

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

G. Schmidt (Referee)

gavin.a.schmidt@nasa.gov

Received and published: 30 December 2015

This paper describes the organization and choices being made in preparation for Phase 6 of the Coupled Model Intercomparison Project. CMIP has been an enormously successful set of projects, and yet has always failed to quite match the full expectations of the community. Thus at each stage, improvements and enhancements have been proposed, some of which have been implemented and some which have not. This iteration is no different, and the increasing scope of the proposals have necessitated a radical overhaul in the organization which is basically reported here.

The ambition of the project is commendable, but it is to be expected that implementa-

C3525

tion will inevitably fall short. Some of these issues are very predictable and I mention below a few of the ones I can foresee. The biggest problem is of course the reliance on ESGF for data delivery, of which more below.

DECK runs: These are a suitable 'entry card' into the process, and the requirements for new DECK entries for variations in physics, resolution, interactivity etc. is necessary (one run that is missing is perhaps a slab ocean equilibrium 2xCO₂ run for coherence with previous estimates of the ECS). However, there are some implications of the DECK/Historical approach that need to be addressed. Specifically, because this is a relatively low barrier to entry, more models and model versions will very likely be archived. Thus instead of ~60 individual model configurations as were available in CMIP3, there will likely be far more DECK entries over the lifetime of the CMIP6 program. I think this will be a good thing scientifically, but people should be ready for this.

AMIP: With the large changes in the Arctic over the AMIP period, particularly in ice thickness, modellers may need to start offering sea ice thickness as well as concentration as an input field. Has this been discussed/considered?

piControl: (p10563) specification of land use components (crops/pasture/irrigation) also have to be set to 1850 conditions. Background volcanic is best set to the mean the 1850-1900 period rather than the open-ended full period - since there is in fact a long term trend related to volcanic forcing (i.e. PAGES2K and associated papers).

Historical Simulation: The CMIP protocol should not be limiting forcings, or specifying what forcings groups have to use. Providing input to help groups without their own capacity to generate ozone datasets etc. is of course helpful, but since the historical runs are the most requested and the most likely to be compared to observations, groups must be free to choose to use their best estimates of all changes that they think important. For instance, orbital forcing is small, but to maintain coherence with past1000k runs, should be included. Irrigation, black carbon on snow, anthropogenic dust, direct

C3526

heating etc. might all be possible forcings next time around for certain groups and this should not be precluded from the design. Similarly, facilities must be made to allow for variations in forcing datasets as a function of real uncertainty, for instance in aerosol composition and distribution through time. The authors should explicitly acknowledge this here, and in the upcoming specific paper related to that experiment.

In the section A1.2, the authors call for a single 'HistoricalMisc'/DAMIP run to be done as well as the historical simulation ('Nat forcing only'). I don't disagree that this is useful, but it elevates the Tier 1 of DAMIP above all other MIPs, and I'm not sure that is sensible. (Additionally, why is this in the description of piControl and not in section A2?). If any MIP should be so elevated, it should be RFMIP (see below).

p10567 line 22. Is it not possible to move this to 2016?

p10568. Is there a recommendation for the interval to use between successive ICs? i.e 20 years? 30 years? The term 'longer' on p10567 is not well-defined.

MIPs: I strongly support the panel's decision to move towards a federation of the MIP organization since it draws in a far wider community of interested parties than just the modelling groups or the CMIP panel. But I am concerned about RFMIP being run as a separate project. One of the key missing analyses in CMIP5 was a coherent test of the forced response across the ensemble. This was hampered because while the specified input files or concentrations over time of atmospheric constituents were available, exactly what the forcing related to those changes was not. The forcing in any specific model depends on the radiative transfer code, the background climatology of water vapour and clouds, and on many model-specific indirect effects and the specific forcing definition. To my knowledge, only GISS have made available full radiative forcing diagnostics for their CMIP5 runs (both iRF and ERF) (Miller et al, 2014; Marvel et al, 2015) and given the importance of this for judging responses, this should be greatly extended in CMIP6. Thus of all the MIPs, RFMIP should be very tightly coordinated with the historical simulations, and indeed, the RF for every Historical run should be

C3527

archived as soon as possible afterwards.

Abrupt4xCO2: p10564 line 23. "effective" ECS, since it is demonstrated at least in some cases that the Gregory method is biased low relative the true ECS (i.e. Schmidt et al, 2014).

1%CO2: previous CMIPs called for stabilized versions (ie 1%CO2 until 2xCO2 or 4xCO2 and then constant thereafter). Has there been a specific decision to not do this? If so it should be stated.

Data requirements: Of the ~3PB in CMIP5, has the panel assessed the downloading and utilization of specific diagnostics? My sense is that while some diagnostics were very heavily used - surface fields, the historical simulations etc., there were many diagnostics that were requested that never got used, not even by the people who requested them in the first place. This might be because the package as a whole was not coherent (for instance the full energy budget) or ultimately, the diagnostic was too obscure or too difficult to compare across models. While it's hard to say that these lesser-used diagnostics will never be useful, the modeling groups would benefit from this ranking as they work to prepare the diagnostic packages for CMIP6.

ESGF: Much of the success of CMIP6 will be tied to the usability and accessibility of the ESGF. This paper takes it for granted that this will be available. Given the intermittent access over the last 6 months, the clunky interface, the notoriously difficult scripting options for systematic downloading, and general unhappiness in the wider community, does the panel want to address a backup option? i.e. a federated set of no-frills ftp sites - one per modeling group perhaps? Ideally, we should be discussing setting up intelligent data analysis sites that sit on top of the datasets to reduce the need for downloads, but I appreciate this goes beyond the scope of this paper.

Additionally, a vital improvement to CMIP and an accelerator for scientific discovery would be providing an archive for derived datasets, and perhaps even code for producing that derived data. Examples would be indices such as global mean temperatures,

C3528

NAO indices, NINO3.4, Max Atl. Overturning, forward modeled brightness temperatures (for MSU + SSU satellite observations), ocean heat content anomalies, etc. I have long pushed for this to be part of ESGF, but this has not happened for a variety of reasons. The CMIP panel however and the authors here should be at the forefront of making this work, and this paper would be a good place to describe their initiatives and aims in this direction.

DOIs: To document the impacts of CMIP6, we should be ensuring that a) every simulation has a doi for the package of diagnostics at the time of deposit, and b) every paper should have a data table listing the doi's used. This will allow forward referencing for every group and simulation, allowing for much improved accountability and feedback. This did not work at all in CMIP5 (because the unscalable bottleneck of individual file-level 'quality control' was (IMHO) a disaster) and we should be ensuring that this does not happen again. This has to be built in to the design explicitly. The only mention of DOI's in the section on p10568 for the forcing datasets and not the simulations which I find very odd. This has to be made explicit right from the get-go and it has to be explicit that this will be a 'on-release' system (as opposed to a 'post QC' system in CMIP5). (Note, if the authors for whatever reason get hung up on the nature of a 'doi' for the simulation package, please replace this acronym with an identifier of their choice that is digital and refers to an object).

Minor edits:

p10541 line 20: will depend on THEIR scientific interests

p10541 line 25: INTERNAL climate variability

p10542 line 9: central element → central INPUT

p10544 "In addition, a monolithic structure to the CMIP design tended to discourage the modelling centres from attempting to design new experiments meant to address specific scientific questions of interest to them." - this might be better phrased as a

C3529

reflection of some peoples opinions, rather than an absolute truth. From our point of view, we did not feel inhibited from expanding the scope of CMIP5 experiments (via HistoricalMisc, different 'physics-versions', forcings etc.) and exploring our scientific interests.

"This in turn contributed to the impression that CMIP was a service that the modelling centres provided to the broader community." - there are many reasons why the interaction is not two-way and there are a number of issues that could be proposed to deal with that. In my opinion, it has very little to do with the monolithic structure, and far more to do with the inability to track where output is used, a lack of archiving possibilities for derived data and code, and a traditional publication schedule that is so long that it makes many analyses obsolete before they are even available.

p10544 "Third, the punctuated structure of CMIP has begun to distort the model development process. Whereas in the past modelling centres developed models based on their own scientific goals and released model versions on their own schedule, the visibility and demands of CMIP were beginning to impose a synchronization of model development with different phases of CMIP." - this is strangely phrased. It is clear that there is a synchronisation (i.e. it hasn't just begun). Indeed, it has been this way since CMIP3. I don't see why this is considered a problem though. Indeed, without external deadlines, I fear models would almost never be released. Frankly this just seems like some people in the community are whining and it detracts from the paper.

p10548 line 23. "the signal FROM THE forced responses (Li et al., 2015)." - Note here that 'forced responses' in AMIP includes forcing by SST/SIC in addition to the external forcing. The authors should be clear the term is being used differently here than elsewhere in the paper.

p10548 line 27. A word perhaps about what 'pre-industrial' means. It is not the same in this context as zero anthropogenic influence. GHGs, LU etc. are all already modified in 1850. There is ongoing discussion about defining it to be in the late 18th Century as

C3530

well - but presumably CMIP is not going to move the start date for the historical runs back to 1750 to account for this.

p10549 line 1: "External human influences on the land surface are likewise excluded. " This cannot be true. You would have a shock to the runs if you had zero LU difference in the piControl and then suddenly jumped to 1850 conditions in the historical transient. Presumably, the authors simply mean that further transient changes to LU are not made in the piControl runs.

p10550 line 5. We should already be aiming to have 2015 forcing included, and for this to be updated on an annual basis.

p10556 line 12. use 'evaluation of the predictions' instead of 'verification of the models'

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