

**Point-by-point reply to the review of gmd-2023-29 manuscript :
„Interactions between atmospheric composition and climate change – Progress in
understanding and future opportunities from AerChemMIP, PDRMIP, and RFMIP“**

We would like to thank the two anonymous reviewers and the editor for the appraisal of our manuscript. Please find below our replies in blue to the comments in black. We include details on how we modified the text in response to the comments and enclose a version of the revised manuscript where we marked all changes in color.

Reviewer #1

I don't think I need to see this manuscript again. My comments can be seen as opinion or suggestion, but not necessarily need to be included in the paper.

Thank you for your comments that helped to improve the manuscript.

Reviewer #2

To the authors,

I appreciate the modifications and improvements to the manuscript, and that review comments have been taken into account. However, I repeat point (2) of my earlier review : "The study is generally well written and structured, although specific sections should be improved for better understanding. I also think that there should be an effort in making the style more homogeneous over the whole manuscript. Also, one sometimes gets the impression that some paragraphs (or sentences) do not really belong in a section : they should be brought more in harmony with the text around them. This is indicated in the detailed list below."

After reading the revised manuscript, I think there are important issues remaining concerning the text of the manuscript :

- some paragraphs contain parts that remain unclear or vague, and some sections are too long;
- ideas should be more precise, without multiple sentences containing very similar messages or information; repetition should be avoided;
- the reasoning within a paragraph should be clear, logical and well-structured; and
- the style should be more homogeneous throughout the text.

I hope the authors will make an effort to improve the manuscript. I am aware that this is a time-consuming task, but I think it is necessary in order for the manuscript to be useful and attractive. I think the manuscript does not need a structural change, but mainly an improvement in how the thoughts, ideas and reasonings are formulated. Is it clearly and concisely written? Is there no repetition? Are all the reasonings clear?

Below is a list of comments mainly related to the clarity of the text. Please take these comments into account when trying to create an updated version of the manuscript.

Thank you for your thorough review. We appreciate your commitment and your time to help making the content more attractive for readers.

ABSTRACT :

- line 16 : "simulation number and length" : maybe "number of simulations and their length".
Changed as suggested.

1. INTRODUCTION

- line 19-20 : "aspects" and "components" : two rather vague terms in one sentence.

Changed to: „This endeavor involves the assessment of numerous spatiotemporally changing variables in the Earth system, which can be determined by multiple, interacting physical, chemical, and biological processes.“

- line 19-20 and line 25-27 : these sentences in the same paragraph are very similar in their message (although one gives more examples). This should be more structured, without repetition.

We removed the sentence in line 25-27 and rephrased the sentence aloft: „For example, changes in irradiance, land use, and atmospheric composition, including for instance aerosols and their precursors, greenhouse gases such as carbon dioxide and methane, perturb the radiation fluxes in and at the top of the atmosphere and hence the Earth’s radiation balance. “

- line 21-22 and line 23-25 : it should be avoided that a definition of ”radiative forcing” via ”is called a radiative forcing” on line 23-25, comes after the first use of ”radiative forcing” on line 21-22.

Resolved, the term radiative forcing is now first stated at the end of the paragraph.

- line 25 : ”power density” : is this the correct term?

Thanks, power density with the unit for radiative forcing in brackets might be misleading without providing further context. Changed to: „and is measured as energy flux in $W\ m^{-2}$ “.

- line 30 : ”... that arise from a climate forcer, e.g., a perturbation in the atmospheric composition.” This is still vague. It does not mention that IRF, in principle, is just a diagnostic, not affecting the temporal evolution of the atmosphere.

Changed to: „ (...) initial change in radiation fluxes that arise from a perturbation in a climate forcer, which could be, for instance, associated with increased greenhouse gas concentrations in the atmosphere due to anthropogenic emissions, in the absence of other changes. IRF excludes any changes in the system other than an imbalance in the Earth’s top-of-the-atmosphere (TOA) radiation budget and is a diagnostic output from Earth System Models (ESMs).”

- line 34-35 : ”Smaller forcings ... than larger perturbations.” I don’t know whether this addition is useful.

Removed.

- line 35-39 : I assume that ”change in atmospheric composition” is the initial perturbation, and ” changes in emissions of reactive trace gases” is a consequence. For the last part, is one thinking of natural responses such as modification in emission changes from ”soil NO_x”, ” lightning NO_x” and ”isoprene/monoterpenes from vegetation”? Is that what is meant?

Yes, natural emissions are meant here. Examples are given further below in the revised paragraph: „Relevant examples are adjustments and feedbacks that modify desert-dust and sea-spray aerosol emission changes.“ The initial perturbation is the change in atmospheric composition, which is technically induced by perturbing concentrations of the atmospheric constituents directly or by perturbing the emission fluxes that lead to a change in atmospheric composition. It depends on the model and the experiment protocol which of the two are done.

- line 42-43 : ”Relevant examples are ... desert-dust and sea-spray aerosol” : maybe add ” emission changes”.

Done, now: „Relevant examples are adjustments and feedbacks that modify desert-dust and sea-spray aerosol emission changes.“

GENERAL : line 48-62 : This paragraph is presenting quite some ideas : ESMs, their differences, MIPs, ensembles, ... I think most of the ideas mentioned here have their place, but it should be written in a more logical and coherent way.

Thanks, we have changed the paragraph as follows in the reply to your specific comments.

- line 48-49 : Understanding and quantification ... are ”assessed” : that is a strange expression. I would rather use ”obtained”.

- line 49 : ”typically” : sounds not very nice

Changes as suggested: „Understanding and quantification of the different segments in the perturbation-response paradigm of climate science are obtained through experiments with Earth System Models (ESMs), (...)“

- line 50 : ”Modern” ESMs ! ”Current” ESMs

Changed as suggested: „Current ESMs (...)“

- line 50 : in their "design" : I would see "design" choices as high level decisions, and implementation as low-level realizations of those. For me "design" remains a strange term to use in this context.

Using design was motivated by the context of „program design“ which addresses the process of high-level decisions for the structure of a program before one starts coding, consistent with your understanding. It is true that also the implementation differs across ESMs. We therefore add: „Current ESMs vary in their design and implementations, e.g., (...)“

- line 50-51 : "Modern ESMs vary in their design ... and boundary data." : the examples listed seem to refer mainly to how things are implemented at low-level, except for "concerning different parameterization schemes", which seems rather vague and could maybe fall under "complexity".

The design, implementation and complexity are not independent of each other. We therefore now change the sentence to better reflect the link in writing: „These imply diversity in the level of complexity for representing physical, chemical, and biological processes, and how represented processes interact.“

- line 53-54 : "can simulate aerosol and their precursor emissions, and transport and deposition of aerosols." What about aerosol growth and coagulation? Maybe use a general term describing that aerosol undergoes various processes in the atmosphere.

We picked these as example and now expand this sentence to: „For example, some ESMs prescribe aerosol properties while models with additional process complexity simulate the complex evolution and interactions of aerosols and their precursors in the atmosphere (...)“

- line 55 : formation of cloud droplets and "ice" ! and "ice crystals"

Changed as suggested: „ice crystals“

- line 55 : "or just part of it" : not so nice

Now more specific: „The simulated aerosols may interact with the radiative transfer and formation of cloud droplets and ice crystals, but not all ESMs simulate all interactions with the cloud microphysics.“

- line 55 : "radiation" transfer ! "radiative" transfer (both are used in the text; maybe choose one)

Thanks for spotting this. Now it is radiative transfer throughout the manuscript.

- line 55 : regularly produces : not so nice

Removed and now: „The climate modeling community creates multi-model ensembles (...)“

- line 56 : "However" : I would skip it. Maybe I would build up the reasoning differently in this sentence.

Now: „The climate modeling community creates multi-model ensembles of a common set of ESM experiments with the same perturbations applied. The simulated climate responses can differ across a multi-model ensemble. This diversity in responses may for instance be due (...)“

- line 58 : "Ensembles of different ESM experimental setups following the same protocol ..." : this is vague

Agreed, now changed to: „Experimental protocols are used to create multi-model ensemble simulations for specific ESM experimental setups. These aim to better understand the reasons for the diversity in climate responses and feedback and to create future climate projections. “

- line 62-63 : The most prominent example is the experimental "protocols" : there is an issue with singular/plural I would think. I also think that a MIP is not the same as a protocol.

Corrected: „The most prominent example is the Coupled Model Intercomparison Project (...)“

- Finally, I would not say that the "protocol" has "informed" the "assessment reports".

Changes to: „that has contributed through multiple phases to the assessment reports of the Intergovernmental Panel on Climate Change (IPCC, Meehl et al., 2023), e.g., the sixth phase of CMIP (CMIP6, Eyring et al., 2016) created experiments that were also used in the sixth IPCC assessment report (IPCC-AR6).“

2 SCIENTIFIC ADVANCEMENT

2.1 MIP'S KEY RESULTS

GENERAL : line 83-96 : I am not so happy with how this paragraph is written (the rest of this section is ok)

We revise the paragraph as follows.

- line 84 : "considering structural differences" : is this the main purpose of the MIPs?

The purpose is the first part of the sentence: „The three MIPs sought to advance the understanding of modern climate change due to anthropogenic influences.“ We split the second part and explain it more as follows: „MIPs address specific research questions and, in comparison to studies with a single ESM, consider structural differences concerning the design and the level of complexity between ESMs. The multi-model spread in the response allows the quantification of a model-based uncertainty for the answer to the MIP's question.“

- line 88-91 : this sentence does not read very well

We now split the content of the sentence as follows: „AerChemMIP also focused on quantifying radiative forcing and responses but addressed more segments in the paradigm. Specifically, all participating models in AerChemMIP simulated atmospheric composition based on emissions, transport, chemical transformations, and deposition, making these models more complex in their process representation and interactions than was necessary for participation in the other two MIPs (...).“

- line 92 : CMIP-class models : strange term

Removed „CMIP-class“.

- line 92-93 : "with a certain class of models in mind" (here the reader thinks that there are 3 different ideas); "i.e., CMIP-class models in all three MIPs" (now they seem equal) and specifically AerChemMIP required (now one seems to be different) : this sentence is not very nice
Agreed, changed to: „ (...) e.g., AerChemMIP required more interactive processes than the other two MIPs.“

- line 94 : "Taken together" : not nice

Removed: „There are (...).“

- line 95 : number : sounds strange

Now: „number of realizations“

- line 98 : "irradiance, sulfate, and black carbon aerosols." ! "irradiance, and sulfate and black carbon aerosols."

Thanks, corrected.

- line 105 : "." should be added at the end of the line

Added.

- line 108-109 : ... to estimate the real-world evolution and timing of emission changes ... : if this refers to anthropogenic emissions, then it is a bit strange (timing of emissions is an externally imposed factor). Should one add "natural"?

Added: „(...) anthropogenic and natural emission changes (...).“

- line 112 : "(hist)" : should be "(hist)"

Corrected.

- line 115 : where the emissions or concentrations ! where "only" the emissions ...

Changed as suggested.

- line 119 : "That is ..." ! "It implied ..."

Changed as suggested.

2.2 MIP's CROSS LINKAGES

- line 147 : the magnitude of a perturbation "in" the radiation budget : is "in" or "on" best representing the meaning?

Revised to: „(...) characterizing the influence on the radiation budget due to a perturbation.“

- line 155 : "the parallel use of preindustrial control experiments" : this is a bit confusing. Does it refer to the fact that both MIPs use the same reference simulation? Or that for most analyses two simulations are used together to derive results? Maybe use "common use".

Changed as suggested: „the common use of pre-industrial control experiments “

- line 158-161 : should one make clear that the "double calls" are used for the IRF calculation and the "triple calls" to separate aerosol direct and cloud-mediated effects?

Yes, that can be useful. Now: „(...) and a better understanding of contributions from different processes to ERF. Double calls typically refer to IRF calculations, whereas the term triple calls is used for separating cloud-mediated effects from direct effects of aerosols.“

- line 166-169 : these two sentences on AerChemMIP fall within a paragraph mainly on RFMIP. Even if these two sentences have their place here, they refer to the first sentence in the paragraph, not to the sentence they immediately follow on. So the structure might be improved.

We moved the content up in the paragraph: „The RFMIP protocol included experiments to diagnose radiative forcing for greenhouse gases and aerosols as bulk quantities with setups parallel to the CMIP6 experiments for the "Diagnostic, Evaluation, and Characterization of Klima" (DECK). The RFMIP tier 1 experiments were carried out by many modeling centers. Some of these contributions, e.g., from UKESM1 and CNRM, arose because the experimental setup was identical to the request in AerChemMIP. It meant that the technical workflow for performing and postprocessing the experiments was already in place such that contributing another variant of such experiments required only little effort. .“

- line 167-169 : this is already said on line 128-130. It is allowed to mention things twice, but one should be careful with it.

Removed here: „As such, RFMIP was able to characterize forcing in CMIP for the first time.“

- line 178 : "(histSST piX)" ! "(histSST-piX)"

Thanks, corrected.

- line 183-184 : "... and the specialized communities for aerosols and atmospheric chemistry that do not participate in CMIP." : my impression is that part of the AeroCOM and CCMI community participates/contributes to CMIP via actually RFMIP and AerChemMIP. So I would maybe not formulate it so strongly.

True, individuals are involved in both communities, although not the entire communities are overlapping. Now removed: „that do not participate in CMIP“

- line 195-196 : "Some key articles based ... deadline." It is unclear what the authors want to stress with this sentence. Now it feels that it stresses that there was little time and some analyses were ready only just in time, whereas it is also possible that the sentence wants to express that there was good collaboration and timing.

Now: „In fact, some key articles based on the experiments were written and submitted very close to the IPCC-AR6 WG1 deadline, which might not have been completed in time if that exchange had not happened.“

- line 197-198 : until a quorum of CMIP7 model output "come" online : "comes".

Thanks, corrected and adjusted to „becomes“.

- line 198 : "online" ! "available"

Now: „becomes available“

3. CHALLENGES IN THE MIP'S RESEARCH

This first paragraph is ok, but there is a lot of repetition of the same words such as "model diversity" and "model differences". I think this paragraph can be improved.

Ok, we improve it as follows.

- line 202-205 : I assume that I understand what the authors want to express - however it is not very clear. Is the line of thought : When a difference in climate response to emissions is observed between models, one cannot say whether it is due to : (i) same emissions give difference in radiative forcing; (ii) same radiative forcings give different response; or (iii) same climate perturbation gives different feedback. Or does one rather want to express that spreads in the radiative forcing (for same emissions), and spreads in the response (for same forcing), cannot just be combined independently?

We are thinking of the former line of thought, although the second interpretation is also true. We now expand this topic in the text for clarity: „Specifically, a model-to-model difference in a climate response might be caused by various segments in the paradigm. For instance, the same emissions can lead to different ERFs, the same ERF can induce different climate responses and the same response can trigger different feedbacks across multi-model ensembles. In multi-model studies, one therefore sees inter-model spreads in forcing for the same change in atmospheric composition and model-dependent climate responses to the same forcing involving different types and magnitudes of feedbacks.“

- line 211 : "elsewhere" (general) "... in the AeroCom community" (specific) : there seems to be a contradiction between "elsewhere" (which sounds general), and the rather specific mentioning of "AeroCom" at the end. I would write : "(in the Aerocom community) (Stier et al., 2013)"

Now: „Such experiments have also been used in the AeroCom community for a better understanding of reasons for model differences in aerosol forcing (...)“

- line 213 : "model-to-model diversity" : whereas on line 207, 208, and 210 "model diversity"

Thanks, now also here: „model diversity“ for brevity.

3.1 COMPUTATIONAL CAPACITY ABYSS

GENERAL : This section is not very nice to read. It is also reasonably long (4 pages is a considerable part of the manuscript). I list here detailed comments, but the section as a whole should be improved.

Thank you, we have now split this section into two subsections. The first one addresses tradeoffs across MIPs and the second one tradeoffs within MIPs.

- line 231-233 : "Simpler and multi-purpose experiments would be useful and less ..." : This sentence falls a little bit out of context. So maybe start with "Also ..." or "In addition to making Tiers, ...".

Agreed, we move the content of these sentences to the conclusion section where it fits better as recommendation for future MIPs.

- line 232-233 : "as long as they facilitate answering the science questions laid out by the MIPs." : I don't know whether this has to be mentioned so explicitly. I would think that the reader assumes this.

Yes, it might be clear for many, but we keep it just in case. The sentence is now included in the conclusions which possibly also less specialised scientists read.

- line 233-234 : "We propose ..." : I don't know whether this conclusion belongs already here.

Moved to conclusions.

. line 235-237 : "Simpler experiments ..." : This sentence is not very nice.

Now: „Nevertheless, multi-purpose experiments can be useful and less burdensome in terms of human and computational resources, as long as they facilitate answering the science questions laid out by the MIPs.“ In the conclusions.

- line 236 : the example seems to suggest that implementing a diagnostic is never a difficult task, although I think it sometimes can be.

A changed setting in the run script within physically sensible ranges does not require an implementation. Now: „in a setting for performing an experiment“ for clarity.

- line 237 : "Another example is an experiment design that needs to implement ..." : grammatically not ok. So maybe: "needs the implementation"

Thanks, changed as suggested: „needs the implementation of“

- line 240 : "and the associated scientific exploitation "I and "do" the scientific exploitation

Changed as suggested.

- line 241 : "run script" : not so nice language

It is a common term in some technical environments, but now: „in a setting for performing an experiment“

- line 242 : A rule of thumb ... : the message in this sentence is rather general, and I don't know whether it is useful.

Removed.

- line 244-245 : "One could say ..." : this is not a nice sentence

Changed to: „A greater number of experiments performed creates more data for statistical analyses and for addressing a variety of research questions, but it is taxing in light of restricted resources. “

- line 244-252 : This part of the text contains a lot of "we"/"our". I would try to limit that.

Agreed, we revised the paragraph accordingly.

- line 253 : we "review" ! we "reviewed"

Changed as suggested.

- line 253 : of "the" number of experiments ! of "a" number of experiments

Now: „current status of the experiments and their usage“

- line 256 : "Model output to assess differences in forcing and response was, however, more restrictive" : is this expressed correctly?

Revised to: „ Available model output to assess differences in forcing and response was, however, limited (...)“

- line 258 : emission-driven ! "SLCF" emission-driven models (to avoid confusion with CO2 emission driven models from C4MIP)

Added as suggested, but we now use the term NTCF throughout the manuscript for consistency with IPCC-AR6. We explain both terms in Section 2.1: „The term NTCF is used in IPCC-AR6 and refers to the same term as short-lived climate forcers (SLCFs) used by (Collins et al., 2017). Both NTCFs and SLCFs refer to radiatively active atmospheric constituents whose climate effects occur primarily within two decades of their emission or formation.“

- line 260-261 : in the "works" of the CMIP climate forcing task team ! "work"

Changed as suggested.

- line 267-268 : with 36 "experiments" ! with 36 "models"

Changed as suggested.

GENERAL line 271-275 : This should be improved.

- line 272-273 : Modelling centers perform the requested experiments "with the ESM which they support." : should this be mentioned explicitly? Isn't this assumed automatically?

The modelling centres do not necessarily own and develop the model on their own or perform the experiments for the scientists in all cases. Now revised to: „Modeling centers provide the resources for the requested experiments with the ESM which they support.“

- line 273 : They "contribute to the decision" ... : sounds strange. Isn't it that they actually decide? Scientists at or external to the modelling centre make a request for experiments to be performed. The modelling centres contribute to that decision since they provide the computing resources, decide who gets how many computing resources and/or what experiments are carried out in which order. We add this now: „They contribute to the decision for which community-driven MIP experiments with the ESM will be conducted, e.g., through granting computational resources and prioritizing experiments to be completed.“

- line 273 : "of" the ESM ! "with" the ESM

Changed as suggested.

- line 274-275 : "Not all ... giving modelers room to make their own choices." This seems to be a slightly different topic than the sentences around.

Now better linked to the previous sentences as follows: „Additional decisions for the experiments are made by the scientists interested in the MIP. There is some room to make their own choices since not all experimental settings are explicitly defined by the MIP's experiment protocols, e.g., they may use a coarser spatial resolution and to some degree less model complexity to reduce the computational burden.“

- line 275-276 : "Taken together, there are inevitable tradeoffs in the final experiment design." Not nice, very general.

Now: „There are inevitable tradeoffs in the final experimental designs for individual MIPs.“ To introduce the concept of the three axes.

GENERAL line 287-309 : This should be improved

- line 287 : "tradeoff" (is already mentioned on line 275)

Now here with a future perspective: „Although computing power continues to grow, tradeoffs along the three axes of experimental design and prioritizations will continue to be necessary. “

- line 290 : "priority" : this is a new aspect, which has not been mentioned before. As it falls a little bit out of the suggested framework, it should be mentioned in the beginning or at the end (not in the main part of the reasoning).

Now included in the first sentence of the paragraph already, again: „Although computing power continues to grow, tradeoffs along the three axes of experimental design and prioritizations will continue to be necessary.“

- line 297 : "restricts the scope for increasing computing resources" ! limits the attribution of computing resources

Changed as suggested.

- line 300 : "by" a certain degree ! "to" a certain degree

Changed as suggested.

- line 300-301 : "while retaining sufficient process detail for the scientific problem that is to be studied" : can maybe be skipped.

Removed as suggested.

- line 305-306 : "Similarly, ... is necessary." Why such a strong statement? Why not as in the former sentence : "... possible that are needed to split ..."

Now as in the previous sentence: „Similarly, a separation of the response in temperature or air quality into a forced signal and a contribution from internal variability is possible.“

- line 306 : "the required ensemble size" : possibly also length of the simulation falls under this category.

Yes, now added: „The required ensemble size and length of experiments (...)"

- line 310 : "data amount" is a derived product. I would think that the amount of simulated years is the determining factor.

Now revised: „necessary number of simulated years“.

- line 310 : "scientific interest" ! "research question"

Changed as suggested.

- line 311 : maybe add at the end of this sentence : "in the current setup/experiment"

Added as suggested: „in the current experiments“.

- line 314 : "exact precision" ! "accuracy"

Precision is the correct term. Precision means to have a sufficiently large number of measurements or in our case data from model experiments for calculating a representative mean value. That mean, however, might still be biased compared to the truth such that accuracy would be a misleading term.

- line 314-315 : "... depends on the model due to model differences in the internal variability ..." ! "... depends on the model's internal variability that induces ..."

Changed as suggested.

- line 318 : in "all" regions ! in "some" regions; or move it to the end of the sentence (then it can remain "in all regions")

Moved to the end as suggested.

- line 318-319 : the aerosol radiative effects : BC, OC, sulfate separately? Or together?

All together, changed to: „the anthropogenic aerosol radiative effects“

- line 322 : "For model responses" : vague

Revised: „For simulated climate responses, (,,)"

- line 322 : the ensemble sizes and lengths ! the ensemble sizes and "simulation" lengths

Changed as suggested.

- line 323-324 : content is very similar to line 326-328

Now combined and shorter: „For simulated climate responses, the ensemble sizes and simulation lengths of the experiments were not sufficient for addressing all research questions of interest in the three MIPs, especially for regional responses.“

- line 324-326 : very similar to line 310-314

Now include in line 310: „The signal-to-variability ratio is for instance sufficiently good for the response of the global mean of precipitation (Myhre et al., 2018, Allen et al., 2020) and the ERF in the global mean for most climate forcings in the current experiments. Specifically, (...)"

- line 328-330 : "Typically larger data amounts ... from the internal variability." : not a nice sentence

Removed.

GENERAL line 331-337 : vague, not well written. It should be improved.

Revised: „Complex models simulating many processes and their interactions are desirable and needed for specific research questions, and also pose challenges for reducing model-based uncertainty in the assessment of the climate response to various forcings.“

GENERAL line 338–358 : vague, not well-written.

Revised: „Model simulations ideally converge to similar solutions for a given question, e.g., how the Earth’s temperature responds to anthropogenic perturbations.“

- line 341-342 : ”emissions, transport, and deposition” : what about ”growth” and ”coagulation”?
Maybe use an expression which covers all the aerosol processing.

Revised: „models simulating the evolution of different aerosol species and their interactions“

- line 349 : ”had more participation” : not nice

Revised: „received more output from model experiments“

- line 349-251 : three times ”hence” in three lines.

Thanks, resolved.

- line 350 : ”had a different structural organization with formal experiment protocols.” This sounds strange. Might the following be better : ”had different structural organizations of their experimental protocols.”?

Changed to: „ (...) had different structural organizations while PDRMIP started earlier and was in comparison more self-organized and flexible in the MIP life cycle.“

- line 351 : ”dynamic” in the MIP life cycle : ”flexible”?

Changed to „flexible“

- line 352-354 : gives the impression that the emission-driven setup was the preferred one (mentioned first and using the expression ”had the capability”), whereas it was the prescribed concentration setup which actually was the preferred one.

Revised: „Specifically, some of the models in PDRMIP performed experiments with prescribed emissions whereas others used concentrations resulting in an ensemble of experiments partially driven by emissions and partially driven by concentrations of climate forcings.“

- line 354-357 : should be improved

Revised: „Yet, MIP experimental protocols do not prescribe the level of process complexity in and the resolution of ESMs. This freedom is well justified since ESMs might otherwise not be able to participate in a MIP if they can not fulfill stricter requirements. A wider participation of ESMs in MIPs ensures a sufficiently large multi-model ensemble needed to robustly quantify forcings and climate responses considering structural model differences.“

- line 357-358 : very general sentence

Revised: „A full exploration of the role of climate-composition feedbacks with focus on biogeochemical processes, however, remains an outstanding challenge due to this difficulty.“

3.2 PROCESS UNDERSTANDING ABYSS

- line 363 : ”Moreover” : Why adding ”moreover”?

Removed.

- line 366 : ... for all aspects. : this is vague.

Removed and point now to the example in the following sentence more explicitly: „(...) as one would hope. For example, (...)“

- line 366 : ”Clouds and circulation are for instance outstanding challenges” : maybe add ”Correctly representing ...” or ”Understanding ...”.

Added as suggested: „correctly representing clouds and circulation (...)“

- line 368-372 : "There is value ..." : not a nice sentence. Maybe just start with "Multi-model inter-comparisons shed light on ...".

Changed as suggested: „Multi-model inter-comparisons shed light on (...)"

- line 371 : which model "output" they ! which model they

Changed as suggested.

- line 372 : since not all "experiments" ! since not all "models"

Changed as suggested.

- line 372 : "e.g., some models might miss processes and interactions that might be crucial to address the research question." Isn't this already mentioned earlier?

Removed.

- line 373 : "Results of MIPs alone can not fully characterize the uncertainty ..." : Possibly start a new paragraph as this is a new thought.

Done.

- line 373-374 : "if it is at all possible" : I don't know if this general thought contributes something to the text.

Removed.

- line 376 : or "ideally" in synergy : I would skip "ideally"

Removed.

- line 380 : would allow us to quantify ... ! would allow us to "first" quantify

Added as suggested.

- line 392-394 : Aren't both sentences expressing the same?

Removed: „However, due to their potential magnitude, these methane feedbacks are important yet uncertain. “

GENERAL line 413-447 : not so nice to read. Please improve.

- line 413-414 : "satisfyingly" simulated : not nice

Changed to: „well simulated“

- line 416 : model consensus ! "erroneous" model consensus?

Yes, added: „erroneous model consensus“

- line 419-420 : no ESM in CMIP5 "and" CMIP6 : no ESM in CMIP5 "or" CMIP6

Changed as suggested.

- line 424-425 : "Winds emit" : not nice

Changed to: „ Winds control the emission and transport of desert-dust aerosols (...)"

- line 427-434 : vague, not nice

Changed to: „ Winds control the emission and transport of desert-dust aerosols and the soil erodibility is influenced by the available moisture from rain events. There is a large diversity in model-simulated regional changes in winds and precipitation in response to warming which in turn introduces uncertainty in simulated dust trends.“

4. OPPORTUNITIES

4.1 AUGMENTED ESMS

4.1.1 MACHINE LEARNING WHERE USEFUL

- line 497 : Training emulators require ! require(s)

Thanks for spotting this. Corrected.

- line 506 : ... remains. ! ... "however" remains.

Added as suggested.

4.1.2 KILOMETER-SCALE EXPERIMENTS

GENERAL : this section is not very well written. The different thoughts should be ordered better and presented more clearly, and repetition should be avoided.

- line 513 : Representing clouds ... ! Representing clouds "correctly" ...

Added as suggested.

- line 518 : "These meteorological variables" : a little bit unclear what it refers to, but the reader will probably assume "clouds and precipitation".

Changed to: „These processes (...)“

- line 519 : "their" : also not so clear what it refers to.

Changed to: „ and associated effects on the atmosphere“

- line 522 : "hold the potential for surprises in understanding climate responses" : the reader is left in doubt what to expect from this. So maybe be more specific.

Added: „ (...) e.g., with respect to future projections of temperatures and rare high-impact events (...)“.

- line 526 : "of a few weeks to years" : "years" (in plural) gives the impression that one can already do reasonably long simulations. Is that correct? Or should one write "weeks to months"?

It depends on the exact resolution and model, but global kilometre-scale experiments have already been done for a few years (Hohenegger et al., 2023). We add the citation as link to further detail.

- line 526-528 : "... in maintaining concentrated emissions, non-linearities in chemistry, and atmospheric transport of pollutants" : my impression is that the first two aspects implicitly assume and express a benefit from high resolution (maintaining strong gradients in concentration, and non-linearities can be resolved), whereas the 3rd aspect misses something, to what the transport improves : e.g., "and fine-scale resolved transport of pollutants"?

We added: „fronts in the atmospheric transport of pollutants“

- line 531-532 : strange sentence

Revised: „Global coupled atmosphere-ocean simulations with a few kilometers resolution can be done and progress has been made in incorporating some representation of atmospheric composition for example the carbon cycle (...)“

- line 531-534 : "... can be done." [in first sentence] "Such model experiments have been carried out ..." [second sentence]. That sounds contradictory between the first and second sentence.

Removed the second sentence.

- line 534 : "...to answer open questions concerning climate change to provide information for societal needs." Why is this rather general information mentioned here?

Removed.

- line 528-529 : "... to advance the understanding of climate and air quality interactions." This is very similar to line 535-536 : "... for the understanding of interactions of atmospheric composition and climate change."

Here, now removed for brevity.

- line 536-542 : "For some questions One question that can better be answered ..." : maybe use "Another question..."

Changed as suggested.

- line 543 : "will not be able to entirely rely on" : this is an understatement I think. With more nuance: "will not be able to mainly rely on"

Changed as suggested.

- line 544 : at a resolution 1 km ! at a resolution "of" 1 km

Added as suggested.

- line 547 : "The first possible and computationally smart way" ! "The first possible way"

Changed as suggested, although this should also be the computationally smarter way, at least for the time being.

- line 553 : our near-future "works" ! our near-future "work"

Changed as suggested.

- line 557-558 : "... UKESM2 and CESM aim also An ensemble of regional composition-climate models exists ...": I would not call a group of two models already an ensemble.

Changed to: „At least two different regional composition-climate models therefore (...)“

- line 557-562 : ... in "our" past MIPs. / As such "our" need for experiments ... / ... although "we" are not averse to the idea ... : I would avoid using too often "we" or "our".

Changed to: „As such a need for experiments with classical global ESMs is retained, at least for CMIP7, although moving towards global kilometer-scale modeling with a sufficient coupling of physical processes to aerosols and chemistry to address the community's research interests will be a goal to aspire to.“

4.2 IMPROVED DIAGNOSTICS AND ANALYSIS

4.2.1 RADIATIVE FORCING CALCULATIONS

- line 592-594 : "Differently from RFMIP, AerChemMIP found ... before calculating the ERF". This is already mentioned earlier, please indicate that.

Added the reference to the section in brackets.

4.2.2 SYNERGY WITH IMPACT ASSESSMENTS

- line 629 : "the opportunity" ! "an opportunity"

Changed as suggested.

- line 634 ; "... a smaller model spread in O₃". (too general) Shouldn't it be "... in tropospheric O₃ burden."?

The tropopause height would also affect the burden calculated for the stratosphere. We nevertheless added „tropospheric“ due to the relevance for air quality: „(...) smaller model spread in tropospheric O₃, which is relevant for air quality assessment.“

5. CONCLUSIONS

- line 657 : "(ESM)" ! "(ESMs)" (as on line 49)

Changed as suggested.

- line 669-671 : this sentence is grammatically a bit unbalanced : "(i) to speed up and improve ..., (ii) to data mining..., and (iii) to develop" Maybe add a verb to "(ii) to data mining" : "to do data mining".

Thanks, added „to do“ as suggested.

- line 674 : or many ensemble members ! or "sufficiently" many ensemble members

Changed as suggested.

- line 681 : closer collaborations with computer "science" ! with computer "scientists"

Changed as suggested.

- line 686 : "a" international vibrant community ! "an"

Changed as suggested.

- line 689-690 : The planning of ... and ... "are" currently ! "is"

Changed as suggested.

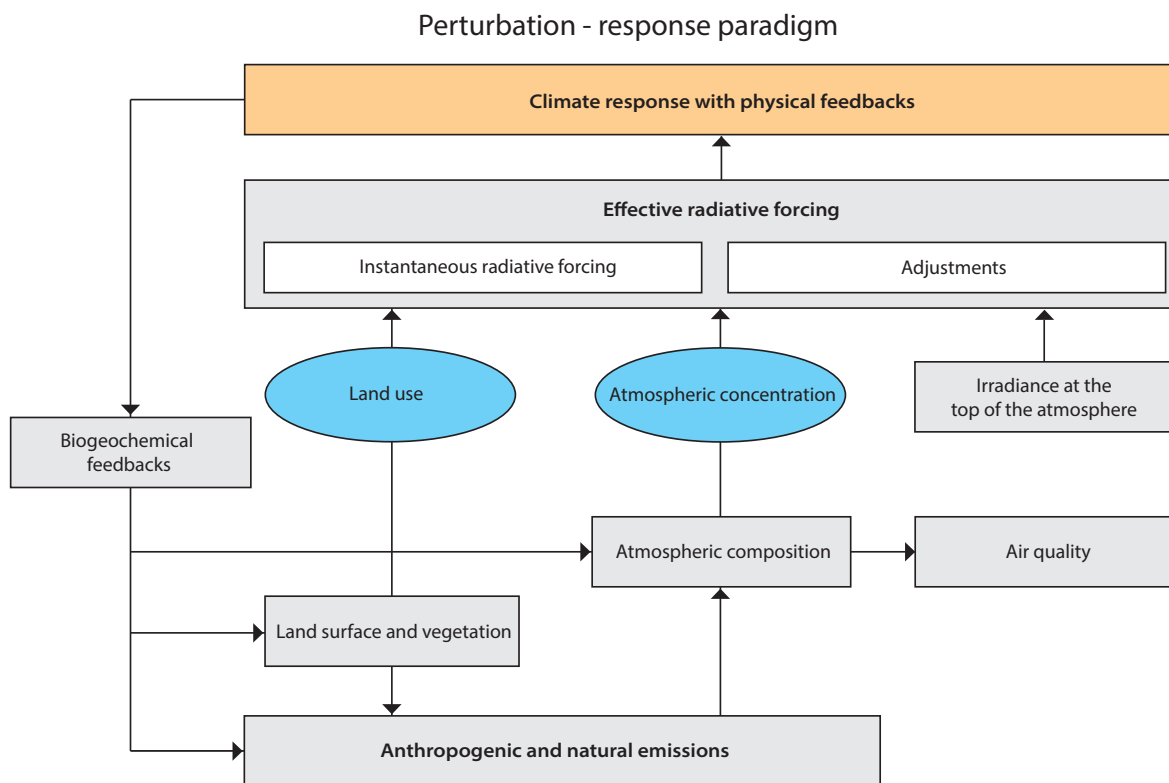
- line 690-693 : sentence sounds a bit strange

Changed to: „Keeping the two MIPs separate has advantages over connecting both initiatives in a single new MIP. The MIP names and the general ideas of RFMIP and AerChemMIP are already known through CMIP6. Moreover, the science questions and the associated experiment protocols are clearer, and the workload for coordination and management is smaller for the separate MIPs.“

FIGURES

Figure 1 : The drawing of the arrows should be improved : no intermediate arrows. Possibly the authors can modify the figure to even better represent the perturbation-response paradigm.

We have redrawn the figure. All arrows are now solid lines, and segments have been rearranged. The depiction has further been modified to include irradiance at the top of the atmosphere, air quality and physical feedbacks to better represent the paradigm.



Figure

3 : Frederiksen et al. [2021] : not in reference list. Possibly homogenize references with the main text (adding", " before year).

Added Frederiksen et al. (2021) in the reference list and adjusted the style of the references in Figure 3 as suggested.

TABLES

Table 1, caption : explain "X" (as it was done in the main text); maybe explain piClim-2xEmS (maybe refer to them also as piClim-2xX, but explain what it means).

Added in the caption: „Listed experiments are fully coupled atmosphere-ocean experiments (*hist-X*), experiments with prescribed transient changes in sea-surface temperatures and sea-ice from a historical experiment (*histSST-X*) and experiments with prescribed sea-surface temperatures and

sea-ice at pre-industrial level (*piClim-X*), where *X* refers to single or a combination of several climate forcers and *piClim-2xX* refers to experiments with prescribed doubled emission fluxes. “

Table 2, 3rd column : “little” change in aerosol forcing between 1970s and 2000s ! “Little” change

...

Changed as suggested.

Table 2, 4th column : O’Connor et al. (2021);” O’Connor et al. (2022) ! ”,”

Changed as suggested.

Table 3, 4th column : The numbers on the 3 lines containing “Total number coupled experiments”, “Total number histSST experiments”, and “Total piClim experiments” possibly do not make a lot of sense. I assume that they are obtained by summing the individual numbers, whereas the number of studies using at least one of the experiments is probably slightly lower.

Yes, we added the numbers from the individual entries to list totals, now explicitly stated in the table caption: „Totals are calculated by adding the individual numbers listed aloft and are generous estimates since some publications used more than one experiment type. “

Table 4 : and their role for climate “changes” ! “change”

Changed as suggested.

Table 4 : Aerosol absorption substantially “differ” across ESMs ! “differs”

Changed as suggested.

Table 6 : Improved diagnostic for “O3” ! “3” in O3 should be a subscript.

Changed as suggested.