Authors' Response to Reviewers' Comments:

Firstly, the authors express their appreciation to the two reviewers and the editor. We believe that their vital comments and suggestions have contributed substantially to improve the presentation of our study, as well as its overall quality and the manuscript. Following, we offer point-by-point replies to the issues and points the reviewers addressed regarding the original manuscript.

**Reviewer #1**

General comments

*This manuscript provides an assessment of the LGM permafrost distribution in coupled AOGCMs on the basis of simulated soil temperatures (direct method). For this purpose, results from the PMIP3 database were used. The analysis is compared with previous work that categorized permafrost from surface air temperatures (indirect method) taken from PMIP2 simulations of the LGM climate. The results certainly provide useful information on the performance of climate models under LGM conditions. It is shown that the models are generally capable of simulating a reasonable distribution of permafrost in the present-day and LGM climates.*

*In my view, the manuscript requires a moderate revision before it can be published. As detailed below, my main points concern the presentation of the results and the need for an evaluation of the inferred LGM permafrost distribution using reconstructions.*

Reply: We are grateful for the reviewer's precise, detailed, and constructive comments and suggestions, after which we have revised the manuscript text, tables, figures, and reference list as described below and in the revised manuscript. We have also elaborated upon our results and discussions with additional analysis and comparison with other field data or reconstructed maps, as suggested.

Main comments

*I) An important part of the manuscript is devoted to a comparison of the direct and indirect methods. It is concluded that both methods have their advantages and disadvantages, and that the direct methods does not necessarily perform better. However, this is mostly based on visual comparison of (small!) maps, making it hard for the reader to assess the performance of the two methods. It would therefore be helpful if Table 2 could be extended by including estimates for different key regions and also include observed modern areas of frozen ground. This would allow a more thorough* *evaluation of the two methods using the 0k simulations. It would also make clear where the models do well and where they have problems.*

Reply: We agree that such an extension of Table 2 is beneficial, and we have augmented Table 2 by adding estimates for the following four regions—Europe (20ºW-60ºE, 40ºN-75ºN), Asia (60ºE-170ºW, 40ºN-75ºN), North America (170ºW-40ºW, 40ºN-75ºN), and the mid-latitudes (0ºN-40ºN), as well as results from the modern observations (column 4).

*II) Page 1569, line 14: ': : : this is a perfect time for assessing the models' ability to reconstruct LGM frozen ground, as there is an effort from the observational side for an Action Group of the International Permafrost Association (IPA) to compile and publish in 2013 an evidence-based map of maximum permafrost extent during the last glaciations period : : :'. Wouldn't it make more sense to wait with the present paper until this evidence-based map is available, so that the model-results can be directly compared with this map? This sounds like a missed opportunity to me. In any case, a map of reconstructed LGM permafrost should be included in the manuscript as a reference to evaluate the LGM simulations. If the IPA-map is not yet available for this purpose, a map should be constructed based on previous work. For instance, Vandenberghe et al. (2012) provides a map with LGM permafrost limits for Eurasia, and similar reconstructions exist for North America. A comparison of the presented LGM permafrost distributions with a map based on reconstructions would certainly have added value. Based on the reconstructed map, also estimates for the LGM could be added to Table 2, similar to what I have suggested for 0k (see my comment I).*

Reply: Unfortunately, the IPA's LPM (Last Permafrost Maximum) map was not yet ready to be included in this paper (in production). We agree that such a reconstruction map of LGM permafrost distribution of global (or even Northern Hemisphere) coverage would be of tremendous benefit as to enable estimation of the areal extent of frozen ground at LGM. At the same time, we are aware that such a new reconstruction from patchy local field-based evidences and/or fragmented regional maps would require overwhelming effort.

In the original manuscript, we already cited the southern boundaries of permafrost reconstructed from evidence referenced at ll.16-25, p.1579. However, we do agree that it is misplaced and also that additional comparison with field-based evidence will be beneficial; therefore, we have moved the citation to section 4.2, and we have included the boundaries and locations of reconstructed permafrost boundaries in Figures 2b, 2d, and 2e.

*III) Page 1578, line 18: 'These results imply larger differences between simulated subsurface thermal regimes among the models, which may be due to differences in implemented physics regarding freeze-thaw processes among the models. The high diversity found in the PMIP2 0 k map owes partly to differences in boundary conditions among the models, such as ice sheets, land/sea mask and orography distribution'. In my view this formulation('may be due to differences', 'owes partly to') is too vague. Please perform a more detailed analysis of the reasons behind the differences between models.*

Reply: As for the simulated subsurface thermal regime, high diversity resulted mainly from the previous "Discontinuous" category—with soil layers frozen for more than half of the ten-year period. This category is very dependent on the total depth of the model's soil column. A model with a shallow soil column tended to show large annual and interannual temperature variations at the bottom layer. Introduction of the revised classification (new equations 1-3, with three categories instead of the previous five) removed much of this fluctuation (cf. new Figs 3a and 3b). We have removed the first sentence in the revised manuscript.

As for the PMIP2 results, correlation analysis showed that diversity among the PMIP2 models was more directly related to the variability of simulated air temperature rather than that of boundary conditions, though the former remains largely influenced or modified by the latter. Hence, we have revised the sentence accordingly.

*IV) Most results are presented in global maps. However, in my opinion these maps are not very well suited, as they are too small to see the sometimes subtle differences and regional details discussed in the text. I would suggest to focus the results on the Northern Hemisphere and to provide circumpolar maps with a polar projection.*

Reply: We have enlarged the main figures and have changed to polar stereographic projection, focusing on the Northern Hemisphere results. In addition, we have provided the global maps as Supplementary Figures.

Minor comments

*- Page 1568, line 25: typo, 'uases' should be 'uses'*

We have revised the typo.

*- Page 1569, lines 8-10: ': : : these comparisons produced mixed results consistent with evidence in some regions but not in others, including north of the Alps etc'. It is not clear from this sentence if the evidence north the Alps is consistent or not. Please*

*rephrase.*
We have revised the sentence to make it clear.

*- Page 1569, line 21: 'The issues this study attempts to address are:'. I would suggest to revise into 'The issues addressed in this study are:'.*
We have revised the sentence accordingly.

*- Page 1569, line 26: 'How is the information regarding modeled surface and subsurface temperatures on a grid box associated with and consolidated into the frozen ground zonation in the area represented by the grid?' For me it is not very clear what is meant here and I had to read this sentence a couple of times. So I suggest to rephrase.*
We have simplified the sentences to clarify as follows: "How do the frozen ground distributions based on air temperature (indirect method) and those from ground temperature (direct method) compare?"

*- Page 1570, line 12: 'Summary of the used models, institutes or groups, and simulations is summarized in Table 1.' The summary is summarized? Please rephrase.*
We have revised the sentence.

*- Page 1571, line 12: 'using the last ten years of the simulations.' Why are only these last 10 years used in the analysis? What is the rationale behind this? Why not take at least 30 years as is normally done for climatological analyses?*
We have redone the entire analysis using the thirty-year data as suggested. We have confirmed that this has not produced any unexpected or substantial changes in the resulting maps and conclusions.

*- Page 1572, line 12: I suggest to make clear that this paragraph is only relevant for the analysis of the PMIP3 results.*
We have revised the paragraph including this sentence to clarify.

*- Page 1572, line 21: 'Continuous': the bottom soil layer is frozen (at or below 0C) for the entire period. 'Discontinuous': the bottom soil layer is frozen for more than half of the period. 'Seasonal': the top soil layer is frozen for more than 30 % of the period. The definitions should be redefined so that the categories exclude each other. For instance, in the definitions as presented here, 'Seasonal' can occur at the same time as 'Continuous', since the bottom layer can be perennially frozen and the top soil can be*

*frozen for more than 30% of the time as well. Obviously, 'Seasonal' only makes sense if the bottom soil layer is not frozen. Please clarify. In addition, please make clear what is meant by 'the period'.*

Following the comments from both reviewers, we have simplified the classification criteria for soil temperature (direct method) to remove the ambiguity, and we have shown clearly that the classification is mutually exclusive.

*- Page 1572, line 28: For the PMIP2 results, the surface air temperatures are used to estimate the distribution of permafrost. Is this the temperature at 2m, or really at the surface? Is the snow pack and its thickness as calculated by the model taken into account? A thick snow pack may isolate the soil from the cold atmosphere above, so if the surface air temperatures are used, this may result in too cold estimates for the ground beneath. Please discuss*

Near-surface air temperature at 2 m was used for the indirect method. We have revised the text to show this clearly. The influence of snow pack on the index method and, consequently, differences in using *air* and *surface* freezing/thawing indices are illustrated in Saito et al. (2013b), which has been referenced in the revised manuscript.

*- Page 1574, line 15: Do all models used in this study have 365 days per year? Some model use 360 days per year (i.e. equal months of 30 days) for efficiency.*

The days in a year used in the simulations have been added to Table 1. All PMIP3 models used a 365-day cycle, while some PMIP2 simulations used a 360-day cycle. Nevertheless, we applied a 365-day cycle in the analysis in order to make comparison of freezing/thawing indices (computed as a summation of a monthly temperature multiplied by the number of days in the month) among models meaningful. We have added a mention of the use of a 365-day cycle.

*- Page 1576, line 16: 'The LGM permafrost maps reconstructed from observational evidence in the previous studies shows: : :' It is not clear to me what previous studies are meant here. Saito et al. (2012, 2013)?*

We meant previous reconstructions based on the field studies (Baulin et al., 1992; Petit-Maire et al., 2000; French, 2007). We have revised the sentence to make this clear.

*- Page 1577, line 20: 'Continuous permafrost increased at 21k relative to 0k in all cases, though the differences varied among the experiments and methods from 2 to 13 million km2, partly due to coarse horizontal resolution.' How does the coarse horizontal*

*resolution affect the permafrost area? Please elaborate.*

The coarse resolution definitely affects local details of permafrost distribution. However, additional analysis performed after the submission revealed that the warmer climate of PMIP2 relative to PMIP3 contributed more to the smaller areal extent of permafrost by PMIP2 than did the horizontal resolution. Therefore, we have revised the sentence accordingly. We have also changed the variables to relative extent (percentage to the land area) rather than absolute extent (million $km^2$), as the areal extent of the entire land varies according to the horizontal resolution and the land/sea mask employed in the simulations, especially in PMIP2.

*- Page 1579, line 29: It would be good to note that there is no 'good correspondence' between the modelled boundaries and field-based evidence for LGM-permafrost in Western Europe. See discussion in Vandenberghe et al. (2012, QSR).*

The PMIP3 models, still not fully compatible with field-based knowledge, have shown great improvement from the PMIP2-generation, and the success of reconstruction of permafrost in the region was already mentioned at l. 28, p. 1576-l. 1, p. 1577 in the original manuscript. We have elaborated to offer more information about this issue, including the result of Vandenberghe et al. (2012).

*- Page 1582, line 20: 'Larger inter-model diversity of soil temperature based distribution has implied that the subsurface regime is still at the development phase'. This sentence is not clear to me. Please explain more clearly what you mean here.*

We have revised the sentence to make it clearer: "That larger inter-model diversity was found in the frozen ground distribution from the direct method (i.e., based on soil temperature) than in that from the indirect method (based on near-surface air temperature) has implied that the implemented subsurface regime in the GCMs is still at the development phase in comparison to the implemented atmospheric processes."

*- Page 1582, line 25: Including snow dynamics is mentioned here as a possible improvement. However, the PMIP3 models already include snow dynamics, implying that in the soil temperature calculation this is already explicitly taken into account.*

We noted the possible improvement of coupling within the physical dynamics, as well as with the biogeochemical ones, and we included "snow dynamics" as an important factor for the latter, but we agree that it was not adequately phrased. We have revised the sentence as below to clarify our intention: "Coupling, such as between thermal and hydrological processes and with biogeochemical processes at and below the surface (including snow-vegetation dynamics),…"

*- The reference list is incomplete. At least three references are missing: Koven et al. (2013), Vandenberghe & Pissart (1993); Vandenberghe et al. (2012).*

We have revised the reference list substantially, following the comments of both reviewers.

*- Figure 2: Please make the legend bars consistent.*

We have made the legends consistent throughout the figures.

*- Supplementary Figures 7-10: What does white shading signify? By the way, I am not convinced that providing the numerous maps in the supplementary figures is necessary.*

We have restructured all Figures and Supplementary Figures, to focus on the main issues (i.e., reconstructed frozen ground distribution). In doing this, we have removed the individual maps for freezing/thawing indices and MAAT (original Supp. Figures 5-10).

**Reviewer #2 (Prof. Jef Vandenberghe)**

General:

*The paper is a clearly presented evaluation of the potentials of PMIP3 modelling in progressing in the reconstruction of the spatial extent of permafrost and the subsurface thermal states in different kinds of permafrost. It is especially positive to see that the numerical results are confronted with 'evidence-based' permafrost reconstruction. The comparison of the new findings with the results based on PMIP2 modelling is interesting. The objectives of the research are well defined at the end of section 1 (p. 1569-1570), while in the last section, apart from the conclusions, suggestions are given for continued future research in the investigated domain (e.g. perspectives to be expected from transient permafrost modelling). As a consequence, I have no hesitation to recommend this relevant paper for publication in CP.*

We appreciate greatly the minute and instructive comments and suggestions by Prof. Jef Vandenberghe. We have elaborated our work to incorporate the suggested analysis and modifications.

One major concern, a few minor comments and suggestions:

1. *The transfer from an areal zonation of frozen ground to temperatures as derived from GCMs is not trivial (p. 1572, from l 11 onwards to the end of the section on p. 1573). In such a case the authors have to be careful with their definitions: on l. 14-15 p. 1572 they state 'if the top soil layers freeze and thaw annually, it is seasonally frozen'. This has no sense as seasonally frozen ground freezes and thaws annually (as correctly written in l. 9-10) while the uppermost soil layers in ALL permafrost ground haw in summer. The next sentence (l. 15) is difficult to understand 'if the temperature remains above 0C, it is not freezing.' Maybe the authors mean 'temperatures all over the year', which is evident?*

   *But more important is the determination of frozen ground zones by Tsl-based criteria from l. 19 on p. 1572 onward. It is very confusing and strange to see appearing now other definitions of continuous- discontinuous-seasonal frozen ground then those used commonly and before in the paper (top os section 2.2 p 1571-1572). Now the zonation is time-based (entire period, half of the period, 30% of the period) instead of area-based. Firstly, it is absolutely not clear why this is done. Secondly, the basic arguments for the new definitions are not given. Thirdly, it is very confusing to introduce definitions that are different from commonly applied ones.*

Reply: Following the comments and suggestions from both reviewers, we have revised the classification of the direct method to make it plain and straightforward, and to eliminate any contradictions.

2. Minor:

*-p 1568 l 9: One of the first papers recognizing the important effects of the water content of he upper soil was by Renssen et al. 2000 in EPSL.*

We have added the suggested citation.

*-p 1568 l 29: I suggest to add reference to French 2007 in addition to a reference to Saito 2013.*

We have revised accordingly.

*-p 1572 l21-25: add the abbreviations you use later on (pr- tr-sf : : :.)*

We have added the abbreviations, though the names of those abbreviations have changed.

*-p 1576 l 2-3: In my opinion, the discrepancies you mention could also be due to your definitions of Tsl at p 1572.*

We have revised our definition of the direct method to eliminate arbitrary factors or shifts from criteria to time. Nevertheless, we agree that the classification criteria (Equations 1-3) has room for improvement. We have shown a result of such an attempt in section 5 (revised Figure 5b).

*-p 1577 l 2 and p 1580 l 3: I suggest to add references to Vandenberghe et al. 2008 and 2012.*

We have revised accordingly.

*-p 1579 l 8-10: The max. ALT is also dependant on local factors as vegetation and snow cover as you mentioned before.*

We have revised the text to include the local factors as suggested.

*-p 1579 l 21: better to replace ‘Vandenberghe et al 2004’ by ‘Vandenberghe et al, 2012’.*

We have revised accordingly.

*-p 1580 l 13: I suggest to insert ‘at large scale’ after ‘distribution’. And I suggest to remove ‘other’.*
We have revised accordingly.

3. - *The results of the modelled areal extent as described on p. 1577 from l. 14 onward are very interesting. But –as a suggestion- the interpretation of the differences between the 2 modelling approaches and the differences between the extent of individual permafrost zones at 0 k and 22 k could be further expanded.*
Following suggestions by both reviewers, we have extended the analysis summarized in Table 2, by examining areal changes in the following Northern Hemisphere regions: Europe (20ºW-60ºE, 40ºN-75ºN), Asia (60ºE-170ºW, 40ºN-75ºN), North America (170ºW-40ºW, 40ºN-75ºN), and the mid-latitudes (0ºN-40ºN),.

*-On p 1581 l 14-17, I suggest to compare the modelled 5ºC increase since 21 k with the 13-15ºC difference reported from ‘evidence-based’ reconstructions in Europe and China (e.g. Vandenberghe et al. 2004; Huijzer and Vandenberghe 1998 in JQS).*
We have added comparison with the evidence-based temperature decreases in Europe and China, according to the suggested references.

Technical remarks:
*-references: In addition to the missing references reported by referee 1, I mention one more missing reference and some inaccuracies and advise to remove many references from the list that are not used in the text: additional missing reference: Peltier et al. 2010.*
We have checked the cited references and revised the reference list.

-textual:
*-p 1567 l 13: change ‘Franzel’ by ‘Frenzel’; p 1568 l 14: change ‘Talor’ by ‘Taylor’;*
*-p 1568 l25: replace ‘uases’ by ‘uses’*
*-p 1569 l19-20: I suggest to replace ‘and H. French, personal communication, 2013 ’ by ‘coord.’*
*-p 1577 l 2 and p 1580 l 3: replace ‘Nachaev’ by ‘Nechaev’.*

*-p 1589 l 25: insert ‘the’ before ‘Introduction’.*
*-p 1581 l 3: insert ‘ºC’ after ‘10’. p 1581 l 4: insert ‘, the’ before ‘southwestern’.*
We have revised accordingly.

*-p 1568 l 16: ‘Koven et al. 2013’: in Ref list it is written ‘2012’.*
*-p 1568 l 20: ‘Boeckl et al’: in the Reference List it is written ‘Boeckli’ and you should indicate whether it is a or b;*
We have checked the cited references and revised the reference list.