Response letter

Nils Weitzel, Heather Andres, Jean-Philippe Baudouin, Marie-Luise Kapsch, Uwe Mikolajewicz, Lukas Jonkers, Oliver Bothe, Elisa Ziegler, Thomas Kleinen, André Paul, and Kira Rehfeld

We thank the Editor for assessing our revised manuscript and the constructive feedback to further improve it. As explained in the response below, we have shortened the methods section by moving Sect. 3.1.1 to 3.1.4 to the supplemental information. However, we retained the most important information from these subsections in the main manuscript. The following is a point-by-point response to the comments by the Editor and from the file validation. Here, the comments by the Editor are shown italicized in teal and our responses are provided in black.

I have now gone through your manuscript and the revisions carefully. I think this is a sound piece of work and a very well written manuscript and that you have correctly addressed most requests of both reviewers.

We thank the Editor for this assessment of our work.

My only concern has to do with the length of the manuscript, already mentioned by one of the reviewers. Readers interested in using the method will make the effort to read it carefully but I am afraid others might not make it through. You argued that instead of shortening the text you would rather give the option to skip certain sections but this is not done. I would also recommend you to move the Methods to the supplementary section but I will leave the decision to you. However, if you decide not to do so you please indicate at the beginning of section 3 that sections 3.1.1-3.1.4, and 3.2-3.4 can be skipped for a less detailed focus on technical aspects.

We understand the concerns of the Editor and the Reviewer. We were hoping that the sentence "We provide technical descriptions of the steps in the next four subsections but first motivate them here" was describing the intentions of the subsections sufficiently, but we realize that this is not the case. Since the manuscript has become longer during the review process to incorporate the suggestions of the Reviewers, we follow the advise of the Editor in our revised manuscript version and have moved Sect. 3.1.1 – 3.1.4 to the supplemental information. We only retain the most crucial information from these subsections in the revised Sect. 3.1, in particular the structure of the PSM and the properties of the IQD. In the beginning of the revised Sect. 3.1, we refer to the supplemental information for an enhance description of the algorithm.

6: I believe you should delete the comma after "allows accounting for non-climatic processes,"

We removed the comma.

21: Define ka

We changed the text in brackets to "*LGM, ∼21 ka where ka stands for 'kilo-annum', i.e., thousands of years ago*" in accordance with the submission guidelines.

27: 100 m is a very low-end estimate, 130 m is more realistic

We apologize for the typo. We changed it to 130 m.

33: references should appear in alphabetical order, here and elsewhere

We have adapted the order of references where needed.

40: time slices - please mention some examples of time periods.

We have added "*such as the LGM and the mid-Holocene*".

70: insert "to" before proxies

We added "to" before proxies.

78: again, delete comma after "time series".

We removed the comma.

111: replace "than the other" by "than in the other"

We added "*in*".

132: delete "do" in "do either"

We assume that this comment refers to line 123 and we removed "do" in this sentence.

130: condense to "3K in TRACE-GHG and in FAMOUS"

We changed the beginning of the sentence to "*With ∼1 K in TraCE-ORB and ∼3 K in TRACE-GHG and FAMOUS, ...*".

141: replace parts by periods

We adapted the sentence as suggested.

142-143: same as in line 107

We were unable to find a comment that this comment could refer back to and, therefore, did not change the text in line 142-143 of the previous tracked changes document.

163: insert comma after "clustered"

We added a comma as suggested.

167-169: As I said above, I agree with reviewer 2 that the methods section is too large. You argue that instead of shortening the text you would rather give the option to skip certain sections but this is not specified. If you decide to go for this option it should be mentioned here.

As described above, we have moved Sect. 3.1.1 – 3.1.4 to the supplemental information and only retain the most important information from them in the revised Sect. 3.1.

181: rephrase "We compare measured and forward-modeled proxy time series" to avoid repetition with previous sentence

We rephrased the sentence to "*We perform this comparison in temperature units ...*".

184-185: when referring to the patterns here please include "temporal"

We added "*temporal*" in the lines suggested by the Editor and afterwards in the sentence "*In contrast, temporal pattern components assess ...*".

194: replace "a simulation could simulate" by "a model could simulate"

As we assess the deviation between simulations and reconstructions, we have changed the sentence to "*For example, a simulation could reproduce the reconstructed spatio-temporal temperature pattern accurately but receive a poor score due to an under-estimation of the LGM-to-Holocene temperature change.*"

198-199: Should the description of the Monte Carlo not be included in (1) already (before the decomposition into patterns and magnitudes)?

We agree with this suggestion by the Editor and have moved the information from the two sentences to (1) in the revised manuscript. To incorporate this change, we have adapted the first two sentences in (3) as follows: "*The decompositions in step 2 result in probability distributions of forward-modeled proxy time series and the corresponding reconstructed SST records because we account for chronological uncertainties and include a noise process in the PSM. We quantify the deviations between these probability distributions with a distance function that takes into account the full probability distributions ...*".

207: Here I assume lack of understanding in how climatic effects translate into proxies (uncertainty in PSMs) should also be taken into account.

Indeed, uncertainties in the PSM formulation can also influence the deviations between forward-modeled proxy time series and reconstructions. We have included these uncertainties by rephrasing the two first two sentences in (4) as follows: "*Deviations between forward-modeled proxy time series and reconstructions can depend strongly on the unknown manifestation of non-climatic influences in the measured proxies and uncertainties in the PSM structure. Assuming that most non-climatic processes and PSM uncertainties are uncorrelated between proxy records, the influence of these processes can be reduced by spatially averaging deviations computed for individual proxy records.*"

216: Why use C and not T to refer to the SST field?

We changed C to T.

220-221: How about regions with strong fronts?

Indeed, the interpolation method can have a non-negligible influence in regions with strong fronts. However, the resolution of most simulations of the deglaciation is coarse (typical grid spacings are between 1° and 5°) such that the representation of small-scale structures in the simulations is limited. Therefore, the representation of regions with strong fronts is likely more affected by the resolution of the employed models than by the interpolation method. We already note in the discussion (Sect. 5.2) that physically-motivated downscaling could help to better represent heterogeneous regions such as coastal upwelling zones. In the revised manuscript, we added the following sentences in the description of the spatial interpolation (note that this subsection has been moved to the supplemental information in the revised manuscript as described in the comments above):

"*Only in areas with strong local heterogeneities such as coastal upwelling zones, the interpolation method could potentially impact the resulting time series. However, the coarse resolution of deglacial simulations likely limits the accurate representation of small-scale structures more than the interpolation method (see also Sect. 5.2 in the main manuscript).*"

236-240: This explanation on why additive noise is used could be moved to the discussion to shorten this section.

In the revised manuscript, this explanation has been shifted to the supplemental information, while we keep the discussion of improvements of the PSM structure in the third paragraph of Sect. 5.2.

274: Equ should be Eq hereafter

We changed Equ to Eq in accordance with the submission guidelines.

285: This sentence explaining the IQD name should appear before and not as an isolated paragraph

We apologize for a rendering issue in the difference document. This sentence was the first sentence of the last paragraph in this section and not an isolated paragraph. Nevertheless, we agree with the suggestion by the Editor and have included this information in a revised version of the first paragraph of this subsection. Note that we have moved this subsection to the supplemental information (see response to comments above).

345: When you introduce the reference simulation you could also introduce already here the ground truth it represents.

We have added "*The temperature time series of the reference simulation at each proxy location serves as the ground truth in the PPE.*" as the second sentence in the paragraph to make the connection between reference simulation and ground truth more explicit. We have kept the description of the FPRR and the ground truth deviations in the last two paragraphs of the section because we think the rationale behind these choices becomes more clear after explaining the structure of the PPEs.

348: "We call it pseudo proxies" should be "We call it a pseudo proxy"

We changed "pseudo proxies" to "a pseudo proxy".

352: Can you specify in the text what you mean by simulation characteristics?

We have added "*such as parameter configurations and the implementation of boundary conditions*".

374-375: Replace "the absolute value of the IQD" by "its absolute value"

We have changed the sentence as suggested.

446: FRPP should be FPRR

Thanks for pointing out the typo.

510: Should this reference to Fig 8a not be to Fig 7a?

The reference should be to Fig. 7a and Fig. 8a because Fig. 7a contains the deviations whereas Fig. 8a shows the absolute magnitudes of the simulates. Accordingly, we have added a reference to Fig. 7a in the revised manuscript.

A few final more general comments follow. To the extent you think these three issues are important, I would recommend you to include a comment on these, possibly in the abstract already, in order to increase the paper's interest for the broader community.
First, the result you point out in lines 539-540 is quite shocking but also revealing. If I understand correctly it implies that over-tuning a model in a given region (for instance via freshwater fluxes) penalises the simulation in others, pointing to the fact that the driving mechanism in the simulation might not be as simple as freshwater fluxes. Lines 696-703 below also relate to this, and I very much agree with this conclusion.

Indeed, the results hint at limitations from over-tuning a model to proxies from a specific region. As stated by the Editor, we discuss this in Sect. 5.2. To emphasize this result more, we added statements in the abstract ("*The ranking of the simulations differs substantially between the considered regions and timescales, which suggests that optimizing for agreement with the temporal patterns of a small set of proxies might be insufficient for capturing the spatial structure of the deglacial temperature variability.*") and the conclusions ("*This suggests that optimizing for agreement with the temporal patterns of a specific proxy or reconstructions from a small region might be an inadequate strategy for capturing the spatial structure of millennial-scale temperature patterns during the deglaciation.*").

Second, I might have skipped something but how does internal climate variability affect the method? What happens if millennial-scale variability is entirely internally driven?

In the case of entirely internally driven millennial-scale variability during the deglaciation, the *millennial pattern* metric does no longer provide meaningful information because it assumes temporal alignment of reconstructions and simulations. This cannot be expected for internally-driven fluctuations. On the other hand, the *millennial magnitude* metric still provides useful information for evaluating the strength of millennial-scale fluctuations. We believe that Sect. 5.2 is the best place to discuss this question and that it is not important enough to highlight it in the abstract or conclusions. While the second-to-last paragraph in Sect. 5.2 already included a discussion of this issue, we rephrased it to emphasize the question of internally driven variability stronger:

"*As the PPEs and the real-world application have shown, the pattern IQDs are sensitive to the timing of timescale-dependent temperature fluctuations. Therefore, they are only meaningful if the goal of a simulation is to reproduce a specific succession of variations observed in reconstructions. Temporal alignment cannot be expected for internally driven variations such as spontaneous millennial-scale fluctuations (...), and in the presence of boundary conditions with large spatio-temporal uncertainties like deglacial meltwater fluxes. In these cases, the magnitude IQDs, which are insensitive to the timing of fluctuations, could be combined with a more insightful measure for temporal patterns, e.g., based on the similarity of spatial relationships in reconstructed and forward-modeled proxy time series (...).*"

Finally, how specific is the method to the last deglaciation? To what extent would it be applicable for example to other periods (give examples)?

The last paragraph of Sect. 5.2 states that the method can be applied to other continuous variables (e.g., oxygen isotopes) and other periods (e.g., penultimate deglaciation, the glacial inception, or the last glacial cycle). To emphasize this point, we have included additional sentences in the abstract ("*Additionally, the algorithm can be applied to variables like oxygen isotopes and climate transitions such as the penultimate deglaciation and the last glacial inception.*") and the conclusions ("*In addition to assessing the temperature evolution during the last deglaciation, the proposed method can be applied to other continuous variables, e.g., oxygen isotopes, and other periods with climate transitions such as the penultimate deglaciation and the last glacial inception.*") to state these possible applications of our proposed method.

Notification to the authors (from file validation):
With the next file upload request, please update the "Competing interests" as follows: At least one of the (co-)authors is a member of the editorial board of Climate of the Past

We apologize for not including this statement before and have added it in the revised manuscript.