First of all, we would like to thank Dr. Fyke for the constructive comments on our manuscript. We will endeavor to integrate them into the revised version. Here is the response to the main concerns.

Niu et al. present a study that identifies a strong sensitivity of glacial-interglacial ice sheet evolution to applied atmospheric forcing fields, particularly summertime temperature. It highlights uncertainty in climate forcing as a major control on long-term evolution (in this case, past evolution).

GENERAL COMMENTS

Recent publications have already highlighted the important/dominant role of climate forcing on long-term ice sheet evolution (e.g. Qing et al., 2014, 10.1007/s00376-013-3002-6; Dolan et al., 2015, 10.5194/cp-11-403-2015; Fyke et al., 2014, 10.1007/s00382-014-2050-7, for example, I know there are others out there). In addition, I'm not sure the finding of summertime temperature (or more accurately, summer energy balance) conditions as the major control on ice sheet extents is highly novel (for example, two early studies that discuss this sensitivity: Gallee et al., 1992 10.1029/92JD01256; Oerlemans, 1991; 10.1177/095968369100100106). Generally, a stronger literature review is necessary, both to recognize these earlier efforts and also to place the present effort in context in terms of originality.

Response: Yes, we agree that the introduction section is too general. The role of climate forcing and model comparison on Greenland ice sheet has been studied not only for paleo time (Dolan et al., 2015), but also for future projections (Yan et al., 2014; Fyke et al., 2013). A more thorough literature review pertaining to the dominant role of climate forcing on long-term ice sheet evolution will be presented in the revised paper. A lot of studies has been done by researchers that for land-based Northern Hemisphere ice sheets, the surface ablation process is of importance (e.g. Gallee et al., 1992). Our main focus is on the sensitivity of ice sheet model simulations on GCM in the PMIP3 framework. We address the highly matching pattern between the summer surface air temperature and the ice sheet extent.

It seems odd to have the first part of the paps focus on COSMOS-AWI, then turn somewhat orthogonally to evaluation of the PMIP3 ensemble. As it stands the manuscript is too disjointed, in part validation of COSMOS-AWI, and in part a PMIP3-based sensitivity experiment. The two need to be better linked if they are to be included as one study - for example, by just including the COSMOS-AWI-based simulations as an additional PMIP3 ensemble member and reserving special evaluation of COSMOS-AWI-based PISM simulations as a separate study.

Response: The COSMOS output was used initially for tuning the ice sheet model to match the sea level record through the last glacial cycle. Over the past several years, COSMOS has been developed for paleo climate simulations, and has been evaluated for different scenarios. We believe it is appropriate as the climate forcing for our PISM simulation and is suitable as our reference simulation. Then we compare the PMIP3 models to show the variability of the model performance. The PMIP3 ensemble members were evaluated by using the same model setup. A better linkage between these two parts will be addressed in the revised version.

The driving of ice sheet models with transient climate model data, where ice sheet topographies, surface types, etc. are prescribed in the climate models to recreated reconstructed ice sheet topographies, is arguably circular. For example, due to the coupled nature of land-surface/near-surface temperatures (including over-ice temperatures, which are strong functions of ice elevation) and albedo feedbacks, temperatures should be much higher in bare-land regions, which strongly limits simulated ice advance past the reconstructed southern ice margins. This is very likely why the PISM ice extents simulated here very consistently do not go southwards of the reconstruction's southern limits. This circularity needs to be addressed, in terms of how it potentially impacts the main results of this paper.

Response: Yes, there is a circularity. A truly coupled version for a glacial cycle simulation is too computationally expensive at this time. We will stress the circularity out more clearly in the revised version.

Recent work has recognized that temperatures alone do not determine the majority of the summertime response - perhaps a dominant aspect comes from direct radiative forcing, at the point where the snow/ice is at the melt point (see van den Berg, 2012, 10.1038/NGEO1245, for the case of Eemian Greenland). Thus, PDD assessments of glacial behavior are missing a major control on summertime ablation conditions. Better justification needs to be given here to why PDD-based assessments of ice sheet sensitivity to glacial climate change are valid, in light of more physically accurate energy balance-based approaches (which is where the broader community is increasingly turning).

Response: Yes, a previous study (e.g. van den Berg et al., 2012) shows that the insolation influences the surface mass balance, this probably needs to be separated, rather than simply just taking surface air temperature as a whole. While on the other hand, the insolation influence is also partly reflected in the surface air temperature.

In order to evaluate the performance of PDD scheme, we compared the surface melt simulated in PISM and the melt computed from COSMOS, which uses a much more sophisticated energy balance scheme but much lower resolution (T31). As is shown in Fig. 1, the plots on the left are surface melt simulated in COSMOS with fixed ice sheet configuration at the LGM (a), preindustrial (d), and Eemian (g; Gierz et al., 2017). The plots on the right panel are the surface melt from COSMOS-AWI simulation in the MS, with reanalysis present day products and COSMOS LGM surface temperature and precipitation as climate forcing. In order to make the comparison more consistent, instead of reanalysis present day products in the previous version, we also ran simulation with the same setup as COSMOS-AWI, but use COSMOS LGM and COSMOS PI as forcing (middle panel) (Also see Fig. 2, and the reply to P5, section 13 below). The results show good agreements between PDD-based approach and energy balance-based approach, especially around the ice sheet margins. At the LGM, all the results show similar melting pattern around the margins of North American ice sheets and Eurasian ice sheets. For present day and Eemian, the snow melt extent in the North American and Asian continents in COSMOS is broader,

while for Greenland ice sheet, the surface melt patterns are still match well. The simulation with reanalysis products show more melt around Greenland, which probably because the observation data contains the warming signal of the previous century (as discussed from Referee 2).
We believe that the PDD-based scheme provides a good representation for the more physically accurate energy balance. More discussion will be added in the revised version.

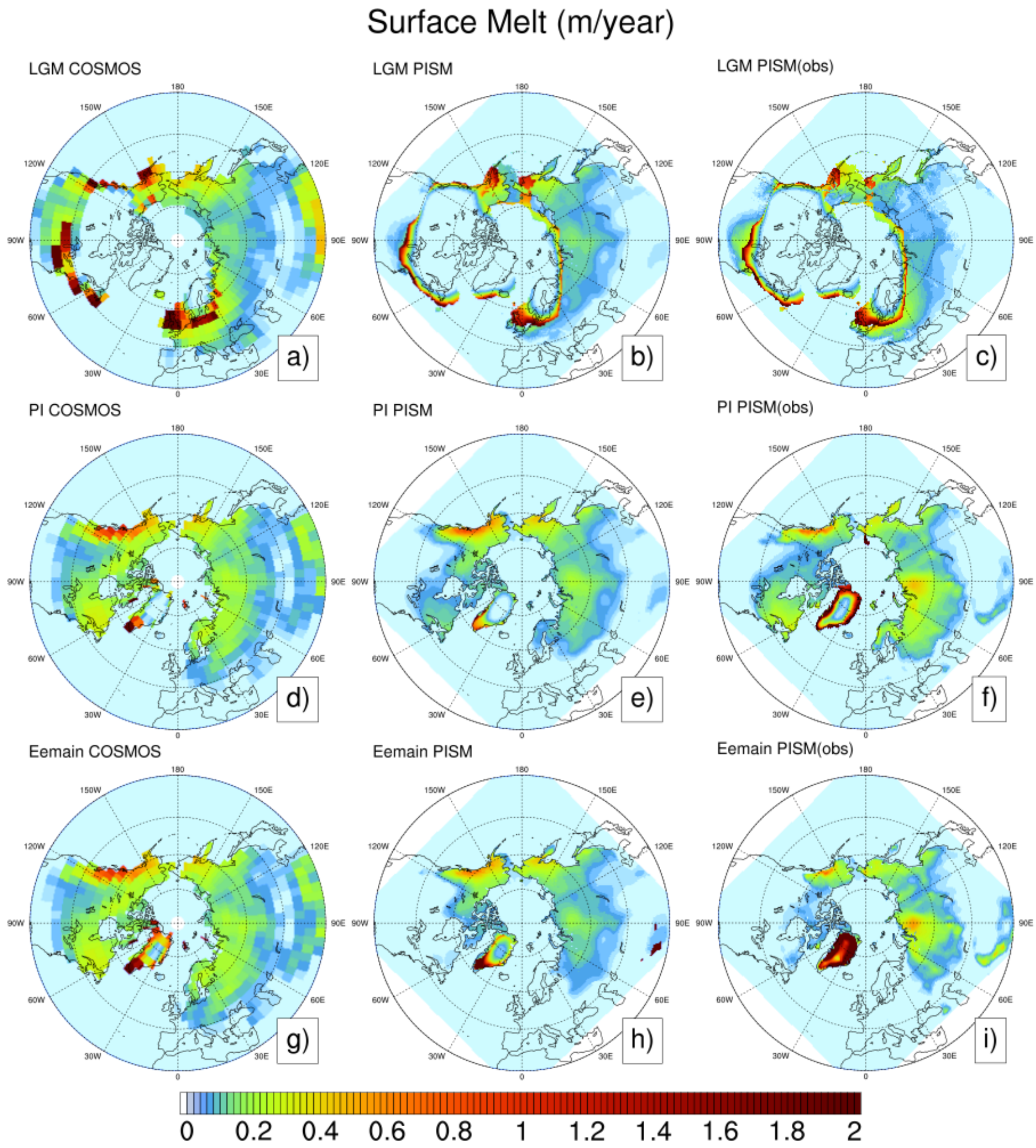

Figure 1. Comparison of surface melt between energy balance-based scheme from COSMOS output (a, d, g) and PDD-based scheme from PISM (b, c, e, f, h, i) at the LGM, present day (PI) and Eemian. (Plots on the right panel are from COSMOS-AWI simulation, by using surface temperature and precipitation from COSMOS LGM and present day reanalysis products as climate forcing. While Plots in middle panel are the simulation with the same setup as COSMOS-AWI, but use COSMOS LGM and COSMOS PI as forcing).

SPECIFIC COMMENTS

P2L15: Also, Fyke et al. (2011), 10.5194/gmd-4-117-2011

Response: Yes, it will be added in the revised version.

P2: A focus on the importance of ice-sheet/climate coupling seems a bit out of place, given this study does not include this coupling. More justification is needed (and I think possible) on the value of standalone modeling, forced with previously-generated climatologies.

Response: As mentioned above, more extensive literature review will be added in the revised version.

P3, section 2.1.1: how is PISM internal temperature initialized (not that it perhaps makes a big difference here, since the main areas of interest have no ice at t=0)

Response: The temperature is set to the solution of a steady one-dimensional differential equation in which conduction and vertical advection are in balance. We found that this is of secondary importance since the main areas of interest have no ice at t=0.

P3, section 2.1.1 & 2.1.3: How are ocean interactions accounted for in the simulations presented here? Is ice removed if floating? This has a bearing for which ice sheet dynamics (SIA vs SSA) are dominant in PISM in these simulations and also on the marine margins of the simulated ice sheets.

Response: The melting at the boundary between ice and ocean is parameterized by assuming a constant value for the ocean temperature and a parameterization scheme from Martin et al. (2011), so an ocean condition from a GCM is not used here. Since the Northern Hemisphere primarily has land-based ice sheets, we regard this assumption as reasonable. The ice is removed by two calving schemes, either as a function of the horizontal strain rates or the thickness of the ice shelf (section 2.1.3, P4L16-26). For ice shelves in PISM, SSA is dominant. More details about this part are added in the revised paper.

P5, L3: It is not immediately clear to the reader how the white noise term is added.

Response: A random temperature with a normal distribution with a mean value of 0 is added to the monthly data. This will be added in the revised paper.

P5, section 2.2 and Figure 2: It is not clear how conditions 'outside' of the LGM/PD range are accounted for, given the 0-1 index range defined between present and 21ka conditions. For example, much of the Holocene, and the late Eemian, and much of the LGM lie outside of the LGM/PD window. Please describe how these conditions at these points are calculated.

Response: As is described in Sec. 2.2, the LGM and PD conditions are treated as the cold and warm reference conditions. Combined with NGRIP $\delta^{18}O$ data, the climate forcing at the other stages can be calculated. For much of the Holocene, the late Eemian and much of the LGM, the NGRIP values are

outside the range of $\delta^{18}O_{LGM} \sim \delta^{18}O_{PD}$, so the index can be larger than 1 or smaller than 0. The calculation of the conditions is still valid by using Eq. 7-9. More description can be added in the revised paper.

P5, section 13 (also, PMIP3 section): The choice to use observed/reanalyzed present-day conditions instead of model conditions for the present-day needs to be better justified. It would seem more consistent to use the GCM's present-day (or more accurately, preindustrial) conditions. This could, for example, limit cases where the LGM is possibly spuriously warmer than the PD due to GCM LGM warm biases. More generally, it seems that for times where the glacial index is closer to 0 (present-day, for example, the frequent glacial interstadials) the signal of GCM forcing is 'diluted' by observations/reanalyses, which adds ambiguity to the main 'sensitivity to GCM forcing' aspect of the study. Can the authors describe in the manuscript how use of GCM PD conditions could increase/decrease the spread of ice sheet conditions in response to different GCM forcing?

Response: Here we fixed the present-day condition by using the observations while by only differing the forcing at the LGM to investigate the sensitivity of the ice sheet to the climate input. In this case, the model bias is also included in the results, which is also part of our conclusion (model deficiencies should be considered carefully) when using the absolute output from GCM (e.g. coupling between GCM and ice sheet model).

We also ran experiments by using the PMIP3 preindustrial output instead of the reanalysis product. Fig. 2 shows the sea level equivalent evolution through the last glacial. Comparing with the simulations with reanalysis product (Fig. 7 in the MS), the curves are more dispersive. The sea level equivalent differences for Greenland at present day can be up to 6-7 meters. The ice thickness differences at the LGM for different models are shown in Fig. 4. For most models, the difference is not pronounced. For MIROC-ESM, the ice thickness difference of Laurentide Ice sheet can be up to 600 m. Comparing with the climate forcing, the summer surface air temperature in MIROC-ESM is warmer than present day reanalysis product (Fig. 5). The present day conditions do influence the simulated results, with warmer climate contributing to reduced ice sheet volume and colder climate result in larger ice sheet volume. We will put the figures into the supplementary materials. More discussion can be added in the revised version.

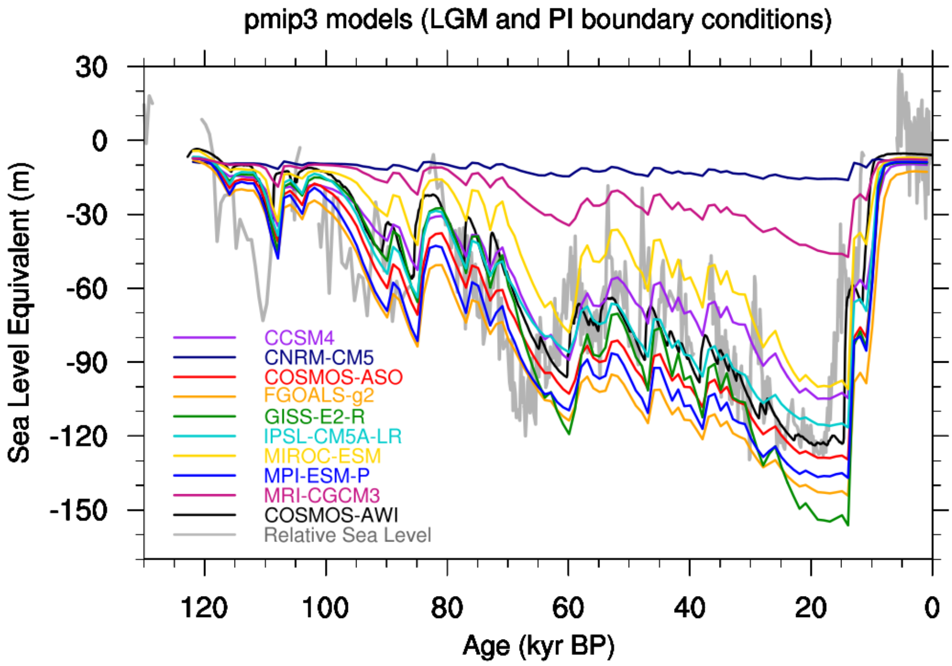

Figure 2. Modelled sea level equivalent of Northern Hemisphere ice sheets change through the last glacial cycle using PMIP3 LGM and PI output.

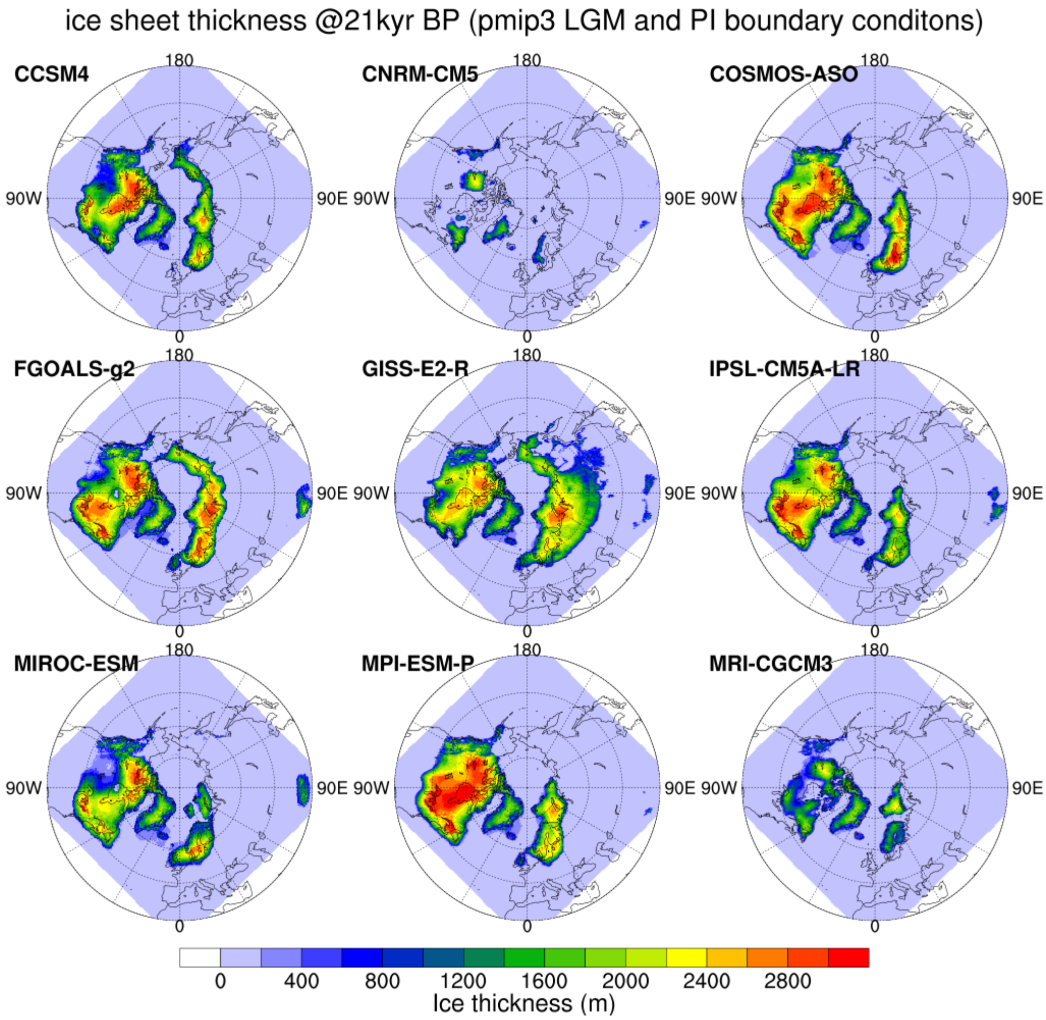

Figure 3. Modelled ice thickness at the LGM using the PMIP3 LGM and PI output

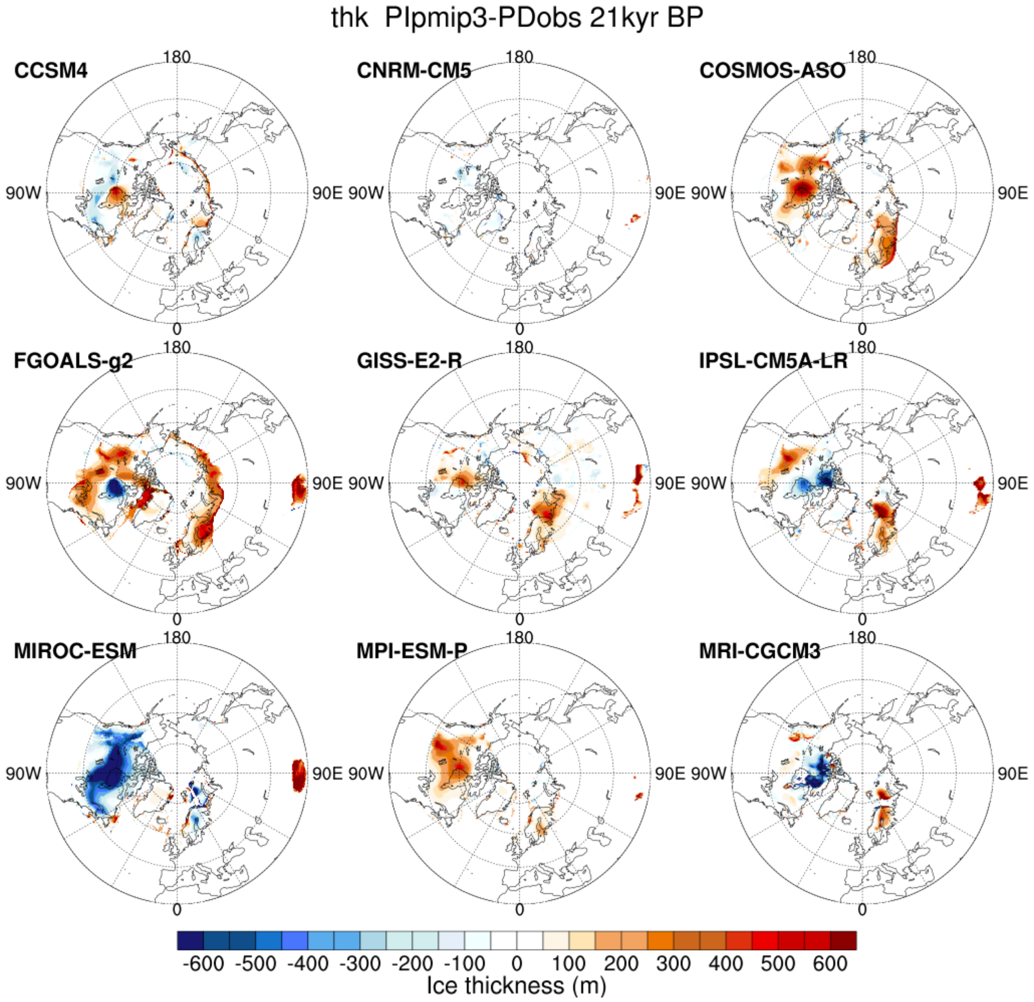

Figure 4. Ice thickness difference between simulations forced by reanalysis product (PDobs) and simulations forced by PMIP3 preindustrial model output (PIpmip3)

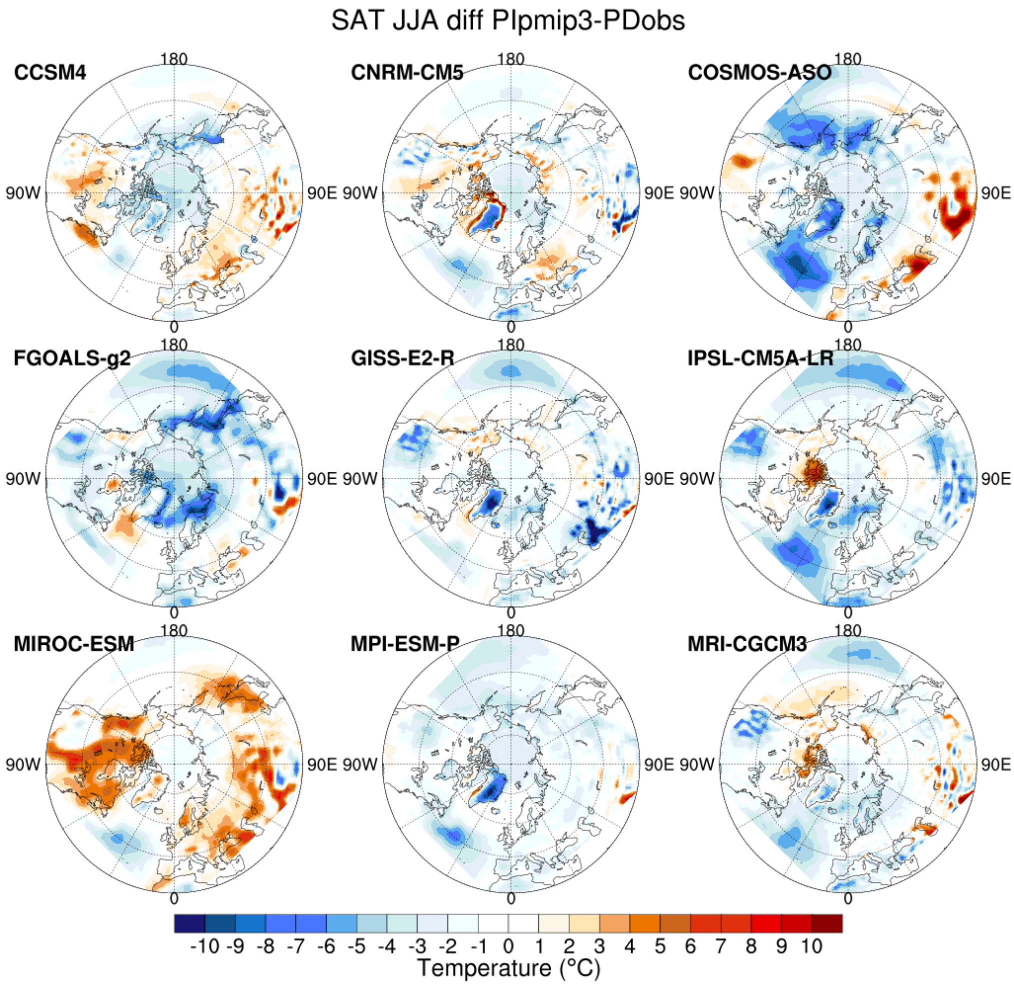

Figure 5. Summer (JJA) surface air temperature differences between simulations forced by reanalysis product (PDobs) and simulations forced by PMIP3 preindustrial model output (PIpmip3)

Section 2.2: is there a lapse rate correction for temperature, in the case where the ISM elevation differs from the GCM elevation? This would seem to be an important feature to include, and is often included (I think) as part of other standalone ice sheet simulations driven by PDD schemes.

Response: In this study, the lapse rate correction for temperature is not included. One of the reasons behind this is that the NGRIP index we use here is assumed to be the surface air temperature variation for the Northern Hemisphere over the glacial cycle. From this aspect, the index itself already accounts for the elevation change factor (the SAT over Greenland should highly correlated to the SAT over other Northern Hemisphere area).

The other reason is that here we want to test the sensitivity of ice sheet to surface temperature from the GCM. Including the temperature lapse rate correction gives the ISM elevation evolution one more degree of freedom. Also, the lapse rate values may differ between different areas or time stages.

Section 2.2: Does the COSMOS-AWI simulation also use the PMIP3 blended ice sheet LGM reconstruction?
Response: Yes, it also uses the PMIP3 blended ice sheet LGM reconstruction. COSMOS-AWI simulation used exactly the same protocol as PMIP3 experiments. This will be described more clearly in the revised version.

P7L18: … "this indicates the abrupt warming in Greenland resulted in negative surface mass balance." It's not clear that this isn't simply a result of the experimental design.
Response: We agree the sentence is a bit misleading, so we propose to delete this sentence.

P7L20: "The total ice volume reached its relative high around 109kyrBP": it's not clear what this means. Clearly ice volume was much bigger around the LGM.
Response: Change to "The total ice volume increased significantly at around 109 kyr BP with a 30 m equivalent sea level fall".

P7L25: could the authors describe mechanistically why they think neglecting SW/LW radiation would cause less variability relative to the reconstructed sea level?
Response: Here we use the same empirical PDD factors based on present day for the simulations over the whole glacial cycle. During different paleo stages, the insolation was different, which may also contribute to ablation. This might cause partial mismatch between the simulation results and the reconstructed sea level, as discussed from Bauer and Ganopolski (2017). This either led to too much ablation during the first phase of the last glacial cycle or too little ablation during the final phase (in our case too much ablation during the first phase).

P7L27: Given that Dansgaard-Oeschger (D-A) cycles recorded in NGRIP were quite possibly driven by ice sheet dynamics, the authors should justify the somewhat circular argument that the PISM response to D-A cycles temperature and precipitation forcing is realistic (since in the real world, D-A cycles were quite possibly a response to ice sheet change, not the reverse).
Response: Yes, we will address the interplay between ice sheet and climate during DO cycles. For stand-alone ice sheet simulations, the results are just a one-way response to the climate forcing.

P8L32/Conclusion #1: The authors should state to what extent similarities between the ice sheet and reconstruction extents at various times (particularly the LGM) relate to the strong control the reconstruction has on temperatures, via elevation (as noted earlier in the manuscript) and potentially in lieu of a temperature lapse rate implementation. It seems likely that ice sheet extents are controlled strongly by where temperature increases due to lower elevation, because of reconstructed ice sheet limits.
Response: The simulated ice sheet extents are controlled by the climate forcing from the GCM, while at the same time, the GCM output is based on a reconstructed ice sheet. This is partly the reason that

there are similarities between the ice sheet and reconstruction extents. The circularity and similarities will be addressed in the conclusion in the revised paper.

P10L7: I think the relationship between summer temperature and ice sheet extent should be shown more quantitatively and compared quantitatively against other seasons/precipitation, via regression/correlation methods, to confirm the qualitative conclusions presented thus far. Also, the line contours of temperature could be overlain on ice sheet extent (i.e. figure 9 overlaid on figure 8) to allow readers to see the similarity more clearly.

Response: Yes, we propose to add the ice sheet extent edge to the temperature pattern to see the similarity more clearly. Also, as suggested from referee 2 in general comments (3), the scatter plot between snowfall and 2 m temperature can be added to show the threshold value of around -5 degree that foster ice accumulation.

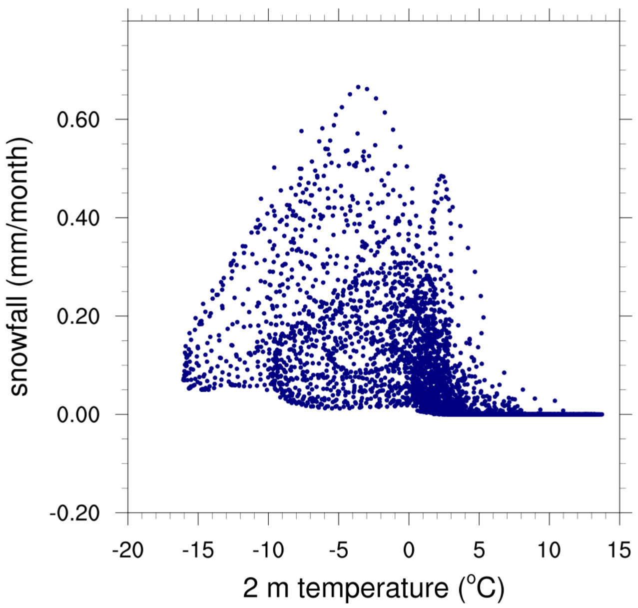

Figure 6. Scatter plot between 2 m temperature and snowfall over Greenland during summer by using the ERA-40 reanalysis product

P11L16: While it is likely that summer SAT is indeed important, the lack of ocean forcing in these experiments is likely a large caveat to the main conclusions. For example, much recent work has highlighted the role of ocean forcing in driving glacial ice sheet variability, yet the simulations presented here do not sample this potential source of climate-driven ice sheet change. This should be clearly noted.

Response: Yes, we will address that in the revised version.

Discussion/Conclusions: an informed discussion on why, exactly, summer temperatures are so different between GCMs would be useful here, to provide more background to the reader (and also to climate modelling groups, as they attempt to calibrate their models for LGM simulations and look for important processes to focus on).

Response: The PMIP3 LGM experiments are all using the same blended ice sheet reconstruction, while the simulated conditions can be different between GCMs. It shows that the climate results are model dependent, which is also the reason why we need the model intercomparison project. The reason behind that might due to the different climate components that included in the GCMs or the different parameterization schemes. For example, some models include the carbon cycle, while the others do not. We will add the discussion in the revised paper.

References:

Bauer, E. and Ganopolski, A.: Comparison of surface mass balance of ice sheets simulated by positive-degree-day method and energy balance approach, Climate of the Past, 13, 819–832, doi:10.5194/cp-13-819-2017, https://www.clim-past.net/13/819/2017/, 2017.

Gallée, H., J. P. Van Yperselb, T. Fichefet, I. Marsiat, C. Tricot, and A. Berger (1992), Simulation of the last glacial cycle by a coupled, sectorially averaged climate-ice sheet model: 2. Response to insolation and CO2 variations, J. Geophys. Res., 97(D14), 15713–15740, doi:10.1029/92JD01256.

Gierz, P. , Werner, M. and Lohmann, G. (2017): Simulating climate and stable water isotopes during the Last Interglacial using a coupled climate-isotope model , Journal of Advances in Modelling Earth Systems . doi: 10.1002/ 2017MS001056

Martin M. A., R. Winkelmann, M. Haseloff, T. Albrecht, E. Bueler, C. Khroulev, and A. Levermann, The Potsdam Parallel Ice Sheet Model (PISM-PIK) –Part 2: Dynamic equilibrium simulation of the Antarctic ice sheet, The Cryosphere, 5 (2011), pp. 727–740.

van de Berg, W. J., van den Broeke, M., Ettema, J., van Meijgaard, E., and Kaspar, F.: Significant contribution of insolation to Eemian melting of the Greenland ice sheet, Nature Geosci, 4, 679–683, http://dx.doi.org/10.1038/ngeo1245, 2011.