

Emulating absent land processes reduces carbon budgets in CMIP6 models

Corresponding Author: Mr Spencer Liddicoat

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

1. Overall recommendation

I recommend acceptance after minor revisions. The manuscript is well written, methodologically innovative, and addresses a highly relevant problem: how missing land processes in CMIP6 ESMs bias TCRE and the carbon budget from 1.5 °C to 2 °C. Using a UKESM1.1 process ensemble to derive multiplicative emulation factors and apply them to 11 CMIP6 models is, in my view, a valuable contribution to both the Earth system modelling and carbon budget communities.

2. Main strengths

- The study is clearly motivated by the central role of TCRE in defining (remaining) carbon budgets.
- The process ensemble design (control plus six single-process perturbations and one all-process configuration) is transparent.
- The results are quantitatively clear: emulating missing processes raises the multi-model mean TCRE by ~21.4% and reduces the 1.5–2.0 °C carbon budget by ~16%, robust across eight different budget formulations.
- The paper makes a useful forward-looking link to CMIP7, arguing that wider inclusion of processes such as nitrogen limitation, fire, and wetland methane will likely further raise TCRE and tighten budgets.

3. Points for clarification (minor)

I see no need for additional simulations or major analysis. I would only suggest a few clarifications, mainly to improve transparency about assumptions:

- a. Transferability of emulation factors Please expand briefly on the assumptions behind applying UKESM-derived multiplicative factors (Fproc, Gproc) to other CMIP6 models that differ in land surface formulation and baseline TCRE. A short paragraph explaining under what conditions these factors are expected to be most reliable, and where they might be less appropriate, would be helpful.
- b. Process independence and interactions The framework assumes that the combined impact of multiple processes can be represented as the product of independent factors. Since you show that rescaled ukesm-ctrl matches ukesm-allprocs very closely, it would be useful to quantify that residual explicitly (e.g. “difference in TCRE and CE is X %”) to reassure readers that non-linear interactions are small in practice.
- c. The description of the UKESM process ensemble and the 1pctCO2 setup is clear at a high level, and I understand that full details are given in Liddicoat et al. (2026a,b) and Mulcahy et al. (2023). For the benefit of readers not closely familiar with these earlier studies, I would encourage the authors to add 1–2 sentences explicitly summarizing (i) the key aspects of the 1pctCO2 configuration used to diagnose TCRE and compatible emissions, and (ii) the main features of the UKESM spin-up / initial conditions relevant for this analysis. Similarly, a brief reminder that the CMIP6 TCRE values come from the standard CMIP6 1pctCO2 experiments as analyzed in Arora et al. (2020) would make the experimental basis more self-contained.

d. Policy-relevant implications

The conclusion that CMIP6-based estimates of remaining carbon budgets may be optimistic because many models omit processes that increase TCRE and reduce budgets is important and well supported. It would be useful to add one or two sentences explicitly connecting this to observational / emergent-constraint estimates of TCRE and to how your adjusted CMIP6 values align or diverge from them, while keeping the focus on your emulation framework.

Reviewer #2

(Remarks to the Author)

The submitted paper by Liddicoat and colleagues examines the potential consequences for the 2°C target remaining carbon budget (RCB, considering 1.5°C as a reference) of adopting six partially unrepresented land processes in eleven CMIP6 ESMs. Building on a process ensemble of UKESM1-0-LL factorial simulations, they build emulation factors (EF) for TCRE and cumulative emissions (CE) to translate the effect of these six processes from the UKESM to the other ESMs. By partially accounting for uncertainty stemming from non-CO₂ gases, unrepresented feedbacks (ESF) and zero-emission effect (ZEC) on climate, they claim that the remaining carbon budget is reduced on average by 16% in the transition from 1.5°C to 2°C global warming average.

The paper is well structured and well written, the methodology is solid, it is explained in a clear manner and the results are convincing, in my opinion. The major claim of the paper is of extremely scientific and policy relevance, since it shows that our current calculations of the RCB based on state-of-the-art ESMs are likely overestimated. First, this paper will be of interest to those modelling groups which are in charge of producing the CMIP7 generation of ESMs, stressing the need to improve processes representation in land surface models in order to have more realistic estimates of TCRE and RCB. However, given the policy relevance of the claims, I was probably expecting some contextualization with other estimates, most simply the Global Carbon Budget 2025, of the RCB obtained from CMIP6 models, eventually. Additionally, I am not surprised by the results themselves, as from the theoretical point of view, the atmospheric CO₂ burden is expected to increase when multiple land processes currently not included in ESMs are accounted for, as elucidated by the authors in line 318-328. Still, it is very nice to have this contribution quantified.

I don't have any major concerns about this submitted paper. However, the following point is a little unclear to me, and I hope the authors could clarify accordingly:

The emulation factors are estimated from UKESM and then applied to all the other ESMs to mimic their response to additional process representation. The underlying assumption is therefore that the inclusion of that process has the exact same impact on TCRE and CE - quantitatively - regardless of the ESMs it is implemented in, something I'm not completely sure about. In line 318-320, the authors state that while the EF are mode-specific, the findings are sound in land carbon sink theory, which I agree, but probably some more discussion on this aspect (and potentially limitation) could help to better understand an additional source of uncertainty.

That said, I don't have any concern about the quality of the work and its fit with the purpose of the Journal. I therefore recommend it for publication.

Reviewer #3

(Remarks to the Author)

Main comments

This manuscript provides valuable insights into updating TCRE and carbon budgets for the upcoming IPCC cycle. By utilizing UKESM and CMIP6 multi-model data, the authors proposed a "decomposition" analysis that effectively attributes the influence of specific biogeochemical processes for vegetation and soil to TCRE and carbon budget estimates.

My primary concern lies in the reliance on concentration-driven rather than emission-driven simulations. While concentration-driven runs were a standard practice for CMIP6, this complicates the narrative and obscures the physical intuition of the results.

TCRE and carbon budgets are high-stake, policy-relevant metrics. It is therefore important to present the analysis with a clear focus on emissions. In addition, as the modeling community transitions toward CMIP7, which will prioritize emission-driven protocols, the current analysis may risk limited applicability. The authors should provide a more rigorous and extended discussion on the limitations of their current approach and how these results can be translated to emission-driven context. For example, how might the TCRE and carbon budget estimates based on compatible emissions be potentially biased?

Furthermore, the distinction between this work and Liddicoat et al. (2026a, 2026b) requires clarification. One is marked as "accepted", while the other is "in prep" with a specific journal name. The authors should explicitly state the unique contributions and novelties of this study to justify its publication as an independent paper in a more transparent manner.

Line-by-line comments

L117

Following my major comment above, this section is an example of where the authors could discuss the limitations of the use of concentration-driven runs, relative to more intuitive and policy-relevant-emission driven runs.

L169-170

The authors rely on an assumption of linear additivity here. While the later comparison with the six-process case suggests this assumption is largely valid, it should be explicitly acknowledged at this point in the text. Explicitly stating this assumption early on, rather than waiting until the end of the manuscript to justify it, would improve the clarity.

L175

It is unclear about how CE' is used in the rest of this paper. Is CE' used to estimate TCRE'?

L191

The authors should provide a clear rationale for the inconsistent treatment of natural feedbacks in their framework. Specifically, emissions from permafrost (CO₂-only?) are accounted for externally to TCRE in variable ESF, whereas methane emissions from wetlands are accounted internally as Fwch₄ within TCRE. Both are natural emissions that can dynamically change with anthropogenic climate change.

L191

In addition, I question the relevancy of including the ZEC in this specific analysis. This analysis discusses doubling CO₂ experiments. Emissions are nowhere near zero.

L194

Please clarify what the uncertainty in the present warming level referred to here is.

L200

Are the ZEC estimates discussed here based on concentration-driven runs? If so, please explicitly state it here.

L230

Carbon budget variables require differentiated notations.

Fig 2

Please clarify why lower-case u does not agree with the control result.

Fig 2

The AR6 mean value could be presented as an additional point of reference, with the best estimate and the uncertainty range.

Fig 4

Why does non-CO₂ have only a small impact? Methane is the second most important GHG at present in terms of radiative forcing.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Thank you for the responses. I am satisfied with the revisions and recommend accepting the manuscript.

Reviewer #2

(Remarks to the Author)

The reviewed manuscript as well as the authors replies to the comments and questions of the reviewers are satisfactory in my opinion. I therefore recommend the paper for publication.

Reviewer #3

(Remarks to the Author)

I highly appreciate the revised paper and the authors response. At places, however, I had struggles to assess whether my previous comments had been properly addressed because the author replies are sometimes not very explicit or do not address the issue directly. Below are my persisting concerns to highlight. If anything arises from a misunderstanding on my part, the authors should feel free to clarify.

First comment

I am still not fully convinced by the authors' argument that emission-driven and concentration-driven runs give an identical TCRE. The underlying temperature pathways at least slightly differ between these two runs, leading to different levels of natural CO₂ release through climate-carbon cycle feedback and resulting in different warming levels. Even though the two runs consider anthropogenic emissions of 1000 PgC for quantifying TCRE, the underlying natural CO₂ emissions should differ at least slightly, leading to different temperature outcomes and affecting the TCRE estimate.

L191

The authors wrote, "ESF accounts for additional greenhouse gas emissions due to natural feedbacks omitted from ESMs". They also wrote, "Including interactive wetland methane emissions influenced TCRE solely through the radiative impact on climate." For clarity, I would write, "ESF accounts for additional CO₂ emissions..." if non-CO₂ is not at all included in ESF.

Furthermore, I would maintain that the treatment of wetland CH₄ emissions represents an inconsistency in the TCRE calculation method. The radiative effect of wetland CH₄ emissions is considered within TCRE, whereas non-CO₂ effects are generally not included in TCRE quantification. The ESF term does not account for wetland CH₄ emissions, despite claiming to account for "future emissions arising from Earth system feedbacks not accounted for in the derivation of TCRE, such as

carbon release from permafrost thaw". While I don't consider this a significant issue, the authors could comment on it.

Version 2:

Reviewer comments:

Reviewer #3

(Remarks to the Author)

The authors have adequately addressed my comments and concerns. I therefore recommend the paper for publication.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Response to Reviewers of “Emulating absent land processes reduces carbon budgets in CMIP6 models” by Liddicoat et al.

Reviewer #1 (Remarks to the Author):

1. Overall recommendation

I recommend acceptance after minor revisions. The manuscript is well written, methodologically innovative, and addresses a highly relevant problem: how missing land processes in CMIP6 ESMs bias TCRE and the carbon budget from 1.5 °C to 2 °C. Using a UKESM1.1 process ensemble to derive multiplicative emulation factors and apply them to 11 CMIP6 models is, in my view, a valuable contribution to both the Earth system modelling and carbon budget communities.

Thank you for taking the time to review the manuscript and for your encouraging comments.

2. Main strengths

- The study is clearly motivated by the central role of TCRE in defining (remaining) carbon budgets.
- The process ensemble design (control plus six single-process perturbations and one all-process configuration) is transparent.
- The results are quantitatively clear: emulating missing processes raises the multi-model mean TCRE by ~21.4% and reduces the 1.5–2.0 °C carbon budget by ~16%, robust across eight different budget formulations.c
- The paper makes a useful forward-looking link to CMIP7, arguing that wider inclusion of processes such as nitrogen limitation, fire, and wetland methane will likely further raise TCRE and tighten budgets.

3. Points for clarification (minor)

I see no need for additional simulations or major analysis. I would only suggest a few clarifications, mainly to improve transparency about assumptions:

- a. Transferability of emulation factors Please expand briefly on the assumptions behind applying UKESM-derived multiplicative factors (Fproc, Gproc) to other CMIP6 models that differ in land surface formulation and baseline TCRE. A short paragraph explaining under what conditions these factors are expected to be most reliable, and where they might be less appropriate, would be helpful.

Yes, this is a fair comment. It is difficult to say with any certainty under which circumstances the multiplying factors would be expected to be most / least reliable. If an ESM has a similar sensitivity to climate and CO₂ to UKESM, it might be expected that the emulation factors would be more reliable than when applied to an ESM with quite different sensitivities, but this is not necessarily the case. Ultimately, how a new process will change a model's sensitivities depends on the structural changes and parametrizations required to implement it: until this is done, it's not easy to predict the impact of a new process within a model, and therefore how reliable our emulation factors will be for any particular model. Although CanESM5 has a positive gamma at 4xCO₂, which might complicate matters if we were applying our scaling factors at 4xCO₂, it and all other ESMs have a negative gamma at 2xCO₂, the level at which TCRE is evaluated and our framework applied, so this complication is avoided. We have

added a coupled of sentences to a paragraph within the Discussion (italicized) in which we justify the use of multiplicative scaling factors rather than additive correction terms:

“The outcome of our emulation framework when applied to another ESM should be considered a first order estimate, rather than a precise prediction, of how its TCRE and CE would change, were the ESM to be developed to include the process. The precise impact of any new process depends on the structural changes and parametrizations of its implementation and how they alter the model’s sensitivity to changing CO₂ and climate. Nevertheless, we can be confident in the overall direction of travel predicted by our emulation framework because our findings are consistent with theoretical prediction of the behaviour of the land carbon sink, and, by extension, cumulative emissions and TCRE. For example, fire...”

b. Process independence and interactions The framework assumes that the combined impact of multiple processes can be represented as the product of independent factors. Since you show that rescaled *ukesm-ctrl* matches *ukesm-allprocs* very closely, it would be useful to quantify that residual explicitly (e.g. “difference in TCRE and CE is X %”) to reassure readers that non-linear interactions are small in practice.

That’s a good point, thank you. The CE matched exactly (to 3 significant figures) at 952 GtC, while the rescaled *ukesm-ctrl* TCRE of 3.13°C/1000 GtC is 0.3% lower than the *ukesm-allprocs* value of 3.14°C/1000 GtC. We’ve added these details at the end of the first section in Results: *“The rescaled ukesm-ctrl values represent a deviation of 0.3% in TCRE, and 0.0% in CE_{2xCO2} from those of ukesm-allprocs.”*

c. The description of the UKESM process ensemble and the 1pctCO₂ setup is clear at a high level, and I understand that full details are given in Liddicoat et al. (2026a,b) and Mulcahy et al. (2023). For the benefit of readers not closely familiar with these earlier studies, I would encourage the authors to add 1–2 sentences explicitly summarizing (i) the key aspects of the 1pctCO₂ configuration used to diagnose TCRE and compatible emissions, and (ii) the main features of the UKESM spin-up / initial conditions relevant for this analysis. Similarly, a brief reminder that the CMIP6 TCRE values come from the standard CMIP6 1pctCO₂ experiments as analyzed in Arora et al. (2020) would make the experimental basis more self-contained.

In the Methods section, we have expanded briefly on the other Liddicoat et al. (2026) studies to clarify their purpose, and therefore how they differ from the present manuscript; described briefly the spin-up protocol; added some more detail on the calculation of TCRE, and explicitly mentioned the Arora et al. (2020) study.

d. Policy-relevant implications

The conclusion that CMIP6-based estimates of remaining carbon budgets may be optimistic because many models omit processes that increase TCRE and reduce budgets is important and well supported. It would be useful to add one or two sentences explicitly connecting this to observational / emergent-constraint estimates of TCRE and to how your adjusted CMIP6 values align or diverge from them, while keeping the focus on your emulation framework.

We considered comparing our 1.5°C to 2°C carbon budget with the equivalent from the Global Carbon Budget 2025 and with the Cox et al (2024) emergent constraints paper. However, the carbon budgets presented in these two studies derive ultimately from historical and SSP experiments with forcing from multiple GHGs, aerosols and land-use change,

compared to our framework based on 1pctCO₂ experiments driven only by changes in CO₂. Therefore, carbon budgets derived from the two experimental approaches are not directly comparable, and we've expanded the final paragraph of the paper to describe how our study complements GCB2025 and Cox et al. without providing a direct numerical comparison:

"Our results make clear the significant impact that differences in process representation within ESMs have on TCRE and carbon budgets calculated from it. This process-based approach complements top-down studies such as the Global Carbon Budget 2025¹⁸ and Cox et al. (2024)³⁷ which use observations-based evidence to constrain remaining carbon budgets across the CMIP6 ESMs' historical and Shared Socioeconomic Pathway experiments, sampling uncertainty in multiple drivers of climate change (greenhouse gases, aerosols and land use forcing). By contrast, we provide evidence of structural uncertainty arising from the process differences embedded within the models. Furthermore, our results suggest that because many models lack important processes like nitrogen limitation and fire-vegetation interactions known to limit long-term land carbon storage, such estimates of remaining carbon budgets may be overly optimistic."

Reviewer #2 (Remarks to the Author):

The submitted paper by Liddicoat and colleagues examines the potential consequences for the 2°C target remaining carbon budget (RCB, considering 1.5°C as a reference) of adopting six partially unrepresented land processes in eleven CMIP6 ESMs. Building on a process ensemble of UKESM1-0-LL factorial simulations, they build emulation factors (EF) for TCRE and cumulative emissions (CE) to translate the effect of these six processes from the UKESM to the other ESMs. By partially accounting for uncertainty stemming from non-CO₂ gases, unrepresented feedbacks (ESF) and zero-emission effect (ZEC) on climate, they claim that the remaining carbon budget is reduced on average by 16% in the transition from 1.5°C to 2°C global warming average.

The paper is well structured and well written, the methodology is solid, it is explained in a clear manner and the results are convincing, in my opinion. The major claim of the paper is of extremely scientific and policy relevance, since it shows that our current calculations of the RCB based on state-of-the-art ESMs are likely overestimated. First, this paper will be of interest to those modelling groups which are in charge of producing the CMIP7 generation of ESMs, stressing the need to improve processes representation in land surface models in order to have more realistic estimates of TCRE and RCB.

Many thanks for your thorough review, as well as for your positive feedback.

However, given the policy relevance of the claims, I was probably expecting some contextualization with other estimates, most simply the Global Carbon Budget 2025, of the RCB obtained from CMIP6 models, eventually.

Yes, this is a fair comment. To reiterate our response to Reviewer #1, we did consider comparing our 1.5°C to 2°C carbon budget with the equivalent from the Global Carbon Budget 2025 and with the Cox et al (2024) emergent constraints paper. However, the carbon budgets presented in these two studies derive ultimately from historical and SSP experiments with forcing from multiple GHGs, aerosols and land-use change, compared to our framework based on 1pctCO₂ experiments driven only by changes in CO₂. Therefore, carbon budgets derived from the two experimental approaches are not directly comparable, and we've expanded the final paragraph of the paper to describe how our study complements GCB2025 and Cox et al. without providing a direct numerical comparison:

“Our results make clear the significant impact that differences in process representation within ESMs have on TCRE and carbon budgets calculated from it. This process-based approach complements top-down studies such as the Global Carbon Budget 2025¹⁸ and Cox et al. (2024)³⁷ which use observations-based evidence to constrain remaining carbon budgets across the CMIP6 ESMs’ historical and Shared Socioeconomic Pathway experiments, sampling uncertainty in multiple drivers of climate change (greenhouse gases, aerosols and land use forcing). By contrast, we provide evidence of structural uncertainty arising from the process differences embedded within the models. Furthermore, our results suggest that because many models lack important processes like nitrogen limitation and fire-vegetation interactions known to limit long-term land carbon storage, such estimates of remaining carbon budgets may be overly optimistic.”

Additionally, I am not surprised by the results themselves, as from the theoretical point of view, the atmospheric CO₂ burden is expected to increase when multiple land processes currently not included in ESMs are accounted for, as elucidated by the authors in line 318-328. Still, it is very nice to have this contribution quantified.

I don’t have any major concerns about this submitted paper. However, the following point is a little unclear to me, and I hope the authors could clarify accordingly:

The emulation factors are estimated from UKESM and then applied to all the other ESMs to mimic their response to additional process representation. The underlying assumption is therefore that the inclusion of that process has the exact same impact on TCRE and CE - quantitatively - regardless of the ESMs it is implemented in, something I’m not completely sure about. In line 318-320, the authors state that while the EF are model-specific, the findings are sound in land carbon sink theory, which I agree, but probably some more discussion on this aspect (and potentially limitation) could help to better understand an additional source of uncertainty.

Thank you for this feedback, we agree that this section would benefit from further clarification. Therefore, we have extended the paragraph within the Discussion in which we justify our use of multiplicative scaling factors rather than additive correction terms by preceding it with two new sentences:

“The outcome of our emulation framework when applied to another ESM should be considered a first order estimate, rather than a precise prediction, of how its TCRE and CE would change, were the ESM to be developed to include the process. The precise impact of any new process depends on the structural changes and parametrizations of its implementation and how they alter the model’s sensitivity to changing CO₂ and climate. Nevertheless, we can be confident in the overall direction of travel predicted by our emulation framework because our findings are consistent with theoretical prediction of the behaviour of the land carbon sink, and, by extension, cumulative emissions and TCRE. For example, fire...”

That said, I don’t have any concern about the quality of the work and its fit with the purpose of the Journal. I therefore recommend it for publication.

Many thanks for your useful comments, and your thorough review.

Reviewer #3 (Remarks to the Author):

Main comments

This manuscript provides valuable insights into updating TCRE and carbon budgets for the upcoming IPCC cycle. By utilizing UKESM and CMIP6 multi-model data, the authors proposed a "decomposition" analysis that effectively attributes the influence of specific biogeochemical processes for vegetation and soil to TCRE and carbon budget estimates.

My primary concern lies in the reliance on concentration-driven rather than emission-driven simulations. While concentration-driven runs were a standard practice for CMIP6, this complicates the narrative and obscures the physical intuition of the results.

TCRE and carbon budgets are high-stake, policy-relevant metrics. It is therefore important to present the analysis with a clear focus on emissions. In addition, as the modeling community transitions toward CMIP7, which will prioritize emission-driven protocols, the current analysis may risk limited applicability. The authors should provide a more rigorous and extended discussion on the limitations of their current approach and how these results can be translated to emission-driven context. For example, how might the TCRE and carbon budget estimates based on compatible emissions be potentially biased?

Thank you for your feedback, the choice of concentration vs emissions-driven runs is always an important consideration in experiments which focus on developing our understanding of the carbon cycle / climate interactions.

The definition of TCRE is the warming caused by emissions of 1000 PgC, and this quantity is independent of pathway. As such it can be diagnosed from either concentration-driven experiments, typically from the warming evaluated at $2\times\text{CO}_2$ in normalised units of $^{\circ}\text{C} / 1000 \text{ GtC}$, as done here, or from emissions-driven simulations as the warming experienced after 1000 GtC have been emitted, i.e. 100 years into the 'flat10' experiment with fixed emissions of 10 GtC/yr, which give equivalent results (see Sanderson et al (2025): GMD - flat10MIP: an emissions-driven experiment to diagnose the climate response to positive, zero and negative CO2 emissions).

Up to and including CMIP6, the 1% concentration-driven simulation has been used to diagnose TCRE. From CMIP7 onwards diagnosis will primarily be based on the flat10 emissions-driven simulations. But in order to provide an update on the TCRE diagnosed from other models it is important that we use the same protocol. In this case, therefore, we use the 1% simulation to derive the emulation factors for application to the CMIP6 models. Note, however, that the concentration-driven 1% CO_2 simulations still form part of C4MIP, so will also be performed at CMIP7, and a direct comparison with the 1pctCO2 experiments performed at CMIP6 will be possible.

From CMIP7 onwards, once multi-model results are available from the Assessment Fasttrack, this approach could be repeated using the flat10 simulations.

Furthermore, the distinction between this work and Liddicoat et al. (2026a, 2026b) requires clarification. One is marked as "accepted", while the other is "in prep" with a specific journal name. The authors should explicitly state the unique contributions and novelties of this study to justify its publication as an independent paper in a more transparent manner.

The two papers are now published and under review, respectively. The published paper focuses solely on the UKESM results, to understand why the processes impact TCRE in the way that they do – see [Role of Earth system processes in the relationship between climate change and cumulative carbon emissions | Nature Communications](#). The other paper does the same for the carbon cycle feedback parameters, β and γ , and is currently under review at the Journal of Climate.

The present manuscript is the only one which applies the results from the UKESM process ensemble to the CMIP6 ESMs, specifically to see how emulating the processes absent from each ESM impacts its TCRE, which we then take further by looking at how the updated TCREs impact the carbon budget from 1.5°C to 2°C. Hence, it is entirely distinct from the first two. We have briefly indicated the above in the manuscript when we first refer to each paper - thank you for pointing out the need for this.

Line-by-line comments

L117

Following my major comment above, this section is an example of where the authors could discuss the limitations of the use of concentration-driven runs, relative to more intuitive and policy-relevant-emission driven runs.

We have added the following sentence to the end of the paragraph starting on line 117: *An alternative approach would be to diagnose TCRE from CO₂ emissions-driven simulations as the change in global mean temperature after 1000 GtC have been emitted, but as we are applying our emulation factors to the CMIP6 ESMs' TCRE taken from Arora et al (2020)⁹, we use the CMIP6 protocol of concentration-driven 1pctCO2 simulations.*

L169-170

The authors rely on an assumption of linear additivity here. While the later comparison with the six-process case suggests this assumption is largely valid, it should be explicitly acknowledged at this point in the text. Explicitly stating this assumption early on, rather than waiting until the end of the manuscript to justify it, would improve the clarity.

Thank you, that's a very useful observation. Since we use our *ukesm-ctrl* vs *ukesm-allprocs* result, we keep the main discussion of this point in the Discussion (i.e. after all results have been presented), but draw attention to this assumption at the point you refer to (immediately following lines 169-170):

“Note that this framework relies on multiplicative emulation factors rather than additive correction terms, which assumes that the impacts of omitted processes are independent and combine linearly. We present our justification for this approach in the Discussion, with reference to the particular case in which we use the framework to emulate the impact on TCRE and CE of the four processes absent from *ukesm-ctrl*.”

L175

It is unclear about how CE' is used in the rest of this paper. Is CE' used to estimate TCRE'?

Figure 2 is modelled on the classic plots from AR5 and AR6 showing the near linear change in GSAT vs cumulative CO₂ emissions which forms the basis of TCRE: Figure 2 shows the

particular case of TCRE and cumulative emissions evaluated when the atmospheric CO₂ has doubled. The CE calculation is primarily used in the figure and table, and isn't used in the calculation of TCRE explicitly. Although our focus is on TCRE, the figure provides visual evidence of how the CE are changed by the processes, e.g. including fire-vegetation interactions shifts the data point to the left, strongly reducing the CE.

L191

The authors should provide a clear rationale for the inconsistent treatment of natural feedbacks in their framework. Specifically, emissions from permafrost (CO₂-only?) are accounted for externally to TCRE in variable ESF, whereas methane emissions from wetlands are accounted internally as F_{wch4} within TCRE. Both are natural emissions that can dynamically change with anthropogenic climate change.

Thank you for your comment, we have altered the text to clarify this point.

Including interactive wetland methane emissions influenced TCRE solely through the radiative impact on climate, not by changing the land carbon sink, airborne fraction and compatible emissions (i.e. the diagnosed cumulative emissions) - this result is included in our first paper, looking at the impact of the processes on TCRE in UKESM.

The ESF term exclusively accounts for emissions of CO₂ from processes not represented in the models from which TCRE is calculated. In the case of the 11 CMIP6 ESMs in this paper, none of them included interactive methane emissions. Additionally, as mentioned above, the inclusion of wetland methane emissions didn't change the cumulative emissions at all in TCRE. So, there is no double-counting when using F_{wch4} to scale TCRE of the CMIP6 models, and ESF when working out the carbon budget from 1.5°C to 2.0°C.

We have added the following clarification (italicized here) to the manuscript:

*“The release of CO₂ to the atmosphere due to Earth system feedbacks not included in the calculation of TCRE is highly uncertain. We use $ESF=7.1$ GtC, converted from the value of 26 ± 97 GtCO₂ provided in Lamboll et al (2023)¹⁹. Since the same ESF value is used for every model it reduces each ESM's carbon budget by the same absolute amount: how this influences the percentage reduction when using the revised TCRE relative to the original depends on the choice of other terms used in the calculation of the carbon budget. Alternatively, since ESF is both small – 7.1 GtC is less than the carbon emitted in one year at 2020s emission rates – and highly uncertain, it could reasonably be omitted. *Note that ESF accounts for additional greenhouse gas emissions due to natural feedbacks omitted from ESMs, primarily permafrost thaw, though wetland methane emissions are sometimes nominally included in such terms⁵. In *ukesm-wch4*, interactive wetland methane emissions change TCRE solely via their radiative influence on climate¹³, without changing land carbon uptake. The use of F_{wch4} in rescaling the CMIP6 ESMs' TCRE is therefore entirely distinct from the use of ESF to account for emissions of CO₂ from unrepresented processes, ensuring no double-counting occurs.*”*

L191

In addition, I question the relevancy of including the ZEC in this specific analysis. This analysis discusses doubling CO₂ experiments. Emissions are nowhere near zero.

Typically, when evaluating the carbon budget required to limit warming to a particular target, the zero emission commitment needs to be considered. Even if emissions were to cease altogether on reaching 2°C, if a further 0.2°C or 0.3°C of warming were inevitable, this effectively reduces the temperature at which emissions must cease in order for the ultimate target (2°C) to not be exceeded, so ZEC must be accounted for in the carbon budget calculation.

L194

Please clarify what the uncertainty in the present warming level referred to here is.

Thank you for pointing out this omission. We have added a sentence quoting the WMO that the 2023-2025 three-year average global mean temperature was $1.48 \pm 0.13^\circ\text{C}$ relative to the 1850-1900 average.

L200

Are the ZEC estimates discussed here based on concentration-driven runs? If so, please explicitly state it here.

The ZEC estimates from MacDougall et al (2020) are from emissions-driven runs after emissions cease, so the atmospheric CO₂ is calculated within the models in response to the land and ocean sinks, and the temperature responds accordingly.

L230

Carbon budget variables require differentiated notations.

Thank you, we've added superscripted indices indicating the particular equation number each refers to (e.g. superscripted 3 indicates equation 8.3).

Fig 2

Please clarify why lower-case u does not agree with the control result.

The lower-case U in Figure 2 refers to UKESM1-0-LL, whereas *ukesm-ctrl* (and all others in the process ensemble) is a version of UKESM1.1, which is the result of some development to improve the historical climate response with bias corrections and minor bug fixes relative to UKESM1-0-LL. The climate response in the two versions of the model is not identical, hence they appear in close proximity, but do not coincide exactly with one another in this figure.

Fig 2

The AR6 mean value could be presented as an additional point of reference, with the best estimate and the uncertainty range.

Thank you for pointing this out, we have updated the figure to include the multi-model mean as a triangle within the boxplots, and updated figure caption accordingly.

Fig 4

Why does non-CO₂ have only a small impact? Methane is the second most important GHG at present in terms of radiative forcing.

The eight different methods of calculating the carbon budget from 1.5°C to 2°C illustrate just how dominating TCRE is in the calculations. The other terms do cause changes to the magnitude of the carbon budget, as indicated by the table included in Figure 4, but the main driver is the TCRE. So when a significantly different value of TCRE is used, in our case the rescaled values, it is that which drives the magnitude of the change in the carbon budget: the 8 different equations accounting for different combinations of ESF, T_{ZEC} and T_{nonCO_2} show that the choice of which of these terms to include is of secondary importance to the changing TCRE value.

In the case of non-CO₂ forcing, the T_{nonCO_2} term is used in the numerator of Equation 8.3 for each ESM, having some impact on the overall carbon budget, as indicated by the table in the bottom right of the figure: accounting for it reduces the model mean CB from 295 GtC to 246 GtC when using the original TCRE, and from 263 GtC to 219 GtC when the rescaled TCREs are used (the data points are slightly lower in panel c) than in panel a)). In each case, this reduction is 16.5%, because the same T_{nonCO_2} value is used for all ESMs and cancels out during the percentage reduction calculation. This doesn't happen when ESF, or the model-specific ZEC, is the sole term used in addition to TCRE. (Note that the figures in the table lower right of Figure 4 are rounded to 3 significant figures for clarity: calculating the reduction from those yields values of 16.6% to 16.7%, while the percentages in each panel are calculated from the raw data).

Response to Reviewers of “Emulating absent land processes reduces carbon budgets in CMIP6 models” by Liddicoat et al.

Reviewer #1 (Remarks to the Author):

Thank you for the responses. I am satisfied with the revisions and recommend accepting the manuscript.

Many thanks for reviewing the changes to the manuscript and for your recommendation of acceptance.

Reviewer #2 (Remarks to the Author):

The reviewed manuscript as well as the authors replies to the comments and questions of the reviewers are satisfactory in my opinion. I therefore recommend the paper for publication.

Many thanks for reviewing the changes to the manuscript and for recommending it for publication.

Reviewer #3 (Remarks to the Author):

I highly appreciate the revised paper and the authors response. At places, however, I had struggles to assess whether my previous comments had been properly addressed because the author replies are sometimes not very explicit or do not address the issue directly. Below are my persisting concerns to highlight. If anything arises from a misunderstanding on my part, the authors should feel free to clarify.

Thank you for again reviewing the manuscript, apologies for the lack of clarity in our previous responses. We'll try to address your remaining concerns below.

First comment

I am still not fully convinced by the authors' argument that emission-driven and concentration-driven runs give an identical TCRE. The underlying temperature pathways at least slightly differ between these two runs, leading to different levels of natural CO₂ release through climate-carbon cycle feedback and resulting in different warming levels. Even though the two runs consider anthropogenic emissions of 1000 PgC for quantifying TCRE, the underlying natural CO₂ emissions should differ at least slightly, leading to different temperature outcomes and affecting the TCRE estimate.

Yes, you're correct that the climate and carbon cycle response in the flat10 (emissions-driven) and 1pctCO₂ (concentration-driven) experiments would not be exactly equivalent, due in part to the rate of change of CO₂ – therefore, TCRE when calculated from the two different experiments would not be expected to be identical, but it has been shown to be extremely similar. Sanderson et al (2025) specifically address this question of the equivalence of the TCRE calculated from the two different experimental configurations. In

their Figure 8a (reproduced below), they plot TCRE calculated after 1000 GtC has been emitted under flat10 forcing (100 years of 10 GtC/yr) against TCRE calculated after 1000 GtC of diagnosed emissions under 1pctCO2 forcing, for ten ESMs which had done both experiments. They find an extremely close match across all ESMs, with the line of best fit being $y=0.08+1.02x$ - very close to a 1:1 equivalence, with a very small y-intercept of 0.08.

So, while it's true that the two approaches won't lead to identical TCREs, nevertheless the TCRE is a fundamental emergent property of the model's climate and carbon cycle response to CO₂ forcing which is robust to the method of calculation and the CO₂ forcing approach adopted.

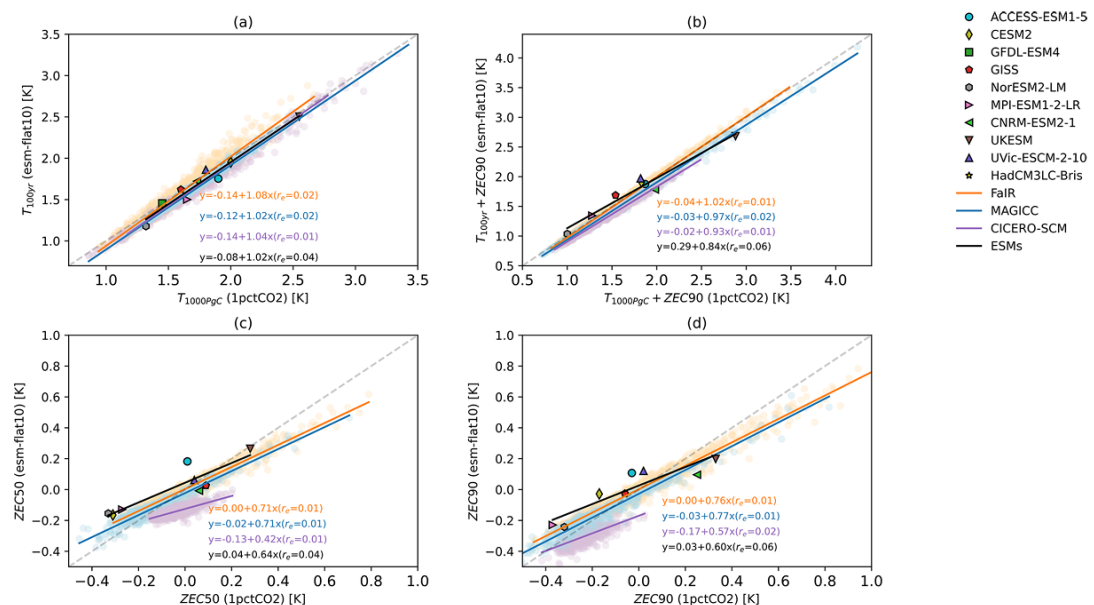


Figure 8. Comparison metrics assessed using the flat10MIP methodology and 1pctCO₂-based experiments. ESM summary metrics are T_{100yr} , ZEC50 and ZEC90 for esm-flat10 and $T_{1000PgC}$ and ZEC50 and ZEC90 for 1pctCO₂. Filled shapes illustrate values assessed from ESMs, and pale dots illustrate members of the simple climate model ensembles for (FaIR, MAGICC, CICERO-SCM) in (orange, blue, purple). Straight lines show the least-squares best fits for the ESMs (black) and SCMs.

Reference: Sanderson, B. M., Brovkin, V., Fisher, R. A., Hohn, D., Ilyina, T., Jones, C. D., Koenigk, T., Koven, C., Li, H., Lawrence, D. M., Lawrence, P., Liddicoat, S., MacDougall, A. H., Mengis, N., Nicholls, Z., O'Rourke, E., Romanou, A., Sandstad, M., Schwinger, J., Séférian, R., Sentman, L. T., Simpson, I. R., Smith, C., Steinert, N. J., Swann, A. L. S., Tjiputra, J., and Ziehn, T.: flat10MIP: an emissions-driven experiment to diagnose the climate response to positive, zero and negative CO₂ emissions, *Geosci. Model Dev.*, **18**, 5699–5724, <https://doi.org/10.5194/gmd-18-5699-2025>, 2025.

L191

The authors wrote, "ESF accounts for additional greenhouse gas emissions due to natural feedbacks omitted from ESMs". They also wrote, "Including interactive wetland methane emissions influenced TCRE solely through the radiative impact on climate." For clarity, I would write, "ESF accounts for additional CO₂ emissions..." if non-CO₂ is not at all included in ESF.

Thank you, we've changed the manuscript accordingly: "Note that ESF accounts for additional CO₂ emissions due to natural feedbacks omitted from ESMs - primarily permafrost thaw, though wetland methane emissions are sometimes nominally included in such terms as a CO₂ equivalent⁵."

Furthermore, I would maintain that the treatment of wetland CH₄ emissions represents an inconsistency in the TCRE calculation method. The radiative effect of wetland CH₄ emissions is considered within TCRE, whereas non-CO₂ effects are generally not included in TCRE quantification. The ESF term does not account for wetland CH₄ emissions, despite claiming to account for "future emissions arising from Earth system feedbacks not accounted for in the derivation of TCRE, such as carbon release from permafrost thaw". While I don't consider this a significant issue, the authors could comment on it.

Thank you, this is a point worth clarifying, so we have added the paragraph below to the Methods section where we describe the 1pctCO₂ experimental protocol.

The important point is that the external forcing applied, i.e. what drives our 1pctCO₂ experiments, is entirely CO₂-only. All other climate drivers – methane emissions, land use, aerosols, and all greenhouse gases except CO₂ are held at the 1850 level: there are no anthropogenic emissions / concentration changes in CH₄ or other non-CO₂ forcing. Methane emissions from wetland, though, do change as a natural feedback in response to climate, and so the methane concentration does deviate from pre-industrial levels in *ukesm-wch4*. This has an additional radiative impact, but it is a response and not an imposed forcing. Essentially, it is this deviation – how much additional methane is added to the atmosphere under 1% CO₂ forcing when the model has the additional degree of freedom afforded by the emission-driven CH₄ configuration, and the impact this has on the TCRE – that we are measuring with *ukesm-wch4*.

New paragraph in Methods:

"In the case of *ukesm-wch4*, the concentration of methane deviates from the pre-industrial level due to the additional coupling of wetland emissions with atmospheric methane¹⁰. Similarly, the aerosol concentration in *ukesm-fire* and *ukesm-bvoc* exceeds pre-industrial levels due to enhanced release of aerosols or aerosol precursors as CO₂ increases¹⁰. In each case the deviation from pre-industrial conditions is a consequence of the additional process representation responding to the standard 1pctCO₂ forcing in which only CO₂ varies, rather than a deviation from the 1pctCO₂ forcing protocol that facilitates calculation of TCRE."