Authors' response to the review by Kirsten Zickfeld

We would like to thank Kirsten Zickfeld for the positive evaluation of our manuscript and for the constructive and helpful criticism. We have revised our manuscript and addressed all points that have been raised as detailed in our point-by-point response below (original comments in grey, italic font). Proposed verbatim changes or additions to the manuscript are highlighted in this colour.

Schwinger et al investigate the reversibility of the Earth system for a range of idealized scenarios that differ with regard to the amount and duration of overshoot. They consider a change to be reversible if the modelled ensemble mean of an overshoot simulation returns to the reference simulation within the range of internal variability. Consistently with earlier studies they find that most Earth system changes are reversible, except for aspects with longer response timescales, such as permafrost carbon and seawater properties of the deep ocean. They do not identify tipping points in their simulations following overshoot.

The manuscript by Schwinger et al. is a valuable contribution to a relatively small body of literature that investigates Earths system reversibility in emissions-driven simulations, allowing for feedback between climate and the carbon cycle. The manuscript is well written, the methodology is adequate and sufficiently documented, and the conclusions are supported by the findings. There are a few instances where the manuscript would benefit from additional explanations, or definition of terminology for an interdisciplinary readership.

Thank you for this positive evaluation.

Specific comments:

l.185: Comparing reversibility for a given year has the disadvantage that it does not allow for a clean separation of the effect of overshoot duration, as differences could also be due to shorter time left to adjust to the final forcing level in the longer vs. shorter overshoot simulations.

Yes, this is true. This point is most relevant for the oceanic zonal mean sections shown in Fig. 10. We have added a figure to the Appendix (Fig. A4) showing the reversibility 95 years after the negative emissions cease (REV_{95}) for the short overshoot. This allows a clean comparison for the effect of overshoot duration when compared to Fig. 10. We have added a few sentences discussing this new figure at the end of section 3.7 as follows: "So far, we have chosen to assess reversibility for the same simulation year, that is, we compare REV_{95} for the long overshoot simulations (95 years after negative emissions ceased, Fig. 10) to REV_{195} for the short overshoot simulations (195 years after negative emissions cease, Figs. A2 and A3). To allow for a clean comparison of the effect of overshoot length on reversibility, Fig. A4 shows REV_{95} for the short overshoots (i.e., reversibility derived from simulation years 290-300). Comparing Fig. A4 and Fig. 10 reveals that the volume of water masses showing irreversible change tends to be larger in the long overshoots, although the spatial patterns are broadly similar. The main difference between the short and long overshoots is that the irreversible changes in the deeper ocean are much more pronounced for the long overshoot duration, indicating a clear benefit of limiting the duration of an overshoot."

l. 265: It could be pointed out that atmospheric CO2 in the overshoot simulations temporarily "undershoots" CO2 levels in the reference simulation.

We followed this suggestion by adding the sentence *"We note that during the negative emission phases, atmospheric CO_2 concentrations decrease below the concentration in the reference simulation in all overshoots. This effect is largest (about 50 ppm) for the large overshoots, which have the fastest rate of CDR."*

l. 283: Before discussing fractional quantities in Fig. 4 I suggest to discuss the cumulative fluxes (Fig. 3), which don't have the denominator changing at the same time and are therefore more intuitive. Here the reason for the decline in the cumulative fluxes during the negative emissions phase could be explained (e.g. the reversal in pCO2 gradient mentioned in l. 387-388.).

We followed this suggestion and reworded/extended the beginning of this paragraph: "During phases with negative emissions in the overshoot simulations, both land and ocean become a source of CO_2 such that their carbon stocks are reduced (Fig. 3c,d). For the ocean this happens as soon as the CO_2 partial pressure difference between atmosphere and ocean becomes negative. For the land, a reduced CO_2 fertilization effect shifts the overall balance between carbon uptake through net primary production and carbon release through heterotrophic respiration towards the latter. However, since these processes are slow and lag the reduction of the cumulative total of emissions, there is a rapid increase of CF_O and CF_L (Fig. 4b,c)."

l. 350: Section 3.5: I suggest to define and explain the meaning of the biogeochemical quantities discussed in this section (preformed vs. remineralized carbon, AOU etc.) to make sure the findings are accessible to an interdisciplinary readership.

We have added text explaining these biogeochemical quantities as follows: "The main carbon reservoirs considered here are vegetation carbon, permafrost carbon, and non-permafrost soil carbon for land, as well as remineralised and preformed dissolved inorganic carbon (DIC) for the ocean. Changes in permafrost carbon are obtained as cumulative carbon fluxes summed over all permafrost grid cells, i.e. those grid cells that are defined as permafrost in the pre-industrial control simulation. Preformed DIC originates from atmospheric CO_2 dissolves in the surface ocean and is transported into the interior by ocean circulation. In contrast, remineralized DIC has been transported into the interior ocean through the biological carbon pump: Biological uptake by planktonic organisms near the ocean surface, sinking to depth as particulate organic matter, and subsequent remineralization by bacterial activity. We note that the remineralization of organic carbon consumes oxygen (if present in sufficient quantity), such that oxic remineralization can be measured by apparent oxygen utilization (AOU), defined as the oxygen deficit in a water parcel relative to its saturated oxygen content."

l. 377: Mention that inclusion of vegetation dynamics could affect reversibility.

This has been added as follows: "We note that the inclusion of vegetation dynamics in our model would most likely affect these results, since changes in biogeography would lead to larger changes in land carbon pools and larger time lags between drivers and response."

l. 387-388: This is the first time this is mentioned. I suggest to discuss this earlier (e.g. in section 3.3.).

This has been done. See our responses to the 2nd and 3rd specific comment above.

l. 390: Irreversibility of thermosteric sea level rise was also investigated in Ehlert & Zickfeld, 2018, https://doi.org/10.5194/esd-9-197-2018.

We have included this paper in the list of cited works.

l. 395 – 398: I don't follow this argument. Perhaps the amount of sea-level rise corresponding to a 1.5C global warming limit was "implicitly accepted", but not the additional sea level rise resulting from overshoot of the warming limit?

We agree that this was not very well explained. Our argument is that the additional sea level rise due to an overshoot remains relatively small compared to the amount of sea level rise committed to in the reference simulation (< 20% except for the most extreme overshoot). We make this clear by rewording: "The additional steric sea level rise due to an overshoot remains relatively small compared to the sea level rise committed to in the reference simulation in our model (<20% at year 400 except for the most extreme overshoot OS_{100}^{1000}). Thus, the rate of sea level rise determines the pace and cost of necessary adaptation for the decades to centuries after an overshoot. Therefore, limiting the rate of sea level rise after an overshoot might arguably be more policy relevant in the context of negative emissions than a relatively limited contribution of the overshoot to the sea level rise itself. "

Figure presentation: Vertical lines showing the positive emissions, negative emissions and zero emissions phases could be included. Also, it would be helpful to have the legend repeated in figures where changes are non-monotonic as a function of overshoot size or duration (e.g. Figs. 5 and 6).

We have added shadings in the background of each time-series figure indicating the different phases of our simulations. We have also added legends to figures 5 and 6 as suggested.