Response to the second set of reviews

One referee had no further comments and recommended acceptance, while another referee with a second reading had further minor comments to address. We thank the referees for their positive recommendations and attention to detail.

Comments to address

In general the authors have answered many of my questions within the response to reviewers document but have not clarified the main text in response. The authors make several useful arguments in the response to reviewers that should be added to the main text prior to publication.

First example, the authors have in the response to reviewers that "The framework is designed to provide a quantitative measure of the different drivers of the ZEC. Without a quantitative measure, one is left making qualitative comparison of thermal and carbon effects when those variables are measured in different ways. Our framework provides a formal way of comparing the relative importance of each driver." This would be very helpful context for readers early on in the MS.

We thank the referee for the additional comments and by recommending that we are more explicit in making these connections.

Added in Abstract, L5-7:

In order to understand these different climate responses, a normalised framework is introduced that quantifies the relative importance of carbon, radiative and thermal drivers of the ZEC.

Expanded in the Introduction, L39-42:

A framework is introduced that formally compares the relative importance of these

thermal, radiative and carbon drivers for the ZEC (Section 2). Without a quantitative measure, only a qualitative comparison of thermal and carbon effects can be made, which is complicated by those variables being measured in different ways. These

drivers for the ZEC are interpreted

Expanded in the Conclusions, L403-404:

In order to gain mechanistic insight as to the controls of the ZEC, a normalised framework is introduced that formally compares the relative importance of thermal, radiative and carbon drivers for the ZEC.

Modified in the Conclusions, L472-474:

In summary, our normalised framework provides a formal comparison of the different thermal, radiative and carbon contributions to the ZEC, and so provides mechanistic insight as to why there are different temperature responses from Earth system models after carbon emissions cease.

Second example, "For example, for the ZECMIP diagnostics for the geometric ZEC, we find that..." (three points follow)

The framework is designed to provide a quantitative measure of the different drivers of the ZEC. Without a quantitative measure, one is left making qualitative comparison of thermal and carbon effects when those variables are measured in different ways. Our framework provides a formal way of comparing the relative importance of each driver.
For example, for the ZECMIP diagnostics for the geometric ZEC, we find that
(i) The intermodal spread in the geometric ZEC is primarily controlled by the intermodal spreads in the normalised thermal contribution and normalised atmospheric carbon concentration, rather than that of the normalised radiative forcing dependence on atmospheric CO2 (Table 1b);
(ii) The intermodal spread of the normalised contribution to the warming dependence on radiative forcing is mainly determined in ZECMIP by the intermodal spread in the fraction of radiative forcing returned to space rather than that of the inverse climate feedback (Table 1c);
(iii) The intermodal spread of the airborne fraction is mainly determined by the intermodal spread of the landborne fraction, rather than the oceanborne fraction (Table 1c).

I don't see these three conclusions clearly stated in the manuscript. This seems like relevant context for presenting this new metric of ZEC - please include it in the main text.

These points were separately explained in the main manuscript, but we have modified the Conclusions to explicitly reiterate these points:

Conclusions, L428-437:

Applying our normalised framework to the ZECMIP diagnostics reveals that relative importance of the different drivers for the inter-model spread of the geometric ZEC: (i) The inter-model spread in the geometric ZEC is primarily controlled by the intermodal spreads in the normalised thermal contribution and normalised atmospheric carbon concentration, rather than that of the normalised radiative forcing dependence on atmospheric CO2 (Table 1b); (ii) The inter-model spread of the normalised contribution to the warming dependence on radiative forcing is mainly determined in ZECMIP by the intermodal spread in the fraction of radiative forcing returned to space rather than that of the inverse climate feedback (Table 1c); (iii) The inter-model spread of the airborne fraction is mainly determined by the intermodal spread of the landborne fraction, rather than the oceanborne fraction (Table~1c).

"We find that the model responses separate into different classes. For the carbon response, the land carbon sink either continues to increase in time or saturates. These different responses appear to be linked to whether there is a nutrient cycle that can inhibit the ability of the land to take up unlimited carbon."

The authors make this point in the abstract and at the end of the conclusions, however it would be helpful to connect this simple summary with the discussion of land sink saturation (implied as due to nutrients ~ line 255).

We have added new text to discuss the land carbon response, see below.

Third example
"For example for the inter-model spread in the climate responses, the framework reveals..." followed by three points.
I can see that these points are made in the text, but not as directly as they are here. I think the MS would benefit from directly stating these findings.

As above we have included these points in the Conclusions, L428-437.

Applying our normalised framework to the ZECMIP diagnostics reveals that relative importance of the different drivers for the intermodel spread of the geometric ZEC: (i) The intermodel spread in the geometric ZEC is primarily controlled by the intermodal spreads in the normalised thermal contribution and normalised atmospheric carbon concentration, rather than that of the normalised radiative forcing dependence on atmospheric CO2 (Table 1b); (ii) The intermodel spread of the normalised contribution to the warming dependence on radiative forcing is mainly determined in ZECMIP by the intermodal spread in the fraction of radiative forcing returned to space rather than that of the inverse climate feedback (Table 1c); (iii) The intermodel spread of the airborne fraction is mainly determined by the intermodal spread of the landborne fraction, rather than the oceanborne fraction (Table~1c).

With regards the quantification of our analyses, we have added new comparisons linked to Table A1 providing estimates of the relative magnitude of the different contributions to the variance of ZEC, the carbon contribution to the ZEC and the thermal contribution to the ZEC:

New paragraph, L236-241

The normalised contributions to ZEC from the thermal, radiative and carbon responses can also be analysed to quantify the contribution of each term to the spread across models. By varying just one model input term in Table A1, the thermal and carbon terms explain 58% and 40% respectively of the variance in ZEC, whereas the radiative term explains only 2% of the variance. This analysis confirms that both the model spread in thermal response and the model spread in carbon sink both contribute significantly to *the spread in ZEC, and both remain high priority research areas to understand in order to reduce uncertainties in ZEC.*

These competing carbon and thermal contributions for the ZEC are next addressed in more detail in terms of their own dependencies.

New paragraph, L274-279

In common with the analysis of Jones and Friedlingstein (2020), by varying just one model input term at a time in Table A1, the land carbon sink is again found to dominate the spread in the carbon sink contribution to ZEC, accounting for 78% of the variance in the carbon sink compared with the ocean sink explaining 22% of the variance at 50 years after net zero. The magnitude of land and ocean sinks are similar on this timescale, but the model spread is greater for the land sink. On longer timescales beyond a century, we expect the land carbon sink to saturate more rapidly and the ocean carbon sink to play a progressively more important role.

Line 306-308

In addition, by varying just one model input at a time in Table A1, the fraction of radiative forcing returning to space explains 67% of the variance in the thermal contribution to the ZEC compared with 33% from the variance in the inverse climate feedback parameter.

Evidence for nutrient limitation:
While nutrient limitation seems like a very plausible hypothesis for saturating land carbon sinks I don't see any analysis or quantification of evidence for that presented in this MS. Please provide evidence or make it clear that this statement is a hypothesis that needs to be tested. For example CESM2 has an increasing sink but includes interactive nutrient cycling. Also I don't see evidence that it is nitrogen specifically causing the nutrient limitation (ACCESS-ESM1.5 also has phosphorus represented). If the authors argue that this is covered already by Aurora et al. 2020 they need to make that argument in the MS more clearly.

The statements about nutrient limitation are more than a hypothesis and are in accord with two other published studies by Zaehle et al. (2015) and Ziehn et al. (2021), both demonstrating the limitation of the land carbon sink through nutrient availability. The models with the smallest landborne fraction are those that have a nutrient limitation. For CESM2, there may still be an increase in the land carbon sink, but the increase in the land carbon sink is less than it would be without nutrient limitation. Models usually represent nutrient limitation in terms of nitrogen limitation (the exception is ACCESS that has phosporus limitation). So the separation between models is more typically represented by whether nitrogen limitation included.

Added new text on the land carbon sink, L260-270

These different relative strengths of the land and ocean carbon sinks are likely due to structural differences in the land carbon model which contributes a much greater model spread than the ocean response (Jones and Friedlingstein., 2020). The three models with higher landborne fractions may be overestimating the possible land carbon sink as they neglect the role of nutrient limitations, while the other models include land nitrogen cycling and limitation of carbon allocation. Models without an explicit terrestrial nitrogen cycle can project unrealistic land carbon sinks which could not be supported by available nutrients (Zaehle et al., 2015). In particular, Ziehn et al. (2021) showed explicitly for the ACCESS model, the important role of nutrients – both nitrogen and phosphorus - in reducing land carbon sinks. Arora et al. (2019) showed this inclusion or absence of nutrient limitation to be the largest systematic difference in the carbon response of CMIP6 Earth system models, with a distinct split in the land carbon response to climate and CO2 between models with and without a nitrogen cycle. Consequently, the inclusion of nitrogen limitation on land acts to reduce the increase in the projected land carbon sink and so acts to increase the resulting ZEC.

Two new references added:

Zaehle, S., Jones, C. D., Houlton, B., Lamarque, J.-F., and Robertson, E.: Nitrogen availability reduces CMIP5 projections of twenty-first-century land carbon uptake, Journal of Climate, 28, 2494–2511, 2015.

Ziehn, T., Wang, Y., and Huang, Y.: Land carbon-concentration and carbon-climate feedbacks are significantly reduced by nitrogen and phosphorus limitation, Environmental Research Letters, 16, 074 043, 2021.

Also added in Conclusions

L441 while other models (ACCESS-ESM1.5, UKESM1) have the land sink eventually saturating due to nitrogen limitation and the ocean sink dominating

discussion of WASP: "There are two reasons for this "coefficient of variation being larger for the landborne and oceanborne fractions than the airborne fraction", one to do with the WASP ensemble and another to do with the system itself..." (and following response) Please add at least a sentence or two making this clarification in the main text.

Agreed. We have added the text L385-387:

This reduced spread of the airborne fraction in WASP is also partly due to how the WASP ensembles are constructed with historic constraints on atmospheric carbon being much narrower than the historic constraints for land and ocean carbon.

I don't find the term "geometric ZEC" intuitive but I am not opposed to this label.
We think that this term is mathematically accurate and avoids the ambiguity of using term of normalised ZEC (as raised previously by the referee), where different choices are possible.

As a note the MS uses intermodel, inter-model, and in the response to reviewers intermodal. Please make consistent.

Changed to inter-model.

line 225 - it would be helpful to add "which are discussed in further detail below" or something similar
Agreed, added on L224

; these contributions are discussed in more detail in the next subsections.

comment about original line 398-402

The text has been modified more than is represented by the highlighted version. Either way I think it would be helpful to add the references to where the evidence for these statements is found in the MS (i.e. Fig 4).

The new changes are highlighted in the pdf.

Added the extra figure references to support the points made:

L4443 (Fig. 4, green line; Fig. 5)

L446 (Fig. 4, red line; Fig. 6);

L448 (Fig. 4),

Figure 9. I understand that the authors don't wish to do further analysis of additional datasets, but I don't see why this couldn't in principle be done for the existing flat10MIP ESM runs.

In principle the analyses can be repeated for the flat 10 scenario, but those analyses would delay the manuscript. The relative importance of the thermal and carbon contributions is unlikely to be determined by the scenario. For example, we do provide an analysis for 1pctCo2 and flat10 for WASP integrations in Figures 9 and 10, and Tables S2 and S3, and the responses only differ in detail.

Added in Conclusions, L463-465

To partly address the above caveats, diagnostics of a large ensemble of an efficient Earth system model (WASP) are also examined for 2 different scenarios (1pctCO2 and flat10). Their ensemble median responses for the ZEC and its contributions are broadly similar for both scenarios and to those of ZECMIP.

L494 We have included thanks for the constructive comments that have strengthened the study.