Reply reviewer 1

We'd like to thank Reviewer 1 for his/her careful review of the manuscript and the valuable suggestions for how to improve it. Please find our detailed replies below (in italics).

This paper describes a new large ensemble with the climate model EC-Earth3 and a relatively unique combination of emissions scenarios. The analysis examples are largely uncontroversial/unsurprising, which is perhaps intentional and adequate for a description paper. The paper is written very clearly, with sound methods and results that support the conclusions. I only have a few small comments that could be considered in a revised version, after which I can recommend publication. I want to express my thanks to the authors for making this data available to the community; it is very valuable.

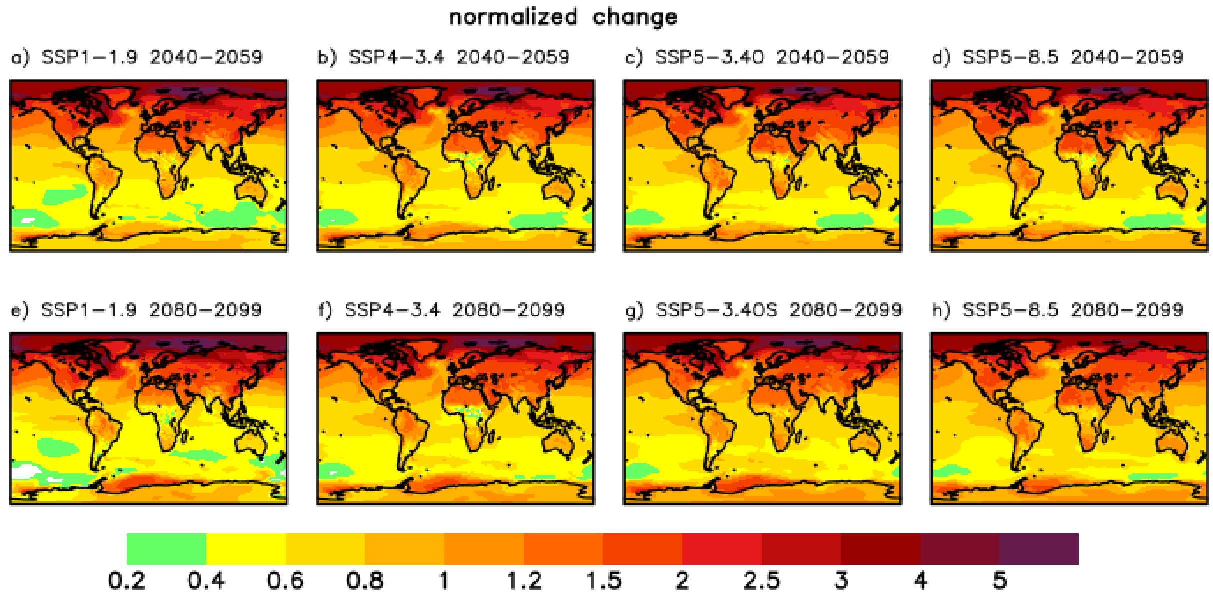
Three requests as a curious reader:

Could the authors make a few intuitive comparisons with CMIP6? In particular, I was under the impression that EC-Earth3 has among the largest decadal global temperature variability of any CMIP6 model (at least in the piControl). How does this affect its signal/noise ratio in comparison to other CMIP6 models and other available large ensembles?

Indeed, the large variability of EC-Earth3 compared to other models and its impact on the signal to noise ratio is definitely worth mentioning. We have added a paragraph to the conclusions in which we emphasise that the results presented in this study are valid for EC-Earth3 and could look different for other models.

It would be interesting to divide the change patterns from the different emissions scenario by global mean temperature, so the pattern become comparable. I actually think that would be a more interesting analysis (and easy to do) than the absolute change maps that are shown here. The authors kind of have to show the absolute change patterns for the description paper, I get that, but we already knew how they would look. I think a comparison of the normalized patterns would be very interesting for the overshoot vs non-overshoot path.

Thank you for this interesting suggestion. We now have looked at the normalised change for the temperature in the different scenarios and found that the patterns are very similar. Therefore we don't think that it is worth including this result in the paper (attach figure).



I encourage a comparison with or at least discussion of Sanderson et al. (2017) and several references therein, which investigate what might have been the first “large” (n=10) ensemble of overshoot simulations, albeit with CMIP5 forcing and not CMIP6 forcing.

We have added a paragraph in the introduction about the Sanderson et al. paper to acknowledge that we are not the first group to look at the effect from an overshoot. However, we cannot do a more detailed comparison as the forcing used by Sanderson et al. is very different from our forcing. Sanderson et al look at overshoot and steadily increasing projections that end with 1.5 deg warming in 2100 which would correspond to a forcing below 2 W/m2 while we look at a substantially larger forcing of 3.4 W/m2 in 2100 and a substantially stronger overshoot around mid-century.

Other comments:

There have been some concerns with the transition of biomass burning (BB) forcing from the satellite period to the future emissions scenarios in the CMIP6 forcing files. The forcing file has high variability during the satellite era and abruptly lower variability thereafter, at least in certain regions of prominent BB. A large ensemble is perfect to investigate whether this has an appreciable effect on the simulated climate during that transition period. It’s not a priori clear whether it affects each model, as aerosol forcing is implemented quite differently across models. It would be valuable information for the community whether there's any sensitivity to this issue in EC-Earth.

This is also a very interesting question but we deem it to be outside the scope of the overview paper. Furthermore, it is not clear if this question can be addressed at all with the EC-Earth3 model that doesn't use an interactive aerosol module but instead the simple aerosol plume model MACv2-SP to prescribe the optical properties of aerosols. MACv2-SP doesn't distinguish explicitly between biomass and other aerosol species, it only tells whether a plume is dominated by biomass without telling the exact contribution from biomass.

Fig. 2: Please add some observations (at least for tas, pr, and sic) as a minimal model validation and to illustrate how the ensemble range compares to real-world variability.

Thanks for the suggestion, we have added observations/reanalyses for tas, pr and sea-ice to Figure 2.

Some map colorbars are a bit unintuitive, as they have a color switch offset from zero (e.g., Fig 3c-j). Also, rainbow or rainbow-like colorbars aren't encouraged.

We have reworked the colorbars in Figs. 3-5 to make them more intuitive. Re rainbow-like color palette: the co-author who has prepared the figures (TK) is heavily colorblind himself and uses these specially designed color tables that allow him to distinguish the different colors while still providing an appealing appearance for non-colorblinds. We therefore don't think that it is necessary to replace the chosen color tables in Figs. 3-5. In Figs 2 and 6 we have replaced the orange color by green to increase the contrast to red. And in Fig.7 we now use a yellow-green color table with increasing intensity. Furthermore, we have also improved the quality of the figures by saving them at higher resolution.

L188: "mainly due to reduced variability of the change across members" sounds confusing to me. Do the authors just mean "due to the larger signal under stronger forcing"?

Changed.

L218: "highly relevant" rather than "highly important"

Changed.

L151: that’s actually quite interesting. Could the authors speculate what could cause this?

The oceanic convection in the Labrador Sea is an important source for variability of the AMOC and this convection shows large natural variability on long time scales in EC-Earth. This results in a large spread of the AMOC across model members in the historical period and the early 21st century. Moving towards a warmer climate with increased ocean surface temperatures and a fresher ocean surface in the Labrador Sea, the convection collapses completely in all model members until the middle of the 21st century (Brodeau and Koenigk 2016, Koenigk et al., accepted). Consequently, the spread of convection in the Labrador Sea and thus the spread of the AMOC is reduced. The figure below shows one standard deviation of convection in the Labrador Sea in March across all model members. Here, the ssp5-8.5 scenario is shown but a similar behaviour occurs in the other scenarios as well. We see that most of the reduction happens already before the differences between the different emission scenarios become dominant.

We added some discussion to the manuscript to explain the decrease in ensemble spread.

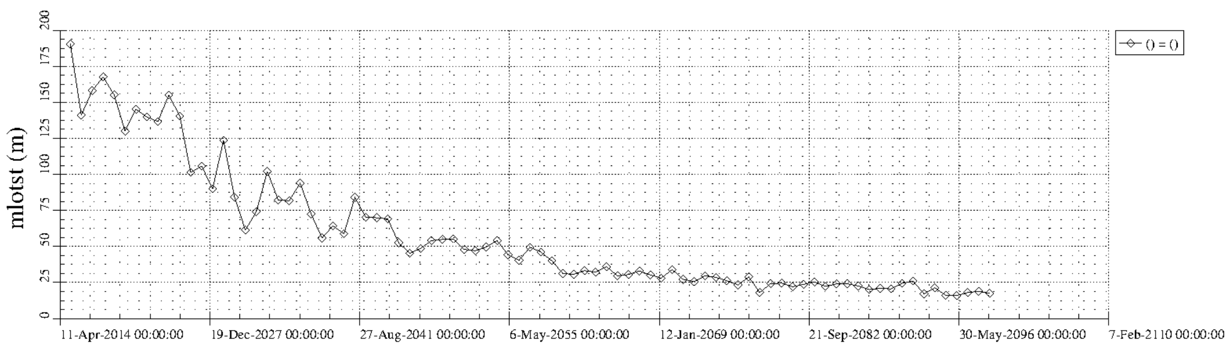


Figure: Time evolution of one standard deviation of March mixed layer depth in the Labrador Sea (in m, averaged over 45-72N, 270-310W), calculated for each March as standard deviation across the 50 model members.

L100: “There are differences”

Changed.

L116: “tell us”

Changed.

References:

Sanderson, B. M., et al., 2017: Community climate simulations to assess avoided impacts in 1.5 and 2°C futures. Earth Syst. Dyn., 8, 827–847.