General response

We thank all three references for their constructive comments and suggestions. Overall, there seem to be two main categories of issues/suggestions:

1) Writing: (a) the aims and conclusions of the research are not communicated clearly enough and (b) the comparison against observations/proxies is too detailed for an EMIC. We agree and will address as detailed below. We also agree that the strengths of the study have not been sufficiently emphasised.

2) Analyses: It is suggested that doing additional analyses would be helpful if feasible (specifically incorporating ^{13}C and performing sensitivity analyses). These analyses would be interesting, but are unfortunately infeasible as the research was performed during a PhD and there are no resources for additional simulations. Note that applying all our ensemble members to the simulation of one stage represents at least 0.5 CPU years of computing. We will instead deal with these helpful suggestions through an existing sensitivity not previously discussed, and re-emphasise caveats as appropriate, and as detailed below.

We note that this paper was intended as part one of two related papers. This paper describing the relationships between ensemble outputs, and the second, currently being finalised for submission, describing dependencies on ensemble parameters to isolate mechanisms. We are happy to provide this as part of the review process as it may help to put this current paper in perspective. Even accounting for the increased focus of the planned revisions, we still believe there to be far too much material overall to cram into a single paper.

Response to referee #1

We thank referee #1 for useful and constructive comments. Our response is provided below in black, with original referee comments in blue.

General comments

I like it very much that the authors provided an comprehensive review of the mechanisms that governing the LGM atm. CO2 drop. And I really appreciate that an extensive body of related work are mentioned during the discussion of model results. However, I do have several major concerns.

1) Research aim:

P1 l15-16 and P3 l19 suggest that the aim of this study is "to investigate the causes of the LGM atmospheric CO2 drop". To me, this research aim is not appropriate. Previous studies have already proposed some hypotheses regarding the mechanisms that governing the decrease of atmospheric CO2 concentration relative to preindustrial (as have been summarised in the Introduction). What is not yet clear is the interplay of these mechanisms and the relative quantitative contribution of each mechanism to the LGM atmospheric CO2 drop. In this study, the authors did not explicitly propose any new hypothesis. Yet, the quantification of the contribution of different mechanisms using GENIE is not possible due to the simplification of the model and due to that many processes are not accounted for.

We agree that the research aim should be reframed. We are using a large ensemble of LGM forced simulations to explore changes in physical and biogeochemical variables thought to influence the drawdown of CO_2. We will make it clearer that the objective is not to explain definitively the causes of CO_2 drop as we strongly believe the inevitable uncertainty surrounding simulations makes it impossible to accurately resolve the relative contributions of individual processes in models of any complexity as a result of irreducible uncertainty regarding processes. Instead our aim is to account for as many sources of model uncertainty as we can and from there describe the model output space we see. Intermediate complexity is demanded for such an uncertainty-based study, which carried out 471x40,000-year simulations, intractable in high complexity models.

I think a more specific research aim/question is needed for this study. When setting up the aim/question, the authors might consider: What are the novel aspects of the model or the EFPC2 ensemble? In this manuscript, is it the first time that an interactive carbon cycle model is applied to LGM? Is it for the first time the sensitivity to process parameters is investigated for LGM climate (Holden et al 2013 pointed out that EMICs are important tools for exploring sensitivities and quantifying uncertainty)? Which mechanism(s) the authors would like to focus on? terrestrial carbon preservation? carbonate weathering?

Agreed, the revision will address this and the introduction will lay out the research objective and novelties of our approach, being:

- Investigating the range of physical and biogeochemical changes (and hence implicitly also specific mechanisms) which may have accompanied the LGM atmospheric CO_2 drop, when taking into account model uncertainty in a large ensemble approach.
- Attempting to simulate the LGM atmospheric CO_2 drop with the simulated CO_2 feeding back to the simulated climate, which is still infrequently done in LGM CO_2 experiments, and the first time it is done with GENIE-1.
- Simulating the burial of land carbon by glacial ice sheets: there have been only a limited number of studies doing this and none that have attempted it in an LGM equilibrium experiment set-up.

2) Key findings/conclusions:
The current version of the Conclusion reads more like a summary of the model results. It’s not clear what are the key findings of this study. I think the key message would become clear once the research aim/question is given.

Agreed, the revision will address this and summarise the key conclusions, being:

- Despite our extensive exploration of model uncertainty, we struggle to achieve -90 ppmv. We attribute this to the potential LGM CO_2 drivers not included in the model as well as error in the model's process representations (and which was not captured by the ensemble). The total effect on ΔCO_2 is estimated at up to ~60 ppmv, yielding an acceptable, or "plausible", ΔCO_2 range of ~-30 to ~90 ppmv.
- Within this plausible ΔCO_2 range, there are multiple potential ΔCO_2 "solutions", simply in terms of the sign and magnitude of physical and biogeochemical changes. However, plausible CO_2 is more frequently associated with some changes than others, namely: decreasing SSTs, increasing sea ice area, a weakening of the AMOC, a strengthening of the AABW cell in the Atlantic Ocean, a decreasing ocean biological productivity, an increasing $CaCO_3$ weathering flux, an increasing terrestrial biosphere carbon inventory and an increasing deep-sea $CaCO_3$ burial flux.
- The paper focuses on these dominant changes: (1) showing that the change in terrestrial carbon is positive both because of ice sheet carbon burial and reduced soil respiration, and (2) suggesting ways in which the other physical and biogeochemical changes may have occurred, based on their spatial patterns and relationship to other changes.
- The dominant changes are also found to be broadly consistent with observations, and based on a qualitative attempt to reconcile the sign of the terrestrial carbon change with carbon isotopes, we show that a positive terrestrial carbon change is not immediately in contradiction. However, a quantitative assessment would be needed to obtain a more definitive answer. Since a detailed and comprehensive assessment of modelled results against observations was not an objective of the current paper, this is a direction for future research.

3) Comparison of model results to field/proxy data or to results of other models:
A large body of text is devoted to the comparison between model results and data and to the explanation of differences between the two. To me this is a bit overdone. Holden et al. (2010), who used the same model and had the same principle for the design of the ensemble, clearly stated that the ensemble is developed to reproduce the main features, but not the precise observation.

We agree and will compare modelled results in less detail in the revision, focusing on spatial patterns at a coarser resolution, and making it more explicit that globally aggregated estimates are to be treated in a similar way.

Specific comments (the following comments mainly concern the suggestions for improving the structure and presentation of model results.)

Abstract

P1 l15-16: It would be helpful to specifically state the research aim here.

We agree that our research aim needs to be clearly laid out at the beginning of the abstract. We propose to briefly introduce the research problem, followed by our aim, and then our key results.

Introduction

P3 l27: "unknown error" – does it mean the discrepancy between e.g. the modelled circulation and that obtained from proxy data? To me it sounds like a mistake one accidentally made in the model.

Our meaning here is that there is no direct two-way, persistent, interaction between ocean and ice sheets (as is the case in some higher complexity models), and the error caused by the absence of this process is unknown.

P3 l28: "unknown error" – Maybe "missing processes" is more suitable here? Agreed, this would be clearer and will be changed.

Please add at the end of the Introduction an overview of the upcoming sections, viz., what will be presented in each section. Yes, thank you for the useful suggestion. We will include a paragraph detailing our main sections: (i) description of the model, ensemble and simulation set-up used; (ii) brief analysis of preindustrial simulations to verify that plausible given that the way we spin-up the model is not exactly the same as in Holden et al., 2013a, and we also wanted to include evaluation of a few metrics not used to constrain the original ensemble; (iii) analysis of LGM simulation results to determine range of physical and biogeochemical changes that may have accompanied LGM CO_2 drop. This section includes direct drivers of CO_2 change (e.g. terrestrial carbon inventory change) as well as indirect drivers (e.g. precipitation change).

Method

P5, Table 1: what is OLR? Outgoing longwave radiation

P6: I think a table or a flow chart summarising the conditions for the four stages would be helpful for readers to understand the set-up of experiments. Agreed, we can include this in the revised version.

How long is the stage 2? 10,000 years

Is the total carbon inventory (that is, sum of atmospheric, terrestrial, ocean and lithospheric carbon inventory) unchanged over the four simulation stage? We checked that carbon is being approximately conserved over stage 3 in PGCAF-16, by calculating the sum of ΔATMC, ΔTERRC, ΔOCEANC and ΔLITHC for each of the 16 ensemble members. We found that each sum, in absolute terms, was less than 10 PgC.

LGM ensemble simulations

P9: Please explain here why the subset PGACF-16 is needed. This can be done by moving p37 l30-32 to section 4.1. We agree that the need for PGACF-16 should be explained more clearly when introducing it. The main reason is to determine whether the patterns identified for PGACF as a whole also appear to hold for the lower end of the ΔCO_2 range more strictly. Assessing variation across the plausible ΔCO_2 spectrum is not a dominant objective of the study, however, and we will update the writing to reflect this.

It would be very helpful if a brief overview of the following text of this section 4 is presented: which variables of which set will be presented. Agreed. In the current manuscript, for all model outputs of interest, we discuss PGACF, PGACF-16 and EFPC2, albeit focusing on the first ensemble. The other two ensembles are discussed to show extent to which similar to PGACF, despite their different ΔCO_2 ranges and number of ensemble members included. In the revised manuscript, we will reduce discussion of PGACF-16 and EFPC2 and have this reflected in the overview section.

P25 l6-7: this is not true because in Table 6 none previous observation/model data study shows negative delta_OceanC. We agree that this statement is misleading as two of the studies make use of soil carbon measurements, and the third is a modelling study which does not explicitly report the change in ocean carbon. The statement was based on the assumption that if ~90% of the atmospheric CO_2 perturbation caused by the reported increase in terrestrial biosphere carbon was removed by oceans and sediments, the change in ocean carbon inventory would still be negative (between ~-9 and ~350 PgC), after adding remaining carbon to be lost from the atmosphere to the ocean. We will rephrase our statement in the revision.

P36 bottom: Fig. 18 should be Fig. 17. We will update the figure number.

P36 l3: "North Atlantic" and "North Pacific" should be switched. We agree that the text is confusing here and we will clarify this by replacing “large increases” with “$\Delta CaCO3_{bur}$ maxima”

Comparison of global-integrated numbers and spatial distribution between model results and data: I suggest to first compare the spatial distribution and then the globalintegrated numbers because the latter is just the sum of the former. Agreed, we will switch the order in the revision.

Colour slots showing spatial distributions of variables, e.g. Fig 3, 6, 7, ...: please add contour line for land-ocean border. These will be included in the revision.

Plots showing standard deviation in e.g. Fig 3, 6, ...: These plots are shown but never mentioned/used/discussed. So please consider move them to a supplementary information file – there are already many figures in the manuscript. Agreed, we included the standard deviation plots for information but they are not central to our arguments. We will move them to a SI file in the revised manuscript.

Conclusions

P39 l15: The positive delta_TerrC has been discussed and justified many times through out the manuscript. Thus, I have the impression that this is one key finding the authors would like to stress. I think this is a bit dangerous because this point is not well supported by data. The authors also seem de-stressing this point several times by stating there are other 4 ways of "achieving a plausible delta_CO2 interns of the sign of individual carbon reservoir changes (although Table 5 suggests those 4 ways are much likely to occur). I have to say I am confused by the above statements.

We agree that the statements are confusing and we will clarify these in the revision as follows: we find that plausible ΔCO_2 can be achieved in 5 different ways in terms of the sign of individual carbon reservoir changes. However, positive $\Delta \mathrm{TerrC}$ combined with positive $\Delta \mathrm{LithC}$ and negative $\Delta \mathrm{OceanC}$ is by far the most common way, encompassing 89% of simulations, and we focus our discussion on these simulations. This includes proposing explanations for the positive $\Delta \mathrm{TerrC}$ and discussing the change in the light of observational constraints.

P39 l39 - P41 l15: I understand that it is a pity that carbon isotopes were not simulated. However, I don't think it is appropriate to extensively present inferred results in the Conclusion section.
Agreed. In the revision, we will move the discussion of carbon isotopes, focused on the positive $\Delta \mathrm{TerrC}$, to the section on carbon reservoir changes.

Response to Pearse Buchanan (referee #2)

We thank Pearse Buchanan for his very detailed and constructive comments. As above, our replies are in black, and original comments in blue.

1 General comments

This study provides some unique perspectives on the glacial drawdown of atmospheric pCO2. This problem has been at the forefront of climate research for many decades. So far, many mechanisms have been proposed, but a recipe of changes that is physical and biogeochemically consistent with proxy records and known mechanisms is still elusive. The authors of this study set out to try and achieve this difficult task.
The authors employ a unique set of methods to attack their exploration of what might plausibly reduce atmospheric pCO2 under glacial conditions. The method involves simulating 315 individual parameter sets using the same Earth System Model (ESM) of intermediate complexity in four stages. From what I can tell, each stage involves the full 315-member ensemble, unless numerical instability problems were encountered. Each ensemble member was therefore independent from another at all stages through the study. Each stage was initialised from the final year solution of its previous stage, except stage 1 which I am unsure about what fields have been used for intialisation. The only boundary conditions that were prescribed to the Earth System Model that I can tell were orbital parameters, aeolian iron deposition rates, detrital flux rate to ocean sediments, and ice-sheet fraction and their orography. The 4 stages are

as follows:
1. Stage 1 (PI 10,000 years): relaxed pCO2 to 278 ppmv; no interaction between carbon reservoirs; conserved alkalinity in the ocean.
2. Stage 2 (PI 10,000 years): freely evolving pCO2; interacting reservoirs; freely evolving ocean alkalinity.
3. Stage 3 (LGM 10,000 years) freely evolving pCO2; interacting reservoirs; freely evolving ocean alkalinity; ice-sheet growth and corresponding sea level loss years 0-1000.
4. Stage 4 (LGM 10,000 years) freely evolving pCO2; interacting reservoirs; freely evolving ocean alkalinity.

We will revise the text to clarify our method, which is the following:
In total, 471 runs were applied to preindustrial and then LGM simulations. At the LGM (stage 3), however, 75 runs either "snowballed" or crashed, leaving 396 ensemble members (EM). By "snowballed" we mean that the runs predicted highly implausible global annual SATs, between ca. -67.8 and -56.8°C, likely as a result of global or near-global sea and land ice cover developing in the simulations. Out of the remaining 396 EM, we then removed those simulations with preindustrial CO2 outside of 280 ± 10 ppmv (18 EM), and subsequently any EM that predicted large abrupt changes in atmospheric CO2 over the LGM simulation that likely caused by instabilities rather than by some physical mechanism (63 EM).

Regarding inititialisation, the model spins up from its default state. Only the CaCO3 weathering fluxes are taken from Holden et al., 2013a, diagnosed from their 25 kyr preindustrial spin-up. However, the prescribed fluxes are automatically and continuously rescaled in the model to balance the modelled CaCO3 burial rate. Thus, their value is not important (Ridgwell, 2017).

The boundary conditions in our simulations are indeed orbital parameters, aeolian iron deposition rates, detrital flux rate to ocean sediments, and ice sheet fraction and their orography. For the latter, we will make it more explicit in the revision that since our ice sheets are as described in Holden et al., 2010b, only the Laurentide and Eurasian Ice Sheets are allowed to change from their preindustrial form and rather than being extracted uniformly, freshwater to build the LGM ice sheets is routed from the Atlantic, Pacific and Arctic, assuming modern topography.

l. 24-26 on p. 4 should also say that a freshwater flux scaling parameter (FFX) value of 1.5 is applied in GENIE to correct for un-modelled isostatic depression at the ice-bedrock interface due to ice sheet growth, and for assuming a fixed land-sea mask. We vary the parameter in the ensemble to capture the uncertainty in the magnitude of the glacial sea level drop and its effects on the carbon cycle.

Following on from this methodology, the authors lead the reader through the results in a methodical manner, taking on changes in some key environmental variables. In each address of a variable, the authors undergo a detailed assessment of PI versus LGM changes in their model and in proxy records. This is done both in a global sense and a regional sense, sometimes being highly specific. The ground-truthing and comparison to observations and other studies is commendable. It one of the largest comparisons that I have come across in a modelling paper. However, the sheer size of comparison makes the paper unwieldy, and loses the major findings amongst the details. It is also unexpected that the authors concentrate on highly specific regional comparisons because they use an ESM of intermediate complexity that will simply not perform well in many ways. Rather, the authors should focus on the sheer size of their parameter spread and in diagnosing the effect that certain changes have on carbon and climate. I therefore suggest that the authors make an effort to reduce the emphasis on comparison, and focus on how their interesting results might explain the LGM carbon cycle in only a broad sense.

Agreed, in the revision we will modify the text to better take into account the complexity of our model, and to better emphasize the links between the changes we see and CO2.

It occurs to me that the authors seem to forget some obvious strengths of their study. Their parameter space is enormous, includes an interactive carbon cycle with carbon-climate feedbacks, as well as many other processes (sea level) that have not been included in other model studies of the LGM. This is of interest to the LGM pCO2 drop problem. I suggest that the authors try and convey these strengths more clearly in their abstract and the final paragraph of the introduction.

Agreed, the abstract will be reduced to one paragraph and the aim and strengths of our study will introduced at the beginning at the paragraph. We will also clarify these in the current final paragraph of the introduction section, and which now focuses on a description of the ensemble.

A particularly interesting result of this study is the increase in terrestrial carbon under LGM conditions. The authors more often than not find that LGM conditions are associated with greater carbon held in soils because (1) ice-sheet growth covers previously fertile regions and traps the carbon in the soil, (2) soil respiration rates decrease more than primary production, and (3) terrestrial carbon in exposed continental shelves due to sea level drop increases. The authors do not simulate point 3, and there is a possible methodological problem with point 1, in that the authors grow the LGM ice sheets from pre-industrial to LGM state in only 1,000 years. It is unclear in the manuscript presented here whether this rapid growth of ice sheets that is prescribed overtakes highly productive regions that have not yet evolved to become arid tundra land, which have lower carbon content in the soils. A rapid trapping of soil carbon by the unrealistic growth of the ice sheets (almost 100x faster than in reality) would likely overestimate the terrestrial carbon reservoir at the LGM. The authors state in the methods that "Sensitivity simulations were performed to verify the simulated equilibrium state is insensitive to the choice of timescale of ice-sheet growth", and yet the authors do not present any evidence of this sensitivity. I am therefore sceptical of the validity of this result, which I would say is their most important and interesting contribution to the field.

We agree that our statement on sensitivity simulations is confusing, and realise here that the following statement may also need updating in the revision to avoid confusion: “The increases in terrestrial biosphere carbon are predominantly due to our choice to preserve rather than destroy carbon in ice sheet areas. However, the ensemble soil respiration also tends to decrease significantly more than net photosynthesis, resulting in relatively large increases in non-burial carbon”. We briefly address the latter point first.

The statement is based on an analysis of terrestrial carbon partitioning between ice-sheet and non-ice sheet carbon pools in PGACF-16 and from there inference about what happens in PGACF. If the LGM burial carbon inventories in PGACF-16 were to be removed, DTERRC would be negative in 13 out of 16 simulations, despite the fact that terrestrial carbon also increases outside of the ice sheet areas in 15 out of 16 simulations. However, it is not strictly correct to attribute the carbon inventories that are buried to the burial itself. It is the combination of our ice sheet carbon stocks increasing rather than decreasing when exposed to LGM climate, and our choice to preserve rather than destroy this carbon. If most of the carbon that was present in ice sheet areas at the end of the preindustrial runs had been lost to climate forcings, it would not matter much whether the remaining stocks had been destroyed or preserved. From the literature, it is not clear how much of the carbon in ice sheet areas is thought to have been lost strictly in response to ice sheet “bulldozing” versus climate impacts. We will update the revised manuscript to reflect this logic.

With regard to the statement "Sensitivity simulations were performed to verify the simulated equilibrium state is insensitive to the choice of timescale of ice-sheet growth",
This refers to the sensitivity of the LGM burial carbon amount to the timescale of ice sheet growth. To test this, we took one ensemble member from PGACF-16 and applied it to two 11,000-year LGM simulations with 1000-year ice sheet growth and 10,000-year ice sheet growth respectively. Both simulations were started from the end of the same equilibrium preindustrial simulation. We then compared the LGM burial carbon inventories in each run and found that these differed by just 34.2 PgC, which, assuming ~90% of removal of atmospheric $\mathrm{CO_2}$ perturbation by ocean and sediments, amounts to a mere 1.6 ppmv $\mathrm{CO_2}$ difference. Our assumption is that applying the same ensemble member to a transient simulation of the full glacial cycle (and therefore a more realistic ice sheet build-up history) would not have yielded a dramatically different burial carbon inventory. Given that we did a sensitivity experiment with just one ensemble member, we are also assuming that the diagnosed sensitivity is roughly representative of the ensemble as a whole. In the revision, we will highlight these as potential caveats.

With regard to the sign of the ice sheet carbon response at the LGM, we argue that it was not necessarily negative and analysis of PGACF-16 indeed suggests that the sign is consistently positive. Most of this increase is due to a reduction in soil respiration as vegetation carbon change is only positive in one simulation. We also find that extending the timescale of ice sheet growth increases rather than decreases the burial carbon inventory. A likely explanation is that the soil carbon inventory was not yet in equilibrium by 1000 years.

Finally, the conclusion needs complete re-writing. It reads as a re-stating of the methods and then the results, which is simply not useful for the reader. The authors do discuss carbon isotope changes, which are appropriate records to discuss given the main result of an increase terrestrial carbon reservoir. I strongly suggest that the authors re-assess their strengths and major findings, present them concisely, and provide some comments regarding how a higher terrestrial carbon reservoir and low ocean reservoir could have occurred despite most studies indicating the opposite. If this is possible, then this would be a useful contribution to the field.

Agreed, our revised conclusions section will summarise our research objective and the novelties of our approach and then the key conclusions, detailed in the author reply on p. 3. Most of the discussion on carbon isotopes will be shifted to the section on carbon reservoir changes.

Major revisions are needed. If the authors can provide some evidence that changes in the timescale over which the ice-sheets are grown do not play a large role in determining how much carbon is present in the terrestrial reservoir, and the manuscript is rewritten addressing the points above and below, then I advocate publication.

Some other general points:

• Regarding the figures, I strongly suggest that the authors overlay the red (PGACF-16) over the yellow (PGACF) bars in the histograms. This would reduce a lot of unnecessary replication in the figures. The authors could use a transparency setting to ensure that the red and yellow are easily seen if they have the same number of experiments (frequency). Agreed, we will include these changes in the revision.

• The writing is generally okay, but this manuscript definitely requires re-writing in some places. There are too many adjectives and unnecessarily difficult acronyms. OK, we will revise this aspect of our writing.

• The presentation of results suffers from the use of opaque acronyms that are easily forgotten after reading the methods section. I strongly suggest that the references to EFPC2, PGACF and PGACF-16 are changed to be more interpretable. PI315, LGM104, LGM16... or equivalent, would be much more helpful. Agreed, the proposed names will be used in the revision.

2 Specific comments

Abstract

The abstract is quite long. The message that this study delivers could be made considerably more poignant if it were condensed for the reader. The authors make some very interesting findings, such as the increase in non-burial carbon in the terrestrail reservoir due to the slow-down in respiration" and "there are 5 different ways to achieve an atmospheric pCO2 drop". These findings (mostly in para 2) should be the focus of the abstract, and I advocate for the more technical aspects (para 1) and comparisons with observations (para 3) be removed. In fact, the entire paragraph 3 of the abstract could be reduced to one important sentence without affecting the findings of this study.

Agreed, we will shorten the abstract, highlighting our key conclusions, with brief sentences on comparison against observations, and before this our research problem and aim.

Introduction

• Page 3, Line 13 - "dissolve organic carbon inventory" of what? the ocean? soils? Ocean

• Page 3, Paragraph 2 - This entire paragraph simply lists the changes that may be assocaited with a glacial ocean. If these mechanisms are to be called upon by the authors, they should be accompanied by at least a brief discussion of why they influence atmospheric pCO2. For any non-specialist of palaeoceanographic literature specifically relating to the LGM, this paragraph is totally opaque. I would either expand on these mechanisms or remove entirely when accounting for them in the discussion of the results. Agreed, we will briefly describe the referenced mechanisms in the revision.

• Page 3, Line 19 - "utilises an ensemble of sets of parameters" a bit clunky. What about "uses a large ensemble (471 parameter sets)"? OK, we will include this change in the revision.

• Page 3, Line 23 - probably no need to mention Holden (2010a) or cite other literature. Just state philosophy. OK.

• Page 3, Line 26 - move this to methodology. Not necessary here. Agreed, this would make the message clearer.

Methods

• Table 2 - define acronym SAT in caption. What is EFPC2 and EFPC? Must define these here or introduce the Table later on. It is true that the acronyms are difficult to remember and we will change EFPC2 and EFPC to PI315 and PI471, and also add surface air temperature to the caption.

• Page 5, Line 5 - please clarify what a closed biogeochemistry system is. Does it mean no interaction between land-ocean-atmosphere-lithorsphere reservoirs? It means $CaCO_3$ weathering and deep sea sediment burial forced into balance, no sediment-ocean interactions. We will clarify this in the revision.

how are these reservoirs initiated in Stage 1? Using the fields from Holden 2013a?

As described above (general comments), the model spins up from its default state and takes the $CaCO_3$ weathering fluxes diagnosed from 25 kyr preindustrial spin-up of Holden et al., 2013a.

• Page 7, Lines 1-3 - I don't think it's useful to mention this. Yes, we agree and will delete the sentences from the revision.

Preindustrial simulations

• Page 7, Line 5 - First of all, it would be helpful to change the title of section 3 to "Results: Preindustrial simulations". OK, we will make the change.

• Page 7, Lines 7-12 - Why present results and talk about EFPC ensemble when the authors only use the EFPC2 315-member ensemble? It seems to be an unnecessary inclusion that confuses the reader, rather than helps understanding. I think given the length of this study that it would be helpful to simply cut any inclusion of EFPC 471-member ensemble and simply present the results of the 315-member ensemble.
We use EFPC to refer to both our 471-member ensemble and that of Holden et al., 2013a. Our EFPC ensemble is mentioned in the context of explaining where the EFPC2 ensemble came from, and we subsequently compare the response of the latter ensemble with that of Holden et al., 2013a (Table 2), to verify that the values taken by the eight modern plausibility metrics are similar to their values in Holden et al., 2013a. To avoid confusion in the revision, we will move the details of the "ensemble filtering" to a SI file, alongside Table 2 as suggested below.

• Table 2 - change "31/12" to 31st Dec. OK, will include in the revision.

• Table 2 - change "wt% CaCO3" to "wt% ocean CaCO3" wt% $CaCO_3$ refers to surface sediment wt% $CaCO_3$. We will clarify in the revision.

• page 7, Lines 15-23 - Tables 2 and 3 could be moved to supplementary material. Table 2 discussion could be reduced to a single sentence saying that the preindustrial simulations of the 315-member ensemble reproduced all aspects of the Holden 2013a simulations. Table 3 discussion could be reduced to note that there was good agreement with observations of ocean carbon inventory, SSTs and sea ice extent relative to known values.
• Based on what I've said above, I suppose that this section could actually be reduced to one paragraph, or completely removed if the authors wished to use address PI conditions via comparisons with LGM conditions in the next section.

Agreed, we will reduce discussion of the results, and move Tables 2 and 3 to a SI file.

LGM ensemble simulations

• Page 9, Line 7 - "104 ensemble members". Are these presented in yellow in Figure 1? If so, mention it here in the text. Yes, “as shown in yellow in Fig. 1” will be added in the revision.

• Page 9, Lines 7-16 - These sentences are confusing for the reader. I understood once reading further on in the paragraph that you do not include these processes in the model, and you are saying that their inclusion would push LGM pCO2 decrease even further, which justifies your choice of a -30 ppmv threshold to define a successful solution of LGM conditions. However, this is not clear. Please re-write.

• Page 9, Lines 16-26 - This needs to be moved to the methods section. From how I understand your thinking: lines 5-16 justify your choice of -30 ppmv threshold; lines 16-26 justify your methodology in treating the LGM simulations.

• Page 9, Lines 22-25 - The lower threshold of -60 ppmv should belong with your choice of the -30 ppmv upper limit. discuss these together, not separated by other sentences and concepts.

With regard to the above 3 comments, we will address these jointly through revision of p9 l6-26: firstly describing our ΔCO_2 results, and secondly justifying our choice of ΔCO_2 ranges to focus on. Our general approach to analyzing the results will be laid out in the introduction.

• Page 9, Lines 28-33 - This is a very interesting results. Why can’t you define the mechanisms that lead or do not lead to the snowball Earth scenario? Surely if you can define a plausible set of mechanisms need to achieve glacial conditions, you can do the same by comparing the 471 PIs, 16 LGMs, and 47 snowballs??? This would mark a significant contribution to the field.

In this paragraph what we were trying to convey, and we will rephrase in the revision, is that in our EFPC2 ensemble we struggle to achieve LGM atmospheric $CO_2 \leq 200$ ppmv (only ~1.6% of simulations). Not included in this ensemble are 47 LGM simulations which completed but which also predicted global annual average SATs between ca. -67.8 and -56.8°C (and which we assumed are the result of global or near-global sea and land ice cover developing in the simulations, i.e. “snowball earth” type conditions). In 96% of these simulations, atmospheric CO_2 drops to ≤ 200 ppmv at least temporarily. We thus hypothesised that the CO_2 and the “snowballing” may be linked. Establishing causal mechanisms would be interesting but analysis of stage 3 CO_2 time series suggests that the CO_2 is far from equilibrium in many of the “snowball earth” simulations by 10 kyr. We expect different dynamics to operate in the snowball and non-snowball earth states. We hope to investigate this further in the future, in a separate manuscript.

• Figure 1 - Overlay red bars on the yellow bars to make 1 panel. We will implement these changes in the revision for this figure, as well as figures 2 and 4 as suggested below.

• Page 11, Line 11 - define SAT . We will specify that SAT means surface air temperature.

• Figure 2 - Again, overlay red on yellow to reduce unnecessary replication. Use transparency perhaps to show where yellow and red are both present at the same frequency.

• Page 12, Lines 4-16 - Comparisons with obs not necessary at this detail given the focus of the work and the fact that you use an ESM of intermediate complexity. I would expect a discussion of global temperature changes, with perhaps a little bit of basin-wide, regional discussion **if** those are important points for later on. Please reduce this paragraph and combine with the next.

We evaluate the spatial distributions of SAT and SST changes as these are likely to influence our CO2 solution through impacts on the solubility pump, land carbon storage etc. We will articulate this more clearly in the revision and make the comparison with observations less detailed to reflect the focus of our work and the complexity of our model. We will also combine SAT and SST evaluations into one paragraph as the two variables are closely linked.

• Page 15, Line 5 - I'd say there is no relationship. Our statement of a weak relationship is based on r greater or equal than 0.12, in line with our chosen 0.05 significance level. We will clarify this in the revision.

• Page 15, Line 10 - too many adjectives. We will change "The PGACF ensemble LGM global annual sea ice area anomaly (ΔSIA) has a mean of" to "the mean LGM global annual sea ice area anomaly in the PGACF ensemble is"

• Figure 4 - overlay red on yellow.

• Page 18, Lines 11-13 - confusing sentence please re-write. What we wanted to convey here is that the observed precipitation decreases are as negative as -1.1 mm day^{-1} in Southern Europe (SE) and Middle East (ME) whereas the ensemble mean decreases are no greater than ~- 0.5 mm day^{-1} in SE and increases of up to +0.1 mm day^{-1} are simulated in ME. We will make the sentences in the precipitation section more concise and also adapt them to better reflect the complexity of model and the focus of our study.

• Page 19, Lines 11-15 - when talking about AABW formation rates, it is better to present this in positive units. Oceanographers are familiar that the units are negative in the calculation of the streamfunction. It is less confusing for the reader to present your changes as negatives if the AABW formation rate declines. THis also removes the need to explain that a positive anomaly is actually a decrease. Agreed, we will convert our values into negative units in the revision.

• Page 20, Line 2-3 - Don't understand. I thought you said that weaker AMOC and stronger AABW was coincident with the glacial runs, these being PGACF-16? Please make this clearer. WHat are you comparing?

There is a typo in the text which we will fix in the revision. The sentence currently reads: "Although not show here, the PGACF-16 ensemble members tend to exhibit a shoaling of the AMOC and enhanced penetration of AABW. With regard to $\Delta\psi_{\mathrm{max}}$ and $\Delta\psi_{\mathrm{min}}$, these tend to be more negative (i.e. weaker AMOC and stronger AABW) than in the PGACF-16. The $\Delta\psi_{\mathrm{max}}$ and $\Delta\psi_{\mathrm{min}}$ in the EFPC2 ensemble tend to conversely be more positive".

The second mention of "PGACF-16" should be "PGACF".

• Page 20, Lines 3-8 - These relationships are made more confusing for the reader because you define AABW formation rates as being stronger when they are negative. It owuld be more helpful if strong = more positive. In the revision, we will make sure that changes and relationships are reported in a way that avoids confusion about direction.

• Page 20, Line 15 - Please see study of Yu et al (2014) Deep ocean carbonate chemistry and glacial-interglacial atmospheric CO2 changes. We will add an appropriate citation.

• Page 20, final sentence - This has already been covered above? The relationship between sea ice and ocean circulation changes has already been discussed so we agree that the sentence is somewhat redundant. The aim was simply to point out that $\mathrm{CO_2}$ may have had an impact on the ocean circulation changes as well as vice versa. We will make this clearer in the revision, and also include discussion of the proposed link between Antarctic sea ice expansion and ocean circulation in Ferrari et al., 2014 cited below.

Also, see Ferrari et al (2014) Antarctic sea ice control on ocean circulation in present and glacial climates. PNAS.

• Figure 9 - please ensure that your colour bars are the same scale! I initially thought that your LGM simulation had strong AMOCs, despite the discussion of weaker AMOCs in the text. The scales of the colour bars will be updated as part of the revision.

• Page 23 - Very interesting result. I think that this is a unique and interesting contribution to the field and should be a focus of this study. We will highlight these and other key results by reorganizing the abstract in manner suggested above, and by rewriting conclusions section.

• Table 5 - why does the order change? Please clarify in caption of correct. There is no particular reason for this. For clarity, we will update the table so that column 1 shows $\Delta\mathrm{TerrC}$, column 2 $\Delta\mathrm{OceanC}$ and column 3 $\Delta\mathrm{LithC}$.

Also,
add the subheadings "Total counts" and "% of total" or their equivalent to the other columns. OK, we will apply the subheadings to all relevant columns in the revision.

• Table 5 - Please quantify the increases and decreases in this table that accompany the scenarios (i.e. mean). Ok, we will include these in the revised table.

• Page 25, Line 6 - Is table 6 mentioned previously? This introduction of table 6 is jarring. It is indeed mentioned for the first time on line 6. Rather than referring to it at the beginning of a new sentence, we will introduce table 6 more explicitly.

• Table 6 - Great Table. Could the acronyms for studies you cite be organised into alphabetical order? Yes, we will change their order.

• Page 25, Lines 9-10 - I'm (and probably many readers) not sure what "Carbonate compensation of the increased terrestrial carbon storage" means. It's not clear whether you refer to carbon compensation mechanism in the land or the ocean. Do you mean a loss in oceanic DIC due to terrestrial carbon storage, causing an increase in alkalinity that increases CaCO3 burial? If changing terrestrial carbon reservoir does have a direct effect on ocean alkalinity and CaCO3 burial, maybe by weathering changes you account for, then please explain more fully.

We mean here carbonate compensation in response to/of the terrestrial carbon uptake: the loss of $\mathrm{CO_2}$ from the ocean leads to an increase in surface $[\mathrm{CO_3^{2-}}]$ and subsequently deep ocean $[\mathrm{CO_3^{2-}}]$, which reduces $\mathrm{CaCO_3}$ dissolution. The latter in turn decreases $[\mathrm{CO_3^{2-}}]$ and increases $[\mathrm{CO_2}]$, which is communicated back to the surface, with a resultant increase in atmospheric $\mathrm{CO_2}$. The modelled change in $\mathrm{CaCO_3}$ deep sea burial flux causes ALK to change (Kohfeld and Ridgwell, 2009). We will clarify the above in the revised manuscript.

• Page 25, Line 17 - I find this hard to believe. "The PGACF ensemble mean ΔTerrC is of the same sign and order of magnitude as the ΔTerrC predicted by Zimov et al. (2006, 2009), Zech et al. (2011) and Zeng (2003, 2007). ΔOceanC is not directly calculated in these studies". The issue here is likely with the second rather than first sentence. We will remove the sentence from the revised manuscript as our meaning here was that $\Delta\mathrm{OceanC}$ was not reported in the modelling-based studies, and $\Delta\mathrm{OceanC}$ was not calculated in the observations-based studies.

• Page 25, Line 22 - and yet present day continental shelves that are inundated are also regions of effective carbon burial through marine export production. Yes, agreed.

• Tables 7 and 8 - Please make "% Total Land" and its relation to Ice-Sheet carbon more clear in the caption. THis takes time to figure out from the reader. In the revised manuscript, we will change the captions to something like “Ice sheet carbon: amount stored (PgC) and % of total land carbon stock”

• Tables 7 and 8 - I think that these tables could be combined to solely show the LGM changes.

We include Table 7 for 2 reasons: (1) to make it possible to compare our preindustrial ice sheet carbon values with those of previous studies (e.g. Zeng, 2003 or O’ishi and Abe-Ouchi, 2013); (2) to allow estimation of what the impact on ΔCO_2 would have been if this carbon was assumed to be released to the atmosphere as has been done previously (e.g. O’ishi and Abe-Ouchi, 2013). For the revision, we propose combining table 7 and table 8 into one table, but keeping the preindustrial ice sheet carbon inventory column. Columns 3-5 of table 7 can be removed as we do not discuss these in-text.

We also note here that the following sentence will be changed for improved understanding (p27 l1-3): “Analysis of the PGACF-16 ensemble members’ terrestrial carbon reservoirs suggests that if the preindustrial ice sheet carbon inventory (the terrestrial biosphere carbon in grid cells to be buried by the LGM ice sheets) were to be destroyed instead of preserved at the LGM, the $\Delta \mathrm{TerrC}$ would be negative in all but 3 simulations (Tables 7 and 8)”. What we mean here is that if we did not bury carbon, $\Delta \mathrm{TerrC}$ would be negative in all but 3 simulations. This includes the amount of carbon initially present in ice sheet areas (preindustrial ice sheet carbon inventory) and the subsequent increase in terrestrial carbon over the ice sheet growth period. If only the preindustrial ice sheet carbon inventory was substracted from $\Delta \mathrm{TerrC}$, the latter would be negative in all but 7 simulations. As a reminder, one of the 16 simulations predicts negative $\Delta \mathrm{TerrC}$ to begin with.

• Page 27, Lines 1-16 - The relationship between ice-sheet carbon, non-ice sheet carbon and soil burial carbon needs to be made clearer. We will clarify this through the use of subscripts: $\Delta \mathrm{TerrC_{ice}}$, $\Delta \mathrm{SoilC_{ice}}$, $\Delta \mathrm{SoilC_{non-ice}}$, etc.

• Page 31, Line 7 - re-write sentence please. We re-write the sentence as “The mean LGM total POC export flux anomaly ($\Delta \mathrm{POC_{exp}}$) in the PGACF ensemble is”

• Page 31, Lines 14-15 - Also the relationship with AABW production and decreased AMOC, which you have just discussed. Agreed, we will update this sentence to include sea ice effects on ocean circulation.

• Page 33, Section 4.5 - This tight description of the main effects is what the other sections should emulate, and would tighten up the manuscript considerably. Agreed. The discussion of spatial changes will be considerably reduced in the revised manuscript.

• Page 35, Line 3 - Again, this sentence is awful to read. Too many adjectives. Agreed. We will replace it with “the mean global deep-sea $\mathrm{CaCO_3}$ burial flux anomaly ($\Delta\mathrm{CaCO3_{bur}}$) in the PGACF ensemble is”

• Page 36, Line 7 - I don’t see a decline in the figure, just no change. The current colour legend indeed makes it difficult to distinguish between no change, small positive and small negative changes. We will center the legend on white and update the text to reflect the plotted changes.

Conclusion

• Page 38, Line 26 - A decrease in ocean POC export is not necessarily associated with an increase in atmospheric pCO2. Please see Sigman et al (2010) The polar ocean and glacial cycles in atmospheric CO2 concentration. Nature. for an explanation. Briefly: it is not total POC export, but the efficiency of carbon fixation relative to outgassing that matters. Agreed, we will clarify in the revision that the assumption here is that the impact of our decrease in POC export is not offset by a decrease in the rate at which remineralised carbon is returned to the surface. We will, however, re-discuss potential caveat of no increase in remineralisation depth with decreasing ocean temperature.

• Page 39, Line 13 - Why are you now talking about terrestrial carbon here?
Terrestrial carbon gets mentioned here as we are summarizing our results. However, the text will become easier to follow as we re-write the conclusions section to summarise our research objective, novelties of our approach and then the key conclusions.

• Page 40, Lines 1-11 - Please see Menviel et al (2017) Poorly ventilated deep ocean at the Last Glacial Maximum inferred from carbon isotopes: A data-model comparison study. Paleoceanography.
Here we attempt to broadly reconcile our positive $\Delta\mathrm{TerrC}$ with the mean ocean $\delta^{13}\mathrm{C}$ change and do not indeed discuss the spatial distribution of $\delta^{13}\mathrm{C}$, which is another useful constraint, as highlighted by the study of Menviel et al., 2017 (and which interestingly suggests weaker, not stronger AABW transport). We will acknowledge this as an added source of uncertainty in the revision.

• Page 40, Lines 13-14 - But atmospheric _13C at the LGM and preindustrial climates were similar at -6.46 ‰ and -6.36 ‰, respectively.
Agreed. We discuss the role that a glacial increase in terrestrial carbon inventory may have played in the glacial-interglacial $\delta^{13}\mathrm{C}$ record but do not attempt to definitely close the $\delta^{13}\mathrm{C}$ cycle as a detailed evaluation against observations was not the focus of the paper. In the revision, we will make the similarity between preindustrial and LGM atmospheric $\delta^{13}\mathrm{C}$ levels more explicit when discussing the deglacial record.

• Page 40, Line 28 - And what do these new records show?
As above, our aim was not to go into the $\Delta^{14}\mathrm{C}$ records in detail but simply acknowledge that there have so far not been attempts to reconcile glacial increases in terrestrial carbon with higher resolution atmospheric $\mathrm{CO_2}$ and $\Delta^{14}\mathrm{C}$ deglacial records. These include a significant

decline in atmospheric $\Delta^{14}C$ around 14.6 kyr BP, which Köhler et al., 2014 attribute to permafrost thawing in high northern latitudes, as well as possibly flooding of the Siberian continental shelf. Marcott et al., 2014 in turn show that a significant fraction of the deglacial CO_2 rise direct radiative forcing occurred in steps of 10-15 ppm, over less than two centuries, and was followed by no notable change in atmospheric CO_2 for ~1000-1500 years.

• Page 40, Line 30 - The deglacial decrease in atmospheric _14C could also be caused by the exchange of highly negative ocean _14C with the atmosphere. The argument here is not clear.

Our argument here, which we will clarify in the revision is that if ^{14}C-depleted carbon was released from the land to the atmosphere during deglaciation, and subsequently absorbed by the ocean, one might expect to see this signal in ocean $\Delta^{14}C$ records. Instead, for the deep ocean at least, we see an increase in $\Delta^{14}C$ and the size of the perturbation has been argued to support an overall positive $\Delta \mathrm{OceanC}$. Hence, there are potential caveats with the enhanced LGM terrestrial carbon hypothesis. However, we also discuss limitations of $\Delta^{14}C$ data: p.41, l5-12.

3 Technical corrections

We will revise the manuscript to include the technical corrections below.

1. Page 7, Line 9 - "did not evidence numerical instability" –> "did not show evidence of numerical instability"
2. Page 7, Line 15 - "has" –> "had". Please use past tense in results sections.
3. Page 9, Line 7 - "impact of error" –> "error"
4. Page 9, Line 7 - "impact of certain" –> "certain"
5. Page 9, Line 7 - ", such as changing" –> ", such as changing" (two spaces)
6. Page 15, Line 4 - "thereis" –> "there is"
7. Page 15, Line 15 - "may have for instance" –> "may have"
8. Page 20, Line 21 - "for instance is that" –> "is that"
9. Page 23, Line 3 - "Most of the ensemble members" –> "Most members"
10. Page 26, first line of table - there is a tab separating -1160 from 530.
11. Page 36, Line 11 - double space –> single space
12. Page 38, Line 19 - "Terr" –> "TerrC"
13. Page 38, Line 27 - remove /

Response to referee #3

We thank referee #3 for the constructive comments and suggestions. As above, our replies are in black, and original comments in blue.

This article presents an ensemble of Last Glacial Maximum (LGM) simulations using the GENIE intermediate complexity model with varying parameter values. The model simulates the carbon cycle allowing the authors to compare the CO2 values obtained from the model with the 90 ppm decrease from ice core data and analyses carbon stocks, in addition to the study of climate variables. They select two subsets of simulations with increasing constraints and analyses the changes obtained in these simulations in terms of temperature, precipitation, sea ice, ocean circulation and carbon stocks with respect to the pre-industrial. The CO2 drawdown in most simulations is due to an increase of carbon storage on land and in the lithosphere while the ocean gets depleted in carbon.

Using an ensemble of simulations to study the change of climate and of the carbon cycle during the LGM is a great idea and GENIE is an adequate model as it is fast enough for the long simulations required. However, concerning the carbon cycle part, the lack of carbon isotopes in the simulations prevents any real conclusion to be drawn on the plausibility of the results obtained and the likelihood of the associated mechanisms and carbon stocks changes. As the carbon isotopes are already incorporated within GENIE, providing that this is feasible, I suggest to rerun the simulations and redo the analyses with the isotopes (at least carbon 13) in comparison to data, which is crucial to properly evaluate the results.

As suggested in replies to previous comments, we will clarify our research aim, which was not to explain definitively the causes of CO_2 but rather take an uncertainty-based approach to exploring the physical and biogeochemical changes which may have accompanied the LGM CO_2 decrease. Given our focus, we seek to compare our simulation results with observations only more broadly.

General comments

1. As I said before, the main point is the absence of carbon isotopes which precludes any strong conclusion to be drawn, it would be best to redo the simulations with the carbon isotopes (at least C13, if possible C14) and compare with ocean and atmosphere data to evaluate which simulations are really plausible. How long would this take?

We address this comment in the general response but also note here that spinning up the $\delta^{13}C$ of the ocean would take > 100,000 years, and that the model ensemble does not simulate $\Delta^{14}C$ and $\delta^{13}C$ in the terrestrial biosphere.

2. From the simulations done so far, we can’t know which processes are responsible for the carbon changes, it would be great to have a few additional sensitivity tests (for example taking one set of parameters from the PGACF ensemble) to evaluate the impact of each process on CO2.

We agree that these experiments would be interesting but are not essential to our study, which focuses on the ensemble as a whole rather than the response of individual ensemble members. We also note again here that while this paper describes the relationships between ensemble outputs, the second (related) paper to be submitted describes dependencies on ensemble parameters to isolate mechanisms.

3. In the figures with maps, it would be good to draw the coastlines of continents to make it easier to see where the changes take place. Agreed, we will add these in the revised manuscript.

4. Â´n conversely Â˙z appears 28 times in the manuscript, it could probably be removed or replaced a few times. There seems to be a typo here which precludes understanding?

5. On the bar charts, all ensembles (grey, yellow and orange) could be drawn on the same plot to avoid having two subplots, which would help the comparison between the ensembles and reduce the space taken by figures. Agreed, we will combine all 3 ensembles in the revision.

6. The hypothesis that carbon stays below the ice sheets is a strong one, it could be interesting to evaluate its impact by doing one (or a few) simulations from the PGACF ensemble without carbon kept under ice sheets.

We agree that it would be interesting to test this directly if there were resources for additional simulations. However, we do show how much carbon is stored below the ice sheets in PGACF-16 and from there one can at least try to estimate what the impact of releasing it would have been on CO_2. Indeed, the LGM burial carbon inventory varies between ~300 and 1300 PgC and if we assume that 90% of the initial atmospheric CO_2 perturbation removed by the oceans and sediments, atmospheric CO_2 would increase by between ~14 and 61 ppmv.

However, it is important to note here (as also emphasized in an earlier reply) that the importance of assumptions regarding what happens to carbon in ice sheet areas depends on how much carbon is there. Our LGM burial carbon estimates include the initial preindustrial carbon inventories, plus carbon accumulated in response to glacial forcings.

Specific comments

p.1-2: The abstract is quite long, and could be better organized, with the problematic explained at the beginning before stating what is the main scientific question raised in this article, how it is raised, and then the main results.

Agreed, the abstract will be reorganized in manner suggested above and reduced to just one paragraph.

p. 2-3: Permafrost is not mentioned in the introduction; it would be good to include it.

Yes, thank you, permafrost growth should here be mentioned as a separate mechanism. The text will be revised accordingly.

It would also be interesting to introduce here which data will be later used to constrain the results.

Agreed. We will include an overview of upcoming sections at the very end of the introduction section, and in this overview describe which model variables will be compared against observations.

p. 6: During the second stage of the simulations, how does CO2 evolve, does it stay stable? We plotted the evolution of atmospheric CO_2 from stage 1 through 4 in PGACF-16: CO_2 stays stable, except for maybe 3 runs where CO_2 changes by ~< 10 ppmv but then reaches an equilibrium again.

p. 7 l. 12: In the EFPC ensemble, are the simulations at equilibrium at the end of stage 3? What is meant here is probably the EFPC2 ensemble. We again have plots of atmospheric CO_2 and surface sediment %wt $CaCO_3$ for a subset of this ensemble (PGACF-16). Both metrics either in or nearing equilibrium by 10 kyr.

p.7 l.21: The SST value is too high compared to data, how does that compare to other models? Is it in the range?

We deem the mean value (18.9 °C) to be comparable to other previous model-based estimates, in that e.g. Kim et al., 2003 predict modern SST of ~18 °C, Zhang et al., 2012 predict preindustrial SST of 17.1 °C. The range is 16.4 to 21.9 °C, which is potentially larger than the range of previous model-based estimates.

p.7 l.22 The sea ice value is given for the Northern and Southern Hemispheres, how is the comparison with data when split between the North and the South?

Our model is set up to output time series of annual average global sea ice area, 31/12 NH and 31/12 SH sea ice areas. The latter is one of the modern plausibility metrics used in Holden et al., 2013a and as shown in Table 2, our estimates are comparable to those of Holden et al., 2013a. For 31/12 NH sea ice area, the mean of the EFPC2 ensemble is 15.1 ± 1.4 million km^2, and the range is 12.6 to 19.2 million km^2. The mean is within the range of typical late winter Arctic sea ice cover today (14-16 million km^2) (NSIDC). Note that the preindustrial NH sea ice extent during the month of maximum (winter) extent simulated by the 13 PMIP2 and PMIP3 models shown in Fig. 4 of Goosse et al., 2013 ranges between ~ 13 and 27 million km^2.

p. 7: The vegetation and soil carbon values are given in table 2 but are not discussed. How does it compare to data ? Is the vegetation distribution ok? Given that it plays an important role in the change of CO2 for the LGM it would be good to know if the preindustrial terrestrial biosphere is well represented or if it has important biases. There is also no discussion of the overturning values given, how does it compare to other models?

We kept discussion of the preindustrial results to a minimum to cut down on the amount of text. It is true that we do not compare any of the values in table 2 (including overturning values) against observations - this is not intentional and we will add a note that there are no major differences between our results (EFPC2) and those of Holden et al., 2013a (EFPC), which meet previously chosen modern plausibility criteria (i.e. modelled values within acceptable distance of observations). We did plot the spatial distribution of vegetation and soil carbon in EFPC2 but did not feel that the discrepancies were large enough to significantly bias our LGM results and therefore did not include the plots in the manuscript. There is however, one potential exception and we mention this on p. 28: "However, it is also noteworthy that, although not shown here, the regions with the largest decreases in terrestrial carbon density, namely northwest North America, Beringia and the Tibetan plateau area, are also the regions with the largest terrestrial carbon densities in the preindustrial simulations". This refers to the terrestrial carbon densities in the EFPC2 ensemble mean and we will clarify in the revision that in the preindustrial EFPC2 ensemble mean: (i) the Tibetan soil carbon peak is overestimated and (ii) the North American soil carbon peak misplaced (compared to observations). We attribute (i) to the lack of soil weathering in the model and the inclusion of land use effects in the observational data-based estimate (Holden et al., 2013b; Williamson et al., 2006). We attribute (ii) to the lack of explicit representation of permafrost (instead the model only attempts to capture the soil respiration rates characteristic of permafrost by utilising a distinct soil respiration temperature sensitivity for land temperatures below freezing) (Williamson et al., 2006) and the absence of moisture control on soil respiration.

p. 9 l. 28-29 I'm not sure I understand or agree with this sentence as the simulations are for the LGM and not the other glacial maxima in terms of orbital parameters.
Agreed, conclusions we draw for the LGM may not be generalizable to other glacial maxima. We will rephrase this in the revision.

p. 10 figure 1: maybe replace PRE by PI and explain it somewhere: Pre-industrial (PI).
We consistently use PRE to denote preindustrial but can change this to PI for improved understanding.

p. 11 l.10 and following: Could you use temperature and salinity data to select ensemble members that are supported by data?
Although a useful suggestion, it goes against our approach of looking at the ensemble more widely and not putting too much emphasis on individual ensemble members/strongly constraining these. This will be clarified in the revised introduction, which will help understand the current presentation of results.

p. 15 line 10: how does sea ice distribution compare with data?

As mentioned above, our model is set up to output time series of 31/12 NH and 31/12 SH sea ice areas, and we also have maps of the annual average spatial distribution of sea ice. There are no obvious observation-based estimates to compare the latter against, or 31/12 NH sea ice area. For 31/12 SH sea ice area, our estimates can be compared against the estimates of Gersonde et al., 2005 and Roche et al., 2012. In the first study, LGM summer sea ice extent is estimated to have increased by between 1-2 million km^2, which is much smaller than our PGACF ensemble mean of 11.4 $\pm$ 6 million km^2, and falls outside of our 3 to 32.6 million km^2 PGACF ensemble range. However, as noted in Gersonde et al., 2005, major uncertainties concern the reconstruction of summer sea ice extent. Roche et al., 2012 predict increases in LGM summer sea ice extent between ~2 and 12 million km^2. We also note here that based on Fig. 4 in Goosse et al., 2013, the LGM change in SH sea ice extent during the month of minimum (summer) extent predicted by the 13 PMIP2 and PMIP3 models ranges between ~ -3 and 25 million km^2.

To estimate how our simulated 31/12 NH sea ice compares with data, we will, in the revision, compare our LGM change in the annual average spatial distribution of sea ice with reconstructed changes in winter and summer sea ice extents in the NH. We note here, however, that the mean LGM change in NH 31/12 sea ice area in the PGCAF ensemble is 7.3 million km2 $\pm$ 2.1 million km^2, and the range is 3.6 million km^2 to 13.2 million km^2. For comparison again, based on Fig. 4 in Goosse et al., 2013, the range of estimates predicted by the 13 PMIP2 and PMIP3 models included therein goes from -7 to 4 million km^2.

We also note here an error in the current manuscript: l13-16 p.12. Contrary to our statement, it is unlikely (or at least not more likely than not) that the LGM increase in annual average SH sea ice is underestimated given that the LGM increase in 31/12 SH sea ice lies at the upper end of observed estimates. Winter SH sea ice also here means 31/12 SH sea ice not austral winter sea ice. We will revise the paragraph.

p. 20 l. 3 Is it really "than in the PGACF-16"? Is this not the ensemble that you are talking about? Thank you, there is a typo, it should indeed say "PGACF".

p. 20 line 10: NADW instead of AABW? We use the brackets here to mean that expanding AABW cell may restrict the AMOC cell (i.e. the upper cell of the AMOC) to lower depths and expanding AMOC cell may restrict AABW to higher latitudes. We will clarify this in the revised manuscript.

Figure 9: It looks like the NADW is stronger for the LGM than the Pre-industrial , while from the text and figure 8 I understood the opposite: : :
We will put the colour bars on the same scale to avoid confusion in the revised manuscript.

Figure 10: could you add the PGACF-16 ensemble? Yes, good suggestion.

p. 23 and following: could you show a map of where the carbon is stored on land ?
We show the spatial distribution of vegetation, soil and total land carbon changes on p.30. As per suggestion of referee #1, we will show the spatial distributions first, then the globally-integrated numbers, in the revised manuscript.

p. 37: the conclusion is long and more descriptive than conclusive, it might be good to re-organize it.

Agreed. We will shorten the current conclusions section, succinctly summarizing the objective of our research and research strengths, and then describe the key conclusions.

References

Gersonde, R., Crosta, X., Abelmann, A. and Armand, L.: Sea-surface temperature and sea ice distribution of the Southern Ocean at the EPILOG Last Glacial Maximum - A circum-Antarctic view based on siliceous microfossil records, Quat. Sci. Rev. 24, 869–896, 2005.

Goosse, H., Roche, D.M., Mairesse, A., Berger, M.: Modelling past sea ice changes, Quat. Sci. Rev., 79, pp. 191–206, 2013.

Holden, P. B., Edwards, N. R., Wolff, E. W., Lang, N. J., Singarayer, J. S., Valdes, P. J. and Stocker, T. F.: Interhemispheric coupling, the West Antarctic Ice Sheet and warm Antarctic interglacials, Clim. Past, 6(4), 431–443, doi:10.5194/cp-6-431-2010, 2010b.

Holden, P. B., Edwards, N. R., Müller, S. A., Oliver, K. I. C., Death, R. M. and Ridgwell, A.: Controls on the spatial distribution of oceanic δ13CDIC, Biogeosciences, 10(3), 1815–1833, doi:10.5194/bg-10-1815-2013, 2013a.

Holden, P. B., Edwards, N. R., Gerten, D. and Schaphoff, S.: A model-based constraint on CO2 fertilisation, Biogeosciences, 10(1), 339–355, doi:10.5194/bg-10-339-2013, 2013b.

Kim, S.-J., Flato, G. and Boer, G.: A coupled climate model simulation of the Last Glacial Maximum, Part 2: approach to equilibrium, Clim. Dyn., 20(6), 635–661, doi:10.1007/s00382-002-0292-2, 2003.

Kohfeld, K. E. and Ridgwell, A.: Glacial-Interglacial Variability in Atmospheric CO2, Surf. Ocean. Atmos. Process., 251–286, doi:10.1029/2008GM000845, 2009.

Köhler, P., Knorr, G. and Bard, E.: Permafrost thawing as a possible source of abrupt carbon release at the onset of the Bølling/Allerød, Nat. Commun., 5, doi:10.1038/ncomms6520, 2014.

Marcott, S. A., Bauska, T. K., Buizert, C., Steig, E. J., Rosen, J. L., Cuffey, K. M., Fudge, T. J., Severinghaus, J. P., Ahn, J., Kalk, M. L., McConnell, J. R., Sowers, T., Taylor, K. C., White, J. W. C. and Brook, E. J.: Centennial-scale changes in the global carbon cycle during the last deglaciation, Nature, 514(7524), 616–619, doi:10.1038/nature13799, 2014.

O'Ishi, R. and Abe-Ouchi, A.: Influence of dynamic vegetation on climate change and terrestrial carbon storage in the Last Glacial Maximum, Clim. Past, 9(4), 1571–1587, doi:10.5194/cp-9-1571-2013, 2013.

Ridgwell, A.: Examples for cGENIE: ‘muffin’ pre-release version. January 5, 2017. http://www.seao2.info/cgenie/docs/cGENIE.muffin.Examples.pdf

Roche, D. M., Crosta, X. and Renssen, H.: Evaluating Southern Ocean sea-ice for the Last Glacial Maximum and pre-industrial climates: PMIP-2 models and data evidence, *Quat. Sci. Rev.* **56,** 99–106, 2012.

Williamson, M. S., Lenton, T. M., Shepherd, J. G. and Edwards, N. R.: An efficient numerical terrestrial scheme (ENTS) for Earth system modelling, Ecol. Modell., 198(3–4), 362–374, doi:10.1016/j.ecolmodel.2006.05.027, 2006.

Zeng, N.: Glacial-interglacial atmospheric CO2 changes – the Glacial Burial Hypothesis. – Advances in Atmospheric Sciences, 20: 677–693, 2003.

Zhang, Z. S., Nisancioglu, K., Bentsen, M., Tjiputra, T., Bethke, I., Yan, Q., Risebrobakken, B., Andersson, C., and Jansen, E.: Pre-industrial and mid-Pliocene simulations with NorESM-L, Geosci. Model Dev., 5, 523-533, 2012.