

Review of responses of authors to comments 1 and 2 of reviewer 2

Robustness of linear regressions

Reviewer 2 pointed out that the linear regression for non-CO₂ emissions and BC forcing could break down for future scenarios and required a sensitivity analysis to test the model's robustness.

The authors extend the dataset to the year 2100 with 1703 scenarios from the IPCC AR6 database. The added figure S1 in the supplement shows the regression predictions for NO_x, VOCs and CO emissions and the forcing of BC for these scenarios. Additionally, the resultant impacts on the surface temperatures are calculated and shown in this figure.

The shown results in figure S1 support the validity and robustness of the chosen linear regressions for the individual species. It is stated that the errors in surface temperature are typically under a few hundredths of a degree.

For clarification, I recommend to mention in the text where the target variables (blue lines) are based on, e.g. Are those also taken from the IPCC AR6 database like the predictor variables? Furthermore, in the text, the statement of the surface temperature errors (few hundredths of a degree) should be compared/related to the absolute temperature changes.

Unexplained Temperature Divergence

Reviewer 2 pointed out the relatively high temperature difference between the models FRIDA-Clim vs. FaIR in case of high-emission scenarios and asked for a better explanation as well as for an isolation of the two effects "carbon cycle" and "EBM parametrization".

The authors provided Figure S6 and additional description in the main text to describe the results of the comparisons between models shown in Figure S6.

The results shown in Figure S6 seem convincing, in the sense that the two effects "carbon cycle" and "EBM parametrization" could be well isolated. Therefore, the explanations in the main text are much better supported, i.e. for the question: *What are the main drivers of the temperature differences between models at a certain scenario background?*

When I first had a look on Figure S6, I struggled with the labels and assignment of the curves to it. Therefore, I recommend to use a similar labelling as in Figure 6 (main paper) where the SSP scenarios have been labelled on the individual columns. That would help the reader to understand the view graph faster.

I would further suggest, to reference back to section 3 where the details of ensemble members and calibration are explained which is used here, to my understanding.