Response to the review of the manuscript "Analysis of the Surface Mass Balance for Deglacial Climate Simulations" submitted to The Cryosphere.

We thank both reviewers for their comprehensive reviews and very useful suggestions to improve the manuscript. We have addressed all of their comments and believe that the changes will significantly improve the manuscript. In the following, reviewer comments are highlighted in blue, author responses in black.

Anonymous Referee #1

General comments

I did enjoy reading this, and it feels like I've spent a lot of time writing it up – apologies for the lateness - I found lots of small things to ask about, see below.

We thank the reviewer for his very comprehensive review and are glad for the very good suggestions to improve the motivation and framing of the manuscript.

I felt there was one wider, general issue that hopefully would be easy enough to address - as I got to the paleoclimate section I didn't think that a solid motivation had really been given for /why/ a simulation of deglacial SMB was being done, and thus I wasn't really sure what to be looking for in how it's been described and analysed. Even by the end this isn't clear, as the Conclusions don't conclude anything that sounds obviously new about the glacial climate system or the model itself. In a couple of places it's said that the SMB fields will be made available to the ice sheet community for their use, but the first part of section 4 says that the simulation isn't going be evaluated, so I don't really buy this as sufficient motivation for the whole exercise - are these modelling groups going to want to use a new SMB product from a climate simulation that hasn't been evaluated for how it compares to evidence of what actually happened?

We agree that the motivation of this analysis did not come out clearly in the manuscript. We changed the abstract and introduction to state our motivation more clearly. The main motivation for our study is that for long term studies of ice sheet changes (fully coupled climate-ice sheet or ice sheets model only simulations) realistic SMBs derived from long Earth System Model (ESM) simulations are needed. Here, we present an EBM that is used to downscale coarse resolution ESM output. We evaluate the EBM and apply it to the first transient simulations of the last deglaciation with the MPI-ESM earth system model and prescribed ice sheets. A misunderstanding is that these simulations are not evaluated. We are currently undertaking an analysis regarding the sensitivity of the experiments to differences in the boundary forcing (Glac1D vs Ice6G). A publication of the results is expected in the near future.

There is some interesting decomposition of how accumulation and melt factors balance differently in different periods and how the (perhaps non-obvious) relationship between total SMB and ELA changes too, but only for GrIS - is it different in different regions? - and no attempt is made at general statements about the global implications or testable hypotheses for underlying principles. Personally, I wanted to know more about how some of the details of the snow physics were being forced by the climate through the deglaciation - albedo, refreezing etc, rather than just the resulting SMB, but that's just one idea.

We shortly discuss these changes in the section for the Laurentide and Fennoscandian ice sheets. However, there are different processes acting across these ice sheets as they respond to both changes in the Atlantic and Pacific. So the relationships are not uniform over the ice sheets. We believe that part of this comment is also related to the confusion of the melt definition (see comment to line 252) - we now defined the melt explicitly in the manuscript, which includes refreezing processes. We also revised the introduction and specifically the motivation for this study, which is to simulate a realistic SMB for long-term studies of ice sheet and climate interactions. Thereby, we hope to have addressed the concerns of the reviewer.

I think the authors should either decide on a clear main goal that's stated at the front of the paper and runs all the way through, or be more clear about the separate aims of each part - either way that should make it easier to see what the paper as a whole is working towards, and clearer what should be in the final conclusions.
We revised the introduction and parts of the conclusion to emphasize our motivation and the aim of the present paper.

Detailed comments:
line 4: "deglacial climate" - quite specifically the /last/ deglaciation
We changed it to 'the last deglaciation'

line 5: "allows to resolve" isn't grammatically great - 'allows the resolution of' perhaps?
Thanks, we changed this.

line 8: The flow of ideas implied by the: 1) <multiple sentences> 2) structure isn't very clear, I'm not sure it's needed
We removed this and rephrase slightly.

line 10 (and throughout): I found the dating notation used throughout off-putting, personally, but that might just be me. In my experience, the convention I've seen most is that kyr = "thousand years" and ka = "thousand years ago" with 1950 usually taken as the reference for 'ago'. "ka BP" thus feels to me like it's mixing conventions and providing both an 'ago' and a 'before'. I could live with the use of "ka BP" if defined here carefully (and used consistently throughout, but sometimes just "ka" seems to be used to indicate an event date, not a period of time - I'm looking at line 447 as I type this, and table 2 has "kas BP"), but I'm afraid the use of just "a" as the shorthand unit for years (and is that actually meant as 'years before 1950'?), spaced away from the number it's attached to so it can be mistaken for the indefinite article 'a', just grated horribly for me - eg line 455, line 359. It looks OK for eg the SMB numbers as (Gt aˆ-1) - to avoid confusion perhaps write the full word 'years' in cases like line 359 where "a" is currently used on its own? Related to the date convention, the only place I saw the 1950 reference date stated for the BP 'present' is in the caption to figure 3, and since this isn't a paleoclimate-specific journal I think it would be helpful for the readership to note this much earlier and more plainly.
Thank you very much for the clarification. We revised the manuscript in respect to the notation mentioned here and introduce the reference year (1950) earlier in the manuscript.

line 11: having said that melt dominates the changes at 13ka, it would be helpful to say which component causes the increase in SMB at 9ka - is it an increase in melt or decrease in accumulation?
We added 'After 13 ka the increase in melt begins to dominate the SMB decrease'.

line 12: would be helpful to note the timescale of this AMOC/SMB variability here, it it related to eg Heinrich events or slower/different variations
We will add 'centennial-scale episodes'.

line 13: not sure a statement on data availability needs to be part of the abstract does it, that intent is implicit in EGU journals now?
It is not and we removed it from the abstract. Thanks for the suggestion.

line 25: a lack of local horizontal resolution is far from the only reason ESMs struggle with ice sheet SMB - large-scale background climate and circulation biases play a big role, and increasing resolution (horizontal and vertical) certainly doesn't solve everything for either Greenland or Antarctic surface simulations in global models.

We do agree with this point and mention these shortcomings by including *"This is specifically challenging, as ESMs exhibit biases and the horizontal resolution is often not sufficient to capture small scale climate features, e.g., sharp topographic gradients at the ice sheet margins as well as cloud, snow and firn processes (e.g. Lenaerts et al. 17, van Kampenhout et al 2017, Fyke et al. 2018)."*

line 28: this paragraph is still supposed to be about ice sheets generally, but all these references are for Greenland SMB. You could throw in some Antarctic studies too to be a bit more general - eg Agosta et al., TC 2019
We added some references about SMB changes over Antarctica (Lenaerts et al., 2012b, van Wessem et al., 2018).

line 29: I spent a while wondering which ice sheets were going to be included in this study, and it's not clarified until much later in the paper. This would be a good opportunity to say this - 'extends the analysis of northern hemisphere SMB changes', perhaps? - and also a good place to note why it's of interest to have a detailed SMB product / analysis for this domain.
Thank you for your suggestion! We included that we are analyzing the northern hemisphere. And we add a short paragraph why we need realistic SMBs specifically for this region and time period. We added in the end of the paragraph: *"These climate changes and the variability associated with the changes in the northern hemispheric ice sheets during the deglaciation and the resemblance to the expected future climate change emphasize the need for a realistic representation of the SMB for past and future stand-alone ice-sheet and coupled climate-ice-sheet model simulations (Fyke et al., 2018)."*

line 37: references for what different proxy studies suggest about AMOC variation during these periods would be useful
We added some references (Heinrich 1988, Keigwin et al 1994, Vidal et al 1997).

line 45: I think this statement is too strong - Bauer and Ganopolski show that using fixed parameter values throughout in their simple PDD model does a poor job, but inverting that to say that *only* EBM models can produce realistic ice volume changes seems too much. I'm a little wary as well of the way the term "EBM" might be taken by readers too, whilst we're on the subject - in my (climate modeller) experience it's most often used in the sense of a much-simplified model of radiation balance for an atmospheric column, but here (I think) its being used in a much more general way for any model that explicitly calculates the components of the energy budget and how they affect temperature and the phase of water at the surface, however complex or simplified. This potential confusion could be eliminated by a disambiguation sentence?
We use the term EBM more specifically as synonym for a model that is used to calculate and downscale the surface mass balance from ESMs independent of its complexity. We are excluding PDD models, which use statistical relationships between melt and temperature patters based on present-day observations. We changed the sentence in response to this review and reformulated the motivation for this study.

line 53: typos: "a EBM" and "Two kind of simulations"
Thanks for pointing these out!

line 62: this is OK for present-day Greenland (although the accumulation isn't a simple function of height), but should this relationship be taken to be equally strong for all ice sheets, all the time? For example, variation in present-day Antarctica SMB components is more significantly controlled by dynamic atmospheric conditions, rather than a local scaling with topographic height I'd say. Might be worth a caveat here, and a sentence in the discussion about the limitations of this sort of downscaling

We added a note in the paragraph that emphasizes that the elevation dependence is not all there is: *"Note, that an elevation dependence of the SMB components is a simplified assumption and valid mainly for present-day Greenland. Atmospheric dynamics also significantly contribute to variations in the SMB components, specifically over present-day Antarctica."*

line 78: can the intervals/level heights be detailed in a table, maybe supplementary material?
We added a Table to the supplementary material.

line 84: it would be helpful to note in this list that the phase of precipitation may be changed too
We added: *"Total precipitation rates (liquid and solid) are corrected under consideration of the height-desertification effect. This halves the precipitation for an orography height difference of 1000 m above a threshold height of 2000 m for each grid point (Budd and Smith, 1979). Note, that snowfall is determined from the total precipitation for height corrected near-surface air temperatures below 0C within the EBM."*

line 105: does the rain that potentially refreezes have a temperature, - perhaps that of the surface - changing the amount of energy required to refreeze it?
We added *"The temperature of rain is assumed to be equal to the height corrected near-surface air temperature."* for clarity.

line 141: I found it confusing to say "the aging process starts", when so far the only aging process described is for fresh snow. 'Another aging process starts', perhaps? Your aging of the surface albedos is purely a function of time, when quite often the rate of increase of snow-grain size and thus lowering albedo is dependent on temperature and density. Do you know how sensitive your results are to the timescales you've chosen, and whether you would expect them or your results to be different for LGM temperature and accumulation rates?
Here, it is not another aging process but the aging of the albedos as described above. Maybe 'the aging process starts from the beginning' would clarify? The parametrization used here goes back to Oerlemans and Knap (1998). We did not do a comprehensive testing regarding the sensitivity to the length scales, as we mainly followed the parametrization in the above mentioned reference. Also computational limitations for the long-term simulations make several attempts with different tuning factors not feasible.

line 155: Going on with the thread of the previous comment, any table of tuned values almost begs the question of how sensitive your model is, and conclusions are, to the choices you have made, especially when they have been tuned to match one type of climate (the present GrIS) and the model is then applied to very different ones. Do you have any insight from your undocumented(?) tuning process that you could use to comment on this sensitivity? More work, but maybe possible if you have already done (or could easily do?) the simulations, would be to not only compare your results with MAR SMB for the present, but also for what other models find for GrIS SMB at 2100, or under some other idealised climate forcing. Freely downloadable results do exist for MAR run to 2100 with the climates from various CMIP6 models, from the ISMIP6 preparations. That wouldn't be an evaluation in the same way, but would give your readers some useful reference points for how sensitive your EBM tuning is to the climate it's in.
The SMB derived here is sensitive to all of these tuning factors, which is the reason why we chose to present the values in the Table. As for any tunable variable in a model we made the choice between being as close as possible to present-day observations and model performance. Due to computational limitations for the long-term simulations we have chosen one set of parameter for the deglacial simulations, hence, we do not know how sensitive the

choice is under a changing climate. That being said, we are not confident that a comparison of the SMBs in a future climate with MAR will allow us to fully evaluate the obtained SMBs for a changing climate. It will be difficult to disentangle the effects on the SMB that arise from differences in the climate forcing itself (MPI-ESM vs. MAR forcing). However, we applied the EBM to a simulation with the MPI-ESM high resolution setup for the SMBMIP inter-comparison and the results showed that the setup was specifically good in representing the trends of the Greenland mass loss between 2003-2012, indicating that not just the SMB mean climate over Greenland, but also changes in the climate are represented reasonably well as compared to observations and RCMs (see Fettweis et al., 2020). We added the following information to the manuscript: *"The same parameters were applied in an EBM simulation forced with output from a high resolution MPI-ESM simulation for historical climate conditions within the scope of a SMB model inter-comparison (SMBMIP; Fettweis et al. 2020). The results showed that the derived SMBs were very similar to observations in terms of the SMB mean climate as well as the SMB trend (2003-2012)."*

line 161: I had trouble following this section, but I think I more or less got there in the end. Might be worth thinking about rephrasing. The two cases handled are labelled 1) and 2), but that numbering isn't used again, and it seems like the first case described ("The inflow of snow and ice from above") is actually case 2). How does the temperature profile shift - in proportion to the amount of mass (solid and liquid?) that has been moved between layers? In the case of surface melting (line 165), the mass has not actually left the snowpack until the percolation/refreezing procedure has been completed, so is the change in density profile delayed until that has been done? Although,surface melting lifts the profile /up/, which sounds like melting at the surface makes the top layer /more/ dense than the reference. How does the surface layer ever reduce again to the reference profile, since more snow falling should further increase the layer density (line 163) or be directed to other layers underneath? Is it said anywhere what the top/bottom densities in this profile, or the density of ice in the virtual under-layer, are?
We have revised the complete paragraph for clarification under consideration of the questions mentioned here. The procedure applied here is instantaneous. We first calculate compaction and then surface melt. Percolation and refreezing can affect the density profile. The temperature profile follows the movement of mass. We also tried to clarify how the density profile changes for different conditions. The top and bottom densities are those of ice and snow. We hope that the rewritten section allows for a better understanding of the used parametrization.

line 190: is even the deep ocean in a true steady state for your LGM initial condition, or are there drifts there to worry about? If it's all a steady state, why are first 5 kyr considered a discardable spin up?
Note that these are transient simulations of the last deglacation, hence, there is no steady-state. We initialized the simulation from a spun-up glacial steady-state at 26 ka. We are only investigating the last 21 kyrs for the model to adjust to the changes and because this is the most interesting time period. 5000 years are enough to capture the drift in the deep ocean. Therefore, we considered the rest as additional spin-up. As this seems to confusing we rephrased.

line 195: Freshwater input forced AMOC variability comes up as a feature later - it would be good to mention specifically how glacial runoff/outburst water is treated in their protocol
We added following information: *"Freshwater from melting ice sheets is calculated from the volume changes in the ice sheet reconstructions for each grid point. For grid cells over land melt water is distributed through the hydrological discharge model, over ocean it is discharged into the adjacent ocean grid cells. "*

line 200: From figure 4 etc, it seems like the ELAs talked about here are not a simple contour on the ice surface demarking where ablation and accumulation balance, but a /potential/ ELA calculatable for the climate in each land surface gridbox regardless of what the actual altitude (or SMB) of the ice sheet was. I'm still not sure what I think that means - precipitation, for instance, is at least partly orographically controlled and either falls in one place or another so it does matter where the surface really is, so not every potential ELA calculated can simultaneously be true...maybe?...but anyway, it would be worth another sentence to explain that the ELA that will be talked about from here on is not a simple contour line on the ice.
We included a sentence that clarifies the ELA used in this study, which is indeed a potential ELA, as it is calculated in each grid point. *"At heights above the ELA it is thermodynamically possible to accumulate snow throughout the year and form an ice sheet or glacier. At elevations below the ELA melt dominates accumulation and no ice sheet can form. Here, the ELA is calculated in each grid point, hence, resembles a potential ELA. It is a proxy for climate changes affecting the ice sheets. "*

line 202: since the ELA is defined as the height where SMB=0 I don't really see what is meant by saying it's less sensitive than SMB
To obtain realistic SMBs we need to downscale the SMB onto a different grid (to the ice sheet topography from the reconstructions), while the ELA can be determined directly on the atmospheric model grid. The downscaled SMB shows enhanced melt patterns specifically at the margins of the ice sheet where the simulated climate from the coarse resolution model and the high resolution ice sheet topography potentially do not fit together (in a fully coupled simulation an ice sheet would likely not survive in these areas). If these values are integrated, the resulting SMB may be underestimated. Integrating the ELA on the native model grid allows for a more model consistent representation of patterns. Hence, it is less sensitive than the SMB. We rephrased this sentence to make it more clear. *"As the ELA estimate is calculated on the native model grid it is more consistent with the model physics and boundary conditions used in the simulations than the downscaled SMB. Hence, integrated values of the ELA are less sensitive to changes in the ice sheet mask than those of the SMB. "*

line 232, 242: structurally, do these short paragraphs need their own section headings? Thinking about it, would it make more sense to put the paleo experiment design section later, next to the paleo results, rather than right up front before all the recent history stuff is described and analysed?
These are good points but we have decided to put it this way around, as our historical simulations are branched off from the deglaciation experiment. So to describe the recent history simulations we believe we need to describe how we initialized them. To avoid too much repetition we therefore decided to put it this way around. We changed the sub headings to "Evaluation Data" to avoid too many subtitles and hope that this is sufficient.

line 238: not just cloud, all precipitation in ERA - and thus your EBM-ERAI GrIS accumulation - is also purely a product of their model physics run at global resolution, so I wouldn't expect that to make for a great comparison with the local MAR precipitation, regardless of how well EBM-ERAI can do the surface energy balance.
We fully agree with the reviewer. This sentence was meant for readers not fully aware of the limitations of reanalysis. We added some examples (precipitation, clouds).

line 252: from here on, neither refreeze nor runoff is mentioned, only "melt". Since we're often talking about total SMB, I'm not sure whether they really mean surface melt (that might refreeze), or the actual liquid mass leaving the snowpack (runoff). Or are they talking about runoff-precipitation - so the proportion of the original mass that has now been lost? I'm interested to see what their refreezing does anyway, and it's a shame that that's not

mentioned, but if they really are talking about just surface melt from here on then I really think the refreezing needs to be shown too.
We believe that we have introduced melt in the introduction *" The SMB is determined by mass gain due to accumulation, as a result of snow deposition, and mass loss by ablation, induced by thermodynamical processes at the surface and subsequent melt-water runoff (Ettema et al., 2009)".* Here the melt is defined as the mass leaving the snowpack, that means melt water that refreezes is not considered in this estimate. We added a sentence to clarify this definition. *"In the following, accumulation is defined as mass gain due to snow deposition and melt as mass loss due to ablation (often referred to as runoff). Refreezing processes are considered in the melt estimate."*

line 269 (and I elsewhere, like line 316): there's a tendency here to simply blame the biases on in the MPI-ESM runs and the EBM SMB derived from them on inadequate spatial resolution and the local topography. That will undoubtedly be part of it, especially at ice sheet margins, but higher resolution doesn't fix everything in complex climate models, in fact it can make it worse (see Kampenhout et al, TC, 2019 for a very geographically relevant case-study). Biases in the large-scale climate, inappropriate initialisation, internal variability, missing physics - the list is of course endless, so it's irksomely simplistic to keep saying that the model resolution is too low and no more. It's not like the EBM_MPI-ESM results would = EBM_ERAI if you just ran MPI-ESM at ERAI resolution.
We are aware that model resolution is not the only reason for the discrepancies between MAR, EBM_ERAI and MPI-ESM-LR and MPI-ESM-CR. In fact, there are many uncertainties in our system, related to biases in the model climate (clouds, precipitation, surface processes, ...), the parametrization used to downscale the SMB (including e.g. a constant lapse rate over the whole ice sheet) and of course the vertical and horizontal interpolations from the elevation classes on coarse resolution to the high resolution ice sheet topography. We revised the manuscript throughout and added specifically *"The comparison between EBM_MPI-ESM-LR and EBM_MPI-ESM-CR indicates that an increase in resolution cannot resolve all biases. This is in line with findings by Kampenhout et al. (2019), showing that a regional grid refinement in simulations with the Community Earth System Model (CESM) did not improve all SMB components. Model biases, e.g. in the large-scale circulation, clouds and precipitation patterns (Mauritsen et al., 2019), as well as uncertainties due to internal variability are exhibited in all ESMs and explain part of the differences seen in the presented comparison."* and also revised other places throughout the manuscript (e.g. line 316). We also believe that we have discussed other possibilities in the discussion section: "*A comparison with SMBs derived from a simulation with the MPI-ESM-LR setup and ERA-Interim reanalysis reveal that discrepancies between the SMBs derived from MPI-ESM-CR and MAR are a result of the coarse resolution of the model (e.g., the extensive melt in the North of the Greenland ice sheet) and due to the quality of the forcing itself (e.g., precipitation patterns). In contrast to ERA-Interim, MPI-ESM evolves freely and does not assimilate any surface observations; hence, differences are to be expected. Further differences are related to underlying topographies of the native models as well as the fact that most fluxes within the EBM are parameterized, as they are not directly available from the model simulations in the required temporal resolution."*

line 275: RACMO hasn't been introduced yet in any way, as a model or an acronym
Thanks, we introduced the acronym here.

line 276: You have some tunable control over the rain/snow partitioning in your EBM, so could this MPI-ESM bias be fixed for your purposes?
Precipitation is considered as snow if the height corrected near-surface temperatures are below 0C. This would be of course a tunable value, but we do not believe that it is very physical to change this relationship.

line 281: I would have thought that larger-scale moisture transports over the ice sheet,

which are also quite resolution dependent, would be likely to play a role here too. Low resolution atmospheres often just drizzle too far north and over too-wide areas because they're overly diffusive everywhere, it's not just a local thing you would fix with a higher ice sheet. Again, maybe this is an area where things in the climate models other than simply local grid resolution/topography play a significant role
This is likely also a contribution but we believe it is rather remarkable how differences in the topography match the differences in the precipitation. This points to the topography being the dominant driver of those difference, but does not rule out other processes. As we added a couple of sentences in response to comment line 269 we believe we have covered this point.

line 295, 300: I didn't follow the logical thread through here. First MPI-ESM-CR has a higher topography than MAR and the EBM produces too much melt, then later the area in the south is higher than the ISMIP topography and there is less melt. Less melt in which model? Is there a consistent physical link being drawn between these two cases?
We apologize for the confusion. The first sentence should read *"has a higher topography than ISMIP..."* - as is shown in Fig. 2. We believe that we explained the physical mechanisms by *"One problem of downscaling melt in these regions is that temperatures are always at the melting point during melting. By projecting the temperatures onto lower elevations, the height corrected temperatures depart significantly from the melting point towards higher temperatures. Hence, the vertical downscaling from higher elevations to low elevations overestimates melting"* and that the reason for the confusion lies in the small error.

line 299: your model has a pretty direct control on how much melting you get when you adjust for a new elevation through the lapse rates - does this point to the lapse rate changes being too high?
This is a good question. The lapse rate is used in the model as tuning parameter and we have chosen a value that results in realistic SMBs for the whole Greenland ice sheet. In the current version of the model it is considered as relatively low (4.6K/km), which certainly is at the lower end for Greenland temperatures. It points towards the challenges in using one constant lapse rate over all ice sheets and certainly is a shortcoming in our EBM.

line 323: I think this statement on the scope of the paper should be in the Introduction, not all the way back here
We moved this sentence up in the introduction. Thanks for the suggestion.

line 329: The Antarctic ice sheet has yet to collapse, thankfully
We fully agree and added 'all northern hemispheric ice sheets'.

line 332: It would be good to say why these particular time-slices are chosen for analysis.
These time slices show the most significant changes during the deglaciation. We added this information. "... in order to indicate the most drastic changes in the northern hemispheric ice sheet configuration."

line 344, 396: the impact of the changes in circulation and the different realisations of AMOC history on the ice sheet SMB sound interesting, it would have been good to hear more about them and how that might have influenced the ice sheet deglaciation. There is some decomposition of whether accumulation or melt is dominating the SMB/ELA trend for the main Glac1D run at different periods - does all this still hold for the ICE6G run? If so, that's an interesting conclusion perhaps that you can say is independent of the ice sheet reconstruction used.
Good point! We added some information on the atmospheric circulation changes (see also comments by reviewer #2). We also see similar melt-accumulation relationship for the Ice6G simulation, although the absolute magnitude differs. Melt is close to zero until about 14 ka

and also reduces during the major AMOC slowdowns. Note, that the AMOC slowdowns are somewhat different in their timings and occurrence, due to uncertainties in the reconstructions - something we are currently investigating in a separate study. As the relationships hold we included a sentence pointing this out. Thanks for this suggestion.

line 394: the final figure of 550 Gt/a doesn't match the values in your table /very/ well. Does the 550 at the end of the transient correspond to the year 1950, rather than _2000? Your recent history run does go through 1950 doesn't it, even if only 1980-2010 were analysed - does this 550 match what the historical run did at that point?
The historical simulation is somewhat different from the transient simulation, as it includes more realistic forcings (e.g. volcanoes, land use, anthropogenic forcings). Therefore the SMB values at the end of the transient simulation (1950) are higher than in the beginning of the historical simulations. The differences in the forcing are also the reason why we performed the last millenniums simulation (starting at 1850 from the deglaciation experiment), in order for the model to adapt to the changes in the forcings. Hence, we have no overlap between the end of the deglaciation and the historical simulation to directly compare the values. We revised the sentence to *"These values are similar to values observed during the 21st-century, although slightly higher as no anthropogenic forcings are considered in the deglaciation simulation (see Fig. 4, Section 3 and Table 3)"* for clarification.

line 403-405: references needed for the decline history of these glacial ice sheets
We added references to this section.

line 450: this information about the meltwater forcing should be in the experiment description
As we included how we determine the freshwater flux from the ice sheet reconstructions in the experiment section in response to an earlier comment we believe that this is covered. The melt water pulse is prescribed by the reconstructions, which are introduced in the experimental description.

figure 1: is this surface melt, or runoff: the caption says ACC = SMB - MELT? It also says that MAR does not provide accumulation, which I don't understand - MAR gives you snowfall and rainfall? The ELA is just the contour of where the SMB on the surface is 0, isn't it? Is it said in the text why CR has so much more melt in far north
Accumulation in our study is defined as the mass gain due to snowfall. While MAR provides snowfall and rainfall it does not explicitly provide how much of the precipitation accumulates over the ice sheet. We believe that calculating it as residual is therefore the right thing to do. We believe that melt in the north of the ice sheet is a result of the underlying topography (and to some extent probably also model biases) and cite from the text: *"Comparisons with EBM_MPI-ESM-LR, which shows less melt in the north and west of the ice sheet as compared to EBM_MPI-ESM-CR (Fig. 1 and 2), confirm that differences in the melt patterns are linked to the underlying topographies of the model versions. MPI-ESM-LR is slightly higher than MPI-ESM-CR and thereby closer to MAR on the northern and western flanks of the ice sheet, hence EBM_MPI-ESM-LR$ shows less melt than EBM_MPI-ESM-CR in these areas."*

fig 2: (top, middle) to denote the second row and (bottom, middle) to denote the third row is not very clear. Ditto (center, left) and (center, right). With this many panels may just label them a)-p)?
We added labels to Fig. 1 and Fig. 2. Thanks for the suggestion.

fig 3: Paleoclimate figures often stack different proxies on the same horizontal time axis but offsetting the vertical axes, rather than overlaying everything completely like this - that might work well here too.

Thanks for the suggestions. We are aware of these plots but think that in the current presentation relationships between e.g. the ELA and CO_2 as well as the MOC are easier to see and interpret. Similarly for melt and accumulation relationships.

fig 4: the 2D ELA surface idea might need explaining in the caption here too. Is the EBM run everywhere? Is there +SMB in the MPI-ESM climates for other places, eg a Siberian ice cap that's just not shown here because masking teh results onto the Glac1D glacier mask?

We clarified the ELA definition and that we masked the SMB and ELA with the glacier mask by including "*The SMB is interpolated on the Glac1D topography for each individual time slice and masked with the Glac1D glacier mask. The ELA, defined as elevation where the SMB equals zero, is calculated for each grid point on the native MPI-ESM-CR model grid from the 3-D SMB in each grid point (see Sections 2.1 and 2.3). The ELA is masked with the glacier mask used in the MPI-ESM-CR simulations.*"