

Interactive comment on “Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organisation” by V. Eyring et al.

V. Eyring et al.

veronika.eyring@dlr.de

Received and published: 11 March 2016

Reply to Anonymous Referee #3

We thank the reviewer for the helpful comments. We have now revised our manuscript in light of these and the other comments we have received. A pointwise reply is given below.

The manuscript is generally clear and well-written but it is difficult to do a proper review without reviewing the CMIP6 set-up itself. I appreciate that a

C4186

lot of thought went into designing CMIP6 through an open process, with the protocol being now largely frozen. The manuscript reflects choices that were made over the last couple of years, and as such I do not expect the authors to make significant changes to CMIP6 at this point. This said, some issues are significant and the authors may still be in a position to improve things and this manuscript is the place to clarify a few things.

I wonder if there is a risk that a long-standing DECK requirement for a standard pi-Control simulation (followed by a historical simulation branched at some point on it) perpetuates the myth that the pre-industrial 1850 climate was at equilibrium. Given the existence of low-frequency climate variability and low-frequency forcings, I don't think there is any such thing as an equilibrium pre-industrial climate. Model complexity increases and models will increasingly include slow components that respond on longer timescales such as permafrost, glaciers and ice sheet. Such model components will likely require long spinup, maybe in the context of realistic millennium simulation. Even though a piControl corresponds to current practices and is quite useful for the interpretation of many other simulations, the piControl + historical set-up may become a bit of a handicap over time and discourage novel (and possibly better ways) to spin up climate models, especially if the CMIP6 protocol is envisaged to last up to CMIP8 (figure 1). I appreciate there is a little bit of discussion on this in appendix A1.2, but I think it is insufficient, and I would be curious to know the authors' thoughts on this.

We added clarification on the *piControl* choice: “The pre-industrial control simulation (*piControl* or *esm-piControl*) is performed under conditions chosen to be representative of the period prior to the onset of large-scale industrialization with 1850 being the reference year.” We have also removed the sentence “The protocols for the DECK are expected to remain essentially unchanged for many years to come.” and

C4187

now include an additional paragraph that the DECK may slowly evolve over time when models are developed to include more complexity.

As indicated in another review of this manuscript, the CMIP6 panel and CMIP6 users should be prepared to a large number of submissions for the DECK with possibly many variations around a given parent model (in terms of resolution, choice of Earth system components, etc). This is not a problem in itself but raises the question of how to construct proper multi-model ensembles when a model flavor may be more represented than another model. Varying numbers of ensemble members across models and a large degree of sharing of some model components by participating modelling groups raise a similar issue. This is probably not for this manuscript to prescribe anything but it is an issue that could be flagged. Meanwhile some thinking may go on to see if it makes sense to form standardized sub-ensembles from all the model run submissions. One may also think of a procedure to flag obsolete model versions and model runs. In a continued process, some groups may want to flag explicitly what is their current “workhorse” model.

These are all good points. Flagging obsolete and the current “workhorse” model is something the WIP can discuss and we encourage you to review their contribution to this Special Issue. The question on proper multi-model ensembles and a possible weighting of a model ensemble based on model interdependence and / or model performance for a specific application is an open question of research that will be addressed scientifically by ScenarioMIP. We have highlighted this briefly in the first paragraph of the summary alongside with examples for other important science questions that were already mentioned.

The forcings used for the Historical simulation (that will be described elsewhere in the Special Issue) are expected to show an increasing level of details

C4188

and will be largely driven by observations. Some of them will include interannual variations (at least this is expected to be the case for tropospheric and stratospheric aerosols, maybe stratospheric ozone as well) which is in part due to the climate variability of the (real-world) historical period. Yet they will be prescribed in climate models that exhibit their own climate variability, which will not be in phase with the climate variability of the (realworld) historical period. This raises an issue on how to interpret the climate models and whether forcing terms should be smoothed out or not. I do not have the answer as what best should be done, but this article is the place where to mention this issue.

The forcings are provided by experts in a given resolution, but we leave it up to the modelling groups to do the level of smoothing. One aspect of having the CMIPX historical simulation is that CMIP in effect defines the state of our understanding of historical forcings, this is a contribution of CMIP to our understanding of the climate system that goes beyond the simulations. As to how to interpret and apply the CMIP-PhaseX forcings, this is something that modelling centres must do based on their own constraints, but an explicit statement has been added asking the centres to document how they apply a forcing.

Knowledge of radiative forcings (instantaneous and effective) is of paramount importance, but other reviewers have already given a rant on this, and I do not have much else to say. Personally I find it much more informative to know the climate sensitivity of a model in $K/(Wm^{-2})$ along with the CO_2 forcing than the climate sensitivity of a model in K for a doubling (or quadrupling CO_2). The DECK will only provide the latter. Surprisingly we seem to make collectively the same mistake CMIP after CMIP.

We agree and we encourage groups now to quantify the forcings, through participation in RFMIP-lite which was constructed specifically for this purpose. RFMIP was not

C4189

included in the DECK because the DECK serves other purposes, and because the success of the RFMIP-lite approach in quantifying the ERF has yet to be demonstrated.

I wonder how CMIP6 is going to police the submission to the DECK prior or simultaneously to a submission to a MIP (page 10559, lines 5-9). Will a modelling group need an authorization from the CMIP panel before submitting data to the ESGF for a MIP? Will the CMIP panel delegate this to the MIP chairs? Or should the system rely on self-policing? This article is the right place to elaborate the procedure but also what constitutes a new model or not (does a bug fix make a new model?).

This is a set of good questions, and we have added the following to the text to address these questions: "CMIP6 is a cooperative effort across the international climate modelling and climate science communities. The modelling groups have all been involved in the design and implementation of CMIP6, and thus have agreed to a set of best practices proposed for CMIP6. Those best practices include having the modelling groups submit to the ESGF the DECK experiments and the CMIP6 Historical Simulation as well as any MIP experiments they choose to run. Additionally, the modelling groups decide what constitutes a new model version. Past experience has indicated that the modelling groups are well aware that their model simulations are under considerable scrutiny. Therefore, we expect that CMIP6 will be similar to past CMIP phases in that the modelling groups will make a best faith effort to provide their highest quality model version that is distinguished from previous versions by substantive improvements in resolution, physics, or simulation skill. The CMIP Panel will work with the MIP co-chairs and the modelling groups to ensure that these best practices are followed."

It is good news that the CMIP6 data protocol follows closely that of CMIP5 so data users can harvest the benefits of their past investment. A few short-

C4190

comings would nevertheless need to be addressed. In particular the time structure of model output files is a nightmare in the CMIP5 archive and should be harmonized. Different models have different start dates and different ways to split their time series. Although some tools exist to make this somewhat transparent to users, CMIP6 would benefit a lot from prescribing this from the outset. I hope this is covered in the WIP manuscript.

Indeed issues like this will be addressed in the WIP contribution which we encourage you to review.

More fundamentally I suspect the CMIP5 archive not to be very friendly to so-called big data analysis. The meteorological and climate communities seem to be ignoring the issue, sometimes with arrogance (it is often heard "we've already been doing big data for years" when in fact big data is not about generating loads of data but more about new methods to extract information). I am not sure what to suggest but some thought could be given in CMIP6 on how to structure a fraction of the data or some diagnostics in a way that could facilitate the use of methods to extract information that climate scientists are generally now familiar with.

Again this is something the WIP could comment on in their contribution.

Other less significant comments

Page 10544, lines17-21: I appreciate that the new CMIP6 format with a DECK was intended to solve that issue, it has not been the case so far!

Yes we agree that this remains a challenge, and we thank the reviewer for supporting our effort to address this issue; let's hope it is successful.

C4191

Page 10548, line 15-16: atmosphere, land or their interactions.

Changed as suggested.

Page 10548, line 21 and elsewhere: historical simulation rather than Historical Simulation for consistency with e.g. piControl.

Changed as suggested

Page 10549, line 16: GHG, spell out.

Changed as suggested.

Page 10549, line 21: gradual should read gradually.

Changed as suggested.

Page 10550, line 4: is “challenges” the right word here? It could be interpreted as CMIP challenged models to reproduce the historical period, rather than to perform a historical simulation. Or do I misunderstand what the authors meant?

Sentence has been changed.

Page 10550, line 22: not so much the carbon cycle but the response of the carbon cycle to anthropogenic emissions of CO₂.

Changed as suggested.

C4192

Page 10552, line 14: why GC and not GSC? acronym is not used consistently throughout the manuscript. Do you really need it?

Now consistently used. The abbreviation ‘GC’ is kept since it is commonly used within WCRP.

Page 10557, line 22: the authors should elaborate on this number (is it compressed or not compressed data). Sounds small to me with the explosion of the MIP.

It is simply our best guess at this time and more details will be given in the WIP contribution.

Page 10557: carbon dioxide, no hyphen.

Changed to CO₂.

Page 10558, line 1: encoded? rooted maybe.

Changed to ‘stored’

Page 10562, line 20-21: but conversely one eliminates from the model the possible (but probably small) long-term trend in sea-level rise that existed in 1850.

Clarified by changing the bullet to: ‘Minimize artefacts in sea level change due to thermal expansion caused by unrealistic mismatches in conditions in the centennial-scale averaged forcings for the pre- and post-1850 periods. Any preindustrial

C4193

multi-centennial trend in global-mean sea level is most likely to be due to slow changes in ice-sheets, which aren't simulated in the CMIP6 model generation.'

Page 10563, lines 21-36: this is important information and should come in the main text.

We prefer to keep together in place the discussion on all the forcing details, including the volcanic forcing. We think it is appropriate to leave it in the Appendix.

Page 10566, line 26: "transients" should be "transient effects"

Changed as suggested.

Please expand Table A1 with more info (length of experiment, recommended ensemble size, etc).

Additional information is already given in Table 2.

Figure 4: here "Experiment" is preferred over "Simulation". Consider harmonize the two terms.

Changed to experiment here.

Interactive comment on Geosci. Model Dev. Discuss., 8, 10539, 2015.