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, τ_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:

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

where C_{dr} is a simple decay rate parameter and ρ_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, σ, 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 1σ 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.