank the reviewer for raising this important point. Indeed, here we use the spatially-distributed data from Buizert et al., (2018) to build 1D climatic indexes for each 5x5 km cell of our 2D fields. To make this clearer, we added further explanations to the relevant methods paragraph:

"To create continuous forcing over remaining data gaps, we use a glacial index approach (Niu et al., 2019; Clark et al., 2022) and linearly scale our climate fields proportionally to variations in independent climate reconstructions in a space-dependent manner i.e. building a glacial index for each individual grid cell (Fig. 5)."

In line with the second comment made, we added a third panel to Figure 5 which shows the time series for different ice core locations but for the mean precipitation rate input data as well. This is a nice addition to our figures and we thank the reviewer for providing this suggestion.

Regarding comparisons with the reconstruction from Badgeley et al. (2020), please refer to our reply to Reviewer 1 comment number 20, in which we describe the additional analysis and figure, and text added to the paper and which provides a direct comparison with the climate reconstruction of Badgeley et al. (2020). Please also refer to our replies to Reviewer 1 specific comments 24, 25, 28, 31, 34 (for discussion on potential climate biases).

***10) Reviewer:***

*"-Line 512: surface salinity: do you mean here you take the 2D equilibrium iCESM simulations for 21, 11, 9, 6, 3 kyr BP and PI and interpolate them linearly?"*

**Author response:**

Yes that is correct, and this is because we do not have an independent proxy-based reconstruction from of ocean salinity and for these timescales around Greenland that we can use as an index scheme. We modified the relevant sentences to add more clarity and details:

"For ocean surface salinity, we use iCESM outputs, following the same methodology as described above. We however use linear interpolation rather than a glacial index scheme to bridge the temporal data-gaps in salinity data, which are located outside of the transient iTRACE data (20-11 kyr BP) and equilibrium iCESM simulations (21, 9, 6, 3 kyr BP and PI)."

***11) Reviewer:***

*"-Line 559: Fixing the parameters to their mid values for the spin-up and then vary them for the ensemble transient runs might bring to some inconsistencies in the first years of simulation after 24 kyr BP. I believe this is not a crucial point, but I would like to see a sentence that discusses this."*

**Author response:**

Many thanks for raising this important point: we agree this can be a limitation of using a single spinup which is however the best we could computationally afford given the chosen resolution and size of the transient ensemble. We added two sentences to the end of the relevant paragraph (2.2) to mention this:

" This spun up GrIS is used as the initial condition for all ensemble transient simulations. The 30 kyr equilibrium spinup limited us computationally to this single initial state at 24 kyr BP with ensemble-varying parameters fixed to mid-range values. Although adjusting parameters in subsequent transient runs can generate instabilities in the first simulation years, equilibrium with parameterisations is likely reached within the first centuries and should not significantly affect the modelled local LGM or deglacial dynamics."

***12) Reviewer:***

*"-Figure 11: could you highlight the best 5 simulations in both panels? Could you also add a mark for the present grounded area (1.7 10^6 km^2)?"*

**Author response:**

We thank the reviewer for this comment. The additional data that is here suggested to add to Figure 11 was initially presented within Figure 22 (very far in the paper due to when it was referenced to), whilst figure 11 was meant to be the raw ensemble results only. In hindsight it appears there is no significant added value for having two separate figures here. Following the suggestion, we thus decided to simply replace Figure 11 by the information in Figure 22 whilst making sure to keep the time information of the various events mentioned in the text which was only included originally in Figure 11 (LGM, local LGM, Late-Glacial etc). This modification also served the purpose of removing a figure, making the paper a bit shorter. Note that "Figure 11" in this comment reply has now also changed to Figure 10 due to addressing other comments.

***13) Reviewer:***

*"-Line 849: please add the reference O Cofaigh et al., 2025 "Shelf-edge glaciation offshore of northeast Greenland during the last glacial maximum and timing of initial ice-sheet retreat" which further supports the maximum extent of the NE Greenland during the LGM. "*

**Author response:**

Thank you for suggesting this valuable addition from this new paper and set of evidence. We have now added it to the references associated with this sentence.

***14) Reviewer:***

*"-Figure 12: I think "basal mass fluxes" should be better replaced by "sub-shelf mass fluxes" to avoid confusion between that and grounded basal mass balance."*

**Author response:**

We thank the reviewer for this valuable insight and have, as a result, modified Figure 12 (now figure 11), its caption, and the relevant text to make sure to replace the term "basal mass flux" by "sub-shelf mass flux" as we here indeed show the data for basal mass flux of floating ice exclusively, and not of grounded ice. Indeed, grounded basal mass fluxes are of significantly lower magnitudes and only account for a small fraction of the GrIS mass changes throughout our simulations, with maximum ensemble-wide flux values reaching -250 Gt $yr^{-1}$ (less than 8% of the maximum ensemble-wide sub-shelf flux values).

***15) Reviewer:***

*"-Line 1077-1081: I would be careful about this paragraph. You don't model a central ice divide migration for NGRIP, but your simulated LGM maximum extent in the NE is significantly underestimated. In fact, as you noted earlier and as several recent publications suggest, the expected extent should likely reach close to the continental shelf break. Although previous modelling work seems not to suggest that the NGRIP summit was migrated during the LGM (e.g Tabone et al., 2024), this discrepancy might imply a more stable and less elevated ice divide in the central-north central region of the ice sheet, even if that may not have been the case. Also, a thickening of the NE stream at the LGM is found from geomorphological records (Lane et al., 2023). This is somehow seen in figure 15, panel b, but it might well be underestimated due to the limited margin extent. Please, add some comments on this."*

**Author response:**

We thank the reviewer for this excellent comment: we have now added more text to the end of this paragraph to put the previous statements within the context of our model-data misfits in the NE and of the results of Tabone et al., 2024:

"Therefore, towards both central (GISP2, GRIP) and northern (NGRIP) GrIS summits, our model results suggest that the local LGM GrIS was not necessarily thicker than today (Fig. 15). A lack of NGRIP summit migration during the LGM was also suggested by the modelling work of Tabone et al. (2024), thus implying a possibly more stable ice divide during glacial-to-interglacial transitions in the central and northern regions of the GrIS, relative to other regions. However, we must remain cautious regarding our results in the NE GrIS region, as our lLGM best-fit simulations still substantially underestimate maximum grounded ice extent in this sector (more discussions in section 5.1.)."

***16) Reviewer:***

*"-Figure 15: could you make the maps within the panels bigger?"*

**Author response:**

In line with this reviewer comment, we have now maximized the sizes of the inset panel maps within each panel of Figure 15 (now figure 14): and have also expanded the size of each panel by removing unnecessary white space, thus hopefully making the figure details easier to visualize.

***17) Reviewer:***

*"-Lines 1120-1134: yes. I would add some comments on the ice discharge in the northeast. Although recent radar measurements suggest that the upper part of the present-day NEGIS was fully developed only during the last 2000 years (Jansen et al., 2024), there is the evidence at least of a paleo ice stream that was flowing before and likely into the Holocene in the northeast (Franke et al., 2022) and this is not captured in your LGM simulations. This suggests that the northeast region of the ice sheet could have been more dynamic as your simulations show."*

**Author response:**

In line with this comment, we have now added an extra sentence to the relevant paragraph that incorporates this valuable extra information:

"During the local LGM, our best-fit simulations model GrIS-wide discharge rates that reach between 1500 and 1900 Gt $yr^{-1}$ (Fig. 19). Such discharge rates are between ~2.8 and ~4.3 times greater than those estimated for the present-day (487 ± 50 Gt $yr^{-1}$ between 2010 and 2019 AD; Mankoff et al., 2020). Moreover, these figures are likely underestimates as our lLGM best-fit simulations do not produce any extensive paleo ice stream in the NE and contemporary NEGIS region whilst there is evidence from radar measurements of widespread streaming during the Holocene in this GrIS region (Franke et al., 2022; Jansen et al., 2024). These higher lLGM GrIS discharge rates have implications for discussing past iceberg production volumes, the contribution of the GrIS to past Heinrich events, and its potential role in former and future AMOC slowdowns (Ma et al., 2024)."

***18) Reviewer:***

*"-Line 1457: as you pointed out earlier, this is due to the onset of sub-shelf melting around 16 kyr BP in some regions? Could you add some comments on the reliability of the temperature forcing in these regions from Osman et al., 2021? Is there any comparison you could make between these simulations and available paleo records, or was it done by Osman et al.?"*

**Author response:**

We thank the reviewer for this comment. In line with this, we added a sentence to indeed raise the possibility that this misfit in timing could also be the consequence of biases and uncertainties in the reconstruction of SST by Osman et al., (2021), which we detail further in section 5.1.. Their reconstruction did assimilate all available paleo records of SST in the region though, but perhaps not as a hard constraint.

"In addition, the onset of modelled GrIS retreat occurs ~2 kyr earlier than suggested by PaleoGrIS 1.0 (Fig. 10). […] As our results show the onset of modelled GrIS retreat during late HS1 and B-A is primarily controlled by sub-shelf melting (see section 3.2.1.), this offset in retreat timing may also reflect uncertainties and biases in the SST reconstruction (Osman et al., 2021; Fig. 6) used as ocean temperature forcing (see section 5.1. for more discussion)."

***19) Reviewer:***

*"-Line 1538: I am not surprised that thinning can be fairly well reproduced for the last 8 kyr BP as generally what is hard to replicate/explain is the thinning at the early Holocene due to the demise of the IIS, LIS, as well as big uncertainties in the climate forcing and ice-sheet dynamics. Please add a sentence commenting on this and citing Lecavalier et al., 2017 (for Camp Century) and Tabone et al., 2024 (for NGRIP)."*

**Author response:**

We thank the reviewer for raising this comment, and have now added two extra sentences to the end of the relevant paragraph to make sure to mention that the 12-8 kyr interval ice thinning histories are expected to be more challenging to replicate with models:

"This finding suggests that best-scoring simulations isolated by model-data comparison using detailed ice-extent reconstructions tend to also result in appropriate Holocene GrIS thinning signals. However, although some ensemble simulations are clearly not in agreement with the ice core - derived thinning curves (Lecavalier et al., 2013), the majority of the ensemble remains close to, or within, the 1$\sigma$ uncertainty bands of these data (Fig. 23). It must also be noted that whilst we here focus on the 8 - 0 kyr BP interval, GrIS thinning histories during the early Holocene (12 – 8 kyr BP) are known to be more challenging to both i) replicate in models and ii) correct for in original ice-core derived data (Lecavalier et al., 2017; Tabone et al., 2024). This is due to the demise of the LIS and IIS and unzipping from the GrIS during this interval, and the important impacts of these events on GrIS thinning and bed isostatic adjustment."

***20) Reviewer:***

*"-Line 1638: I would add here a discussion on how your transient precipitation forcing compares with Badgeley et al., 2020, as they suggest different low- to high- precipitation scenarios for the early Holocene."*

**Author response:**

We thank the reviewer for this suggestion. There are several reviewer comments from both reviewers 1 and 2 suggesting to make comparisons of our climate forcing with the ice-core-derived and TrACE-21-derived temperature and precipitation reconstructions of Badgeley et al. (2020). We here provide an answer that attempts to address these several comments together, to minimize redundancy in our reply.

We have, in response to these comments, produced a new data-processing analysis and a new Figure (25) which has been added to the end of the main manuscript and which makes a direct comparison of our mean annual temperature and precipitation forcing with the data presented in Badgeley et al. (2020) figures 8 and 13 (which also includes data from TRACE-21ka and Buizert et al., (2018): and which combined show data for both the Kangerlussuaq region and GrIS summit.  Please find here the new caption for this figure:

"Comparisons between our input mean annual temperature and precipitation forcings (orange time series) with the climate reconstructions of Badgeley et al. (2020), Buizert et al. (2018), and raw TraCE-21ka data (Liu et al., 2009). More specifically, these panels present the same data as shown in Figures 8 and 13 in Badgeley et al. (2020). Note that precipitation fractions and temperature anomalies are here expressed with reference to the mean of 1850–2000 AD for all datasets except this study's input climate data (orange), instead expressed with reference to the mean of 1750-1850 AD, caused by our most recent iCESM simulation and datapoint being 1850 AD."

In two paragraphs of section 5.1: which discusses potential reasons for remaining model-data misfits and already mentioned biases in our climate forcing as a possible explanation, we have now added more sentences that links to this new comparison with Badgeley et al., (2020) and new figure: we present these additions directly below:

-Section 5.1, paragraph 4:

"Our simulations may also underestimate grounded ice extent in the NO and NE due to too low accumulation rates, largely controlled by our input precipitation forcing. Throughout these regions, iCESM-derived forcing suggests precipitation rates below 20 mm per month during HS1 (Fig. 6). We note that although iTRACE […] may still be subject to CESM biases that can sometimes misrepresent present-day and former precipitation rates over certain GrIS regions (van Kampenhout et al., 2020; Lofverstrom et al., 2020). Although we use an ensemble-varying parameter introducing precipitation perturbations of up to +200% (Table 1), such an increase is not space-dependent and may still be too low over NE Greenland. […] Thus, better model-data scores at the *local-LGM extent* test could potentially be achieved with precipitation offset values above +200%. We compared our precipitation forcings with the paleoclimate data assimilation reconstruction of Badgeley et al. (2020), who extended ice-core derived climate reconstructions across Greenland using TRACE-21 (Liu et al., 2009), and also made comparisons with raw data from TraCE-21ka and Buizert et al. (2018)'s reconstruction. This analysis indeed suggests notably lower precipitation rates in our iTRACE-derived climate forcing during HS1, and this in numerous regions across Greenland (Fig. 25b)."

-Section 5.2, paragraph 2:

"As mentioned above, biases in iTRACE-derived climate are possible, especially towards the margins of the former GrIS. For instance, an overestimation of the ice thickness and extent reconstruction used as forcing within iCESM (ICE-6G: Peltier et al., 2015) during the last deglaciation in NO, NE, and CE Greenland, would lead to unrealistically high albedo feedbacks impeding the atmospheric warming required to model appropriate GrIS thinning and retreat rates. Our experiment features an ensemble-varying temperature offset parameter (Table 1) with maximum space-independent warming of up to +3.5 °C, along with ensemble-varying snow and ice PDD melt factors that can reach 5 and 12 mm *w.e.* $d^{-1}$ $°C^{-1}$, respectively. However, if significant input climate biases exist in the regions of concern, these perturbations may still underestimate the resulting surface melt during deglaciation. We note that a cold temperature bias during the Late-Glacial and early-to-mid Holocene is not supported by comparison against the climate reconstruction (and its associated uncertainty range) of Badgeley et al. (2020), which instead suggests that our forcing produce relatively warm mean annual temperature anomalies towards the GrIS summit and NO, NE, and CE GrIS regions, between 15 and 5 kyr BP (Fig. 25c). On the other hand, this comparison reveals that our iTRACE and iCESM – derived climate forcing results in significantly higher (up to ~100%) precipitation rates during the entire Holocene towards the GrIS summit and its vicinity, than is obtained in ice-core-data-informed reconstructions from Badgeley et al. (2020) and Buizert et al. (2018) (Fig. 25d). Although our experiment features an ensemble-varying precipitation offset scheme with possible reductions down to 20% input precipitations, this potential bias may be responsible for too high Holocene precipitation in many of our ensemble simulations, thus impeding GrIS retreat in certain regions and causing ice extent overestimations during the modelled deglaciation (Fig. 25d)."

***21) Reviewer:***

*"-Lines 2045-2054: there is a modelling work able to reproduce the retreat of the NE Greenland since the LGM fairly well, from a fully expanded NE Greenland to its present day margin (Tabone et al., 2024). In their work the onset of the paleo NEGIS at the early Holocene is key for the NE retreat. Some comparison between your work and theirs could be useful to investigate the causes of your run-data misfits during the deglaciation."*

**Author response:**

We thank the reviewer for this valuable comment and insight. We agree that adding more reference to the work of Tabone et al. (2024) represents a valuable addition. In reply to this comment and also other comments from both Reviewer 2 and Reviewer 1 arguing in favour of more comparison between our results and the study by Tabone et al. (2024), we have now added more text to our manuscript in several locations, and which makes such comparisons. Please refer to our replies to Reviewer 2 specific comments 15 and 19, and to Reviewer 1 specific comments 8, 13, and 23: for more details regarding these changes.

Regarding this specific paragraph of the discussion on NEGIS, here mentioned by Reviewer 2 in this specific comment (21), we have now added an extra sentence on the results of Tabone et al. (2024):

"Due to its significant impact on ice flux of the entire NE GrIS region, modelling an accurate NEGIS configuration throughout the Late-Glacial and Holocene periods would produce higher regional-mean discharge and thinning rates. Over millennial timescales, this may help model greater and more data-consistent GrIS margin retreat rates during deglaciation. This is moreover supported by the results of Tabone et al. (2024) which suggest that an early-Holocene activation of a present-like NEGIS, achieved through highly targeted parameterization of low basal friction along the ice stream, is crucial to drive deglacial ice thinning over the central and northern GrIS. Therefore, it is likely that not fully reproducing NEGIS may contribute to increasing model-data misfits in NE Greenland relative to other GrIS regions, where ice streams are generally less challenging to model accurately."

***22) Reviewer:***

*"-Robert et al., 2024a and Roberts et al., 2024 b are the same publication. Same for O Cofaigh et al., 2023 a and b. Please correct."*

**Author response:**

We thank the reviewer for spotting these mistakes. Changes were made accordingly. These were due to errors from the Mendeley MS Word plugin.