Reply to Norman Steinert

General comments

This manuscript analyzes a new set of climate change experiments utilizing the ACCESS-ESM-1.5 Earth system model to simulate climate evolution under net-zero emissions. Seven simulations spanning 1000 years were branched of the SSP5-8.5 simulation at 5-year intervals starting in 2030 to assess climate stabilization in line with the Paris Agreement and higher warming scenarios. Significant findings are provided for continued Southern Ocean warming, reversal of mid-latitude drying trends, and varying sea ice extent trajectories in the Arctic and Antarctic. The study also examines the El Niño-Southern Oscillation (ENSO), noting reduced amplitude and frequency of events under climate stabilization. The results show that regional changes persist for centuries post-emission cessation, indicating long-term climate impacts despite net-zero emissions. The results are presented by temporal evolution and by global warming levels, where climatic changes in early and late periods after emission cessation are compared against transient warming under SSP5-8.5. The manuscript provides relevant and novel results such as insights into the climatic effects of delayed emission cessation, regional timing of peak warming after emission cessation, and temperature and precipitation trend differences between transient and stabilized global warming levels.

I do have a few questions and comments, which I consider overall minor, that I would like the authors to address.

Thank you for your constructive feedback. We believe we can address all your comments as we outline below.

The manuscript is well written, structured and generally reads fluently. While the manuscript is quite long, the flow of the text does help its understanding. Specific suggestion for sections to be adjusted are listed below.

This study provides a unique dataset as the stabilization simulations are run in emission-driven mode and because they are provided for an entire millennium, which is specifically relevant for 'slow' climate subsystems such as the ocean or cryosphere elements. The authors highlight an important point in line 96, that is relevant for the consideration of climate change experiments for policy under emission budget frameworks. Hence, the experimental design provided here should be inspiration for other modeling groups to at least replicate a sub-ensemble of the experiments described here for a better inter-model comparison.

Thanks. We also hope that the experimental design is of use to others.

Given that these simulations are emission driven, I was a bit surprised that no carbon cycle response was analyzed. I understand that ocean and land carbon cycle

responses will affect atmospheric CO_2 concentration (which is provided) and therefore its effect on temperature (and precipitation) is shown – ignoring physical feedbacks here, but given the length of these simulations, there would be an opportunity to analyze long-term carbon cycle responses under emission cessation. Since the paper is already long in itself by focusing on temperature and precipitation, I would not expect any additional analysis here, as the authors also refer to further analysis in the last sentence of Section 4. Nonetheless, I would appreciate one or two sentences discussing carbon cycle implications in these emission-driven simulations in the Conclusions Section.

There is a lot of scope for examining carbon cycle responses in these simulations and there is a follow-up study underway. To do the topic justice and adequately explain the changes we identify we decided a whole other study was needed. We can certainly add a brief discussion of carbon cycle implications of net-zero emissions though.

Further do I miss a bit a more refined discussion of the results in the frame of other ESM responses (such as under ZECMIP). While the authors make clear that the results can only be interpreted under the specific climate response of ACCESS-ESM-1.5, its differences to other models, or e.g., the ZECMIP multi-model mean, can affect the interpretation of the results presented here. I suggest to either add a few comments in appropriate places in the main text or discuss this issue a bit more in the Conclusions Section.

We agree that contextualising our findings with reference to multi-model experiments like ZECMIP is useful. We will add additional text to do this.

An overall limitation of this study is that a single climate model is used, which is very relevant for and interpretation of the result provided here. This is acknowledged by the authors in various places of the main text. However, I think this is a key part to keep in mind when assessing the results of this study. I made specific comments below where I think the manuscript could benefit from some more detail in the framing of the results in the context of other climate models.

We agree and will make changes in response to your comments below.

Specific comments

From the abstract, I find it is not entire clear whether the focus of the analysis is on the comparison of stabilization vs. transient climate responses, or long-term stabilization at different warming levels only. The reader might appreciate if this can be better clarified.

There are elements of both. We will clarify the purpose of the study in the revision.

Line 25 – ‘mid-latitude drying trends’ – it might be useful for the reader to know here that this refers to extremes provided in the analysis.

Yes, on reflection this is quite a vague statement that isn’t very helpful so we will edit this and perhaps remove this from the Abstract.

Line 42 – ‘near-linear relationship between cumulative carbon dioxide emissions’ – there are indications that this might only be true for transient positive emissions.

Indeed. We can add a clarification of this to make it clear we mean under continued global warming only.

Line 46 – ‘near-zero global mean temperature change for the following century’ – that’s on a multi-model mean, but individual models can show quite large positive or negative ZEC after emissions cessation.

Indeed. We can add a point emphasising model uncertainty on this statement.

Figure 1 – I find it not quite intuitive to present results in panels c) and specifically d) together with the conceptual panels of the experiments in a) and b). I would further like to see a panel showing cumulative emissions in these experiments. I therefore suggest moving panel d) to Figure 3 as the first panel, and exchange it with a panel for cumulative emissions in Figure 1. Also, the panels c) and d) have no x-axis labels, while panels a) and b) do.

We agree that there are different ways of presenting this information and we think there are merits in different options. We chose figures c and d as approximate parallels to a and b. Admittedly, a and b are drawn to represent a multi-model ensemble and we then plot only data for the ACCESS model in c and d. If we were to plot a cumulative emissions as a new Figure 1d then this would be quite different to Figure 1b. We propose that we largely retain Figure 1 in its current form but add a new Appendix figure with the cumulative emissions graph. We will add axes labels to c and d.

Lines 115-116 – ‘… with the starting point chosen…’ – As I understand it, I am not convinced by the transferability of this approach to other ESMs because: 1) each ESM will have different long-term climate response – positive or negative ZEC, which vastly affects the branching points from the transient warming experiment, and thus may have different policy implications for when net-zero need to be reached, and 2) reaching temperature targets under non-stable climate response under zero emissions will strongly depend on the intended simulation length, as there is no stabilization (as illustrated in Fig. 1d).

Indeed, we agree that models would exhibit different responses to net-zero emissions in terms of GMST and other changes. In King et al. (2021) we suggest that prior knowledge of the model could be used to choose different starting points to those that we have

applied here for the ACCESS model. We can add some text to clarify this point and make the caveats that you raise.

Line 150 – '… representation of different ENSO flavors appears to be worse in ACCESS-ESM-1.5 than in other CMIP6 models' – given that a part of the analysis focuses on the ENSO response, I miss some discussion of this point. How does it differ from other ESM? This part could be provided at this point of the main text. And in which way does this affect the results later discussed in the main Results Section and provided in Figure 15? This part could be incorporated in Section 3.4 line 495, where this is already referred to by the authors.

We agree, some additional discussion would be useful as justification for our analysis. ACCESS-ESM-1.5 has some biases in simulating the spatial pattern of Central and East Pacific ENSO events and lacks seasonality in SST variability in the west-central Pacific seen in observations and some other models. We propose to add some discussion of this point to section 2.3.3 which explains in more detail this model deficiency and explains how we design our analysis in the context of known model biases.

Line 159 – '… mean surface temperature (GMST) slightly decreases in the first 20-50 years…' – could the authors please add an explanation as to why there is this initial drop?

Yes, we can add a brief discussion. We believe that the initial decline in temperature is due to the choices we made when transitioning from the SSP5-8.5 run to the stabilised period. At the branching point we set all non-CO2 GHGs back to their pre-industrial values (including aerosols). Some initial analysis has shown that the change in CH4 concentrations back to 1850 levels dominates the initial temperature response, resulting in the observed decrease over the first 50 years. We will include this short explanation in the revised version.

Line 160 'reduced non-CO2 and aerosol concentrations has taken place…' – The combination of these sentences is slightly confusing as non-CO2 and aerosols were set to 1850 levels as indicated in line 154. Please adjust for clarity.

Indeed. We will edit this for clarity as we agree this is not currently well worded.

Lines 157-168 – These lines describe results provided in Figure 1 (see my suggestion on Figure 1 above). I think this paragraph is better placed in the Results Section, e.g. line 281, which would go along nicely with moving Figure 1 panel d) to Figure 3. Please adjust for clarity of the Methods Section.

In a sense, we are showing initial results here so we agree that this could be moved to the start of the Results section. We propose to move some of the text to the Results section and keep only a brief description of the GMST results here because we would like to

retain Figure 2 as a part of the Methods and that requires at least a brief description of GMST changes before we can talk about sampling of GWLs.

Line 172 – 'quasi-stabilized' – I find this a bit misleading, as emissions are entirely stabilized, but temperatures are clearly not. Perhaps remove.

Yes, this maybe isn't the best descriptor. We will edit this.

Section 2.3.1 – This entire section is a listing of analyses that have been performed. This is a bit a question of taste, but I would suggest removing this Section entirely. The reader might have difficulties remembering methodological details when reading about the results later. Rather I would like to see the relevant bits incorporated into the results section where appropriate, which would make following and interpreting the results much easier.

A similar comment is made in the other review and we agree this is a lengthy section that is perhaps hard to read without also seeing results. We propose to keep a short summary and move some of this text to the Results section accordingly.

Line 189 – '… with the first period…' – disconnected sentence that sound strange, please revise.

We agree and will make a suitable edit.

Line 193 – given that there is often a focus on reaching 1.5°C, the reader might appreciate an extra Figure A3 that is as Figure A2 but keeping NZ2030 as reference (i.e., NZ2035-NZ2030, NZ2040-NZ2030, …).

Yes, we agree. We can include this figure the revised version of the paper.

Figure 2 – the definition of 'early stable' and 'late stable' should be indicated more clearly to be based on *stable emissions,* as temperature are clearly not stable – both in the Figure caption and in the text between lines 220-223.

Yes, the wording choice here has been challenging, but we agree and will make suitable edits.

Figure 2 – The reader might appreciate horizontal reference lines for the three warming levels of 1.5, 2 and 3 °C. Also, to me these periods highlighted in bold do not average out at these warming levels, which I would explain by the fact that a range of +/-0.2°C was taken for the definition of the target GWL in these timeseries.

We agree that this would be useful to include and will add these lines in the revised version as well as a brief explanation of the +/-0.2°C window causing slight differences in averages.

Figure 2 – If the intention of defining warming levels from the time series was to collect as many years as possible from the available simulation that comply with the +/-0.2°C range around the target GWL, I don't quite understand why simulated period are only chosen from specific simulations. There are clearly periods from other simulations, that reach into this range, but the specific years/decades were not considered here apparently. For example, the simulation branched at 2040 (third grade of orange) drops into the range of temperatures in the bold period of the simulation branched at 2035 (second grade of orange), but the respective years were not considered in the analysis. Such instances can be found for almost all periods highlighted in bold. Could the authors please clarify why this approach has been taken. If it was simply for simplicity of selecting the timeseries for the analysis, this should be made transparent in the main text.

Yes, we agree that this should be clear. There would be validity to different approaches, but we thought this was a simple way to take advantage of the fact that these runs roughly span the target GWL in the desired early and late time periods. Use of additional runs could inadvertently weight the samples more towards the start or end of the 200-year windows, e.g. using the NZ2035 run to sample the late 1.5°C warming period would result in more weighting nearer 800 than 1000 years post net-zero. This was a minor reason for this choice. We will add some further explanation in the text.

Lines 224-226 – '..., but results may... GWL ensembles.' – I think this a key part in the assessment and interpretation of the results given the methodology used here. The reader might benefit from the authors providing more detail (or some speculation) on how the described differences may look like and how relevant they are with respect to the results presented here.

Yes, we agree. We can add some speculative commentary on this point noting this is hard to say confidently what the differences may be.

Line 227 – 'The method...' – I suggest moving this sentence to line 216.

Yes, we can make this change.

Line 228 – 'The use of...' – I suggest moving this sentence to line 218.

Yes, we can make this change.

Figure 3 – apart from adding the global mean temperature pane as indicated in the comments above, I would suggest plotting panel a as the land-ocean ratio instead of anomalies. Further, I would plot panels c) and d) as anomalies.

Thanks for this suggestion. For Figure 3a, the difference rather than ratio is plotted because the ratio is very unstable right at the beginning due to small ocean changes and the fact that 1850-1900 is the baseline and is itself included in the plot. We should

explain this choice more because we know others often use the ratio rather than difference. For c) and d) we can make these anomalies in the revised version.

Line 284 – ', and increase…' – worth noting that is seems to even stabilize at a non-zero difference.

Yes, we agree and will make a suitable edit.

Lines 222-223 – '…, suggest that there are complex changes occurring through the next-zero simulations…' – In line with the interpretation provided in lines 341-342, could the authors please comment on – and perhaps incorporate in the main text – whether this is purely because of the regional temperature changes shown in Fig. 5 (i.e. relatively stable Arctic Amplification, and much accelerated Antarctic regional warming due to slow Southern Ocean response), or whether there is indication and/or contribution of other irreversible processes such as non-linear ice sheet responses and feedbacks. Comments in that direction are made in the Conclusions Section, but it might be worth to also discuss this here.

We can add a comment on this point, although it would be a little speculative too. We would also make the point that more focussed analysis of sea ice changes is needed beyond the overview analysis performed for this paper.

Lines 471-472 – This sentence is a repetition of lines 462-464.

Yes, we can rewrite this to make it less repetitive.

Line 495 – see comment for line 150. It might be good to add some discussion on how ACCESS-ESM-1.5 differs to other models in the context of the ENSO response and how relevant this is for the analysis here.

Indeed, we agree this is important to include and can briefly comment on this here too.

Technical corrections

We will make technical corrections to make Figures easier to read and fix and minor typos.

Axes and/or legend labels are too small in Figures 1, 5, 7, 8, 9, 10, 11, 13, 14, 15. Figure 3 has no x-axis labels, but I assume they would appear too small if they were there, too.

Line 160 – subscript '2' in 'CO2'.

Line 189 – '… temperature difference values …' – add 'between the two periods'.

Line 199 – 'The changing pattern' – remove 'changing'.

Line 305 – '… in the Arctic, also continues…' – remove 'also'.

Line 361 and Line 362 – 'Figure 4X-X' should be 'Figure 7X-X'.

Line 465 – 'G)' should be 'g)'.

Line 503 – 'are not projected' – please check the presence of 'not' in this sentence. For how I understand this sentence, it should be removed.

Other

1. Does the paper address relevant scientific questions within the scope of ESD?

   Yes
2. Does the paper present novel concepts, ideas, tools, or data?

   Yes
3. Are substantial conclusions reached?

   Yes
4. Are the scientific methods and assumptions valid and clearly outlined?

   Yes
5. Are the results sufficient to support the interpretations and conclusions?

   Yes
6. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)?

   Yes
7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution?

   Yes
8. Does the title clearly reflect the contents of the paper?

   Yes
9. Does the abstract provide a concise and complete summary?

   Yes

10. Is the overall presentation well structured and clear?

    Yes, see some specific recommendations for the Methods Section

11. Is the language fluent and precise?

    Yes

12. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?

    Yes

13. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?

    Yes, see some specific recommendations for the Methods Section

14. Are the number and quality of references appropriate?

    Yes

15. Is the amount and quality of supplementary material appropriate?

    Yes, with the suggestion to add one panel