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 δ^{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."

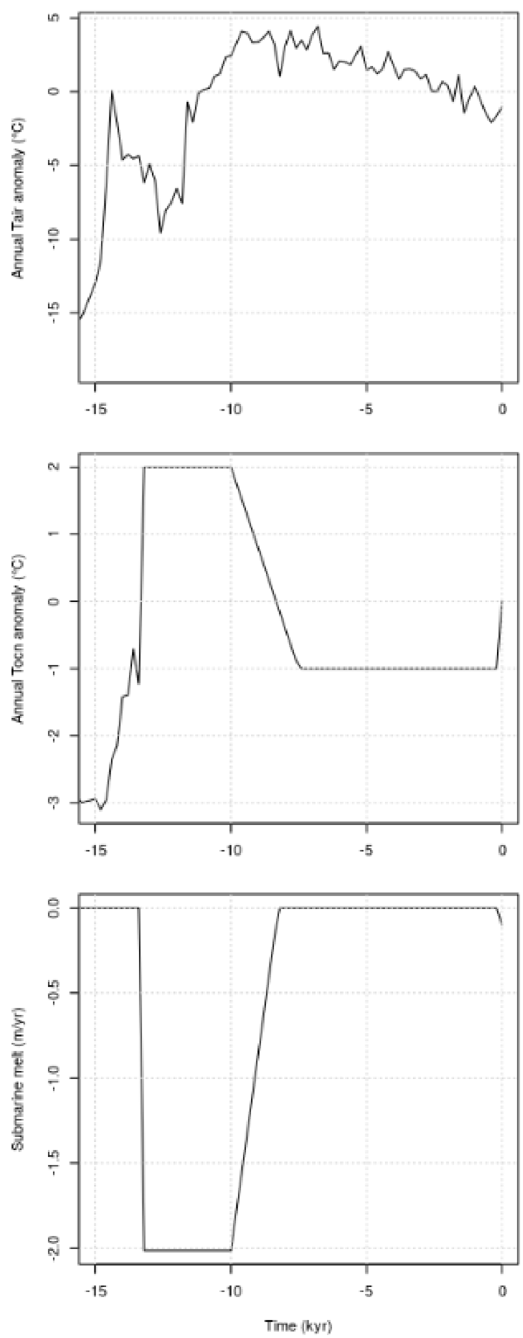


Figure S3. Atmospheric temperature anomaly (upper panel), ocean temperature anomaly (middle panel) and submarine melt (lower panel) used to force the ice-sheet model in the northeast outer shelf.

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)."

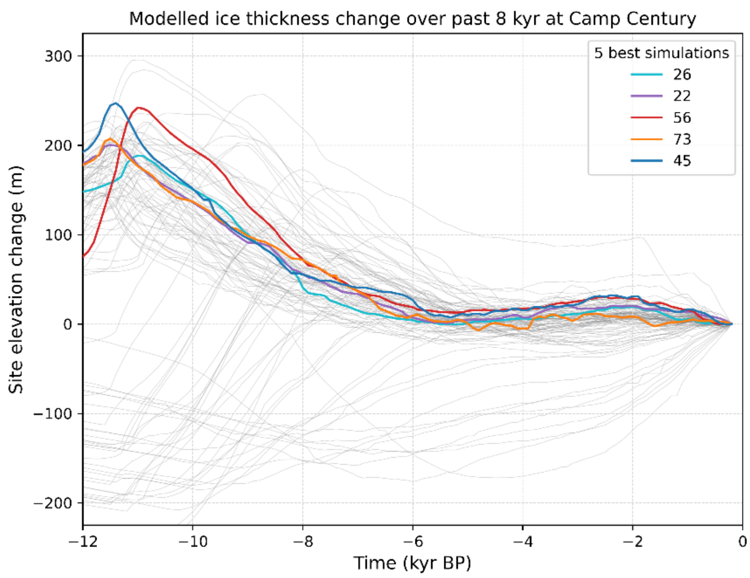
22) Reviewer:

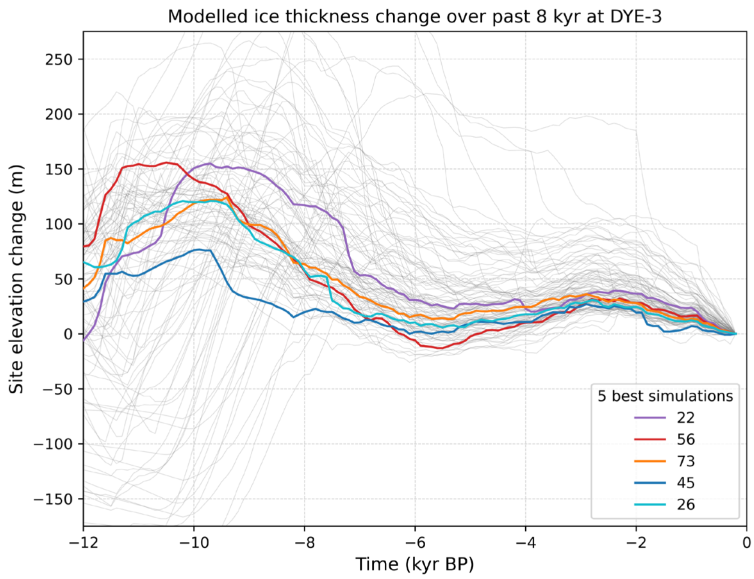
"*Lines 1521-1544: The thinning curves go back to ~11.5 ka. Any reason why the analysis was only for the 8kyr to 1850CE period and not back further since this section does include discussion about the early Holocene? It would have been nice to see if any models simulated the early Holocene thickening.*"

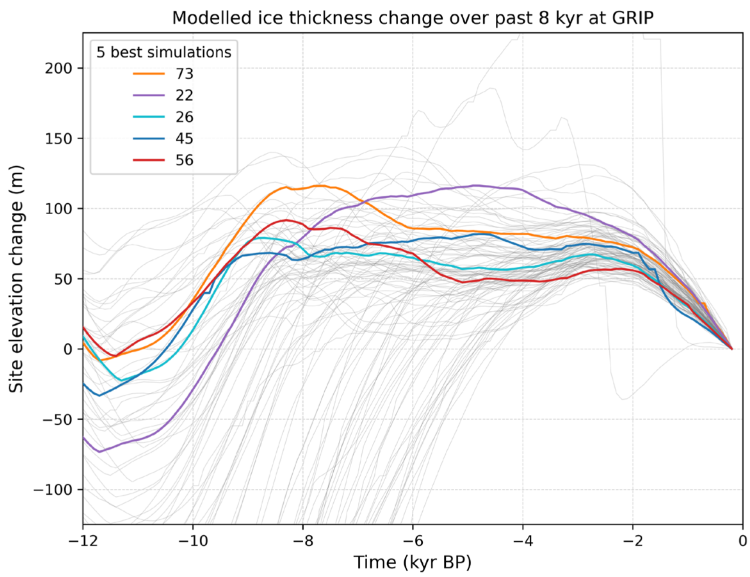
Author response:

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
Site elevation change (m)
Time (kyr BP)
5 best simulations
26
22
56
73
45
300
200
100
0
−100
−200
−12
−10
−8
−6
−4
−2
0


Modelled ice thickness change over past 8 kyr at DYE-3
Site elevation change (m)
Time (kyr BP)
5 best simulations
22
56
73
45
26
250
200
150
100
50
0
−50
−100
−150
−12
−10
−8
−6
−4
−2
0


Modelled ice thickness change over past 8 kyr at GRIP
Site elevation change (m)
Time (kyr BP)
5 best simulations
73
22
26
45
56
200
150
100
50
0
−50
−100
−12
−10
−8
−6
−4
−2
0

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)."