Response to the editor and reviewers

We thank the editor and the three reviewers for the critical assessment of our work and their very helpful and constructive comments. We have addressed all comments point by point and revised our manuscript accordingly.

Reviewer 3

Comments on "AERA-MIP: Emission pathways, remaining budgets and carbon cycle dynamics compatible with 1.5oC and 2oC global warming stabilization" by Silvy et al.

Overall, I think this is a really nice paper and an important contribution that shows a new way of using ESMs to understand remaining emissions budgets at different reference warming levels. The authors have done a great job of describing a complex set of experimental results. I do have a couple concerns that I describe below, but overall I think it is a very worthwhile publication.

We thank the reviewer for the positive assessment of our manuscript.

My first concern is semantic but I think it is important. I object to the usage of the word "target" in the sense of "global warming target" starting in the very first sentence of manuscript, and its many instances thereafter. The Paris agreement (English-language version) uses the word "target" only in article 4 to describe emissions reduction targets, it never refers to warming levels as "targets". Nor do any of the IPCC AR6 SPMs. So the first sentence is not correct: the Paris agreement does not primarily focus on global warming "targets". And it is a mistake to think of warming levels like 1.5C or 2C as targets; just because we do not wish to exceed them does not mean that these warming levels are where we wish the climate to arrive at or remain, either as a peak temperature or after a peak-and-decline. So in this paper, there are basically two distinct usages of the word "target": in the control-theory sense of "setpoint", and in the policy sense of reference warming level, and I advocate that the authors do not conflate them. I.e., there is a 1.5 degree global warming level as specified in the Paris agreement, and because one is interested in both the impacts at that or another policy-relevant global warming level, as well as the emissions pathway that is compatible with that warming level and how it might evolve over time, one uses the AERA-MIP protocol with a setpoint at that level to drive the simulations to that setpoint. I think that it is very important *not* to use the word "target" for the setpoint value, because doing so implies a value judgement that the setpoint warming level is in fact a desired goal. So I would strongly encourage the authors to replace every single instance in the entire manuscript of the word "target" for either of the more neutral words "level" or "setpoint", or other similar wording where appropriate, throughout (including when used within figures such as figure 1).

We thank the reviewer for highlighting this important nuance. In the revised version of the manuscript, including Figure 1, we have replaced the word '*target*' with more neutral terms such as global warming '*level*'.

My second major concern is about how this manuscript treats ACCESS as an outlier. The overall hypothesis of the manuscript seems to be that it is possible to apply the AERA in a set of full-complexity ESMs, but then the one model that does not work as intended is not fully included in the analysis. The implications of this would seem to be pretty important, as it implies that certain types of physical uncertainties might lead to a failure of climate policies that are structured like the AERA approach. The explanation on lines 132-136 makes sense, but what is the evidence for it? How confident are we that this mismatch in non-CO2 forcing is not pointing to a real uncertainty that we need to consider in remaining emissions budgets? So I would advocate keeping ACCESS within the ensemble statistics unless a stronger case is made that the non-CO2 forcing estimate difference represents an unphysical model artifact.

We note that the ACCESS model also converges to both prescribed temperature levels, but substantially later than all other models. Therefore, we do not fully agree that, according to ACCESS, this would indicate a failure of climate policies. In AERA, an existing mismatch between actual non-CO2 emissions and those supplied to the algorithm is not corrected for. As a policy instrument, scientists would revise initially biased non-CO2 emissions estimates over time, allowing temperatures to converge faster.

We have, however, now included the ACCESS model in a new figure in the Appendix (Figure A3), which depicts the multi-model mean CO2-fe emissions pathways with and without ACCESS. For example in the 1.5ºC scenario, the multi-model mean CO2-fe emissions in year 2050 are 0.03 PgC yr^{-1} with ACCESS and 0.16 PgC yr^{-1} without ACCESS. Between 2100 and 2150, the average CO2-fe emissions are 0.79 PgC yr^{-1} with ACCESS and 1.66 PgC yr^{-1} without ACCESS. Thus, the multi-model mean emissions pathways are nearly identical. We have added the following to the text: "*The inclusion of the ACCESS model in the ensemble statistics may alter our quantitative results, but not qualitatively*".

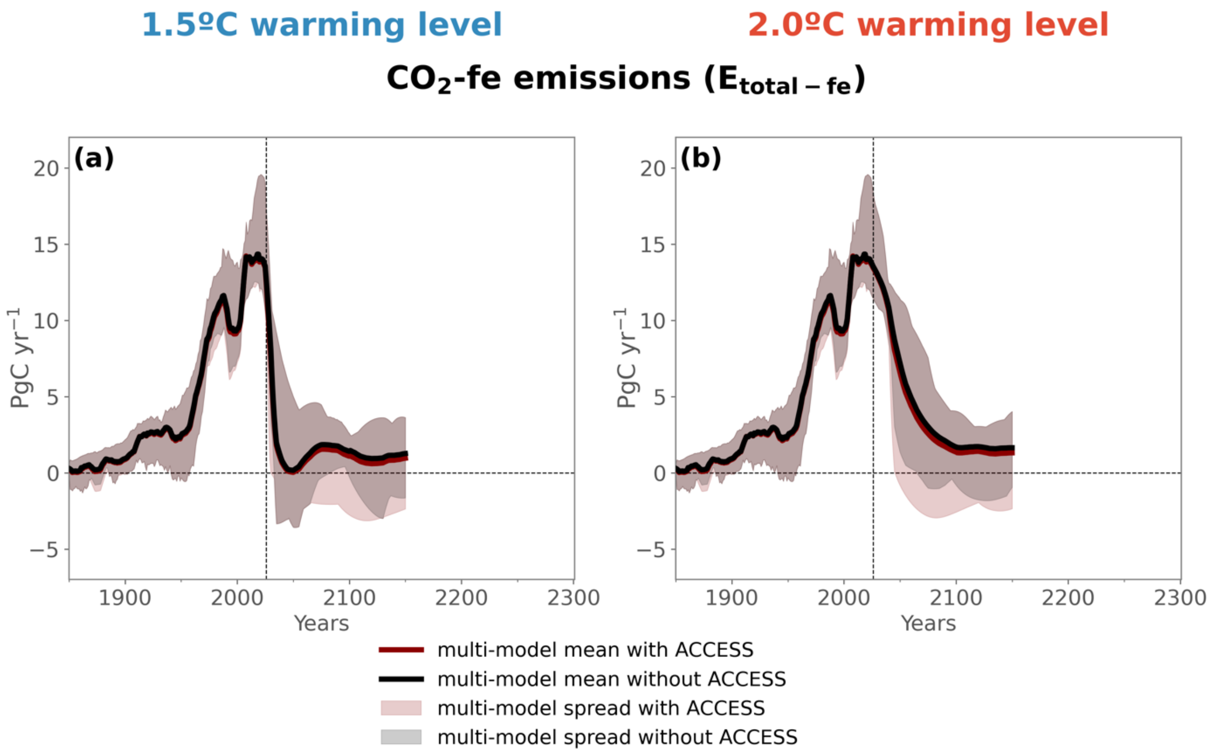

Figure A3: Simulated multi-model mean CO_2-fe emissions for the 1.5°C and 2.0°C warming levels with and without including the ACCESS model.

I am also a bit confused about the implications of removing the historical warming differences between the models by lining them up so that they all pass through 1.22 degrees at year 2020. How was this handled for the models with multiple ensemble members -- were the ensemble members each lined up at their own year 2020 temperature, and thus they have different absolute temperature setpoints? If so, is there any correlation between the REBs and the absolute temperature at 2020 across the ensemble members? Or was the ensemble-mean global temperature used for the year 2020 value so that all members have the same absolute temperature setpoint? Given that there is real uncertainty in the current level of global warming due to internal variability (e.g. as pointed out on line 218), what are the implications of this choice on the uncertainty in REBs? Some further discussion of this would be helpful.

Yes, each ensemble member was aligned to its own anthropogenic temperature in 2020, resulting in slightly different absolute temperatures for that year. However, the maximum difference between ensemble members in 2020 is very small. For instance, in the 5-member GFDL ensemble simulation, this difference is only 0.067°C. We have added to the text: "*Each ensemble member has its own anthropogenic temperature in year 2020, resulting in very small differences in absolute temperatures for that year (maximum differences across GFDL ESM2M ensemble members of 0.067°C).*"

Regarding the comment about the REB, by definition:

$$REB(t_{st}) = \frac{remaining\ warming\ (t_{st})}{TCRE - fe}$$

with t_{st} the year of the current stocktake, for example 2025 for the first stocktake. The remaining warming at each stocktake is model and simulation dependent. It is defined by the difference between the absolute temperature target and the absolute temperature at the current stocktake: $T_{target} - T(t_{st})$, with T the temperature fit (extended 31-year running mean). In the "relative warming" case presented in this study, the temperature target is defined by: $T_{target} = T(2020) + observed\ remaining\ warming$. The observed remaining warming is 0.28ºC for the 1.5ºC scenario and 0.78ºC for the 2ºC scenario (see Methods). Thus, for example for the 1.5ºC scenario, the remaining warming is equal to $T(2020) - T(t_{st}) + 0.28$.

On the other hand, the TCRE-fe is defined by the ratio of the temperature anomaly relative to preindustrial over the cumulative emissions: $TCRE - fe = \frac{T(t_{st}) - T_{piC}}{\sum_{1850}^{t_{st}} E_{total-fe}}$.

Which finally gives:

$$REB(t_{st}) = (\,T(2020) - T(t_{st}) + 0.28)\frac{\sum_{1850}^{t_{st}} E_{total-fe}}{T(t_{st}) - T_{piC}}$$

Thus, while the REB depends on the absolute temperature in 2020, it also depends on the absolute temperature of the current stocktake, which means we do not find a clear linear relationship between the two across the GFDL ensemble members (see Figure R2).

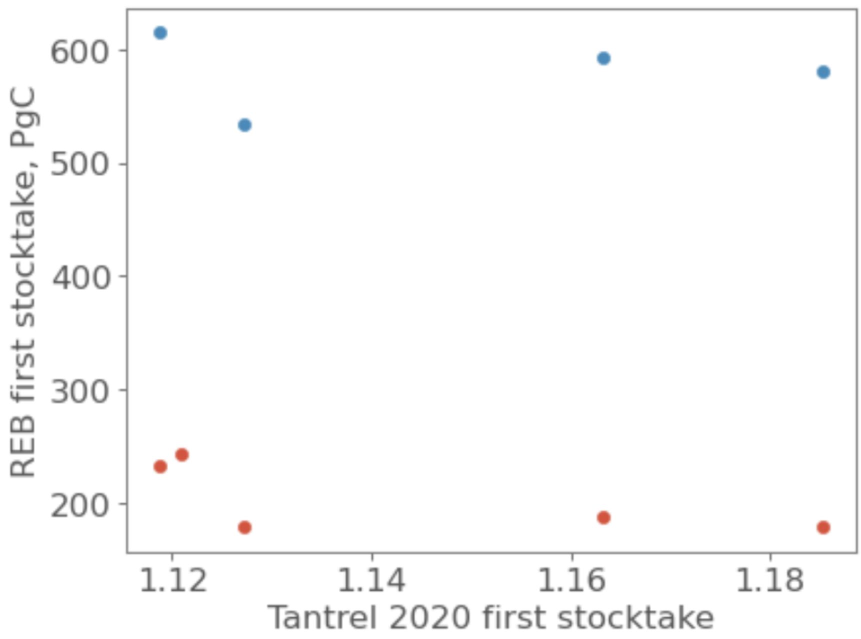

Figure R2: Scatterplot between simulated anthropogenic temperature in year 2020 and the remaining emissions budget at the first stocktake for the individual ensemble members for both the 1.5°C warming scenario (red dots) and the 2.0°C warming scenario (blue dots). The 2°C scenario consists only four members in this figure, as the metadata from the AERA algorithm was missing for one member.

Lastly, I recommend revising the schematic in figure 1 bottom-right panel to show that the minimum in temperature spread is at present-day, and then increases prior to that, rather than what is shown where the temperature spread increases over time starting from preindustrial conditions.

We decided not to change Figure 1. Its purpose is to highlight the difference between the traditional IPCC forward approach (prescribed emission pathway → temperature response) and the new AERA approach (prescribed temperature level → adaptive emissions pathway). Introducing the concept of the relative temperature target would confuse the general reader.

Why is the time evolution of the ensemble spread of CO2-fe in figure 3b and 3c so different between GFDL-ESM2M and EC-Earth?

The number of ensemble members is very small for both GFDL ESM2M (5 ensemble members) and EC-Earth (3 ensemble members). Therefore, the magnitude of the ensemble spread and its temporal evolution should be interpreted with caution. For example, we cannot determine with confidence if the ensemble spread changes in magnitude over time.

Line 172-175: The E_LUC won't be quite the same as in the references, since the temperature and CO2 pathways will differ, and thus the loss of additional sink capacity and weakening of carbon-climate feedbacks will be different. Can you quantify that effect, and does it matter?

Correct. But we cannot quantify this effect. The difference may not be that large because all models converge relatively rapidly despite this mismatch between the real E_LUC and the E_LUC prescribed to the AERA. No changes are made to the text.

Lines 216-217. Is the +/- 10% correspond to the light red and blue shading in fig. 1? Or is it the +/- 0.2C referred to on line 217? Either way, please note that in the figure caption.

We cannot find the +/-10% statement the reviewer is referring to. In addition, we assume that the reviewer refers to Fig. 2 and not Fig. 1, as blue and red shading is only shown in Fig. 2. We have added to the caption of Figure 2: "The horizontal shading in (a) and (b) indicates the uncertainty with which anthropogenic warming can be determined (± 0.2°C)"

In fig. B3, why is the difference between the AERA-derived and "true" estimate of TCRE so systematic over the early years of the experiment? I.e. the multi-model ensemble looks roughly to have about the same difference as individual models. Is that due to some systematic path dependence in aerosol, land-use, and other SLCF dynamics?

This issue is related to the temperature fit and not due to aerosols or other factors. We explained this in the revised Figure caption: "*Solid lines show TCRE-fe calculated by AERA at each stockake year. Dotted lines indicate the TCRE-fe calculated a posteriori with the final ("true") estimate of the 31-year running mean timeseries of GSAT. The deviation between the solid and dotted lines occurs because the temperature fit becomes less accurate when the temperature increase slows toward stabilization. As the warming rate decreases, the 31-year mean will also decrease. Therefore the "true" estimate of the 31-year mean will be smaller than the estimate at each stocktake. Consequently, the TCRE will be smaller than estimated at the stocktake during periods of slowing down temperature increases. In addition, the reconstructed TCRE-fe (dotted lines) is less variable than the TCRE-fe calculated by AERA, which calculates ΔT based on the last year of the extended 31-year running mean timeseries until each stocktake year, which introduces some noise compared to the true final estimate.*"