

Public justification (visible to the public if the article is accepted and published):

I thank the authors for their detailed responses to the community and referee comments. The revisions made to the manuscript are generally adequate. However, a few issues remain that should be addressed before the manuscript can be considered for publication. Most importantly, the manuscript does not sufficiently engage with or cite relevant literature on the central topic of the paper, namely overshoot and reversibility. There is a rich body of literature on this topic that is not cited. Even if most of those studies are based on CO₂-concentration-driven simulations (e.g. CDRMIP reversibility or SSP5-3.4OS experiments), they are still relevant in this context as reversibility of changes in physical climate variables (AMOC, SLR, precipitation) are largely driven by changes in radiative forcing or temperature.

Thanks Kirsten for your support and thorough read through of the paper. I trust that the updates are sufficient for publication in ESD.

- p. 3 l. 81 “idealized CO₂-only trajectories”: Cite relevant studies (e.g. Tokarska et al., 2015; Schwinger et al., 2022b)

Now cited.

- p. 4, l. 99: A large part of the introduction reviews literature on scenarios and motivates the need for emissions-driven simulations but does not introduce the research questions of the study. To address this, I suggest to move the research questions to the introduction and include a paragraph that reviews the literature on overshoot and reversibility relevant to this study. The review does not need to be exhaustive but should go beyond the few studies (Pfleiderer et al. 2024; Roldán- Gómez et al. 2025) currently cited. For example, Koven et al., 2022 explore the multi-century response of a range of variables (GMSAT, AMOC, carbon cycle) in SSP5-3.4OS.

Research questions have been moved to the introduction and made clearer with additional supporting text moved to the following paragraph. The Koven paper has been cited here too.

- L. 135: Acronym used for surface air temperature in text (GSAT) is inconsistent with acronym in figures (TSA). Please ensure consistency.

Now fixed in figures 3, 8, 9 and 10.

- L. 140: A potential reason the two models exhibit different warming rates in the positive emission phase could be different climate sensitivities. It is mentioned later in the discussion that the models have similar climate equilibrium sensitivities but it would be helpful to have this information earlier.

A short discussion of the ECS and TCR in the two ESMs is added here.

- L. 150: “in this model”. Clarify what model you are referring to.

NorESM2-LM; now clarified

References not already cited

Koven, C., V.K. Arora, P. Cadule, RA. Fisher, C. D. Jones, D. M. Lawrence, J. Lewis, K. Lindsey, S. Mathesius, M. Meinshausen, M. Mills, Z. Nicholls, B.M. Sanderson, N.C. Swart, W.R. Wieder, and K. Zickfeld, Multi-century dynamics of the climate and carbon cycle under both high and net negative emissions scenarios, *Earth Syst. Dyn.*, 13, 885-909, <https://doi.org/10.5194/esd-13-885-2022>.

Tokarska, K.B., and Zickfeld, K., 2015, The effectiveness of net negative carbon dioxide emissions in reversing anthropogenic climate change, *Environmental Research Letters*, 10, 094013, <https://doi.org/10.1088/1748-9326/10/9/094013>.