The authors would like to thank the editor Thomas Mölg, the two anonymous reviewers, and Irina Rogozhina and Raymond Sellevold for their critical reviews that have have lead to significant improvements of this manuscript.

In addition to a few stylistic changes to the text, we have made the following slightly more substantial changes based on the reviews:

1. We have added a paragraph in the introduction that motivates the simplified modeling approach and puts some of the compromises we make in perspective with the real glacial cycle.
2. We have significantly revised and extended the discussion section to put our results in perspective with previous modeling studies, and to dig a bit deeper into reasons for the increased temperature signal that hinders ice from forming in certain key regions.
3. All figures have been re-plotted on a stereographic projection and with a better suited color scales, and we have added a figure (new Fig. 1) of 500 hPa eddy streamfunction, zonal wind, and total cloudiness to illustrate how the atmospheric dynamics and physics respond to the resolution change.
4. The table has been extended to also include information about the computational demand at the different resolutions (explored in the discussion section).
5. A supplementary document has been added to show fields supporting our conclusions, but that not imperative to follow the main storyline.

Editor:

I would like to thank the two anonymous reviewers and Irina Rogozhina and Raymond Sellevold for their comments and critical reading. My following remarks are thought to help in putting together your final responses.

Based on all the comments, there will have to be some expansions to the manuscript, however without having to give up on the general character of this study of being presented in a rather concise format (which I support). I suggest to pay attention to two points in particular.

1) A bit more analyses of the results (as I assumed in my access review). – I would in any case suggest to tackle the lapse rate point by sensitivity analyses, and elaborate a bit more on the atmospheric model differences (causes).

Response:
A new figure has been added to the main text (new Fig. 1), illustrating how the atmospheric “dynamics” (500 hPa eddy streamfunction and zonal wind) and “physics” (total cloudiness) respond to the horizontal resolution. These fields are discussed in relation to previous studies of resolution dependence in Earth-system models. We have also added a supporting document that presents a number of additional figures that help the interpretation (but not essential to

follow the narrative of the main text) of the climate response at the different horizontal resolutions.

We do not fully agree with the criticism regarding the choice of atmospheric lapse rate. First of all, the value 4.6 K/km is misleading as it appears to be based on a subset of the dataset presented in the cited paper (Fausto et. al 2009). When the authors include all data points they get a July lapse rate of 5.5 K/km over the Greenland ice sheet. Secondly, it is not clear that modern day Greenland is a good analog for the continental ice sheets in the last glacial cycle. The glacial climate was generally colder and drier than present, which should shift the temperature lapse rate in the direction of the dry adiabat (9.8 K/km). Case in point, Loomis et al. 2017 (http://advances.sciencemag.org/content/3/1/e1600815) argues that the atmospheric lapse rate in equatorial Africa shifted from 5.8 K/km in the modern climate, to 6.7 K/km at the LGM. Note that this example is from the tropics where both the relative and absolute humidity are expected to be higher than over mid- to high-latitude ice sheets. Thirdly, there is no consensus in the literature that one lapse rate is a better choice than another, as models typically have a range of different tuning parameters that can be used to offset the influence of the atmospheric lapse rate. For example, Stone et al (2010) reproduced the modern day Greenland geometry using lapse rates as different as 4 and 8 C/km, by instead adjusting the largely unconstrained PPD factors accordingly. Given this uncertainty and the fact that we don't actually know what the conditions were like over the Laurentide and Eurasian ice sheets, using the standard (global average) atmospheric lapse rate seems like a reasonable compromise that is in line with the simplified theme of this study. For that reason, running sensitivity experiments with different choices of lapse rates is not deemed necessary as the same argument can be used for all other unconstrained tunable parameters in both the atmosphere and ice sheet models, which would require a huge amount of work to explore properly, and no doubt obscure the main storyline. Having said that, we have added a paragraph in the discussion section on how the choice of lapse rate may influence the climate forcing.

2) More justification on the study design, see in particular the comment by Irina Rogozhina (and partly RC2). – I would like to add to this issue that the paper would also profit from clarifying the term "coupled" in the given context. Your study is clearly "oneway" (or "standalone", as you say in the paper) with regard to the ice sheet model. Most probably, for the general reader "coupled" would imply "interactive". Therefore, (a) be cautious with using "coupled" here (e.g., not in the title, see RC2 as well) and (b) provide a short argument early in the paper why standalone simulations make sense and/or are still state-of-the-art in face of the comments by Irina Rogozhina. RC2 also suggests in this regard you could argue with the objective of your study.

Response:
This is a good point. We have changed the title accordingly.

A paragraph has also been added at the end of the introduction to motivate the experimental design, and to highlight some of the simplifications we make in relation to the last glacial cycle. In short, because of the challenges of running coupled climate-ice-sheet model experiment over glacial cycles (identified in the introduction), we scale down on the realism and resort to a simplified experiment design. Although this abstraction makes the study somewhat academic (and unfortunately somewhat esoteric), it is actually beneficial in order to cleanly isolating the influence of the atmospheric resolution on the subsequent ice sheet model experiment. We feel that the addition of this paragraph is a great improvement of the study, as the objectives and general approach are motivated at an early stage.

Small things to consider:

(i) Please explain/argue briefly why only 12 years are simulated with the AGCM. For the general reader, this might be unclear since you talk in the introduction about ice sheet response time scales of hundreds to thousands of years.

Response:
As stated above, a paragraph has been added to the introduction to motivate the experimental design and put this in perspective with the real glacial cycle. A number of reminders of the study objective have also been added to the text to make it clear that the study should be viewed in a somewhat abstract light.

(ii) At which time step does the ice sheet model receive its input (daily or monthly)? Please state it explicitly

Response:
We have added a clarifying statement that the ice sheet model receives monthly climatological forcing data.

Reviewer 1:

This short manuscript presents an assessment of the influence of atmospheric model resolution in coupled climate-ice-sheet simulations. It shows that the atmospheric resolution matters enormously for an accurate simulation of the major LGM ice sheets. The manuscript is concise, clearly written, easily readable and presents an important, albeit un-surprising, result. This study can prove to be an easy to read and quick to fall back on article when introducing new and old glaciologists fresh into these kinds of simulations.

My main concern with the manuscript is whether the differences between the different atmospheric resolutions is of dynamical/physical nature, or just a matter of resolution and the model topography. The authors argue that the main cause of accurate/inaccurate simulation of the ice sheets is an inaccurate temperature field. Differences in precipitation are rather small. I think that temperature is a very straightforward parameter to model correctly: it largely depends on elevation, which directly depends on the model resolution. Therefore, this conclusion could have been found without doing any of the model simulations presented.

Response:
Simulating the temperature is actually harder than one might think as it depends on a range of physical and dynamical processes, such as the surface energy balance, and heat-flux convergence (essentially temperature advection) from the large scale atmospheric flow. Even comparatively small changes in these fields can a have large influence on both the local and global surface temperature profiles. The revised manuscript includes a new figure showing how fields typically associated with the model dynamics (500 hPa eddy streamfunction and zonal wind) and the model physics (total cloudiness) change with the horizontal resolution (new Fig. 1). We have also added a number of supplementary figures exploring sources for the temperature increase at the lower resolutions.

The supplementary figures show, e.g., how the surface temperature over the LGM ice sheets compare to the T85 case when being extrapolated to the modern topography using two different lapse rate corrections (the 6.5 C/km used here, and a lower value of 4.6 C/km that has been observed over the Greenland ice sheet in modern times). This extrapolated temperature is essentially what the ice-sheet model sees upon initialization. The extrapolated temperature, which effectively eliminates the elevation change between the different resolutions, shows a generally similar pattern as the full temperature field (Fig. 2 in the revised manuscript). The lower resolution cases typically have warmer temperatures over the areas covered by the LGM ice sheets, which we interpret as being related to increased downwelling of longwave radiation at the surface due to a general increase in cloudiness. The atmospheric dynamics appears to play a more secondary role for the surface temperature signal (also shown in the supplementary material).

How robust are the temperature anomalies shown in Fig. 2? Are these resolution differences also found for other atmospheric models (used for ice-sheet forcing)? Or is it CAM3-specific? The authors should make this finding more convincing by comparing (more) the results to other studies. Currently, in the discussion section, much attention is on the quality of the T21 forcing. I would like to see some more focus, in this section, on the intermodel resolution differences. In order to warrant publication this concern should be addressed and made less qualitative.

Response:
This comment hits at the very heart of the study. As stated in the introduction and in the discussion section, the apparent degradation of the model climate we find between the T31 and T21 resolutions appears to be largely independent of model physics, and therefore rather general among models. Polvani et al (2004) showed that numerical convergence for baroclinic waves is compromised somewhere between the T42 and T21 resolutions, Magnusdottir and Haynes (1999) narrowed it down further and showed that the limit for a "realistic" representation of planetary wave propagation falls somewhere between T31 and T21, and Lofverstrom et al (2016) showed that even the climatological (time mean) circulation is compromised at the T21 resolution. Also, the apparent similarity in simulated ice sheets between our T21 case and CLIMBER-2 suggests that this particular model may suffer from similar shortcomings as we discuss here.

In the revised manuscript we have extended the discussion section to point out similarities with previous studies (we could only find studies that have looked at resolution dependence in the pre-industrial climate). We have also extended the discussion section to be more quantitative, and added a supplementary document where additional comparisons of the different resolutions have been made.

Possibly some of the following questions could be addressed in more detail: What do the results of this study say about current studies of coupled atmosphere-icesheet models? Have other studies been conducted with inaccurate climate forcing; is this a big issue or not? Have other studies attacked and/or addressed the simple temperature discrepancies; possibly by additional topography (down)scaling techniques, spectral diffusivity, lapse rate corrections? Does the current glaciological community realize that the atmospheric resolution is as important for the results of these types of simulations? How large is the trade-off between "accuracy" and "speed"?

Response:
The revised discussion section presents a more general discussion on the trade off between "accuracy" and "speed"; the table that has been extended to compare the computational demand at the different resolutions (numerical operations per simulated year, normalized by the T21 case).

In addition, we have also included a discussion on: (1) the robustness of these results with respect to other models (this appears to be the first study on this particular topic so we compare our results with similar studies of the modern climate), and (2) shortcomings of this particular study, and that more refined methods may help improve some of these results. However, we also raise the important point that the quality of the atmospheric climate forcing is ultimately determined by the model resolution, hence it is doubtful that a more realistic ice sheet would be simulated in, e.g., the T21 case, even if using a more sophisticated coupling between the climate and ice sheet models, as the model climate is heavily compromised at this resolution.

Whether or not previous studies have used inaccurate climate forcing is hard to say. All models are simplifications of reality, so on one level the answer is definitely yes. However, models can be inaccurate in different ways (and to different degrees depending on their complexity) as they often use vastly different physics parameterisations. The point we are trying to make when comparing the apparent similarities between our T21 case and CLIMBER-2 is that this particular model appears to suffer from similar deficiencies (perhaps related to the low resolution) that comprehensive circulation models also show when run at very coarse horizontal grids. Whether or not this is an accurate assessment is hard to say without doing a more thorough comparison of these models.

A different approach might be to use different atmospheric climatologies at several model resolutions to make the results more robust, but I understand if this might prove to be beyond the work of this study.

Response:
Extending the study to include atmospheric simulation on different types of grids (Guassian, finite difference, finite volume, spectral element, etc.) would allow us to say something more conclusive about the importance of model resolution in these types of experiments. However, that is far beyond the purpose of this study and would require a huge computational effort --- comparable to, or perhaps even beyond the scale of PMIP --- to carry out. Forcing the ice-sheet model with atmospheric data from the PMIP archive would be possible, but it is not obvious how that would benefit this particular study as the participating institute only used one resolution of their specific model. This is however an excellent topic for a potential follow up study.

Because I do like the brevity of this manuscript and the concise and to the point explanation of this problem, I do suggest publication after minor revisions. Hopefully the authors can tackle my concerns adequately. I have no additional minor comments, except for a typo on p6, line 6: considerable-> considerably.

Response:
Done

Reviewer 2:
Review of “A note on the influenced of atmospheric resolution in coupled climateâA˘Tice-sheet simulations”: The manuscript examines the effect of atmospheric resolution on ice sheet modeling forced with climate model output. The theme is certainly relevant for the emerging research on coupled ice-sheet/climate modeling, both in the context of future climate projections as well as paleo-research, and timely due to launch of international collaborative projects such as ISMIP6. To my knowledge, this topic has not been directly addressed in a systematic way like this before.

The method applied consists in forcing an ice sheet model (SICOPOLIS, using the Shallow Ice Approximation), with climate output of different resolutions corresponding to the LGM. The ice sheet model is started from zero thickness. PDDs are applied for the surface mass balance calculation.

General comment: I would like to see more analysis of the climate model in addition to see the results of applying it as forcing to the ice sheet model. I'll explain in the following. The outcome of the study, namely the identification of a threshold resolution beyond what the climate simulation quality is compromised, is a very interesting result. For this reason, I would urge the authors to expand on the reasons (that is, physical processes lacking, misrepresented, and/or well-captured in the model at each resolution) for this threshold. In other words, what makes the low-resolution model unable to capture the essence of the LGM climate? Introduction: The topic is very well introduced/motivated and the literature review is a great piece of work.

Response:
A thorough analysis of why the model climate breaks down somewhere between the T31 and T21 resolutions is not an easy task with a model of this complexity. The model used here employs the same parameterizations (solves the same equations) regardless of horizontal resolution, hence the abrupt changes that occur between the T31 and T21 grids appear to show a fundamental limitation of these types of models, rather than changes in the model equations. Similar results have been observed in experiments with both primitive equation models (dry dynamical core without any physics) and comprehensive circulation models (sophisticated physics), suggesting that the numerical breakdown is somehow related to the model dynamics. Since the dynamic equations are scale independent, we speculate that it instead has something to do with the precision of the spherical harmonic transform functions that these dynamical cores rely on. Although we are unable to point to one specific cause of this phenomenon, it has been suggested that the increased diffusion required to ensure numerical stability on coarse horizontal grids is the main culprit. We can neither confirm nor deny this hypothesis from these experiments, but since a similar resolution limit has been shown to influence the development of individual cyclones (Polvani et al. 2004), the way planetary waves propagate (Magnusdottir and Haynes 1999), and also the climatological circulation (Lofverstrom et al. 2016), the model dynamics appears to be compromised on all temporal and spatial scales, and it is therefore reasonable to assume that diffusion is to blame.

We have extended the analysis in the revised manuscript to better support our conclusions. In addition to improve the figures shown in the original submission, we have also added a new figure (Fig. 1) illustrating how the atmospheric “dynamics” (500 hPa stationary waves and zonal jet stream) and “physics” (cloud cover) respond to the horizontal resolution. The figure is discussed and put in perspective with previous modeling studies investigating the effects of resolution in climate-modeling experiments. We have also added a number of supplementary figures to illustrate how other fields respond to the grid resolution.

Comment on methodology: The method relies in strong assumptions and approximations. The simulation target is to reproduce the reconstructed LGM ice sheets (by Peltier et al.) by building them from zero thickness under a steady-state LGM forcing. In reality, there was a history of building up these ice sheets, so they are not the effect of a constant LGM climate. However, the method seems an efficient reasonable approach for the objective of the manuscript, and seems to work in the identification of a threshold for “minimum required resolution”.

Response:
That is correct. The simplified modeling approach takes a few steps away from reality as we ignore the glacial history, variations in insolation and greenhouse gas concentrations and vegetation, as well as changes in the atmosphere and ocean circulation over the ~100 kyrs from inception to the LGM. However, since the objective of the study is to illustrate that both the ice sheet expansion and the quality of the atmospheric forcing data are strongly controlled by the atmospheric model resolution, it is actually beneficial to resort to a simplified experiment design as many of the challenges with coupled climate--ice-sheet models can be circumnavigated. The drawback is of course that the study becomes somewhat academic (perhaps even esoteric), but that is deemed a necessary compromise to illustrate this phenomenon. A paragraph motivating the study objective and experiment design has been added to the introduction.

Other comments: It is difficult to follow the precipitation discussion due to the choice for the color bar. Polar latitudes have low precipitation rates, please use a suitable color bar, albeit the loss of resolution for the tropical area.

Response:
All figures have been revised to have stereographic projections and better suited color scales to highlight the main points of the study.

I would remove the words “a note” from the title after expanding the manuscript with further climate model analysis. Also, the study does not include “fully” coupled climateâA˘Tice-sheet simulations in the sense ˘ that the climate model is not influenced by the ice sheet model in any way. The authors probably chose the wording “in coupled (. . .) simulations” in the context of motivation, but the title can be misleading about the

content of the actual study. I would replace the title for e.g. "On the influence of atmospheric resolution on climate-model-forced ice sheet simulations"

Response:
That is a very good suggestion. We have changed the title accordingly.

Irina Rogozhina:
I have major concerns about the experimental design of this study and how the authors' choices affect the main conclusions of their manuscript. One of such choices is the lapse rate correction. A temperature lapse rate correction is used to derive the temperature forcing at the beginning of ice sheet simulations, which are initiated from the present-day ice-free topography (excluding the Greenland Ice Sheet), and to adjust the temperature forcing to the growth of ice masses throughout the simulations. I question both (1) the choice of the lapse rate correction and (2) the assumption of the initial ice-free conditions, which are undermining the very core of this study.

Problem (1): While the authors discuss potential impacts of PDD parameters, they do not question their choice of the temperature lapse rate (6.5C/km). Fausto et al. (2009) measured lapse rates as low as 4.7 and 4.6C/km during the June and July months, respectively (the months of the strongest ablation) on the Greenland ice sheet. This is nearly 2C/km below the value that the authors use throughout the year. The temperature lapse rate is used by the authors to correct the model-based air (or surface) temperatures for the difference between their LGM topography in the atmospheric simulations and ice-free present-day topography assumed at the beginning of their ice sheet simulations. This initial temperature forcing is crucial to the development of an ice sheet: If it is excessively above 0C during the summer period, an ice sheet will not build. A difference of ~2C/km would reduce the corrections of the near-surface temperatures over the areas covered by the former British-Irish and Fennoscandian ice sheets by 3-4C during the months that matter most for the ice sheet surface mass balance, but would not impact Arctic Siberia. Quoting from the study of Löfverström and Liakka, page 1, line 9: "Sensitivity experiments using different surface mass balance parameterizations improve the simulations of the Eurasian ice sheet in the T42 case, but the compromise is a substantial buildup in Siberia". This compromise does not have to be made: The choice of a lower lapse rate correction will trigger a buildup of ice masses over the British Isles and Fennoscandia when using the T42 & T31 climate datasets, reasonable choices of PDD parameters and a higher ice sheet model resolution (see below) but will still keep Arctic Siberia ice free.

Response:
Although we agree that an annually fixed lapse rate is artificial, it is not at all obvious that 4.6 C/km would be a better motivated choice than using the standard atmospheric lapse rate of 6.5 C/km. First or all, the comment is misleading as the value 4.6 C/km is based on a subset of the data presented in Fausto et al. (2009). When accounting for all data points the authors yield a lapse rate of 5.4-5.5 C/km for the summer months (their Table 3). Second, it is reasonable to assume that the glacial climate was drier than present (response to cooler temperatures via the Clausius-Clapeyron equation), which shifts the lapse rate towards the dry adiabat (9.8 C/km). Note that both proxy data and modeling studies show that the LGM climate was significantly cooler than pre-industrial (the fully coupled LGM simulation in PMIP2 bracket a cooling of -3.7 to -5.8 degC relative to pre-industrial; Braconnot et al. 2007). For example Loomis et al. 2017 (http://advances.sciencemag.org/content/3/1/e1600815) showed evidence that the LGM lapse

rate in tropical Africa was around 6.7 C/km, which is to be compared with the modern value of around 5.8 C/km. Also, even though modeling studies have shown that the lapse-rate correction may be used as a tuning parameter to get "reasonable" ice geometries, there is no consensus in the literature that one value would be a better choice than another. For example Stone et al (2010) simulated a realistic modern day Greenland geometry using lapse rates as different as 4 and 8 C/km, by instead treating the PPD factors as tuning knobs. Given this uncertainty, we maintain that this comment is largely unsubstantiated and that the standard atmospheric lapse rate is an appropriate choice for the objective of this study.

Having said that, there is no doubt that the atmospheric lapse rate may be used to improve the ice sheet expansion in different regions. However, the point of the study is not to be as realistic as possible, but rather to show that the resolution of the atmospheric model (all else being equal) is a potential source of uncertainty in experiments of this type (coupled or semi-coupled climate-ice-sheet simulations). Also, because of the challenges running a comprehensive model over glacial timescales (raised in the introduction), we instead resort to a highly idealized setup to demonstrate this point. It is regrettable that this makes the study somewhat esoteric, but that is unfortunately the nature of the problem.

Lastly, we have extended the discussion section to include a paragraph on ways to improve coupled climate--ice-sheet model experiments. We have also added a (supplementary) figure of the JJA surface temperature extrapolated down to the modern topography (effectively the surface temperature the ice sheet model sees upon initialization) in the different cases using the 6.5 and 4.6 C/km lapse rates. Positive JJA temperatures are found in Eurasia at the T31 and T21 resolutions even when using the lower atmospheric lapse rate.

Problem (2): A more important question is whether such corrections should be applied at all. While talking about "the influence of atmospheric model resolution in coupled climate-ice sheet simulations" (quoted from the title of the paper), one would rather think about whether the use of lower atmospheric model resolution contaminates the climate state in such a way that it becomes inconsistent with the modeled/prescribed ice sheet geometries (included as a topographic boundary condition in a climate model). The question is not whether this climate forcing can build ice sheets if it's heavily modified using lapse rate corrections but whether it can maintain reasonable ice sheet geometries when unmodified atmospheric model outputs are used.

Response:
We agree that it would be advantageous not having to use lapse-rate corrections at all, however this is not a viable option as it would require the use of a synchronously coupled climate-ice-sheet model, where all components are discretized on the same grid (projection and resolution), using a realistic SMB calculation based on the local energy balance. Any deviation from this setup will require interpolation and/or lapse rate corrections to be made. However, as explicitly stated in the first part of the introduction, such a simulation is far beyond the current

computational limit of all Earth-system models, and this is not expected to change in any foreseeable future.

This second part of this comment seems to be missing the point of the study. Whether or not the LGM ice sheets can be maintained when using the climate simulated at different horizontal resolutions is not what we are examining here. This is however an excellent topic for a follow-up study.

While Problem (1) can be easily tackled by performing additional sensitivity experiments, Problem (2) is more challenging to resolve. The authors could test T42, T31 and T21 datasets using ice sheet modeling results from the T85 dataset "reproduced to a high accuracy" (quoted from Löfverström and Liakka, page 1, line 8) to address the question, whether the degradation of the atmospheric model resolution results in ice sheet collapses consistent with their current conclusions. However, their modeled ice sheets in the T85 simulation are 1.5 to 2 km too thick relative to the existing reconstructions of the LGM ice sheet geometries. Even without additional lapse rate corrections (introduced to reconcile the difference between the ice sheets prescribed in the atmospheric simulations and derived from the T85 simulation), removing such thick ice sheets would be a difficult task for the T42, T31 and even T21 data sets. At this point a question arises: Why are the modeled ice sheets so unrealistically thick? I envision several potential causes of such unrealistic model performance: (i) The spin-up of the ice sheet model: Running an ice sheet model to an equilibrium with the LGM climate over 150 thousand years is not in line with the existing evidence. Most of the former ice sheets were short-lived (tens of thousands of years from buildup to decay) (ii) Shallow ice approximation (excluding ice stream dynamics) in combination with excessively low resolution of the ice sheet model (80 km) fails to approximate the rapid ice flow and routing of ice masses towards the ocean. (iii) The isostatic adjustment scheme may cause an exaggerated bending of the bedrock surface under the weight of growing ice sheets (I have not seen the Local Lithosphere and Relaxing Asthenosphere method being used in years).

Response:
Not exactly. The figure (Fig. 4 in the revised manuscript) shows ice thickness, not topography above sea level. This means that approximately 30% of the quantity shown is depressed below the modern bedrock elevation. With a maximum North American ice thickness of about 5-6 km, this translates to an actual topographic height of 3.5-4 km, which is in broad agreement with modern reconstructions of the LGM Laurentide ice sheet. For example, the highest point over the Laurentide ice-sheet topography in ICE-6G is around 3900 m.

The suggested improvements are orthogonal to the main objective of the study. We are not trying to make a realistic simulation of the ice evolution over the last glacial cycle, but rather to show a previously unexplored source of bias (importance of atmosphere model resolution) that we think that the research community interested in climate--ice-sheet model experiments should be aware of. That our methodology deviates from reality is no secret and we state this explicitly in the paper, however whether or not these simplifications lead to false conclusions is something future research will show.

The overall quality of the study could improve if the authors address problems (1) and (2). It can also benefit from the use of SICOPOLIS v3.3 that includes options for higher resolution, more realistic treatment of ice streams and glacial isostatic adjustment. Finally I strongly suggest that the authors improve their figures. The adopted projection strongly distorts the Arctic region, which is the main focus of the present study.

Response:
For the scope of the study we do not think that a newer version of SICOPOLIS would be advantageous. The projection and color scales were arguably somewhat lacking in the original submission and we have re-plotted all figures on stereographic projection with a refined color scale.

Raymond Sellevold:

You mention in the introduction how horizontal diffusion does not only influence horizontal motions, but may also impact vertical transport and convection. I wonder how important resolving convection adequately is for building/removing an ice sheet. It would be informative if you more specifically related resolution dependent dynamics/physics with possible shortcomings in building/removing the ice sheets.

Response:
The precipitation in the midlatitude storm tracks is predominantly large scale (stratiform), so a weakening/breakdown of the convection (used loosely here as these types of clouds are highly parameterised in climate models) is probably not a major influence on the surface mass balance by itself. However, both atmospheric convection and large-scale vertical transport are important components of cyclogenesis, so the large-scale precipitation field may be altered via more indirect processes related to vertical motions. Also, the fact that the quality of virtually all atmospheric fields is compromised at sufficiently coarse horizontal grids suggests that there may be a fundamental problem with low-resolution models that transcends both model tuning and complexity.

We have added figures comparing the large scale circulation to illustrate how the different resolutions compare to one another. Fields related to both model dynamics and physics are shown and discussed in relation to previous studies.

In your method section, you describe that present day non-glaciated areas are prescribed with modern day vegetation cover. Do you think this assumption is valid? Do you think it may have a large influence on the simulations?

Response:
The pre-industrial vegetation is almost certainly impacting the ice extension, even though the influence should be largely comparable at all resolutions and therefore not of first order importance for our main conclusions. No reliable reconstruction of the LGM biome exists, so the modern vegetation cover is actually used in the official PMIP boundary conditions (https://pmip4.lsce.ipsl.fr/doku.php/exp_design:lgm), unless the model can simulate vegetation changes interactively (not an option in CAM3).

Note that the AGCM simulations have prescribed ice sheets in the areas indicated by the black contours in the figures, hence the surface albedo is high in those areas regardless of the model resolution. Differences in the simulated ice sheets (e.g. in western Eurasia and to some extent in the interior of North America) is therefore not to first order a response to the (modern) vegetation cover.

We have added a clarifying statement about the use of modern vegetation in the text, and also cited the paper outlining the official PMIP4 LGM simulation design.

You present the smoothed topography as a reason why some areas are warmer when the resolution is lower. After that, areas that are colder with respect to the T85 case are pointed out with no explanation why they may be colder.

Response:
There can be any number of explanation for local positive and negative anomalies in the difference fields. In addition to differences in the simulation quality at the different resolutions (note that the objective of the study is to illustrate that there is a general degradation of the simulation quality on coarser horizontal grids), the short (10 year) climatologies are certainly one possible explanation for these anomalies. Also, comparing gridded data on different resolutions can naturally give rise to small scale "inconsistencies" like that. Note that there is only one value of temperature in each grid cell, so even if the 4x4 cluster on the T85 grid that covers one cell on the T21 grid has the same average temperature as the T21 cell, grid cells one side of the cluster might be cooler than the average, while the other side might be warmer. Subtracting the two would then yield local warm and cold anomalies even if the area means are the same. Examining this further is however beyond the scope of the study as we try to focus more on the large scale structure, rather than local scale variations in the difference fields.

The tropical and midlatitude precipitation fields are well covered in the text, but there is not much mentioning of precipitation over the ice sheets. It is also hard to see the difference of this field between different resolutions, since the colors starts at 200 mm/yr. A possible solution would be to make a non-linear color scale to better resolve the low-precipitation areas (such as the ice sheets).

Response:
That is a fair point. We have changed the projection and the color scale to better illustrate the changes over the ice sheets. We have also made appropriate changes to the text where we describe the precipitation field.

"The ice sheets forming under the high resolution atmospheric climatology (T85; panel 3a) are in close resemblance with the target extent (indicated by solid contours; Kleman et al., 2013). There is essentially only too much ice extending along the Siberian Arctic coast." Why is there too much ice extending there? Because it is very cold, because the precipitation is very high, or maybe a combination of both?

Response:
It is probably a combination of both, even though the temperature field is the most important contributor. Explaining the cause is however outside the scope of the study and therefore not examined further.

”However, the T21 resolution only has ”functional support”, which means that boundary conditions are provided but the model climate has not been tuned to the same standard as the other resolutions (the resolution dependent tuning parameters are broadly the same as in the T42 case).” How is the 1850-present climate in this low resolution as it is not tuned? Was there any attempt made to tune it? If no, why not?

Response:
Well, the climate is tuned but not to the same rigour as the other resolutions. Also the modern (pre-industrial) climate simulated on the T21 grid is quite different from observations (not shown).

No attempt was made to tune the T21 climate. Even though CAM3 is a comparatively simple model with modern day standards (e.g. CAM6, the newest addition to the NCAR model family, is much more complex), it is still a comprehensive AGCM with twelve independent tuning knobs. Re-tuning such a model is a herculean task that is extremely difficult and therefore typically carried out by a team of experts.