

In the responses below, the original reviewer reports are in black, while all our comments are in blue. We have also numbered all the reviewer comments and our replies for clarity. We have quoted text from the manuscript in grey italics.

Reviewer 3 Comment 1

The paper aims to introduce the motivation and rationale for a model intercomparison of reduced-complexity models. These models are commonly used to interpret (mostly global mean) temperature observations and complex model simulations. The paper introduces scientific questions which can be answered with this type of models, the experimental design and diagnostics, participating models, and shows first analyses of modeled temperatures for the historical period and scenarios for the next century.

Reviewer 3 Reply 1

Thank you for your review of our paper. We greatly appreciate the time you have put in and have found your comments very helpful, particularly to better define the scope of our paper. We have produced an updated manuscript which we feel is greatly improved thanks to your suggestions. We hope it is a useful first step to helping the 'the big group of people who are confronted with RC output but don't know how to evaluate them' (as well as us as model developers).

Reviewer 3 Comment 2

One important contribution to the ongoing discussion of differences between CMIP5 and CMIP6 is the finding, that about 46% of the additional warming at the end of the 21st century in CMIP6 compared to CMIP5 models stems from differences in the radiative forcing in SSPs and RCPs.

Reviewer 3 Reply 2

Thank you for your comment. In response to Reviewer 2 (specifically their Comments 3 and 14) and after discussion with the editor, we have removed this discussion from this paper. Following your Comment 6 as well, we have removed this discussion and will save it for a paper which has the room to explore it in the detail it deserves. Instead, we have simply presented the difference between the SSPs and RCPs in the sample results section and stated that there is a difference in the results submitted to date but further evaluation is required to fully understand why because these are only preliminary results.

Reviewer 3 Comment 3

I applaud the endeavor to conduct an RCMIP. This will be very useful, both for people using single RCs but also the big group of people who are confronted with RC output without knowing how to evaluate them (both the impact user side and the GCM modeler side). However, the paper in its current format is weak and obscure and does not allow me to draw clear conclusions. It should not be too much effort to improve the paper, as it is mostly "just" improving the presentation, explanation, arguments, clarity. No new simulations are necessary.

Reviewer 3 Reply 3

Thank you for your positive comments. We agree that the paper required significant updates and have done so in the revised manuscript. In particular, we have focussed solely on the presentation of the MIP, leaving comprehensive evaluation of the results for future study.

Reviewer 3 Comment 4

Major comments 1) The differentiation between possible/future research questions and the ones addressed (and answered?) in this paper is unclear. I think possible questions do not belong into a paper. Anybody can come up with some vage questions. I read a paper to learn about what has been done and how I can used this for my own research. What somebody (who?) might be doing/planning/considering can be discussed in conferences etc. not in a scientific paper.

Reviewer 3 Reply 4

Thank you for your comment. We agree that these comments do not belong in a paper and have removed most of them, leaving a few relevant ones for the Extension section.

Reviewer 3 Comment 5

2) From a model intercomparison paper, I'd like to learn how I can use the output, which criteria have been used to select the models, which experiments have been conducted, . . . technical parameters, what can I learn from your effort. There are a lot of MIP-explaining papers out there. I suggest to imitate one of these in the structure and focus of the paper.

Reviewer 3 Reply 5

Thank you for your comment. We have updated the paper to read like a MIP-explaining paper (most closely following the structure of the RCEMIP paper) and hope that this makes the purpose of RCMIP and how it can be used much clearer.

Reviewer 3 Comment 6

3) The interpretation of the differences between CMIP5 and CMIP6 scenarios is a major scientific contribution to the ongoing discussion. It is extremely relevant for writing the IPCC report. As such it belongs into a more visible, less technical journal and it needs to be highlighted. Here, this finding is buried towards the end of the paper and I get the sense that this is because the science behind this finding is actually not really well understood, at least I don't from reading the paper. What's the relationship of this finding with the paper of Forster et al. 2019, who's estimate for the impact of the different scenarios (in CMIP5 vs 6) to surface temperature is much smaller.

Reviewer 3 Reply 6

Thank you for your comment. We agree that such a discussion requires a much more in depth discussion and have accordingly removed it, saving it for future research (see also Reply 2).

Reviewer 3 Comment 7

Minor comments Title: “initial observations” - these are modeling results, reformulate

Reviewer 3 Reply 7

Thank you for your comment. We have updated the title to better reflect the revised manuscript’s focus on MIP description.

Reviewer 3 Comment 8

line 12: output - in what

Reviewer 3 Reply 8

This was indeed unclear. We have reworded the sentence for clarity.

Reviewer 3 Comment 9

line 12: change in scenario - please explain much more thorough throughout the paper: Why has the scenario been changed? Were they not supposed to be traceable (i.e. SSP8.5 approx RCP8.5? How does this fit with the 46% additional warming due to different scenarios

Reviewer 3 Reply 9

Thank you for the comment. Following your comments and Reviewer 2 we have removed the (attempted) discussion of the percentage difference between the SSPs and RCPs and instead simply presented raw results from the models along with the caveat that these results are not comprehensive. We agree that exploring this change in more detail is required and hope future study can do so.

Reviewer 3 Comment 10

line 15: “as first anticipated” - by whom and why?

Reviewer 3 Reply 10

This expression was indeed vague and has been removed.

Reviewer 3 Comment 11

line 16: “provide results available . . .” which results? (the authors can pick/will find the results they need themselves. It’s not a scientific finding to plan to provide results.)

Reviewer 3 Reply 11

Thank you for highlighting this odd formulation, we have removed this line from the revised manuscript.

Reviewer 3 Comment 12

line 28: “exploring interacting uncertainties” - explain which parts of the climate system can interact in these models - mostly they only contain surface temperature and some forcing agents and parameterized ocean heat uptake?

Reviewer 3 Reply 12

Thank you for this suggestion. Many of the models contain parameterised representations of the carbon cycle, non-CO2 gas cycles and land surface, all of which can interact and represent (in a parameterised way) many of the key feedbacks in the Earth System. We now clarify this with some examples in the revised manuscript and hope to provide a follow up paper with much more detail on the models' different structures in future (given we do not have the space to do so here).

Reviewer 3 Comment 13

line 41: useful statistics - useful for what?

Reviewer 3 Reply 13

Thanks for pointing out this unclear sentence, it has been removed during the revisions.

Reviewer 3 Comment 14

line 60: . . . to understand their strengths, weaknesses and limitation so that we can make more confident, informed conclusions from their quantitative results” - yes, great, it would be good if all these points were indeed discussed in the conclusion in a clear manner.

Reviewer 3 Reply 14

Thank you for the comment. We have expanded the discussion in the conclusion and put the areas for further examination (as there are many we have not yet covered) in the extension section.

Reviewer 3 Comment 15

line 75: what's a lifetime of an RCMIP?

Reviewer 3 Reply 15

Thank you for the comment, we have removed this unclear phrasing.

Reviewer 3 Comment 16

line 75 following: For this paper, specify the questions you are actually answering. Here is mixed list is given of what is and could be done. Could/should/would is used much more in this paper than in usual scientific literature. For a MIP paper, a discussion of future/possible questions is fine, but these should be listed in one concise place and not dominate the paper.

Reviewer 3 Reply 16

Thanks for pointing this out, during the restructuring of the manuscript we have consolidated the most important of these future questions into one extensions section.

Reviewer 3 Comment 17

line 77: “some aspects will receive less attention here than others” - and later: more precise language would help to make this a scientific and useful paper and not an opinion piece.

Reviewer 3 Reply 17

Thank you for the comment. We have removed this line and clarified the scope of the paper, pushing all the extensions into a single extensions section which we hope helps to restore the scientific tone of the paper.

Reviewer 3 Comment 18

line 78: “what can they tell us about” ?

Reviewer 3 Reply 18

We agree this is an awkward, colloquial phrasing, it has been revised.

Reviewer 3 Comment 19

line 84: Experimental design: This section is not actually describe the experimental design fully. Or at least, after the section, I wouldn't be able to replicate what you did. From the title, I expect a list of experiments, their input, assumptions, rational, in a clear understandable fashion. Right now the section is a collection of random issues (a lot of detail about emission some specific scenarios, non about others)

Reviewer 3 Reply 19

Thank you for the comment. We have overhauled the experimental design so it actually describes what we did in a clear, reproducible fashion.

Reviewer 3 Comment 20

line 96: What's the standard set of inputs from CMIP5 and CMIP6?

Reviewer 3 Reply 20

Thank you for the comment. We agree this is not sufficiently clear and have clarified in much more detail in the revised experimental design section.

Reviewer 3 Comment 21

line 120: Diagnostics: I expect to learn how I can use this data. What's the available output? What's the rational for it?

Reviewer 3 Reply 21

Thank you for the comment, we have updated the manuscript to include a standalone output and diagnostics section which outlines the available output and why it is requested.

Reviewer 3 Comment 22

line 128: How is an RC model defined? What's the criteria to be included in your comparison? Table one is a nice overview. I suggest to move even more information from the text into the table: What are the input variables and assumptions about them? On what data are they tuned? Add a paragraph about similarities and differences among the models. What do they all share? Some of them seem to use the same basic equations. Are there classes of RCs? Could you draw a genealogy? Which ones are structurally more similar? Which ones are fully independent? From the text, I e.g. do not get a good sense of the difference between FaIR and CICERO-SCM.

Reviewer 3 Reply 22

Thank you for your comment. We have updated the text to provide clarity.

In the RCMIP community call (available at rcmip.org) RCMs were broadly defined as follows: “[...] RCMIP is aimed at reduced complexity, simple climate models and small emulators that are not part of the intermediate complexity EMIC or complex GCM/ESM categories.” In practice, we encourage any group in the scientific community who identifies with the label of RCM to join RCMIP.

We agree that the topic of differences and similarities between RCMs is an important one. However, we do not feel that we have sufficient space within this MIP description paper to do it justice hence, after consulting with the editor, have removed all but the most important details from this paper. We hope to provide a follow up paper which does discuss model details and genealogies in far more detail.

Reviewer 3 Comment 23

line 156 what does it imply to have two or three timescales?

Reviewer 3 Reply 23

We have removed discussion of the intricacies of different model setups due to space constraints. We hope to provide a follow up paper which can cover this topic with the detail it deserves in future.

Reviewer 3 Comment 24

line 433-436 I do not understand the sentence “These probabilistic . . . “

Reviewer 3 Reply 24

Thank you for the comment. The key difference is that the probabilistic distributions derived from RCMs are designed to capture the likelihood that different warming levels are reached under a specific emissions scenario (e.g. 50% and 66%) based on the combined available evidence hence are quite different from an ensemble of multiple model outputs, which have been produced independently with no relative relationship in mind. Hence the probabilistic

distributions provide extra information on top of what is provided by the raw CMIP6 output, which we described as 'extending'. We agree this was not as clear as it should have been and have clarified in the revised text.

Reviewer 3 Comment 25

line 442-444 "Given that . . . " is pure speculation. Somebody might be doing these experiments, who knows, maybe not, . . . what's the purpose of this "information" here? Is this a call to the community that these experiments should be done? Are you planning to do them? Can I expect the results in phase 2? Maybe this is all about precision of formulation only? It's so vague, I don't know what to do with this information.

Reviewer 3 Reply 25

Thank you for the comment. We agree it was far too vague. We have moved it to the new extensions section and made clear that we are calling on the community to do such experiments in future.

Reviewer 3 Comment 26

line 448: Developing a method . . ." "Such results would enhance ..." same as the point above: Are you suggesting to do this? Are you doing this? Should I do it? Why don't you do the research first and then tell me about the outcome?

Reviewer 3 Reply 26

Thank you for the comment, like for Comment 25 we have removed the text and clarified in the extensions.

Reviewer 3 Comment 27

line 464 following: Isn't this way too important to be buried in the Supplemental Material?

Reviewer 3 Reply 27

Thank you for the comment. As discussed in Reply 2, we have revised the manuscript to present the raw data here and caveated the results by acknowledging that these are raw outputs and further evaluation is required to make strong conclusions.

Reviewer 3 Comment 28

line 469: monotonic relationship?

Reviewer 3 Reply 28

We have removed this text from the manuscript. For clarity, in this context monotonic simply means that if CO₂ concentrations increase, CO₂ effective radiative forcing increases.

Reviewer 3 Comment 29

line 472 “At this stage, this residual is most likely explained . . .” and in two months you might change your mind or interpretation? Why not waiting with writing a paper until clear results and their interpretation materialize?

Reviewer 3 Reply 29

Thank you for the comment. We agree that the speculation was not helpful so have presented only raw results in the revised manuscript with appropriate caveats (see also Reply 2).

Reviewer 3 Comment 30

line 476 “A number of experiments have not been discussed here. . .” . . .?

Reviewer 3 Reply 30

Thank you for the comment. We agree this is vague and not helpful. We have revised our results section into an ‘Illustrative results’ section to make clear that the presented results are to illustrate the usefulness of the MIP, rather than being intended as detailed evaluation.

Reviewer 3 Comment 31

line 480 I can’t follow. Your “Conclusion” is an Outlook. Both, a conclusion and an outlook would be useful. I suggest to re-write the entire paper and discuss solely the models and (clarified) experimental set up and the *results* and then have one dedicated Outlook section with all your if/when/could/should/might items and maybe a clear plan for phase 2.

Reviewer 3 Reply 31

Thank you for the recommendation, we have found it very helpful. We have re-written the entire paper as suggested to make it much more focussed on the MIP description. After discussions with the editor, we decided to remove the discussion of the models as there is not sufficient space to discuss both the models and the MIP to the level of detail required. We hope to provide a detailed model description paper in future.

Reviewer 3 Comment 32

Fig.1 suggestions: Shade CMIP5 and CMIP6 models? There’s too much information in this plot, I can’t differentiate the lines. Maybe add panels with each RC to the SM?

Reviewer 3 Reply 32

Thank you for the suggestion, we have updated the plots to make clear the different lines.

Reviewer 3 Comment 33

Fig.2 a again, I can’t see which information is relevant here. e.g. why are there not “hector” for SSP126.

Reviewer 3 Reply 33

Thank you for the comment, we have updated the plots to highlight the key story (i.e. the degree to which RCMs reproduce the target CMIP6 model’s behaviour). In the revised manuscript, we

clarify that the calibrations depend on each RCM development team's individual capacity hence there is no Hector output for ssp126.

Reviewer 3 Comment 34

Stretch all plots into the horizontal.

Reviewer 3 Reply 34

Thank you for the suggestion, we have updated the plots.

Reviewer 3 Comment 35

Why do the RCM lines stop earlier than the GCM line in panel c)? Use year 1, 2, 3 instead of 1850, 2000, . . . this is very confusing for idealized experiments.

Reviewer 3 Reply 35

Thank you for the comment. We only requested the abrupt-4xCO₂ experiment be run until 2500 in our experiment protocol but the CMIP6 run has continued longer. We agree this is confusing and have updated the plot to remove this confusion as well as to use a more standard time axis.

Reviewer 3 Comment 36

Table 2 "[TO DO . . .]"

Reviewer 3 Reply 36

Thank you for picking this up, we have fixed the reference.

Reviewer 3 Comment 37

Fig.4 It's hard to see the point here (SM fig. 3 and 4 are much clearer). maybe change shading/colors? The information of the historical is not needed at this point anymore, the figures could start at year 2000 or so and then stretched. Maybe this would help to make the information more digestible? Adding versions of SM Fig. 3 and 4 could help to make this point stronger in the paper.

Reviewer 3 Reply 37

Thank you for the comment. We have updated the figure as suggested (and in line with Reply 2 and our softening of conclusions). The difference between the RCP and SSP-based scenario pairs is now much clearer.