Dear Referee #2,

Thank you for your time and effort to provide constructive feedback on our manuscript. We have replied to each of your comments and concerns (in italic text) below. Specifically, we will include a separate discussion section in a revised version of our manuscript. In this section, the missing references that the reviewer pointed out will be included. We expect this change will address the reviewers' concerns about the clarity and structure of the text, as well as improve the discussion of potential implications of our results and the comparison with existing literature.

Yours sincerely,
Anne Morée and co-authors

General comments
The authors use an ocean-sea-ice model (NorESM-OC) that also includes biogeochemistry, δ13C carbon isotopes and radiocarbon, to quantify the role of the efficiency of the LGM biological pump in obtain the best agreement between model simulations and proxy data. Their results indicate that the efficiency should be doubled to obtain the smallest model-proxy mismatch.
The model setup is novel, properly thought through, overall well-described, and can certainly be used to provide useful insight on long-standing questions about the role of ocean circulation and biogeochemistry in driving glacial-interglacial changes in ocean carbon storage. However, the structure and clarity of parts of the manuscript need to be substantially improved. Some additional simulations/sensitivity experiments may also need to be included, or at least their potential implications need to be better discussed and compared with the existing literature. A few highly relevant studies, and all very recent, are also missing in the references.
Setting up these simulations must have involved a substantial amount of work and this should be acknowledged and this framework will also be useful to investigate other research questions. This study makes valuable contributions to the topic and definitely deserves to be published, but several issues need to be addressed first, as described in the comments below.

Specific comments
Abstract, page 2, line 33: This statement is a bit too strong. The LGM is indeed a good test case for models and their evaluation and process-based understanding, but it can't be considered a necessary "requirement" for their reliability for future projections. I get the point and I agree, but this need to be rephrased.
Author response: We think the reviewer refers to p. 1 l. 33 here. We will change the statement as well as shorten this part of the abstract (see also our response to reviewer (#1, response to comment on Line 26-35). Our results underline that only those coupled climate models that contain the processes and/or components that realistically change both ocean circulation and biogeochemistry will be able to simulate an LGM ocean in satisfactory agreement with proxy data. Such a simulation is also a test for Earth system models for their ability to reproduce natural climate variations adequately as a basis for reliable future projections, including human-induced forcing. I.e., a satisfactory fidelity of Earth System Models in reproducing orbitally forced climate variations will increase our confidence in these models as tools for projecting future anthropogenic climate change.
Changes in the manuscript: Shorten abstract lines 26-35 and move part of text to discussion/conclusion sections.

Page 2, lines 10 and 23: Add references to Stein et al. (2020) and Marzocchi and Jansen (2019), especially since these studies both address directly the role of physical changes on glacial carbon storage, which is not really done in this manuscript. These also needs to be discussed further with the results – see later comments.
Author response: Agreed.

Changes in the manuscript: The references Stein et al. (2020) and Marzocchi and Jansen (2019) will be added and discussed.

Methods

The simulations are integrated for a long period of time. Nonetheless, it would still be useful to show some of the LGM ocean state equilibrium/drift in the Supplement. Perhaps some timeseries of T and S and/or AMOC and Drake Passage transport, which are already mentioned in the text.

The Bern3D model part of the study needs to be introduced and explained, at least briefly, in this section – with proper reference to the Supplement for the rest of the details.

Author response:

We will include a time series over the last 1000 years of the LGM and PI simulations in the supplement (for S, T, AMOC and Drake Passage transport) to give a visual impression of the equilibration/drift. Regarding the Bern3D model, we see that besides the information on the Bern3D model in Sect. 3.3 and SM3, the model and its application in the context of our study should be introduced in the methods section as well for which we will include a new section.

Changes in the manuscript: Addition of a new equilibration time series figure in the supplement, as well as a new methods Section (2.5) to describe the setup and use of the Bern3D model in this study.

Results and discussion

This part of the manuscript needs some substantial restructuring and improvements. Parts of it are quite confusing, which takes away from the key findings and the main points that the authors are trying to get across.

Perhaps separate more clearly parts of the results that are more of a “model evaluation” and then for each of these have a subsection that discuss the reasons for the biases, to give some separation between results and discussion, especially where comparisons to observations and other studies are also discussed.

All of this is already in the text, but currently quite mixed up all together, making several parts a little hard to follow. I am not against having results and discussion together, but the structure needs to be clearer and easier to follow.

Author response: In order to improve the clarity and structure of section 3, we can agree that the inclusion of a separate discussion section would help. We would be able to lift some of the model-data comparison discussion points that are currently spread throughout the Sect. 3 text into such a new section, as well as provide a dedicated section for the discussion of the remaining model-data mismatch after adjustment of the efficiency of the biological pump.

Changes in the manuscript: Inclusion of a separate discussion section at the end of the manuscript that focuses on the discussion around model-proxy data mismatches. This section can consist of two parts, one discussing the model-data mismatch of the original simulation, and a second dedicated to the remaining model-data mismatch after adjustment of the efficiency of the biological pump (i.e., p. 12 l. 12-29).

Section 3.1 is a little hard to follow without any figures. . .maybe add some in the Supplement?

Author response: We see that no reference is made in Sect. 3.1 to Fig. S5, which shows the PI physical state and could already be referred to here (currently done in Sect 3.2). The biogeochemical state of the PI simulation is described in detail for the C isotopes in Tjiputra et al. (2020). Otherwise, our focus is on the change in the biogeochemical marine state (LGM-PI), which is shown in Fig. 2. To address the reviewers comment further, we can provide supplementary figures of PI temperature (section), PP (vertically integrated), and regenerated phosphate (section) compared to observational estimates, since these are mainly discussed in Sect. 3.1.

Changes in the manuscript: Include reference to Fig. S5 in Sect 3.1 as well as new supplementary figures of temperature, PP and regenerated PO4 compared to observations.

Page 8
Discuss the radiocarbon ages also with respect to the results of Burke et al. (2015)
Author response: We will include the Burke et al. (2015) reference in our discussion.
Changes in the manuscript: Add and discuss Burke et al. (2015).

Line 31: add references to Jansen (2017) and Marzocchi and Jansen (2019) to support this statement on the importance of atmospheric temperatures for both LGM water masses and biogeochemistry, respectively.
Author response: We agree these references should be cited here and will do so in a revised version of the manuscript.
Changes in the manuscript: Add Jansen (2017) and Marzocchi and Jansen (2019) references.

Line 35: this needs to be discussed a little further (i.e. the underestimation of negative buoyancy fluxes) – for instance, compare Klockmann et al. (2018) – this is an example of where I think a separate Discussion section is missing. Alternatively, this could be picked up again in the conclusions as one of the potentially important biases. The abyssal cell actually looks weaker at the LGM? (Figure S5) This also needs to be discussed, perhaps here.
Author response: As described in our response to the general comment on the result and discussions section, we propose to include a separate discussion section in a revised version of the manuscript. In the first part of this new section, where we want to discuss the model-proxy data mismatch of the original simulation we would be able to discuss our results in more detail regarding buoyancy fluxes and the abyssal cell strength (which indeed weakens). We assume the reviewer means Klockmann et al. (2016) here (as in their reference list), and will include the findings of Klockmann et al. (2016) in our discussion.
Changes in the manuscript: Discuss our simulation with regard to (Southern Ocean) buoyancy fluxes and the strength of the abyssal cell in the new discussion section. Specifically, include and discuss Klockmann et al. (2016).

Page 9
Line 5: add reference to Marzocchi and Jansen (2019) and Stein et al. (2020) where the link to ocean carbon storage is actually tested.
Author response: See our response to p. 2, l 10 and 23 above.
Changes in the manuscript: We will include these references here and in other appropriate locations where their findings are useful for our discussion.

Section 3.2.2
Lines 10-26: This result (i.e. reduced LGM biological pump efficiency but lower pCO2 concentrations) is not dissimilar from what discussed in Marzocchi and Jansen (2019), despite a very different model setup. So this is worth discussing further – perhaps think about this in the context of the carbon pump decomposition. This may mean that there is something we simply don't understand in this part of the mechanism. Can your study clarify this apparent discrepancy further? Can you make this clearer/highlight it better?
Author response: The lowered atmospheric pCO2 is expected from the combined effect of the ocean volume decrease and increased CO2 solubility due to decreased ocean temperatures (p. 9 l. 19-22). That is, mostly the physical C pump is represented in our study and driving down atmospheric pCO2 (as evidenced by increases in DICpref and DICsat, not shown (for definitions see also Sect 3.1 in Tjiputra et al., (2020) and references therein)). The lack (and actually decreased efficiency) of a soft tissue pump strengthening is discussed in Sect. 3.3. The inability of our model to simulate the strengthening of the soft tissue pump is expected from earlier results for ESMs and our model setup (f.e. summary point 5 in Galbraith and Skinner, 2020; p. 9 l. 22-26) - and indeed indicates that some biogeochemical processes/mechanisms are lacking in these models. Pinning down the exact processes of this strengthening is an ongoing challenge, and beyond the scope of our study. Nevertheless, we would be able to decompose the LGM-PI change in DIC into DICsoft, DICpref, DICsat, DICbio,

DICcarb and DICdiss (definitions in Sect 3.1 in Tjiputra et al., (2020) and references therein) and add a figure of this to the supplement in order to visualize their individual contributions. We will highlight this result and its discussion in the new discussion section.
Changes in the manuscript: Clarify the atmospheric pCO2 drawdown in the context of the LGM-PI changes in the different C pump components (DICsoft, DICpref, DICsat, DICbio, DICcarb and DICdiss) in a new supplement figure. Discuss this and specifically the lack of a contribution from the soft tissue pump on simulated LGM atmospheric pCO2 in more detail in the new discussion section.

Page 10
Lines 10-19: this is another example where this is a discussion part, but it's somewhat "thrown" in the middle of some other text. So again this needs restructuring to make it easier for the reader to follow.
Author response: As described in our response to the general comment on the result and discussions section, we propose to include a separate discussion section in a revised version of the manuscript. The discussion on p. 10 l. 10-19 could be moved to such a new section to improve the structure of Sect. 3.
Changes in the manuscript: Include p. 10 l. 10-19 in a new discussion section.

Line 25: here the reference is Marzocchi and Jansen (2019) rather than Jansen (2017).
Author response: Thank you for noting this, we see that Marzocchi and Jansen (2019) is more appropriate here than Jansen (2017) and will adjust the manuscript accordingly
Changes in the manuscript: As suggested

Page 11
Lines 2-21: This part about the Bern3D ESM comes a bit out of the blue and I can't say that this is explained well enough and entirely clear. Make better reference to the Supplement and better introduce the setup in the Methods (as noted before), where the goals of this additional step need to be better clarified and introduced. Then it will come less out of the blue here in the results.
Author response: As described above, we will add a new methods section 2.5 on the Bern3D model. Here we will pay specific attention to clarifying why and how the Bern3D model was used.
Changes in the manuscript: Addition of a new methods section 2.5 on the Bern3D model.

Page 12
Lines 12-29: this is again a somewhat self-standing discussion part that should perhaps be a subsection.
Author response: As described in our response to the general comment on the result and discussions section, we propose to include a separate discussion section in a revised version of the manuscript. The second part of this new section will be dedicated to discussing the remaining model-proxy data error after adjustment of the efficiency of the biological pump, which is essentially p. 12 l. 12-29.
Changes in the manuscript: Restructure the text to include a new discussion section, where p. 12 l. 12-29 would be a second section that discusses the remaining model-proxy data error after adjustment of the efficiency of the biological pump.

Here, and/or earlier, you should discuss the results of Odalen et al. (2019). Actually, would it be feasible to test their variable C/P ratio in your simulations?
Author response: We assume that the reviewer refers to the paper Ödalen et al. (2020). Their results, which could decrease in d13C while keeping (regenerated) PO4 constant, could indeed be included in our discussion on the remaining proxy-data mismatch. With regard to the feasibility to test variable C:P ratios in our model setup: Our model is computationally more

demanding than the cGENIE model as employed by Ödalen et al (2020) and we do not have the resources for repeating long runs for this currently.
Changes in the manuscript: Discuss the results on variations in the C:P ratio by Ödalen et al. (2020) and their potential implications for our remaining model-data mismatch.

Also could you quantify the dependence of your results to your model initial state, as discussed in Odalen et al. (2018)? [this reference is already cited in the manuscript].
Author response:
We did not carry out experiments with vastly different initial states. It is known since long, that different initial conditions for temperature and salinity can result in different circulation modes. However, in our case we assume that the initial conditions for the glacial ocean circulation would not be too different from preindustrial conditions and not fully different. Due to the high computational demand of our model, we cannot carry out multiple spin-ups (as 10,000 years done in the cGENIE model) with different initial conditions or tunings. This would not be feasible given currently available computational resources.
Changes in the manuscript: No changes will be made.

Conclusions
Add a reference to Rae et al. (2019) when discussing the importance of southern-sourced waters. This should probably also be discussed earlier in the results/discussion.
Author response: We assume that the reviewer refers to the paper Rae et al. (2018). As our paper does not deal with pH changes, we do not specifically discuss this article, but can include this reference as it highlights the central role of SSW as the reviewer points out.
Changes in the manuscript: Rae et al. (2018) will be cited at page 2, lines 5 and 23, and added to the reference list.

Technical corrections
Abstract Line 17: ocean model state? Do you mean "equilibrium simulations"? Clarify. Ocean model state is not the best term to use here.
Author response: We will adjust the text as proposed below
Changes in the manuscript: replace sentence 'We prepared a PI and LGM ocean model state (NorESM-OC) with full biogeochemistry (including the carbon isotopes δ13C and radiocarbon) and dynamic sea ice.' with 'We prepared a PI and LGM equilibrium simulation using model NorESM-OC with full biogeochemistry (including the carbon isotopes δ13C and radiocarbon) and dynamic sea ice.'

Line 23: "we explore the theoretical effects" doesn't quite make sense. This could just be "we explore/test the effects".
Author response: We think clarifying that our approach is exploring the potential effects only (i.e. it is an approximation as no actual simulation is done) is important here, and we therefore propose to replace 'theoretical' with 'potential (offline)' in the abstract.
Changes in the manuscript: Replace 'theoretical' with 'potential (offline)' on p.1 l.23.

Line 29: again "theoretical" is not quite the right word. Just say "our approach". Same in the rest of the manuscript (e.g. page 7, 10, 13). Perhaps do just call it "offline".
Author response: See our response to the previous comment.
Changes in the manuscript: Replace 'theoretical' with 'potential (offline)' or 'our approach' throughout the text to clarify our intention to explore the potential (offline) effects whenever we describe our approach.

Page 10, line 30: miss-match should be mismatch.
Author response: Thank you for noting this mistake, we will adjust the manuscript as suggested.
Changes in the manuscript: change p. 10, l. 30 miss-match to mismatch.

Everywhere: “Southern Source” should really be “southern-sourced”.
Author response: We revisited the literature and see that both southern source water (e.g., Adkins, 2013; Curry and Oppo, 2005; Roberts et al., 2010) and southern-sourced water (Howe et al., 2016; Pöppelmeier et al., 2018) are commonly used. We therefore feel the current use of Southern Source Water (SSW) throughout the manuscript can be maintained.
Changes in the manuscript: None.

References
Burke, A., Stewart, A.L., Adkins, J.F., Ferrari, R., Jansen, M.F. and Thompson, A.F., 2015. The glacial middepth radiocarbon bulge and its implications for the overturning circulation. Paleoceanography, 30(7), pp.1021-1039.
Klockmann, M., Mikolajewicz, U. and Marotzke, J., 2016. The effect of greenhouse gas concentrations and ice sheets on the glacial AMOC in a coupled climate model. Climate of the Past, 12, pp.1829-1846.
Marzocchi, A. and Jansen, M.F., 2019. Global cooling linked to increased glacial carbon storage via changes in Antarctic sea ice. Nature Geoscience, 12(12), pp.1001.
Ödalen, M., Nycander, J., Ridgwell, A., Oliver, K.I., Peterson, C.D. and Nilsson, J., 2019. Variable C/P composition of organic production and its effect on ocean carbon storage in glacial model simulations. Biogeosciences Discussions, pp.1-33. (accepted) DOI: https://doi.org/10.5194/bg-2019-149
Stein, K., Timmermann, A., Kwon, E.Y. and Friedrich, T., 2020. Timing and magnitude of Southern Ocean sea ice/carbon cycle feedbacks. Proceedings of the National Academy of Sciences, 117(9), pp.4498-4504

References of the response
Adkins, J. F.: The role of deep ocean circulation in setting glacial climates, Paleoceanography, 28, 539-561, 10.1002/palo.20046, 2013.
Burke, A., Stewart, A. L., Adkins, J. F., Ferrari, R., Jansen, M. F., and Thompson, A. F.: The glacial mid-depth radiocarbon bulge and its implications for the overturning circulation, Paleoceanography, 30, 1021-1039, 10.1002/2015PA002778, 2015.
Curry, W. B., and Oppo, D. W.: Glacial water mass geometry and the distribution of δ13C of ΣCO2 in the western Atlantic Ocean, Paleoceanography, 20, 10.1029/2004PA001021, 2005.
Galbraith, E. D., and Skinner, L. C.: The Biological Pump During the Last Glacial Maximum, Annual Review of Marine Science, 12, 559-586, 10.1146/annurev-marine-010419-010906, 2020.
Klockmann, M., Mikolajewicz, U., and Marotzke, J.: The effect of greenhouse gas concentrations and ice sheets on the glacial AMOC in a coupled climate model, Clim. Past, 12, 1829-1846, 10.5194/cp-12-1829-2016, 2016.
Roberts, N. L., Piotrowski, A. M., McManus, J. F., and Keigwin, L. D.: Synchronous Deglacial Overturning and Water Mass Source Changes, Science, 327, 75, 10.1126/science.1178068, 2010.
Howe, J. N. W., Piotrowski, A. M., Noble, T. L., Mulitza, S., Chiessi, C. M., and Bayon, G.: North Atlantic Deep Water Production during the Last Glacial Maximum, Nature Communications, 7, 11765, 10.1038/ncomms11765, 2016.
Jansen, M. F.: Glacial ocean circulation and stratification explained by reduced atmospheric temperature, Proceedings of the National Academy of Sciences, 114, 45-50, 10.1073/pnas.1610438113, 2017.
Marzocchi, A., and Jansen, M. F.: Global cooling linked to increased glacial carbon storage via changes in Antarctic sea ice, Nature Geoscience, 12, 1001-1005, 10.1038/s41561-019-0466-8, 2019.
Pöppelmeier, F., Gutjahr, M., Blaser, P., Keigwin, L. D., and Lippold, J.: Origin of Abyssal NW Atlantic Water Masses Since the Last Glacial Maximum, Paleoceanography and Paleoclimatology, 33, 530-543, 10.1029/2017PA003290, 2018.

Rae, J. W. B., Burke, A., Robinson, L. F., Adkins, J. F., Chen, T., Cole, C., Greenop, R., Li, T., Littley, E. F. M., Nita, D. C., Stewart, J. A., and Taylor, B. J.: CO2 storage and release in the deep Southern Ocean on millennial to centennial timescales, Nature, 562, 569-573, 10.1038/s41586-018-0614-0, 2018.
Stein, K., Timmermann, A., Kwon, E. Y., and Friedrich, T.: Timing and magnitude of Southern Ocean sea ice/carbon cycle feedbacks, Proceedings of the National Academy of Sciences, 117, 4498, 10.1073/pnas.1908670117, 2020.
Ödalen, M., Nycander, J., Ridgwell, A., Oliver, K. I. C., Peterson, C. D., and Nilsson, J.: Variable C/P composition of organic production and its effect on ocean carbon storage in glacial-like model simulations, Biogeosciences, 17, 2219–2244, https://doi.org/10.5194/bg-17-2219-2020, 2020.