

The experiment description paper *Reduced Complexity Model Intercomparison Project (Phase 1)* has changed in a fundamental way. Thank you for your effort. My point of view is external and I have to assess the present manuscript. Having a model RCM intercomparison is a great idea but the present manuscript does not meet the quality standards of GMD. I suggest a major revision. Otherwise, the authors should cancel the GMD publication process. In general, the authors should elaborate on the scientific goals and research questions that are associated with their intercomparison project. Next, the authors should elaborate on a consistent and unified experimental strategy. Finally, the authors should improve the format of the paper. To date, the present manuscript is a collection of interesting thoughts rather than a coherent text to describe a scientific idea.

The referees have given a variety of advices during the first phase of the review process. These advices are general comments on how to elaborate on the scientific goals or research questions that are associated with the model intercomparison project as well as specific comments on the wording in single sentences. Concerning the maturity of the present manuscript, I do not provide comments on single sentences or the wording which must improve, because I think the authors should rewrite or delete entire sections. At the same time, I am convinced of the scientific idea and think that a RCM intercomparison project is very valuable. In that respect, I would like to provide comments on every section.

The title of the paper is unspecific and I do not know what *Phase 1* actually means. The authors should introduce the experimental design and strategy of their RCM intercomparison project, and the title should be somehow related to this stage of development. The abstract is imprecise in the sense that the content of the abstract does not put forward the main messages of the main body of the manuscript. It is not about the experimental design and strategy. The content of the introduction should be related to intercomparison projects such as CMIP or scenario-MIP in order to establish common ground and explain why it is necessary to have a RCM intercomparison.

Section 2 is crucial and about the scientific focus of the RCM intercomparison project. However, it is unspecific and the authors should use common language such as *scientific goals* or associated *research questions*. I suggest that the authors spend some effort into specifying the research questions in order to highlight the actual variables or quantities that are evaluated. The RCM intercomparison should be consistent in the sense that the specific research questions and variables apply to the full range of RCMs considered here.

Section 3 is a mix of the organization of the RCM intercomparison project and the experimental strategy. In this connection, I do not think that the section title *simulation design* is appropriate. The authors should elaborate on section 3.1 *model configuration* and say in a direct way how the different RCMs compare and how the different RCMs are fitted to complex model output. I think having the equilibrium climate sensitivity tuned to 3°C is a good start. I would propose to focus on additional constraints such as changes in the energy budget if possible.

Section 3.2 is about the forcing that drives the temperature evolution of the different RCMs. It is a collection of different RCM drivers that can be associated with CMIP projects. I think a RCM intercomparison should be as simple as possible because of the great variety of RCMs. In that respect, the authors should establish common ground or common language and introduce the radiative forcing concept. I would propose to focus on CO₂ concentrations and emissions in the first place or select specific emission scenarios in order to make the RCM intercomparison tangible. Irrespective of the latter advice, the authors should explain why they use the different setups. The setups presented in this section should apply to the full range of RCMs considered here.

Section 3.3 and 4 is about the experimental design or organization. I think that sections on the input format, output specifications and data sources do not belong to the main body of the manuscript. They should be briefly described in the appendix. Moreover, the first RCM intercomparison should be limited to a small set of variables or quantities, and these variables should be common to the full range of RCMs. The authors should focus on the experimental strategy, and explain specifically why it is necessary to consider the idealized experiments and scenarios presented in this section. The experiments presented in this section should apply to the full range of RCMs. I would propose to focus on a set of experiments that are most important to the authors and generate the most important insights. Please also elaborate on the section on probabilistic outputs in case you still wish to include this section. It is unclear to me how these probabilistic ensembles are generated.

Section 5 presents illustrative results. A paper should be based on solid findings that emerge from a consistent and unified procedure. There are great figures. The experiment description paper should focus on the scientific goals, research questions and

experimental strategy. In that respect, the results should be based on the definition of specific research questions and the associated experimental strategy. I would propose to elaborate on the experimental strategy and present the most important results based on that experimental strategy. Furthermore, the results should be presented in an explicit way with respect to the research questions, and the results should be related to the full range of RCMs considered in the RCM intercomparison project. In this connection, section 6 raises different issues and does not relate the future research questions to the current experimental strategy or stage of development. Finally, the figures and tables of the appendix should be somehow related to the main body of the manuscript. A table which describes the different models and their structural differences is crucial.