Geosci. Model Dev. Discuss., 8, C4224–C4225, 2016
www.geosci-model-dev-discuss.net/8/C4224/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific Model Development Discussions

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 Julia Hargreaves (Editor)

I called quite a few reviewers for this paper, as I thought it would be good to have a wide range of opinions, and we obtained 3 official reviews in the end. I am very glad to see that so many have added their own comments. The MetOffice are to be particularly congratulated for their industry in this regard! Looking through the comments I do not see any that stand out as obviously foolish. Rather I see many points raised that I am curious to see the response to. So please respond

to them all, point-by-point in the same manner as the official reviews.

There are quite a lot of details but nothing of major concern so please just let us know if you need extra time to do the revision. I look forward to receiving your revised manuscript in due course.

On the topic of the code/data availability section, the issue with the positioning of the section has now been resolved with Copernicus. Please move the section back to the original position, directly following the Conclusions and before the Appendices.

We thank the editor for finding three good reviewers and for the prompt guidance during the submission of our manuscript. The reviews and short comments are very constructive indeed. We have now revised our manuscript in light of all comments we have received. A pointwise reply is given in our responses to each review and short comment.

Thanks also for the clarification on the 'Data Availability' section. As said in our response to Astrid Kerkweg's comment, this section was correctly placed in our submission, but it has been mistakenly moved by the editorial office during the typesetting. It is now in its original position, directly following the Conclusions and before the Appendix.

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

Geosci. Model Dev. Discuss., 8, C3427–C3428, 2015
www.geosci-model-dev-discuss.net/8/C3427/2015/
© Author(s) 2015. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific Model Development Discussions

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: 21 December 2015

Dear Astrid,

In response to your comment:

"Dear authors, In my role as executive editor I ask you to move the Data Availability Section to its regular place after the conclusion but in front of the Appendix when revising your article. Thanks, Astrid Kerkweg"

The Data Availability Section was placed as you request after the Conclusions and before the Appendix and Acknowledgements in the version we have submitted, and

therefore followed the guidelines that are given on the GMD website and that we have received from our editor:

"Inclusion of Code and/or data availability sections is mandatory for all papers and should be located at the end of the article, after the conclusions, and before any appendices or acknowledgments."

However, this section has then been moved by the editorial office during the typesetting. I had asked to move it back to where it was when I have received the proofs, but have received the following response from the editorial office:

"Please note that according to our house standard, the section "Data availability" remains unnumbered and is placed after the Appendix and before Supplement, Author contribution, and Acknowledgements."

This differs from the guidelines on the GMD website. Please clarify this within the editorial office and provide guidance where to place the Data Availability Section in the final version.

Thanks and best wishes, Veronika Eyring

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

Geosci. Model Dev. Discuss., 8, C4169–C4181, 2016
www.geosci-model-dev-discuss.net/8/C4169/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific Model Development Discussions

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 Gavin Schmidt (Referee #1)

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.

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 implementation 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?

We agree that some models now might make use of sea ice thickness. To our knowledge, however, there is not a single standard sea ice thickness monthly mean data set agreed upon by the international community as a standard. There is a “synthetic” sea ice thickness data set that has been used in the past, but it is not based on any direct observations (see http://www-pcmdi.llnl.gov/projects/amip/RESOURCES/synice.php). Until a “reference” dataset has been developed, we do not expect to supply sea ice thickness for AMIP simulations.

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).

Land use: The sentence 'Unless indicated otherwise (e.g., the background volcanic forcing), experiment conditions should be representative of Earth ca. 1850.' is implying that this should be done for all forcings. However, we now explicitly specify the land use component in a separate bullet: land use is set to 1850 conditions but land use changes are not included in *piControl*.

Background volcanic: The average volcanic RF over the CMIP historical period is quite close to the average over the last millennium in the forcing dataset of Crowley (2000); they are -0.22 Wm^{-2} and -0.18 Wm^{-2} respectively (Gregory et al., 2013). Of course there is uncertainty about this but the similarity means that it is reasonable to use the historical period average in *piControl* on the assumption that it is typical. Given the uncertainty in the observations this assumption seems okay and it has the advantage of being well-defined, implying that the model will by design not have a volcanic spin-up drift in ocean heat content during the historical period. This can also be tested by running historicalNat (assuming that solar forcing has a smaller long-term effect) or historicalVol (even better, but not part of anyone's Tier 1). We therefore stick to the historical mean period, but have added additional explanation to the text.

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 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.

The forcing datasets for the DECK and the CMIP6 historical simulations will indeed be made available through the ESGF for common use in the simulations. One reason to do this is obviously that it is convenient for the participating modelling groups that datasets are provided by the corresponding experts since it is a lot of work to produce forcings datasets. We don't believe it is helpful to purposefully conflate uncertainty in the specification of forcing with uncertainty in the response to a given forcing, and rather encourage groups to sample the latter through supplementary simulations.

It is true that in the future other forcings might be required. But it could also be the other way round, some of the forcings that are provided as boundary condition to the runs might disappear since they are simulated interactively by the models. In any case, any deviations from the specified forcings as well as any additional forcings used in the models should be documented, as we state in the manuscript.

The last point on uncertainty will be addressed by some of the MIPs (e.g. DAMIP and DCPP). We have also asked the groups who produce the forcings to discuss uncertainty in the forcing datasets in their contributions to the GMD Special Issue.

The discussion on page 10547 has been changed and extended in response to this comment.

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).

We discuss this here because it directly relates to the choice of volcanic forcing in the control run. We have reworded the paragraph and now the reference to DAMIP is just informational and the explanation of the value to understanding the *piControl* and historical runs is the focus.

In Section 3 we have now expanded the discussion of the problem in CMIP5 that forcing was not well quantified and we now specifically encourage modeling groups to do the most important of the RFMIP experiments.

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

We have extensively discussed this with all groups involved, in particular those that produce the forcing datasets mostly in form of voluntary and unfunded work. It

is already a success that they can provide forcings extending until the end of 2014, made available early in 2016. Forcings for the CMIP6 historical simulations are as much as possible based on observations. In the CMIP6 timeline, DECK and the CMIP6 historical simulations can be run in 2016 (see Figure 4). It will take substantial time to update the forcings to end of 2015 which would cause a significant delay of CMIP6 from the start. It would also delay the harmonization with the future scenarios that takes several months, so this would just mean that the entire CMIP6 process is substantially delayed. Therefore no, this cannot be moved. However, several groups are working towards updating their forcings on a more regular basis, and modelling groups might get updates during the next coming years to extend the historical simulation. But the "start of the future in CMIP6" will remain to be 2015 and the period from 2015 through near-present will be referred to as "historical-extension" as said in the manuscript.

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.

The statement as stands indicates that the larger the interval between ICs the more independent the resulting simulations will be. It is difficult to say much beyond this. There are many considerations, both practical and scientific, that might affect the interval. If, for example, a model has a strong 50 yr oscillation in the AMOC then using a 50 yr interval could lead to problems. It is therefore hard to give a general recommendation but we request that groups document what they did and why and have added this recommendation to the text.

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 archived as soon as possible afterwards.

We had extensive discussions with the modelling groups and MIP co-chairs on the design of CMIP6, and in particular on the aspect which simulations to include in the DECK. In the interest of keeping the DECK small and restricted to those simulations the modelling groups do anyway as part of their model development cycles, the four DECK simulations were selected, and it was decided to have all other experiments except the CMIP6 historical simulations in one of the CMIP6-Endorsed MIPs. We agree that the quantification of forcings and feedbacks in RFMIP (and also AerChemMIP) is essential, which is reflected in the CMIP6 design since it is one of the three main CMIP6 science questions. In our revised version, we encourage the participation in RFMIP-lite to fill gaps identified in CMIP5.

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).

'effective' has been added.

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.

Most important for analysis of the 1pctCO2 experiment is that the 1% CO_2 increase is continued until at least CO_2 quadrupling after 140 years. We are not aware that the stabilization period after 4xCO2 has been analysed and have decided to simplify the experiment to avoid mistakes in the setup of the simulation. The text has been extended to clarify this change.

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.

The way the CMIP6 Data Request is created differs from how this was done in CMIP5. All CMIP6-Endorsed MIPs were asked to specify the variables they need for their own simulations, from other MIP simulations, and from the DECK and CMIP6 historical simulations. The MIPs also commit to analyse the data they request which will avoid that large amount of data that is requested will never be looked at.

ESGF: Much of the success of CMIP6 will be tied to the usability and ac-

cessibility 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.

We agree that the shut-down of CMIP5 ESGF for an extended period was unacceptable. As you note, this has little to do with the experiment design for CMIP6, which is the focus of this paper. The ESGF will be described and discussed further in the WIP contribution to this Special Issue.

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, 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.

Indeed many users would benefit from enhanced ESGF capabilities to perform server side calculations for derived variables such as global mean temperatures and precipitation, zonal or annual means. This was mentioned also in the CMIP5 survey, has been encouraged during WGCM meetings, and is something that the ESGF teams will try to establish. Details on what actually is planned to achieve for CMIP6 will be

given in the WIP contribution to this Special Issue. In addition, evaluation tools that will be installed at some ESGF nodes can also be used to compute derived variables or indices alongside the ESGF, and their output could be provided back to the distributed ESGF archive. We have added one sentence at the end of Section 3.3. that addresses this.

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 filelevel '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 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).

Good point, thanks. Again this will be further detailed in the WIP contribution, but we have added a paragraph making a few points about this to the 'Data Availability' section.

Minor edits:

p10541 line 20: will depend on THEIR scientific interests

Changed as suggested.

p10541 line 25: INTERNAL climate variability

Changed as suggested, here and in abstract.

p10542 line 9: central element -> central INPUT

Changed as suggested.

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 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.

This section has been considerably revised in light of the reviewers comments.

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.

See previous comment.

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.

We have removed the explicit mentioning of forcings in this sentence and replaced it with 'to improve the signal to noise ratio'.

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

This has been clarified in the revised version.

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.

Thanks for spotting this. LU is set to 1850 conditions as specified now in the appendix. Land use changes are however excluded in *piControl*. The sentence has been deleted to avoid misinterpretation.

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

See response to your comment on p10567 line 22 above.

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

Changed as suggested.

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

Geosci. Model Dev. Discuss., 8, C4182–C4185, 2016
www.geosci-model-dev-discuss.net/8/C4182/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific
Model Development
Discussions

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 #2

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.

This paper describes the changes made for the next stage of the CMIP, CMIP6 and the rationale behind the changes. In CMIP6 there is clear "minimum

entry" of basic runs used to calibrate modes (the" DECK" and a simulation with historical forcings) to participate in the rest of CMIPs. The many MIPs have been rationalised by choosing those which map most clearly onto the WCRP "Grand Challenges "(Giving a simpler structure which should also enable a much more efficient use of resources).

Main Comments The paper is generally well written, informative and clear. Given that one of the stated main aims if CMIP6 is to understand "How does the earth system respond to radiative forcing", I would expect the intercomparison of radiative forcing to be part of the minimum requirement, rather than a separate MIP (RFMIP). In CMIP5 there is wide range of estimates of aerosol forcing in particular over the historical period (even when nominally identical aerosol forcing agents are included) which makes evaluating the relative importance of aerosols problematic. How can one understand the response of models to changes in radiative forcing when a substantial part of the forcing has not been calculated adequately?

See response to a similar comment from Gavin Schmidt. We are now encouraging the modelling groups to participate in RFMIP-lite, but have not changed the DECK which has been agreed with the community in extensive discussions.

Line16 The punctuation (distortion?) of model development cycles may have more to do with the cycle of IPCC Assessments rather than CMIP- I suspect some modelling centres feel under pressure to produce updated models for each Assessment. Aligning CMIP with IPCC assessments, whether deliberately or accidently may re-enforce this pressure.

This section has been considerably revised and somewhat streamlined in response to the reviewers comments. We hope that doing so helps address the reviewers point;

suffice to say we wanted to avoid getting into a discussion as to what extent CMIP timelines are implicitly set by expectations related to IPCC reports.

Minor comments

Can a model which has appeared in CMIP5 and submits the correct DECK and Historical runs is eligible for CMIP6.

YES.

(It is mandatory like the DECK experiments Why is the Historical control not part of the DECK?)

This has been clarified in the text in the second paragraph in Section 3.

Why does the data section on data come after the summary? It should either be part of the main text before the summary. Or in an Appendix

There are specific rules where to place the 'Data availability' section in a GMD paper. We have followed the GMD requirements but the editorial office moved it during the proof stage (see our response to the Executive Editor's comment by Astrid Kerkweg). This should now be solved (see our response to the Editor's comment by Julia Hargreaves).

Line 129 It would be useful to have a bit more general information on the MIPs. Is there the opportunity to introduce further MIPS should the need arise during the lifetime of CMIP6? Are there any criteria for winding up MIPs? Will MIPs time expire or at least be reviewed at the end of CMIP6?

This has been added.

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

Geosci. Model Dev. Discuss., 8, C4186–C4194, 2016
www.geosci-model-dev-discuss.net/8/C4186/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific Model Development Discussions
Open Access
EGU

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

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 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

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 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 ESFG 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-

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.

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 CO_2.

Changed as suggested.

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 CO_2.

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 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.

Geosci. Model Dev. Discuss., 8, C4195–C4199, 2016
www.geosci-model-dev-discuss.net/8/C4195/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific Model Development Discussions

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 Chris Jones

Thanks Chris 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.

[note - as a co-chair of C4MIP this has a carbon-cycle centric view of the world, but these are my own comments and not an "official" C4MIP contribution.]

Overall a good overview and description of CMIP/CMIP6 and DECK runs. The experiment descriptions and rationale were clear. Just a few minor comments which hopefully will be helpful.

My main comment is that I had feared some confusion around the need for emissionsdriven PictI and historical runs if you are running emissions-driven runs in C4MIP. I think you handle this well, but you could add a couple more specific details, and also site the C4MIP documentation paper (which will be Jones et al 2016 GMD – currently in prep). A couple of things are worth being explicit:

We have added a few more details, see also below. We are not referring to any specific CMIP6-Endorsed MIP paper since they are currently still in preparation or just submitted. However, we make clear that the MIPs are detailed in these papers and have done so for C4MIP here.

- for both control and hist, the C-driven and E-driven variants should be identical in all forcings except the treatment of CO_2. This sounds obvious but worth saying it clearly.

This has been added for clarification.

- The name "emissions driven control run" may be confusing as in fact there are no emissions. I can't think of a better name, so you should explain that what this means is:
a) atmospheric CO_2 concentration evolves prognostically in response to natural land and ocean carbon fluxes
b) external input of CO_2 from either fossil fuel or land use is prescribed to be

zero

c) CO_2 is therefore free to evolve and should be stable in the long term but will have some internal variability. In the C4MIP paper we define a desirable level of drift in the control run as within 5 ppm per century in the atmospheric CO_2. We recommend for spin-up that the concentration-driven control run is spun-up first and then can be used as a start point for any final spin-up in emissions-driven configuration. It might be useful to include this level of technical detail in the Appendix on the control run.

We have added this clarification at the end of Section A1.2 and refer to the C4MIP contribution for further details.

A few minor comments:

- p. 10549, line 1. As per Gavin's review, land-use is not excluded from Pictl, but it is held fixed at 1850 so there is no land use CHANGE. 1850-level crop, pasture and management activities etc should be held fixed so there is no long term change in any land surface properties

Sentence deleted and land use further specified in the appendix.

- p.10549, line 16, and later in the appendix p.10565 line 6. What month is recommended for the quadrupling? January? You mention an ensemble of runs starting at different dates – how much were these looked at last time? I know there is some dependence possibly on the date of quadrupling so presumably all models should choose the same date. If you request an ensemble, can you specify what you want? 12 runs from a different month, 4 runs quarterly?

Sentence changed to 'In the first, the CO_2 concentration is immediately and

abruptly quadrupled from January 1850 values.' to be more precise.

We have added a sentence at the end of section A1.3 recommending that the start-months be spaced evenly throughout the year for ensembles.

- p. 10556, line 15. Better to say "isolating" than "looking". Land-use was included as a forcing in CMIP5 and there are papers which look at it. Rather, LUMIP is the first time a set of experiments have been designed around this as a focal point.

Changed as suggested.

- p. 10559. lines 19-21. A1.1 AMIP. You say that having land carbon diagnostics on for the AMIP run will be valuable for evaluation as the surface climate will be closer to observed. I think we need some caution here as the carbon cycle will be dependent on how it is initialised at 1979. The long memory of land carbon stores mean it is out of equilibrium by then, so can't just be spun up be repeating the 1970s over a few times. You certainly don't want to repeat a single year multiple times as suggest on line 27. May be OK for atmospheric variables but not land carbon. The carbon fluxes will be dependent on the carbon stores and the only robust way to initialise these (obs don't exist at the scale we need) would be to take from a historical run. In general, if we want a more realistic surface climate to drive the land-carbon fluxes a much better option is offline land runs (as per the TRENDY activity, and LS3MIP) – here land models can be run with observed meteorology but run for the full 20th century so don't have the initialisation problem in 1979. So overall I'm not convinced evaluation of land carbon from AMIP will be that useful.

We have removed the sentence 'This will enable evaluation of the carbon cycle component of the model when climate conditions are more similar to the observed than in coupled atmosphere-ocean simulations.' in response to this comment.

- p.10566, line 5. instead of "fluxes" can you say "fluxes and stores". The carbon pools are actually more important (IMHO) than the fluxes, but always get sidelined...

Changed as suggested.

- p. 10567, line 20. Can you mention harmonisation of the forcing data between historical and future scenarios? Where will this be fully documented? In the forcings paper? In ScenarioMIP? It's one of those essential cross-cutting things which everyone might leave to someone else...

The harmonization is considered in the timeline to produce the forcings for CMIP6. The harmonization is now added as suggested on page 10547 where forcings are first mentioned.

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

Geosci. Model Dev. Discuss., 8, C4200–C4206, 2016
www.geosci-model-dev-discuss.net/8/C4200/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific Model Development Discussions

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 Gareth Jones

Thanks Gareth 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 paper is a clear and concise description of the CMIP6 design. It is extremely useful to modelling centres to have such a document to describe the

experimental set up of the DECK and historical simulations expected for CMIP6.

Below are some comments that I hope the authors will consider and find helpful.

Section 3 Why the historical (1850 to present) simulation is not part of the DECK is not explained. It has been a curious discrepancy that I have not seen an adequate explanation for. Given the AMIP experiment is supposed to have the same radiative forcing factors as the historical experiment it would seem obvious to include the historical simulation within DECK too. Apart from simplifying the design, it would make it less cumbersome in talks/reports/papers/documentation for modellers and users to just call it 'DECK'.

The text has been expanded (2nd paragraph in Section 3) to better explain this, see also our response to a similar comment by Referee #2.

Page 10546, Lines 21-25 For the cases of models in CO2 emission mode, won't there also be a requirement for there to be a piControl that is in CO2 concentration mode, to enable analysis of the amip, 1ptco2 and abrupt4xco2 experiments?

This is already implicit in the CMIP design which asks for the DECK being performed for all model configurations (see page 10546).

Page 10547 Line 27 to Page 10548 Line 1 The consistency of experiments over time will not strictly hold for amip. The forcing factors, and to a lesser extent the SSTs, will change from CMIPx to CMIPy. It would be helpful to clarify this.

This is true and is strictly speaking also true for the *piControl*. But our leading argument what to include in the DECK is that these simulations are routinely done and possibly repeated many times over a long period of model development by the groups with whatever forcings are available. We still want the AMIP simulation that is then submitted for a specific phase to use the exact forcings of that phase to be compliant with the historical simulation, see also our response to the above comment.

Page 10549 Line 13 It would be really helpful to have a minimum length for the *piControl*. Elsewhere it is recommended to have the same length as whatever experiments are spun off, but that could mean for a model just providing the DECK (and historical experiment) the piControl may be only a 160 years long.

We agree and have deleted the sentence 'The length of this 'spin-up' period is model and resource dependent." since a minimum length is specified in the Appendix.

Page 10550, Lines 9-13 It would be a worrying precedent to promote the use of the historical simulations as a tool to constrain the uncertainties in the forcing factors that are then put in the models. As historical simulations are used in model validation (Page 10566 Lines 12-14) such a policy will introduce circular reasoning [Rodhe, 2000]. Such circular reasoning will reduce the usefulness of the models in understanding past climate changes and also reduce confidence in projections.

The sentence on page 10566 has been reworded.

Additionally the authors are being over enthusiastic when they say that historical simulations have "proven essential in reducing uncertainty in radiative forcing associated with short lived species such as the atmospheric aerosol" Apart from the inherent circular reasoning, many aerosol climate

scientists may strongly dispute how 'essential' this is.

Sentence removed.

Page 10563 Lines 12-14 It will be obvious to many, but not all, that the implementation of a background volcanic aerosol in piControl has implications for historical and historical natural experiments. During periods of no volcanic activity the radiative forcing from volcanoes, with respect to the control, will be positive. This has been surprising to some in the past, so it might be useful to highlight this.

We have revised our discussion of volcanic forcing following the bulleted summary of forcing specifications. A paragraph is included explaining the rationale for imposing a background aerosol in the *piControl*.

Page 10563 Lines 21-26 "Modelling groups are urged to perform this experiment ... as doing so will most effectively separate the role of natural vs. anthropogenic drivers of climate change and variability since 1850." If it is so important it obviously makes sense to make the natural experiment part of the DECKplus (DECK + historical) suite. Asking modellers to do one experiment for a MIP (in this case DAMIP) but not the other experiments within a MIP is confusing.

We do not think that encouraging modelling groups to run a specific CMIP6-Endorsed MIP experiment in this paper justifies inclusion of this experiment in the DECK. We simply encourage modelling groups to perform this experiment because then they will be able to determine whether in their model the specification of control-run volcanic forcing leads to an artefact in sea level changes during the historical period. We have modified the text somewhat to explain this better.

Additionally has there been any consideration to the usefulness of doing just one natural simulation? The signal to noise is usually very low for many climate diagnostics from a simulation driven with natural forcings. One ensemble member will not be helpful for the interpretation of the results of a single model. However, it may be useful as part of a multi-model analysis.

These questions will be addressed by DAMIP.

Page 10564 Lines 6-7 This is an important point. More guidance about what to do would be helpful here. For instance HadGEM2-ES gradually ramped up the aerosol in the stratosphere for the RCPs [Fig 14 in Jones 2011], but the way this was done may not be the best way. If all the models reintroduce the volcanic aerosols in the scenarios in a similar way, it could introduce a signal to a multi-model mean which may be misinterpreted. Given relatively small volcanic eruptions possible role in reducing short temperature trends, how future background aerosols are re-introduced should be discussed.

Since we are not discussing the forcings beyond the DECK and the CMIP6 historical simulation here, we refer to the ScenarioMIP contribution and the volcanic aerosol forcing paper in this Special Issue, where the treatment of volcanic aerosol in the future scenarios will be specified. But we have clarified, also in response to the comment by Alistair, that a non-zero value will be used.

Page 10566 Lines 12-14 The definition here is a little worrying. "The CMIP6 historical simulation is meant to reproduce observed climate and climate change". That implies the simulations have to be tuned to match the observations. Rather the historical simulations are designed to be compared with observations for various hypothesis testing

Agreed and sentence changed.

Page 10567 Lines 10-18 It is useful that a "historical-extension" is recommended, but it would seem logical to include it as part of DECKplus. Where would it be allocated to if not? Documentation about what was forcing the extension runs - which was largely missing from CMIP5 - will be crucial, but an agreed scenario for the historical-extension would be even more helpful.

The idea is that the extension would be based on observations, not a scenario. It is impractical with current resources to promise that a standard post 2014 observational dataset will be made available, so we'll have to live with documented non-standard datasets. This extension is considered included in the DECK, but the details on how it will be handled by the infrastructure are still under discussion, so we can only indicate at this time that the runs will be named *historical-ext* or *esm-hist-ext* (for the emissions-driven run).

Page 10568 Lines 9-11 It seems curious that a requirement from DAMIP is singled out here (see also Table 2). If extra ensemble members are really so strongly required outside of DAMIP, it would seem logical to include them in the DECKplus.

We do not think that encouraging modelling groups to run a specific CMIP6-Endorsed MIP experiment in this paper justifies inclusion of this experiment in the DECK. We simply encourage modelling groups to perform additional ensemble members and not pointing to DAMIP that has the additional ensemble members in their Tier 1 would unnecessarily remove existing information. A model group could still run ensemble members and submit them to the ESGF even if it doesn't participate in DAMIP.

A final comment. A central reference point, with DOI, that lists all the models

and subsequent links to their documentations would be a great addition to CMIP6. The increasing demands on authors and the increasing size of CMIP6 would be countered somewhat by not having the requirement to reference every model in every paper.

The WIP is considering different formats and different granularities to provide DOI's, which are still under discussion, so we are unable to say anything more specific at this time.

References

Jones, C.D. et al., The HadGEM2-ES implementation of CMIP5 centennial simulations, Geosci. Model Dev., 4,543-570, 2011 Rodhe, H.; Charlson, R.J. and Anderson, T.L., Avoiding circular logic in climate modeling, Climatic Change, 44, 419-422, 2000.

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

Geosci. Model Dev. Discuss., 8, C4207–C4210, 2016
www.geosci-model-dev-discuss.net/8/C4207/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific Model Development Discussions

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 Alistair Sellar

Thanks Alistair 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.

Most of these comments are concerned with making the design specification more explicit and removing ambiguity which could lead to avoidable

differences between models. Thanks to Martin Andrews for one of the comments.

A general point for all experiments is to state the start and end date, not just the year. Should all experiments begin on 1st January of a given year? All experiments should continue until at least December of the end year. A lack of clarity here caused problems for HadGEM2 data submissions in CMIP5.

Thanks for pointing this out. We have added a sentence to the caption of Table 2 "All experiments are started on 1 January and end at 31 December of the specified years." that clarifies this point.

p10550 line 18 "...a prescribed CO_2 concentration and a prescribed emissions simulation (accounting explicitly for fossil fuel combustion), in which concentrations are then "predicted" by the model...." For absolute clarity (though it makes it more repetitive), it would be worth inserting "CO_2" again before "emissions" and "concentrations" and stating that the treatment of other GHGs should be identical in both simulations.

Changed as suggested.

p10561 line 8: "The it piControl used in CMIP begins at this point and generally continues for at least a few hundred years." There is a more precise statement about the it piControl length on the following page. Suggest dropping the second half of this sentence.

Sentence deleted.

p10563 line 2: "With that understanding, here are the recommendations for

the imposed conditions on the it piControl " Suggest "spin-up and it piControl".

Changed as suggested.

p10563 line 15: "Models without interactive ozone chemistry should specify ozone as in the mean of the first decade of the CMIP Historical Simulation" Suggest aligning the ozone meaning period with the solar meaning period for consistency, even if it doesn't make a significant difference.

The Chemistry-Climate Model Initiative (CCMI) indeed followed as similar approach as that suggested by the reviewer. The text is now more explicit: 'Models without interactive ozone chemistry should specify the pre-industrial ozone fields from the corresponding dataset, which is produced from a pre-industrial control simulation that uses 1850 emissions and a mean solar forcing averaged over solar cycles 8-10, representative of the mean mid-19th century solar forcing."

p10564 line 1: "The forcing specified in the it piControl also has implications for simulations of the future, when solar variability and volcanic activity will continue to exist, but at unknown levels. These issues need to be borne in mind when designing and evaluating future scenarios, as a failure to include volcanic forcing in the future will cause future warming and sea-level rise to be over-estimated relative to a it piControl experiment in which a non-zero volcanic forcing is specified. This could be addressed by re-introducing the mean volcanic forcing for the it piControl into the scenarios." I would make this statement stronger and state at the very least that ScenarioMIP *will* use a time-constant non-zero volcanic forcing. I presume that this known already.

Changed as suggested.

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

Geosci. Model Dev. Discuss., 8, C4211–C4213, 2016
www.geosci-model-dev-discuss.net/8/C4211/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific
Model Development
Discussions
Open Access
EGU

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 Helene Hewitt

The comments below reflect the views of the CLIVAR Ocean Model Development Panel (OMDP). There is general agreement among the panel that spin-up is an issue which requires further attention. While analysis of spin-up does not usually appear in peer-reviewed publications, being able to analyse the spin-up enables a deeper understanding of the simulations and may lead to improved guidance on spin-up for the future. Archiving the model spin-up could

also be important for comparison with HighResMIP where control and transient experiments are likely to be run close to the model initial conditions.

The paper states (p10549, line 5) that the spin-up 'is usually performed and discarded'. In the discussion of *piControl* (p10561, line 7), it is stated that 'the length of the spin-up period should be documented'. OMDP suggest that:

-Text on p10549 is modified to read 'is usually performed, with the length of this 'spinup' period being model and resource-dependent. It is recommended that the spin-up length and procedures are documented and the spin-up archived as far as possible.'

-Text on p10561 is amended to read 'This spin-up period can be as long as several hundred years. Spin-up length and procedures (coupling/forced, bug corrections, retuning, etc) should be documented and the spin-up archived as far as possible'

Regarding archiving requirements, we suggest that the spin-up should be archived with decadal means of 3D ocean temperature, salinity, carbon-related quantities, 2D surface fluxes and zonally integrated ocean heat and freshwater transports. Archiving these fields would be sufficient to enable a better understanding of climate drift.

Helene Hewitt on behalf of CLIVAR OMDP

Thanks Helene and Ocean Model Development Panel for the helpful comments. We have changed the text accordingly and the WIP will define specific names that clearly mark this.

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

Geosci. Model Dev. Discuss., 8, C4214–C4221, 2016
www.geosci-model-dev-discuss.net/8/C4214/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific Model Development Discussions

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 Elena Shevliakova

Thanks Elena 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.

These comments are concerned with the CMIP6 description of ESM experiments, particularly a requirement of "constant" land use in the pre-industrial

control (*piControl*) experiment and its implications for the ESM historical experiment initialization. Echoing a comment by the reviewer #3 (regarding perpetuating a myth of an equilibrium preindustrial 1850 physical climate), we would like to raise a concern that the proposed *piControl* for the CO_2-emission-driven ESMs and inferred fluxes in CO_2-concentration-driven ESMs will promote an unjustified assumption, that the pre-industrial 1850 carbon system, particularly land, was in equilibrium.

It appears that the manuscript focuses mainly on the CO_2-concentration driven experiments under the assumption that the additional CO_2-emission-driven *piControl* and historical experiments will require a simple switch from CO_2 concentration to CO_2 emissions as is typically done with short-lived atmospheric species. The DECK experiments, including those with ESMs, are expected to remain unchanged in future CMIPs. However, the experimental design of the CMIP6 ESM spin-ups, controls, and historical runs have a number of specific challenges, often not considered in AOGCMs experiments, and which are not discussed in the manuscript. I hope that the authors could expand and comment on such challenges in the manuscript.

Specific comments

1. The authors could clarify that the state of land carbon in the ESM *piControl* (i.e. equilibrated climate and carbon cycle) ignores the long term impacts of several centuries of secular change in vegetation and soil carbon storage and sustained land-use carbon emissions prior to 1850 due to land use practices such as clearing of primary lands for croplands and pasture, shifting cultivation, logging, fuel-wood extraction, and associated regrowth.

Previous CMIPs have initialized historical AOGCM simulations from a pre-

industrial control. In this manuscript the authors describe the period chosen as ". . .prior to the onset of the large-scale industrialization. . ." with ". . .no secular changes in forcing, so the concentrations/or sources of atmospheric constituents (e.g. GHGs and other forcing) are held fixed. . ." (p 10548, l26-29). Unlike emissions of fossil fuels and cement production, CO_2 emissions from agricultural activities, biomass burning, and wood harvesting were not fixed and were not small prior to 1850, particularly in the Northern hemisphere. It's well established in the literature that in the 1850s the land was a sustain source of carbon (0.6 PgC/yr) based on a number of modeling approaches (Houghton 2010). This imbalance is well outside of the proposed "equilibrium" tolerance of 0.1 PgC/yr in the comment by Chris Jones, and thus represents an inconsistency for initialization of historical runs.

Furthermore there was secular changes in both agricultural expansion and in the amount of wood harvested for fuel and logging and those trends are documented in CMIP5 land-use change reconstruction for 1500-2005 (Hurtt et al, 2011). While one could argue that the implications of such changes (i.e. biophysical feedback, mostly from agricultural conversion) on the physical climate was small globally, these changes have major carbon cycle implications for vegetation, litter and soil carbon storage which were not in equilibrium before the onset of the industrial revolution in the 1850s. Therefore, it is only in a highly idealized context that one can interpret the state of land carbon from an ESM control run with constant land forcing as pre-industrial (before 1850).

The reference year for the piContol was discussed in detail at the WGCM meetings and the conclusion after long community consultation was not to move backward in time for computational reasons at this stage. Rather it was decided that single model simulations should be used to further quantify the omission of pre-1850 CO_2 emissions from land use and land use changes on the historical and present-day climate. Results

from such studies could be used to give further guidance on the interpretation of the CMIP DECK and CMIP6 historical simulations and further recommendations for future efforts could be given. CMIP6 includes many different research topics and many model groups with different scientific foci. Compromises like this have to be made. To consider this comment, we have extended the text and now further address the caveats of a 1850 *piControl* choice explicitly, also for ESMs.

2. Because i) a large diversity in implementation of ESMs' land components, including land use and its interactions with carbon components, and ii) a lack of detailed analysis of how such differences may affect initialization and evolution of global carbon cycling in historical simulations, the authors should add discussion about whether the historical ESM experiment initialization should or should not follow the AOGCMs' practice of initializing historical runs from a proposed ESM equilibrated control and how much flexibility modeling centers may have in deviating from that AOGCM practice. Our experience is that a discontinuity in land carbon between a control and historical simulation is necessary to "bridge" an idealized control to accurate historical carbon cycle evolution (Sentman et al., 2009). How a modeling center could document such discontinuity and how to archive possible "bridge" experiments, could be added in the discussion section.

A number of publications (Hoffman et al 2013, Brovkin et al 2013, Jones et al 2013) show that the CMIP5 ESMs are dramatically diverse in their implementations of vegetation dynamics, soil biogeochemistry and, particularly, land use and management components. Most ESMs in CMIP5 have ignored harvesting of wood and shifting cultivation, which have been shown to play a significant role in altering natural forest dynamics, forest age structure, and carbon uptake on time scales from decades to centuries (Houghton, 2010). Some CMIP5 ESMs included crops as plant functional types, others pastures as plant functional

types, and a few treated their carbon dynamics differently from natural grasslands (i.e., no harvesting or grazing). Some ESMs transferred harvested or cleared carbon from agricultural practices to the atmosphere directly, others deposited cleared or harvested carbon to anthropogenic pools, with the release time scales varying from a year to a decade or century. Still others returned harvested or cleared carbon directly to soils.

In CMIP5, differences in vegetation dynamics and land use models also led to diverse practices in ESMs' spin-ups. Examples include a fixed crop/pasture fraction from a dataset of choice, with or without wood harvesting, or potential vegetation without any land use. A similar variety of strategies was used in controls or idealized experiments. We do not expect that land models, including implementation of land use, will be less diverse in CMIP6 and future CMIPs. All such model differences have implications for a) how much carbon an ESM is going to converge to in an equilibrated state, including vegetation, litter and soil carbon pools, b) for how these pools are going to respond to warming in idealized experiments and to the atmospheric CO_2 increases, and, importantly, c) for historical simulations to be compared with observations.

These are all good points which should be considered in the analysis and interpretation of the ESM *piControl* and *historical*. We are however not commenting here on the diversity of the ESMs that we expect in CMIP6. This is something that will be described in the model documentation papers. We have noted the caveat that 1850 is not strictly speaking pre-industrial condition (see also our response above).

There are then indeed no secular changes in the *piControl* because the forcings are held constant (for AOGCMs and ESMs). To consider this comment, we have added a paragraph at the end of Section A1.2 that specifies what is done for the *piControl* for ESMs in CO_2-emission driven mode: external input of CO_2 from either

fossil fuel or land use is prescribed to be zero in the *piControl*, to be able to remove the model drift that arises from the uptake of CO_2 by the ocean and land even in the absence of CO_2 emissions.

This then indeed creates a small discontinuity in land carbon between the control and the historical simulation when moving from 1850 to 1851 and throughout the historical simulation within the lifetime of this effect which we now state in Section A2. In Section A2 we now also comment on strategies to account for the fact that land-surface was not in equilibrium in 1850 ("bridging" experiments). Due to the wide diversity of modelling approaches for land carbon in the ESMs, the actual method applied by each group to account for these effects will differ so cannot be further specified here. But we request that it needs to be well documented.

3. As Gavin Schmidt already pointed out in his review, a clarification would be helpful about how to interpret "constant" land for *piControl*, particularly in CO_2 emission-driven simulations, especially how to implement "constant" wood harvesting and shifting cultivation (secondary lands in Hurtt et al 2011). As more models attempt to capture forest age structure distribution, which is important for simulating the rate of carbon uptake, it's not clear how to initialize that age distribution from the PI control. In a concentration-driven AOGCMs the implications of "constant" land use for physical climate are different than the implications of the same treatment of land use in an ESM for carbon storage or for simulated atmospheric CO_2 concentration. As land components of ESMs are still rapidly changing and implications of a particular "constant" land–use treatment for the DECK and historical experiments are not clear, we suggest that modeling centers should have some flexibility in how to interpret "constant" land use in AOGCM and ESMs (including an option of not having any land use in control), as long as their documentation manuscripts report clearly details of spin-up and PI controls and any extra experiments they may do for initialization

of idealized or historical experiments.

We now give a more detailed recommendation for the treatment of land-use in the *piControl*. The land use forcing dataset will be described in the corresponding contribution to this Special Issue. We also clearly state that any deviations from the recommendations given here need to be documented, so this point is covered already.

4. A minor point, it would be helpful if the manuscript defined which models are AOGCMs and which are ESMs for the purpose of the CMIP6, as there are many definitions of ESMs in the literature. Section 3 opens by stating that the DECK comprises four base experiments plus *historical*. It would be helpful to clarify from the beginning that the above applies only to AOGCMs. For ESMs, the DECK comprises 5 experiments and 2 historical simulations.

ESM now defined and where required, the discussion has been extended with explicit mentioning that applies to CO_2-emission driven simulations.

The way we have defined this is that the DECK and the historical simulation need to be performed with all model configurations, so the number of simulations is implicit. C4MIP includes both CO_2-concentration and -emission driven experiments, so that participating in C4MIP requires repeating the DECK and the CMIP6 historical simulations with both model configurations. We stick to this definition and have not changed the text.

Brovkin, Victor, et al. "Effect of anthropogenic land-use and land-cover changes on climate and land carbon storage in CMIP5 projections for the twenty-first century." Journal of Climate 26.18 (2013): 6859-6881.

Jones, Chris, et al. "Twenty-first-century compatible CO_2 emissions and

airborne fraction simulated by CMIP5 earth system models under four representative concentration pathways." Journal of Climate 26.13 (2013): 4398-4413.

Sentman, Lori T., et al. "Time scales of terrestrial carbon response related to landuse application: Implications for initializing an Earth system model." Earth Interactions 15.30 (2011): 1-16.

Hurtt, G. C., et al. "Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands." Climatic Change 109.1-2 (2011): 117-161.

Hoffman, Forrest M., et al. "Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models." Journal of Geophysical Research: Biogeosciences 119.2 (2014): 141-162.

Houghton, Richard A. "How well do we know the flux of CO_2 from landarea use change?". Tellus B 62.5 (2010): 337-351.

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

Geosci. Model Dev. Discuss., 8, C4222–C4223, 2016
www.geosci-model-dev-discuss.net/8/C4222/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific Model Development Discussions

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 Martine Michou

I would recommend to update Figure 4 (and also the text of the manuscript) as DECK and historical experiments cannot start before the final design is out, together with the release of all CMIP6 forcing fields and of the final data request (mid 2016?).

Thanks Martine for pointing this out. The figure on the timeline has been slightly

changed to reflect that the *piControl* can be started only end of April 2016 due to the small delay in the provision of the forcings and the creation of the CMIP6 Data Request compared to the original schedule. We have changed 'Around April 2016 the forcings for the historical simulations should be ready' to 'By May 2016 the forcings for the DECK and the CMIP6 historical simulation should be ready' to make the text more precise.

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

Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organisation

V. Eyring[1], S. Bony[2], G. A. Meehl[3], C. Senior[4], B. Stevens[5], R. J. Stouffer[6], and K. E. Taylor[7]

[1] Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany

[2] Laboratoire de Météorologie Dynamique, Institut Pierre Simon Laplace (LMD/IPSL), CNRS, Université Pierre et Marie Curie, Paris, France

[3] National Center for Atmospheric Research (NCAR), Boulder, USA

[4] Met Office Hadley Centre, Exeter, UK.

[5] Max-Planck-Institute for Meteorology, Hamburg, Germany

[6] Geophysical Fluid Dynamics Laboratory/NOAA, Princeton, NJ, USA

[7] Program for Climate Model Diagnosis and Intercomparison (PCMDI), Lawrence Livermore National Laboratory, Livermore, CA, USA

Abstract. By coordinating the design and distribution of global climate model simulations of the past, current and future climate, the Coupled Model Intercomparison Project (CMIP) has become one of the foundational elements of climate science. However, the need to address an ever-expanding range of scientific questions arising from more and more research communities has made it necessary to revise the organization of CMIP. After a long and wide community consultation, a new and more federated structure has been put in place. It consists of three major elements: (1) a handful of common experiments, the DECK (Diagnostic, Evaluation and Characterization of Klima ~~experiments~~) and ~~the~~ CMIP ~~Historical Simulation~~historical simulations (1850 – near-present) that will maintain continuity and help document basic characteristics of models across different phases of CMIP, (2) common standards, coordination, infrastructure and documentation that will facilitate the distribution of model outputs and the characterization of the model ensemble, and (3) an ensemble of CMIP-Endorsed Model Intercomparison Projects (MIPs) that will be specific to a particular phase of CMIP (now CMIP6) and that will build on the DECK and ~~the~~ CMIP ~~Historical Simulation~~historical simulations to address a large range of specific questions and fill the scientific gaps of the previous CMIP phases. The DECK and CMIP ~~Historical Simulation~~historical simulations, together with the use of CMIP data standards, will be the entry cards for models participating in CMIP. The participation in the CMIP6-Endorsed MIPs will be at the discretion of the modelling groups, and will depend on their scientific interests and priorities. With the Grand Science Challenges of the World Climate Research Programme (WCRP) as its scientific backdrop, CMIP6 will address three broad questions: (i) How does the Earth system respond to forcing?, (ii) What are the origins and consequences of systematic model biases?, and (iii) How can we assess future climate changes given internal climate variability, predictability and uncertainties in scenarios? This CMIP6 overview paper presents the background and rationale for the new structure of CMIP, provides a detailed description of the DECK and ~~the~~ CMIP6 ~~Historical Simulation~~historical simulations, and includes a brief introduction to the 21 CMIP6-Endorsed MIPs.

1. Introduction

The Coupled Model Intercomparison Project (CMIP) organized under the auspices of the World Climate Research Programme's (WCRP) Working Group on Coupled Modelling (WGCM) started twenty years ago as a comparison of a handful of early global coupled climate models performing experiments using atmosphere models coupled to a dynamic ocean, a simple land surface, and thermodynamic sea ice ~~(Meehl et al., 1997)~~(Meehl et al., 1997). It has since evolved over five phases into a major international multi-model research activity ~~(Meehl et al., 2000; Meehl et al., 2007; Taylor et al., 2012)~~(Meehl et al., 2000; Meehl et al., 2007; Taylor et al., 2012) that has not only introduced a new era to climate science research, but has also become a central element of national and international assessments of climate change ~~(IPCC, 2013).~~(e.g., IPCC (2013)). An important part of CMIP is to make the multi-model output publically available in a standardized format for analysis by the wider climate community and users. The standardization of the model output in a specified format, and the collection, archival, and access of the model output through the Earth System Grid Federation (ESGF) data replication centres have facilitated multi-model analyses.

The objective of CMIP is to better understand past, present and future climate change arising from natural, unforced variability or in response to changes in radiative forcings in a multi-model context. Its increasing importance and scope is a tremendous success story, but this very success poses challenges for all involved. Coordination of the project has become more complex as CMIP includes more models with more processes all applied to a wider range of questions. To meet this new interest and to address a wide variety of science questions from more and more scientific research communities, reflecting the expanding scope of comprehensive modelling in climate science, has put pressure on CMIP to become larger and more extensive. Consequently, there has been an explosion in the diversity and volume of requested CMIP output from an increasing number of experiments causing challenges for CMIP's technical infrastructure ~~(Williams et al., 2015)~~(Williams et al., 2015). Cultural and organizational challenges also arise from the tension between expectations that modelling centres deliver multiple model experiments to CMIP yet at the same time advance basic research in climate science.

In response to these challenges, we have adopted a more federated structure for the sixth phase of CMIP (i.e., CMIP6) and subsequent phases. Whereas past phases of CMIP were usually described through a single overview paper, reflecting a centralized and relatively compact CMIP structure, this GMD Special Issue describes the new design and organization of CMIP, the suite of experiments, and its forcings, in a series of invited contributions. In this paper, we provide the overview and backdrop of the new CMIP structure as well as the main scientific foci that CMIP6 will address. We begin by describing the new organizational form for CMIP and the pressures that it was designed to alleviate (Section 2). It also contains a description of a small set of simulations for CMIP which are intended to be common to all participating models (Section 3), details of which are provided in an Appendix. We then present a brief overview of CMIP6 ~~serving~~that serves as an introduction to the other contributions to this Special Issue (Section 4~~)~~), and we close with a summary.

2. CMIP design - a more continuous and distributed organization

In preparing for ~~the sixth phase of CMIP~~CMIP6, the CMIP Panel (the authors of this paper), which traditionally has the responsibility for direct coordination and oversight of CMIP, initiated ~~an extensive~~a two year process of community consultation~~ that spanned two years.~~. This consultation ~~not only~~ involved the modelling centres whose contributions form the substance of CMIP~~, but also the~~ as well as communities that rely on CMIP model output for their work. ~~This consultation included the organization of special~~Special meetings were organized to reflect on the ~~success and science~~ successes of CMIP5 as well as the scientific gaps that remain or have since emerged~~ from CMIP5. It~~. The consultation also sought input through a community survey, the scientific results of which are described by ~~Stouffer et al. (2015). This process identified four~~Stouffer et al. (2015)[1]. Four main issues related to the overall structure of CMIP were identified.

~~First, there was increasing recognition that results from different CMIP phases should not necessarily be used in isolation. For instance, as pointed out by Rauser et al. (2014), a common set of simulations across different phases of CMIP could be used to construct more rational ensembles than just a single phase ensemble.~~First, we identified a growing appreciation of the scientific potential to use results across different CMIP phases. Such approaches however require an appropriate experimental design to facilitate the identification of an ensemble of models with particular properties drawn from different phases of CMIP (e.g., Rauser et al. (2014)). At the same time it was recognized that an increasing number of Model Intercomparison Projects (MIPs) were being organized independent of CMIP, the data structure and output requirements were often inconsistent, and the relationship between the models used in the various MIPs was often difficult to determine, in which context measures to help establish continuity across MIPs or phases of CMIP would also be welcome.

[1] Stouffer, R. J., Eyring, V., Meehl, G. A., Bony, S., Senior, C., Stevens, B., and Taylor, K. E.: CMIP5 Scientific Gaps and Recommendations for CMIP6, BAMS, submitted, 2015.

Second, the scope of CMIP was taxing the resources of modelling centres making it impossible for many to consider contributing to all the proposed experiments. ~~At the same time, relatively little guidance was available~~By providing a better basis to help modelling centres decide exactly which subset of experiments to perform~~. This led to a more~~ it was thought that it might be possible to minimize fragmented participation in ~~CMIP5, as compared to earlier phases and created the impression that perhaps modelling centre resources were being used sub-optimally. In addition, a monolithic structure to the CMIP design tended to discourage the~~ CMIP6. A more federated experimental protocol could also encourage modelling centres ~~from attempting to design new experiments meant to address specific scientific questions of interest to them. This in turn contributed to the impression that CMIP was a service that the modelling centres provided to the broader community~~to develop intercomparison studies based on their own strategic goals.

Third, some centres expressed the view that the punctuated structure of CMIP ~~has~~had begun to distort the model development process. ~~Whereas in the past~~Defining a protocol that allowed 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~~ to decouple their model development ~~with different phases of CMIP. Though this might have seemed desirable to those who relied on the different phases of CMIP to measure and pace progress in model development, such a view poorly reflected the reality of model development. The resulting stress this placed on~~from the CMIP schedule would offer additional flexibility, and perhaps encourage modelling centres ~~often resulted in unnecessary delays in the provision of model output, as it was conditioned on the finalization of new model releases and risked redefining the role of modelling centres in ways that did not best exploit~~to finalize their ~~scientific potential.~~models and submit some of their results sooner on their own schedule.

Fourth and finally, ~~there was the~~many groups expressed a desire for particular phases of CMIP to be more than just a collection of MIPs, but rather to reflect the strategic goals of the climate science community, as ~~developed through the~~for instance articulated by WCRP. By focusing a particular phase of CMIP around ~~a few~~specific scientific issues, it was felt that the modelling ~~centres~~resources could ~~better contribute~~be more effectively applied to ~~selected~~those scientific questions that had matured to a point where coordinated activities ~~could~~were expected to have substantial impact~~, thereby more rapidly advancing the science and promoting a cohesive strategy across WCRP~~.

A variety of mechanisms were proposed and intensely debated to address these issues. The outcome of these discussions is embodied in the new CMIP structure, which has three major components.

First, the identification of a handful of common experiments, the DECK (Diagnostic, Evaluation and Characterization of Klima~~ experiments~~) and ~~the ~~CMIP ~~Historical Simulation~~historical simulations, which can be used to establish model characteristics and serves as its "entry card" for participating in one of CMIP's phases or in other MIPs organized between CMIP phases, as depicted in Fig. 1. Second, common standards, coordination, infrastructure and documentation that facilitate the distribution of model outputs and the characterization of the model ensemble, and third, the adoption of a more federated structure, building on more autonomous CMIP-Endorsed MIPs.

Realising the idea of a particular phase of CMIP being centred on a collection of more autonomous MIPs required the development of procedures for soliciting and evaluating MIPs~~,~~ in light of the scientific focus chosen for CMIP6. These procedures were developed and implemented by the CMIP Panel. The responses to the CMIP5 survey helped inform a series of workshops and resulted in a draft experiment design for CMIP6. This initial design for CMIP6 was published in early 2014 ~~(Meehl et al., 2014)~~(Meehl et al., 2014) and was open for comments from the wider community until mid-September 2014. In parallel to the open review of the design, the CMIP Panel distributed an open call for proposals for MIPs in April 2014. These proposals were broadly reviewed within WCRP with the goal to encourage and enhance synergies among the different MIPs, to avoid overlapping experiments, to fill gaps, and to help ensure that the WCRP Grand Science Challenges would be addressed. Revised MIP proposals were requested and evaluated by the CMIP Panel in summer 2015. The selection of MIPs was based on the CMIP Panel's evaluation of ten endorsement criteria (Table 1). To ensure community engagement, an important criterion was that enough modelling groups (at least eight) were willing to perform all of the MIP's highest priority (Tier 1) experiments and providing all the requested diagnostics needed to answer at least one of its leading science questions. For each of the selected CMIP6-Endorsed MIPs it turned out that at least ten modelling groups indicated their intent to participate in at least Tier 1 experiments, thus attesting to the wide appeal and level of science interest from the climate modelling community.

3. The DECK and CMIP ~~Historical Simulation~~historical simulations

The DECK comprises four baseline experiments: (a) a historical Atmospheric Model Intercomparison Project (~~*AMIP*~~*amip*) simulation, (b) a pre-industrial control simulation (*piControl* or *esm-piControl*), (c) a simulation forced by an abrupt quadrupling of CO_2 (~~*abrupt4xCO2*~~*abrupt-4xCO2*) and (d) a simulation forced by a 1% yr^{-1} CO_2 increase (*1pctCO2*). CMIP also includes a ~~Historical Simulation~~historical simulation (*historical* or *esm-hist*) that spans the period of extensive instrumental temperature measurements from 1850 to the present. In naming the experiments, we distinguish between simulations with CO_2 concentrations calculated and anthropogenic sources of CO_2 prescribed (*esm-picontrol* and *esm-hist*) and simulations with prescribed CO_2 concentrations (all others). Hereafter, models that can calculate atmospheric CO_2 concentration and account for the fluxes of CO_2 between the atmosphere, the ocean, and biosphere are referred to as Earth System Models (ESMs).

The experiments chosen to be included in the DECK are well suited for evaluating models and for understanding important climate change response characteristics. For these reasons, these experiments are already commonly performed by modelling groups as part of their model development cycle. Modelling groups also commonly perform simulations of the historical period, but reconstructions of the external conditions imposed on historical runs (e.g., land-use changes) continue to evolve significantly, influencing the simulated climate. In order to distinguish among the historical simulations performed under different phases of CMIP, the historical simulations are labelled with the phase (e.g., "CMIP5 *historical*" or "CMIP6 *historical*"). Note that in AMIP runs, the dominating role of sea surface temperatures and the focus on recent decades means that for most purposes runs from different phases of CMIP can be compared near the Earth's surface despite some differences in other imposed conditions.

The persistence and consistency of the DECK will make it possible to track changes in performance and response characteristics over future generations of models and CMIP phases. Although this core set of experiments is not expected to evolve much, additional experiments may become well enough established as benchmarks (routinely run by modelling groups as they develop new model versions) so that in the future they might be migrated into the DECK. The common practice of including the DECK in model development efforts means that models can contribute to CMIP without carrying out additional computationally burdensome experiments. All of the DECK and CMIP historical simulations were included in the core set performed under CMIP5 (Taylor et al., 2012), and all but the *abrupt-4xCO2* simulation were included in even earlier CMIP phases.

Under CMIP, ~~the~~ credentials of the participating atmospheric-ocean general circulation models (AOGCMs) and ~~Earth System Models (~~ESMs~~)~~ are established by performing the DECK and ~~the~~ CMIP ~~Historical Simulation~~historical simulations, so these experiments are required from all models. Together these experiments document the mean climate and response characteristics of models. They should be run for each model configuration used in a CMIP-Endorsed MIP. A change in ~~'model configuration'~~model configuration includes any change that might affect its simulations other than "noise" expected from different realizations. This would include, for example, a change in model resolution, physical processes, or atmospheric chemistry treatment. ~~If a model will be run in emission-driven mode as part of a CMIP6-Endorsed MIP (e.g., C[4]MIP) then the *piControl* and the CMIP Historical Simulation should be run in emission-driven mode too, whereas the AMIP, 1% yr^{-1} and switch-on 4xCO2 simulations would also be required as part of the DECK, but they would be driven by prescribed concentrations. Similarly, if a model will be used in both emission-driven mode and~~If an ESM is used in both CO_2 emission-driven mode and CO_2 concentration-driven mode in subsequent CMIP6-Endorsed MIPs, then both emission-driven and concentration-driven ~~*piControl*~~control and ~~Historical Simulations~~historical simulations should be done~~.~~ and they will be identical in all forcings except the treatment of CO_2.

The ~~forcings~~ forcing datasets that ~~are used in~~will drive the DECK and ~~in the~~ CMIP6 ~~Historical Simulation~~ historical simulations are described separately in a series of invited contributions to this Special Issue~~, and are~~. These articles also include some discussion of uncertainty in the datasets. The data will be provided by the respective author teams~~. These include: (1) historical~~ and made publicly available through the ESGF using common metadata and formats.

The historical forcings are based as far as possible on observations and cover the period 1850 to 2014. These include:

- emissions of short-lived species and long-lived greenhouse gases (GHGs),~~(2) historical~~
- GHG concentrations,~~(3)~~
- global gridded land-use forcing datasets,~~(4)~~
- solar forcing,~~(5)~~
- stratospheric aerosol dataset (volcanoes),~~and (6)~~
- AMIP sea surface temperatures (SSTs) and sea-ice concentrations (SICs~~). In addition~~),
- for simulations with prescribed aerosols~~(7)~~ a new approach to prescribe aerosols in terms of optical properties and fractional change in cloud droplet effective radius to provide a more consistent representation of aerosol forcing, and ~~for models without ozone chemistry (8) time-varying gridded ozone concentrations are provided.~~
- for models without ozone chemistry time-varying gridded ozone concentrations and nitrogen deposition.

Some models might require additional forcing datasets (e.g., black carbon on snow or anthropogenic dust). Allowing model groups to use different forcing[2] datasets might better sample uncertainty, but makes it more difficult to assess the uncertainty in the response of models to the best estimate of the forcing, available to a particular CMIP phase. To avoid conflating uncertainty in the response of models to a given forcing, it is strongly preferred for models to be integrated with the same forcing, and for forcing uncertainty to be sampled in supplementary simulations. In any case it is important that all forcing datasets are documented and are made available alongside the model output on the ESGF. Likewise to the extent modelling centres simplify forcings, for instance by regridding or smoothing in time or some other dimension, this should also be documented.

For the future scenarios selected by ScenarioMIP, forcings are provided by the integrated assessment model (IAM) community for the period 2015 to 2100 or to 2300 for the extended simulations. For atmospheric emissions and concentrations as well as for land use these are harmonized across IAMs and scenarios similar to the CMIP5 procedure (van Vuuren et al., 2011) to ensure consistency with historical forcing datasets and between the different forcing categories. They are described elsewhere in this Special Issue, while the underlying IAM scenarios are described in a Special Issue in Global Environmental Change.

An important gap identified in CMIP5, and in previous CMIP phases, was a lack of careful quantification of the radiative forcings from the different specified external forcing factors (e.g., GHGs, sulphate aerosols) in each model (Stouffer et al., 2015). This has impaired attempts to identify reasons for differences in model responses. The "effective radiative forcing" or ERF component of the Radiative Forcing MIP (RFMIP) includes "fixed SST" simulations to diagnose the forcing ('RFMIP-lite'), which are further detailed in the corresponding contribution to this Special Issue. Although not included as part of the DECK, in recognition of this deficiency in past phases of CMIP we encourage all CMIP6 modelling groups to participate in RFMIP-lite. This modest effort would enable the radiative forcing to be characterized for both historic and future scenarios across the model ensemble and would lead to a step change in the understanding of the spread of model responses for CMIP6.

---

[2] Here we distinguish between an applied input perturbation (e.g. the imposed change in some model constituent, property, or boundary condition), which we refer to somewhat generically as a "forcing", and *radiative* forcing, which can be precisely defined. Even if the forcings are identical, the resulting *radiative* forcing depends on a model's radiation scheme (among other factors) and will differ among models.

An overview of the main characteristics of the DECK and ~~the~~ CMIP6 ~~Historical Simulation is given~~historical simulations appears in Table 2. Here we briefly describe these experiments. Detailed specifications for the DECK and ~~the~~ CMIP6 ~~Historical Simulation~~historical simulations are provided in Appendix ~~A1 and A2, respectively,~~A and are summarized in Table A1.

3.1. The DECK

The AMIP and pre-industrial control simulations of the DECK provide opportunities for evaluating the atmospheric model and the coupled system, and in addition they establish a baseline for performing many of the CMIP6 experiments. Many experiments branch from, and are compared with, the pre-industrial control. Similarly, a number of diagnostic atmospheric experiments use AMIP as a control. The idealized CO_2-forced experiments in the DECK (1% yr^{-1} CO_2 and abrupt $4xCO_2$ increases), despite their simplicity, can reveal fundamental forcing and feedback response characteristics of models. ~~The experiments of the DECK are already commonly performed by modelling groups in the course of developing a new model, and their protocols are expected to remain essentially unchanged for many years to come. The persistence and consistency of these experiments will make it possible to track over future generations of models any changes in performance and response characteristics. Although this core set of experiments is not expected to evolve much, additional experiments may become well enough established as benchmarks (routinely run by modelling groups as they develop new model versions) so that in the future they might be migrated into the DECK. The common practice of including the DECK in model development efforts means that models can contribute to CMIP without carrying out additional computationally burdensome experiments. All experiments of the DECK were included in the core set performed under CMIP5 (Taylor et al., 2012), and all but the *abrupt4xCO2* simulation were included in even earlier CMIP phases.~~

For nearly three decades, AMIP simulations ~~(Gates et al., 1999)~~(Gates et al., 1999) have been routinely relied on by modelling centres to help in the evaluation of the atmospheric component of their models. In AMIP simulations, the SSTs and SICs are prescribed based on observations. The idea is to analyse and evaluate the atmospheric and land components of the climate system when they are constrained by the observed ocean conditions. These simulations can help identify which model errors originate in the atmosphere, land, or ~~land components~~their interactions, and they have proven useful in addressing a great variety of questions pertaining to recent climate changes. The AMIP simulations performed as part of the DECK cover at least the period from January 1979 to December 2014. The end date will continue to evolve as the SSTs and SICs are updated with new observations. Besides prescription of ocean conditions in these simulations, realistic forcings are imposed that should be identical to those applied in the CMIP ~~Historical Simulation.~~historical simulations. Large ensembles of AMIP simulations are encouraged as they can help ~~separate~~to improve the signal ~~of forced responses~~to noise ratio ~~(Li et al., 2015)~~(Li et al., 2015).

~~The remaining three experiments in the DECK are premised on the coupling of the atmospheric and oceanic circulation. The pre-industrial control simulation (*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. There are no secular changes in forcing, so the concentrations and/or sources of atmospheric constituents (e.g., GHGs and other forcings) are held fixed, as are Earth's orbital characteristics. External human influences on the land surface are likewise excluded. Because of the absence of both naturally occurring changes in forcing (e.g., volcanoes, orbital or solar changes) and human-induced changes, the *piControl* simulation gives insight into the unforced internal variability of the climate system.~~

The remaining three experiments in the DECK are premised on the coupling of the atmospheric and oceanic circulation. 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. Historically, the industrial revolution began in the 18th century, and in nature the climate in 1850 was not stable as it was already changing due to prior historical changes in radiative forcings. In CMIP6, however, as in earlier CMIP phases, the control simulation is an attempt to produce a stable quasi-equilibrium climate state under 1850 conditions. When discussing and analysing historical and future radiative forcings, it needs to be recognized that the radiative forcing in 1850 due to anthropogenic greenhouse gas increases alone was already around 0.25 W/m^2 (Cubasch, 2013) although aerosols might have offset that to some extent. In addition, there were other pre-1850 secular changes, for example in land use (Hurtt et al., 2011), and as a result, global net annual emissions of carbon from land use and land-use change already were responsible in 1850 for about 0.6 PgC/yr (Houghton, 2010). Under the assumptions of the control simulation, however, there are no secular changes in forcing, so the concentrations and/or sources of atmospheric constituents (e.g., GHGs and emissions of short-lived species) as well as land use are held fixed, as are Earth's orbital characteristics. Because of the absence of both naturally occurring changes in forcing (e.g., volcanoes, orbital or solar changes) and human-induced changes, the control simulation can be used to study the unforced internal variability of the climate system.

An initial climate "spin-up" portion of a control simulation, during which the climate begins to come into balance with the forcing, is usually performed~~ and discarded. The length of this "spin-up" period is model and resource dependent.~~. At the end of the "spin-up" period, the *piControl* starts. The *piControl* serves as a baseline for experiments that branch from it. To account for the effects of any residual drift, it is required that the *piControl* simulation extends as far beyond the branching point as any experiment to which it will be compared. Only then can residual climate drift in an experiment be removed, so that it is not misinterpreted as part of the model's forced response. The recommended minimum length for the *piControl* is 500 years.

~~The two DECK 'climate change' experiments branch from some point in the *piControl* and are designed to document basic aspects of the climate system response to GHG forcing. In the first, the CO_2 concentration is immediately and abruptly quadrupled. This *abrupt4xCO2* simulation has proven to be useful for characterizing the radiative forcing that arises from an increase in atmospheric CO_2 as well as changes that arise indirectly due to the warming. It can also be used to estimate a model's equilibrium climate sensitivity (ECS, Gregory et al. (2004)). In the second, the CO_2 concentration is increased gradual at a rate of 1% per year. This experiment has been performed in all phases of CMIP since CMIP2, and serves as a consistent and useful benchmark for analysing model transient climate response (TCR). The TCR takes into account the rate of ocean heat uptake which governs the pace of all time-evolving climate change (e.g., Murphy and Mitchell (1995)). In addition to the TCR, the 1% CO_2 integration with ESMs that include explicit representation of the carbon cycle allows the calculation of the transient climate response to cumulative carbon emissions (TCRE) defined as the transient global average surface temperature change per unit of accumulated CO_2 emissions (IPCC, 2013). Despite their simplicity, these experiments provide a surprising amount of insight into the behaviour of models subject to more complex forcing (e.g., Bony et al. (2013); Geoffroy et al. (2013)).~~

The two DECK 'climate change' experiments branch from some point in the 1850 control simulation and are designed to document basic aspects of the climate system response to greenhouse gas forcing. In the first, the CO_2 concentration is immediately and abruptly quadrupled from January 1850 values. This *abrupt-4xCO2* simulation has proven to be useful for characterizing the radiative forcing that arises from an increase in atmospheric CO_2 as well as changes that arise indirectly due to the warming. It can also be used to estimate a model's equilibrium climate sensitivity (ECS, Gregory et al. (2004)). In the second, the CO_2 concentration is increased gradually at a rate of 1% per year. This experiment has been performed in all phases of CMIP since CMIP2, and serves as a consistent and useful benchmark for analysing model transient climate response (TCR). The TCR takes into account the rate of ocean heat uptake which governs the pace of all time-evolving climate change (e.g., Murphy and Mitchell (1995)). In addition to the TCR, the 1% CO_2 integration with ESMs that include explicit representation of the carbon cycle allows the calculation of the transient climate response to cumulative carbon emissions (TCRE), defined as the transient global average surface temperature change per unit of accumulated CO_2 emissions (IPCC, 2013). Despite their simplicity, these experiments provide a surprising amount of insight into the behaviour of models subject to more complex forcing (e.g., Bony et al. (2013); Geoffroy et al. (2013)).

3.2. CMIP Historical Simulationhistorical simulations

In addition to the DECK, CMIP challenges models to simulate the historical period, defined to begin in 1850 and extend to the near present (i.e., 2014 in CMIP6). The CMIP Historical Simulation branches*historical* simulation and its CO_2-emission-driven counterpart, *esm-hist*, branch from the *piControl* and *esm-*piControl and is, respectively (see details in A1.2). These simulations are forced, based on observations, by evolving, externally-imposed forcings such as solar variability, volcanic aerosols, and changes in atmospheric composition (GHGs, and aerosols) caused by human activities. The CMIP Historical Simulation provideshistorical simulations provide rich opportunities to assess model ability to simulate climate, including variability and century time-scale trends (e.g., Flato et al. (2013)), and it has also proven essential in reducing uncertainty in radiative forcing associated with short lived species such as the atmospheric aerosol (e.g., Stevens (2015)).Flato et al. (2013)). When supplemented with additional experiments, the Historical Simulationhistorical simulations can be used in detection and attribution studies (e.g., Stott et al. (2006)Stott et al. (2006)) to help interpret the extent to which observed climate change can be explained by different causes.

As in the *piControl* simulationperforming control simulations, models that include representation of the carbon cycle should normally perform two different CMIP Historical Simulations: a historical simulations: one with prescribed CO_2 concentration and athe other with prescribed CO_2 emissions simulation (accounting explicitly for fossil fuel combustion), in which). In the second CO_2 concentrations are then "predicted" by the model. The treatment of other GHGs should be identical in both simulations. Both types of simulation are useful in evaluating how realistically the model represents the response of the carbon cycle anthropogenic CO_2 emissions, but the prescribed concentration simulation enables these more complex models to be evaluated fairly against those simpler models without representation of carbon cycle processes.

3.3. Common standards, infrastructure and documentation

A key to the success of CMIP and one of the motivations for incorporating a wide variety of coordinated modelling activities under a single framework in a specific phase of CMIP (now CMIP6) is the desire to reduce duplication of effort, minimize operational and computational burdens, and establish common practices in producing and analysing large amounts of model output. To enable automated processing of output from dozens of different models, CMIP has led the way in encouraging adoption of data standards (governing structure and metadata) that facilitate development of software infrastructure in support of coordinated modelling activities. The ESGF has capitalized on this standardization to provide access to CMIP model output hosted by institutions around the world. As the complexity of CMIP has increased and as the potential use of model output expands beyond the research community, the evolution of the climate modelling infrastructure requires enhanced coordination. To help in this regard, the WGCM Infrastructure Panel (WIP) was set up (see details in the corresponding contribution to this Special Issue), and is now providing guidance on requirements and establishing specifications for model output, model and simulation documentation, and archival and delivery systems for CMIP6 data.

A more routine benchmarking and evaluation of the models is envisaged to be a central part of CMIP6. As noted above, one purpose of the DECK and ~~the~~ CMIP ~~Historical Simulation~~historical simulations is to provide a basis for documenting model simulation characteristics. Towards that end an infrastructure is being developed to allow analysis packages to be routinely executed whenever new model experiments are contributed to the CMIP archive. These efforts utilize observations served by the ESGF contributed from the obs4MIPs ~~(Teixeira et al., 2014)~~(Ferraro et al., 2015; Teixeira et al., 2014) and ana4MIPs projects. Examples of available tools that target routine evaluation in CMIP include the PCMDI metrics software ~~(Gleckler et al., 2015) and the Earth System Model Evaluation Tool (ESMValTool, Eyring et al. (2015)~~(Gleckler et al., 2016) and the Earth System Model Evaluation Tool (ESMValTool, Eyring et al. (2015)), which brings together established diagnostics such as those used in the evaluation chapter of IPCC AR5 ~~(Flato et al., 2013)~~(Flato et al., 2013). The ESMValTool also integrates other packages, such as the NCAR Climate Variability Diagnostics Package ~~(Phillips et al., 2014)~~(Phillips et al., 2014), or diagnostics such as the cloud regime metric ~~(Williams and Webb, 2009)~~(Williams and Webb, 2009) developed by the Cloud Feedback MIP (CFMIP) community. These tools can be used to assess new models, and can help inform users of model output, as well as the modelling centres, as to the strengths and weaknesses of the simulations, including the extent to which long-standing model errors remain evident in newer models. Building such a community-based capability is not meant to replace how CMIP research is currently performed but rather to complement it. These tools can also be used to compute derived variables or indices alongside the ESGF, and their output could be provided back to the distributed ESGF archive.

4. CMIP6

4.1. Scientific focus of CMIP6

In addition to the DECK and ~~the~~ CMIP ~~Historical Simulation~~historical simulations, a number of additional experiments will colour a specific phase of CMIP, now CMIP6. These experiments are likely to change from one CMIP phase to the next. To maximize the relevance and impact of CMIP6, it was decided to use the Grand Science Challenges (GCs) of the WCRP as the scientific backdrop of the CMIP6 experimental design. By promoting research on critical science questions for which specific gaps in knowledge have hindered progress so far, but for which new opportunities and more focused efforts raise the possibility of significant progress on the timescale of 5-10 years, these GCs constitute a main component of the WCRP strategy to accelerate progress in climate science ~~(Brasseur and Carlson, 2015)~~(Brasseur and Carlson, 2015). Five such GCs have been identified, and two additional ones are under consideration. They relate to advancing (1) understanding of the role of clouds in the general atmospheric circulation and climate sensitivity ~~(Bony et al., 2015)~~(Bony et al., 2015), (2) assessing the response of the cryosphere to a warming climate and its global consequences, (3) understanding the factors that control water availability over land ~~(Trenberth and Asrar, 2014)~~(Trenberth and Asrar, 2014), (4) assessing climate extremes, what controls them, how they have changed in the past and how they might change in the future ~~(Alexander et al., 2015)~~(Alexander et al., 2015), (5) understanding and predicting regional sea-level change and its coastal impacts, (6) improving near-term climate predictions, and (7) determining how biogeochemical cycles and feedbacks control greenhouse gas concentrations and climate change.

These GCs will be using the full spectrum of ~~observation~~observational, modelling and ~~analysis~~analytical expertise across the WCRP, and in terms of modelling most GCs will address their specific science questions through a hierarchy of numerical models of different complexities. Global coupled models obviously constitute an essential element of this hierarchy, and CMIP6 experiments will play a prominent role across all GCs by helping to answer the ~~three~~ following three CMIP6 science questions: How does the Earth system respond to forcing? What are the origins and consequences of systematic model biases? How can we assess future climate change given internal climate variability, climate predictability, and uncertainties in scenarios?

These three questions will be at the centre of CMIP6. They will be addressed through a range of CMIP6-Endorsed MIPs that are organized by the respective communities and overseen by the CMIP Panel (Fig. 2). Through these different MIPs and their connection to the GCs, the goal is to fill some of the main scientific gaps of previous CMIP phases. This includes in particular facilitating the identification and interpretation of model systematic errors, improving the estimate of radiative forcings in past and future climate change simulations, facilitating the identification of robust climate responses to aerosol forcing during the historical period, better ~~taking into account~~accounting of the impact of short-term forcing agents and land-use on climate, better understanding the mechanisms of decadal climate variability, ~~and~~along with many other issues ~~that could~~ not~~ be~~ addressed satisfactorily in CMIP5 (Stouffer et al~~.~~., 2015). In endorsing a number of these MIPs the CMIP panel acted to minimize overlaps among the MIPs and to reduce the burden ~~of~~on modelling groups, while maximizing the scientific complementarity and synergy among the different MIPs.

4.2. The CMIP6-Endorsed MIPs

Close to ~~thirty~~30 suggestions for CMIP6 MIPs ~~were~~have been received so far of which 21 MIPs were eventually endorsed and invited to participate (Table 3). Of those not selected some were asked to work with other proposed MIPs with overlapping science goals and objectives. Of the 21 CMIP6-Endorsed MIPs, four are diagnostic in nature, which means that they define and analyse additional output, but do not require additional experiments. In the remaining 17 MIPs, a total of around 190 experiments have been proposed resulting in 40,000 model simulation years with around half of these in Tier 1. The CMIP-Endorsed MIPs show broad coverage and distribution across the three CMIP6 science questions, and ~~links~~all are linked to the WCRP Grand Science Challenges (Fig. 3).

Each of the 21 CMIP6-Endorsed MIPs ~~will be~~is described in a separate invited contribution to this Special Issue. These contributions will detail the goal of the MIP and the major scientific gaps the MIP is addressing, and will specify what is new compared to CMIP5 and previous CMIP phases. The contributions will include a description of the experimental design and scientific justification of each of the experiments for Tier 1 (and possibly beyond), and will link the experiments and analysis to the DECK and ~~the~~ CMIP6 ~~Historical Simulation~~historical simulations. They will additionally include an analysis plan to fully justify the resources used to produce the various requested variables, and if the analysis plan is to compare model results to observations, the contribution will highlight possible model diagnostics and performance metrics specifying whether the comparison entails any particular requirement for the simulations or outputs (e.g. the use of observational simulators). In addition, possible observations and reanalysis products for model evaluation ~~will be~~are discussed and the MIPs are encouraged to help facilitate their use by contributing them to the obs4MIPs/ana4MIPs archives at the ESGF (see Section 3.3). In some MIPs additional forcings beyond those used in the DECK and CMIP6 ~~Historical Simulation will be~~historical simulations are required, and these ~~will be~~are described in the respective contribution as well.

A number of MIPs are developments and/or continuation of long standing science themes within CMIP. These include MIPs specifically addressing science questions related to cloud feedbacks and the understanding of spatial patterns of circulation and precipitation (CFMIP), carbon cycle feedbacks and the understanding of changes in carbon fluxes and stores (C^{4}MIP), detection and attribution (DAMIP) that newly includes 21st-century GHG-only simulations allowing the projected responses to GHGs and other forcings to be separated and scaled to derive observationally-constrained projections, and paleoclimate (PMIP) that), which assesses the credibility of the model response to forcing outside the range of recent variability. These MIPs reflect the importance of key forcing and feedback processes in understanding past, present and future climate change and have developed new experiments and science plans focused on emerging new directions that will be at the centre of the WCRP Grand Science Challenges. A few new MIPs have arisen directly from gaps in understanding in CMIP5 (Stouffer et al., 2015),(Stouffer et al., 2015), for example poor quantification of radiative forcing (RFMIP), better understanding of ocean heat uptake and sea-level rise (FAFMIP)), and understanding of model response to volcanic forcing (VolMIP).

Since CMIP5, other MIPs have emerged as the modelling community has developed more complex ESMs with interactive components beyond the carbon cycle. These include the consistent quantification of forcings and feedbacks from aerosols and atmospheric chemistry (AerChemMIP), and, for the first time in CMIP, modelling of sea-level rise from land-ice sheets (ISMIP6).

Some MIPs specifically target systematic biases focusing on improved understanding of the sea-ice state and its atmospheric and oceanic forcing (SIMIP), the physical and biogeochemical aspects of the ocean (OMIP), land, snow and soil moisture processes (LS3MIP), and improved understanding of circulation and variability with a focus on stratosphere-troposphere coupling (DynVar). With the increased emphasis in the climate science community on the need to represent and understand changes in regional circulation, systematic biases are also addressed on a more regional scale by the Global Monsoon MIP (GMMIP) and a first co-ordinatedcoordinated activity on high resolution modelling (HighResMIP).

For the first time future scenario experiments, previously co-ordinatedcoordinated centrally as part of the CMIP5 ‘core’ experiments, will be run as a MIP ensuring clear definition and well-coordinated science questions. ScenarioMIP will run a new set of future long-term (century time scale) integrations engaging input from both the climate science and Integratedintegrated assessment modelling communities. The new scenarios that are based on the shared socioeconomic pathways (SSPs, O’Neill et al. (2015)O’Neill et al. (2015)) - Representative Concentration Pathways (RCP) matrix span the same range as the CMIP5 RCPs (Moss et al., 2010)(Moss et al., 2010), but fill critical gaps for intermediate forcing levels and questions, for example, on short-lived species and land-use. The near-term experiments (10–30 years) ~~will be~~are coordinated by the decadal climate prediction project (DCPP) with improvements expected for example from the initialization of additional components beyond the ocean and from a more detailed process understanding and ~~verification~~evaluation of the ~~models~~predictions to better identify sources and limits of predictability.

Other MIPs include specific future mitigation options, e.g. the land use MIP (LUMIP) that is for the first time in CMIP ~~looking at~~isolating regional land management strategies to study how different surface types respond to climate change and direct anthropogenic modifications, or the geoengineering MIP (GeoMIP~~) that~~), which examines climate impacts of newly proposed radiation modification geoengineering strategies.

The diagnostic MIP CORDEX will oversee the downscaling of CMIP6 models for regional climate projections. Another historic development in our field that provides, for the first time in CMIP, an avenue for a more formal communication between the climate modelling and user community is the endorsement of the vulnerability, impacts and adaptation and climate services advisory board (VIACS AB). This diagnostic MIP requests certain key ~~outputs from CMIP6 models to deliver~~variables of interest to the VIACS ~~communities~~community be delivered in ~~rapid time for application to~~a timely manner to be used by climate services and in impact studies.

All MIPs define output streams in the centrally coordinated CMIP6 data request for each of their own experiments as well as the DECK and CMIP6 ~~Historical Simulations~~historical simulations (see the CMIP6 data request contribution to this Special Issue for details). This will ensure that the required variables are stored at the frequency and resolution required ~~so that~~to address the specific science questions and evaluation needs of each MIP ~~can be addressed,~~ and to enable a broad characterization of the performance of the CMIP6 models~~ and ensemble can be made~~.

We note that only the Tier 1 MIP experiments are overseen by the CMIP Panel, but additional experiments are proposed by the MIPs in Tier 2 and 3. We encourage the modelling groups to participate in the full suite of experiments beyond Tier 1 to address in more depth the scientific questions posed.

The call for MIP applications for CMIP6 is still open and new proposals will be reviewed at the annual WGCM meetings. However, we point out that the additional MIPs suggested after the CMIP6 data request has been finalized will have to work with the already defined model output from the DECK and CMIP6 historical simulations, or work with the modelling group to recover additional variables from their internal archives. We also point out that some experiments proposed by CMIP6-Endorsed MIPs may not be finished until after CMIP6 ends.

5. Summary

CMIP6 continues the pattern of evolution and adaptation characteristic of previous phases of CMIP. To ~~address the importance of broadly centring~~center CMIP at the heart of activities within climate science~~, yet link~~ and encourage links among activities within the World Climate Research Programme (WCRP~~) more specifically,~~), CMIP6 has been formulated scientifically around three specific themes, amidst the backdrop of the WCRP's seven Grand Science Challenges. To meet the increasingly broad scientific demands of the climate-science community, yet be responsive to the individual priorities and resource limitations of the modelling centres, CMIP has adopted a new, more federated organizational structure.

CMIP has now evolved from a centralized activity involving a large number of experiments to a federated activity, encompassing many individually designed MIPs. CMIP6 comprises 21 individual CMIP6-Endorsed MIPs~~,~~ and the DECK and ~~the~~ CMIP6 ~~Historical Simulations~~historical simulations. Four of the 21 CMIP6-Endorsed MIPs are diagnostic in nature, meaning that they require additional output from models, but not additional simulations. The total amount of output from CMIP6 is estimated to be between 20 and 40 Petabytes, depending on model resolution and the number of modelling centres ultimately participating in CMIP6. Questions addressed in the MIPs are wide ranging, from the climate of distant past to the response of turbulent cloud processes ~~influence by the response~~ to radiative forcing, from how the terrestrial biosphere influences the uptake of ~~carbon-dioxide~~CO_2 to how much predictability is ~~encoded~~stored in the ocean, from how to best project near-term to long-term future climate changes while considering interdependences and differences in model performance in the CMIP6 ensemble, and from what regulates the distribution of tropospheric ozone, to the influence of land-use changes on water availability.

The last two years have been dedicated to conceiving and then planning what we now call CMIP6. Starting in 2016, the first modelling centres are expected to begin performing the DECK and uploading output on the ESGF. ~~Around April~~By May 2016 the forcings for the DECK and CMIP6 historical simulations ~~should~~will be ready, and by the end of 2016 the diverse forcings for different scenarios of future human activity will become available. Past experience suggests that most centres will complete their CMIP simulations within a few years while the analysis of CMIP6 results will likely go on for a decade or more (Fig. 4).

Through an intensified effort to align CMIP with specific scientific themes and activities we expect CMIP6 to continue CMIP's tradition of major scientific advances. CMIP6 simulations and scientific achievements are expected to support the IPCC Sixth Assessment Report (AR6) as well as other national and international climate assessments or special reports. Ultimately scientific progress will be the best measure of the success of CMIP6. Measures of success will include improved understanding of how the climate system works through the quantification of forcings and feedbacks, improved understanding and interpretation of systematic model biases and corresponding identification of ways to alleviate them for model improvements, and robust climate projections and uncertainty estimates for adaptation and mitigation policies.

Data availability

The model output from the DECK and ~~the~~ CMIP6 ~~Historical Simulations~~historical simulations described in this paper will be distributed through the Earth System Grid Federation (ESGF~~).~~) with digital object identifiers (DOIs) assigned. As in CMIP5, the model output will be freely accessible through data portals after registration. In order to document CMIP6's scientific impact and enable ongoing support of CMIP, users are obligated to acknowledge CMIP6~~ and~~, the participating modelling groups, and the ESGF centres (see details on the CMIP Panel website at http://www.wcrp-climate.org/index.php/wgcm-cmip/about-cmip). Further information about the infrastructure supporting CMIP6, the metadata describing the model output, and the terms governing its use ~~will be~~are provided by the WGCM Infrastructure Panel (WIP) in their invited contribution to this Special Issue. Along with the data itself, the provenance of the data will be recorded, and DOI's will be assigned to collections of output so that they can be appropriately cited. This information will be made readily available so that published research results can be verified and credit can be given to the modelling groups providing the data. The WIP is coordinating and encouraging the development of the infrastructure needed to archive and deliver this information. In order to run the experiments, datasets for natural and anthropogenic forcings are required. These forcing datasets ~~will be~~are described in separate invited contributions to this Special Issue. The forcing datasets will be made available through the ESGF with version control and ~~digital object identifiers (DOI's) assigned, and can additionally be provided as a supplement to the corresponding documentation paper. Links to all forcings datasets will be made available via the CMIP Panel website~~DOIs assigned.

~~**Acknowledgements**. We thank the scientific community for their engagement in the definition of CMIP6, in particular for the broad participation in the CMIP5 survey in 2013. We thank the co-~~

~~chairs and steering committee members of the CMIP6-Endorsed MIPs for their continuous engagement in defining CMIP6, and the modelling groups and wider community for reviewing the CMIP6 design and organisation.~~ ~~We thank the WGCM Infrastructure Panel (WIP) for overseeing the CMIP6 infrastructure, Martin Juckes for taking the lead in preparing the CMIP6 data request, and Jonathan Gregory for raising awareness about the treatment of volcanic forcing in the preindustrial control experiment and its consequence for sea level changes. Norbert Noreiks is thanked for help in drafting the figures.~~

Appendix A. Experiment Specifications

A1 Specifications for the DECK

Here we provide information needed to perform the DECK, including specification of forcing and boundary conditions, initialization procedures, and minimum length of runs. This information is largely consistent with but not identical to the specifications for these experiments in CMIP5 ~~(Taylor et al., 2009)~~(Taylor et al., 2009).

The DECK and ~~the~~ CMIP6 ~~*Historical Simulation* is~~historical simulations are requested from all models participating in CMIP. The expectation is that this requirement will be met for each model configuration used in the subsequent CMIP6-Endorsed MIPs (an entry card). In the special case where the burden of the entry card simulations are prohibitive but the scientific case for including a particular model simulation is ~~strong~~compelling (despite only partial completion of the entry card simulations), an exception to this policy can be granted on a model by model basis ~~based on a specific recommendation to~~by the CMIP Panel ~~made by~~, which will seek advice from the chairs of the affected CMIP6-Endorsed MIP.

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 the DECK experiments and the CMIP6 historical simulations to the ESGF, as well as any CMIP6-Endorsed-MIP experiments they choose to run. Additionally, the modelling groups decide what constitutes a new model version. Modelling groups are well aware that their model simulations are under considerable scrutiny. Therefore, we expect that as in the past, modelling groups will in good faith provide their highest quality model version and that it will differ 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.

A1.1 AMIP simulation

As in the first simulations performed under the Atmospheric Model Intercomparison Project (AMIP, ~~Gates et al. (1999)~~Gates et al. (1999)), SSTs and SICs in AMIP experiments are prescribed consistent with observations (see details on this forcing dataset in the corresponding contribution to this Special Issue). Land models should be configured as close as possible to that used in the CMIP6 ~~Historical Simulation~~historical simulation including transient land use and land cover. Other external forcings including volcanic aerosols, solar variability, GHG concentrations, and anthropogenic aerosols should also be prescribed consistent with those used in the CMIP6 ~~*Historical Simulation*~~historical simulation (see ~~Appendix~~Section A2 below). Even though in AMIP simulations models with an active carbon cycle will not be fully interactive, surface carbon fluxes should be archived over land. ~~This will enable evaluation of the carbon cycle component of the model when climate conditions are more similar to the observed than in coupled atmosphere-ocean simulations.~~

AMIP integrations can be initialized from prior model integrations or from observations or in other reasonable ways. Depending on the treatment of snow cover, soil water content, the carbon cycle, and vegetation, these runs may require a spin-up period of several years. One might establish quasi-equilibrium conditions consistent with the model by, for example, running with ocean conditions starting earlier in the 1970's or cycling repeatedly through year 1979 before simulating the official period. Results from the spin-up period (i.e., prior to 1979) should be discarded, but the spin-up technique should be documented.

For CMIP6, AMIP simulations should cover at least the period from January 1979 through December 2014, but modelling groups are encouraged to extend their runs to the end of the observed period. Output may also be contributed from years preceding 1979 with the understanding that surface ocean conditions were less complete and in some cases less reliable then.

The climate found in AMIP simulations is largely determined by the externally-imposed forcing, especially the ocean conditions. Nevertheless, unforced variability ("noise") within the atmosphere introduces some non-deterministic variations that hamper unambiguous interpretation of apparent relationships between, for example, the year-to-year anomalies in SSTs and their consequences over land. To assess the role of unforced atmospheric variability in any particular result, modelling groups are encouraged to generate an ensemble of AMIP simulations. For most studies a three-member ensemble, where only the initial conditions are varied, would be the minimum required, with larger size ensembles clearly of value in making more precise determination of statistical significance.

A1.2 Multi-century pre-industrial control ~~simulation~~simulations

Like laboratory experiments, numerical experiments are designed to reveal cause and effect relationships. A standard way of doing this is to perform both a "control" experiment and a second experiment where some externally-imposed experiment condition has been altered. For many CMIP experiments, including the rest of the experiments discussed in this Appendix, the "control" is a simulation with atmospheric composition and other conditions prescribed and held constant, consistent with best estimates of the forcing~~ estimated~~ from the historical period.

Ideally the pre-industrial control (*piControl*) experiment for CMIP would represent a near-equilibrium state of the climate system under the imposed conditions. In reality, simulations of hundreds to many thousands of years would be required for the ocean's depths to equilibrate and for biogeochemical reservoirs to fully adjust. Available computational resources generally preclude integrations long enough to approach equilibrium, so in practice shorter runs must suffice. Usually, a *piControl* simulation is initialized from the control run of a different model or from ~~near present-day~~ observations, and then run until at least the surface climate conditions stabilize using 1850 forcings (see ~~Stouffer et al. (2004)~~Stouffer et al. (2004) for further discussion). This spin-up period~~, which is discarded,~~ can be as long as several hundred years~~. Note~~ and variables that can document the spin-up behaviour should be archived (under the experiment labels *piControl-spinup* or *esm-piControl-spinup*). At the very least the length of the spin-up period should be documented. ~~The *piControl* used in CMIP begins at this point and generally continues for at least a few hundred years.~~

Although equilibrium is generally not achieved, the changes occurring after the spin-up period are usually found to evolve at a fairly constant rate that presumably decreases slowly as equilibrium is approached. After a few centuries, these "drifts" of the system mainly affect the carbon cycle and ocean below the main thermocline, but they are also manifest at the surface in a slow change in sea level. The climate drift must be removed in order to interpret experiments that use the pre-industrial simulation as a control. The usual procedure is to assume that the drift is insensitive to CMIP experiment conditions and to simply subtract the control run from the perturbed run to determine the climate change that would occur in the absence of drift.

Besides serving as ~~a "control~~"controls" for numerical experimentation, the *piControl* ~~is~~and *esm-piControl* are used to study the naturally occurring, unforced variability of the climate system. The only source of climate variability in a control arises from processes internal to the model, whereas in the more complicated real world, variations are also caused by external forcing factors such as solar variability and changes in atmospheric composition caused, for example, by human activities or volcanic eruptions. Consequently, the physical processes responsible for unforced variability can more easily be isolated and studied using the control run of models, rather than by analysing observations.

A DECK ~~*piControl*~~control simulation is required to be long enough to extend to the end of any perturbation runs initiated from it so that climate drift can be assessed and possibly removed from those runs. If, for example, a historical simulation (beginning in 1850) were initiated from the beginning of the control simulation and then were followed by a future scenario run extending to year 2300, a control run of at least 450 years would be required. As discussed above, control runs are also used to assess model-simulated unforced climate variability. The longer the ~~*piControl*~~control, the more precisely can variability be quantified for any given time scale. A control simulation of many hundreds of years would be needed to assess variability on centennial time-scales. For CMIP6 it is recommended that the control run should be at least 500 years long (following the spin-up period), but of course the simulation must be long enough to reach to the end of the experiments it spawns. It should be noted that those analysing CMIP6 simulations might also require simulations longer than 500 years to accurately assess unforced variability on long time-scales, so modelling groups are encouraged to extend their ~~*piControl*~~control runs well beyond the minimum recommended number of years.

Because the climate was very likely not in equilibrium with ~~its~~the forcing ~~in the year~~of 1850~~,~~ and because different components of the climate system differentially ~~measure~~respond to the ~~effect~~effects of the forcing prior to that time, there is some ambiguity in deciding on what forcing to apply for the control. For CMIP6 we recommend a specification of this forcing that attempts to balance ~~partially~~ conflicting objectives to

- Minimize artificial climate responses to discontinuities in radiative forcing at the time a historical simulation is initiated.
- 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. Note that any preindustrial multi-centennial observed trend in global-mean sea level is most likely to be due to slow changes in ice-sheets, which are likely not to be simulated in the CMIP6 model generation.

The first consideration above implies that radiative forcing in the ~~*piControl*~~control run should be close to that imposed at the beginning of the CMIP ~~Historical Simulation~~historical simulation (i.e., 1850). The second implies that a background volcanic aerosol and time-averaged solar forcing should be prescribed in the control run, since to neglect it would cause an apparent drift in sea-level associated with the suppression of heat uptake due to the net effect of, for instance, volcanism after 1850, and this has implications for sea level changes ~~(Gregory, 2010; Gregory et al., 2013)~~(Gregory, 2010; Gregory et al., 2013). We recognize that it will be impossible to entirely avoid artefacts and artificial ~~transients~~transient effects, and practical considerations may rule out conformance with every ~~aspect~~detail of the ~~*piControl*~~control simulation protocol stipulated here. With that understanding, here ~~are~~is a summary of the recommendations for the imposed conditions on the ~~*piControl*~~spin-up and control runs, followed by further clarification in subsequent paragraphs:

- Conditions must be time-invariant except for those associated with the mean climate (notably the seasonal and diurnal cycles of insolation).
- Unless indicated otherwise (e.g., the background volcanic forcing), experiment conditions (e.g., greenhouse gas concentrations, ozone concentration, surface land conditions) should be representative of Earth ~~ca.~~around the year 1850.
- Orbital parameters (eccentricity, obliquity, and longitude of the perihelion) should be held fixed at their 1850 values.
- Land use should not change in the control run and should be fixed according to reconstructed agricultural maps from 1850. Due to the diversity of model approaches in ESMs for land carbon, some groups might deviate from this specification, and again this must be clearly documented.
- The solar constant should be fixed at its mean value (no 11 year solar cycle) over the first two solar cycles of the historical simulation (i.e., the 1850 – ~~1871~~1873 mean).
- A background volcanic aerosol should be specified that results in radiative forcing matching, as closely as possible, that experienced, on average, during the historical simulation (i.e., 1850-2014 mean).
- Models without interactive ozone chemistry should specify ~~ozone as in the mean of the first decade of the CMIP Historical Simulation~~the pre-industrial ozone fields from a dataset produced from a pre-industrial control simulation that uses 1850 emissions and a mean solar forcing averaged over solar cycles 8-10, representative of the mean mid-19th century solar forcing.

There are some special considerations that apply to control simulations performed by "emission-driven" ESMs (i.e. runs with atmospheric concentrations of CO_2 calculated prognostically rather than being prescribed). In the *esm-piControl* simulation, emissions of CO_2 from both fossil fuel combustion and land use change are prescribed to be zero. In this run any residual drift in atmospheric CO_2 concentration that arises from an imbalance in the exchanges of CO_2 between the atmosphere and the ocean and land (i.e. by the natural carbon cycle in the absence of anthropogenic CO_2 emissions) will need to be subtracted from perturbation runs to correct for a control state not in equilibrium. It should be emphasized that the *esm-piControl* is an idealized experiment and is not meant to mimic the true 1850 conditions, which would have to include a source of carbon of around 0.6 PgC/yr from the already perturbed state that existed in 1850.

Due to a wide variety of ESMs and the techniques they use to compute land carbon fluxes, it is hard to make statements that apply to all models equally well. A general recommendation, however, is that the land carbon fluxes in the emission and concentration driven control simulations should be stable in time and in approximate balance so that the net carbon flux into the atmosphere is small (less than 0.1 PgC/yr). Further details on ESM experiments with a carbon cycle are provided in the C^4MIP contribution to this Special Issue.

The historical time-average volcanic forcing stipulated above for the control run is likely to approximate the much longer term mean. Crowley's (2000) estimates of volcanic aerosol radiative forcing for the historical period and the last millennium are -0.18 W m^{-2} and -0.22 W m^{-2}, respectively. Because the mean volcanic forcing between 1850 and 2014 is small, the discontinuity associated with transitioning from a mean forcing to a time-varying volcanic forcing is also expected to be small. Even though this is the design objective, it is likely that it will be impossible to eliminate all artefacts in quantities such as historical sea level change~~ will not be entirely prevented.~~. For this reason, and because some models may deviate from these specifications, it is recommended ~~for all~~that groups~~ to~~ perform an additional simulation of the historical period but with only natural forcing included. ~~This natural-only historical simulation~~With this additional run, which is already called for under DAMIP~~. Modelling groups are urged to perform this experiment even if they elect not to participate in DAMIP as doing so will most effectively separate~~, the ~~role of natural versus~~purely anthropogenic ~~drivers of climate~~effects on sea-level change ~~and variability since 1850.~~can be isolated.

The forcing specified in the *piControl* also has implications for simulations of the future, when solar variability and volcanic activity will continue to exist, but at unknown levels. These issues need to be borne in mind when designing and evaluating future scenarios, as a failure to include volcanic forcing in the future will cause future warming and sea-level rise to be over-estimated relative to a *piControl* experiment in which a non-zero volcanic forcing is specified. This ~~could be addressed~~ is accounted for by ~~re-~~introducing a time-invariant non-zero volcanic forcing (e.g., the mean volcanic forcing for the *piControl*) into the scenarios. This is further specified in the ScenarioMIP contribution to this Special Issue.

These issues, and the potential of different modelling centres adopting different approaches to account for their particular constraints, highlight the paramount importance of adequately documenting ~~how~~the conditions under which this and the other DECK experiments ~~were~~are performed.

A1.3 Abruptly quadrupling CO_2 simulation

Until CMIP5, there were no experiments designed to quantify the extent to which forcing differences might explain differences in climate response. It was also difficult to diagnose and quantify the feedback responses, which are mediated by global surface temperature change ~~(Sherwood et al., 2015)~~(Sherwood et al., 2015). In order to examine these fundamental characteristics of models – CO_2 forcing and climate feedback – an abrupt 4xCO_2 simulation was included for the first time as part of CMIP5. Following Gregory et al. (2004), the simulation branches ~~from~~in January of the CO_2-concentration driven *piControl* and abruptly the atmospheric CO_2 concentration is ~~abruptly~~ quadrupled and ~~then~~ held ~~constant~~fixed. As the system subsequently evolves toward a new equilibrium, the imbalance in the net flux at the top of the atmosphere can be plotted against global temperature change. ~~As Gregory et al. (2004)~~As Gregory et al. (2004) showed, it is then possible to diagnose both the effective radiative forcing due to a quadrupling of CO_2 and also effective equilibrium climate sensitivity (ECS). Moreover, by examining how individual flux components evolve with surface temperature change, one can learn about the relative strengths of different feedbacks, notably quantifying the importance of various feedbacks associated with clouds.

In the ~~*abrupt4xCO2*~~*abrupt-4xCO2* experiment, the only externally-imposed difference from the *piControl* should be the change in CO_2 concentration. All other conditions should remain as they were in the *piControl*, including any background volcanic aerosols. By changing only a single factor, we can unambiguously attribute all climatic consequences to the increase in CO_2 concentration.

The minimum length of the ~~*abrupt4xCO2*~~*abrupt-4xCO2* simulation should be 150 years, but longer simulations ~~with CO_2 held constant after quadrupling are~~would enable investigations of ~~value for investigating~~ longer-time scale responses. Also there is value, as in CMIP5, in performing an ensemble of short (~5-year) simulations initiated at different times throughout the year~~, as called for as part of a Tier 1 set of experiments in CMIP5. An~~ (in addition to the required January run). Such an ensemble would reduce the statistical uncertainty with which the effective CO_2 radiative forcing could be quantified and would allow more detailed and accurate diagnosis of the fast responses of the system under an abrupt change in forcing ~~(Bony et al., 2013; Gregory and Webb, 2008; Kamae and Watanabe, 2013; Sherwood et al., 2015)~~(Bony et al., 2013; Gregory and Webb, 2008; Kamae and Watanabe, 2013; Sherwood et al., 2015)~~.~~. Different groups will be able to afford ensembles of different sizes, but in any case each realization should be initialized in a different month and the months should be spaced evenly throughout the year.

A1.4 1% CO_2 increase simulation

The second idealized climate change experiment was introduced in the early days of CMIP (Meehl et al., 2000)(Meehl et al., 2000). It is designed for studying model responses under simplified but somewhat more realistic forcing than an abrupt increase in CO_2. In this experiment, the simulation is branched from the *piControl*, and CO_2 concentration is gradually increased at a rate of 1% yr^{-1} (i.e., exponentially). A minimum length of 150 years is requested so that the simulation goes beyond the quadrupling of CO_2 after 140 years. Note that in contrast to previous definitions, the experiment has been simplified so that the 1% CO_2 increase per year is applied throughout the entire simulation rather than keeping it constant after 140 years as in CMIP5. Since the radiative forcing is approximately proportional to the logarithm of the CO_2 increase, the radiative forcing linearly increases over time. Drawing on the estimates of effective radiative forcing (for definitions see Myhre et al. (2013)Myhre et al. (2013)) obtained in the *abrupt4xCO2abrupt-4xCO2* simulations, analysts can scale results from each model in the 1% CO_2 increase simulations to focus on the response differences in models, largely independent of their forcing differences. In contrast, in the CMIP6 Historical Simulationhistorical simulations (see Section A2), the forcing and response contributions to model differences in simulated climate change cannot be easily isolated.

As in the *abrupt4xCO2*experiment*abrupt-4xCO2* experiment, the only externally-imposed difference from the *piControl* should be the change in CO_2 concentration. The omission of changes in aerosol concentrations is the key to making these simulations easier to interpret. The 1% yr^{-1} CO_2 increase simulation should be run for a minimum of 150 years (ten years after the time of quadrupling).

Models with a carbon cycle component will be driven by prescribed CO_2 concentrations, but terrestrial and marine surface fluxes and stores of carbon will become a key diagnostic from which one can infer emission rates that are consistent with a 1% yr^{-1} increase in model CO_2 concentration. This DECK baseline carbon cycle experiment is built upon in C^4MIP to diagnose the strength of model carbon climate feedback and to quantify contributions to disruption of the carbon cycle by climate and by direct effects of increased CO_2 concentration.

A2 The CMIP6 Historical Simulationhistorical simulations

The CMIP6 Historical Simulation is meant to reproduce observed climate and historical simulations of climate change starting inover the yearperiod 1850 and extending to the present (i.e.,through 2014 in CMIP6). It servesare forced by common datasets that are largely based on observations. They serve as an important benchmark for assessing model performance. through evaluation against observations. The historical integration should be initialized from some point in the ~~*piControl*~~control integration (with *historical* branching from the *piControl* and the *esm-hist* branching from *esm-piControl*) and be forced by time-varying, externally-imposed conditions that are based on observations. Both naturally-forced changes (e.g., due to solar variability and volcanic aerosols) and changes due to human activities (e.g., CO_2 concentration, aerosols, and land-use) will lead to climate variations and evolution. In addition, there is unforced variability which can obscure the forced changes and lead to expected differences between the simulated and observed climate variations ~~(Deser et al., 2012)~~(Deser et al., 2012).

The externally-imposed forcing datasets that should be used in CMIP6 cover the period 1850 through the end of 2014 are described in detail in various other contributions to this Special Issue. Recall from section A1.2 that the conditions in the ~~*piControl*~~control should generally be consistent with the forcing imposed near the beginning of the CMIP ~~Historical Simulation.~~historical simulation. This should minimize artificial ~~transients~~transient effects in the first portion of the CMIP ~~Historical Simulation~~historical simulation. An exception is that for the CO_2-emission driven experiments, the zero CO_2 emissions from fossil fuel and the land use specifications for 1850 in the *esm-piControl* could cause a discontinuity in land carbon at the branch point.

As described in Section A1.2, the 1850 *esm-piControl* should be developed for an idealized case that is stable in time and balance so that the net carbon flux into the atmosphere is small. Meanwhile, the start of the *esm-hist* in 1850 should be as realistic as possible and attempt to account for the fact the land-surface was not in equilibrium in 1850 due to prior land-use effects (Houghton, 2010; Hurtt et al., 2011). Some modelling groups have developed methods to achieve these twin goals in a computationally efficient manner, for example by performing pre-1850 off-line land model simulations to account for the land carbon cycle disequilibrium before 1850 and to adequately simulate carbon stores at the start of the historical simulation (Sentman et al., 2011). Due to the wide diversity of modelling approaches for land carbon in the ESMs, the actual method applied by each group to account for these effects will differ and needs to be well documented.

As discussed earlier, there will be a mismatch in the specification of volcanic aerosols between control and historical simulations that especially affect estimates of ocean heat uptake and sea level rise in the historical period. This can be minimized by prescribing a background volcanic aerosol in the ~~*piControl*~~pre-industrial control that has the same cooling effect as the volcanoes included in the CMIP6 ~~Historical Simulation~~historical simulation. Any residual mismatch will need to be corrected, which requires a special supplementary simulation (see Section A1.2) that should be submitted along with the CMIP6 ~~Historical Simulation.~~ historical simulation.

For model evaluation and for detection and attribution studies (the focus of DAMIP) there would be considerable value in extending the CMIP6 ~~Historical Simulations~~historical simulations beyond the nominal 2014 ending date. To include the more recent observations in model evaluation, modelling groups are encouraged to document and apply forcing data sets representing the post-2014 period. For short extensions (up to a few years) it may be acceptable to simply apply forcing from one of the future scenarios defined by ScenarioMIP. To distinguish between the portion of the historical period when all models will use the same forcing data sets (i.e., 1850-2014) from the extended period where different data sets might be used, the experiment for 1850 through 2014 will be ~~referred to as "~~labelled *historical*~~"~~ (*esm-hist* in the case of the emissions-driven run) and the period from 2015 through near-present will likely be ~~referred to as "~~labelled *historical*-~~extension".~~*ext* (*esm-hist-ext*).

Even if the CMIP6 ~~Historical Simulations~~historical simulations are extended beyond 2014, all future scenario simulations (called for by ScenarioMIP and other MIPs) should be initiated from the end of year 2014 of the CMIP6 ~~*Historical Simulation*. The~~historical simulation since the "future" in CMIP6 begins in 2015.

Due to interactions within and between the components of the Earth system, there is a wide range of variability on various time and space scales ~~(Hegerl et al., 2007)~~(Hegerl et al., 2007). The time scales vary from shorter than a day to longer than several centuries. The magnitude of the variability can be quite large relative to any given signal of interest depending on the time and space scales involved and on the variable of interest. To more clearly identify forced signals emerging from natural variability, multiple model integrations (comprising an "ensemble") can be made where only the initial conditions are perturbed in some way~~.~~, which should be documented. A common way to do this is to simply branch each simulation from a different point in the control run. Longer intervals between branch points will ensure independence of ensemble members on longer time-scales. By averaging many different ensemble members together, the signal of interest becomes clear because the natural variations tend to average out if the ensemble size and averaging period are long enough. If the variability in the models is realistic, then the spread of the ensemble members around the ensemble average is caused by unforced (i.e., ~~"~~internal~~"~~)) variability. To minimize the number of years included in the entry card simulations, only one ensemble member is requested here. However, we strongly encourage model groups to submit at least three ensemble members ~~for the~~of their CMIP ~~*Historical Simulation*~~historical simulation as requested in DAMIP.

Acknowledgements. We thank the scientific community for their engagement in the definition of CMIP6 and for the broad participation in the CMIP5 survey in 2013. We thank the co-chairs and steering committee members of the CMIP6-Endorsed MIPs for their continuous engagement in defining CMIP6, and the modelling groups and wider community for reviewing the CMIP6 design and organisation. We thank the WGCM Infrastructure Panel (WIP) for overseeing the CMIP6 infrastructure, Martin Juckes for taking the lead in preparing the CMIP6 data request, Jonathan Gregory for raising awareness about the treatment of volcanic forcing in the pre-industrial control experiment and its consequence for sea level changes, and Pierre Friedlingstein, George Hurtt, Chris Jones, and David Lawrence for help in defining carbon cycle and land use specifications. Norbert Noreiks is thanked for help in drafting the figures. Thanks to Gavin Schmidt and the two anonymous reviewers, and to everyone who contributed to the open discussions for constructive comments. GM was supported by the Regional and Global Climate Modeling Program (RGCM) of the U.S. Department of Energy's Office of Biological & Environmental Research (BER) Cooperative Agreement # DE-FC02-97ER62402, and the U.S. National Science Foundation. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

References

Alexander, L. V., Kumar, A., Naveau, P., Seneviratne, S. I., Sivakumar, M. V. K., Zhang, X., and Zwiers, F. W.: The World Climate Research Program Grand Challenge on Extremes – WCRP-ICTP Summer School on Attribution and Prediction of Extreme Events 2015.

Bony, S., Bellon, G., Klocke, D., Sherwood, S., Fermepin, S., and Denvil, S.: Robust direct effect of carbon dioxide on tropical circulation and regional precipitation, Nat Geosci, 6, 447-451, 2013.

Bony, S., Stevens, B., Frierson, D. M. W., Jakob, C., Kageyama, M., Pincus, R., Shepherd, T. G., Sherwood, S. C., Siebesma, A. P., Sobel, A. H., Watanabe, M., and Webb, M. J.: Clouds, circulation and climate sensitivity, Nature Geosci, 8, 261-268, 2015.

Brasseur, G. and Carlson, D.: Future directions for the World Climate Research Programme, Eos, Transactions American Geophysical Union, 96, 2015.

Cubasch, U., D. Wuebbles, D. Chen, M.C. Facchini, D. Frame, N. Mahowald, and J.-G. Winther: Introduction. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Ed.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Deser, C., Knutti, R., Solomon, S., and Phillips, A. S.: Communication of the role of natural variability in future North American climate, Nat Clim Change, 2, 775-779, 2012.

Eyring, V., Righi, M., Evaldsson, M., Lauer, A., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I., Davin, E. L., Deser, C., Ehbrecht, C., Friedlingstein, P., Gleckler, P., Gottschaldt, K. D., Hagemann, S., Juckes, M., Kindermann, S., Krasting, J., Kunert, D., Levine, R., Loew, A., Mäkelä, J., Martin, G., Mason, E., Phillips, A., Read, S., Rio, C., Roehrig, R., Senftleben, D., Sterl, A., van Ulft, L. H., Walton, J., Wang, S., and Williams, K. D.: ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth System Models in CMIP, Geosci. Model Dev. Discuss., 8, 7541-7661, 2015.

Ferraro, R., Waliser, D. E., Gleckler, P., Taylor, K. E., and Eyring, V.: Evolving obs4MIPs to Support the Sixth Coupled Model Intercomparison Project (CMIP6), B Am Meteorol Soc, doi: 10.1175/BAMS-D-14-00216.1, 2015. 2015.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Ed.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Gates, W. L., Boyle, J. S., Covey, C., Dease, C. G., Doutriaux, C. M., Drach, R. S., Fiorino, M., Gleckler, P. J., Hnilo, J. J., Marlais, S. M., Phillips, T. J., Potter, G. L., Santer, B. D., Sperber, K. R., Taylor, K. E., and Williams, D. N.: An Overview of the Results of the Atmospheric Model Intercomparison Project (AMIP I), B Am Meteorol Soc, 80, 29-55, 1999.

Geoffroy, O., Saint-Martin, D., Olivie, D. J. L., Voldoire, A., Bellon, G., and Tyteca, S.: Transient Climate Response in a Two-Layer Energy-Balance Model. Part I: Analytical Solution and Parameter Calibration Using CMIP5 AOGCM Experiments, J Climate, 26, 1841-1857, 2013.

Gleckler, P. J., Doutriaux, C., Durack P. J., Taylor K. E. , Zhang, Y., Williams, D. N., Mason, E., and Servonnat, J.: A More Powerful Reality Test for Climate Models, Eos Trans. AGU, in press, 2016.

Gregory, J. and Webb, M.: Tropospheric adjustment induces a cloud component in CO2 forcing, J Climate, 21, 58-71, 2008.

Gregory, J. M.: Long-term effect of volcanic forcing on ocean heat content, Geophys. Res. Lett., 37, L22701, 2010.

Gregory, J. M., Bi, D., Collier, M. A., Dix, M. R., Hirst, A. C., Hu, A., Huber, M., Knutti, R., Marsland, S. J., Meinshausen, M., Rashid, H. A., Rotstayn, L. D., Schurer, A., and Church, J. A.: Climate models without preindustrial volcanic forcing underestimate historical ocean thermal expansion, Geophys Res Lett, 40, 1600-1604, 2013.

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., Lowe, J. A., Johns, T. C., and Williams, K. D.: A new method for diagnosing radiative forcing and climate sensitivity, Geophys Res Lett, 31, 2004.

Hegerl, G. C., Zwiers, F. W., Braconnot, P., Gillett, N. P., Luo, Y., Marengo Orsini, J. A., Nicholls, N., Penner, J. E., and Stott, P. A.: Understanding and Attributing Climate Change. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (Ed.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Houghton, R. A.: How well do we know the flux of CO_2 from land-use change?, Tellus B, 62, 337-351, 2010.

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., Fisk, J. P., Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Goldewijk, K. K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D. P., and Wang, Y. P.: Harmonization of land-use scenarios for the period 1500-2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Climatic Change, 109, 117-161, 2011.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Kamae, Y. and Watanabe, M.: Tropospheric adjustment to increasing CO_2: its timescale and the role of land-sea contrast, Clim Dynam, 41, 3007-3024, 2013.

Li, C., Stevens, B., and Marotzke, J.: Eurasian winter cooling in the warming hiatus of 1998-2012, Geophys Res Lett, 42, 8131-8139, 2015.

Meehl, G. A., Boer, G. J., Covey, C., Latif, M., and Stouffer, R. J.: The Coupled Model Intercomparison Project (CMIP), B Am Meteorol Soc, 81, 313-318, 2000.

Meehl, G. A., Boer, G. J., Covey, C., Latif, M., and Stouffer, R. J.: Intercomparison makes for a better climate model, Eos, Transactions American Geophysical Union, 78, 445-451, 1997.

Meehl, G. A., Covey, C., Taylor, K. E., Delworth, T., Stouffer, R. J., Latif, M., McAvaney, B., and Mitchell, J. F. B.: THE WCRP CMIP3 Multimodel Dataset: A New Era in Climate Change Research, B Am Meteorol Soc, 88, 1383-1394, 2007.

Meehl, G. A., Moss, R., Taylor, K. E., Eyring, V., Stouffer, R. J., Bony, S., and Stevens, B.: Climate Model Intercomparisons: Preparing for the Next Phase, Eos Trans. AGU, 59, 77, 2014.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J.: The next generation of scenarios for climate change research and assessment, Nature, 463, 747-756, 2010.

Murphy, J. M. and Mitchell, J. F. B.: Transient-Response of the Hadley-Center Coupled Ocean-Atmosphere Model to Increasing Carbon-Dioxide .2. Spatial and Temporal Structure of Response, J Climate, 8, 57-80, 1995.

Myhre, G., D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, Takemura, T., and Zhang, H.: Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Ed.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

O’Neill, B. C., Kriegler, E., Ebi, K. L., Kemp-Benedict, E., Riahi, K., Rothman, D. S., van Ruijven, B. J., van Vuuren, D. P., Birkmann, J., Kok, K., Levy, M., and Solecki, W.: The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century, Global Environmental Change, doi: 10.1016/j.gloenvcha.2015.01.004, 2015. 2015.

Phillips, A. S., Deser, C., and Fasullo, J.: Evaluating Modes of Variability in Climate Models, Eos Trans. AGU, 95(49), 453-455, 2014.

Rauser, F., Gleckler, P., and Marotzke, J.: Rethinking the Default Construction of Multimodel Climate Ensembles, B Am Meteorol Soc, 96, 911-919, 2014.

Sentman, L. T., Shevliakova, E., Stouffer, R. J., and Malyshev, S.: Time Scales of Terrestrial Carbon Response Related to Land-Use Application: Implications for Initializing an Earth System Model, Earth Interactions, 15, 1-16, 2011.

Sherwood, S. C., Bony, S., Boucher, O., Bretherton, C., Forster, P. M., Gregory, J. M., and Stevens, B.: Adjustments in the Forcing-Feedback Framework for Understanding Climate Change, B Am Meteorol Soc, 96, 217-228, 2015.

Stott, P. A., Mitchell, J. F. B., Allen, M. R., Delworth, T. L., Gregory, J. M., Meehl, G. A., and Santer, B. D.: Observational constraints on past attributable warming and predictions of future global warming, J Climate, 19, 3055-3069, 2006.

Stouffer, R. J., Weaver, A. J., and Eby, M.: A method for obtaining pre-twentieth century initial conditions for use in climate change studies, Clim Dynam, 23, 327-339, 2004.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of Cmip5 and the Experiment Design, B Am Meteorol Soc, 93, 485-498, 2012.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: A Summary of the CMIP5 Experiment Design, http://cmip.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf, 2009.

Teixeira, J., Waliser, D., Ferraro, R., Gleckler, P., Lee, T., and Potter, G.: Satellite Observations for CMIP5: The Genesis of Obs4MIPs, B Am Meteorol Soc, 95, 1329-1334, 2014.

Trenberth, K. and Asrar, G.: Challenges and Opportunities in Water Cycle Research: WCRP Contributions, Surv Geophys, 35, 515-532, 2014.

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J. F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., and Rose, S. K.: The representative concentration pathways: an overview, Climatic Change, 109, 5-31, 2011.

Williams, D. N., Balaji, V., Cinquini, L., Denvil, S., Duffy, D., Evans, B., Ferraro, R., Hansen, R., Lautenschlager, M., and Trenham, C.: A Global Repository for Planet-Sized Experiments and Observations, B Am Meteorol Soc, doi: 10.1175/bams-d-15-00132.1, 2015. 150904101253006, 2015.

Williams, K. and Webb, M.: A quantitative performance assessment of cloud regimes in climate models, Clim Dynam, 33, 141-157, 2009.

Table 1. Main criteria for MIP endorsement as agreed with representatives from the modelling groups and MIPs at the WGCM 18th Session in Grainau, Germany in October 2014.

Nr	MIP Endorsement Criterion
1	The MIP and its experiments address at least one of the key science questions of CMIP6.
2	The MIP demonstrates connectivity to the DECK experiments and the CMIP6 ~~Historical Simulation~~historical simulations.
3	The MIP adopts the CMIP modelling infrastructure standards and conventions.
4	All experiments are tiered, well-defined, and useful in a multi-model context and ~~don't~~do not overlap with other CMIP6 experiments.
5	Unless a Tier 1 experiment differs only slightly from another well-established experiment, it must already have been performed by more than one modelling group.
6	A sufficient number of modelling centres (~8) are committed to performing all of the ~~MIP‘s~~MIP's Tier 1 experiments and providing all the requested diagnostics needed to answer at least one of its science questions.
7	The MIP presents an analysis plan describing how it will use all proposed experiments, any relevant observations, and specially requested model output to evaluate the models and address its science questions.
8	The MIP has completed the MIP template questionnaire.
9	The MIP contributes a paper on its experimental design to the GMD CMIP6 Special Issue.
10	The MIP considers reporting on the results by co-authoring a paper with the modelling groups.

Table 2. Overview of DECK and ~~the~~ CMIP6 ~~Historical Simulation~~historical simulations providing the experiment short names, the CMIP6 ~~label, a~~labels, brief experiment ~~description~~descriptions, the forcing methods as well as the start and end year and minimum number of years per ~~experiments~~experiment and its major purpose. The DECK and CMIP6 ~~Historical Simulation~~historical simulation are used to characterize the CMIP model ensemble. Given resource limitations, these entry card simulations for CMIP include only one ensemble member per experiment. However, we strongly encourage model groups to submit at least three ensemble members for the CMIP ~~*Historical Simulation*~~historical simulation as requested in DAMIP. Large ensembles of AMIP simulations are also encouraged. ~~"All" in~~In the "forcing methods" column, "All" means "volcanic, solar and anthropogenic forcings". All experiments are started on 1 January and end at 31 December of the specified years.

Experiment short name	**CMIP6 label**	**Experiment description**	**Forcing methods**	**Start Year**	**End Year**	**Minimum # Years Per Simulation**	**Major purpose**
DECK Experiments							
~~Historical~~ AMIP	*amip*	Observed SSTs and SICs prescribed	All; CO_2 concentration ~~-driven~~ prescribed	1979	2014	36	Evaluation, variability
~~Pre~~pre-industrial control	*piControl or esm-piControl*	Coupled atmosphere/ocean pre-industrial control	CO_2 ~~emission- or~~ concentration ~~-driven~~ prescribed or calculated	~~1850~~n/a	n/a	500	Evaluation, unforced variability
~~Quadruple~~abrupt quadrupling of CO_2 ~~abruptly, then hold fixed~~concentration	~~*abrupt4xCO2*~~*abrupt-4xCO2*	CO_2 abruptly quadrupled and then held constant	CO_2 concentration ~~-driven~~ prescribed	n/a	n/a	150	Climate sensitivity, feedbacks, fast responses
1% yr^{-1} CO_2 concentration increase	*1pctCO2*	CO_2 prescribed to increase at 1% yr^{-1}	CO_2 concentration ~~-driven~~ prescribed	n/a	n/a	150	Climate sensitivity, feedbacks, idealized benchmark
CMIP6 ~~Historical Simulation~~historical simulation							

~~Past~~past ~1.5 centuries	*historical* or *esm-hist*	Simulation of the recent past	All; CO_2 ~~emission- or~~ concentration ~~-driven~~ prescribed or calculated	1850	201 4	165	Evaluation

1099 **Table 3.** List of CMIP6-Endorsed MIPs along with the long name of the MIP, the primary goal(s)
1100 and the main CMIP6 science theme as displayed in Fig. 2. Each of these MIPs is described in more
1101 detail in a separate contribution to this Special Issue. MIPs marked with * are Diagnostic-MIPs.

Short name of MIP	Long name of MIP	Primary Goal(s) in CMIP6	Main CMIP6 Science Theme
AerChemMIP	Aerosols and Chemistry Model Intercomparison Project	a) Diagnosing forcings and feedbacks of tropospheric aerosols, tropospheric ozone precursors and the chemically reactive WMGHGs; b) Documenting and understanding past and future changes in the chemical composition of the atmosphere; c) Estimating the global to regional climate response from these changes.	Chemistry / Aerosols
C[4]MIP	Coupled Climate Carbon Cycle Model Intercomparison Project	Understanding and quantifying future century-scale changes in the global carbon cycle and its feedbacks on the climate system, making the link between CO_2 emissions and climate change.	Carbon cycle
CFMIP	Cloud Feedback Model Intercomparison Project	Improved assessments of cloud feedbacks via a) improved understanding of cloud-climate feedback mechanisms and b) better evaluation of clouds and cloud feedbacks in climate models. Also improved understanding of circulation, regional-scale precipitation and non-linear changes.	Clouds / Circulation
DAMIP	Detection and Attribution Model Intercomparison Project	a) Estimating the contribution of external forcings to observed global and regional climate changes; b) Observationally constraining future climate change projections by scaling future GHG and other anthropogenic responses using regression coefficients derived for the historical period.	Characterizing forcings
DCPP	Decadal Climate Prediction Project	Predicting and understanding forced climate change and internal variability up to 10 years into the future through a coordinated set of hindcast experiments, targeted experiments to understand the physical processes, and the ongoing production of skilful decadal predictions.	Decadal prediction
FAFMIP	Flux-Anomaly-Forced Model Intercomparison Project	Explaining the model spread in climate projections of ocean climate change forced by CO_2 increase, especially regarding the geographical patterns and magnitude of sea-level change, ocean heat uptake and thermal expansion.	Ocean / Land / Ice
GeoMIP	Geoengineering Model Intercomparison Project	Assessing the climate system response (including on extreme events) to proposed radiation modification geoengineering schemes by evaluating their efficacies,	~~Geo-engineering~~Geoengineering

		benefits, and side effects.	
GMMIP	Global Monsoons Model Intercomparison Project	a) Improve understanding of physical processes in global monsoons system; b) better simulating the mean state, interannual variability and long-term changes of global monsoons.	Regional phenomena
HighResMIP	High Resolution Model Intercomparison Project	Assessing the robustness of improvements in the representation of important climate processes with "weather-resolving" global model resolutions (~25km or finer), within a simplified framework using the physical climate system only with constrained aerosol forcing.	Regional phenomena
ISMIP6	Ice Sheet Model Intercomparison Project for CMIP6	Improving confidence in projections of the sea level rise associated with mass loss from the ice sheets of Greenland and Antarctica.	Ocean / Land / Ice
LS3MIP	Land Surface, Snow and Soil Moisture	Providing a comprehensive assessment of land surface, snow, and soil moisture-climate feedbacks, and diagnosing systematic biases in the land modules of current ESMs using constrained land-module only experiments.	Ocean / Land / Ice
LUMIP	Land-Use Model Intercomparison Project	Quantifying the effects of land use on climate and biogeochemical cycling (past-future), and assessing the potential for alternative land management strategies to mitigate climate change.	Land use
OMIP	Ocean Model Intercomparison Project	Provide a framework for evaluating, understanding, and improving ocean, sea-ice, and biogeochemical, including inert tracers, components of climate and ~~earth~~Earth system models contributing to CMIP6. Protocols are provided to perform coordinated ocean/sea-ice/tracer/~~biogeochemisty~~biogeochemistry simulations forced with common atmospheric datasets.	Ocean / Land / Ice
PMIP	Paleoclimate Modelling Intercomparison Project	a) Analysing the response to forcings and major feedbacks for past climates outside the range of recent variability; b) Assessing the credibility of climate models used for future climate projections.	Paleo
RFMIP	Radiative Forcing Model Intercomparison Project	a) Characterizing the global and regional effective radiative forcing for each model for historical and 4xCO_2 simulations; b) Assessing the absolute accuracy of clear-sky radiative transfer parameterizations; c) Identifying the robust impacts of aerosol radiative forcing during the historical period.	Characterizing forcings
ScenarioMIP	Scenario Model	a) Facilitating integrated research on the	Scenarios

	Intercomparison Project	impact of plausible future scenarios over physical and human systems, and on mitigation and adaptation options; b) addressing targeted studies on the effects of particular forcings in collaboration with other MIPs; c) help quantifying projection uncertainties based on multi-model ensembles and emergent constraints.	
VolMIP	Volcanic Forcings Model Intercomparison Project	a) Assessing to what extent responses of the coupled ocean-atmosphere system to strong volcanic forcing are robustly simulated across state-of-the-art coupled climate models; b) Identifying the causes that limit robust simulated behaviour, especially differences in their treatment of physical processes	Characterizing forcings
CORDEX*	Coordinated Regional Climate Downscaling Experiment	Advancing and coordinating the science and application of regional climate downscaling (RCD) through statistical and dynamical downscaling of CMIP DECK, CMIP6 ~~Historical Simulation~~historical, and ScenarioMIP output.	Impacts
DynVar*	Dynamics and Variability of the Stratosphere-Troposphere System	Defining and analysing diagnostics that enable a mechanistic approach to confront model biases and understand the underlying causes behind circulation changes with a particular emphasis on the two-way coupling between the troposphere and the stratosphere.	Clouds / Circulation
SIMIP*	Sea-Ice Model Intercomparison Project	Understanding the role of sea-ice and its response to climate change by defining and analysing a comprehensive set of variables and process-oriented diagnostics that describe the sea-ice state and its atmospheric and ocean forcing.	Ocean / Land / Ice
VIACS AB*	Vulnerability, Impacts, Adaptation and Climate Services Advisory Board for CMIP6	Facilitating a two-way dialogue between the CMIP6 modelling community and VIACS experts, who apply CMIP6 results for their numerous research and climate services, towards an informed construction of model scenarios and simulations and the design of online diagnostics, metrics, and visualization of relevance to society.	Impacts

1102

Table A1. Specifications in the DECK and CMIP6 ~~Historical Simulation~~historical simulations.

Experiment	Volcanic Stratospheric Aerosol	Solar Variability	Anthropogenic forcings
amip	Time-dependent observations	Time-dependent observations	Time-dependent observations
piControl	Background volcanic aerosol that results in radiative forcing matching, as closely as possible, that experienced, on average, during the historical simulation (i.e., 1850-2014 mean)	Fixed at its mean value (no 11 year solar cycle) over the first two solar cycles of the historical simulation (i.e., the 1850 – ~~1871~~1873 mean)	Given that the historical ~~starts~~simulations start in 1850, the *piControl* should have fixed 1850 atmospheric composition, not true pre-industrial
esm-piControl	As in *piControl*	As in *piControl*	As in *piControl* but with CO_2 concentration calculated, rather than prescribed. CO_2 from both fossil fuel combustion and land use change are prescribed to be zero.
~~*abrupt4xCO2*~~*abrupt-4xCO2*	As in *piControl*	As in *piControl*	As in *piControl* except CO_2 that is four times *piControl*
1pctCO2	As in *piControl*	As in *piControl*	As in *piControl* except CO_2 that is increasing at 1%/yr^{-1}
historical	Time-dependent observations	Time-dependent observations	Time-dependent observations
esm-hist	As in *historical*	As in *historical*	As in *historical* but with CO_2 emissions prescribed and CO_2 concentration calculated (rather than prescribed)

FIGURES

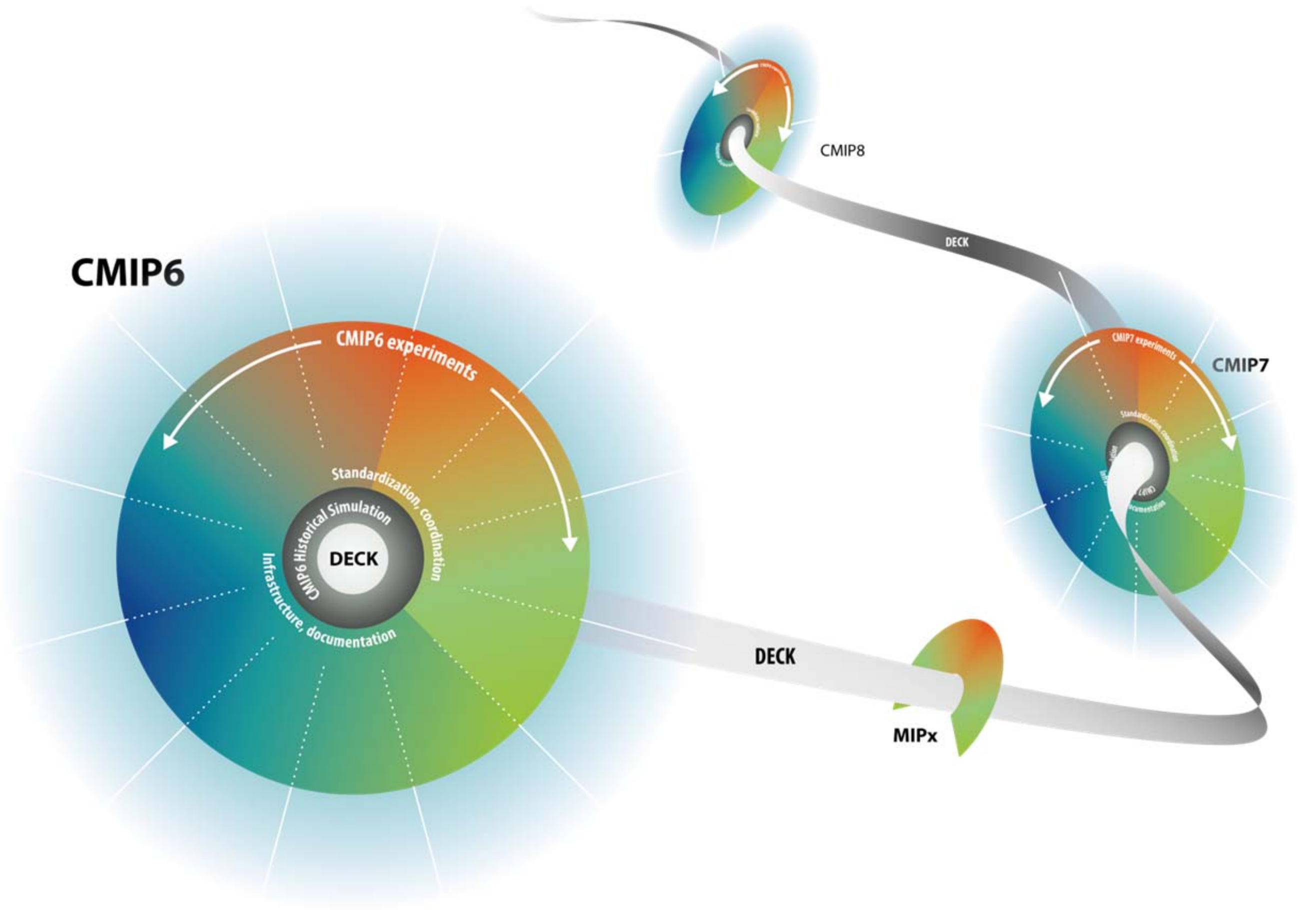


Figure 1. CMIP evolution. CMIP will evolve but the DECK will provide continuity across ~~different~~ phases.

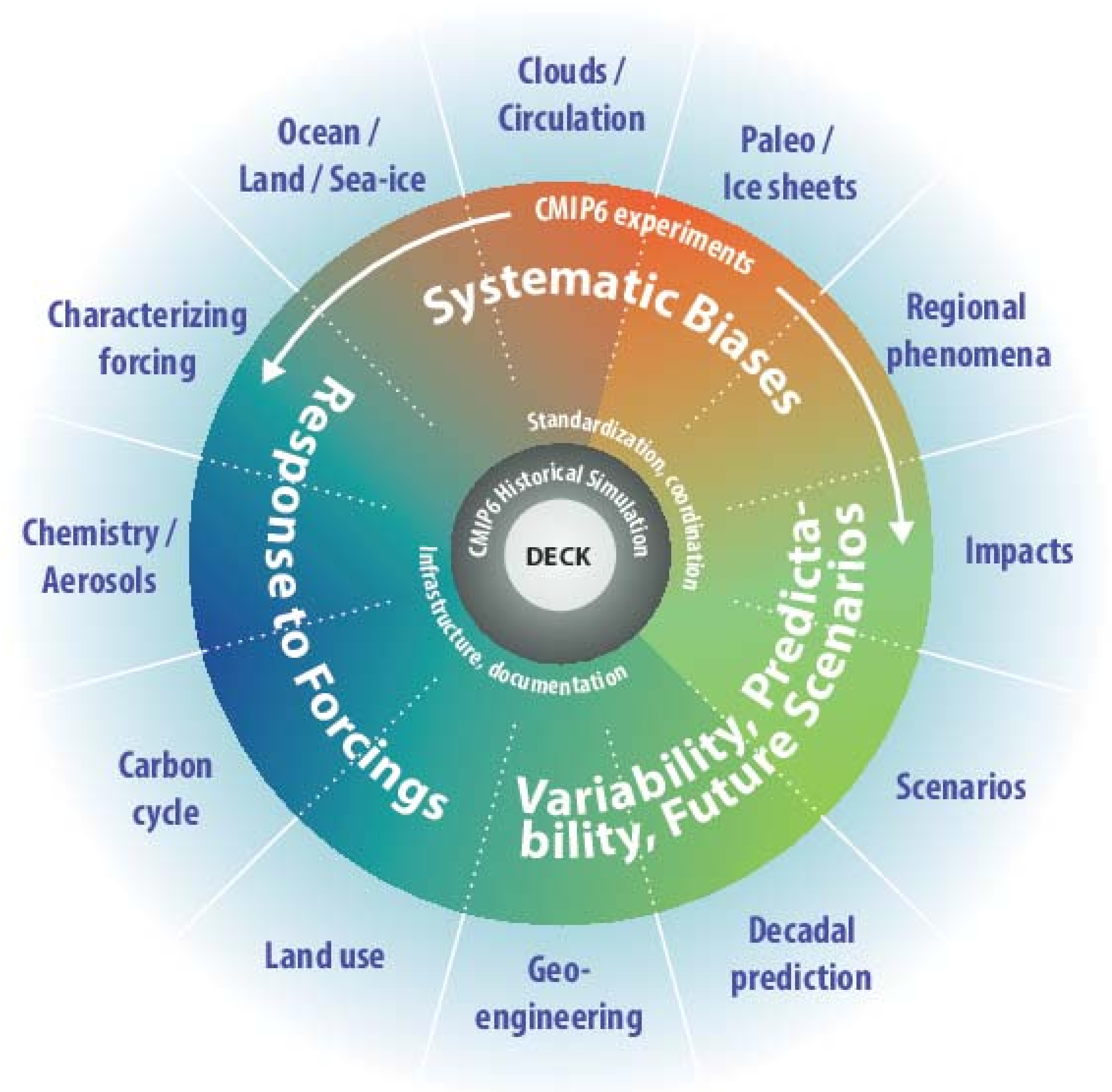


Figure 2. Schematic of the CMIP/CMIP6 experiment design. The inner ring and surrounding white text involve standardized functions of all CMIP DECK experiments and the CMIP6 ~~Historical Simulation.~~historical simulation. The middle ring shows science topics related specifically to CMIP6 ~~to be~~that are addressed by the CMIP6-Endorsed MIPs, with MIP topics shown in the outer ring. This framework is superimposed on the scientific backdrop for CMIP6 which are the seven WCRP Grand Science Challenges.

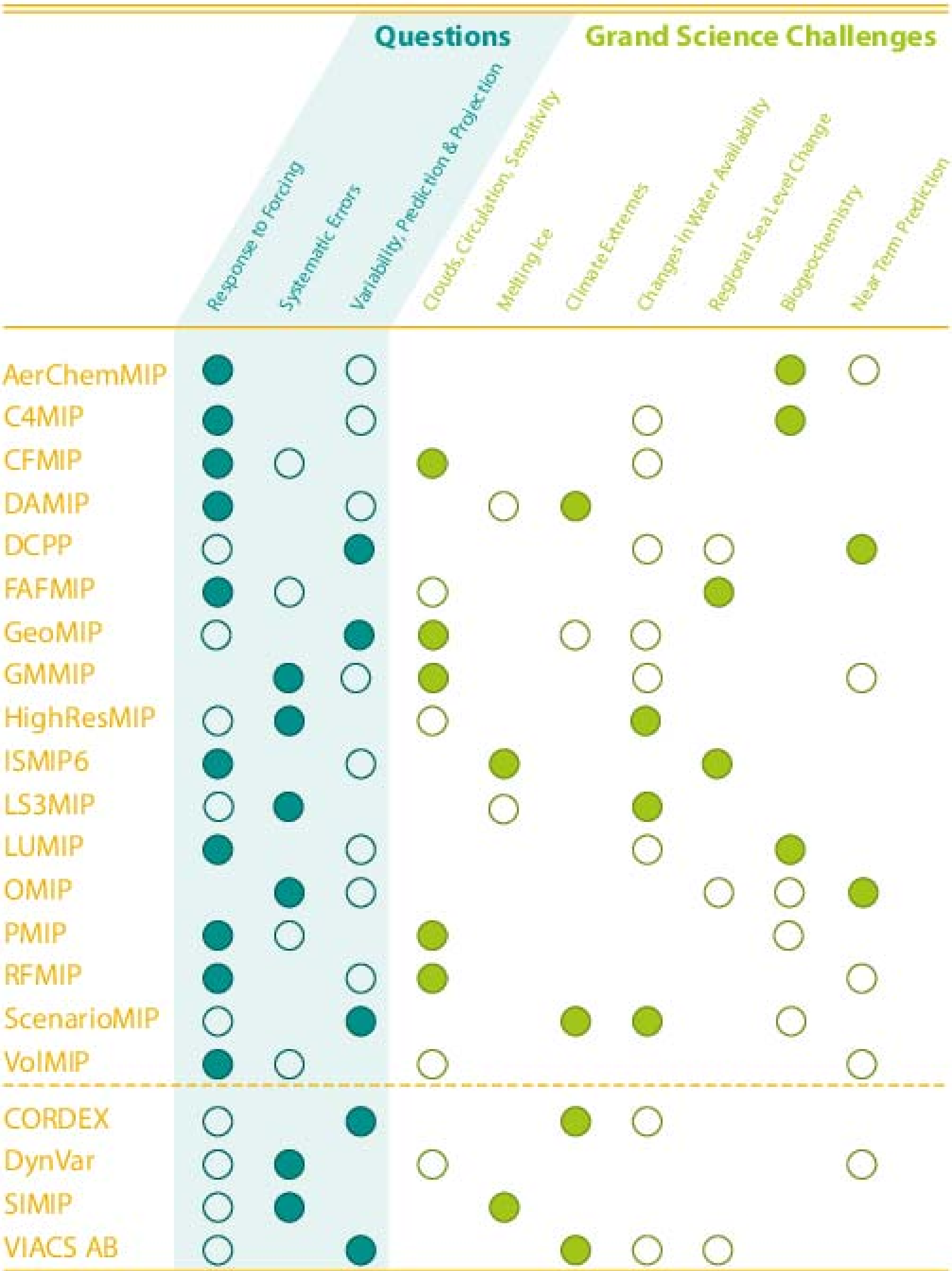
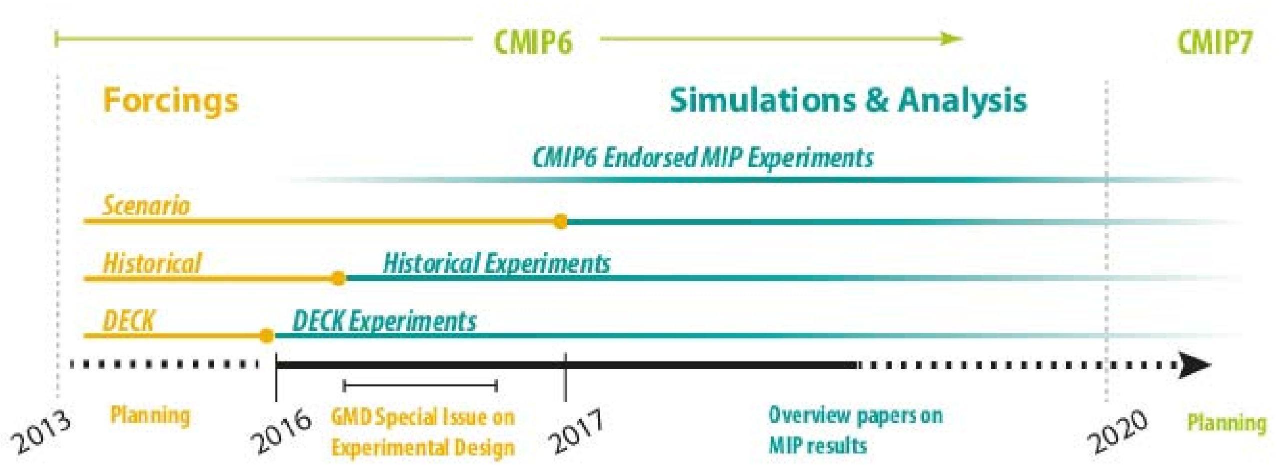


Figure 3. Contributions of CMIP6-Endorsed MIPs to the three CMIP6 science questions and the WCRP Grand Science Challenges. A filled circle indicates highest priority and an open circle, second highest priority. Some of the MIPs additionally contribute with lower priority to other CMIP6 science questions or WCRP Grand Science Challenges.



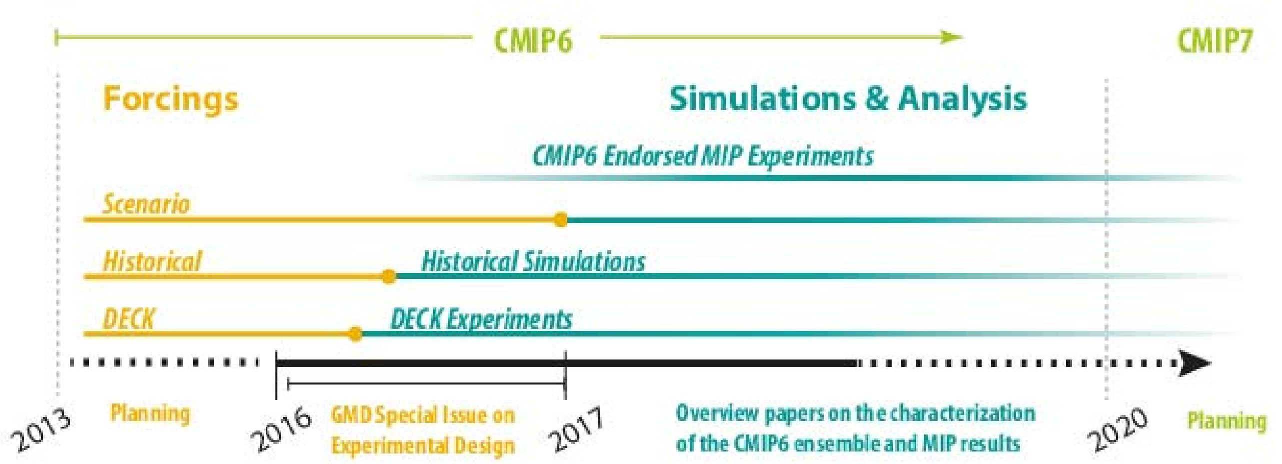


Figure 4. CMIP6 ~~Timeline~~timeline for the preparation of forcings, the realization of ~~simulations~~experiments and their analysis.