Reply to Reviewer 1

We thank the referee for the detailed and constructive criticism. In the following, we address the comments on a point-to-point basis. We also indicate the changes that we anticipate will be implemented in our manuscript. In summary, in response to the reviews 1 and 2 we will

- introduce our goals as key hypotheses to frame the manuscript
- reword individual sentences and shorten text in Methods and Results.
- remove Fig. 11 from the manuscript, include melt-water forcing in an additional panel in Fig. 1, rearrange Fig. 12 and modify labels in Figs. 2 and 3.

This will allow us to improve the manuscript substantially, and increase its impact.

Response to scientific comments

Referee: "The manuscript presents an analysis of global climate variability during the last deglaciation based on multi-model and reanalysis data. The authors collected climate model simulations of the last deglaciation with different climate models of different levels of complexities or experimental protocols. The authors analyzed global temperature and precipitation variabilities using multiple indicators for variabilities. The authors found increased climate variabilities during the last deglaciation than the LGM or Holocene with specific timescales and regions. The authors also find that the variability during the last deglaciation is affected by the complexity of the climate models or experimental design protocol.

I think this study's topic is well-suited for Climate of the Past, and the method and analysis of this study, particularly for introducing multiple variability indicators and analyzing both temperature and precipitation, is unique."

Authors: *We thank the reviewer for the thorough assessment of our study and appreciate that she/he noted that the study fills a gap in the literature.*

Anticipated changes: None.

Referee: "However, the manuscript needs additional work to improve the readability particularly for the following two points. Firstly, many figure panels, including supplemental figures (S41), are referenced in the manuscript (but some supplemental figures are not referenced), making it hard to follow. A multi-model study with global analysis may need many figure panels, but I had tough time understanding figures (Do Figs 6-11 need 24 or 27 panels?). I wonder if there is a better way to show figures in a more structured way to help readers."

Authors: *We agree with the reviewer that Figures 6-11 present a wealth of information, highlighting the differential role of model type, version, boundary conditions, and forcing on climate variability in transient simulations. In Fig. 6, for example, we investigated regional effects of meltwater and volcanic forcing on centennial-scale variability. Here, we were able to compute difference fields between different protocols, and could reduce the number of panels. Standard deviation patterns, for temperature, were sufficiently similar (and previously described in the literature, e.g. Rehfeld et al., ESD 2020), that we condensed these onto a latitudinal view in Fig. 6. However, to allow for intercomparison between Figs. 7-11, so skewness and kurtosis of temperature (7,8), as well as standard deviation and skewness of precipitation (9,10) we would prefer to keep the same Figure layout for a better comparison. Centennial-scale kurtosis of precipitation appears largely model-insensitive, and we therefore will remove Fig. 11. To enhance readability we will combine Fig. 25 (interannual kurtosis) and Fig. S24 (decadal kurtosis) and Fig. 11 (centennial kurtosis) for MPI-ESM r7 into one 3x3 plot, and refer to this in the discussion instead.*

Anticipated changes: We will remove Fig. 11 from the main manuscript (kurtosis of precipitation fields) and adjust the discussion.

Referee:"Secondly, the introduction section seems to lack information on what has been done regarding climate variability during the last deglaciation and what the knowledge gap is. As in the discussion section of this manuscript, there's a proxy study (e.g. Rehfeld et al. 2018) and climate modelling study (e. g. Zhu et al. 2019; Shi et al. 2022) on climate variability during the LGM or the last deglaciation. I think their methodology and results can be summarized in the introduction, and the authors can clarify what knowledge is lacking and what this study's strengths are. I also think stating a hypothesis in the introduction will help clarify the key points of this study."

Authors: *We thank the reviewer for this suggestion, which we agree would better frame the study, from the introduction to the discussion. We may hypothesize that (1) patterns of surface climate variability are state-dependent for the quasi-equilibrium conditions of the LGM and the Holocene, (2) the deglaciation, with the Earth system transitioning between a cold and a warm(er) climate state, stands out in the higher moments of precipitation and temperature distributions, (3) state-, model- and forcing-induced changes in variability are amplified with increasing timescales.*

Anticipated changes: We will include a paragraph on the current knowledge gap on LGM and glacial variability in the introduction, and reframe the last part of the introduction with the key hypotheses.

Specific Comments:

Referee: "L8-L9: The phrase "largely unexplored" might be too general. This sentence can be more specific based on previous knowledge gap or strength of this study."

Authors: *We note that, indeed, this sentence is very general, and a clear framing of (i) what we denote as climate variability and (ii) what the key knowledge gaps are improves the impact of the study.*

Anticipated changes: We will rewrite lines 8-9 of the abstract to highlight the focused definition of climate variability and the knowledge gap.

Referee: "L27-L28: I'm not sure what is unclear. Do you mean it is unclear whether LGMR (Osman et al. 2021) simulates accurate spatial patterns of climate variability?"

Authors: *The LGM reanalysis by Osman et al. draws on model simulations and a proxy dataset to reconstruct a spatio-temporal evolution of surface temperature changing from the LGM to the Holocene. Indeed, the LGMR reconstruction shows similar climate variability in the global mean to most models. However, by definition, the data assimilation procedure merges properties of the proxy data and the underlying model ensemble, following an algorithm. It is therefore not quite clear which one of these three aspects dominates in a reconstruction for a given region and timescale. However, patterns of centennial temperature variability change, in particular over the deglaciation, are smoother, and closer to normally distributed in LGMR than they are for most of the model simulations. Still, we cannot at this point clarify whether these patterns are accurate or not. Research into palaeoclimate data assimilation must clarify the impact of reconstruction methods on the higher moments of temperature, and ideally also precipitation, distributions.*

Anticipated changes: We will rephrase the statement in L27-28 as "A reanalysis of the LGM exhibits similar global mean variability to most of the ensemble. However, paleoclimate data assimilation combines model and proxy data information using a Kalman filter-based algorithm. More research is needed to disentangle their relative impact on reconstructed levels of variability."

Referee: "L72-L79: I'm not sure what the point of this paragraph is. I wonder if L71 and L80 can be directly connected to state the importance of climate variability and what proxy says on climate variabilities in the last deglaciation."

Authors: *We understand that the reviewer would prefer us to introduce climate variability more concisely and pragmatically with the deglaciation and proxy data in mind first. We agree that this would likely increase readability, and will re-order the paragraph so that it better leads to the abovementioned key hypotheses.*

Anticipated changes: We will reframe the paragraph in the introduction leading to the key hypotheses.

Referee: "L134-L140: I understand that one strength of skewness is that it can be an indicator of abrupt climate change, according to this paragraph. There would be a discussion paragraph on whether skewness in the deglaciation simulations can be an indicator of abrupt climate changes. "

Authors: *Indeed, we find clearly outstanding patterns of skewness and kurtosis change over the deglaciation in the transient simulations. Our simulation ensemble undergoes large-scale, sometimes abrupt, changes due to prescribed boundary conditions and forcing. Regionally there may be abrupt change due to internal dynamics of atmosphere, sea-ice, ocean or land surface. In light of this, focused future work could investigate this by drawing on joint analyses of high-resolution proxy data for key variables of tipping elements, with coupled model simulations.*

Anticipated changes: We will include this consideration in the introduction and discussion.

Referee: "L141-L151: As far as I understand, applying skewness and kurtosis to paleoclimate is new in this study, which can be emphasized."

Authors: *Indeed, we are not aware of other studies that, systematically or otherwise, investigated higher moments than standard deviation for paleoclimate.*

Anticipated changes: We will include a statement to the effect that we are, to our knowledge, the first to investigate higher-order moments in paleoclimate in the abstract and conclusion.

Referee: "L181: Is dd/m always used as GMP, global mean precipitation? Please clarify."

Authors: *Units are always in Celsius (for temperature) or mm/d for precipitation. In the text this is always noted. It is also the case for all analyses and figures.*

Anticipated changes: We will add a sentence clarifying that we always use Celsius/$\frac{mm}{d}$ as units for temperature/precipitation.

Referee: "L187: I don't understand what is different between MPI-ESM r1&r6 and r2&r5, as all columns in Table 1 are the same. Are they from simulations with different model parameters in Kapsch et al. (2022)? One way is to add a reference column in Table 1."

Authors: *This is correct. These simulations are described in Kapsch et al., (2022) and they differ in parameter choice.*

Anticipated changes: As suggested, we will clarify the differences between the simulations in the text by adding one sentence in the model section, and we will include a reference column in Table 1. For the description, we will change L189-191 to: "They use different sets ice sheet reconstructions – GLAC1-D or ICE-6G_C (in the following ICE6G, Peltier et al., 2015) – and vary by meltwater scenario. Further, a parameter for cloud formation was changed in r5–r7 to remove a cold bias found in r1–r4 (as detailed in the supporting information of Kapsch et al. 2022)."

Referee: "L465 & L470 "centennial" instead of "decadal and centennial"? Because Figure 6 say centennial standard deviation."

Authors: *We thank the reviewer for pointing this out. This should read "centennial".*

Anticipated changes: Correction to "centennial timescales".

Referee: "L513-L523: Based on Figures 7, 8 it is discussed that volcanic forcing impacts skewness and kurtosis during Holocene based on MPI-ESM r6 and r7 simulations. However, Figures 7, 8 and S12 make me feel that with the volcanic forcing, MPI-ESM simulations resemble HadCM3 simulations despite HadCM3 not having volcanic forcing. Are there any discussions for this model difference? "

Authors: *We indeed discuss in Sect. 4.3.3 the difference that volcanic forcing on the simulated higher-order moments of the surface climate distributions makes. Surprisingly, perhaps, the patterns generated by HadCM3 intrinsically (i.e., without volcanic forcing) are similar to those simulated by MPI-ESM with volcanic forcing. We have not spelled this out previously.*

Anticipated changes: We will explicitly mention the similarity of MPI-ESM r7 to the HadCM3 simulations in Sect. 4.3.3.

Referee: "L614-L618: Is it because (a) volcanic forcing or inter-annual to centennial variability, or (b) volcanic forcing does not correspond to timescale variation, but it can induce inter-annual to centennial variability?"

Authors: *Volcanic forcing both acts to stimulate temperature variability linearly (i.e., on the timescales on which it is active), as well as on timescales that are longer. This was noted, amongst others, in Ellerhoff et al., 2022, and we do as well, in the discussion, albeit without reference to the literature and without distinguishing the linear and nonlinear impact.*

Anticipated changes: We will rephrase the inital statement to "Volcanic forcing strongly impacts the spectrum of simulated surface temperatures. The MPI-ESM r7 run, which includes it, has the largest PSD [...]". We will further add the appropriate reference in the discussion.

Referee: "L694-L714 and Figure 13: I couldn't understand the point of this subsection, and why Figure 13 is necessary for discussing variability uncertainty. Please add some introduction."

Authors: *This section, and Fig. 13, are geared to give context for the expected mean-state change between the LGM and the Holocene. This is important, because one of our key goals is to understand whether the patterns in higher-order moments we see are state-dependent, i.e. dependent on the global-mean change. This section thus provides a backdrop to the discussion of variability change. We will revise the wording of this section to align with the hypotheses outlined in the introduction to improve readability.*

Anticipated changes: Rewording of title and introductory text of Sect. 5.1.

Referee: "L735-737: You mean that the meridional temperature gradient is enhanced during LGM as in Shi et al. (2022), but the variance is not increased like Shi et al. (2022)? If so, isn't it a significant result worth emphasizing and discussing further?"

Authors: *Changes in the meridional surface temperature gradient have been suggested to drive interannual temperature variability change between the LGM and the Holocene on multi-centennial to millennial temperature variability (Rehfeld et al. 2018). Furthermore, Shi et al., 2022 suggested that, for the extratropics, there is a significant spatial correlation between the gradient change, and the variability pattern that they link to interannual variability. They also, indeed, suggest that temperature variability is increased by some 20% compared to the Holocene based on a PMIP3/PMIP3 model ensemble, which is consistent with the 25-31% increase found in the PMIP3 ensemble considered in Rehfeld et al., (2018). In our ensemble we found that, for some models and some regions, this notion holds, whereas for other regions and models there is no correlation between the temperature gradient and variability change. Overall we do find an enhanced temperature variability, and we do find an enhanced temperature gradient in the LGM, which is consistent with the previous* *studies. We will emphasize this in the revision.*

Anticipated changes: Revision of text in Sect.5.2.1 to expand on and better present the results with respect to the relationship between the change in temperature gradient, and the change in surface temperature variability.

Referee: "L744-L745: In addition to long-term memory, there's transient forcing during 23 to 19 ka (Ivanovic et al. 2016), unlike equilibrium LGM simulation at 21ka."

Authors: *We thank the referee for pointing out that the transient change in forcing over the deglaciation is another external source of variability on centennial to millennial timescales. We missed to include it in this statement because the previous paragraph was referring to PMIP3/PMIP4 simulations, where these do not play a role. For the revision we will more clearly distinguish the two types of simulations in this section.*

Anticipated changes: We will reword the sentence to read "Since the differences are especially apparent on longer timescales, this might point towards long-term memory effects or transient forcing missing in such equilibrium simulations.".

Referee: "L884 (minor): Why "using an EMIC more focused on atmospheric dynamics" , unlike [using a GCM when focusing on climate variability] "

Authors: *Indeed, dependent on the timescale of key interest a GCM may be computationally efficient enough to be suitable for simulation of centennial to millennial climate variability (as evidenced by the selection of simulations we draw on in our study). We meant to emphasize the fact that most EMICs have a reduced atmospheric complexity, which limits their simulated variablity (e.g., Schillinger, Ellerhoff et al., 2022).*

Anticipated changes: We will rephrase the sentence to "The EMICs included here have reduced atmospheric complexity. This will affect simulated variability and could be different in other EMICs as it is for GCMs."

Referee: "L892-L893 (minor): Each simulation from previous articles used in this study focused specifically on the atmosphere or ocean processes of the last deglaciation, which is one primary reason the model complexity or experimental design differs. Even so, it's a great opportunity to discuss good choices on the scientific question of climate variability."

Authors: *We appreciate this suggestion to highlight a) reasons for the diversity of simulations and models and b) make a suggestion for an experimental design geared towards climate variability.*

Anticipated changes: We will add "Due to substantial differences in forcing [and boundary condition] protocols [inherently arising from different research foci], it can be hard to identify the source of difference between simulations from different models of similar complexity" in l888-889. Furthermore, we will rephrase in l892-893 to "An experimental design geared towards understanding the roles of feedbacks on surface climate variability must take into account external forcing and boundary condition changes, distinguishing interactive effects and prescribed changes in boundary conditions which may, or may not, be physically consistent with the climate evolution. Given the impact of meltwater forcing and its uncertainties, simulations with interactive ice sheets are of particular interest to the study of climate extremes in response to mean changes."

Referee: "L895-L925 (minor): The sentences overlap with the first paragraph of the discussion section. Please consider describing brief conclusions. (or merged with the discussion section?)"

Authors: *We thank the reviewer for this suggestion. We will revise the first sentences in the Conclusions focusing on closing the hypothesis-bracket from the introduction. This should reduce the overlap.*

Anticipated changes: We will rephrase the first two sentences in the conclusions.

Referee: "Figure 1a: EPICA Dome C (Jouzel et al. 2007) and NGRIP (Andersen et al. 2004) presents local temperature change at the ice core site, so it looks strange the vertical axis is represented as GMST. Please clarify the vertical axis."

Authors: *We thank the reviewer for the attention to detail. For easier visualization of both, Antarctic and Greenland temperature change, ice-core records were scaled to GMST assuming a polar amplification of 2. We will include the details of the calibration and scaling in the Figure caption.*

Anticipated changes: We will add the relevant calibration detail in the caption of Fig. 1.

Referee: "Figure 1c: this panel presents sea-level change, but it would fit better including meltwater input as the timeseries of meltwater. While meltwater input would differ between models (e.g. Snoll et al., 2024), it provides an essential information as the meltwater is discussed as the critical factor in climate variabilities. "

Authors: *We thank the reviewer for this suggestion. Meltwater input should, in principle, be related to the time-derivative of the sea-level curve and ice-sheet extent. We will include Meltwater forcing timeseries calculated from ICE-6G and GLAC1-D in the plot to provide the reference for the climate variability discussion.*

Anticipated changes: We will add a panel in Fig. 1 with meltwater forcing.

Referee: "Figure 12: What does "PMIP3" mean? Does it come from Li et al. (2013)? Please clarify in the caption and results section."

Authors: *We thank the referee for the careful reading. In Fig. 12 we denote by PMIP3 the interannual to decadal scale change in climate variability computed in Rehfeld et al., 2018 based on a PMIP3 simulation ensemble. The bar denoted 'reconstructed' refers to the multicentennial to millennial-scale change in variability reconstructed from (mostly) marine palaeoclimate proxy datasets.*

Anticipated changes: We will adapt the last sentence of the figure caption to read: "In panel d, Rehfeld et al. (2018)'s estimated range of the multi-centennial to millennial LGM-to-Holocene variance ratio based on proxy reconstructions (reconstructed) and interannual variability based on the PMIP3 ensemble (PMIP3) are marked for comparison."

References

Ellerhoff, B., Kirschner, M. J., Ziegler, E., Holloway, M. D., Sime, L., and Rehfeld, K.: Contrasting State-Dependent Effects of Natural Forcing on Global and Local Climate Variability, Geophysical Research Letters, 49, e2022GL098 335, `https://doi.org/10.1029/2022GL098335`, 2022.

Rehfeld, K., Hébert, R., Lora, J. M., Lofverstrom, M., and Brierley, C. M.: Variability of Surface Climate in Simulations of Past and Future, Earth System Dynamics, 11, 447–468, `https://doi.org/10.5194/esd-11-447-2020`, 2020.